

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Pomalidomid (Imnovid[®])

Celgene GmbH

Modul 4 A

Imnovid[®] ist in Kombination mit Dexamethason indiziert für die Behandlung des rezidivierten oder refraktären multiplen Myeloms bei erwachsenen Patienten, die mindestens zwei vorausgegangene Therapien, darunter Lenalidomid und Bortezomib, erhalten haben und unter der letzten Therapie eine Progression gezeigt haben.

**Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen**

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	7
Abkürzungsverzeichnis	8
4 Modul 4 – allgemeine Informationen	12
4.1 Zusammenfassung der Inhalte von Modul 4.....	13
4.2 Methodik.....	22
4.2.1 Fragestellung.....	22
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	23
4.2.3 Informationsbeschaffung.....	24
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	25
4.2.3.2 Bibliografische Literaturrecherche.....	25
4.2.3.3 Suche in Studienregistern.....	26
4.2.3.4 Selektion relevanter Studien.....	26
4.2.4 Bewertung der Aussagekraft der Nachweise.....	27
4.2.5 Informationssynthese und –analyse.....	29
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	29
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	30
4.2.5.3 Meta-Analysen.....	37
4.2.5.4 Sensitivitätsanalysen.....	38
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	39
4.2.5.6 Indirekte Vergleiche.....	41
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	42
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	42
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	42
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	42
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	44
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	47
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	47
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	48
4.3.1.2.1 Studiendesign und Studienpopulationen.....	48
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	54
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	54
4.3.1.3.1 Endpunkt – RCT.....	55
4.3.1.3.1.1 Mortalität (Gesamtüberleben) – RCT.....	56
4.3.1.3.1.2 Morbidität – RCT.....	59
4.3.1.3.1.3 Lebensqualität – RCT.....	75
4.3.1.3.1.4 Sicherheit – RCT.....	87
4.3.1.3.2 Subgruppenanalysen – RCT.....	97

4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	101
4.3.2	Weitere Unterlagen.....	105
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	105
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	105
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	105
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	106
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	106
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	108
4.3.2.2	Nicht randomisierte vergleichende Studien	108
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	109
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	109
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	110
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	110
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	111
4.3.2.3	Weitere Untersuchungen.....	111
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	111
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	112
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	112
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	112
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	113
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	113
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	113
4.4.1	Beurteilung der Aussagekraft der Nachweise	113
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	115
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	125
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	126
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	127
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	127
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	127
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	127
4.5.4	Verwendung von Surrogatendpunkten	128
4.6	Liste der eingeschlossenen Studien.....	130
4.7	Referenzliste.....	130
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		139
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		140
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund		141
Anhang 4-D : Liste der abgebrochenen Studien		142
Anhang 4-E : Liste der laufenden Studien.....		143

Anhang 4-F : Methodik der eingeschlossenen Studien – RCT..... 145
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 155

Tabellenverzeichnis

	Seite
Tabelle 4-1: Übersicht der Kriterien für den Studieneinschluss	13
Tabelle 4-2: Zusammenfassung der Ergebnisse der Endpunkte Gesamtüberleben, Progressionsfreies Überleben, Ansprechen und Sicherheit.....	21
Tabelle 4-3: Einschlusskriterien für die Studienbewertung	24
Tabelle 4-4: Anforderungen an die Aussage zur Beleglage.....	29
Tabelle 4-5: Zensierungsmethodik bei dem Endpunkt PFS.....	38
Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	43
Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	44
Tabelle 4-8: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	47
Tabelle 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	48
Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	49
Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	50
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	50
Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	54
Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	55
Tabelle 4-15: Operationalisierung von Mortalität.....	56
Tabelle 4-16: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel.....	56
Tabelle 4-17: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel.....	57
Tabelle 4-18: Operationalisierung von Morbidität.....	59
Tabelle 4-19: Bewertung des Verzerrungspotenzials für PFS in RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-20: Ergebnisse für PFS aus RCT mit dem zu bewertenden Arzneimittel.....	62
Tabelle 4-21: Progression als signifikante Determinante reduzierter Lebensqualität.....	63
Tabelle 4-22: Progression als signifikante Determinante für unerwünschte Ereignisse	64
Tabelle 4-23: Bewertung des Verzerrungspotenzials für TTP in RCT mit dem zu bewertenden Arzneimittel	65

Tabelle 4-24: Ergebnisse für TTP aus RCT mit dem zu bewertenden Arzneimittel	66
Tabelle 4-25: Bewertung des Verzerrungspotenzials für die Ansprechrate gemäß den IMWG-Kriterien in RCT mit dem zu bewertenden Arzneimittel	67
Tabelle 4-26: Ergebnisse für die Ansprechrate gemäß den IMWG-Kriterien aus RCT mit dem zu bewertenden Arzneimittel.....	68
Tabelle 4-27: Mindestens partielles Ansprechen gemäß den IMWG-Kriterien als Determinante verbesserter Lebensqualität	69
Tabelle 4-28: Bewertung des Verzerrungspotenzials für die Ansprechrate gemäß den EBMT-Kriterien in RCT mit dem zu bewertenden Arzneimittel	70
Tabelle 4-29: Ergebnisse für Ansprechrate gemäß der EBMT-Kriterien aus RCT mit dem zu bewertenden Arzneimittel	71
Tabelle 4-30: Bewertung des Verzerrungspotenzials für Klinisches Ansprechen in RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-31: Ergebnisse für Klinisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-32: Operationalisierung von Lebensqualität	75
Tabelle 4-33: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel	76
Tabelle 4-34: Adhärenzraten bei der Lebensqualitätserhebung.....	77
Tabelle 4-35: Mixed Model für die Veränderung der Lebensqualität	79
Tabelle 4-36: Signifikante Ergebnisse für Lebensqualität aus RCT mit dem zu bewertenden Arzneimittel pro Zyklus.....	81
Tabelle 4-37: Mediane Zeit bis zur klinisch relevanten Verschlechterung der Lebensqualität	85
Tabelle 4-38: Operationalisierung von Sicherheit	87
Tabelle 4-39: Bewertung des Verzerrungspotenzials für Sicherheit in RCT mit dem zu bewertenden Arzneimittel	88
Tabelle 4-40: Übersicht über Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel.....	90
Tabelle 4-41: Detailübersicht der Unerwünschten Ereignisse aus RCT mit dem zu bewertenden Arzneimittel	91
Tabelle 4-42: Übersicht der SPM in RCT mit dem zu bewertenden Arzneimittel	93
Tabelle 4-43: Inzidenzraten für ausgewählte UE.....	94
Tabelle 4-44: Ergebnisse der Interaktionstests der Endpunkte PFS, OS, TTP, Ansprechen gemäß den IMWG-Kriterien, Ansprechen gemäß den EBMT-Kriterien.....	98
Tabelle 4-45: Ergebnisse der Interaktionstests der Sicherheitsendpunkte.....	99
Tabelle 4-46: Detaillierte Ergebnisse der Subgruppenanalysen des Endpunkts UE mit Grad 3/4 nach Krankheitspopulation.....	100
Tabelle 4-47: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	106

Tabelle 4-48: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	106
Tabelle 4-49: Operationalisierung von <Endpunkt xxx>.....	107
Tabelle 4-50: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	107
Tabelle 4-51: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	108
Tabelle 4-52: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	109
Tabelle 4-53: Operationalisierung von <Endpunkt xxx>.....	110
Tabelle 4-54: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	110
Tabelle 4-55: Operationalisierung von <Endpunkt xxx>.....	112
Tabelle 4-56: Zusammenfassung der Ergebnisse.....	117
Tabelle 4-57: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	126
Tabelle 4-58: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	127
Tabelle 4-59(Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel .	142
Tabelle 4-60 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	143
Tabelle 4-61(Anhang): Studiendesign und -methodik für Studie MM-003.....	145
Tabelle 4-62(Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MM-003	155

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	46
Abbildung 4-2: Kaplan-Meier Kurve für das Gesamtüberleben (7. September 2012), inkl. des HD-Dex-Arms, der für den Effekt des <i>cross-overs</i> adjustiert wurde (Celgene Europe Limited, 2013b).....	58
Abbildung 4-3: Kaplan-Meier Kurve für die Zeit bis zur Verschlechterung der QLQ-C30 Skala Insomnie	95
Abbildung 4-4: Kaplan-Meier Kurve für die Zeit bis zur Verschlechterung der QLQ-MY20 Skala Nebenwirkungen der Behandlung.....	96
Abbildung 4-5(Anhang): Flussdiagramm der eingeschlossenen Studie - RCT	154

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AM	Arzneimittel
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
bzgl.	bezüglich
bzw.	beziehungsweise
CHMP	<i>Committee for Medicinal Products for Human Use</i>
CONSORT	<i>Consolidated Standards of Reporting Trials</i>
CPMP	<i>Committee for Proprietary Medicinal Products</i>
CR	Komplettes Ansprechen (<i>Complete Response</i>)
CTC	<i>Common Terminology Criteria</i>
CTCAE	<i>Common Terminology Criteria for Adverse Events</i>
d	Tag (<i>day</i>)
Dex	Dexamethason
DGHO	Deutsche Gesellschaft für Hämatologie und Onkologie
d. h.	das heißt
DIMDI	Deutsches Institut für Medizinische Dokumentation
DKG	Deutsche Krebsgesellschaft
dl	Deziliter
EBMT	<i>European Group for Blood and Marrow Transplantation</i>
ECOG	<i>Eastern Cooperative Oncology Group</i>
EG	Europäische Gemeinschaft
EMA	<i>European Agency for the Evaluation of Medicinal Products</i> (seit 2009: EMA)
EMA	<i>European Medicines Agency</i>
EORTC	<i>European Organisation for Research and Treatment of Cancer</i>
EQ-5D	EuroQol 5 Dimensions
EPO	Erythropoetin
et al.	et alii
etc.	et cetera
e.V.	eingetragener Verein
FDA	<i>Food and Drug Administration</i>
g	Gramm

Abkürzung	Bedeutung
G-BA	Gemeinsamer Bundesausschuss
Gem..	Gemäß
ggf.	gegebenenfalls
h	Stunde (<i>hour</i>)
HD-Dex	hoch-dosiertes Dexamethason (<i>High-Dose Dexamethasone</i>)
HIV	humanes Immundefizienzvirus (<i>Human Immunodeficiency Virus</i>)
HR	<i>Hazard Ratio</i>
IDMC	<i>Independent Data Monitoring Committee</i>
i.d.R.	in der Regel
IMiD	immunmodulierende Substanz (<i>Immunomodulatory Drug</i>)
IMWG	<i>International Myeloma Working Group</i>
inkl.	inklusive
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRAC	<i>Independent Response Adjudication Committee</i>
ITT	<i>Intention-to-treat</i>
IVRS	Interaktives Voice-Response System
IWRS	Interaktives Web-Response System
J	Jahre
Kap.	Kapitel
KI	Konfidenzintervall
l	Liter
LD-Dex	niedrig-dosiertes Dexamethason (<i>Low-Dose Dexamethasone</i>)
m	männlich
MCID	<i>Minimum Clinically Important Difference</i>
MedDRA	<i>Medical Dictionary for Regulatory Activities</i>
mg	Milligramm
min	Minute
mind.	mindestens
ml	Milliliter
MM	Multiples Myelom
mmol	Millimol
Mo	Monat

Abkürzung	Bedeutung
MR	minimales Ansprechen (<i>Minimal Response</i>)
MTC	<i>Mixed Treatment Comparison</i>
MW	Mittelwert
n. a.	nicht anwendbar
n.b.	nicht berichtet
n. e.	nicht erreicht
n.s.	nicht signifikant
NCI	<i>National Cancer Institute</i>
Nr.	Nummer
NYHA	<i>New York Heart Association</i>
OR	<i>Odds Ratio</i>
OS	Gesamtüberleben (<i>Overall Survival</i>)
p	p-Wert
PD	Krankheitsprogression (<i>Progressive Disease</i>)
PFS	Progressionsfreies Überleben (<i>Progression-Free Survival</i>)
p. o.	per os
Pom	Pomalidomid
PR	partiellles Ansprechen (<i>Partial Response</i>)
PS	<i>Performance Status</i>
qd	quaque die
QLQ-C30	<i>Quality of Life Questionnaire Core 30</i>
QLQ-MY20	<i>Quality of Life Questionnaire Multiple Myeloma Module 20</i>
RCT	<i>Randomized Controlled Trial</i>
RPSFT	<i>Rank preserving structural failure time</i>
rrMM/RRMM	rezidiertes/refraktäres Multiples Myelom
sCR	stringentes komplettes Ansprechen (<i>Stringent Complete Response</i>)
SD	stabile Erkrankung (<i>Stable Disease</i>)
SEM	Standardfehler des Mittelwerts (<i>Standard Error of Mean</i>)
SGB	Sozialgesetzbuch
SGOT	Serum-Glutamat-Oxalacetat-Transferase
SGPT	Serum-Glutamat-Pyruvat-Transferase

Abkürzung	Bedeutung
SmPC	Zusammenfassung der Merkmale des Arzneimittels (<i>Summary of Product Characteristics</i>)
SPM	sekundäres Primärmalignom (<i>Second Primary Malignancy</i>)
STD	Standardabweichung (<i>Standard Deviation</i>)
STE	<i>Surrogate Threshold Effects</i>
STROBE	<i>Strengthening the Reporting of Observational Studies in Epidemiology</i>
SUE	Schwerwiegendes unerwünschtes Ereignis
TREND	<i>Transparent Reporting of Evaluations with Non-Randomized Design</i>
TTF	Zeit bis zum Therapieversagen (<i>Time to Treatment-Failure</i>)
TTP	Zeit bis zur Progression (<i>Time to Progression</i>)
u. a.	unter Anderem
UE	Unerwünschtes Ereignis
UK	<i>United Kingdom</i>
USA	<i>United States of America</i>
VerfO	Verfahrensordnung
VGPR	sehr gutes partielles Ansprechen (<i>Very Good Partial Response</i>)
vs.	versus
vsl.	voraussichtlich
w	weiblich
WHO	<i>World Health Organization</i>
Wo	Woche
z. B.	zum Beispiel

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Der medizinische Zusatznutzen von Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) bei der Behandlung von Patienten mit rezidiviertem oder refraktärem Multiplem Myelom, die mindestens zwei Vortherapien, darunter Bortezomib und Lenalidomid, erhalten haben und die unter der letzten Therapie eine Progression gezeigt haben, soll im Vergleich zur Behandlung mit hoch-dosiertem Dexamethason (HD-Dex) als Monotherapie untersucht werden. Als Zielp Parameter werden patientenrelevante Endpunkte in den Bereichen *Mortalität*, *Morbidität*, *Lebensqualität* und *Sicherheit* berücksichtigt. Die Beantwortung der Fragestellung erfolgt auf Basis randomisierter kontrollierter klinischer Studien.

Datenquellen

Als Datenquellen dienten zum einen interne Daten- und Informationssysteme von Celgene, zum anderen die Studienregister *clinicaltrials.gov*, *clinicaltrialsregister.eu* und *International Clinical Trials Registry Platform*. Hierbei wurde nach abgebrochenen, abgeschlossenen und laufenden Studien gesucht, bei denen das Unternehmen Celgene Sponsor bzw. finanziell beteiligt war. Zusätzlich wurde eine Suche nach Studien „Dritter“ durchgeführt. Eine bibliographische Literaturrecherche wurde nicht durchgeführt, da weltweit keine Zulassung von Pomalidomid vorliegt, die mindestens zwölf Monate vor Dossiereinreichung erfolgte und außerdem kein indirekter Vergleich für die Nutzenbewertung von Pomalidomid herangezogen wurde. Daher ist eine bibliographische Literaturrecherche nicht erforderlich.

Ein-/Ausschlusskriterien für Studien

Die Einschlusskriterien für die Auswahl der Studien sind in Tabelle 4-1 ersichtlich.

Tabelle 4-1: Übersicht der Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Indikation: Patienten mit rezidiviertem oder refraktärem Multiplem Myelom, die mindestens zwei Vortherapien, darunter Lenalidomid und Bortezomib, erhalten haben und unter der letzten Therapie eine Progression gezeigt haben.
E2	Intervention: Pomalidomid (4 mg) in Kombination mit niedrig-dosiertem Dexamethason
E3	Kontrollintervention: hoch-dosiertes Dexamethason
E4	Mindestens einer der in 4.2.2 formulierten patientenrelevanten Endpunkte wird berichtet
E5	Randomisierte kontrollierte Studie
E6	Studiendauer: keine Mindestdauer anwendbar, da die Studiendauer ereignisgesteuert ist

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

Die Bewertung der Aussagekraft der Ergebnisse der eingeschlossenen Studien erfolgte in zwei Schritten.

Zuerst wurden methodische Charakteristika randomisierter klinisch kontrollierter Studien endpunktübergreifend **auf Studienebene** untersucht. Diese Beurteilung betraf das Studiendesign, darunter die Erzeugung der Randomisierungssequenz, die Verdeckung der Gruppenzuweisung sowie die Verblindung der Studienteilnehmer und Behandler. Weiterhin sind Verzerrungen bei der Berichterstattung und andere verzerrende Aspekte in die Beurteilung eingeflossen. Das Verzerrungspotenzial auf Studienebene wurde als *niedrig* oder *hoch* eingestuft. Eine Einteilung als *niedrig* wurde vorgenommen, wenn eine Verzerrung der Ergebnisse unwahrscheinlich und keine Anhaltspunkte für verzerrende Aspekte vorhanden waren, die bei Behebung die Beurteilung in einer qualitativen Weise verändern würden.

In einem weiteren Schritt wurde die Aussagekraft der Ergebnisse **auf Endpunktebene** beurteilt. Hierbei wurden u. a. die Verblindung der Endpunkterheber, die Umsetzung des *Intention-to-Treat*-(ITT)-Prinzips, die ergebnisgesteuerte Berichterstattung der Ergebnisse und weitere potenziell verzerrende Punkte bewertet. Das Verzerrungspotenzial wurde als *niedrig* eingestuft, wenn sich keine Anhaltspunkte für eine Verzerrung fanden, die die Aussage der Ergebnisse bei Behebung grundlegend verändern würden.

Die eingeschlossenen Studien wurden studienbezogen anhand der Items 2b bis 14 des *Consolidated Standards of Reporting Trial*-(CONSORT)-*Statements 2010* beschrieben (Tabelle 4-61). In diesem Rahmen erfolgte eine Beschreibung von Studienziel, Methodik (Design, Studienteilnehmer, Interventionen, Zielkriterien, Fallzahl, Erzeugung und Geheimhaltung der Randomisierungssequenz [*Allocation Concealment*], Durchführung der Randomisierung, Verblindung, statistische Methoden) und der Ergebnisdarstellung (Patientenfluss, Anzahl Studienteilnehmer, Aufnahme/Rekrutierung).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Grundlage der Bewertung des therapeutischen Nutzens und Zusatznutzens von Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) war eine randomisierte klinische Studie der Phase III von Pom+LD-Dex gegen hoch-dosiertes Dexamethason als Monotherapie (HD-Dex). Die Bewertung des therapeutischen Nutzens und Zusatznutzens erfolgte in den patientenrelevanten Nutzendimensionen *Mortalität*, *Morbidität*, *Lebensqualität* und *Sicherheit* und den jeweiligen weitergehenden Operationalisierungen.

Der Zusatznutzen gilt durch die Orphan Drug Zulassung gemäß § 35a Absatz 1 Satz 10 SGB V als **belegt** (Kommission der Europäischen Gemeinschaften, 2009, European Medicines Agency (EMA), 2013b). Celgene hat sich, als das verantwortliche pharmazeutische Unternehmen, entschieden, ein über das geforderte Mindestmaß hinausgehendes Dossier einzureichen. Die Entscheidung wurde aufgrund der Güte der Datenlage getroffen, die aus

Unternehmenssicht bei Offenlegung aller relevanten Studienergebnisse einen erheblichen Zusatznutzen für Pomalidomid in der zugelassenen Indikation belegt.

Für den primären Endpunkt Progressionsfreies Überleben (*Progression-Free Survival*, PFS) wurde eine Interimsanalyse (vorzeitiger Abbruch nur bei wahrscheinlicher Erfolglosigkeit der Studie (*Futility*)) und eine finale Analyse durchgeführt. Primär wird der Zusatznutzen auf Basis der Daten zum Zeitpunkt der finalen PFS Analyse (7. September 2012) dargestellt. Als sekundärer Endpunkt wurde unter anderem das Gesamtüberleben (OS) ausgewertet. Hierfür war die Studie ebenfalls gepowert (dies bedeutet, auch OS wurde konfirmatorisch unter Einhaltung des globalen 5 % Signifikanzniveaus getestet). Zum Zeitpunkt der finalen PFS Analyse wurde eine Interimsanalyse für OS durchgeführt. Diese Interimsanalyse war bereits erfolgreich und die Studie wurde erfolgreich beendet. Die geplante finale Auswertung des Gesamtüberlebens erfolgte am 1. März 2013. Im Zuge dessen wurden ebenfalls das Progressionsfreie Überleben, die Zeit bis zur Progression, die Ansprechrate gemäß den IMWG- und EBMT-Kriterien (letzteres nur für die Remissionskategorie *minimales Ansprechen*) und die Sicherheit ausgewertet. Diese Daten werden unterstützend in diesem Dossier berichtet.

Um einer potentiellen Verzerrung durch das unverblindete Studiendesign vorzubeugen, wurden die Ereignisse Progression und Ansprechen für den Auswertungszeitpunkt 7. September 2012 verblindet durch ein *Independent Response Adjudication Committee* (IRAC) beurteilt. Nachdem die Studie sowohl den primären Endpunkt (PFS) als auch die Interimsanalyse für das Gesamtüberleben erfolgreich erreicht hatte, empfahl das *Independent Data Monitoring Committee* (IDMC) den Patienten des HD-Dex Arms die Behandlung mit Pomalidomid zu ermöglichen. Die nachfolgende Analyse (1. März 2013) fokussierte auf dem langfristigen Gesamtüberleben; für das Ansprechen und die Progression basieren die Auswertungen auf der Bewertung der behandelnden Prüfarzte.

Mortalität

Für die Kombination von Pomalidomid mit LD-Dexamethason wurde eine statistisch signifikante und therapeutisch relevante Verlängerung des Gesamtüberlebens gezeigt. Im Interventionsarm trat dabei bis zum Zeitpunkt der Interimsanalyse des Gesamtüberlebens eine nicht ausreichende Anzahl von Ereignissen auf, um das mediane Überleben berechnen zu können (nicht erreicht vs. 34,0 Wochen). Die Behandlung mit Pom+LD-Dex war im Vergleich zu einer Therapie mit HD-Dex mit einer Verringerung des Sterberisikos von 47 % verbunden (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$). Studienteilnehmer des Kontrollarms konnten nach Krankheitsprogression in eine Begleitstudie (MM-003/C) zu einer Pomalidomid-Monotherapie wechseln, so dass auch die Teilnehmer der Kontrollgruppe von der Studienmedikation und der Verlängerung des Gesamtüberlebens durch Pomalidomid profitierten. Dies führt zu einer Verzerrung der beobachteten Ergebnisse, da sich der beobachtete Effekt zwischen den Studienarmen verringert. Die Verzerrung ist jedoch zuungunsten des Interventionsarms, da hiermit eine Überschätzung des Therapieeffektes im Kontrollarm einhergeht und somit der Therapievorteil von Pomalidomid unterschätzt wird. Die Unterbindung eines *cross-overs* ist aus ethischen Gründen nicht vertretbar. In MM-003 war die Gabe von Pom+LD-Dex trotz dieses *cross-overs* dennoch mit einer statistisch signifikanten Verlängerung des Gesamt-

überlebens gegenüber dem Kontrollarm verbunden. Eine statistische Modellierung nach Robins und Tsiatis (1991) ergab, dass der Median von HD-Dex ohne den Vorteil eines *cross-overs* bei 26,8 Wochen liegen würde (Robins und Tsiatis, 1991).

Das oben dargestellte Ergebnis beruht auf der Interimsanalyse des Gesamtüberlebens (7. September 2012), die bereits ein signifikantes Ergebnis zeigte und zu einem vorzeitigen erfolgreichen Ende der Studie führte. Das IDMC empfahl zu diesem Zeitpunkt die Beendigung der Studie und die Behandlung der Patienten des Kontrollarms mit Pomalidomid (Nachbeobachtung der Patienten dauert an). Zusätzlich wurde das Gesamtüberleben auch zum Zeitpunkt der ursprünglich geplanten finalen Analyse des Gesamtüberlebens ausgewertet (1. März 2013). Hier wurde trotz erheblichen *cross-overs* (50,0 % der Patienten im LD-Dex-Arm) weiterhin ein signifikanter Vorteil von Pom+LD-Dex mit einem medianen Gesamtüberleben von 54,6 Wochen im Pom+LD-Dex-Arm und 34,8 Wochen im HD-Dex-Arm (HR [95 %-KI]: 0,74 [0,56-0,97]; $p < 0,028$) beobachtet. Dies bedeutet einen Überlebensvorteil von 4,6 Monaten (19,8 Wochen).

Die hier gezeigte Verlängerung des Gesamtüberlebens ist auch vor dem Hintergrund der besonderen Therapiesituation (es handelt sich um vielfach vortherafierte Patienten ohne weitere effektive Therapieoption) bemerkenswert: Pomalidomid verbessert die Prognose dieser Patienten deutlich.

Morbidität

Bezogen auf den primären Endpunkt *Progressionsfreies Überleben* zeigte sich ebenfalls eine deutliche Überlegenheit der Interventionstherapie. Das vom IRAC bewertete Progressionsfreie Überleben (PFS) der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen der Kontrollgruppe im Median um 7,7 Wochen (15,7 vs. 8,0 Wochen) verlängert. Die hiermit assoziierte relative Risikoreduktion für das Eintreten von Krankheitsprogression oder Tod betrug 55 % (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$), und war statistisch signifikant.

Die Auswertung zum Zeitpunkt 1. März 2013 bestätigte den Vorteil für Pom+LD-Dex in Bezug auf das Progressionsfreie Überleben nach Prüfarztbewertung (17,2 vs. 8,2 Wochen; HR [95 %-KI]: 0,48 [0,39-0,60]; $p < 0,001$). Zum Vergleich ergab das PFS nach Prüfarztbewertung zum Zeitpunkt 7. September 2012 eine relative Risikoreduktion von 59 % (HR [95 %-KI]: 0,41 [0,32-0,53]; $p < 0,001$). Eine derartige Verlängerung des Progressionsfreien Überlebens ist vor dem Hintergrund der besonderen Therapiesituation bemerkenswert und therapeutisch relevant.

Die vom IRAC bewertete *Zeit bis zur Progression* der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen des Kontrollarms um 11,8 Wochen verlängert. Der Unterschied (20,1 vs. 8,3 Wochen) war statistisch signifikant und mit einer relativen Risikoreduktion für das Auftreten einer Progression um 58 % verbunden (HR [95 %-KI]: 0,42 [0,31-0,56]; $p < 0,001$). Zum Auswertungszeitpunkt 1. März 2013 unterschied sich die Zeit bis

zur Progression nach Prüfarztbewertung weiterhin signifikant mit 20,2 Wochen unter Pom+LD-Dex und 9,0 Wochen unter HD-Dex.

Auch bezogen auf den Endpunkt *Gesamtansprechen* nach IRAC zeigte sich eine Überlegenheit von Pom+LD-Dex gegenüber einer Behandlung mit HD-Dex.

Gemäß IMWG-Kriterien wurde zum Auswertungszeitpunkt 7. September 2012 bei keinem Studienteilnehmer ein stringentes komplettes Ansprechen beobachtet. Studienteilnehmer der Interventionsgruppe erreichten jedoch häufiger ein komplettes und ein sehr gutes partielles Ansprechen als jene des Kontrollarms (0,3 vs. 0,0 % und 1,3 vs. 0,7 %).

Das *Gesamtansprechen gemäß den IMWG-Kriterien* war in der Interventionsgruppe zum Auswertungszeitpunkt 7. September 2012 mehr als viermal höher als im HD-Dex-Arm (16,6 vs. 3,9 %; OR [95 %-KI]: 4,86 [2,03-11,61]; $p < 0,001$).

Das *Ansprechen gemäß den European Group for Blood and Marrow Transplantation (EBMT) Kriterien* bestätigt die Daten aus der Auswertung des Ansprechens gemäß den IMWG-Kriterien. Das Gesamtansprechen gemäß den EBMT-Kriterien war im Interventionsarm mehr als viermal höher als im Kontrollarm (14,9 vs. 3,3 %; OR [95 %-KI]: 5,18 [2,01-13,35]; $p < 0,001$).

Für den Auswertungszeitpunkt 1. März 2013 betrug das Gesamtansprechen nach Prüfarztbewertung gemäß den IMWG-Kriterien 31,0 % im Pom+LD-Dex-Arm und 10,0 % im HD-Dex-Arm (OR [95 %-KI]: 4,22 [2,35-7,58]; $p < 0,001$). Patienten im Interventionsarm erreichten häufiger ein komplettes oder sehr gutes partielles Ansprechen als jene des Kontrollarms (1,0 vs. 0,0 % und 5,0 vs. 0,7 %). Ein Patient im Interventionsarm (0,3 %) erreichte eine stringente komplette Remission. Das Ansprechen gemäß den EBMT-Kriterien wurde zu diesem Auswertungszeitpunkt nur für die Kategorie minimales Ansprechen (MR) erhoben; dieses wurde von Patienten im Interventionsarm nicht signifikant häufiger erreicht (8,0 vs. 6,0 %; $p = 0,565$).

Da für den primär dargestellten Auswertungszeitpunkt 7. September 2012 sowohl Progression wie auch das Ansprechen verblindet durch ein *Independent Response Adjudication Committee (IRAC)* beurteilt wurden, liegt kein Hinweis für ein erhöhtes Verzerrungspotenzial für die Endpunkte *Progressionsfreies Überleben*, *Zeit bis zur Progression* und *Ansprechen* vor. Die Ergebnissicherheit ist hoch.

Bezogen auf den Endpunkt *Klinisches Ansprechen* zeigte sich ein Vorteil für die Therapie mit Pom+LD-Dex bzgl. der Verbesserung des *ECOG Performance Status*. Signifikant mehr Studienteilnehmer des Interventionsarms profitierten von einer Verbesserung als jene der Kontrollgruppe (20,7 vs. 11,1 %; OR [95 %-KI]: 2,09 [1,09-4,02]; $p = 0,027$). Ebenso wurde für mehr Patienten im Pom+LD-Dex-Arm eine Verbesserung von Knochenschmerzen beobachtet. Studienteilnehmer des Interventionsarms profitierten häufiger von einer Verbesserung von Knochenschmerzen als jene der Kontrollgruppe (38,2 vs. 32,6 %; OR [95 %-KI]: 1,28 [0,75-2,18]; $p = 0,422$). Dieser Unterschied war statistisch nicht signifikant. Das Verzerrungspotenzial für diesen Endpunkt ist hoch, da die Studie nicht verblindet durchgeführt wurde.

Lebensqualität

Die Lebensqualität wurde durch den *European Organisation for Research and Treatment of Cancer* (EORTC) QLQ-C30 Fragebogen, den krankheitsspezifischen EORTC QLQ-MY20 Fragebogen und den EuroQoL (EQ)-5D-Fragebogen jeweils an Tag 1 eines jeden Behandlungszyklus' à 28 Tage, sowie zum Zeitpunkt des Behandlungsabbruchs bzw. Zeitpunkt der Progression erhoben.

Die Adhärenz, der zum jeweiligen Erhebungszeitpunkt noch in der Studie befindlichen Studienteilnehmer, war im Studienverlauf über alle Zyklen für alle drei Erhebungsinstrumente hoch und war in beiden Gruppen vergleichbar.

Die Ergebnisse zur Lebensqualität waren heterogen. Zusammenfassend lässt sich feststellen, dass die Kombination der zwei Substanzen Pomalidomid und niedrig-dosiertes Dexamethason im Vergleich zur Therapie mit hoch-dosiertem Dexamethason als Einzelsubstanz in Summe die Lebensqualität der Patienten erhielt. In einzelnen Aspekten war vor allem in frühen Behandlungszyklen ein Hinweis auf einen positiven Effekt gegenüber dem Kontrollarm zu beobachten. Insbesondere konnte gezeigt werden, dass das unter Pom+LD-Dex verbesserte Ansprechen sowie die signifikante Verzögerung der Progression zu einer Verbesserung der Lebensqualität bzw. einer Verringerung von Symptomen führen. Das Verzerrungspotenzial für diesen Endpunkt ist hoch, da die Studie nicht verblindet durchgeführt wurde.

Sicherheit

Bei nahezu allen Studienteilnehmern beider Arme wurde mindestens ein unerwünschtes Ereignis (UE) (Pom+LD-Dex 97,0 vs. HD-Dex 93,3 %) und bei etwa der Hälfte der Studienteilnehmer beider Gruppen mindestens ein schwerwiegendes unerwünschtes Ereignis (51,0 vs. 50,3 %) beobachtet. Die Unterschiede zwischen den Gruppen waren jeweils statistisch nicht signifikant. Die Gesamtrate an Studienteilnehmern mit mindestens einem unerwünschten Ereignis mit *Common Terminology Criteria* (CTC) Grad 3/4 war in beiden Gruppen vergleichbar (78,0 vs. 75,8 %). Der Anteil der Studienteilnehmer, die mindestens ein zum Absetzen des Studienarzneimittels führendes unerwünschtes Ereignis hatten, war in der Interventionsgruppe nicht signifikant erhöht (9,7 vs. 5,4 %). Der Anteil jener Teilnehmer, die innerhalb von 30 Tagen nach Gabe der letzten Dosis des Studienarzneimittels verstarben, war in der Interventionsgruppe niedriger als im Kontrollarm (12,7 vs. 14,1 %). Der Unterschied war statistisch nicht signifikant.

Bei Studienteilnehmern, die mit Pom+LD-Dex behandelt wurden, wurden signifikant häufiger Neutropenien (41,7 vs. 14,8 %), Leukopenien (8,7 vs. 3,4 %), febrile Neutropenien (6,7 vs. 0,0 %) und hämatopoetische Leukopenien (47,7 vs. 15,4 %) der Grade 3 und 4 beobachtet. Es ist anzumerken, dass eine Leukopenie in den meisten Fällen auf eine Verringerung der Zahl der neutrophilen Granulozyten, also eine Neutropenie, zurückzuführen ist. Diese hämatologischen Nebenwirkungen sind mit Dosisreduktionen und/oder dem Einsatz von Wachstumsfaktoren im Allgemeinen gut handhabbar. Neutropenien wirken sich in der überwiegenden

Mehrzahl der Fälle erst dann negativ auf die Lebensqualität der Patienten aus, wenn sie mit infektionsbedingten Symptomen einhergehen.

Ansonsten gab es keine statistisch signifikanten Unterschiede bezogen auf CTC Grad 3/4 UE zwischen Interventions- und Kontrollgruppe. In beiden Armen entwickelten bis zum 7. September 2012 0,7 % der Studienteilnehmer mindestens ein sekundäres Primärmalignom. Die beobachteten sekundären Primärmalignome waren nicht-invasive Basalzellkarzinome der Haut; invasive solide oder hämatologische Neoplasien wurden nicht beobachtet. Zum 1. März 2013 hatten zwei Patienten im Pom+LD-Dex-Arm ein invasives, solides SPM entwickelt. Es gibt bislang – insbesondere vor dem Hintergrund der multiplen Vortherapien der Patienten – keine Hinweise auf einen Zusammenhang zwischen Pomalidomid und der Entstehung von SPM. Im HD-Dex-Arm wurden keine weiteren SPM beobachtet.

Nur 6,6 % der Patienten im Interventionsarm brachen die Therapie aufgrund von unerwünschten Ereignissen ab (Kontrollarm: 5,9 %; statistisch nicht signifikant unterschiedlich). Die Auswertung der unerwünschten Ereignisse bis zum 1. März 2013 bestätigte das beschriebene Sicherheitsprofil.

Das Verzerrungspotenzial für die Nutzendimension *Sicherheit* ist hoch, da die Studie nicht verblindet durchgeführt wurde.

Subgruppen

Subgruppenanalysen wurden für die Faktoren *Geschlecht, Alter, Krankheitspopulation¹, Anzahl vorheriger Myelomtherapien, zytogenetische Risikogruppe* und *Durie-Salmon-Stadium* durchgeführt. Es ergab sich kein Anhaltspunkt für einen unterschiedlichen Behandlungseffekt in den unterschiedlichen Subgruppen. Dies bestätigt auch die EMA im *CHMP Assessment Report* (European Medicines Agency (EMA), 2013a): *“PFS and OS analyses were consistently favourable for Pom+LD-dex group in all subgroups”* (PFS und OS Analysen zeigten durchgehend Vorteile für den Pom+LD-Dex-Arm in allen Subgruppen).

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Der Bewertung des Ausmaßes und der Wahrscheinlichkeit des Zusatznutzens liegt die Studie MM-003 zugrunde. Hierbei handelt es sich um eine randomisierte kontrollierte Studie mit hoher Ergebnissicherheit und Evidenz, bei der sich keine Anhaltspunkte für ein Verzerrungspotenzial auf Studienebene finden. Die Studie MM-003 entspricht den Kriterien für eine adäquate pivotale Studie zum Beleg eines Zusatznutzens.

Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) führt im Vergleich zur alleinigen Anwendung von hoch-dosiertem Dexamethason (HD-Dex) zu einer

¹ *Gruppe 1:* Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. *Gruppe 2:* Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten nach Ende der Lenalidomid- und/oder Bortezomibtherapie. *Gruppe 3:* Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach \geq 2 Bortezomibzyklen

patientenrelevanten Reduktion des Mortalitätsrisikos von 47 % (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$). Auch zytogenetische Hochrisikopatienten und ältere Patienten > 65 Jahre profitieren. Die finale Auswertung des Gesamtüberlebens zeigte trotz erheblichem *cross-over* Anteil (50,0 %) einen signifikanten Überlebenszeitgewinn unter Pom+LD-Dex von 4,6 Monaten (54,6 vs. 34,8 Wochen; HR [95 %-KI]: 0,74 [0,56-0,97]; $p < 0,028$). Gemäß Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) § 5 Absatz 7 liegt ein erheblicher Zusatznutzen u. a. bei einer erheblichen Verlängerung der Überlebensdauer vor (Bundesministerium der Justiz, 2010). Pom+LD-Dex erzielt diesbezüglich gegenüber der Vergleichstherapie eine große, bisher nicht erreichte Verbesserung. Für die Nutzendimension *Mortalität* ist der Zusatznutzen – gerade vor dem Hintergrund der ungünstigen Prognose und besonderen Therapiesituation der Patienten - damit **erheblich**.

Unter Pom+LD-Dex verdoppelt sich das Progressionsfreie Überleben nahezu (15,7 vs. 8,0 Wochen). Das Risiko für das Eintreten von Krankheitsprogression oder Tod wird durch eine Therapie mit Pom+LD-Dex im Vergleich zu HD-Dex um 55 % gesenkt (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$). Dieser Vorteil blieb auch nach längerer Nachbeobachtungszeit (median 10 Monate, 1. März 2013) erhalten. Zudem hatten die Studienteilnehmer unter Pom+LD-Dex gemäß den IMWG-Kriterien nach IRAC Bewertung eine 4,86-fach höhere Wahrscheinlichkeit für ein mindestens partielles Ansprechen (OR [95 %-KI]: 4,86 [2,03-11,61]; $p < 0,001$). Die Auswertung zum 1. März 2013 nach Prüfarztbewertung ergab eine Gesamtansprechrate im Interventionsarm von 31,0 % und 10,0 % im Kontrollarm. Unter Pom+LD-Dex erreichten mehr Patienten eine komplette oder sehr gute partielle Remission (1,0 vs. 0,0 % und 5,0 vs. 0,7 %). Eine Auswertung des Ansprechens gemäß den EBMT-Kriterien bestätigte die deutliche Verbesserung der Ansprechrate durch Pom+LD-Dex. Da für Patienten mit fortgeschrittener Tumorerkrankung auch bereits ein geringfügiger Rückgang der Tumorlast relevant sein kann, wurde ebenfalls das Minimale Ansprechen bewertet. Das sich daraus ergebende Objektive Ansprechen (\geq MR) zeigt eine deutliche Überlegenheit von Pom+LD-Dex im Vergleich zu HD-Dex (26,8 vs. 6,5 %). Für den Endpunkt *Klinisches Ansprechen* ergab sich ein weiterer Vorteil für die Therapie mit Pom+LD-Dex gegenüber dem Kontrollarm: unter Pom+LD-Dex wurde für signifikant mehr Patienten eine Verbesserung des *ECOG PS Scores* als unter HD-Dex erreicht. Weiterhin wurde für mehr Patienten im Pom+LD-Dex eine Verbesserung des Symptoms *Knochenschmerzen* beobachtet.

Gemäß AM-NutzenV § 5 Absatz 7 liegt ein erheblicher Zusatznutzen u. a. beim Erreichen einer langfristigen Freiheit von schwerwiegenden Symptomen vor (Bundesministerium der Justiz, 2010). Pom+LD-Dex erzielt im Vergleich zu HD-Dex signifikant verlängerte Progressionsfreiheit und signifikant erhöhte Ansprechraten, was zu einer Verminderung der krankheitsbedingten Symptome beiträgt. Dies zeigt sich auch darin, dass mehr Patienten unter Pom+LD-Dex Verbesserungen in Bezug auf Knochenschmerzen zeigen. Aufgrund dieser Verbesserung ist der Zusatznutzen in der Nutzendimension *Morbidität* **erheblich**.

Die Ergebnisse zur Lebensqualität waren heterogen. Zusammenfassend lässt sich feststellen, dass die Kombination der zwei Substanzen Pomalidomid und niedrig-dosiertes Dexamethason im Vergleich zur Therapie mit hoch-dosiertem Dexamethason als Einzelsubstanz in Summe

die Lebensqualität der Patienten erhielt. In einzelnen Aspekten war vor allem in frühen Behandlungszyklen ein positiver Effekt gegenüber dem Kontrollarm zu beobachten.

Bezieht man die deutliche Verlängerung der Zeit bis zur Progression sowie die maßgeblich erhöhte Ansprechrate unter Pom+LD-Dex und den Effekt dieser Einflussvariablen auf die Lebensqualität mit in die Gesamtschau ein, so ist von einem verlängerten Erhalt der Lebensqualität unter Pom+LD-Dex im Vergleich zu HD-Dex auszugehen.

Insgesamt ist es bemerkenswert, dass unter der Zweierkombination Pom+LD-Dex die Verlängerung des Gesamtüberlebens ohne Einschränkungen der Lebensqualität im Vergleich zur Monotherapie mit HD-Dex einhergeht. Die Kombinationstherapie Pom+LD-Dex im Vergleich zur Monotherapie mit HD-Dex verursacht also keinen erhöhten Schaden in Bezug auf die Lebensqualität der Patienten.

Angesichts der Schwere der Erkrankung und der bestehenden Vorbelastung der Patienten durch multiple Rezidive und Vortherapien zeigte sich insgesamt eine gute Verträglichkeit von Pom+LD-Dex. Dies wird zusätzlich dadurch unterstrichen, dass nur 6,6 % der Patienten im Interventionsarm die Therapie aufgrund von unerwünschten Ereignissen abbrachen (unter HD-Dex: 5,9 %; statistisch nicht signifikant unterschiedlich). Eine Zusammenfassung der Ergebnisse der Endpunkte *Gesamtüberleben*, *Progressionsfreies Überleben*, *Ansprechen gemäß den IMWG-Kriterien* und *Sicherheit* ist in Tabelle 4-2 dargestellt. Der Endpunkt Lebensqualität ist in der folgenden Tabelle nicht dargestellt. Für diesen wurde kein relevanter Unterschied zwischen den beiden Behandlungsarmen gefunden.

Tabelle 4-2: Zusammenfassung der Ergebnisse der Endpunkte Gesamtüberleben, Progressionsfreies Überleben, Ansprechen und Sicherheit

Studie	Effektschätzer HR/OR [95 %-KI]; p-Wert							
	Gesamt- über- leben	Progres- sions- freies Überle- ben (PFS)	Anspre- chen (IMWG)	Studienteilnehmer mit mindestens				
				1 UE	1 SUE	1 zum Absetzen des Stu- dien-AM führen- des UE	Tod ≤30 d nach letzter Dosis Studien- AM	1 UE CTC Grad 3/4
MM-003	0,53 [0,37- 0,74]; <0,001	0,45 [0,35- 0,59]; <0,001	4,86 [2,03- 11,61]; <0,001	2,326 [0,924- 5,854]; 0,0814	1,027 [0,693- 1,521]; 0,9204	1,886 [0,840- 4,234]; 0,1454	0,884 [0,498- 1,568]; 0,6595	1,130 [0,710- 1,796]; 0,6332

HR: Hazard Ratio; OR: Odds Ratio; KI: Konfidenzintervall; IMWG: International Myeloma Working Group; UE: unerwünschtes Ereignis; SUE: schwerwiegendes unerwünschtes Ereignis; AM: Arzneimittel; CTC: Common Terminology Criteria

Ein Anhaltspunkt auf Unterschiede im Therapieeffekt zwischen verschiedenen Subgruppen liegt nicht vor.

Gemäß AM-NutzenV § 5 Absatz 7 liegt ein erheblicher Zusatznutzen u. a. vor, wenn die Überlebensdauer erheblich verlängert, eine langfristige Freiheit von Symptomen erzielt oder eine weitgehende Vermeidung schwerwiegender Nebenwirkungen gegenüber der Vergleichstherapie erreicht wird (Bundesministerium der Justiz, 2010).

Pom+LD-Dex erzielt sowohl eine in der Indikation bisher nicht erreichte, erhebliche Verlängerung des Gesamtüberlebens bei gleichzeitig gegenüber HD-Dex nicht relevant erhöhter Toxizität. Die Symptomlast der Patienten wird durch die verlängerte Progressionsfreiheit und die höheren Ansprechraten gesenkt, was sich in erhöhten Verbesserungsraten bezüglich des ECOG PS sowie der Knochenschmerzen widerspiegelt. Demgemäß zeigen die Ergebnisse der Zulassungsstudie MM-003 in Anbetracht der Relevanz der Endpunkte Mortalität, Morbidität, des Ausmaßes der Verringerung der Risiken für Progression und Tod sowie des weitgehenden Erhalts der Lebensqualität und der guten Verträglichkeit einen **erheblichen** Zusatznutzen für Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason gegenüber der alleinigen Anwendung von hoch-dosiertem Dexamethason.

Der Zusatznutzen gilt durch die Orphan Drug Zulassung gemäß § 35a Absatz 1 Satz 10 SGB V (Kommission der Europäischen Gemeinschaften, 2009, European Medicines Agency (EMA), 2013b), die Evidenzkraft der zugrunde liegenden Studie sowie durch die Güte des Datenmaterials als **belegt**. Aus der Sicht von Celgene reicht die Bindungswirkung der Orphan Drug-Ausweisung von Pomalidomid soweit, dass sie Aussagen zu einem erheblichen Zusatznutzen durch die Entscheidung der Europäischen Kommission rechtsverbindlich umfasst.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Der medizinische Zusatznutzen von Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) bei der Behandlung von Patienten mit Multiplem Myelom, die mindestens zwei Vortherapien, darunter Bortezomib und Lenalidomid, erhalten haben und die unter der letzten Therapie einen Progress gezeigt haben, soll im Vergleich zur Behandlung mit hoch-dosiertem Dexamethason (HD-Dex) als Monotherapie untersucht werden. Als Zielparameter werden patientenrelevante Endpunkte in den Bereichen *Mortalität*, *Morbidität*, *Lebensqualität* und *Sicherheit* berücksichtigt. Die Beantwortung der Fragestellung erfolgt auf Basis randomisierter kontrollierter klinischer Studien.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Tabelle 4-3: Einschlusskriterien für die Studienbewertung

Kriterium	Einschlusskriterien	Begründung
Patientenpopulation	Erwachsene Patienten mit rezidiertem oder refraktärem Multiple Myelom, die mindestens zwei Vortherapien, darunter Lenalidomid und Bortezomib, erhalten haben und unter der letzten Therapie eine Progression gezeigt haben.	Konsistent mit SmPC/ Fachinformation
Intervention	Pomalidomid (4 mg) in Kombination mit niedrig-dosiertem Dexamethason	Zieltechnologie der Nutzenbewertung
Vergleichstherapie	Hoch-dosiertes Dexamethason	Kontrollarm der Zulassungsstudie als maßgebliche Vergleichstherapie gemäß § 35a Absatz 1 Satz 10 SGB V (Bundesministerium der Justiz, 2013), § 12 Absatz 1 Satz 1 Nr. 1 VerfO 5. Kapitel (Gemeinsamer Bundesausschuss, 2011) und auf Grundlage des Beschlusses in der Sitzung des Gemeinsamen Bundesausschuss (G-BA) vom 15. März 2012 (siehe Modul 3, Abschnitt 3.1)
Endpunkt(e)*	Mortalität Morbidity Lebensqualität Sicherheit	Auswahl der Nutzendimensionen nach Vorgaben der VerfO
Studientyp	RCT	Gem. § 5 Absatz 3 VerfO; methodisch adäquat durchgeführte RCTs sind mit der geringsten Ergebnisunsicherheit behaftet
Studiendauer	Nicht eingeschränkt	Es handelt sich bei klinischen Studien in der Onkologie um ereignisgesteuerte Studien, so dass eine Mindeststudiendauer nicht praktikabel ist.

*Operationalisierungen siehe Tabelle 4-15, Tabelle 4-18, Tabelle 4-32 und Tabelle 4-38

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist in folgenden Fällen erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt:

- Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.
- Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

Sofern eine bibliografische Literaturrecherche erforderlich ist, soll diese mindestens in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

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

Für die Zusatznutzenbewertung von Pomalidomid ist keine bibliographische Literaturrecherche erforderlich, da

- 1.) die erste Zulassung für den Wirkstoff Pomalidomid weltweit weniger als zwölf Monate vor Dossiereinreichung erfolgte (USA: *Food and Drug Administration (FDA)*-Zulassung am 08.02.2013).
- 2.) keine indirekten Vergleiche für die Nutzenbewertung von Pomalidomid herangezogen werden mussten.

Folglich entfällt auch der Abschnitt 4.3.1.1.2.

4.2.3.3 Suche in Studienregistern

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

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

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

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

Für die Identifikation von relevanten Studien wurden die Studienregister *clinicaltrials.gov*, *clinicaltrialsregister.eu*, und *International Clinical Trials Registry Platform Search Portal* einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie nach abgeschlossenen, abgebrochenen und laufenden Studien durchsucht. Die Suche erfolgte ohne generelle Einschränkungen. Das Studienregister *clinicalstudyresults.org* ist nicht mehr verfügbar und wurde entsprechend nicht durchsucht.

Die detaillierte Suchstrategie und deren Ergebnisse sind in Anhang 4-B dokumentiert.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das

Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Liste der über die Suche in den Studienregistern identifizierten Studien wurde von zwei Personen unabhängig voneinander unter Berücksichtigung der in Tabelle 4-1 dargestellten Einschlusskriterien im Hinblick auf ihre Relevanz überprüft. Im ersten Schritt wurden Studien ausgeschlossen, deren Titel eindeutig eine Einstufung als *nicht relevant* zuließen. In nachfolgenden Schritten wurden ggf. die weiteren Angaben betrachtet, um festzustellen, ob die Einschlusskriterien erfüllt waren. Diskrepanzen in der Bewertung zwischen den Reviewern wurden durch Diskussion gelöst.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-G):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie des Behandlers
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Bewertung der Aussagekraft der Ergebnisse der eingeschlossenen Studien erfolgte in zwei Schritten.

Zuerst wurden methodische Charakteristika randomisierter klinisch kontrollierter Studien endpunktübergreifend **auf Studienebene** untersucht. Diese Beurteilung betraf das Studiendesign, darunter die Erzeugung der Randomisierungssequenz, die Verdeckung der Gruppenzuweisung, die Verblindung der Studienteilnehmer und Behandler. Weiterhin sind Verzerrungen bei der Berichterstattung und andere verzerrende Aspekte eingeflossen. Das Verzerrungspotenzial auf Studienebene wurde als *niedrig* oder *hoch* eingestuft. Eine Einteilung als *niedrig* wurde vorgenommen, wenn eine Verzerrung der Ergebnisse unwahrscheinlich und keine Anhaltspunkte für verzerrende Aspekte vorhanden waren, die bei Behebung die Beurteilung in einer qualitativen Weise verändern würden.

In einem nachfolgenden Schritt wurde die Aussagekraft der Ergebnisse **auf Endpunktebene** beurteilt. Hierbei wurden u. a. die Verblindung der Endpunkterheber, die Umsetzung des ITT-Prinzips, die ergebnisgesteuerte Berichterstattung der Ergebnisse und weitere potenziell verzerrende Punkte bewertet. Das Verzerrungspotenzial wurde als *niedrig* eingestuft, wenn sich keine Anhaltspunkte für Verzerrungen fanden, die die Aussage der Ergebnisse bei Behebung grundlegend verändert hätten.

Nach der Beurteilung des Verzerrungspotenzials erfolgte auf Einzelstudien- und Endpunktebene eine Kategorisierung des Ausmaßes der Ergebnissicherheit in eine hohe Ergebnissicherheit, wenn das Ergebnis aus einer randomisierten Studie mit niedrigem Verzerrungspotenzial resultiert, eine mäßige Ergebnissicherheit, wenn das Ergebnis aus einer randomisierten Studie mit hohem Verzerrungspotenzial kam und eine geringe Ergebnissicherheit, wenn die Grundlage des Ergebnisses eine nicht randomisiert vergleichende Studie war.

Im Anschluss wurde die wissenschaftliche Evidenz für jeden vordefinierten Endpunkt mit *Beleg*, *Hinweis* oder *Anhaltspunkt* bewertet, um abschließend die Beleglage für den Zusatznutzen ableiten zu können.

Tabelle 4-4: Anforderungen an die Aussage zur Beleglage

Aussage	Anforderung		
	Anzahl der Studien	Ergebnissicherheit	Effekt
Beleg	≥2	mehrheitlich hoch	gleichgerichtet
Hinweis	≥2	mehrheitlich mäßig	gleichgerichtet
	1	hoch	statistisch signifikant
Anhaltspunkt	≥2	mehrheitlich gering	gleichgerichtet
	1	mäßig	statistisch signifikant

Quelle: (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011b)

Eine einzelne Studie genügt ebenfalls für einen Nutzen*beleg*, wenn sie die besonderen Qualitätskriterien gemäß „CPMP. Points to consider on application with: 1. Meta-analysis; 2. one pivotal study“ erfüllt (European Agency for the Evaluation of Medicinal Products (EMA) - Committee for Proprietary Medicinal Products (CPMP), 2001). Dies ist auch in den IQWiG-Methoden “Allgemeine Methoden - Version 4.0” dargelegt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011b).

Hierbei ist zu beachten, dass der Zusatznutzen bereits durch die Orphan Drug Zulassung gemäß § 35a Absatz 1 Satz 10 SGB V als **belegt** gilt.

4.2.5 Informationssynthese und –analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)². Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-³ bzw. STROBE-Statements⁴ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

² Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

³ Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

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

Die Informationen der eingeschlossenen Studien wurden anhand der Items 2b bis 14 des CONSORT-Statements 2010 in Anhang 4-F beschrieben. Dazu wurden Angaben zu Studienziel, Methodik (Design, Studienteilnehmer, Interventionen, Zielkriterien, Fallzahl, Erzeugung und Geheimhaltung der Randomisierungssequenz [*Allocation Concealment*], Durchführung der Randomisierung, Verblindung, statistische Methoden) und der Ergebnisdarstellung (Patientenfluss, Anzahl Studienteilnehmer und Aufnahme/Rekrutierung) dargestellt. Der Patientenfluss wurde in einem Flow-Chart dargestellt. Die Methodik der eingeschlossenen Studien, die Interventionen sowie die Charakteristika der Studienpopulationen sind in Tabelle 4-10, Tabelle 4-11 und Tabelle 4-12 (Abschnitt 4.3.1.2.1) detailliert beschrieben.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

Patientencharakteristika

Folgende Charakteristika wurden zur Beschreibung der Studienpopulation dargestellt (Tabelle 4-12):

- Alter
- Geschlecht
- Anzahl vorangegangener Therapien
- Art der vorangegangenen Therapien

- Krankheitspopulation⁵
- Zeit seit Diagnose
- Durie-Salmon-Stadium (Schweregrad der Erkrankung, siehe Modul 3, Abschnitt 3.2.2)
- vorangegangene Stammzelltransplantation
- Refraktärität auf die letzte vorhergehende Myelomtherapie
- Therapie-Refraktärität (Thalidomid, Lenalidomid, Bortezomib)
- ECOG Performance-Status
- zytogenetische Risikokategorie

Endpunkte, zu denen Ergebnisse dargestellt werden:

Für die nachfolgenden Vergleiche werden adäquate statistische Verfahren verwendet und es wird von einem statistisch signifikanten Ergebnis gesprochen, wenn der resultierende p-Wert $<0,05$ ist. Für Häufigkeiten werden absolute und relative Häufigkeiten angegeben und die *Odds Ratio* als Effektmaß verwendet. Für stetige Messgrößen werden die Veränderungen zur *Baseline* angegeben und mit einem t-Test verglichen. Für Überlebenszeitanalysen werden Kaplan-Meier Schätzer und *Hazard Ratios* angegeben. Im Falle von konkurrierenden Ereignissen (*competing events*) erfolgt die Interpretation der Ergebnisse primär anhand des *Hazard Ratios*.

Mortalität

Gesamtüberleben

Die Gesamtüberlebenszeit gilt in der Onkologie als harter, patientenrelevanter Endpunkt.

Morbidität

Als morbiditätsbezogene klinische Effektparameter wurden *Progressionsfreies Überleben*, *Zeit bis zur Progression*, *Ansprechen* (gemäß den IMWG- und EBMT-Kriterien), *Verbesserung des ECOG Performance Status Scores* sowie *Verbesserung von Knochenschmerzen* für die Bewertung herangezogen. Auch Ergebnisse von Lebensqualitätsanalysen fließen in die Bewertung des Therapienutzens in Hinblick auf die Morbidität ein.

⁵ *Gruppe 1*: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. *Gruppe 2*: Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten nach Ende der Lenalidomid- und/oder Bortezomibtherapie. *Gruppe 3*: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥ 2 Bortezomibzyklen.

Ansprechen und Progression

Sowohl das Ansprechen (in Form einer möglichst tiefgehenden Remission) als auch eine möglichst lange Verzögerung der Krankheitsprogression sind patientenrelevante Therapieziele, da sie die Morbidität der Patienten senken und damit die Lebensqualität der Patienten bewahren oder verbessern (Böhrer und Pönisch, 2001, Smith et al., 2005). Dieser Tatsache kommt vor dem Hintergrund, dass es sich beim Multiplen Myelom um eine unheilbare Tumorerkrankung handelt, besonderes Gewicht zu. Das Multiple Myelom ist durch eine Symptomatik mit Auswirkungen u. a. auf die Nierenfunktion und die Knochenstruktur gekennzeichnet. Es geht mit Nierenschädigungen einher und kann aufgrund auftretender Knochenläsionen zu Knocheninstabilitäten, Frakturen und zu Hyperkalzämie führen. Diese Morbiditäten nehmen mit dem weiteren Fortschreiten der Erkrankung zu, so dass die Patienten einen fortschreitenden Verlust an Lebensqualität erleiden.

Der klinische Endpunkt *Ansprechen* ist ein Parameter, der Auskunft über die tumortoxische Wirksamkeit einer Substanz gibt. In der Studie MM-003 wurde das *Ansprechen* gemäß den IMWG-Kriterien, das *Ansprechen* gemäß den EBMT-Kriterien und die *Krankheitsprogression* wiederum gemäß den IMWG-Kriterien beurteilt (Bladé et al., 1998, Durie et al., 2006).

Aus diagnostischer Sicht ist der etablierte Leitparameter zur Messung des Ansprechens auf eine Therapie und des Fortschreitens der Erkrankung das M-Protein (Bladé et al., 1998). M-Proteine sind funktionslose Immunglobuline, die bei der großen Mehrheit der Patienten mit Multiplem Myelom von den malignen Plasmazellen sezerniert werden (Stringer et al., 2011). M-Proteine setzen sich aus schweren und/oder leichten Ketten zusammen. Freie Immunglobulin-Leichtketten werden glomerulär filtriert und führen in hoher Konzentration zu tubulointerstitiellen Schäden (Stringer et al., 2011, Dimopoulos et al., 2008). Dies resultiert in tubulointerstitieller Nephritis und einer Cast-Nephropathie, der sogenannten Myelom-Niere (Hutchison et al., 2012). Diese pathologischen Prozesse können zu Nierenversagen und dadurch bedingt zur Dialysepflicht führen. Bis zu 50 % der Myelom-Patienten weisen bereits bei Diagnosestellung eine Nierenschädigung auf, etwa 10 % der Patienten leiden unter einer schweren, unbehandelt letalen Ausprägung der Niereninsuffizienz (Stringer et al., 2011). Nierenversagen ist eine häufige Todesursache bei Patienten mit Multiplem Myelom (Herrera et al., 2004). Eine Reduktion der M-Proteinkonzentration, wie sie mit einem Ansprechen gemäß den IMWG- und EBMT-Kriterien einhergeht, kann also das Entstehen bzw. Fortschreiten einer myelombedingten Niereninsuffizienz verhindern oder zumindest verzögern (siehe hierzu auch Abschnitt 4.5.4). Das Erreichen einer möglichst tiefen Remission wird auch deshalb als grundsätzliches Therapieziel bei der Behandlung des Multiplen Myeloms gesehen, und Studiendaten zeigen, dass eine tiefergehende Remission in der Rezidivsituation mit einer verbesserten Prognose einhergeht (Harousseau et al., 2010, Straka et al., 2012, Anderson et al., 2008). In der Indikation des rezidierten Multiplen Myeloms zeigen Studien auch, dass bereits ein minimales Ansprechen (Minimal Response, MR) einen klinischen Nutzen für die Patienten bedeutet. So hatten in der Studie von Niesvizky et al. diejenigen Patienten, welche eine MR erreichten, ein signifikant längeres Gesamtüberleben sowie eine signifikant längere Zeit bis zur Progression als solche Patienten, die kein Ansprechen erreichten (Niesvizky et al.,

2008). Patienten mit partiellem Ansprechen hatten wiederum einen signifikant höheren klinischen Nutzen als diejenigen mit MR. Die Autoren stellen fest, dass es eine eindeutige Beziehung zwischen Ansprechen und klinischen Nutzen gibt. In einer anderen Arbeit wird geschlossen, dass das Ansprechen ein wichtiger Indikator für einen Patientennutzen bei rezidiviertem und refraktärem Multiplem Myelom ist (Anderson et al., 2008). Dies wird durch die Lebensqualitätsdaten der hier vorliegenden Studie MM-003 untermauert, die zeigt, dass ein mindestens partielles Ansprechen eine Determinante für eine Verbesserung der Lebensqualität ist (Abschnitt 4.3.1.3.1.2).

Aus klinischer Sicht manifestieren sich myelombedingte Veränderungen der Knochenstruktur durch Neuauftreten oder symptomatische Verschlechterung bereits bestehender Knochenläsionen. Die Definition des Ansprechens sieht unter anderem vor, dass das Ansprechen mit einem Stillstand der Knochensymptomatik des Multiplen Myeloms einhergeht. Bei einem klinischen Ansprechen liegt somit keine Vergrößerung oder Vermehrung lytischer Knochenläsionen vor (Bladé et al., 1998). Solche Läsionen betreffen 70-80 % aller Patienten mit Multiplem Myelom und gehen mit Knochenschmerzen, Frakturen und bei Befall der Wirbelkörper mit Lähmungen durch Nervenkompression einher (Breitkreutz et al., 2008). Skelettbezogene Ereignisse stellen eine häufige Ursache für Morbidität bei Myelom-Patienten dar und können ihre Lebensqualität erheblich reduzieren (Berenson et al., 1998, Yeh und Berenson, 2006, Costa et al., 2008). Der gesteigerte Knochenabbau kann zudem Hyperkalzämien induzieren, was wiederum zu Symptomen bis hin zu Verwirrung, Somnolenz und Koma führen kann (Ralston et al., 1990).

Ein Ansprechen auf die Therapie und eine Verzögerung der Zeit bis zur Progression, bemessen in der Reduktion der M-Proteinkonzentration, der Vermeidung von wachsenden oder neu entstehenden Knochenläsionen sowie auftretender Hyperkalzämien, sind daher patientenrelevant.

Der Endpunkt *Progressionsfreies Überleben* war der primäre Endpunkt der Zulassungsstudie MM-003 und wird hier ebenfalls vorgestellt. Hierbei handelt es sich um einen kombinierten Endpunkt, bei dem die Ereignisse „Tod“ und „Progression“ bewertet werden. Das Ereignis „Tod“ ist als solches *per definitionem* patientenrelevant. Das Ereignis *Progression* ist für den betroffenen Patienten ebenfalls schwerwiegend und patientenrelevant. So wurde das Ereignis *Progression* für die Indikation *Multiples Myelom* durch das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) im Abschlussbericht zur Nutzenbewertung der Stammzelltransplantation beim Multiplem Myelom, Seite iv, als patientenrelevantes Ereignis beschrieben (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011a). Studien untermauern dies. So stellen Migkou et al. fest, dass Patienten, die schnell rezidivieren, einen schlechteren weiteren Krankheitsverlauf aufweisen als solche Patienten, die längere progressionsfreie Intervalle aufweisen („[...] *our paper shows that patients who relapse early have a poor outcome.*“) (Migkou et al., 2011). Sogar wenn eine Progression nicht symptomatisch ist, sondern durch einen schnellen Anstieg des M-Proteins charakterisiert ist, kann Therapiebedarf bestehen, da „sich bei starker Progression der Tumormasse noch nicht vorhandene Einschränkungen von Organfunktionen schnell entwickeln können“ (Straka

et al., 2012). Weiterhin zeigte eine Studie bei Patienten mit neu-diagnostiziertem Multiplem Myelom, dass Progression ein prognostischer Faktor für eine Verschlechterung der Lebensqualität ist (Dimopoulos et al., 2011). Dies wird durch die Lebensqualitätsdaten der hier diskutierten Studie MM-003 bestätigt, anhand derer Progression ebenfalls als negativer prognostischer Faktor für Lebensqualität und das Auftreten von Krankheitssymptomen identifiziert wurde (Abschnitt 4.3.1.3.1.2). Schlussendlich wird eine Verlängerung des Progressionsfreien Überlebens bei Patienten mit rrMM als wichtiges Therapieziel genannt, insbesondere aufgrund des chronischen, durch multiple Rezidive charakterisierten Krankheitsverlaufs, der einer Heilung entgegensteht (Straka et al., 2012, Smith et al., 2005).

ECOG Performance Status-Score

Der *ECOG Performance Status* (PS) bemisst den globalen Leistungsstatus von Patienten. Dieser beschreibt die Fähigkeit von Patienten alltäglichen Aktivitäten nachzugehen. Kriterien sind die Gehfähigkeit, die Fähigkeit, körperlicher Arbeit nachzugehen und für sich selbst zu sorgen. Der ECOG PS umfasst sechs Kategorien (0 bis 5), die narrativ beschrieben werden (Oken et al., 1982).

- 0: normale uneingeschränkte Aktivität wie vor der Erkrankung
- 1: Einschränkung bei körperlicher Anstrengung, aber gehfähig; leichte körperliche Arbeit bzw. Arbeit im Sitzen (z. B. leichte Hausarbeit oder Büroarbeit) möglich.
- 2: gehfähig, Selbstversorgung möglich, aber nicht arbeitsfähig; kann mehr als 50 % der Wachzeit aufstehen.
- 3: nur begrenzte Selbstversorgung möglich; 50 % oder mehr der Wachzeit an Bett oder Stuhl gebunden.
- 4: völlig pflegebedürftig, keinerlei Selbstversorgung möglich; völlig an Bett oder Stuhl gebunden.
- 5: Tod

Aufgrund des direkten Bezuges zu den Grundbedürfnissen eines selbstbestimmten Lebens ist jegliche Verschlechterung des ECOG PS Scores von unmittelbarer Patientenrelevanz.

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität gilt *per definitionem* als patientenrelevant. Dies bestätigt zudem auch das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) im Abschlussbericht zur Nutzenbewertung der Stammzelltransplantation beim Multiplem Myelom, Seite iv, (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011a).

Die gesundheitsbezogene Lebensqualität wurde mittels des EORTC QLQ-C30, des EORTC QLQ-MY20 sowie des EuroQoL EQ-5D beurteilt. Der EORTC QLQ-C30-Fragebogen wurde zur Erhebung der Lebensqualität von Krebspatienten erstellt. Er ist etabliert und hat eine hohe Validität und Reliabilität (Aaronson et al., 1993). Der EORTC QLQ-MY20-Fragebogen wurde speziell zur Erhebung der Lebensqualität von Myelom-Patienten entworfen. Er ist in dieser Indikation etabliert und hat ebenfalls eine hohe Validität und Reliabilität (Cocks et al., 2007). Der EQ-5D ist ein standardisiertes und gebräuchliches Messinstrument, um indikationsunabhängig die Lebensqualität zu erheben. Für die Auswertung des EQ-5D wurden die deutschen Populationsgewichte verwendet (Greiner et al., 2003). Neben dem EORTC QLQ-MY20 wurden sowohl der EORTC QLQ-C30 als auch der EuroQoL EQ-5D innerhalb der Indikation des Multiplen Myeloms validiert (Wisloff et al., 1996, Kvam et al., 2011).

Sicherheit

Die Bedeutung unerwünschter Arzneimittelwirkungen ist abhängig von ihrer Schwere, der Häufigkeit, ihrer Behandelbarkeit bzw. der Reversibilität ihrer Symptome. Insbesondere schwerwiegende und irreversible unerwünschte Arzneimittelwirkungen sind Gegenstand der Entscheidung über die Verkehrsfähigkeit von Arzneimitteln im Rahmen ihrer arzneimittelrechtlichen Zulassung. Unerwünschte Arzneimittelwirkungen sind aus mehreren Gründen patientenrelevant. Zum einen kann ihr Auftreten zu einer Veränderung der initial gewählten Therapie oder ihrem Abbruch führen, womit eine Verminderung der Effektivität einer Behandlung einhergehen kann. Zum anderen können unerwünschte Arzneimittelwirkungen die Adhärenz bei der Medikamenteneinnahme vermindern und auf diese Weise mittelbar die Wahrscheinlichkeit eines Behandlungserfolgs reduzieren. Unerwünschte Arzneimittelwirkungen können zudem in Abhängigkeit von ihrer Schwere und Häufigkeit die Lebensqualität von Patienten sowohl unmittelbar als auch mittelbar beeinträchtigen (Deutsche Krebsgesellschaft (DKG) e.V., 2011). Die Lebensqualität kann dabei unmittelbar aufgrund der Symptomatik der unerwünschten Arzneimittelwirkungen beeinflusst werden, während mittelbare negative Einflüsse auf die Lebensqualität durch eine Verminderung der Effektivität der Behandlung und eine Progredienz der Grunderkrankung auftreten können. Unerwünschte Arzneimittelwirkungen sind in der Realität nur schwer von unerwünschten Ereignissen, die nicht aufgrund des Medikaments auftreten, zu unterscheiden. Deshalb ist es gängige Praxis, alle unerwünschten Ereignisse, unabhängig von ihrer möglichen Ursache, zu berichten und den Unterschied zum Kontrollarm zu interpretieren.

Folgende unerwünschten Ereignisse werden im Detail vorgestellt:

- unerwünschte Ereignisse
- schwerwiegende unerwünschte Ereignisse
- unerwünschte Ereignisse, die zum Therapieabbruch führten
- Tod \leq 30 Tage nach letzter Studienmedikation

- unerwünschte Ereignisse mit CTC Grad ≥ 3
- spezielle unerwünschte Ereignisse (CTC Grad 3/4 mit Häufigkeit ≥ 5 % sowie *Adverse Events of Special Interest* gemäß Studienprotokoll):

Neutropenie, Anämie, Thrombopenie, Leukopenie, febrile Neutropenie, Pneumonie, allgemeine körperliche Gesundheitsverschlechterung, Asthenie, Hyperkalzämie, Hyperglykämie, Knochenschmerzen, hämatopoetische Leukopenie, Infektionen, hämatopoetische Thrombopenie, akutes Nierenversagen, Diabetes mellitus, kardiale Arrhythmie, Hämorrhagie, hämodynamisches Ödem, venöses thromboembolisches Ereignis, Depression und Suizid, Muskelschwäche, periphere Neuropathie, arterielles thromboembolisches Ereignis, hämatopoetische Zytopenie bei >1 Zelltyp, unspezifisches gefäßassoziiertes arterielles und venöses thromboembolisches Ereignis, Linsenerkrankung

- Sekundäre Primärmalignome (SPM)

Für die Endpunkte *Mortalität*, *Morbidität* und *Lebensqualität* werden alle randomisierten Patienten in den Analysen dargestellt (ITT-Population). Für den Endpunkt *Sicherheit* werden alle Patienten in den Analysen dargestellt, die mindestens eine Studiendosis erhalten haben (Safety-Population). Dies entspricht dem prä-spezifizierten Vorgehen im Studienbericht der verwendeten Studie MM-003.

4.2.5.3 Meta-Analysen

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

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁶ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁷ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁸ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Es wurden keine Meta-Analysen durchgeführt, da nur eine Studie eingeschlossen wurde.

⁶ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

⁷ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

⁸ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

4.2.5.4 Sensitivitätsanalysen

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

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

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

In der dargestellten Studie MM-003 wurde für den Endpunkt *Progressionsfreies Überleben* (PFS) eine Sensitivitätsanalyse mit einem alternativen Zensierungsalgorithmus durchgeführt. Die Originalanalyse basierte auf einer Leitlinie der *European Medicines Agency* (EMA) (European Medicines Agency (EMA), 2008). Die Sensitivitätsanalyse basierte auf einer Leitlinie der FDA (U.S. Department of Health and Human Services - Food and Drug Administration et al., 2007). Die Unterschiede zwischen den beiden Zensierungsmethoden sind in folgender Tabelle 4-5 dargestellt.

Tabelle 4-5: Zensierungsmethodik bei dem Endpunkt PFS

Szenario	Primäre Analyse (EMA)	Sensitivitätsanalyse (FDA)
Progression wurde zwischen geplanten Bewertungszeitpunkten (Studienvisiten) dokumentiert	Datum der Progression als Ereigniszeitpunkt	Letzter geplanter Bewertungszeitpunkt als Ereigniszeitpunkt. Falls kein adäquater Bewertungszeitpunkt existiert wird ein Tag nach Randomisierung als Ereigniszeitpunkt verwendet.
Neue Myelomtherapie oder verbotene Behandlung (laut Studienprotokoll) beginnt vor der Progression	Datum der Progression als Ereigniszeitpunkt	Zensiert zum letzten adäquaten Bewertungszeitpunkt mit <i>stable disease</i> vor der neuen Behandlung. Falls kein adäquater Bewertungszeitpunkt vorliegt wird der Patient zum Zeitpunkt der Randomisierung zensiert.
Tod oder Progression während der Behandlungsphase nach einem Zeitraum mit mindestens zwei nicht durchgeführten Bewertungszeitpunkten (Studienvisiten)	Datum des Todes oder der Progression als Ereigniszeitpunkt	Zensiert zum letzten adäquaten Bewertungszeitpunkt mit <i>stable disease</i> vor der neuen Behandlung. Falls kein adäquater Bewertungszeitpunkt vorliegt wird der Patient zum Zeitpunkt der Randomisierung zensiert.

Eine weitere Sensitivitätsanalyse für den primären Endpunkt PFS untersuchte die Zeit bis zum Therapieversagen (*Time to Treatment Failure*, TTF). Hier wurde die Zeitspanne bis zum ersten Eintreten eines der folgenden Ereignisse analysiert:

- Tod
- Progression gemäß den IMWG- oder EBMT-Kriterien (festgelegt durch den behandelnden Arzt und bestätigt durch das *Independent Response Adjudication Committee* (IRAC))
- Studienabbruch
- Beginn einer neuen Myelomtherapie

Diese Analysen wurden durchgeführt, um die Robustheit der Ergebnisse des primären Endpunkts PFS zu untersuchen.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

Sollten sich aus den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Um mögliche Unterschiede im Therapieeffekt in Subgruppen zu überprüfen, wurden die Ergebnisse hinsichtlich potenzieller Effektmodifikation untersucht.

Alle im Studienprotokoll präspezifizierten Subgruppen wurden untersucht:

- Alter (≤ 75 Jahre vs. > 75 Jahre),
- Krankheitspopulation⁹,
- Geschlecht,
- Anzahl vorheriger Myelomtherapien (2 vs. > 2) und
- zytogenetische Risikogruppe (High risk vs. Non-high risk).

Um zusätzlich einen Einfluss des Schweregrads der Erkrankung zu untersuchen, wurde weiterhin eine Subgruppenanalyse nach

- Durie-Salmon-Stadium (I vs. II vs. III)

durchgeführt. Die Trennpunkte für die quantitativen Merkmale wurden entsprechend der Spezifikationen im Studienprotokoll gewählt. Im Studienprotokoll wurden zudem Analysen nach dem Ansprechen-Status am Ende des Zyklus 2 spezifiziert. Da es sich um keine Subgruppe handelt, die zu Beginn der Therapie festgestellt werden kann, werden diese Analysen nicht im Dossier dargestellt.

Ein Interaktionstest zwischen der Behandlung und dem subgruppenbildenden Faktor wurde durchgeführt. Hierfür wurde entsprechend des Endpunktes ein logistisches bzw. Cox Modell mit den festen Effekten Behandlungsgruppe, dem subgruppenbildenden Faktor und der Interaktion Behandlung*subgruppenbildender Faktor durchgeführt. Ein zum Niveau $\alpha = 0,05$ signifikantes Ergebnis des Interaktionsterms wurde entsprechend der IQWiG Methoden 4.0 (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011b) als Beleg

⁹ Gruppe 1: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. Gruppe 2: Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten nach Ende der Lenalidomid- und/oder Bortezomibtherapie. Gruppe 3: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥ 2 Bortezomibzyklen.

unterschiedlicher Effekte, ein zum Niveau $\alpha = 0,20$ signifikantes Ergebnis wird als Hinweis auf unterschiedliche Effekte in den Gruppen gewertet. Im Falle eines p-Wertes von kleiner 0,2 wurde der Endpunkt im Detail für die entsprechenden Subgruppen dargestellt.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich¹⁰. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison (MTC)-Meta-Analysen“¹¹, „Multiple-Treatment-Meta-Analysen“¹² oder auch „Netzwerk-Meta-Analysen“¹³, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹⁴.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹⁵.

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

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

¹⁰ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

¹¹ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹² Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹³ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

¹⁴ B. Schöttker, D. Lüthmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹⁵ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Der Vergleichsarm der Zulassungsstudie ist die zu verwendende Vergleichstherapie und daher wurden keine indirekten Vergleiche durchgeführt.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

Benennen Sie nur randomisierte, kontrollierte Studien, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

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

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen/abgebrochen/laufend)	Studiendauer	Therapiearme
MM-003	ja	laufend	03/2011- Nachbeobachtung noch andauernd	a) Pom+LD-Dex b) HD-Dex
2009-013319-36, 2009_28/0924	nein	laufend	10/2009- vsl.08/2013	a) Pom an 21 d eines 28 d Zyklus+Dex b) Pom an 28 d eines 28 d Zyklus+Dex
MCC-16705, PO-MM-PI-0039	nein	laufend	11/2008- vsl.10/2013	a) Pom+Dex b) Pom+Dex+ Cyclophosphamid
HIC 1011007607	nein	laufend	06/2011- vsl.11/2014	a) Pom an d 1-28 eines 28 d Zyklus +Dex b) Pom an d 1-21 eines 28 d Zyklus +Dex
MM-002	nein	laufend	05/2008- vsl.09/2015	a) Pom+Dex b) Pom
12-128	nein	laufend	12/2012- vsl.12/2015	a) Pom+Dex+ Stammzelltrans- plantation b) Pom+Dex
CC-4047-MM-007	nein	laufend	12/2012- vsl.01/2017	a) Pom+Dex+ Bortezomib b) Dex+Bortezomib

d: Tag(e); vsl. voraussichtlich

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

Stand der Informationen: 22.07.2013

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

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

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
2009-013319-36, 2009_28/0924 (Leleu et al., 2013)	<u>Intervention/Vergleichstherapie</u> (Pom an 21 d vs. Pom an 28 d) entspricht nicht den Einschlusskriterien
MCC-16705, PO-MM-PI-0039	<u>Intervention/Vergleichstherapie</u> (Pom+Dex vs. Pom+Dex+Cyclophosphamid) entspricht nicht den Einschlusskriterien
HIC 1011007607	<u>Intervention/Vergleichstherapie</u> (Pom d 1-28 vs. Pom d 1-21) entspricht nicht den Einschlusskriterien
MM-002 (Richardson et al., 2013)	<u>Intervention/Vergleichstherapie</u> (Pom+Dex vs. Pom) entspricht nicht den Einschlusskriterien
12-128	<u>Intervention/Vergleichstherapie</u> (Pom+Dex+Stammzelltransplantation vs. Pom+Dex) entspricht nicht den Einschlusskriterien
CC-4047-MM-007	<u>Intervention/Vergleichstherapie</u> (Pom+Dex+Bortezomib vs. Dex+Bortezomib) entspricht nicht den Einschlusskriterien

Die Studien 2009-013319-36 (IFM-2009-02) und MM-002 sind weder Zulassungsstudien, noch entsprechen sie den Einschlusskriterien für die Einbeziehung von RCT in die Nutzenbewertung. Sie wurden jedoch von der EMA im *CHMP Assessment Report* als unterstützende Studien (*supportive studies*) diskutiert (European Medicines Agency (EMA), 2013a, Leleu, 2012, Backstrom und Celgene Corporation, 2012).

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

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

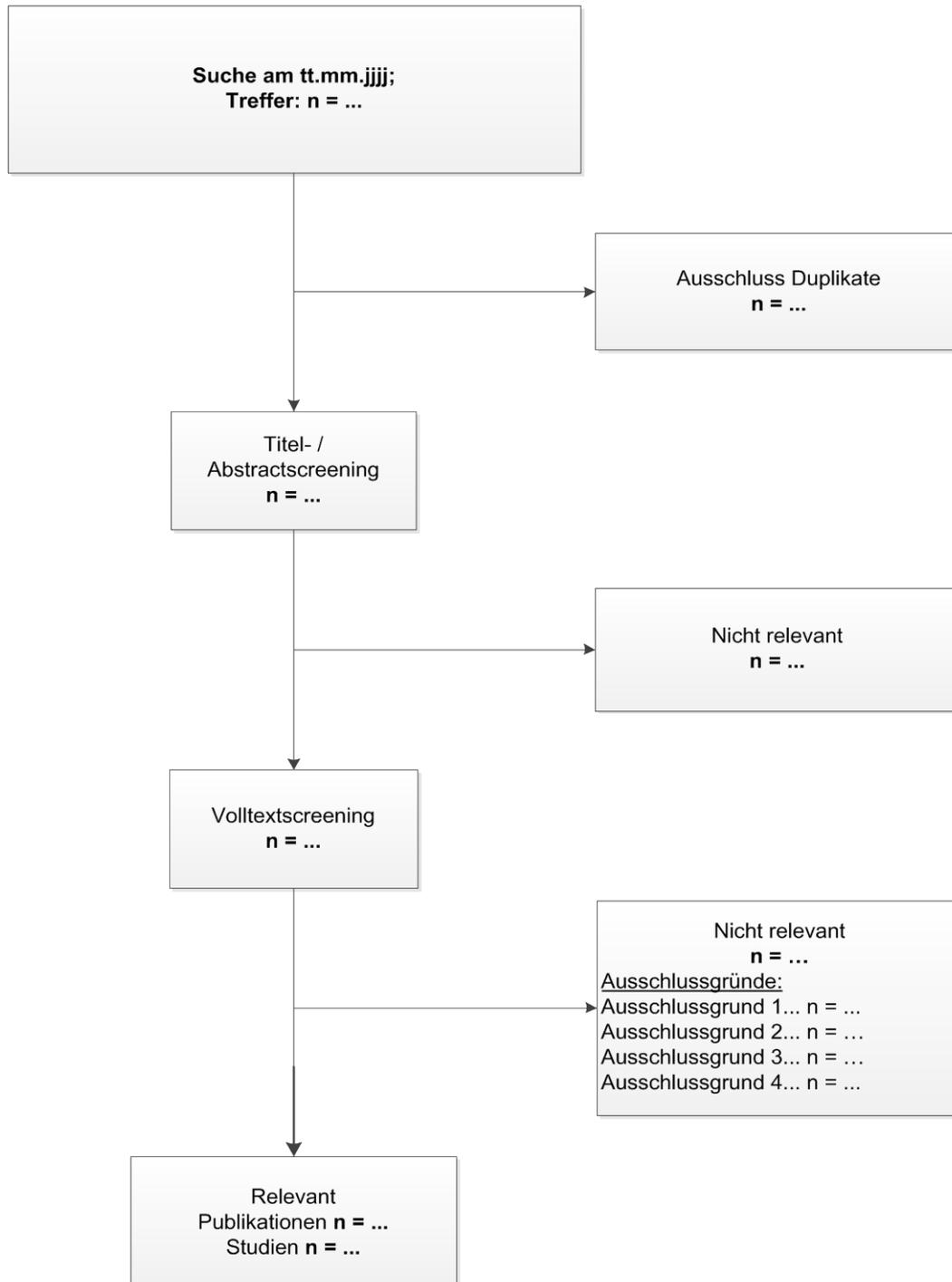


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Dieser Abschnitt entfällt, da für die Zusatznutzenbewertung von Pomalidomid keine bibliographische Literaturrecherche erforderlich ist (Abschnitt 4.2.3.2).

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (Tabelle 4-6) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

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

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

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

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein/n. d. ^b)
MM-003	<u>Clinicaltrials.gov:</u> NCT01311687 <u>Clinicaltrialsregister.eu:</u> EudraCT: 2010-019820-30 ICTRP: NCT01311687 EUCTR2010-019820-30-BE	ja	n. d.
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, bei-

des). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

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

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponser-te Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studien-bericht (ja/nein [Zitat])	Register-eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebokontrolliert						
	n. a.					
aktivkontrolliert, Vergleichstherapie(n)						
MM-003	ja	ja	nein	ja (Jacques und Celgene Corporation, 2013)	ja <u>Clinicaltrials.gov</u> : ^d NCT01311687 <u>Clinicaltrialsregister.eu</u> : ^d EudraCT: 2010-019820-30 <u>ICTRP</u> : ^d NCT01311687 EUCTR 2010-019820-30-BE	Publikation zur Veröffentlichung akzeptiert* (San Miguel et al., 2013)
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. d: Keine weiteren Zitate in den Registereinträgen vorhanden * Die Publikation der Studie MM-003 wurde von <i>Lancet Oncology</i> zur Veröffentlichung akzeptiert, der Veröffentlichungszeitpunkt ist noch unbekannt.</p>						

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
MM-003	RCT, offen, parallel, geplante Interimsanalyse (<i>Futility-Analyse</i>) für PFS nach Tod oder PD bei 50 % der geplanten Ereignisse (=121); geplante Interimsanalyse für OS bei 106 Todesfällen (und frühestens zum Zeitpunkt der finalen PFS-Analyse).	≥18 Jahre mit refraktärem oder rezidivierendem und refraktärem MM, die mind. 2 Myelom-Therapien (inkl. Lenalidomid und Bortezomib) erhalten haben; ECOG PS ≤2	Pom+LD-Dex (n=302) HD-Dex (n=153)	Ereignisgesteuert (Behandlung bis Krankheitsprogression oder Studienabbruch) Nachbeobachtung bis zu 5 Jahre nach Randomisierung	93 Zentren in Europa (58), Australien (10), UK (10), Kanada (10), Russland (4), USA (1) 03/2011-09/2012	Primärer Endpunkt: PFS* Sekundäre Endpunkte: OS*, Lebensqualität, TTP, Ansprechrate gemäß den IMWG-Kriterien, Ansprechrate gemäß den EBMT-Kriterien, Klinisches Ansprechen, Sicherheit

*MM-003 war auf PFS und OS gepowert (d. h. auch OS wurde konfirmatorisch unter Einhaltung des globalen 5 % Signifikanzniveaus getestet)

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

Studie	Pom+LD-Dex	HD-Dex	<i>ggf. weitere Spalten mit Behandlungscharakteristika</i> <i>z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
MM-003	Zyklen von 28 d <u>Teilnehmer ≤75 J:</u> 4 mg Pom p. o., qd (d 1-21) + 40 mg Dex p. o., qd (d 1,8,15 und 22) <u>Teilnehmer >75 J:</u> 4 mg Pom p. o., qd (d 1-21) + 20 mg Dex p. o., qd (d 1,8,15 und 22)	Zyklen von 28 d: <u>Teilnehmer ≤75 J:</u> 40 mg Dex p. o., qd (d 1-4, d 9-12 und d 17-20) <u>Teilnehmer >75 J:</u> 20 mg Dex p. o., qd (d 1-4, d 9-12 und d 17-20)	Thromboseprophylaxe mit Acetylsalicylsäure, niedermolekularem Heparin oder äquivalenter Medikation bei Teilnehmern im Interventionsarm und allen Teilnehmern mit tiefen Venenthrombosen oder Lungenembolien in der Vorgeschichte und nach Ermessen des Prüfarztes Gabe anderer Myelomtherapien, chronische Anwendung von Steroiden und Immunsuppressiva verboten Anwendung von Strahlentherapie zur Behandlung von pathologischen Bruchstellen oder Schmerzen nach Frakturen erlaubt Vermeiden von korrigierter QT-Zeit-verlängernden Arzneimitteln

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Arm	N	Alter (Jahre)	Geschlecht w/m (%)	<i>ggf. weitere Spalten mit Populationscharakteristika</i> <i>z. B. Dauer der Erkrankung, Schweregrad, weitere Basisdaten projektabhängig</i>
MM-003 Pom+LD-Dex	302	63,6 ± 9,33 (MW+STD)	w: 40,1 m: 59,9	Alter (%): ≤75 J: 92,1 >75 J: 7,9 Anzahl vorangegangener Therapien (median): 5,0 Vorangegangene Therapien (%): Kortikosteroide: 100,0 Dexamethason: 97,4 IMiDs [®] : 99,7 Lenalidomid: 99,7 Thalidomid: 57,3 Proteasom-Inhibitoren: 99,7 Bortezomib: 99,7 Alkylanzien: 99,0 Anzahl vorangegangener Therapien (%): 2: 5,6 >2: 94,4

Studie Arm	N	Alter (Jahre)	Geschlecht w/m (%)	ggf. weitere Spalten mit Populationscharakteristika z. B. Dauer der Erkrankung, Schweregrad, weitere Basisdaten projektabhängig
				<p>Krankheitspopulation* (%): Gruppe 1: 82,5 Gruppe 2: 2,6 Gruppe 3: 14,9</p> <p>Zeit seit Diagnose (Jahre, median): 5,3</p> <p>Durie-Salmon-Stadium (%): I: 7,0 II: 31,5 III: 58,6 Fehlende Daten: 3,0</p> <p>Vorangegangene Stammzelltransplantation (%): 70,9</p> <p>Therapie-Refraktärität (%): Thalidomid: 29,8 Lenalidomid: 93,0 Thalidomid und Lenalidomid: 27,8 Bortezomib: 78,1 Lenalidomid und Bortezomib: 73,2</p> <p>Thalidomid, Lenalidomid und Bortezomib: 23,2</p> <p>ECOG PS (%): 0: 36,4 1: 45,7 2: 17,2 3: 0,0 Fehlende Daten: 0,7</p> <p>Zytogenetische Risiko-Kategorie (%): <i>High Risk**</i>: 43,0 <i>(Modified High Risk***)</i>: 25,5 <i>Non-high Risk</i>: 30,1</p>
MM-003 HD-Dex	153	63,7 ± 9,56 (MW+STD)	w: 43,1 m: 56,9	<p>Alter (%): ≤75 J: 92,2 >75 J: 7,8</p> <p>Anzahl vorangegangener Therapien (median): 5,0</p> <p>Vorangegangene Therapien (%): Kortikosteroide: 100,0 Dexamethason: 99,3 IMiDs[®]: 99,3 Lenalidomid: 99,3 Thalidomid: 60,8 Proteasom-Inhibitoren: 100,0 Bortezomib: 100,0 Alkylanzien: 98,0</p>

Studie Arm	N	Alter (Jahre)	Geschlecht w/m (%)	ggf. weitere Spalten mit Populationscharakteristika z. B. Dauer der Erkrankung, Schweregrad, weitere Basisdaten projektabhängig
				<p>Anzahl vorangegangener Therapien (%): 2: 5,2 >2: 94,8</p> <p>Krankheitspopulation* (%): Gruppe 1: 81,7 Gruppe 2: 3,3 Gruppe 3: 15,0</p> <p>Zeit seit Diagnose (Jahre, median): 6,1</p> <p>Durie-Salmon-Stadium (%): Stadium I: 7,8 Stadium II: 24,2 Stadium III: 67,3 Fehlende Daten: 0,7</p> <p>Vorangegangene Stammzelltransplantation (%): 69,3</p> <p>Therapie-Refraktärität (%): Thalidomid: 32,0 Lenalidomid: 90,2 Thalidomid und Lenalidomid: 29,4 Bortezomib: 77,1 Lenalidomid und Bortezomib: 70,6 Thalidomid, Lenalidomid und Bortezomib: 23,5</p> <p>ECOG PS (%): 0: 23,5 1: 56,2 2: 16,3 3: 2,0 Fehlende Daten: 2,0</p> <p>Zytogenetische Risiko-Kategorie (%): <i>High Risk**</i>: 37,3 <i>(Modified High Risk***)</i>: 22,9 <i>Non-high Risk</i>: 30,7</p>

Anmerkung: Aufgrund mitunter fehlender Angaben/Werte addieren sich nicht alle Prozentwerte zu 100 % (z.B. Zytogenetik)

*Gruppe 1: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. Gruppe 2: Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten. Gruppe 3: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach \geq 2 Bortezomibzyklen.

**Jegliche zytogenetische Anomalie in 13q14, 17p13, 4p16/14q32 oder 14q32/16q23

***Jegliche zytogenetische Anomalie in 17p13 oder 4p16/14q32 (Untergruppe von High Risk)

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

Die Studie MM-003 wurde in einem randomisierten, offenen und multizentrischen Design durchgeführt. Es wurden Studienteilnehmer mit einem Lebensalter von mindestens 18 Jahren mit refraktärem oder rezidiviertem und refraktärem Multiplem Myelom eingeschlossen, die mindestens zwei Vortherapien erhalten hatten (inklusive Lenalidomid- und Bortezomib-basierter Behandlungen). Studienteilnehmer mussten einen ECOG Performance Score von 0, 1 oder 2 aufweisen. Das Randomisierungsverhältnis der beiden Studienarme war 2:1. Die Studienpopulation wies einen höheren Anteil an männlichen Teilnehmern auf. Das durchschnittliche Lebensalter (\pm STD) betrug für den Pom+LD-Dex-Arm 63,6 Jahre (\pm 9,33) bzw. für die Kontrollgruppe 63,7 Jahre (\pm 9,56). Die Länge der Behandlungszyklen betrug 28 Tage; die Gabe der Studienarzneimittel erfolgte bis zur Krankheitsprogression oder Studienabbruch. Studienteilnehmer der Kontrollgruppe hatten bei Krankheitsprogression die Möglichkeit in eine Begleitstudie zu wechseln und mit Pomalidomid als Monotherapie behandelt zu werden. Als primärer Endpunkt wurde *Progressionsfreies Überleben* (PFS) untersucht, die Studie war jedoch zudem auf den Endpunkt *Gesamtüberleben* (OS) gepowert (dies bedeutet, auch OS wurde konfirmatorisch unter Einhaltung des globalen 5 % Signifikanzniveaus getestet). Als sekundäre Endpunkte wurden zusätzlich zum *Gesamtüberleben* auch *Zeit bis zur Progression*, *Ansprechrate gemäß den IMWG-Kriterien*, *Ansprechrate gemäß den EBMT-Kriterien*, *Klinisches Ansprechen*, *Lebensqualität* und *Sicherheit* erhoben. Der Stichtag für die Datenauswertung war der 7. September 2012, weil zu diesem Zeitpunkt die benötigte Anzahl von Ereignissen zur Analyse des primären Endpunktes PFS aufgetreten war. Die Daten dieser Auswertung bilden die Grundlage für die Zulassung durch die EMA. Detaillierte Angaben zur Operationalisierung der Endpunkte und zu statistischen Methoden, die zur Auswertung der Ergebnisse verwendet wurden, sind in Tabelle 4-15, Tabelle 4-18, Tabelle 4-32, Tabelle 4-38, und Tabelle 4-61 aufgeführt.

Am 7. September 2012 erfolgte ebenfalls die OS Interimsanalyse. Aufgrund dieser erfolgreichen Interimsanalyse empfahl das IDMC die Behandlung der Patienten des Kontrollarms mit Pomalidomid fortzusetzen. Die ursprünglich geplante finale Auswertung des Gesamtüberlebens erfolgte am 1. März 2013. Im Zuge dessen wurden ebenfalls das Progressionsfreie Überleben, die Zeit bis zur Progression, die Ansprechrate gemäß den IMWG- und EBMT-Kriterien (letzteres nur für die Remissionskategorie *minimales Ansprechen*) und die Sicherheit ausgewertet. Diese Daten finden sich nicht im Studienbericht, werden jedoch unterstützend in diesem Dossier berichtet. Sie werden im Rahmen der von *Lancet Oncology* zur Veröffentlichung akzeptierten Publikation dargestellt (San Miguel et al., 2013). Die Auswertung der Daten zum 1. März 2013 müssen unter Berücksichtigung des ausgeprägten *cross-overs* (50,0 % der Patienten im LD-Dex-Arm erhielten bereits Pomalidomid) interpretiert werden.

Die EMA bestätigte im CHMP Assessment Report, dass die Auswahl der Endpunkte und statistischen Methoden in der MM-003 adäquat sind, um einen Nutzen von Pom+LD-Dex gegen den angemessenen Komparator zu zeigen („*The phase 3 study CC-4047-MM-003 was adequately designed to demonstrate the efficacy of pomalidomide in combination with low dose*

dexamethasone against an appropriate comparator (HD-Dex) in the proposed RRMM target population”) (European Medicines Agency (EMA), 2013a).

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
MM-003	ja	ja	nein	nein	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der Studie MM-003 erfolgte die Randomisierung mit computergestützter Gruppenzuteilung. Die Studie wurde offen durchgeführt. Eine Verzerrung durch Kenntnis der Gruppenzugehörigkeit wird nicht auf Studienebene diskutiert, sondern separat für jeden Endpunkt, da die Endpunkte nicht in gleichem Maße von einem offenen Studiendesign betroffen sind. Der Patientenfluss ist transparent und nachvollziehbar dargestellt. Es lassen sich auf Studienebene keine Anzeichen ergebnisgesteuerter Berichterstattung oder sonstige Hinweise auf verzerrende Aspekte identifizieren. Darüber hinaus finden sich keine Anhaltspunkte für eine inhaltliche oder formal verzerrende Darstellung der Ergebnisse. Das Verzerrungspotenzial auf Studienebene für MM-003 ist daher niedrig. Die Studie ist, falls keine endpunktspezifischen Verzerrungsaspekte dagegen sprechen, geeignet Ergebnisse mit hoher Ergebnissicherheit zu liefern.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mor- talität	Morbidity					Lebens- qualität	Sicherheit
		PFS	TTP	Ansprech- rate gemäß den IMWG- Kriterien	Ansprech- rate gemäß den EBMT- Kriterien	Klinisches Ansprechen		
MM-003	ja	ja	ja	ja	ja	ja	ja	ja

4.3.1.3.1 Endpunkt – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

4.3.1.3.1.1 Mortalität (Gesamtüberleben) – RCT

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

Tabelle 4-15: Operationalisierung von Mortalität

Studie	Operationalisierung
MM-003	Gesamtüberleben: Zeitraum von Randomisierung bis Tod jeglicher Ursache; Zensierung von Studienteilnehmern, die zum Zeitpunkt der Analyse nicht verstorben waren oder die aus den Studien ausgeschieden sind, bevor der Tod dokumentiert wurde

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

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

Tabelle 4-16: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	nein	ja	nein	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Daten zum Gesamtüberleben wurden unverblindet erhoben. Der Eintritt des Todes jedweder Ursache ist ein objektives Kriterium, dessen Feststellung frei von Subjektivität erfolgt und dementsprechend keiner Verblindung bedarf. Eine unverblindete Erhebung dieses Endpunktes hat somit keine Auswirkungen auf das Verzerrungspotenzial. Es existieren keine Hinweise für eine ergebnisgesteuerte Berichterstattung der Ergebnisse. Die Auswertung der Daten erfolgte konservativ nach dem ITT-Prinzip. Studienteilnehmer der Kontrollgruppe hatten bei Krankheitsprogression die Möglichkeit in eine Begleitstudie zu wechseln (*cross-over*) und mit Pomalidomid als Monotherapie behandelt zu werden. Die Unterbindung eines *cross-overs* ist aus ethischen Gründen nicht vertretbar. Der Anteil Studienteilnehmer, der zu einer Pomalidomid-basierten Therapie gewechselt ist, lag zum 7. September 2012 bei 29 % (1. März 2013: 50 %), so dass hierdurch ein erhöhtes Verzerrungspotenzial vorliegt. Die Verzerrung ist jedoch zuungunsten des Pom+LD-Dex-Arms, da hiermit eine Überschätzung des

Therapieeffektes im Kontrollarm einhergeht und somit der Therapievorteil von Pomalidomid unterschätzt wird. Aus diesen Gründen ist von einer hohen Ergebnissicherheit auszugehen.

Zum Zeitpunkt der OS Analyse (7. September 2012) lag die mediane Nachbeobachtungszeit im Pom+LD-Dex Arm bei 19,2 Wochen und im HD-Dex Arm bei 16,3 Wochen.

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

Tabelle 4-17: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	Gesamtüberleben (Wochen, median)	HR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	302	n. e.	0,53 [0,37-0,74]; <0,001
	HD-Dex	153	34,0	

* p-Wert basierend auf nicht stratifiziertem *Log-Rank-Test*;
n. e.: nicht erreicht; HR: *Hazard Ratio*

Für die Kombination von Pomalidomid mit LD-Dexamethason wurde eine statistisch signifikante und therapeutisch relevante Verlängerung des Gesamtüberlebens gezeigt. Im Interventionsarm traten dabei bis zum Zeitpunkt der Analyse eine nicht ausreichende Anzahl von Ereignissen auf, um das mediane Überleben berechnen zu können (nicht erreicht vs. 34,0 Wochen). Es wurde jedoch auf mindestens 48 Wochen veranschlagt, da dies die untere Grenze des Konfidenzintervalls bildete. Die Behandlung mit Pom+LD-Dex war im Vergleich zu einer Therapie mit HD-Dex mit einer Verringerung des Sterberisikos von 47 % verbunden (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$).

Unter Berücksichtigung des *cross-overs* (siehe oben) wird der tatsächliche Behandlungseffekt von Pom+LD-Dex im Vergleich zu HD-Dex unterschätzt. Dies bedeutet, dass das Sterberisiko vermutlich noch stärker vermindert und die Überlebenszeit noch stärker verlängert wird. Um dies zu bewerten wurde ein *rank preserving structural failure time* (RPSFT) Modell nach Robins und Tsiatis (1991) gerechnet (Robins und Tsiatis, 1991). Dieses Modell berechnet die Überlebenszeiten, die ohne *cross-over* aufgetreten wären. Hierdurch verkürzte sich der Median im HD-Dex-Arm auf 26,8 Wochen und die *Hazard Ratio* verbessert sich auf 0,477 [95 %-KI]: [0,328; 0,686] (Celgene Europe Limited, 2013b). Dies bedeutet eine Verringerung des Sterberisikos von mehr als 50 %. In Abbildung 4-2 ist die Überlebenswahrscheinlichkeit für den Pom+LD-Dex-Arm und den HD-Dex-Arm (adjustiert für *cross-over* und unadjustiert) dargestellt.

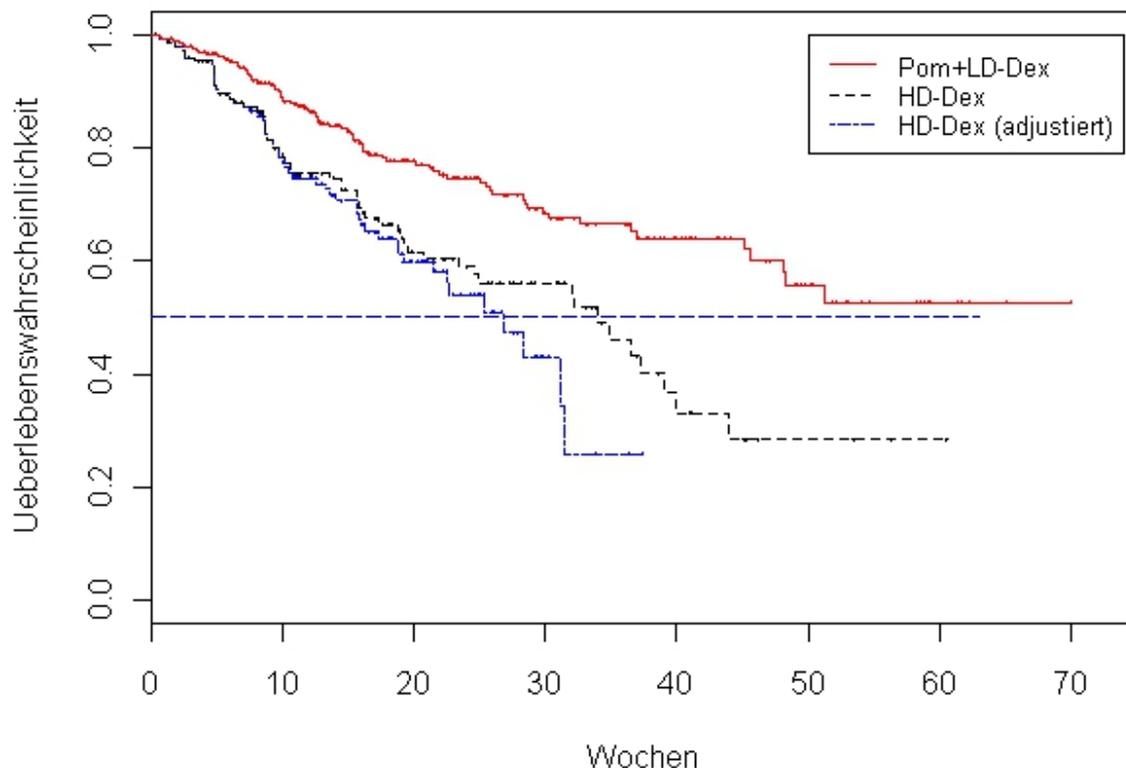


Abbildung 4-2: Kaplan-Meier Kurve für das Gesamtüberleben (7. September 2012), inkl. des HD-Dex-Arms, der für den Effekt des *cross-overs* adjustiert wurde (Celgene Europe Limited, 2013b)

Die hier gezeigte Verlängerung des Gesamtüberlebens ist auch vor dem Hintergrund der besonderen Therapiesituation (schwer vorbehandelte, oft therapierefraktäre Patienten) bemerkenswert: Pomalidomid verbessert die Prognose dieser Patienten deutlich.

Das oben dargestellte Ergebnis beruht auf der Interimsanalyse des Gesamtüberlebens, die bereits ein signifikantes Ergebnis zeigte und zu einem vorzeitigen erfolgreichen Ende der Studie führte (Anmerkung: die Nachbeobachtung der Patienten dauert noch an). Das *Independent Data Monitoring Committee* (IDMC) empfahl zu diesem Zeitpunkt die Beendigung der Studie und die Behandlung der Patienten des Kontrollarms mit Pomalidomid (Nachbeobachtung der Patienten dauert an). Die Bedingungen, die zu einem potentiellen vorzeitigen erfolgreichen Ende der Studie führen konnten, waren vorher im Studienprotokoll festgelegt.

Zusätzlich wurde das Gesamtüberleben auch zum Zeitpunkt der ursprünglich geplanten finalen OS-Analyse ausgewertet (1. März 2013). Hier wurde trotz fortgeschrittenem *cross-over* (bereits 50 % der Patienten im HD-Dex-Arm wurden mit Pomalidomid behandelt) weiterhin ein signifikanter Vorteil von Pom+LD-Dex mit einem medianen Gesamtüberleben von 54,6 Wochen im Pom+LD-Dex-Arm und 34,8 Wochen im HD-Dex-Arm (HR [95 %-KI]:

0,74 [0,56-0,97]; $p < 0,028$) beobachtet. Dies bedeutet einen Überlebensvorteil von 4,6 Monaten (19,8 Wochen).

Eine Adjustierung für das *cross-over* anhand eines RPSFT Modells für diesen Datenstand ergab einen Median im HD-Dex-Arm von 29,3 Wochen und die *Hazard Ratio* verbesserte sich auf 0,489 (95 %-KI: 0,329; 1,001) (Celgene Europe Limited, 2013b). Dies bedeutet weiterhin eine Verringerung des Sterberisikos von mehr als 50 %. Hier ergibt sich durch die längere Nachbeobachtungszeit ein deutlich größerer *cross-over* Anteil, was dazu führt, dass das Ergebnis mit einer höheren Unsicherheit behaftet ist (breiteres Konfidenzintervall). Eine robustere RPSFT Analyse ist die Analyse auf dem Datenstand zum Zeitpunkt der finalen PFS-Analyse (7. September 2012).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

4.3.1.3.1.2 Morbidität – RCT

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

Tabelle 4-18: Operationalisierung von Morbidität

Studie	Operationalisierung
MM-003	<p><i>PFS</i>: Zeitspanne von Randomisierung bis zum Tod oder zur Krankheitsprogression, festgelegt durch das IRAC (<i>Independent Response Adjudication Committee</i>); Definition der Krankheitsprogression gemäß der IMWG-Kriterien (<i>Durie et al., 2006</i>)</p> <p>Mind. eines der folgenden Kriterien:</p> <ul style="list-style-type: none"> - ≥ 25 % Anstieg der Serum-M-Komponente (absoluter Anstieg $\geq 0,5$ g/dl) und/oder - ≥ 25 % Anstieg der Urin-M-Komponente (absoluter Anstieg ≥ 200 mg/24 h) und/oder - Nur bei Patienten ohne messbares M-Protein in Serum und Urin: Differenz zwischen beteiligten und unbeteiligten freien Leichtketten-Werten (absoluter Anstieg > 10 mg/dl) - Absoluter Anstieg der Plasmazellen im Knochenmark ≥ 10 % - Entstehung neuer oder eindeutige Vergrößerung bestehender Knochenläsionen oder Weichteilplasmozytome - Entwicklung einer Hyperkalzämie (korrigierter Serum Kalzium-Spiegel $> 11,5$ mg/dl oder $2,65$ mmol/l), die nicht anderen Gründen zugeschrieben werden kann <p><i>TTP</i>: Zeitspanne von Randomisierung bis zur ersten dokumentierten Krankheitsprogression (bestätigt durch das IRAC)</p> <p><i>Ansprechrate gemäß IMWG</i>: basierend auf den Kriterien der IMWG (<i>Durie et al., 2006</i>) (festgelegt durch das IRAC)</p> <ul style="list-style-type: none"> • sCR: CR-Kriterien plus

Studie	Operationalisierung
	<ul style="list-style-type: none"> - Normales Verhältnis freier Leichtketten und - Abwesenheit von klonalen Plasmazellen im Knochenmark, nachgewiesen durch Immunhistochemie oder Immunfluoreszenz • CR: <ul style="list-style-type: none"> - Negative Immunfixation in Serum und Urin und - Verschwinden jeglicher Weichteilplasmozytome und - ≤ 5 % Plasmazellen im Knochenmark • VGPR: <ul style="list-style-type: none"> - M-Protein in Serum und Urin durch Immunfixation, aber nicht durch Elektrophorese nachweisbar oder Reduktion des M-Proteins in Serum um ≥ 90 % und M-Proteinspiegel im Urin < 100 mg/24 h • PR: <ul style="list-style-type: none"> - ≥ 50 % Reduktion des M-Protein im Serum und um ≥ 90 % oder < 200 mg im Urin/24 h - Bei nicht messbaren M-Proteinspiegeln in Serum und Urin: ≥ 50 % Reduktion der Differenz zwischen beteiligten und unbeteiligten freien Leichtketten-Werten; bei nicht messbarer Differenz zwischen beteiligten und unbeteiligten freien Leichtketten: ≥ 50 % Reduktion der Plasmazellen (vorausgesetzt Plasmazellen zur Baseline ≥ 30 %); zusätzlich ≥ 50 % Reduktion der Größe von Weichteilplasmozytomen <p><u>Ansprechrate gemäß EBMT:</u> basierend auf den Kriterien der EBMT (Bladé et al., 1998) (festgelegt durch das IRAC)</p> <ul style="list-style-type: none"> • CR: <ul style="list-style-type: none"> - Mind. 6 Wo kein Nachweis des M-Proteins im Serum und/oder Urin - < 5 % Plasmazellen im Knochenmarksaspirat und Stanzbiopsie (falls Stanzbiopsie durchgeführt) - Kein Anstieg lytischer Knochenläsionen in Größe und Anzahl - Verschwinden von Weichteilplasmozytomen • PR: <ul style="list-style-type: none"> - Mind. 6 Wo ≥ 50 % Reduktion des M-Proteinspiegels im Serum - Mind. 6 Wo Reduktion der Ausscheidung von Leichtketten im Urin über 24 h um ≥ 90 % oder auf < 200 mg - Bei Patienten mit nicht-sekretorischem Myelom: Mind. 6 Wo ≥ 50 % Reduktion der Plasmazellen in Knochenmarksaspirat und Stanzbiopsie (falls Stanzbiopsie durchgeführt) - ≥ 50 % Reduktion der Größe von Weichteilplasmozytomen - Kein Anstieg lytischer Knochenläsionen in Größe und Anzahl <p><u>Klinisches Ansprechen:</u></p> <ul style="list-style-type: none"> • Anzahl Patienten mit einer Verbesserung von Knochenschmerzen/ Zeit bis zur Verbesserung von Knochenschmerzen Eine Verbesserung ist definiert als eine Verbesserung um mindestens eine Kategorie im Vergleich zur Baseline bei der Frage 1 (Knochenschmerzen) des QLQ-MY20 Fragebogens. • Anzahl Patienten mit einer Verbesserung des ECOG PS/ Zeit bis zur Verbesserung des ECOG PS Eine Verbesserung ist definiert als eine Verbesserung um mindestens eine Kategorie im Vergleich zur Baseline im ECOG <i>Performance Status</i>.

Progressionsfreies Überleben (PFS) – RCT

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

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

Tabelle 4-19: Bewertung des Verzerrungspotenzials für PFS in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Eine Progression wurde durch ein verblindetes Komitee (IRAC) auf Basis der IMWG-Kriterien identifiziert, so dass eine Verzerrung der Befunderhebung aufgrund Kenntnis der Gruppenzugehörigkeit ausgeschlossen werden kann. Die Auswertung der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben zum Endpunkt *PFS* in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Das Verzerrungspotenzial ist aus diesen Gründen für *PFS* als niedrig eingestuft und mit einer hohen Ergebnissicherheit assoziiert. Das *cross-over* führt für diesen Endpunkt für den primär dargestellten Datenstand vom 7. September 2012 nicht zu einer Verzerrung, da die HD-Dex Patienten erst nach einer Progression in den Pom+LD-Dex-Arm wechseln konnten.

Die Daten zum Auswertungszeitpunkt 1. März 2013 werden lediglich supportiv dargestellt. Für diese liegt lediglich die Bewertung der Prüfarzte für die Progressionen vor.

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

Tabelle 4-20: Ergebnisse für PFS aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	PFS Zensierungsregeln EMA [FDA] (Wochen, median)	HR Zensierungsregeln EMA [FDA] [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	302	15,7 [15,6]	0,45 [0,35-0,59]; <0,001 [0,45 [0,34-0,58];<0,001]
	HD-Dex	153	8,0 [7,7]	

*Studienname fett; PFS = primärer Endpunkt

**p-Wert basierend auf stratifiziertem Log-Rank-Test

EMA: European Medicines Agency; FDA: Federal Drug Agency; HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall;

PFS: Progressionsfreies Überleben

Bezogen auf den primären Endpunkt *Progressionsfreies Überleben* zeigte sich eine deutliche Überlegenheit der Interventionstherapie. Das Progressionsfreie Überleben der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen der Kontrollgruppe um 7,7 Wochen (15,7 vs. 8,0 Wochen) verlängert. Die hiermit assoziierte relative Risikoreduktion für das Eintreten von Krankheitsprogression oder Tod betrug 55 % (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$) und war statistisch signifikant.

Eine derartige Verlängerung des Progressionsfreien Überlebens ist vor dem Hintergrund der besonderen Therapiesituation (es handelt sich um vielfach vortherafierte Patienten ohne weitere effektive Therapieoption) besonders bemerkenswert.

Die Primäranalyse des Endpunktes PFS beruhte auf Zensierungsregeln der EMA. Eine durchgeführte Sensitivitätsanalyse auf Basis von Zensierungsregeln der FDA (HR [95 %-KI]: 0,45 [0,34-0,58]; $p < 0,001$) bestätigte die Ergebnisse der Primäranalyse.

Eine weitere Sensitivitätsanalyse für den primären Endpunkt PFS analysierte die Zeit bis zum Therapieversagen (*Time to Treatment Failure*, TTF). Hier wurde die Zeitspanne bis zum ersten Eintreten eines der folgenden Ereignisse analysiert:

- Tod
- Progression gemäß den IMWG- oder EBMT-Kriterien, bestimmt durch das IRAC oder den behandelnden Arzt
- Studienabbruch
- Beginn einer neuen Myelomtherapie

Auch diese Sensitivitätsanalyse unterstützt den Behandlungsvorteil (HR [95 %-KI]: 0,44 [0,35-0,56]; $p < 0,001$).

Insgesamt bestätigen also die Sensitivitätsanalysen die Robustheit des beobachteten Nutzens von Pom+LD-Dex bzgl. des Progressionsfreien Überlebens.

Auch die Auswertung nach längerer Beobachtungszeit zum 1. März 2013 bestätigte den Vorteil unter Pom+LD-Dex für PFS nach Prüfarztbewertung (17,2 vs. 8,2 Wochen; HR [95 %-KI]: 0,48 [0,39-0,60]; $p < 0,001$). Zum Vergleich ergab das PFS nach Prüfarztbewertung zum Zeitpunkt 7. September 2012 eine relative Risikoreduktion von 59 % (HR [95 %-KI]: 0,41 [0,32-0,53]; $p < 0,001$).

Einfluss einer Progression auf Lebensqualität und ausgewählte UE

Um den Einfluss einer Progression auf die Lebensqualität zu analysieren, wurden folgende Analysen durchgeführt. Für alle Skalen/Items wurde ein multivariates *Mixed Model* verwendet. Für dieses wurden nur Werte bis einschließlich Zyklus 5 eingeschlossen, da nur bis Zyklus 5 mindestens 20 Patienten in jedem Arm noch in der Studie waren. Der Faktor *Patient* ging als zufälliger Effekt in das Modell ein und der Zyklus (zeitlicher Verlauf) und die Einflussfaktoren Behandlungsgruppe, Alter, Geschlecht und der jeweilige Skalenwert zur Baseline als fester Effekt. Zusätzlich war die Interaktion Behandlungsgruppe*Zyklus im Modell enthalten. Das Auftreten einer Progression wurde als zeitabhängige Variable in dem Modell untersucht, ebenso wie ein mindestens partielles Ansprechen.

In der folgenden Tabelle werden die Skalen/Items dargestellt, für die sich ein signifikanter Einfluss der Progression gezeigt hat.

Tabelle 4-21: Progression als signifikante Determinante reduzierter Lebensqualität

Einfluss der Progression auf	Koeffizient	p-Wert#
EQ-5D↑	-0,076	<0,001
QLQ-C30		
Allgemeiner Gesundheitszustand / Lebensqualität ↑	-3,815	0,001
Soziale Funktion↑	-3,815	0,048
Physische Funktion↑	-4,130	0,001
Dyspnoe↓	+6,538	0,001
Schmerz↓	+3,811	0,036

↓Ein höherer Score indiziert eine schlechtere Lebensqualität (höheres Ausmaß an Symptomen)

↑Ein höherer Score indiziert eine höhere Funktionalität/bessere Lebensqualität

#p-Wert aus dem *Mixed Model* für den Einfluss der Progression auf die Skala/Symptomatik

Quelle: (Celgene Europe Limited, 2013k)

Eine Progression der Erkrankung ist also ein prognostischer Indikator für eine Verschlechterung der Lebensqualität bzgl. des EQ-5D, des allgemeinen Gesundheitszustands/Lebensqualität, der Sozialen Funktion und der Physischen Funktion sowie den Symptomen Schmerz und Dyspnoe.

Der Einfluss einer Progression wurde ebenfalls für ausgewählte relevante UE untersucht. Hierbei wurden alle UE mit CTC Grad 3/4 ausgewählt, die laut Studienbericht der MM-003 „merklich häufiger“ (*notably more frequent*) in einem der beiden Behandlungsarme auftraten (Asthenie, Knochenschmerzen, febrile Neutropenie, Hyperglykämie, Infektionen, Myopathie¹⁶, und Neutropenie). Zusätzlich wurden Fatigue und Nierenversagen analysiert, da diese in der Indikation relevant sind. In der folgenden Tabelle wird angegeben, ob das Auftreten einer Progression die Wahrscheinlichkeit für das Auftreten des jeweiligen UE (z. B. Nierenversagen) beeinflusst, d. h. die *Hazard Ratio* für Progression vs. keine Progression. Hierbei wird für den Behandlungseffekt adjustiert, d. h. die Behandlungsgruppe ist im Cox Modell zusätzlich zur zeitabhängigen Progression enthalten. Ein *Hazard Ratio* das hierbei signifikant größer als 1 ist, ist in dem Sinne zu interpretieren, dass eine Progression die Wahrscheinlichkeit für das Auftreten des UE signifikant erhöht.

Tabelle 4-22: Progression als signifikante Determinante für unerwünschte Ereignisse

Einfluss der Progression	HR [95 %-KI] p-Wert*
Nierenversagen	3,37 [1,262-8,987] 0,015

*p-Wert basierend auf nicht stratifiziertem *Log-Rank*-Test;
 Quelle: (Celgene Europe Limited, 2013j)

Diese Analyse zeigt, dass das Auftreten einer Progression die Wahrscheinlichkeit eines Nierenversagens für den Patienten um mehr als das Dreifache erhöht. Für die weiteren geprüften UE konnte kein signifikanter Einfluss der Progression entdeckt werden.

Diese Daten unterstreichen zusammenfassend die Bedeutung der Verlängerung des Progressionsfreien Überlebens bei Patienten mit rrMM. Die durch Pom+LD-Dex erzielte Verlängerung des progressionsfreien Intervalls bedeutet somit für die Patienten einen Erhalt der Lebensqualität sowie eine Senkung der Symptomlast.

¹⁶ Myopathie konnte nicht analysiert werden, da Grad 3/4 Ereignisse nur im HD-Dex Arm auftraten und das Model daher nicht berechenbar war.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

Zeit bis zur Progression (TTP)– RCT

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

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

Tabelle 4-23: Bewertung des Verzerrungspotenzials für TTP in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Vorliegen einer Progression wurde durch ein verblindetes Komitee (IRAC) auf Basis der IMWG-Kriterien identifiziert, so dass eine Verzerrung der Befunderhebung aufgrund Kenntnis der Gruppenzugehörigkeit hierdurch ausgeschlossen werden kann. Die Auswertung der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben zum Endpunkt TTP in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Das Verzerrungspotenzial ist aus diesen Gründen für TTP als niedrig eingestuft und mit einer hohen Ergebnissicherheit assoziiert. Das *cross-over* führt für diesen Endpunkt für den primär dargestellten Datenstand vom 7. September 2012 nicht zu einer Verzerrung, da die HD-Dex Patienten erst nach einer Progression in den Pom+LD-Dex-Arm wechseln konnten.

Die Daten zum Auswertungszeitpunkt 1. März 2013 werden lediglich supportiv dargestellt. Für diese liegt lediglich die Bewertung der Prüfarzte für die Progressionen vor.

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

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

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	TTP (Wochen, median)	HR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	302	20,1	0,42 [0,31-0,56]; <0,001
	HD-Dex	153	8,3	

*p-Wert basierend auf stratifiziertem Log-Rank-Test

HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; TTP: Time to Progression

Die Zeit bis zur Progression der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen des Kontrollarms um 11,8 Wochen verlängert. Der Unterschied (20,1 vs. 8,3 Wochen) war statistisch signifikant und war mit einer relativen Risikoreduktion für das Auftreten einer Progression um 58 % verbunden (HR [95 %-KI]: 0,42 [0,31-0,56]; p<0,001). Zum Auswertungszeitpunkt 1. März 2013 unterschied sich die Zeit bis zur Progression nach Prüfarztbewertung weiterhin signifikant mit 20,2 Wochen unter Pom+LD-Dex und 9,0 Wochen unter HD-Dex.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

Ansprechrate gemäß den IMWG-Kriterien – RCT

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

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

Tabelle 4-25: Bewertung des Verzerrungspotenzials für die Ansprechrate gemäß den IMWG-Kriterien in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Ansprechrate gemäß den IMWG-Kriterien wurde durch ein verblindetes Komitee (IRAC) auf Basis der IMWG-Kriterien festgelegt, so dass eine Verzerrung der Befunderhebung aufgrund Kenntnis der Gruppenzugehörigkeit minimiert wurde. Die Auswertung der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben zum Endpunkt *Ansprechrate gemäß den IMWG-Kriterien* in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Das Verzerrungspotenzial ist aus diesen Gründen für die *Ansprechrate gemäß den IMWG-Kriterien* als niedrig eingestuft und ist mit einer hohen Ergebnissicherheit assoziiert.

Die Daten zum Auswertungszeitpunkt 1. März 2013 werden lediglich supportiv dargestellt. Für diese liegt lediglich die Bewertung der Prüfarzte für das Ansprechen vor.

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

Tabelle 4-26: Ergebnisse für die Ansprechraten gemäß den IMWG-Kriterien aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	Ansprechen gemäß den IMWG-Kriterien (%)				Dichotomisiertes Ansprechen gemäß den IMWG-Kriterien (%)		
			sCR	CR	VGPR	PR	sCR, CR, VGPR oder PR	SD, PD oder nicht evaluierbar	OR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	302	0,0	0,3	1,3	14,9	16,6	83,4	4,86 [2,03-11,61]; <0,001
	HD-Dex	153	0,0	0,0	0,7	3,3	3,9	96,1	

*p-Wert basierend auf stratifiziertem *Cochran-Mantel-Haenszel-Test*

IMWG: *International Myeloma Working Group*; ITT: *Intention-to-treat*; KI: *Konfidenzintervall*; OR: *Odds Ratio*;

PD: *progressive disease*; (s)CR: (*stringent*) *complete response*; SD: *stable disease*; (VG)PR: (*very good*) *partial response*

Bezogen auf den Endpunkt *Ansprechraten gemäß den IMWG-Kriterien* zeigte sich eine klinische Überlegenheit von Pom+LD-Dex gegenüber einer Behandlung mit HD-Dex.

Für keinen Studienteilnehmer wurde ein stringentes komplettes Ansprechen beobachtet. Studienteilnehmer der Interventionsgruppe erreichten häufiger ein komplettes und ein sehr gutes partielles Ansprechen als jene des Kontrollarms (0,3 vs. 0,0 % und 1,3 vs. 0,7 %). Ein partielles Ansprechen wurde im Pom+LD-Dex-Arm mehr als viermal so häufig erreicht als in der Vergleichsgruppe (14,9 vs. 3,3 %).

Die Ansprechraten gemäß den IMWG-Kriterien war in der Interventionsgruppe mehr als viermal höher als im HD-Dex-Arm (16,6 vs. 3,9 %; OR [95 %-KI]: 4,86 [2,03-11,61]; p<0,001).

Für den Auswertungszeitpunkt 1. März 2013 betrug das Gesamtansprechen nach Prüfarztbewertung gemäß den IMWG-Kriterien 31,0 % im Pom+LD-Dex-Arm und 10,0 % im HD-Dex-Arm (OR [95 %-KI]: 4,22 [2,35-7,58]; p<0,001). Patienten im Interventionsarm erreichten häufiger ein komplettes oder sehr gutes partielles Ansprechen als jene des Kontrollarms (1,0 vs. 0,0 % und 5,0 vs. 0,7 %). Ein Patient im Interventionsarm (0,3 %) erreichte eine stringente komplette Remission. Zum Vergleich betrug das Gesamtansprechen nach Prüfarztbewertung zum 7. September 2012 gemäß den IMWG-Kriterien 26,2 % im Pom+LD-Dex-Arm und 5,9 % im HD-Dex-Arm (OR [95 %-KI]: 5,67 [2,76-11,65]; p<0,001).

Einfluss des Ansprechens auf Lebensqualität

Um den Einfluss eines mindestens partiellen Ansprechens auf die Lebensqualität zu analysieren, wurde innerhalb desselben Mixed Models pro Skala ein analoges Vorgehen gewählt wie für den Nachweis des Einflusses der Progression (siehe Seite 63).

Tabelle 4-27: Mindestens partielles Ansprechen gemäß den IMWG-Kriterien als Determinante verbesserter Lebensqualität

Ansprechen \geq PR	Koeffizient	p-Wert#
EQ-5D \uparrow	+0,134	<0,001
QLQ-C30		
Allgemeiner Gesundheitszustand / Lebensqualität \uparrow	+5,391	0,008
Rollenfunktion \uparrow	+9,304	<0,001
Physische Funktion \uparrow	+7,074	<0,001
Soziale Funktion \uparrow	+7,086	0,008
Schmerz \downarrow	-7,444	0,003
Dyspnoe \downarrow	-6,250	0,022
Fatigue \downarrow	-7,265	0,002
Insomnie \downarrow	-10,89	< 0,001
QLQ-MY20		
Krankheitssymptome \downarrow	-6,470	<0,001

\uparrow Ein höherer Score indiziert eine höhere Funktionalität/bessere Lebensqualität

\downarrow Ein höherer Score indiziert eine schlechtere Lebensqualität (höheres Ausmaß an Symptomen)

p-Wert aus dem *Mixed Model* für den Einfluss der Progression auf die Skala/Symptomatik

Quelle: (Celgene Europe Limited, 2013k)

Diese Daten unterstreichen die Bedeutung eines Ansprechens für Patienten mit rrMM. Spricht die Erkrankung auf die Therapie an, so können die Patienten von einem Erhalt der Lebensqualität bzgl. des EQ-5D, des allgemeinen Gesundheitszustands/Lebensqualität, der Rollenfunktion, der Physischen und der Sozialen Funktion profitieren. Weiterhin reduziert das Ansprechen die Symptomlast für die Patienten bzgl. der Skalen/Items Schmerz, Dyspnoe, Fatigue, Insomnie gemäß QLQ-C30 und der Krankheitssymptome gemäß QLQ-MY20.

Da die Ansprechraten unter Pom+LD-Dex signifikant höher sind als unter HD-Dex und auch tiefere Remissionen bis hin zu Komplettremissionen erzielt werden können, zeigt sich hier ein zusätzlicher Nutzenaspekt für Pom+LD-Dex in Bezug auf die Nutzendimension *Morbidität*.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

Ansprechrate gemäß den EBMT-Kriterien – RCT

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

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

Tabelle 4-28: Bewertung des Verzerrungspotenzials für die Ansprechrate gemäß den EBMT-Kriterien in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Ansprechrate gemäß EBMT wurde durch ein verblindetes Komitee (IRAC) auf Basis der EBMT-Kriterien festgelegt, so dass eine Verzerrung der Befunderhebung aufgrund Kenntnis der Gruppenzugehörigkeit hierdurch minimiert wurde. Die Auswertung der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben zum Endpunkt *Ansprechrate gemäß den EBMT-Kriterien* in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Das Verzerrungspotenzial ist aus diesen Gründen für die *Ansprechrate gemäß den EBMT-Kriterien* als niedrig eingestuft und ist mit einer hohen Ergebnissicherheit assoziiert.

Die Daten zum Auswertungszeitpunkt 1. März 2013 werden lediglich supportiv dargestellt. Für diese liegt lediglich die Bewertung der Prüfärzte für das Ansprechen vor.

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

Tabelle 4-29: Ergebnisse für Ansprechraten gemäß der EBMT-Kriterien aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	Ansprechen gemäß den EBMT-Kriterien (%)		Dichotomisiertes Ansprechen gemäß den EBMT-Kriterien (%)		
			CR	PR	CR oder PR	MR, SD, PD oder nicht evaluierbar	OR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	302	0,3	14,6	14,9	85,1	5,18 [2,01-13,35]; <0,001
	HD-Dex	153	0,0	3,3	3,3	96,7	

*p-Wert basierend auf stratifiziertem Cochran-Mantel-Haenszel-Test

Die Ansprechraten gemäß den EBMT-Kriterien bestätigen die Ergebnisse aus der Analyse des Ansprechens gemäß den IMWG-Kriterien.

Ein Teilnehmer des Pom+LD-Dex-Arms erreichte ein komplettes Ansprechen, jedoch kein Teilnehmer in der Kontrollgruppe (0,3 vs. 0,0 %). Studienteilnehmer des Interventionsarms erreichten häufiger ein partielles Ansprechen, als jene der Kontrollgruppe (14,6 vs. 3,3 %). Die Ansprechraten gemäß den EBMT-Kriterien waren mehr als viermal höher als im HD-Dex-Arm (14,9 vs. 3,3 %; OR [95 %-KI]: 5,18 [2,01-13,35]; $p < 0,001$).

Untersucht wurde ebenfalls das Objektive Ansprechen (*Objective Response*; CR+PR+MR). Hier wird auch das Minimale Ansprechen (*Minimal Response*; MR) mit einbezogen (Bladé et al., 1998). Dieses war im Interventionsarm signifikant höher als im Kontrollarm (26,8 vs. 6,5 %, OR [95 %-KI]: 5,24 [2,63-10,45]; $p < 0,001$). Zum Auswertungszeitpunkt 1. März 2013 wurde gemäß den EBMT-Kriterien nur die Kategorie minimales Ansprechen (MR) erhoben; dieses wurde von Patienten im Interventionsarm nicht signifikant häufiger erreicht (8,0 vs. 6,0 %; $p = 0,565$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

Klinisches Ansprechen – RCT

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

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

Tabelle 4-30: Bewertung des Verzerrungspotenzials für Klinisches Ansprechen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Zeit bis zur ersten Verbesserung von Knochenschmerzen/ Anzahl Patienten mit einer Verbesserung von Knochenschmerzen						
MM-003	niedrig	nein	ja	nein	nein	hoch
Zeit bis zur ersten Verbesserung des ECOG PS/ Anzahl Patienten mit einer Verbesserung des ECOG PS						
MM-003	niedrig	nein	ja	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Knochenschmerzen und der ECOG PS wurden unverblindet von den behandelnden Prüfärzten erhoben. Die Analyse der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben zu diesen Endpunkten in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Durch die unverblindete Erhebung kann ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine daraus resultierende Verzerrung nicht ausgeschlossen werden. Aus diesen Gründen ist das Verzerrungspotenzial für die Endpunkte *Verbesserung von Knochenschmerzen* und *Verbesserung des ECOG PS* als hoch eingestuft und mit einer mäßigen Ergebnissicherheit assoziiert.

Im Studienbericht wurden für die beiden Endpunkte die Zeit bis zur Verbesserung anhand eines Wilcoxon Rangsummentests verglichen. Diese Analyse kann zu Verzerrungen führen, wenn eine Behandlungsgruppe deutlich länger beobachtet wird. In dieser Gruppe können die Verbesserungen auch zu einem späteren Zeitpunkt auftreten, und somit die Zeit bis zur Verbesserung durch die Rangordnung unbegründet als schlechter (später) bewertet werden. Um

dieser Thematik gerecht zu werden, werden im Dossier primär die Anteile an Patienten mit Verbesserungen verglichen. Die Analysen aus dem Studienbericht werden ebenfalls der Vollständigkeit halber dargestellt.

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

Tabelle 4-31: Ergebnisse für Klinisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl Studienteilnehmer	Patienten mit einer Verbesserung		OR [95 %-KI] p-Wert*	Mediane Zeit bis zur Verbesserung (Wochen)	p-Wert**
			N	%			
Verbesserung von Knochenschmerzen							
MM-003	Pom+LD-Dex	199	76	38,2	1,28 [0,75-2,18] 0,422	4,9	0,007
	HD-Dex	86	28	32,6		4,1	
Verbesserung des ECOG PS							
MM-003	Pom+LD-Dex	246	51	20,7	2,09 [1,09-4,02] 0,027	5,4	0,272
	HD-Dex	117	13	11,1		4,1	

Patienten mit fehlenden Daten wurden aus der Analyse ausgeschlossen.

*Fishers-Exact-Test

**p-Wert basierend auf einem Wilcoxon-Rangsummen-Test

Quelle: (Celgene Europe Limited, 2013l)

Studienteilnehmer des Interventionsarms profitierten häufiger von einer Verbesserung von Knochenschmerzen als jene der Kontrollgruppe (38,2 vs. 32,6 %; OR [95 %-KI]: 1,28 [0,75-2,18]; p=0,422). Dieser Unterschied war statistisch nicht signifikant (Celgene Europe Limited, 2013l).

Gemessen am *ECOG Performance Status* profitierten signifikant mehr Studienteilnehmer des Interventionsarms von einer Verbesserung als jene der Kontrollgruppe (20,7 vs. 11,1 %; OR [95 %-KI]: 2,09 [1,09-4,02]; p=0,027) (Celgene Europe Limited, 2013l).

Die mediane Zeit bis zur Verbesserung wurde im Studienbericht analysiert und wird hier dargestellt. Aus methodischen Gründen (siehe oben Abschnitt Verzerrungspotential) ist diese Analyse wenig belastbar und wird daher hier nicht detailliert besprochen.

Die Lebensqualitätsanalysen der MM-003 zeigten einen weiteren morbiditätsrelevanten Effekt: Während unter der Therapie mit HD-Dex die mediane Zeit bis zur ersten Verschlechterung von Fatigue, einem relevanten Krankheitssymptom bei Patienten mit rrMM, 60 Tage betrug [95 %-KI]: [57-113], war diese Zeit unter Pom+LD-Dex mit 113 Tagen [95 %-KI]: [71-169] signifikant verlängert ($p=0,038$). Dies zeigt den patientenrelevanten Effekt von Pom+LD-Dex in Bezug auf die Vermeidung krankheitsbedingter Symptome.

In der Studie MM-002 wurde ebenfalls die Wirksamkeit von Pom+LD-Dex in Bezug auf die klinischen Nutzenparameter (*Clinical Benefit Parameters*) Thrombozytenzahl, Serum-Kalziumspiegel, Serum-Kreatininspiegel, Serum-Albuminspiegel, Hämoglobinspiegel, Immunglobulinspiegel sowie ECOG Performance Status untersucht (Lonial et al., 2012). Im Verlauf von acht Zyklen Pom+LD-Dex ($n=112$) verbesserten 62 % der Patienten ihre Thrombozytenzahl, 93 % ihren Serum-Kalziumspiegel, 42 % ihren Serum-Kreatininspiegel, 15 % ihren Serum-Albuminspiegel und 50 % ihren Hämoglobinspiegel. Bei nur 5 % der Patienten wurde in dieser Zeitspanne eine Verschlechterung der Thrombozytenzahl bzw. bei 8 % eine Verschlechterung des Serum-Kreatininspiegels beobachtet. Der Immunglobulinspiegel verbesserte sich bei 15 % der Patienten, bei zwei von drei Patienten verschlechterte er sich jedoch, was den weit fortgeschrittenen Status ihrer Erkrankung widerspiegelt. Der ECOG Performance Status verbesserte sich bei 33 % der Patienten unter Pom+LD-Dex um mindestens einen Punkt.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

4.3.1.3.1.3 Lebensqualität– RCT

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

Tabelle 4-32: Operationalisierung von Lebensqualität

Studie	Operationalisierung
MM-003	<p>Erhebung am Tag 1 jedes Zyklus (inkl. Zyklus 1) und bei Behandlungsabbruch oder Progression durch:</p> <ul style="list-style-type: none"> • EORTC QLQ-C30: 30 Fragen aufgeteilt auf <ul style="list-style-type: none"> ○ 5 funktionale Skalen (Physische Funktion, Rollenfunktion, Emotionale Funktion, Soziale Funktion, Kognitive Funktion) ○ 3 Symptomskalen (Fatigue, Übelkeit und Erbrechen, Schmerz) ○ 1 globale Skala (Allgemeiner Gesundheitszustand/Lebensqualität) ○ 6 einzelne Items (Dyspnoe, Insomnie, Appetitlosigkeit, Obstipation, Diarrhoe, Finanzielle Schwierigkeiten) • EORTC QLQ-MY20: 20 Fragen aufgeteilt auf 4 Skalen <ul style="list-style-type: none"> ○ Krankheitssymptome ○ Nebenwirkungen der Behandlung ○ Zukunftsperspektiven ○ Körperbild • EQ-5D: Berechnung eines <i>Utility Index Scores</i> mit deutschen Populationsgewichten auf Basis folgender 5 Dimensionen <ul style="list-style-type: none"> ○ Mobilität ○ Selbstversorgung ○ Allgemeine Tätigkeiten ○ Schmerz ○ Angst

Anmerkung: Im Studienbericht der MM-003 wird in Abschnitt 11.6 auf einen separaten Bericht zur Lebensqualität verwiesen. Dieser Bericht liegt noch nicht vor, da die Analyse der gesamten Lebensqualitätsdaten der Studie MM-003 noch andauert. Aus diesem Grund werden einzelne Analysen individuell in Form statistischer Übersichtstabellen im Rahmen des eingereichten Nutzendossiers referenziert.

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

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

Tabelle 4-33: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MM-003	niedrig	nein	ja	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Daten zur Lebensqualität wurden unverblindet erhoben. Die Analyse der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ergebnisse zur *Lebensqualität* sind im Studienbericht nicht enthalten. Für dieses Dossier wurden die vorhandenen Daten für den Datenstand der finalen PFS Analyse 7. September 2012 analysiert.

Die Adhärenzrate ist definiert als die Anzahl von Patientenfragebögen mit mindestens 50 % beantworteter Items zu einem jeweiligen Erhebungszeitpunkt, geteilt durch die Anzahl der noch in der Studie befindlichen Studienteilnehmer („Rücklaufquote“). Die Adhärenz ist ein wichtiges Maß um die Aussagekraft und die Vergleichbarkeit der Ergebnisse in den Studienarmen zu bewerten. Die Adhärenz bzgl. der Lebensqualität in MM-003 ist in Tabelle 4-34 dargestellt.

Tabelle 4-34: Adhärenzraten bei der Lebensqualitätserhebung

Studie	Zeitpunkt	Noch in der Studie verbleibende Anzahl Patienten		Adhärenzrate (%)*					
				EORTC QLQ-C30		EORTC QLQ-MY20		EQ-5D	
		Pom+LD-Dex	HD-Dex	Pom+LD-Dex	HD-Dex	Pom+LD-Dex	HD-Dex	Pom+LD-Dex	HD-Dex
MM-003	Zyklus 1	300	148	81,0	81,1	84,0	83,8	79,7	77,0
	Zyklus 2	258	118	77,5	71,2	79,5	74,6	74,4	70,3
	Zyklus 3	190	57	75,3	78,9	76,8	78,9	73,2	78,9
	Zyklus 4	134	31	79,9	77,4	81,3	80,6	76,9	77,4
	Zyklus 5	104	25	83,7	84,0	83,7	84,0	83,7	80,0
	Zyklus 6	86	14	82,6	92,9	82,6	92,9	79,1	85,7
	Zyklus 7	66	12	78,8	91,7	80,3	91,7	75,8	91,7
	Zyklus 8	51	6	70,6	66,7	72,5	66,7	70,6	66,7
	Zyklus 9	36	5	72,2	80,0	75,0	80,0	69,4	80,0
	Zyklus 10	25	4	80,0	50,0	80,0	50,0	80,0	50,0
	Zyklus 11	18	2	77,8	50,0	77,8	50,0	77,8	0,0
	Zyklus 12	13	2	76,9	100,0	76,9	100,0	76,9	100,0
	Zyklus 13	8	0	75,0	-	75,0	-	75,0	-
	Zyklus 14	6	0	50,0	-	50,0	-	50,0	-
	Zyklus 15	3	0	33,3	-	33,3	-	33,3	-
	Zyklus 16	1	0	0,0	-	0,0	-	0,0	-

*Anteil erhaltener Fragebögen (mit einem Ausfüllungsgrad ≥ 50 %) an der Gesamtzahl zu den jeweiligen Zeitpunkten erwarteten Fragebögen
Quelle: (Celgene Europe Limited, 2013a)

Die Adhärenzrate war über alle Zyklen und für alle drei Erhebungsinstrumente hoch. Für Patienten im Pom+LD-Dex-Arm lag die Adhärenzrate bis einschließlich Zyklus 13 mit einer Ausnahme (EQ-5D Zyklus 9: 69,4 %) durchgehend über 70 %. Für Teilnehmer der Kontrollgruppe war dies bis einschließlich Zyklus 7 der Fall. Die Adhärenzraten unterschieden sich für keines der drei Erhebungsinstrumente zu keinem Erhebungszeitpunkt statistisch signifikant zwischen den beiden Behandlungsgruppen (Celgene Europe Limited, 2013a).

Die Ergebnisse zur Lebensqualität sind also aufgrund der hohen Adhärenz sowohl aussagekräftig als auch zwischen den Behandlungsarmen vergleichbar. Zu beachten ist hierbei, dass die Gesamtzahl der in der Studie verbleibenden Patienten im HD-Dex-Arm aufgrund der auftretenden Progressionen und dem damit verbundenen Studieneende schnell sinkt.

Durch die unverblindete Erhebung kann ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine daraus resultierende Verzerrung jedoch nicht ausgeschlossen wer-

den. Aus diesen Gründen ist das Verzerrungspotenzial für den Endpunkt *Lebensqualität* als hoch eingestuft und mit einer mäßigen Ergebnissicherheit assoziiert.

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

Die Lebensqualität wurde durch den EORTC QLQ-C30 Fragebogen, den krankheitsspezifischen EORTC QLQ-MY20-Fragebogen und den EuroQoL EQ-5D-Fragebogen jeweils an Tag 1 eines jeden Behandlungszyklus' à 28 Tage sowie zum Zeitpunkt des Behandlungsabbruchs bzw. Zeitpunkt der Progression erhoben. Ergänzend zur Erhebung bei Baseline ergaben sich insgesamt 240 Messpunkte (20 Skalen der einzelnen Fragebögen mit Erhebungen an 12 Zyklen für beide Arme) an denen der Behandlungseffekt in beiden Behandlungsarmen verglichen werden kann. Für die visiten-basierten Analysen (Tabelle 4-36) wurden alle Fragebögen, die beim Studienabbruch eines Patienten zwischen zwei geplanten Auswertungszeitpunkten ausgefüllt wurden, dem nächsten geplanten Auswertungszeitpunkt zugeordnet. Dies ermöglicht, dass auch diese Informationen in die Analyse eingehen.

Auch wenn *Minimum Clinically Important Differences* (MCID) im Rahmen anderer MM-Studien diskutiert wurden (Kvam et al., 2010), so weisen Daten aus der eigenen Studie immer einen höheren Präzisionsgrad auf als externe Quellen ((Sonneveld et al., 2013), Tabelle 5). Deshalb wurde die klinische Relevanz dieser Veränderungen für die einzelnen Skalen anhand des jeweiligen Standardfehlers des Mittelwerts (*Standard Error of Mean*, SEM) beurteilt (Wyrwich et al., 1999). Dieses Verfahren zur Bestimmung der klinischen Relevanz wurde für Patienten mit multiplem Myelom für die Fragebögen EORTC QLQ-C30 und EORTC QLQ-MY20 beispielsweise in Dimopoulos 2013 verwendet (Dimopoulos et al., 2013).

Für die Items *Körperbild*, *Dyspnoe*, *Insomnie*, *Obstipation*, und *Finanzielle Schwierigkeiten* des EORTC QLQ-C30 konnte kein SEM berechnet werden, da es sich hierbei um einzelne Items handelt, der SEM sich jedoch auf Basis der Reliabilität mehrerer Items innerhalb einer Skala errechnet. Für oben genannte Skalen wurde die klinische Relevanz daher anhand der halben Standardabweichung ($0,5 \cdot \text{STD}$) bewertet (Celgene Europe Limited, 2013h).

Um die Veränderung der Lebensqualität über die gesamte Studienzeit zu bewerten, wurde ein *Mixed Model (random intercept/slope model)* analysiert. Hierbei gingen der *Intercept* und die Zeit (Zyklus) als zufällige Variablen, die Behandlungsgruppe als fester (*fixed*) Effekt und die Interaktion Behandlungsgruppe*Zyklus in das Modell ein. In Tabelle 4-35 wird die Veränderung der Lebensqualität über die Zeit in beiden Armen verglichen. Ein positiver Wert bedeutet, dass sich der Lebensqualitätswert im Pom+LD-Dex-Arm schneller erhöht, bzw. weniger schnell verringert als im HD-Dex-Arm. Hierbei ist zu beachten, dass eine Erhöhung der Werte für manche Skalen eine Verbesserung (\uparrow) und für andere Skalen eine Verschlechterung (\downarrow) bedeutet. Dies ist jeweils durch Pfeilsymbole in der Tabelle markiert.

Tabelle 4-35: Mixed Model für die Veränderung der Lebensqualität

Studie	Skala	Veränderung unter Pom+LD-Dex im Vergleich zu HD-Dex pro Zyklus	p-Wert
EORTC QLQ-C30			
MM-003	Allgemeiner Gesundheitszustand / Lebensqualität ↑	0,267	0,701
	Physische Funktion↑	1,606	0,023
	Rollenfunktion↑	1,348	0,161
	Emotionale Funktion↑	0,926	0,186
	Soziale Funktion↑	-1,060	0,250
	Kognitive Funktion↑	0,979	0,152
	Fatigue↓	-1,394	0,089
	Übelkeit und Erbrechen↓	1,092	0,021
	Schmerz↓	-0,260	0,791
	Dyspnoe↓	0,902	0,324
	Insomnie↓	-0,945	0,284
	Appetitlosigkeit↓	-0,0374	0,962
	Obstipation↓	1,762	0,088
	Diarrhoe↓	-1,270	0,056
Finanzielle Schwierigkeiten↓	-0,0867	0,919	
EORTC QLQ-MY20			
MM-003	Krankheitssymptome↓	0,339	0,551
	Nebenwirkungen der Behandlung↓	-0,435	0,307
	Zukunftsperspektiven↑	0,550	0,524
	Körperbild↑	0,147	0,858
EQ-5D			
MM-003	Utility Index Score↑	0,003	0,665

° Die Differenz der Veränderungen unter Pom+LD-Dex vs HD-Dex wird dargestellt. Ein positiver Wert bei Parametern mit ↑ bedeutet, dass sich die Lebensqualität unter Pom+LD-Dex mehr verbessert bzw. weniger verschlechtert als unter HD-Dex. Für Parameter mit ↓ ist die Interpretation umgekehrt, da ein negativer Wert eine Verringerung der Symptomlast (Verbesserung der Lebensqualität) bedeutet.

p-Wert aus dem *Mixed Model* für die Interaktion Behandlungseffekt*Zyklus

Ergebnis fett gedruckt = statistisch signifikant

↑ Für Zukunftsperspektiven, Körperbild, Allgemeiner Gesundheitszustand/Lebensqualität, Physische Funktion, Rollenfunktion, Emotionale Funktion, Kognitive Funktion und Soziale Funktion indiziert ein höherer Score eine höhere Funktionalität/bessere Lebensqualität (Score zwischen 0 und 100; bzw. für den EQ-5D zwischen -0,2 und 1)

↓ Für Nebenwirkungen der Behandlung, Übelkeit und Erbrechen, Insomnie, Obstipation, Fatigue, Schmerz, Finanzielle Schwierigkeiten und Dyspnoe indiziert ein höherer Score eine schlechtere Lebensqualität (höheres Ausmaß an Symptomen; Score zwischen 0 und 100)

Quelle: (Celgene Europe Limited, 2013i)

Es zeigt sich für die Physische Funktion ein signifikanter Vorteil von Pom+LD-Dex gegenüber HD-Dex über den ganzen Verlauf der Studie. Für die Skala *Übelkeit und Erbrechen* wird ein signifikanter Vorteil für den HD-Dex-Arm beobachtet. Für alle anderen Skalen wurde bei dieser Analyse kein signifikanter Unterschied gefunden.

Um die Veränderungen noch detaillierter zu jedem Zyklus einzeln zu bewerten, wurde zusätzlich an allen 240 möglichen Messzeitpunkten (20 Skalen zu jeweils 12 Zyklen) der Behandlungseffekt der beiden Arme verglichen. Für 25 Messzeitpunkte wurden signifikante Unterschiede zwischen den beiden Armen bzgl. der Lebensqualitätsveränderungen im Vergleich zur Baseline beobachtet (Tabelle 4-36). Ein positiver Wert bedeutet, dass sich der Lebensqualitätswert im Pom+LD-Dex-Arm im Vergleich zur Baseline schneller erhöht, bzw. weniger schnell verringert als im HD-Dex-Arm zum jeweiligen analysierten Zyklus. Hierbei ist zu beachten, dass eine Erhöhung der Werte für manche Skalen eine Verbesserung (↑) und für andere Skalen eine Verschlechterung (↓) bedeutet. Dies ist jeweils durch Pfeilsymbole in der Tabelle markiert.

An 16 Messzeitpunkten war Pom+LD-Dex überlegen, davon waren neun Unterschiede klinisch relevant. An neun Messzeitpunkten war HD-Dex überlegen und die Unterschiede waren an sieben Messzeitpunkten klinisch relevant (Tabelle 4-36). Die vollständigen Ergebnisse der Lebensqualitätserhebungen sind in Modul 5 beigelegt (Celgene Europe Limited, 2013g).

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

Tabelle 4-36: Signifikante Ergebnisse für Lebensqualität aus RCT mit dem zu bewertenden Arzneimittel pro Zyklus

Studie	Skala (MCID)	Pom+LD-Dex				HD-Dex				Mittlere Differenz ^o der Veränderung zur Baseline; [95 %-KI]; p-Wert #
		Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	
EORTC QLQ-C30 Zyklus 2, Tag 1										
MM-003	Allgemeiner Gesundheitszustand / Lebensqualität (6,65) ↑	242 49,76± 23,13	201 51,74± 21,26	187	1,07± 22,82	119 48,95± 21,22	85 45,10± 23,55	78	-6,09± 25,11	7,16 [0,92–13,40]; 0,0247(*)
	Übelkeit und Erbrechen (9,96) ↓	243 8,02± 15,25	201 9,70± 18,74	188	1,86± 19,90	120 9,31± 16,28	85 4,71± 11,96	78	-3,85± 14,45	5,71 [1,40–10,02]; 0,0098
	Insomnie (15,14) ↓	243 29,90± 30,34	201 30,02± 29,63	188	-0,00± 33,60	120 25,56± 30,17	85 40,39± 36,77	78	13,25± 34,96	-13,25 [-22,26– -4,23]; 0,0041
	Obstipation (13,87) ↓	243 13,85± 25,26	201 24,38± 30,69	188	9,40± 31,41	120 21,94± 32,19	85 18,43± 26,98	78	-1,71± 25,71	11,11 [3,80–18,42]; 0,0031
EORTC QLQ-C30 Zyklus 3, Tag 1										
MM-003	Physische Funktion (9,38) ↑	243 63,76± 24,05	149 64,12± 23,55	137	-1,61± 20,56	120 60,07± 24,91	55 53,82± 27,93	50	-13,73± 23,19	12,13 [5,19–19,07]; 0,0007(*)
	Rollenfunktion (10,03) ↑	243 55,35± 33,45	149 54,36± 30,96	137	-2,31± 29,51	119 51,96± 34,93	55 45,45± 34,01	50	-15,33± 33,98	13,02 [3,00–23,05]; 0,0112(*)
	Emotionale Funktion (9,44) ↑	243 69,17± 23,99	149 71,87± 21,21	137	0,49± 20,48	120 67,71± 26,14	55 59,55± 27,05	50	-11,83± 27,62	12,32 [3,78–20,86]; 0,0053(*)

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

Studie	Skala (MCID)	Pom+LD-Dex				HD-Dex				Mittlere Differenz ^o der Veränderung zur Baseline; [95 %-KI]; p-Wert #
		Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	
MM-003	Kognitive Funktion (13,60) ↑	243 77,64± 24,21	149 75,17± 22,73	137	-6,08± 20,89	120 74,44± 26,55	55 66,36± 28,60	50	-15,00± 23,63	8,92 [1,86–15,98]; 0,0136
	Fatigue (9,24) ↓	243 46,91± 27,51	149 48,84± 25,73	137	3,65± 26,92	120 53,10± 29,23	55 58,38± 28,22	50	13,56± 27,79	-9,91 [-18,76– -1,06]; 0,0285(*)
	Schmerz (11,53) ↓	243 42,94± 33,43	149 40,49± 30,09	137	-1,09± 30,73	120 46,11± 34,55	55 49,70± 34,17	50	9,33± 29,96	-10,43 [-20,38– -0,48]; 0,0401
	Insomnie (15,14) ↓	243 29,90± 30,34	149 27,29± 29,52	137	-3,89± 33,84	120 25,56± 30,17	55 40,61± 36,11	50	17,33± 31,04	-21,23 [-32,02– -10,43]; 0,0001(*)
	Finanzielle Schwierigkeiten (14,82) ↓	243 20,58± 30,26	149 15,44± 26,99	137	-5,60± 26,37	120 17,22± 28,33	55 15,76± 28,58	50	3,33± 24,51	-8,93 [-17,37– -0,49]; 0,0382
EORTC QLQ-C30 Zyklus 5, Tag 1										
MM-003	Kognitive Funktion (13,60) ↑	243 77,64± 24,21	89 76,78± 21,85	83	-3,41± 22,51	120 74,44± 26,55	22 69,70± 25,01	18	-16,67± 24,92	13,25 [1,42–25,09]; 0,0285
	Insomnie (15,14) ↓	243 29,90± 30,34	89 26,97± 28,37	83	-6,43± 30,11	120 25,56± 30,17	22 39,39± 41,96	18	11,11± 41,22	-17,54 [-34,20– -0,88]; 0,0393(*)

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

Studie	Skala (MCID)	Pom+LD-Dex				HD-Dex				Mittlere Differenz ^o der Veränderung zur Baseline; [95 %-KI]; p-Wert #
		Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	
EORTC QLQ-C30 Zyklus 6, Tag 1										
MM-003	Soziale Funktion (12,55) ↑	243 65,23± 31,26	74 65,09± 30,74	68	-5,88± 30,31	120 60,83± 32,10	18 66,67± 22,14	15	7,78± 18,76	-13,66 [-26,04– -1,28]; 0,0316(*)
	Übelkeit und Erbrechen (9,96) ↓	243 8,02± 15,25	75 11,78± 22,39	69	3,62± 22,66	120 9,31± 16,28	18 2,78± 6,39	15	-8,89± 16,51	12,51 [0,19–24,83]; 0,0466(*)
	Dyspnoe (15,08) ↓	243 30,18± 29,13	75 33,78± 31,71	69	9,66± 28,64	120 35,28± 32,13	18 37,04± 27,75	15	-8,89± 29,46	18,55 [2,24–34,86]; 0,0263(*)
EORTC QLQ-C30 Zyklus 7, Tag 1										
MM-003	Dyspnoe (15,08) ↓	243 30,18± 29,13	52 41,03± 33,41	51	17,65± 25,26	120 35,28± 32,13	12 38,89± 34,33	10	-10,00± 27,44	27,65 [9,93–45,36]; 0,0028(*)
EORTC QLQ-C30 Zyklus 8, Tag 1										
MM-003	Soziale Funktion (12,55) ↑	243 65,23± 31,26	40 61,25± 25,43	39	-8,97± 25,32	120 60,83± 32,10	7 76,19± 18,90	5	20,00± 29,81	-28,97 [-53,69– -4,26]; 0,0227(*)
	Dyspnoe (15,08) ↓	243 30,18± 29,13	40 40,83± 34,99	39	17,09± 27,44	120 35,28± 32,13	7 38,10± 35,63	5	-20,00± 29,81	37,09 [10,57–63,62]; 0,0073 (*)

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

Studie	Skala (MCID)	Pom+LD-Dex				HD-Dex				Mittlere Differenz ^o der Veränderung zur Baseline; [95 %-KI]; p-Wert #
		Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	Wert zur Baseline (n, MW, ±STD)	Wert zum Zyklus (n, MW, ±STD)	Anzahl analysierter Studienteilnehmer	Veränderung (MW, ±STD)	
MM-003	Insomnie (15,14) ↓	243 29,90± 30,34	40 22,50± 23,13	39	-11,11± 34,48	120 25,56± 30,17	7 42,86± 46,00	5	26,67± 43,46	-37,78 [-71,75–-3,81]; 0,0301(*)
EORTC QLQ-C30 Zyklus 9, Tag 1										
MM-003	Dyspnoe (15,08) ↓	243 30,18± 29,13	26 41,03± 35,66	26	14,10± 23,42	120 35,28± 32,13	4 25,00± 31,91	3	-22,22± 19,25	36,32 [7,37–65,28]; 0,0158(*)
EORTC QLQ-MY20 Zyklus 3, Tag 1										
MM-003	Nebenwirkungen der Behandlung (7,80) ↓	252 21,80± 16,13	152 23,25± 14,22	143	3,40± 14,30	124 22,90± 15,58	55 29,34± 17,53	53	10,45± 15,99	-7,04 [-11,73–-2,36]; 0,0034
	Zukunftsperspektiven (10,75) ↑	251 51,35± 29,44	152 60,38± 26,26	143	7,93± 24,66	124 50,00± 28,53	55 46,46± 27,14	53	-3,56± 27,80	11,49 [3,39–19,59]; 0,0057(*)
	Körperbild (15,94) ↑	251 72,78± 31,52	152 75,44± 30,14	143	2,56± 29,89	123 72,90± 32,60	55 64,85± 37,08	53	-11,32± 37,52	13,88 [2,48–25,29]; 0,0177

^o Die Differenz der Veränderungen unter Pom+LD-Dex vs HD-Dex wird dargestellt. Ein positiver Wert bei Parametern mit ↑ bedeutet, dass sich die Lebensqualität unter Pom+LD-Dex mehr verbessert oder zumindest weniger verschlechtert als unter HD-Dex. Für Parameter mit ↓ ist es die Interpretation umgekehrt, da ein negativer Wert eine Verringerung der Symptomlast (Verbesserung der Lebensqualität) bedeutet.

p-Wert des t-Tests

(*) klinisch relevante Veränderung

Ergebnis fett gedruckt = Wert zugunsten von Pom+LD-Dex

↑Für Zukunftsperspektiven, Körperbild, Allgemeiner Gesundheitszustand / Lebensqualität, Physische Funktion, Rollenfunktion, Emotionale Funktion, Kognitive Funktion und Soziale Funktion indiziert ein höherer Score eine höhere Funktionalität/bessere Lebensqualität (Score zwischen 0 und 100, bzw. für den EQ-5D zwischen -0,2 und 1)

↓Für Nebenwirkungen der Behandlung, Übelkeit und Erbrechen, Insomnie, Obstipation, Fatigue, Schmerz, Finanzielle Schwierigkeiten und Dyspnoe indiziert ein höherer Score eine schlechtere Lebensqualität (höheres Ausmaß an Symptomen; Score zwischen 0 und 100)

Quelle: (Celgene Europe Limited, 2013g)

Auf Basis der in Tabelle 4-36 dargestellten Ergebnisse fällt auf, dass signifikante Unterschiede zugunsten von Pom+LD-Dex insbesondere in frühen Behandlungszyklen dokumentiert wurden. Bis einschließlich Zyklus 5 wurden insgesamt 15 der 16 signifikanten Unterschiede zugunsten von Pom+LD-Dex gemessen und acht der neun klinisch relevanten Unterschiede. Umgekehrt wurden ab Zyklus 6 sieben der insgesamt neun signifikanten Unterschiede zugunsten von HD-Dex dokumentiert und alle sieben waren klinisch relevant. Hierbei ist anzumerken, dass die Ergebnisse ab Zyklus 6 nur noch auf sehr wenigen Patienten im HD-Dex-Arm basieren (weniger als 15 Patienten ab Zyklus 6; weniger als sieben Patienten ab Zyklus 8) und daher nur von eingeschränkter Aussagekraft sind.

Angesichts der schwerwiegenden unheilbaren Erkrankung verschlechtern sich die Werte der Lebensqualität der Patienten zumeist langfristig gesehen für diese Patienten (Mols et al., 2012). Um zu analysieren welche Therapie die Verschlechterung der Lebensqualität am längsten verzögern kann, wurde die Zeit bis zur klinisch relevanten Verschlechterung analysiert. Pro Skala wurde eine klinisch relevante Verschlechterung als eine Überschreitung des skalen-spezifischen MCID definiert (Celgene Europe Limited, 2013h). Ein *Hazard Ratio* größer 1 bedeutet hier, dass das Risiko einer klinisch relevanten Verschlechterung unter HD-Dex erhöht ist im Vergleich zu Pom+LD-Dex.

Tabelle 4-37: Mediane Zeit bis zur klinisch relevanten Verschlechterung der Lebensqualität

Studie	Skala	Mediane Zeit bis zur Verschlechterung (Tage, median)		Hazard Ratio (Risiko unter HD-Dex / Risiko unter Pom+LD-Dex)		
		Pom+LD-Dex	HD-Dex	HR	[95 %-KI]	p-Wert*
EORTC QLQ-C30						
MM-003	Allgemeiner Gesundheitszustand/ Lebensqualität	114	85	1,36	[0,98-1,89]	0,058
	Physische Funktion	174	106	1,37	[0,95-1,97]	0,088
	Rollenfunktion	116	66	1,35	[0,96-1,90]	0,070
	Emotionale Funktion	190	124	1,53	[1,05-2,24]	0,023
	Soziale Funktion	113	87	1,13	[0,80-1,59]	0,484
	Kognitive Funktion	135	64	1,34	[0,95-1,87]	0,084
	Fatigue	113	60	1,40	[1,01-1,95]	0,038
	Übelkeit und Erbrechen	316	n. e.	0,56	[0,32-0,97]	0,034
	Schmerz	147	113	1,26	[0,87-1,83]	0,203
	Dyspnoe	114	169	0,84	[0,58-1,23]	0,354
	Insomnie	259	n. e.	1,58	[1,07-2,32]	0,017
	Appetitlosigkeit	233	n. e.	1,04	[0,66-1,64]	0,867
	Obstipation	192	n. e.	0,61	[0,38-0,97]	0,031
	Diarrhoe	295	n. e.	1,06	[0,64-1,76]	0,820
Finanzielle Schwierigkeiten	n. e.	n. e.	0,99	[0,56-1,75]	0,977	

Studie	Skala	Mediane Zeit bis zur Verschlechterung (Tage, median)		Hazard Ratio (Risiko unter HD-Dex / Risiko unter Pom+LD-Dex)		
		Pom+LD-Dex	HD-Dex	HR	[95 %-KI]	p-Wert*
EORTC QLQ-MY20						
MM-003	Krankheitssymptome	169	141	1,29	[0,88-1,90]	0,180
	Nebenwirkungen der Behandlung	142	113	1,49	[1,05-2,12]	0,023
	Zukunftsperspektiven	211	113	1,45	[0,99-2,13]	0,048
	Körperbild	225	n. e.	1,46	[0,97-2,20]	0,064
EQ-5D						
MM-003	Utility Index Score**	338	n.e.	1,19	[0,74-1,94]	0,467

n.e. = nicht erreicht

*p-Wert basierend auf nicht stratifiziertem *Log-Rank*-Test

**deutsche Populationsgewichtung

Ergebnis fett: = signifikanter Unterschied

Quelle: (Celgene Europe Limited, 2013m)

Ein signifikanter Vorteil für Pom+LD-Dex bzgl. der Verlängerung der Zeit bis zu einer Verschlechterung der Lebensqualität wurde für die fünf Skalen Emotionale Funktion, Fatigue, Insomnie, Nebenwirkungen der Behandlung und Zukunftsperspektiven identifiziert. Ein signifikanter Vorteil für HD-Dex zeigte sich bei den beiden Skalen „Übelkeit und Erbrechen“ und Obstipation. Für die anderen Skalen wurden keine signifikanten Unterschiede festgestellt. Auch hier zeigt sich ein heterogenes Bild, wobei die Behandlung mit Pom+LD-Dex die Verschlechterung der Lebensqualität bzgl. einer größeren Anzahl von Skalen aufhalten kann als die Behandlung mit HD-Dex.

Zusammenfassend lässt sich feststellen, dass die Kombination der zwei Substanzen Pomalidomid und niedrig-dosiertes Dexamethason im Vergleich zur Therapie mit hoch-dosiertem Dexamethason als Einzelsubstanz in der Summe die Lebensqualität der Patienten erhält. In einzelnen Aspekten war vor allem in frühen Behandlungszyklen sogar ein Hinweis auf einen positiven Effekt gegenüber dem Kontrollarm zu beobachten.

Bezieht man die deutliche Verlängerung der Zeit bis zur Progression sowie die signifikant erhöhte Ansprechrate unter Pom+LD-Dex und den Effekt dieser Einflussvariablen auf die Lebensqualität mit in die Gesamtschau ein, so ist von einem verlängerten Erhalt der Lebensqualität unter Pom+LD-Dex im Vergleich zu HD-Dex auszugehen (Abschnitt 4.3.1.3.1.2).

Insgesamt ist bemerkenswert, dass unter der Zweierkombination Pom+LD-Dex die Verlängerung des Gesamtüberlebens ohne Einschränkungen der Lebensqualität im Vergleich zur Monotherapie mit HD-Dex einhergeht. Die Kombinationstherapie Pom+LD-Dex im Vergleich zur Monotherapie mit HD-Dex verursacht also keinen erhöhten Schaden in Bezug auf die Lebensqualität der Patienten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

4.3.1.3.1.4 Sicherheit – RCT

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

Tabelle 4-38: Operationalisierung von Sicherheit

Studie	Operationalisierung
MM-003	<p><u>Unerwünschte Ereignisse:</u></p> <ul style="list-style-type: none"> • Studienteilnehmer mit: <ul style="list-style-type: none"> • Mind. 1 UE • Mind. 1 SUE • Mind. 1 zum Absetzen des Studien-AM führendem UE • Tod \leq 30 Tage nach letzter Dosis Studien-AM • Mind. 1 UE \geq CTC Grad 3/4 • Neutropenie • Anämie • Thrombopenie • Leukopenie • Febrile Neutropenie • Pneumonie • Allgemeine körperliche Gesundheitsverschlechterung • Asthenie • Hyperkalzämie • Hyperglykämie • Knochenschmerzen • Hämatopoetische Leukopenie • Infektionen • Hämatopoetische Thrombopenie • Akutes Nierenversagen • Hyperglykämie/neuer Diabetes Mellitus • Kardiale Arrhythmie • Hämorrhagie • Hämodynamisches Ödem, Erguss und Hypervolämie • Venöses thromboembolisches Ereignis • Depression und Suizid/Selbstverletzung • Muskelschwäche

Studie	Operationalisierung
	<ul style="list-style-type: none"> • Periphere Neuropathie • Arteriell-thromboembolisches Ereignis • Hämatopoetische Zytopenie bei >1 Zelltyp • Unspezifisches gefäßassoziiertes arterielles und venöses thromboembolisches Ereignis • Linsenstörung <p>Zusätzlich dargestellte unerwünschte Ereignisse: Studienteilnehmer mit:</p> <ul style="list-style-type: none"> • Sekundären Primärmalignomen <p>Die Schwere der unerwünschten Ereignisse wurde nach <i>National Cancer Institute Common Terminology Criteria</i> (NCI CTC) (Version 4.0) klassifiziert. Alle unerwünschten Ereignisse wurden gemäß des <i>Medical Dictionary for Regulatory Activities</i> (MedDRA) Version 14.0 kodiert.</p>

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

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

Tabelle 4-39: Bewertung des Verzerrungspotenzials für Sicherheit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Mind. 1 UE						
MM-003	niedrig	nein	ja	nein	ja	hoch
Mind. 1 SUE						
MM-003	niedrig	nein	ja	nein	ja	hoch
Mind. 1 zum Absetzen des Studien-AM führendes UE						
MM-003	niedrig	nein	ja	nein	ja	hoch
Tod ≤ 30 d nach letzter Dosis Studien-AM						
MM-003	niedrig	nein	ja	nein	ja	hoch
Mind. 1 UE ≥ CTC Grad 3/4						
MM-003	niedrig	nein	ja	nein	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die unerwünschten Ereignisse (UE) wurden unverblindet von den behandelnden Prüfarzten erhoben und gemäß des *Medical Dictionary for Regulatory Activities* (MedDRA) Version 14.0 kodiert. Die Analyse der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, ist transparent beschrieben. Ein Abgleich der Angaben in Studienprotokoll, Studienbericht und den Studienregisterdatenbanken ergab keinen Hinweis auf eine ergebnisgesteuerte Berichterstattung oder sonstige verzerrende Aspekte. Durch die unverblindete Erhebung kann ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine daraus resultierende Verzerrung nicht ausgeschlossen werden. Außerdem war die Beobachtungszeit im Pom+LD-Dex-Arm durch die längere Zeit bis zur Progression länger, was ebenfalls zu einer Verzerrung bzgl. der Häufigkeit des Auftretens von UE führen kann. Aus diesen Gründen ist das Verzerrungspotenzial für alle sicherheitsbezogenen Endpunkte als hoch eingestuft und ist mit einer mäßigen Ergebnissicherheit assoziiert.

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

Im Folgenden werden alle unerwünschten Ereignisse berichtet, die bezogen auf den Schweregrad 3/4 mit einer Häufigkeit von mindestens 5 % aufgetreten sind. Zudem werden die im Studienbericht als *Adverse Events of Special Interest* gekennzeichneten Ereignisse tabellarisch dargestellt. Zusätzlich wird die Häufigkeit sekundärer Primärmalignome aufgezeigt.

Tabelle 4-40: Übersicht über Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer (Safety Population)	Studienteilnehmer mit UE (%)									
			Mind. 1 UE	OR [95 %-KI]; p-Wert*	Mind. 1 SUE	OR [95 %-KI]; p-Wert*	Mind. 1 zum Absetzen des Studien-AM führendes UE	OR [95 %-KI]; p-Wert*	Tod ≤30 d nach letzter Dosis Studien-AM	OR [95 %-KI]; p-Wert*	Mind. 1 UE CTC Grad 3/4	OR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	300	97,0	2,326 [0,924-5,854]; 0,0814	51,0	1,027 [0,693-1,521]; 0,9204	9,7	1,886 [0,840-4,234]; 0,1454	12,7	0,884 [0,498-1,568]; 0,6595	78,0	1,130 [0,710-1,796]; 0,6332
	HD-Dex	149	93,3		50,3		5,4		14,1		75,8	

*p-Wert basierend auf einem Fishers-Exact-Test

Tabelle 4-41: Detailübersicht der Unerwünschten Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer Safety Population	Studienteilnehmer mit UE (%)							
			CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*
			Neutropenie		Anämie		Thrombopenie		Leukopenie	
MM-003	Pom+LD-Dex	300	41,7	4,123 [2,482-6,849]; <0,0001	27,0	0,912 [0,589-1,411]; 0,7368	20,7	0,818 [0,512-1,305]; 0,3985	8,7	2,732 [1,027-7,266]; 0,0465
	HD-Dex	149	14,8		28,9		24,2		3,4	
			Febrile Neutropenie		Pneumonie		Allgemeine körperliche Gesundheitsverschlechterung		Asthenie	
MM-003	Pom+LD-Dex	300	6,7	n. a.; 0,0004	9,0	1,241 [0,598-2,575]; 0,7190	5,3	0,783 [0,346-1,770]; 0,5293	3,3	0,536 [0,213-1,350]; 0,2141
	HD-Dex	149	0,0		7,4		6,7		6,0	
			Hyperkalzämie		Hyperglykämie		Knochenschmerzen		Hämatopoetische Leukopenie	
MM-003	Pom+LD-Dex	300	3,7	0,671 [0,264-1,705]; 0,4569	3,0	0,430 [0,171-1,082]; 0,0814	6,3	2,451 [0,819-7,338]; 0,1148	47,7	4,989 [3,030-8,216]; <0,0001
	HD-Dex	149	5,4		6,7		2,7		15,4	
			Infektionen		Hämatopoetische Thrombopenie		Akutes Nierenversagen		Hyperglykämie/neuer Diabetes Mellitus	
MM-003	Pom+LD-Dex	300	24,0	1,068 [0,671-1,701]; 0,8144	23,0	0,873 [0,553-1,377]; 0,5586	7,0	1,326 [0,573-3,070]; 0,6838	5,3	0,643 [0,296-1,397]; 0,3009
	HD-Dex	149	22,8		25,5		5,4		8,1	

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

Studie	Arm	Anzahl analysierter Studienteilnehmer Safety Population	Studienteilnehmer mit UE (%)							
			CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*	CTC Grad 3/4	OR [95 %-KI]; p-Wert*
			Kardiale Arrhythmie		Hämorrhagie		Hämodynamisches Ödem, Erguss und Hypervolämie		Venöses thromboembolisches Ereignis	
MM-003	Pom+LD-Dex	300	3,7	0,907 [0,329-2,502]; 0,7997	2,7	0,556 [0,198-1,563]; 0,2734	2,0	0,588 [0,176-1,958]; 0,5175	1,3	n.a.; 0,3068
	HD-Dex	149	4,0		4,7		3,4		0,0	
			Depression und Suizid/Selbstverletzung		Muskelschwäche		Periphere Neuropathie		Arteriell thromboembolisches Ereignis	
MM-003	Pom+LD-Dex	300	1,0	0,366 [0,081-1,657]; 0,2274	1,0	0,366 [0,081-1,657]; 0,2274	1,0	0,742 [0,123-4,491]; 0,6686	0,7	0,993 [0,089-11,04]; 1,0000
	HD-Dex	149	2,7		2,7		1,3		0,7	
			Hämatopoetische Zytopenie bei >1 Blutzelltyp		Unspezifisches gefäßassoziiertes arterielles und venöses thromboembolisches Ereignis		Linsenstörung			
MM-003	Pom+LD-Dex	300	0,7	0,493 [0,069-3,537]; 0,6027	0,3	0,495 [0,031-7,969]; 0,5541	0,3	n.a.; 1,0000		
	HD-Dex	149	1,3		0,7		0,0			

*p-Wert basierend auf einem Fishers-Exact-Test

Tabelle 4-42: Übersicht der SPM in RCT mit dem zu bewertenden Arzneimittel

Studie	Arm	Anzahl analysierter Studienteilnehmer ITT	Studienteilnehmer (%) mit			
			≥1 SPM	OR [95 %-KI]; p-Wert	Basalzellkarzinom	OR [95 %-KI]; p-Wert*
MM-003	Pom+LD-Dex	300	0,7	0,993 [0,089-11,04]; 1,0000	0,7	0,993 [0,089-11,04]; 1,0000
	HD-Dex	149	0,7		0,7	

*p-Wert basierend auf einem Fishers-Exact-Test

Bei nahezu allen Studienteilnehmern beider Arme wurde mindestens ein unerwünschtes Ereignis (Pom+LD-Dex 97,0 vs. HD-Dex 93,3 %) und bei etwa der Hälfte der Studienteilnehmer beider Gruppen mindestens ein schwerwiegendes unerwünschtes Ereignis (51,0 vs. 50,3 %) beobachtet. Die Unterschiede zwischen den Gruppen waren jeweils statistisch nicht signifikant. Die Gesamtrate an Studienteilnehmern mit mindestens einem unerwünschten Ereignis mit CTC Grad 3/4 war in beiden Gruppen vergleichbar (78,0 vs. 75,8 %). Der Anteil jener, die mindestens ein zum Absetzen des Studienarzneimittels führendes unerwünschtes Ereignis hatten, war in der Interventionsgruppe nicht signifikant erhöht (9,7 vs. 5,4 %). Der Anteil jener Teilnehmer, die innerhalb von 30 Tagen nach Gabe der letzten Dosis des Studienarzneimittels verstarben, war in der Interventionsgruppe niedriger als im Kontrollarm (12,7 vs. 14,1 %). Der Unterschied war statistisch nicht signifikant.

Bei Studienteilnehmern, die mit Pom+LD-Dex behandelt wurden, wurden signifikant häufiger Neutropenien (41,7 vs. 14,8 %), Leukopenien (8,7 vs. 3,4 %), febrile Neutropenien (6,7 vs. 0,0 %) und hämatopoetische Leukopenien (47,7 vs. 15,4 %) der Grade 3 und 4 beobachtet. Es ist anzumerken, dass eine Leukopenie in den meisten Fällen auf eine Verringerung der Zahl der neutrophilen Granulozyten, also eine Neutropenie, zurückzuführen ist.

Darüber hinaus gab es keine statistisch signifikanten Unterschiede bezogen auf CTC Grad 3/4 UE zwischen Interventions- und Kontrollgruppe. In beiden Armen entwickelten bis zum 7. September 2012 0,7 % der Studienteilnehmer mindestens ein sekundäres Primärmalignom. Die beobachteten sekundären Primärmalignome waren nicht-invasive Basalzellkarzinome der Haut; invasive solide oder hämatologische Neoplasien wurden nicht beobachtet. Zum 1. März 2013 hatten zwei Patienten im Pom+LD-Dex-Arm ein invasives, solides SPM entwickelt. Es gibt bislang - insbesondere vor dem Hintergrund der multiplen Vortherapien der Patienten - keine Hinweise auf einen Zusammenhang zwischen Pomalidomid und der Entstehung von SPM. Im HD-Dex-Arm wurden keine weiteren SPM beobachtet. Desweiteren bestätigte die Auswertung zum 1. März 2013 das beschriebene Sicherheitsprofil.

Um die unterschiedlich langen Beobachtungszeiten in die Bewertung mit einzubeziehen, wurden für einige relevante UE die Inzidenzraten berechnet. Hierbei wurden alle UE mit CTC

Grad 3/4 ausgewählt, die laut Studienbericht der MM-003 „merklich häufiger“ in einem der beiden Behandlungsarme auftraten. Zusätzlich wurden Fatigue und Nierenversagen analysiert, da diese in der Indikation relevant sind.

Die Inzidenzrate berechnet sich als Quotient aus der Anzahl der Patienten mit dem jeweiligen UE geteilt durch die gesamte Beobachtungsdauer im jeweiligen Behandlungsarm. Hierbei geht für Patienten mit dem jeweiligen UE nur die Zeit bis zum Auftreten des Ereignisses in die Beobachtungszeit ein.

Tabelle 4-43: Inzidenzraten für ausgewählte UE

Studie	UE	Pom+LD-Dex (N=300)			HD-Dex (N=149)		
		Patienten mit UE	Beobachtungszeit (Jahre)	Inzidenzrate [95 %-KI]	Patienten mit UE	Beobachtungszeit (Jahre)	Inzidenzrate [95 %-KI]
MM-003	Asthenie	10	102,7	9,74 [5,24-18,10]	9	33,4	26,96 [14,03-51,82]
	Fatigue	14	103,0	13,60 [8,05-22,96]	7	34,0	20,60 [9,82-43,21]
	Febrile Neutropenie	20	98,1	20,39 [13,16-31,61]	0	34,9	0 [n.b.-n.b.]
	Hyperglykämie	9	103,0	8,74 [4,55-16,79]	10	33,7	29,71 [15,98-55,21]
	Infektionen	72	87,9	81,95 [65,05-103,24]	34	30,8	110,34 [78,84-154,42]
	Knochen-schmerzen	19	99,8	19,03 [12,14-29,84]	4	34,2	11,69 [4,39-31,16]
	Myopathie	0	105,1	0 [n.b.-n.b.]	5	33,6	14,89 [6,20-35,77]
	Nieren-versagen	17	103,4	16,44 [10,22-26,44]	6	34,6	17,34 [7,79-38,59]
	Neutropenie	125	66,1	189,10 [158,69-225,33]	22	31,3	70,31 [46,30-106,79]

n.b. nicht berechenbar

Quelle: (Celgene Europe Limited, 2013f)

Die Adjustierung für unterschiedliche Beobachtungsdauern zeigt, dass Neutropenie und febrile Neutropenie die einzigen untersuchten UE sind, die unter Pom+LD-Dex signifikant häufiger auftreten als unter HD-Dex. Es wird ebenfalls bestätigt, dass trotz der höheren Neutropenie-Inzidenz das Auftreten von Infektionen unter Pom+LD-Dex nicht häufiger ist als unter HD-Dex. Die Analyse bestätigt zudem auch das bekannte Nebenwirkungsprofil von HD-Dex (erhöhtes Auftreten von Myopathie, Hyperglykämie).

Daten aus den Lebensqualitätsanalysen von MM-003 zeigten weitere relevante Effekte in Bezug auf die Nutzendimension *Sicherheit*. Die Zeit bis zur ersten Verschlechterung von Obstipation und „Übelkeit und Erbrechen“ war unter HD-Dex signifikant länger.

Für Insomnie ist zu beachten, dass der Pom+LD-Dex-Arm signifikant besser ist, obwohl der Median im HD-Dex-Arm noch nicht erreicht ist (Abbildung 4-3).

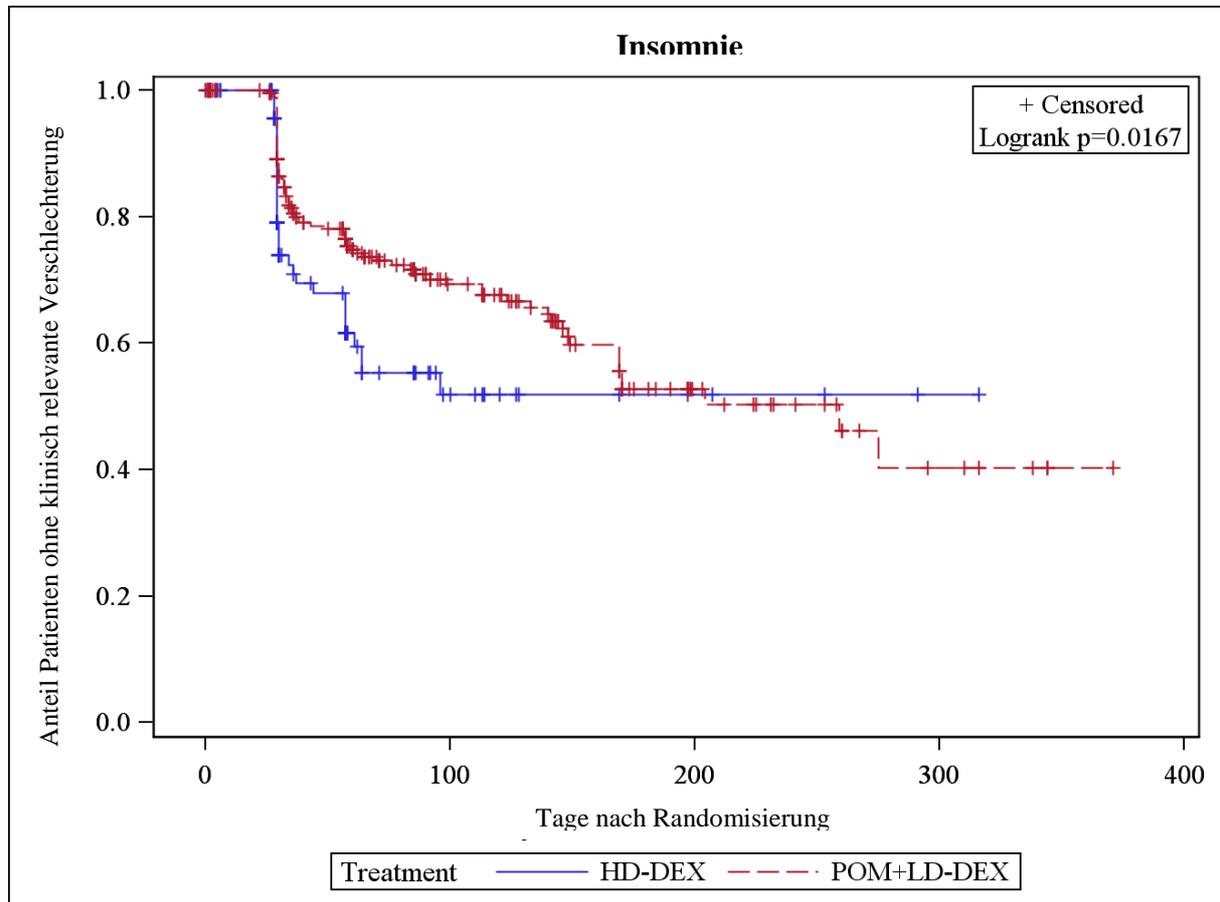


Abbildung 4-3: Kaplan-Meier Kurve für die Zeit bis zur Verschlechterung der QLQ-C30 Skala Insomnie

Bezogen auf die QLQ-MY20 Skala *Nebenwirkungen der Behandlung* zeigt sich hingegen, dass die mediane Zeit bis zur ersten Verschlechterung unter Pom+LD-Dex signifikant länger ist als unter HD-Dex (142 vs. 113 Tage, $p=0,023$). Dies ist ein Beleg, dass die deutlich wirksamere Kombinationstherapie Pom+LD-Dex in Bezug auf unerwünschte Nebenwirkungen keinen zusätzlichen Schaden im Vergleich zu HD-Dex induziert.

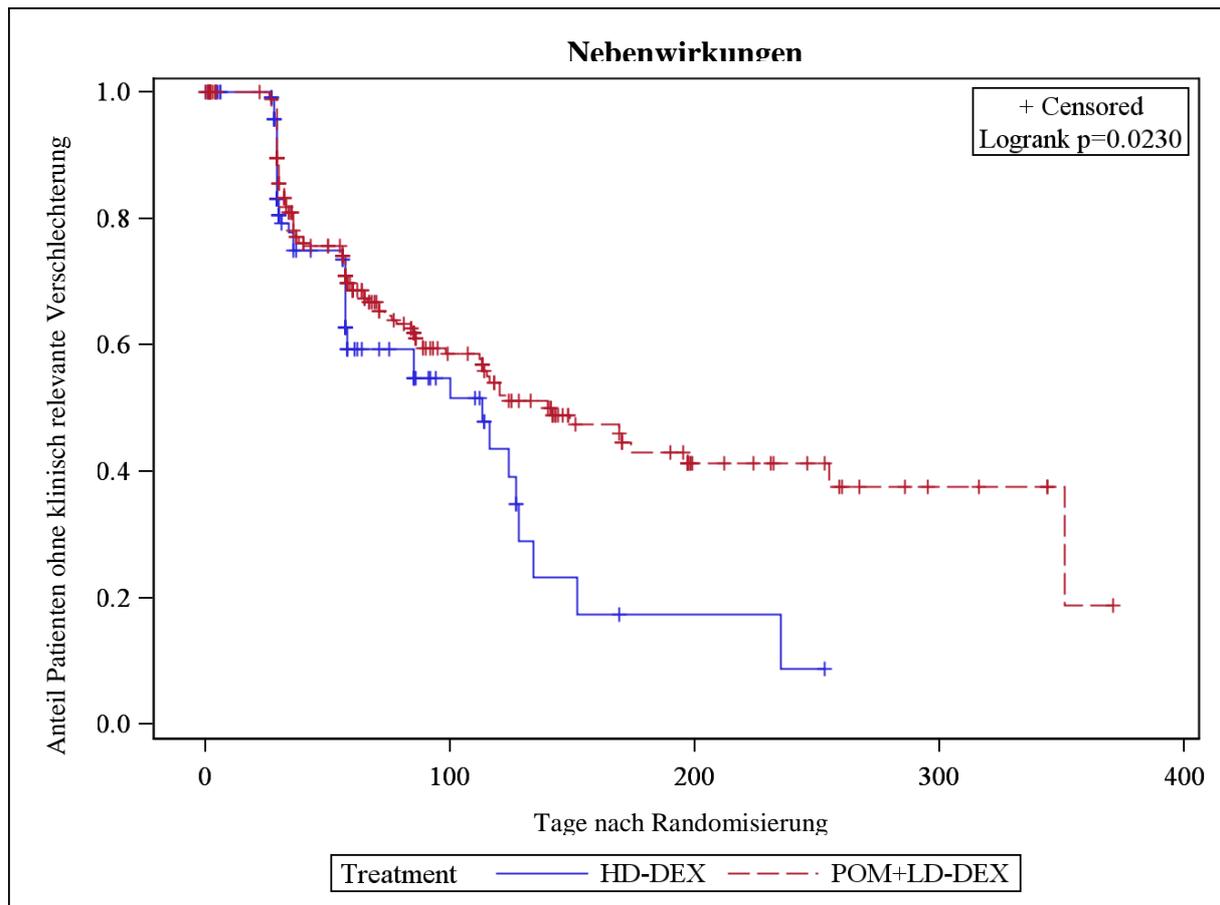


Abbildung 4-4: Kaplan-Meier Kurve für die Zeit bis zur Verschlechterung der QLQ-MY20 Skala Nebenwirkungen der Behandlung

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde. Eine Meta-Analyse ist aus diesem Grund nicht durchführbar.

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

4.3.1.3.2 Subgruppenanalysen – RCT

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

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

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

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

Für die Endpunkte *Progressionsfreies Überleben, Gesamtüberleben, Zeit bis zur Progression, Ansprechrate gemäß den IMWG-Kriterien, Ansprechrate gemäß den EBMT-Kriterien, mind. ein UE, mind. ein SUE, mind. ein zum Absetzen des Studien-Arzneimittel führendes UE, Tod ≤30 Tage nach letzter Dosis Studien-Arzneimittel und mind. ein UE CTC Grad 3 oder 4* wurden Subgruppenanalysen durchgeführt. Diese Endpunkte wurden getrennt nach Geschlecht, Alter, Krankheitspopulation, Anzahl vorheriger Myelomtherapien, zytogenetische Risikogruppe und Durie-Salmon-Stadium zur Baseline dargestellt und die Ergebnisse der Interaktionstests angegeben (Tabelle 4-44 und Tabelle 4-45).

Tabelle 4-44: Ergebnisse der Interaktionstests der Endpunkte PFS, OS, TTP, Ansprechen gemäß den IMWG-Kriterien, Ansprechen gemäß den EBMT-Kriterien

	PFS	OS	TTP	Ansprechrate nach IWMG-Kriterien	Ansprechrate gemäß den EBMT-Kriterien
Alter (≤75 J, >75 J)	0,453	0,335	0,455	0,973 ^a	0,974 ^a
Krankheits-Population (1, 2, 3)*	0,354	0,435	0,441	0,877 ^a	0,736 ^a
Vorherige Myelomtherapien (2, >2)	0,801	0,730	0,467	0,978 ^a	0,979 ^a
Geschlecht (m, w)	0,466	0,343	0,684	0,725	0,383
Zytogenetische Risikogruppe (High risk**, non-high risk***)	0,611	0,624	0,755	0,977 ^a	0,978 ^a
Durie-Salmon-Stadium an Baseline (I, II, III)	0,861	0,555	0,897	0,963 ^a	0,909 ^a

Cox Proportional Hazard Modell mit Behandlung nach Subgruppe als fixem Term.

*Gruppe 1: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. Gruppe 2: Rezidierte und refraktäre Erkrankung mit Ansprechen ≥PR und Progression innerhalb von 6 Monaten. Gruppe 3: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥2 Bortezomibzyklen.

**Jegliche zytogenetische Anomalie in 13q14, 17p13, 4p16/14q32 oder 14q32/16q23

***Jegliche zytogenetische Anomalie in 17p13 oder 4p16/14q32 (Untergruppe von High Risk)

^aAufgrund der Nullzählung (*zero count*) in einer der Interaktions-Subgruppen ist die Validität der Modellrechnung nicht gesichert. Die gezeigten Ergebnisse basieren auf der letzten Maximum Likelihood Iteration.

Quelle: (Celgene Europe Limited, 2013e)

Tabelle 4-45: Ergebnisse der Interaktionstests der Sicherheitsendpunkte

	Mind. 1 UE	Mind. 1 UE Grad 3 oder 4	Mind. 1 SUE	Mind. 1 zum Absetzen des Studien-AM führendes UE	Tod ≤30 d nach letzter Dosis Stu- dien-AM
Alter (≤75 J, >75 J)	0,998 ^a	0,656	0,596	0,998 ^a	0,842
Krankheits-Population (1, 2, 3)*	0,998 ^a	0,191^a	0,376	0,999 ^a	0,535
Vorherige Myelom- therapien (2, >2)	0,502	0,544	0,656	0,974 ^a	0,641
Geschlecht (m, w)	0,584	0,742	0,431	0,402	0,857
Zytogenetische Risikogruppe (High risk**, non-high risk***)	0,864	0,466	0,592	0,789	0,340
Durie-Salmon-Stadium an Baseline (I, II, III)	0,822 ^a	0,571	0,335	0,963 ^a	0,497

Cox Proportional Hazard Modell mit Behandlung nach Subgruppe als fixem Term.

*Gruppe 1: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. Gruppe 2: Rezidierte und refraktäre Erkrankung mit Ansprechen ≥PR und Progression innerhalb von 6 Monaten. Gruppe 3: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥2 Bortezomibzyklen.

**Jegliche zytogenetische Anomalie in 13q14, 17p13, 4p16/14q32 oder 14q32/16q23

***Jegliche zytogenetische Anomalie in 17p13 oder 4p16/14q32 (Untergruppe von High Risk)

^aAufgrund der Nullzählung (*zero count*) in einer der Interaktions-Subgruppen ist die Validität der Modellrechnung nicht gesichert. Die gezeigten Ergebnisse basieren auf der letzten Maximum Likelihood Iteration.

Quelle: (Celgene Europe Limited, 2013c)

Im Folgenden werden für alle Subgruppen, für die mindestens ein Hinweis auf eine Effektmodifikation vorliegt (p-Wert des Interaktionstests <0,2), die detaillierten Ergebnisse vorgestellt. Dies ist nur für ein UE mit CTC Grad 3/4 nach Krankheitspopulation der Fall. Kein Beleg (p-Wert des Interaktionstests <0,05) wurde gefunden.

Tabelle 4-46: Detaillierte Ergebnisse der Subgruppenanalysen des Endpunkts UE mit Grad 3/4 nach Krankheitspopulation

Studie	Subgruppe	Pom+LD-Dex			HD-Dex			OR [95 %-KI p-Wert]**
		Anzahl Studien- teilnehmer	Patienten mit UE		Anzahl Studien- teilnehmer	Patienten mit UE		
			Anzahl	%		Anzahl	%	
Mind. 1 UE Grad 3 oder 4								
MM-003	Krankheitspopulation*							
	1	247	193	78,1	121	87	71,9	1,397 [0,849-2,299]; 0,1954
	2	8	7	87,5	5	5	100	n. b. [n. b.; n. b.]; 1,000
	3	45	34	75,6	23	21	91,3	0,294 [0,059; 1,461]; 0,1925

Cox Proportional Hazard Modell mit Behandlung nach Subgruppe als fixem Term.

n. b. nicht berechenbar; OR: *Odds ratio*; KI: Konfidenzintervall

**Gruppe 1*: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. *Gruppe 2*: Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten. *Gruppe 3*: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥ 2 Bortezomibzyklen.

**p-Wert basierend auf dem *Fisher Exact Test*

Quelle: (Celgene Europe Limited, 2013d)

Die Ergebnisse der Subgruppenanalysen sind unter folgenden Gesichtspunkten zu beurteilen. Die Studie MM-003 war nicht auf Subgruppenanalysen gepowert und aufgrund der Vielzahl der durchgeführten Interaktionstests kann das Auftreten signifikanter Werte zufällig erfolgen. Bei der Interpretation der Resultate sind zudem die kleinen Fallzahlen innerhalb der Subgruppen zu berücksichtigen.

Nur für eine Kombination (UE mit Grad CTC 3/4 in den unterschiedlichen Krankheitspopulationen) wurde ein p-Wert kleiner 0,2 beobachtet ($p=0,191$). Ein Einfluss der Krankheitspopulation auf den Behandlungseffekt bzgl. anderer Endpunkte wurde nicht gefunden.

Unter Berücksichtigung der Vielzahl der durchgeführten Subgruppenanalysen (*Multiple Tests*) ist ein einzelner p-Wert kleiner 0,2 nicht auffällig, und da dieser Hinweis bzgl. der Krankheitspopulation nicht bei weiteren Endpunkten gefunden wurde, kann gefolgert werden, dass die durchgeführten Subgruppenanalysen keinen Anhaltspunkt für einen unterschiedlichen Behandlungseffekt in den a-priori geplanten Subgruppen ergeben.

Dies bestätigt auch die EMA im *CHMP Assessment Report* für die Endpunkte PFS und OS: “PFS and OS analyses were consistently favourable for Pom+LD-dex group in all subgroups” (PFS und OS Analysen zeigten durchgehend Vorteile für den Pom+LD-Dex-Arm in allen Subgruppen) (European Medicines Agency (EMA), 2013a).

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Der medizinische Nutzen und medizinische Zusatznutzen wird auf Basis der randomisierten kontrollierten Studie MM-003 bewertet. Die betrachteten Nutzendimensionen sind *Mortalität, Morbidität, Lebensqualität und Sicherheit* mit ihren weitergehenden Operationalisierungen. Primärer Endpunkt der Studie war das *Progressionsfreie Überleben*. Die Studie war ebenfalls für den Endpunkt *Gesamtüberleben* gepowert, dieser wurde aber als sekundärer Endpunkt klassifiziert. Weitere patientenrelevante Endpunkte waren *Ansprechen* (gemäß den IMWG- und EBMT-Kriterien), *Zeit bis zur Progression*, *Klinisches Ansprechen*, *Lebensqualität* und *unerwünschte Ereignisse*. Meta-Analysen wurden nicht durchgeführt, da eine einzelne RCT die Grundlage der Nutzenbewertung bildet.

Mortalität

Für die Kombination von Pomalidomid mit LD-Dexamethason wurde eine statistisch signifikante und therapeutisch relevante Verlängerung des Gesamtüberlebens gezeigt. Im Interventionsarm trat dabei bis zum Zeitpunkt der Interimsanalyse eine nicht ausreichende Anzahl von Ereignissen auf, um das mediane Gesamtüberleben berechnen zu können (nicht erreicht vs. 34,0 Wochen). Das mediane Gesamtüberleben wurde zu diesem Zeitpunkt auf mindestens 48 Wochen veranschlagt, da dies die untere Grenze des Konfidenzintervalls bildete. Die Behandlung mit Pom+LD-Dex war im Vergleich zu einer Therapie mit HD-Dex mit einer Verringerung des Sterberisikos von 47 % verbunden (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$). Diese Daten sind vor dem Hintergrund des *cross-overs* im Kontrollarm zu betrachten (29 % der Patienten aus dem Kontrollarm wechselten bei Progress zu einer Pomalidomid-Monotherapie). Der tatsächliche Behandlungseffekt von Pom+LD-Dex im Vergleich zu HD-Dex wird daher unterschätzt. Dies bedeutet, dass das Sterberisiko vermutlich noch stärker vermindert und die gewonnene Überlebenszeit noch stärker verlängert wird. Eine statistische Modellierung nach Robins und Tsiatis (1991) ergab, dass der Median im HD-Dex ohne den Vorteil eines *cross-overs* sogar nur bei 26,8 Wochen liegen würde (Robins und Tsiatis, 1991).

Das oben dargestellte Ergebnis beruht auf der Interimsanalyse des Gesamtüberlebens, die bereits ein signifikantes Ergebnis zeigte und zu einem vorzeitigen erfolgreichen Ende der Studie führte (Anmerkung: die Nachbeobachtung der Patienten dauert noch an). Zusätzlich wurde das Gesamtüberleben auch zum Zeitpunkt der ursprünglich geplanten finalen Analyse ausgewertet (1. März 2013). Hier wurde trotz erheblichem *cross-over* (50,0 % der Patienten im HD-Dex-Arm wurden mit Pomalidomid behandelt; das IDMC empfahl bereits zum Zeitpunkt der OS Interimsanalyse die Behandlung der Patienten des Kontrollarms mit Pomalido-

mid fortzusetzen) weiterhin ein signifikanter Vorteil von Pom+LD-Dex beobachtet mit einem medianen Gesamtüberleben von 54,6 Wochen im Pom+LD-Dex-Arm und 34,8 Wochen im HD-Dex-Arm (HR [95 %-KI]: 0,74 [0,56-0,97]; $p < 0,028$). Dies bedeutet einen Überlebensvorteil von 4,6 Monaten (19,8 Wochen). Für März 2013 führte eine Adjustierung nach Robins und Tsiatis für HD-Dex zu einem medianen OS von 29,3 Wochen.

Morbidität

Bezogen auf den primären Endpunkt Progressionsfreies Überleben zeigte sich ebenfalls eine deutliche Überlegenheit der Interventionstherapie. Das Progressionsfreie Überleben der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen der Kontrollgruppe im Median um 7,7 Wochen (15,7 vs. 8,0 Wochen) verlängert. Die hiermit assoziierte relative Risikoreduktion für das Eintreten von Krankheitsprogression oder Tod betrug 55 % (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$), war statistisch signifikant und kann als therapeutisch relevant angesehen werden. Die Auswertung zum Zeitpunkt 1. März 2013 bestätigte den Vorteil für Pom+LD-Dex in Bezug auf das Progressionsfreie Überleben nach Prüfarztbewertung (17,2 vs. 8,2 Wochen; HR [95 %-KI]: 0,48 [0,39-0,60]; $p < 0,001$).

Die Zeit bis zur Progression der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen des Kontrollarms um 11,8 Wochen verlängert. Der Unterschied (20,1 vs. 8,3 Wochen) war statistisch signifikant und mit einer relativen Risikoreduktion für das Auftreten einer Progression um 58 % verbunden (HR [95 %-KI]: 0,42 [0,31-0,56]; $p < 0,001$). Zum Auswertungszeitpunkt 1. März 2013 unterschied sich die Zeit bis zur Progression nach Prüfarztbewertung weiterhin signifikant mit 20,2 Wochen unter Pom+LD-Dex und 9,0 Wochen unter HD-Dex.

Bezogen auf den Endpunkt *Gesamtansprechen* zeigte sich eine klinische Überlegenheit von Pom+LD-Dex gegenüber einer Behandlung mit HD-Dex.

Das *Gesamtansprechen gemäß den IMWG-Kriterien* war in der Interventionsgruppe zum Auswertungszeitpunkt 7. September 2012 mehr als viermal höher als im HD-Dex-Arm (16,6 vs. 3,9 %; OR [95 %-KI]: 4,86 [2,03-11,61]; $p < 0,001$). Bei keinem Studienteilnehmer wurde ein stringentes komplettes Ansprechen beobachtet. Studienteilnehmer der Interventionsgruppe erreichten jedoch häufiger ein komplettes und ein sehr gutes partielles Ansprechen als jene des Kontrollarms (0,3 vs. 0,0 % und 1,3 vs. 0,7 %).

Das *Ansprechen gemäß den EBMT-Kriterien* bestätigt die Daten aus der Auswertung des Ansprechens gemäß den IMWG-Kriterien. Das Ansprechen gemäß den EBMT-Kriterien war zum Auswertungszeitpunkt 7. September 2012 im Interventionsarm mehr als viermal höher als im Kontrollarm (14,9 vs. 3,3 %; OR [95 %-KI]: 5,18 [2,01-13,35]; $p < 0,001$).

Untersucht wurde ebenfalls das sogenannte Objektive Ansprechen (*Objective Response*; CR+PR+MR), bei dem das „Minimale Ansprechen“ (*minimal response*; MR) mit einbezogen wird (Bladé et al., 1998). Es war im Interventionsarm signifikant höher als im Kontrollarm (26,8 vs. 6,5 %; OR [95 %-KI]: 5,24 [2,63-10,45]; $p < 0,001$).

Für den Auswertungszeitpunkt 1. März 2013 betrug das Gesamtansprechen nach Prüfarztbewertung gemäß den IMWG-Kriterien 31,0 % im Pom+LD-Dex-Arm und 10,0 % im HD-Dex-Arm (OR [95 %-KI]: 4,22 [2,35-7,58]; $p < 0,001$). Patienten im Interventionsarm erreichten häufiger ein komplettes oder sehr gutes partielles Ansprechen als jene des Kontrollarms (1,0 vs. 0,0 % und 5,0 vs. 0,7 %). Ein Patient im Interventionsarm (0,3 %) erreichte eine stringente komplette Remission. Das Ansprechen gemäß den EBMT-Kriterien wurde zu diesem Auswertungszeitpunkt nur für die Kategorie minimales Ansprechen (MR) erhoben; dieses wurde von Patienten im Interventionsarm nicht signifikant häufiger erreicht (8,0 vs. 6,0 %; $p = 0,565$).

Bezogen auf den Endpunkt *Klinisches Ansprechen* zeigte sich ein Vorteil für die Therapie mit Pom+LD-Dex bzgl. der Verbesserung des *ECOG Performance Status*. Signifikant mehr Studienteilnehmer des Interventionsarms profitierten von einer Verbesserung als jene der Kontrollgruppe (20,7 vs. 11,1 %; OR [95 %-KI]: 2,09 [1,09-4,02]; $p = 0,027$). Ebenso wurde für mehr Patienten im Pom+LD-Dex-Arm eine Verbesserung von Knochenschmerzen beobachtet. Studienteilnehmer des Interventionsarms profitierten häufiger von einer Verbesserung von Knochenschmerzen als jene der Kontrollgruppe (38,2 vs. 32,6 %; OR [95 %-KI]: 1,28 [0,75-2,18]; $p = 0,422$). Dieser Unterschied war statistisch nicht signifikant.

Lebensqualität

Die Lebensqualität wurde durch den EORTC QLQ-C30 Fragebogen, den krankheitsspezifischen QLQ-MY20-Fragebogen und den EuroQoL EQ-5D-Fragebogen jeweils an Tag 1 eines jeden Behandlungszyklus' à 28 Tage sowie zum Zeitpunkt des Behandlungsabbruchs bzw. Zeitpunkt der Progression erhoben.

Die Adhärenz der zum jeweiligen Erhebungszeitpunkt noch in der Studie befindlichen Studienteilnehmer war im Studienverlauf über alle Zyklen für alle drei Erhebungsinstrumente hoch und war in beiden Gruppen vergleichbar.

Die Ergebnisse zur Lebensqualität waren heterogen. Zusammenfassend lässt sich feststellen, dass die Kombination der zwei Substanzen Pomalidomid und niedrig-dosiertes Dexamethason im Vergleich zur Therapie mit hoch-dosiertem Dexamethason als Einzelsubstanz in Summe die Lebensqualität der Patienten erhielt. In einzelnen Aspekten war vor allem in frühen Behandlungszyklen ein Hinweis auf einen positiven Effekt gegenüber dem Kontrollarm zu beobachten. Insbesondere konnte gezeigt werden, dass das unter Pom+LD-Dex signifikant verbesserte Ansprechen sowie die signifikante Verzögerung der Progression zu einer Verbesserung der Lebensqualität bzw. einer Verringerung von Symptomen führen.

Sicherheit

Bei nahezu allen Studienteilnehmern beider Arme wurde mindestens ein unerwünschtes Ereignis (Pom+LD-Dex 97,0 vs. HD-Dex 93,3 %) und bei etwa der Hälfte der Studienteilnehmer beider Gruppen mindestens ein schwerwiegendes unerwünschtes Ereignis (51,0 vs. 50,3 %) beobachtet. Die Unterschiede zwischen den Gruppen waren jeweils statistisch nicht signifikant. Die Gesamtrate an Studienteilnehmern mit mindestens einem unerwünschten Ereignis mit CTC Grad 3/4 war in beiden Gruppen vergleichbar (78,0 vs. 75,8 %). Der Anteil jener, die mindestens ein zum Absetzen des Studienarzneimittels führendes unerwünschtes Ereignis hatten, war in der Interventionsgruppe nicht signifikant erhöht (9,7 vs. 5,4 %). Der Anteil jener Teilnehmer, die innerhalb von 30 Tagen nach Gabe der letzten Dosis des Studienarzneimittels verstarben, war in der Interventionsgruppe niedriger als im Kontrollarm (12,7 vs. 14,1 %). Der Unterschied war statistisch nicht signifikant.

Bei Studienteilnehmern, die mit Pom+LD-Dex behandelt wurden, wurden signifikant häufiger Neutropenien (41,7 vs. 14,8 %), Leukopenien (8,7 vs. 3,4 %), febrile Neutropenien (6,7 vs. 0,0 %) und hämatopoetische Leukopenien (47,7 vs. 15,4 %) der Grade 3 und 4 beobachtet. Es ist anzumerken, dass eine Leukopenie in den meisten Fällen auf eine Verringerung der Zahl der neutrophilen Granulozyten, also eine Neutropenie, zurückzuführen ist. Diese hämatologischen Nebenwirkungen sind mit Dosisreduktionen und/oder dem Einsatz von Wachstumsfaktoren im Allgemeinen gut handhabbar. Unter Berücksichtigung der längeren Beobachtungszeit von Pom+LD-Dex vs. HD-Dex wurde lediglich für Grad 3/4 Neutropenie und febrile Neutropenie ein signifikanter Unterschied beobachtet.

Ansonsten gab es keine statistisch signifikanten Unterschiede bezogen auf CTC Grad 3/4 UE zwischen Interventions- und Kontrollgruppe. In beiden Armen entwickelten bis zum 7. September 2012 0,7 % der Studienteilnehmer mindestens ein sekundäres Primärmalignom. Die beobachteten sekundären Primärmalignome waren nicht-invasive Basalzellkarzinome der Haut; invasive solide oder hämatologische Neoplasien wurden nicht beobachtet. Zum 1. März 2013 hatten zwei Patienten im Pom+LD-Dex-Arm ein invasives, solides SPM entwickelt. Es gibt bislang – insbesondere vor dem Hintergrund der multiplen Vortherapien der Patienten – keine Hinweise auf einen Zusammenhang zwischen Pomalidomid und der Entstehung von SPM. Im HD-Dex-Arm wurden keine weiteren SPM beobachtet.

Nur 6,6 % der Patienten im Interventionsarm brachen die Therapie aufgrund von unerwünschten Ereignissen ab (5,9 % im Kontrollarm; statistisch nicht signifikant unterschiedlich). Die Auswertung der unerwünschten Ereignisse bis zum 1. März 2013 bestätigte das beschriebene Sicherheitsprofil.

Subgruppen

Subgruppenanalysen wurden für die Faktoren *Geschlecht, Alter, Krankheitspopulation¹⁷, Anzahl vorheriger Myelomtherapien, zytogenetischer Risikogruppe und Durie-Salmon-Stadium* durchgeführt. Es ergab sich kein Anhaltspunkt für einen unterschiedlichen Behandlungseffekt in den einzelnen Subgruppen. Dies bestätigt auch die EMA im *CHMP Assessment Report: “PFS and OS analyses were consistently favourable for Pom+LD-dex group in all subgroups”* (PFS und OS Analysen zeigten durchgehend Vorteile für den Pom+LD-Dex-Arm in allen Subgruppen) (European Medicines Agency (EMA), 2013a).

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht anwendbar.

Der Vergleichsarm der Zulassungsstudie ist die zu verwendende Vergleichstherapie und daher wurden keine indirekten Vergleiche durchgeführt.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

¹⁷ *Gruppe 1:* Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. *Gruppe 2:* Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten nach Ende der Lenalidomid- und/oder Bortezomibtherapie. *Gruppe 3:* Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach \geq 2 Bortezomibzyklen

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-47: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Le- bensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

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

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

Tabelle 4-48: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	Intervention	<Vergleichs- therapie 1>	<Vergleichs- therapie 2>	<Vergleichs- therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2>	•		•	
	<Studie 3>	•		•	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

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

Tabelle 4-49: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

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

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

Tabelle 4-50: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

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

Tabelle 4-51: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

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

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

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

Nicht anwendbar, da keine indirekten Vergleiche durchgeführt wurden.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht anwendbar. Nicht randomisierte vergleichende Studien wurden nicht in die Bewertung mit einbezogen.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-52: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Andere Aspekte, die Verzerrungen verursachen können
			Patient	Behandler		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Nicht anwendbar. Nicht randomisierte vergleichende Studien wurden nicht in die Bewertung mit einbezogen.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-53: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Nicht anwendbar. Nicht randomisierte vergleichende Studien wurden nicht in die Bewertung mit einbezogen.

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

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

Tabelle 4-54: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

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

Nicht anwendbar. Nicht randomisierte vergleichende Studien wurden nicht in die Bewertung mit einbezogen.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

Nicht anwendbar. Nicht randomisierte vergleichende Studien wurden nicht in die Bewertung mit einbezogen.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern* Resultierender Studienpool aus den einzelnen Suchschritten

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungspotenzial.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen**4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen**

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

Tabelle 4-55: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

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

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

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

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

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

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

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

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht anwendbar. Keine weiteren Untersuchungen wurden in die Bewertung einbezogen.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Der Zusatznutzen für die Therapie mit Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason im Vergleich zu hoch-dosiertem Dexamethason gilt durch die Orphan Drug Zulassung als **belegt**. Für den zusätzlichen Daten-basierten Nachweis zum Ausmaß des Zusatznutzens wurden randomisierte, kontrollierte Studien berücksichtigt, deren Methodik internationalen Standards und evidenzbasierter Medizin entspricht. Die hier relevante Studie MM-003 entspricht der Evidenzstufe Ib. Sie vergleicht die Intervention Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) mit der Vergleichstherapie hoch-dosiertes Dexamethason (HD-Dex).

Eine einzelne Studie genügt für einen Nutzenbeleg, wenn sie die besonderen Qualitätskriterien gemäß „CPMP. Points to consider on application with: 1. meta-analyses; 2. one pivotal study“ erfüllt. Dies ist auch in den IQWiG-Methoden „Allgemeine Methoden – Version 4.0“ dargelegt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011b, European Agency for the Evaluation of Medicinal Products (EMA) - Committee for Proprietary Medicinal Products (CPMP), 2001). MM-003 erfüllt die dort genannten Kriterien, wie aufgrund der Zulassung von Pomalidomid auf Basis dieser Studie durch die EMA bestätigt wird.

Die klinische Studie MM-003 wurde anhand des *CONSORT-Statements* (Tabelle 4-61) und der *Cochrane-Kriterien* (Tabelle 4-62) zur Bewertung des Verzerrungspotenzials der Ergebnisse weiterhin bezüglich der Planungs-, Durchführungs- und Auswertungsqualität, überprüft.

Die Auswertung der Daten erfolgte nach dem ITT-Prinzip. Der Anteil an Studienteilnehmern, die die Studie vorzeitig beendeten ohne den primären Endpunkt zu erreichen, unterschied sich in beiden Gruppen nicht maßgeblich und ist transparent beschrieben. Zwar wurde die Studie unverblindet durchgeführt, jedoch wurde das Ansprechen sowie die Progression durch ein *Independent Response Adjudication Committee* (IRAC, Unabhängiges Komitee zur Beurteilung des Ansprechens) für den Auswertungszeitpunkt 7. September 2012 verblindet verifiziert, um eine verzerrungsfreie Beurteilung der Daten zu gewährleisten.

Es existieren keine Hinweise für eine ergebnisgesteuerte Berichterstattung der Ergebnisse für die Endpunkte *PFS/TTP*. Ebenso existieren keine sonstigen Anhaltspunkte für eine verzerrende inhaltliche oder formale Darstellung der Ergebnisse. Die Ergebnisse dieser Endpunkte zum Zeitpunkt der Entblindung sind daher mit einer hohen Ergebnissicherheit assoziiert. Das Gleiche gilt für das Ansprechen auf die Therapie. Auch dieses wurde durch das IRAC verblindet festgelegt. Insofern ist auch für den Endpunkt *Ansprechen* (sowohl gemäß IMWG als auch EBMT-Kriterien) von einer hohen Ergebnissicherheit auszugehen.

Für den Endpunkt *Gesamtüberleben* liegt ein erhöhtes Verzerrungspotenzial vor, da die Patienten im Kontrollarm nach einer Progression durch Wechsel in eine Begleitstudie (MM-003/C) Pomalidomid erhalten konnten. Dies geschah bei 29 % der Patienten in diesem Arm (Datenstand 7. September 2012). Aufgrund der zu diesem Zeitpunkt durchgeführten erfolgreichen OS-Interimsanalyse empfahl das IDMC die Behandlung der Patienten des Kontrollarms mit Pomalidomid fortzusetzen. Bei der finalen Auswertung des Gesamtüberlebens (1. März 2013) hatten dadurch bereits 50 % der Patienten im Kontrollarm Pomalidomid erhalten. Die Verzerrung ist zuungunsten des Pom+LD-Dex-Arms, da hiermit eine Überschätzung des Therapieeffektes im Kontrollarm einhergeht und somit der Therapievorteil von Pomalidomid unterschätzt wird. Aus diesen Gründen ist von einer hohen Ergebnissicherheit auszugehen. Die Tatsache, dass das Gesamtüberleben unverblindet erhoben wurde, hat keine Auswirkungen auf das Verzerrungspotenzial, da der Eintritt des Todes ein objektives Kriterium ist und somit dessen Feststellung frei von jeglicher Subjektivität erfolgt.

Für die Endpunkte *Klinisches Ansprechen*, *Sicherheit* und *Lebensqualität* liegt hingegen aufgrund der unverblindeten Erhebung ein erhöhtes Verzerrungspotenzial und eine mäßige Ergebnissicherheit vor.

Für die Bewertung des Zusatznutzens wurden nur patientenrelevante Endpunkte verwendet. Das Gesamtüberleben ist patientenrelevant und der am meisten geforderte und belastbarste Endpunkt in der Hämatologie und Onkologie. Die Studie MM-003 war auch für den Endpunkt *Gesamtüberleben* gepowert. Das Ereignis *Progression* für die Indikation *Multiple Myelom* wurde durch das IQWiG als patientenrelevant beschrieben (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011a). Progression ebenso wie das Ansprechen wurden anhand der etablierten Kriterien der *International Myeloma Working Group*

und der *European Group for Blood and Marrow Transplant* erhoben. Die damit zusammenhängenden Endpunkte sind aufgrund der Krankheitsmechanistik patientenrelevant, da sie die Morbidität der Patienten senken (z.B. Reduktion des nephrotoxischen M-Proteins, Milderung von Knochenschmerzen, Verhinderung weiterer Osteolysen), und primäre Therapieziele beim rezidierten und refraktären Multiplen Myelom sind (Straka et al., 2012, Böhrer und Pönisch, 2001). Wegen des direkten Bezuges zu den Grundbedürfnissen eines selbstbestimmten Lebens ist ebenso jegliche Verbesserung des ECOG PS-Scores von unmittelbarer Patientenrelevanz. Das Auftreten von Knochenschmerzen schränkt die Lebensqualität von Myelompatienten ein, was dem Endpunkt *Verbesserung von Knochenschmerzen* ebenfalls Patientenrelevanz verleiht. Durch die Erfassung unerwünschter Ereignisse wird die Verträglichkeit eines Arzneimittels untersucht, die ebenso als patientenrelevant einzustufen ist. Der für onkologische Therapeutika immer wieder deutlich eingeforderte Endpunkt *Lebensqualität* ist *per definitionem* patientenrelevant (§ 2 Absatz 3 AM-NutzenV), was auch vom IQWiG explizit bestätigt wurde (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 2011a, Bundesministerium der Justiz, 2010).

Zusätzlich zum Beleg des Zusatznutzens durch den Orphan Drug Status ist im Hinblick auf die Evidenzstufe, die Studienqualität, die Validität und Patientenrelevanz der Endpunkte der direkte Vergleich zwischen Pom+LD-Dex und HD-Dex ebenfalls geeignet, *Belege* des medizinischen Nutzens und Zusatznutzens abzuleiten.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

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

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

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

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

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

Die Bewertung des medizinischen Zusatznutzens erfolgt in Bezug auf die patientenrelevanten Nutzendimensionen *Mortalität*, *Morbidität*, *Sicherheit* und *Lebensqualität*, die in der dem Dossier zugrunde liegenden Studie weiter operationalisiert wurden. Die Grundlage für die Nutzenbewertung bildet die Phase III-Studie MM-003. Der Zusatznutzen wird anhand des Datenstandes vom 7. September 2012 abgeleitet. Unterstützend werden die Daten zum Zeitpunkt der finalen Auswertung des Gesamtüberlebens (1. März 2013) für die Endpunkte Gesamtüberleben, Progressionsfreies Überleben, Zeit bis zur Progression, Ansprechen gemäß der IMWG- und EBMT-Kriterien (letzteres nur für die Remissionskategorie minimales Ansprechen) sowie Sicherheit zusammenfassend dargestellt.

Tabelle 4-56: Zusammenfassung der Ergebnisse

Endpunkte	Effektschätzer (Pom+LD-Dex vs HD-Dex)	Zusatznutzen
Mortalität Gesamtüberleben	Datenstand: 7. September 2012 HR [95 %-KI]: 0,53 [0,37-0,74]; p<0,001 Median: nicht erreicht vs. 34,0 Wochen Datenstand: 1. März 2013 HR [95 %-KI]: 0,74 [0,56-0,97]; p<0,028 54,6 vs. 34,8 Wochen	erheblich (Beleg)
Morbidität		
Progressionsfreies Überleben	OR [95 %-KI]: 0,45 [0,35-0,59]; p<0,001 Median: 15,7 vs. 8,0 Wochen	erheblich (Beleg)
Zeit bis zur Progression	OR [95 %-KI]: 0,42 [0,31-0,56]; p<0,001 Median: 20,1 vs. 8,3 Wochen	
Ansprechen (IMWG)	OR [95 %-KI]: 4,86 [2,03-11,61]; p<0,001 Häufigkeit: 16,6 vs. 3,9 %	
Verbesserung von Knochenschmerzen des ECOG PS	OR [95 %-KI]: 1,28 [0,75-2,18]; p=0,422 Häufigkeit: 25,2 vs. 18,3 % OR [95 %-KI]: 2,09 [1,09-4,02]; p=0,027 Häufigkeit: 16,9 vs. 8,5 %	
Lebensqualität		
EORTC QLQ-C30 EORTC QLQ-MY20 EuroQoL EQ-5D	kein signifikanter Trend zu erkennen	kein zusätzlicher Schaden belegt
Sicherheit		
Anzahl Patienten mit: einem UE einem SUE einem zum Absetzen des Studien-AM föhrendes UE Tod≤30 d nach letzter Dosis Studien-AM 1 UE CTC Grad 3/4	OR [95 %-KI]: 2,33 [0,92-5,85]; p=0,081 Häufigkeit: 97,0 vs. 93,3 % OR [95 %-KI]: 1,03 [0,69-1,52]; p=0,920 Häufigkeit: 51,0 vs. 50,3 % OR [95 %-KI]: 1,89 [0,84-4,23]; p=0,145 Häufigkeit: 9,7 vs. 5,4 % OR [95 %-KI]: 0,88 [0,50-1,57]; p=0,660 Häufigkeit: 12,7 vs. 14,1 % OR [95 %-KI]: 1,13 [0,71-1,80]; p=0,633 Häufigkeit: 78,0 vs. 75,8 %	kein zusätzlicher Schaden belegt

AM: Arzneimittel; CTC: *Common Terminology Criteria*; ECOG PS: *Eastern Cooperative Oncology Group Performance Status*; EORTC: *European Organisation for Research and Treatment of Cancer*; HR: *Hazard Ratio*; IMWG: *International Myeloma Working Group*; KI: *Konfidenzintervall*; OR: *Odds Ratio*; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis;

Mortalität

Für die Kombination von Pomalidomid mit LD-Dexamethason wurde eine statistisch signifikante und therapeutisch relevante Verlängerung des Gesamtüberlebens gezeigt. Die Studie MM-003 war auf den Endpunkt *Gesamtüberleben* (OS) gepowert (d. h. auch OS wurde konfirmatorisch unter Einhaltung des globalen 5 % Signifikanzniveaus getestet).

Im Interventionsarm traten bis zum Zeitpunkt der Analyse eine nicht ausreichende Anzahl von Ereignissen auf, um das mediane Überleben berechnen zu können (nicht erreicht vs. 34,0 Wochen). Die Behandlung mit Pom+LD-Dex war im Vergleich zu einer Therapie mit HD-Dex mit einer Verringerung des Sterberisikos von 47 % verbunden (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$). Diese Daten sind vor dem Hintergrund des *cross-overs* im Kontrollarm zu betrachten (29 % der Patienten aus dem Kontrollarm wechselten bei Progress zu einer Pomalidomid-Monotherapie), so dass der tatsächliche Behandlungseffekt von Pom+LD-Dex im Vergleich zu HD-Dex daher unterschätzt wird. Dies bedeutet, dass das Sterberisiko vermutlich noch stärker vermindert und die gewonnene Überlebenszeit noch länger werden würde. Eine statistische Modellierung nach Robins und Tsiatis (1991) ergab, dass der Median im HD-Dex-Arm ohne den Vorteil eines *cross-overs* sogar nur bei 26,8 Wochen liegen würde (Robins und Tsiatis, 1991).

Die hier gezeigte Verlängerung des Gesamtüberlebens ist auch angesichts der besonderen Therapiesituation (schwer vorbehandelte Patienten mit fortgeschrittener Erkrankung) besonders bemerkenswert: Pomalidomid verbessert die Prognose dieser Patienten deutlich.

Die Patienten, die zytogenetisch als Hochrisikogruppe klassifiziert werden, stellen eine besonders herausfordernde Patientengruppe mit schlechter Prognose dar. Die Studie MM-003 zeigt, dass die Kombination aus Pom+LD-Dex auch dort wirksam ist (51,3 vs. 24,6 Wochen, HR [95 %-KI]: 0,52 [0,32-0,87]). Das Gesamtüberleben der Hochrisikopatienten war damit nicht signifikant unterschiedlich von dem der ITT-Population. Dies ist hervorzuheben, da andere in der Myelomtherapie eingesetzte Substanzen bei dieser Risikogruppe oft nur eingeschränkt wirksam sind (Avet-Loiseau et al., 2010, Kapoor und Rajkumar, 2011, Klein et al., 2011, Reece et al., 2009).

Auch ältere Patienten (>65 Jahre), die aufgrund gleichzeitig bestehender Komorbiditäten eine herausfordernde Patientengruppe darstellen, profitierten ebenso von Pom+LD-Dex wie die ITT-Population (nicht erreicht vs. 24,9 Wochen, HR [95 %-KI]: 0,56 [0,34-0,91]).

Die oben dargestellten Ergebnisse beruhen auf der Interimsanalyse des Gesamtüberlebens, die bereits zu einem signifikanten Ergebnis und deshalb zu einem vorzeitigen erfolgreichen Ende der Studie führte (Anmerkung: die Nachbeobachtung der Patienten dauert noch an). Zusätzlich wurde das Gesamtüberleben auch zum Zeitpunkt der ursprünglich geplanten finalen Analyse ausgewertet. Hier wurde trotz fortgeschrittenem *cross-over* (bereits 50,0 % der Patienten im HD-Dex-Arm wurden mit Pomalidomid behandelt; das IDMC empfahl bereits zum Zeitpunkt der OS Interimsanalyse die Behandlung der Patienten des Kontrollarms mit Pomalido-

mid fortzusetzen) weiterhin ein signifikanter Vorteil von Pom+LD-Dex mit einem medianen Gesamtüberleben von 54,6 Wochen im Pom+LD-Dex-Arm und 34,8 Wochen im HD-Dex-Arm (HR [95 %-KI]: 0,74 [0,56-0,97]; $p < 0,028$) beobachtet. Dies bedeutet einen Überlebensvorteil von 4,6 Monaten (19,8 Wochen). Für März 2013 ergab die *cross-over*-Adjustierung nach Robins und Tsiatis für HD-Dex ein medianes Gesamtüberleben von 29,3 Wochen.

Gemäß Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV § 5 Absatz 7) liegt ein erheblicher Zusatznutzen u. a. bei einer erheblichen Verlängerung der Überlebensdauer vor (Bundesministerium der Justiz, 2010). Pom+LD-Dex erzielt diesbezüglich gegenüber der Vergleichstherapie eine bisher nicht erreichte große Verbesserung. Für die Nutzendimension Mortalität ist der Zusatznutzen damit **erheblich** und gilt als **belegt**.

Morbidität

Bezogen auf den primären Endpunkt *Progressionsfreies Überleben* zeigte sich ebenfalls eine deutliche Überlegenheit der Interventionstherapie. Das Progressionsfreie Überleben der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen der Kontrollgruppe nahezu verdoppelt (15,7 vs. 8,0 Wochen). Die hiermit assoziierte relative Risikoreduktion für das Eintreten von Krankheitsprogression oder Tod betrug 55 % (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$) und war statistisch signifikant. Die Auswertung zum Zeitpunkt 1. März 2013 bestätigte den Vorteil für Pom+LD-Dex in Bezug auf das Progressionsfreie Überleben nach Prüfarztbewertung (17,2 vs. 8,2 Wochen; HR [95 %-KI]: 0,48 [0,39-0,60]; $p < 0,001$). Eine derartige Verlängerung des Progressionsfreien Überlebens ist auch vor dem Hintergrund der besonderen Therapiesituation (es handelt sich um vielfach vortherapierte Patienten ohne weitere befriedigende Therapieoption) besonders bemerkenswert.

Die *Zeit bis zur Progression* der Studienteilnehmer, die mit Pom+LD-Dex behandelt wurden, war gegenüber jenen des Kontrollarms um 11,8 Wochen verlängert. Der Unterschied (20,1 vs. 8,3 Wochen) war statistisch signifikant und mit einer relativen Risikoreduktion für das Auftreten einer Progression um 58 % verbunden (HR [95 %-KI]: 0,42 [0,31-0,56]; $p < 0,001$). Zum Auswertungszeitpunkt 1. März 2013 unterschied sich die Zeit bis zur Progression nach Prüfarztbewertung weiterhin signifikant mit 20,2 Wochen unter Pom+LD-Dex und 9,0 Wochen unter HD-Dex.

Bezogen auf den Endpunkt *Gesamtansprechen* zeigte sich eine Überlegenheit von Pom+LD-Dex gegenüber einer Behandlung mit HD-Dex.

Das *Gesamtansprechen* gemäß *IMWG-Kriterien* war in der Interventionsgruppe zum Auswertungszeitpunkt 7. September 2012 mehr als viermal höher als im HD-Dex-Arm (OR [95 %-KI]: 4,86 [2,03-11,61]; $p < 0,001$). Es ist anzumerken, dass diejenigen Patienten mit mindestens einer partiellen Remission eine mediane Ansprechdauer (und damit eine Zeit ohne relevante Proliferation des Tumors) von 32,0 Wochen hatten.

Das *Ansprechen* gemäß *EBMT-Kriterien* bestätigt die Daten aus der Auswertung des Ansprechens gemäß *IMWG-Kriterien*. Das Ansprechen gemäß den *EBMT-Kriterien* war zum Aus-

wertungszeitpunkt 7. September 2012 im Interventionsarm ebenfalls mehr als viermal höher als im Kontrollarm (14,9 vs. 3,3 %; OR [95 %-KI]: 5,18 [2,01-13,35]; $p < 0,001$).

Untersucht wurde ebenfalls das sogenannte Objektive Ansprechen (*Objective Response*; CR+PR+MR), bei dem das „Minimale Ansprechen“ (*Minimal Response*; MR) mit einbezogen wird (Bladé et al., 1998): Es war zum Datenstand 7. September 2012 im Interventionsarm signifikant höher als im Kontrollarm (26,8 vs. 6,5 %; OR [95 %-KI]: 5,24 [2,63-10,45]; $p < 0,001$).

Für den Auswertungszeitpunkt 1. März 2013 betrug das Gesamtansprechen nach Prüfarztbewertung gemäß den IMWG-Kriterien 31,0 % im Pom+LD-Dex-Arm und 10,0 % im HD-Dex-Arm (OR [95 %-KI]: 4,22 [2,35-7,58]; $p < 0,001$). Patienten im Interventionsarm erreichten häufiger ein komplettes oder sehr gutes partielles Ansprechen als jene des Kontrollarms (1,0 vs. 0,0 % und 5,0 vs. 0,7 %). Ein Patient im Interventionsarm (0,3 %) erreichte eine stringente komplette Remission. Das Ansprechen gemäß den EBMT-Kriterien wurde zu diesem Auswertungszeitpunkt nur für die Kategorie minimales Ansprechen (MR) erhoben; dieses wurde von Patienten im Interventionsarm nicht signifikant häufiger erreicht (8,0 vs. 6,0 %; $p = 0,565$).

Auch die Lebensqualitätsanalysen der MM-003 bestätigen einen morbiditätsrelevanten Effekt von Pom-LD-Dex. Während unter der Therapie mit HD-Dex die mediane Zeit bis zur ersten Verschlechterung von Fatigue, einem relevanten Krankheitssymptom bei Patienten mit rrMM, 60 Tage betrug [95 %-KI]: [57-113], war diese Zeit unter Pom+LD-Dex mit 113 Tagen [95 %-KI]: [71-169] signifikant verlängert ($p = 0,038$).

Bezogen auf den Endpunkt *Klinisches Ansprechen* zeigte sich ein Vorteil für die Therapie mit Pom+LD-Dex bzgl. der Verbesserung des *ECOG Performance Status*. Signifikant mehr Studienteilnehmer des Interventionsarms profitierten von einer Verbesserung als jene der Kontrollgruppe (20,7 vs. 11,1 %; OR [95 %-KI]: 2,09 [1,09-4,02]; $p = 0,027$). Ebenso wurde für mehr Patienten im Pom+LD-Dex-Arm eine Verbesserung von Knochenschmerzen beobachtet. Studienteilnehmer des Interventionsarms profitierten häufiger von einer Verbesserung von Knochenschmerzen als jene der Kontrollgruppe (38,2 vs. 32,6 %; OR [95 %-KI]: 1,28 [0,75-2,18]; $p = 0,422$). Dieser Unterschied war statistisch nicht signifikant.

Gemäß AM-NutzenV § 5 Absatz 7 liegt ein erheblicher Zusatznutzen u. a. beim Erreichen einer langfristigen Freiheit von schwerwiegenden Symptomen vor (Bundesministerium der Justiz, 2010). Pom+LD-Dex erzielt im Vergleich zu HD-Dex eine signifikant verlängerte Progressionsfreiheit und höhere Ansprechraten, was zu einer Verminderung der krankheitsbedingten Symptome beiträgt. Dies zeigt sich darin, dass mehr Patienten unter Pom+LD-Dex Verbesserungen in Bezug auf Knochenschmerzen und ECOG PS Score zeigen. Aufgrund dieser bisher nicht erreichten Verbesserung ist der Zusatznutzen in der Nutzendimension Morbidität **erheblich** und gilt als **belegt**.

Lebensqualität

Die Lebensqualität wurde anhand der Fragebögen EORTC QLQ-C30, EORTC QLQ-MY20 und EuroQoL EQ-5D untersucht.

Die Adhärenz der zum jeweiligen Erhebungszeitpunkt noch in der Studie befindlichen Studienteilnehmer war im Studienverlauf über alle Zyklen für alle drei Erhebungsinstrumente hoch und war in beiden Gruppen vergleichbar.

Die Ergebnisse zur Lebensqualität waren heterogen. Zusammenfassend lässt sich feststellen, dass die Kombination der zwei Substanzen Pomalidomid und niedrig-dosiertes Dexamethason im Vergleich zur Therapie mit hoch-dosiertem Dexamethason als Einzelsubstanz in Summe die Lebensqualität der Patienten erhielt. In einzelnen Aspekten war vor allem in frühen Behandlungszyklen sogar ein Hinweis auf einen positiven Effekt gegenüber dem Kontrollarm zu beobachten.

Bezieht man die deutliche Verlängerung der Zeit bis zur Progression sowie die maßgeblich erhöhte Ansprechrate unter Pom+LD-Dex und den Effekt dieser Einflussvariablen auf die Lebensqualität mit in die Gesamtschau ein, so ist von einem verlängerten Erhalt der Lebensqualität unter Pom+LD-Dex im Vergleich zu HD-Dex auszugehen.

Insgesamt ist es bemerkenswert, dass unter der Zweierkombination Pom+LD-Dex die Verlängerung des Gesamtüberlebens ohne Einschränkungen der Lebensqualität im Vergleich zur Monotherapie mit HD-Dex einhergeht. Die Kombinationstherapie Pom+LD-Dex verursacht im Vergleich zur Monotherapie mit HD-Dex also keinen erhöhten Schaden in Bezug auf die Lebensqualität der Patienten.

Sicherheit

Bei nahezu allen Studienteilnehmern beider Arme wurde mindestens ein unerwünschtes Ereignis (Pom+LD-Dex 97,0 vs. HD-Dex 93,3 %) und bei etwa der Hälfte der Studienteilnehmer beider Gruppen mindestens ein schwerwiegendes unerwünschtes Ereignis (51,0 vs. 50,3 %) beobachtet. Die Unterschiede zwischen den Gruppen waren jeweils statistisch nicht signifikant. Die Gesamtrate an Studienteilnehmern mit mindestens einem unerwünschten Ereignis mit CTC Grad 3/4 war in beiden Gruppen vergleichbar (78,0 vs. 75,8 %). Der Anteil jener, die mindestens ein zum Absetzen des Studienarzneimittels führendes unerwünschtes Ereignis hatten, war in der Interventionsgruppe nicht signifikant erhöht (9,7 vs. 5,4 %). Der Anteil jener Teilnehmer, die innerhalb von 30 Tagen nach Gabe der letzten Dosis des Studienarzneimittels verstarben, war in der Interventionsgruppe niedriger als im Kontrollarm (12,7 vs. 14,1 %). Der Unterschied war statistisch nicht signifikant.

Bei Studienteilnehmern, die mit Pom+LD-Dex behandelt wurden, wurden signifikant häufiger Neutropenien (41,7 vs. 14,8 %), Leukopenien (8,7 vs. 3,4 %), febrile Neutropenien (6,7 vs. 0,0 %) und hämatopoetische Leukopenien (47,7 vs. 15,4 %) der Grade 3 und 4 beobachtet. Es ist anzumerken, dass eine Leukopenie in den meisten Fällen auf eine Verringerung der Zahl

der neutrophilen Granulozyten, also eine Neutropenie, zurückzuführen ist. Unter Berücksichtigung der längeren Beobachtungszeit von Pom+LD-Dex vs. HD-Dex wurde lediglich für Grad 3/4 Neutropenie und febrile Neutropenie ein signifikanter Unterschied beobachtet.

Ansonsten gab es keine statistisch signifikanten Unterschiede bezogen auf CTC Grad 3/4 UE zwischen Interventions- und Kontrollgruppe. Periphere Neuropathie vom Grad 3/4, ein die Lebensqualität drastisch einschränkendes UE, welches mit verschiedenen Myelomtherapien assoziiert ist (u. a. Bortezomib, Thalidomid) (Delforge et al., 2010, Richardson et al., 2011), trat kaum auf (1,0 vs. 1,3 %). Ebenso waren sowohl venöse als auch arterielle thromboembolische Ereignisse Grad 3/4 sehr selten, was auf die Effektivität der in der Studie MM-003 angewandten Thromboseprophylaxe hinweist. Sekundäre Primärmalignome wurden kaum beobachtet. Die Auswertung zum 1. März 2013 bestätigte das beschriebene Sicherheitsprofil. Insgesamt war die Inzidenz unerwünschter Nebenwirkungen unter Pom+LD-Dex in den ersten Behandlungszyklen am höchsten und verringerte sich im weiteren Verlauf der Therapie (European Medicines Agency (EMA), 2013a).

Insgesamt ist es bemerkenswert, dass unter der Zweierkombination Pom+LD-Dex nur wenig mehr unerwünschte Ereignisse auftraten als unter der Monotherapie mit HD-Dex. Dies insbesondere vor dem Hintergrund, dass die Studienteilnehmer meist vielfach vortheraPIert waren, was eine erhöhte Belastung durch kumulative Toxizitäten in dieser Population erwarten ließ. Die häufigste Nebenwirkung unter Pom+LD-Dex, die Neutropenie, ist in der Mehrzahl der Fälle als reiner Laborwert zu verstehen, der keine negativen Auswirkungen auf die Lebensqualität der Patienten hat. Erst wenn die Neutropenie mit infektionsbedingten Symptomen einhergeht, wird sie relevant. Grundsätzlich lassen sich Neutropenien durch Dosisreduktionen und/oder die Gabe von Wachstumsfaktoren gut handhaben. Der Etablierung einer Infektion kann mittels Antibiotikagabe vorgebeugt werden (Gay und Palumbo, 2010). Das gute Sicherheitsprofil von Pom+LD-Dex wird auch daran ersichtlich, dass nur bei 6,6 % der Patienten im Interventionsarm die Therapie aufgrund von unerwünschten Ereignissen beendet wurde (Kontrollarm: 5,9 %, statistisch nicht signifikant unterschiedlich).

Die Daten zeigen also, dass die zusätzliche Gabe von Pomalidomid keinen zusätzlichen Schaden bzgl. der Sicherheit für die Patienten bedeutet.

Subgruppen

Subgruppenanalysen wurden für die Faktoren *Geschlecht, Alter, Krankheitspopulation¹⁸, Anzahl vorheriger Myelomtherapien, zytogenetischer Risikogruppe und Durie-Salmon-Stadium* durchgeführt. Es ergab sich kein Anhaltspunkt für einen unterschiedlichen Behandlungseffekt in den unterschiedlichen Subgruppen. Dies bestätigt auch die EMA im *CHMP Assessment Report*: “PFS and OS analyses were consistently favourable for Pom+LD-dex group in all

¹⁸ Gruppe 1: Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib. Gruppe 2: Rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten nach Ende der Lenalidomid- und/oder Bortezomibtherapie. Gruppe 3: Refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach \geq 2 Bortezomibzyklen

subgroups“ (PFS und OS Analysen zeigten durchgehend Vorteile für den Pom+LD-Dex-Arm in allen Subgruppen) (European Medicines Agency (EMA), 2013a).

Abschließende Bewertung

Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason (Pom+LD-Dex) führt im Vergleich zur alleinigen Anwendung von hoch-dosiertem Dexamethason (HD-Dex) zu einer patientenrelevanten Reduktion des Mortalitätsrisikos von 47 % (HR [95 %-KI]: 0,53 [0,37-0,74]; $p < 0,001$). Auch zytogenetische Hochrisikopatienten und ältere Patienten > 65 Jahre profitieren diesbezüglich. Bei der finalen OS-Analyse im März 2013 wurde weiterhin ein signifikanter Vorteil von Pom+LD-Dex mit einem medianen Gesamtüberleben von 54,6 Wochen im Pom+LD-Dex-Arm und 34,8 Wochen im HD-Dex-Arm (HR [95 %-KI]: 0,74 [0,56-0,97]; $p < 0,028$) beobachtet. Dies bedeutet einen Überlebensvorteil von 4,6 Monaten (19,8 Wochen).

Das Progressionsfreie Überleben verdoppelt sich nahezu (15,7 vs. 8,0 Wochen). Das Risiko für das Eintreten von Krankheitsprogression oder Tod wird durch eine Therapie mit Pom+LD-Dex im Vergleich zu HD-Dex um 55 % gesenkt (HR [95 %-KI]: 0,45 [0,35-0,59]; $p < 0,001$). Das Gesamtansprechen gemäß den IMWG-Kriterien war in der Interventionsgruppe zum Auswertungszeitpunkt 7. September 2012 mehr als viermal höher als im HD-Dex-Arm (16,6 vs. 3,9 %; OR [95 %-KI]: 4,86 [2,03-11,61]; $p < 0,001$).

Eine Auswertung des Ansprechens gemäß den EBMT-Kriterien bestätigte die deutliche Verbesserung der Ansprechrate durch Pom+LD-Dex. Da für Patienten mit fortgeschrittener Tumorerkrankung auch bereits ein geringfügiger Rückgang der Tumorlast relevant sein kann, wurde ebenfalls das Minimale Ansprechen bewertet. Das sich daraus ergebende Objektive Ansprechen (\geq MR) zeigt eine deutliche Überlegenheit von Pom+LD-Dex im Vergleich zu HD-Dex (26,8 vs. 6,6 %).

Der Patientennutzen durch die Pomalidomid-induzierte Kontrolle der Erkrankung, d. h. die Verbesserung des Ansprechens sowie die Verlängerung der progressionsfreien Zeit, manifestiert sich auch darin, dass unter Pom+LD-Dex mehr Patienten eine Verbesserung des ECOG PS Scores sowie des Symptoms Knochenschmerzen erfahren. Die allgemeine Symptomlast der Patienten wird somit durch Pom+LD-Dex im Vergleich zu HD-Dex gesenkt.

Die Lebensqualität der Patienten ist unter Pom+LD-Dex vergleichbar zum Kontrollarm, teilweise, insbesondere in frühen Zyklen, sogar besser. Das verbesserte Ansprechen und die verlängerte Zeit bis zur Progression unter Pom+LD-Dex erwiesen sich als Einflussfaktoren für eine verbesserte Lebensqualität.

Vor dem Hintergrund der Schwere der Erkrankung und der bestehenden Vorbelastung der Patienten durch multiple Rezidive und Vortherapien zeigte sich insgesamt eine gute Verträglichkeit von Pom+LD-Dex. Dies wird zusätzlich dadurch unterstrichen, dass nur 6,6 % der Patienten im Interventionsarm die Therapie aufgrund von unerwünschten Ereignissen abbrechen (Kontrollarm: 5,9 %; statistisch nicht signifikant unterschiedlich).

Die Auswertung nach längerer Nachbeobachtungszeit (1. März 2013) bestätigte die genannten Daten.

Gemäß AM-NutzenV § 5 Absatz 7 liegt ein erheblicher Zusatznutzen u. a. vor, wenn die Überlebensdauer erheblich verlängert, eine langfristige Freiheit von Symptomen erzielt oder eine weitgehende Vermeidung schwerwiegender Nebenwirkungen gegenüber der Vergleichstherapie erreicht wird (Bundesministerium der Justiz, 2010).

Pom+LD-Dex erzielt eine in der Indikation bisher nicht erreichte, erhebliche Verlängerung des Gesamtüberlebens bei gleichzeitig gegenüber HD-Dex nicht relevant erhöhter Toxizität. Die Symptomlast der Patienten wird durch die verlängerte Progressionsfreiheit und die höheren Ansprechraten gesenkt, was sich in erhöhten Verbesserungsraten bezüglich Knochenschmerzen und ECOG PS Score widerspiegelt. Demgemäß zeigen die Ergebnisse der Zulassungsstudie MM-003 in Anbetracht der Relevanz der Endpunkte Mortalität, Morbidität, des Ausmaßes der Verringerung der Risiken für Progression und Tod sowie des weitgehenden Erhalts der Lebensqualität und der guten Verträglichkeit einen **erheblichen** Zusatznutzen für Pomalidomid in Kombination mit niedrig-dosiertem Dexamethason gegenüber der alleinigen Anwendung von hoch-dosiertem Dexamethason.

Der oben nachgewiesene Zusatznutzen gilt dem Grunde nach auch durch die Zulassung als Orphan Drug als belegt. Aber auch das beschriebene Ausmaß des Zusatznutzens als „erheblich“ erfährt eine rechtliche Begründung wie nachfolgend gezeigt wird.

Die Ausweisung eines Arzneimittels als Orphan Drug durch die Europäische Kommission ist eine behördliche Feststellung über Qualität, Wirksamkeit und Unbedenklichkeit eines Arzneimittels, der die Verordnungseinschränkungen oder -ausschlüsse des G-BA nicht widersprechen dürfen. Insoweit entfaltet diese Feststellung Drittbindungswirkung für den G-BA. Die Bindung reicht soweit, wie die Orphan Drug-Ausweisung und die frühe Nutzenbewertung des G-BA über die gleichen Tatsachen nach den gleichen Kriterien entscheiden, also im Umfang der Entscheidungsgleichheit. Tatsächlich ist die frühe Nutzenbewertung durch den G-BA im Hinblick auf den zu ermittelnden Zusatznutzen regelmäßig entscheidungsgleich mit der vorherigen Orphan Drug-Ausweisung, wie sich aus den Anforderungen an die klinische Überlegenheit eines Arzneimittels für seltene Leiden gemäß Verordnung (EG) Nr. 847/2000 der Kommission vom 27. April 2000 (Artikel 3 Absatz 3 d) ergibt: Danach erfolgt die Ausweisung als Orphan Drug, wenn eine größere Wirksamkeit als ein bisheriges Arzneimittel für seltene Leiden (Wirkung ist an einem klinisch bedeutungsvollen Endpunkt in geeigneten und ordnungsgemäß kontrollierten klinischen Prüfungen zu bewerten) gefunden wurde und darin ein bedeutender Beitrag zur Diagnose und Behandlung von Patienten liegt.

Diese weitgehende Entscheidungsgleichheit betreffend des Zusatznutzens folgt daraus, dass der G-BA im Rahmen der Nutzenbewertung auf der gleichen Tatsachengrundlage entscheidet wie die Europäische Kommission im Rahmen der Orphan Drug-Ausweisung. Ebenso entscheidet der G-BA im Rahmen der Nutzenbewertung über den Zusatznutzen im Wesentlichen nach den gleichen Kriterien wie die Europäische Kommission (BT-Drs. 17/3698, Seite 73) (Deutscher Bundestag, 2010).

Allerdings kann die Tatsachengrundlage in einzelnen Fällen auch abweichen und bei den Kriterien können sich im Detail ebenso einzelne Unterschiede ergeben. Da die Orphan Drug-Ausweisung vor Zulassung des Arzneimittels stattfindet und der Nutzenbewertung darum zeitlich vorgelagert ist, ist dies insbesondere der Fall, wenn sich die Tatsachen nach der Orphan Drug-Entscheidung aber noch vor der Nutzenbewertung geändert haben, etwa inzwischen zusätzliche Studien zur Verfügung stehen. Da die Orphan Drug-Entscheidung aber mit der Erteilung der arzneimittelrechtlichen Zulassung nochmals geprüft und bestätigt wird, fallen beide Entscheidungen zeitlich nur auseinander im Zeitraum von der Zulassungserteilung bis zur tatsächlichen Bewertung durch den G-BA.

Die an die Entscheidungsgleichheit gekoppelte Bindungswirkung gilt insbesondere, wenn das betreffende Arzneimittel als Orphan Drug ausgewiesen wurde, weil es einen erheblichen Nutzen gegenüber existierenden zufriedenstellenden Methoden hat. Dies ist bei Pomalidomid der Fall. Im Rahmen dieser Entscheidungsgleichheit ist der G-BA an die Orphan Drug-Ausweisung und ihre tragenden Gründe gebunden. Er hat daher davon auszugehen, dass das Orphan Drug einen erheblichen Zusatznutzen im Sinne von § 5 Nr. 7 des 5. Kap. G-BA VerFO hat.

Im Rahmen der weitgehenden Entscheidungsgleichheit betreffend des Zusatznutzens sollte der G-BA die Orphan Drug-Ausweisung bei seiner Nutzenbewertung daher beachten. Dies folgt aus dem europarechtlichen Grundsatz, dass eine Entscheidung für andere Entscheidungsträger hinsichtlich des Entscheidungssatzes und den tragenden Gründen Drittbindungswirkung entfaltet. Wir verweisen an dieser Stelle auch auf die veröffentlichte Literatur (Willhöft und Lietz, 2012, Schickert und Schmitz, 2012a, Schickert und Schmitz, 2012b). Aus der Sicht von Celgene reicht die Bindungswirkung der Orphan Drug-Ausweisung von Pomalidomid soweit, dass sie Aussagen zu einem erheblichen Zusatznutzen durch die Entscheidung der Europäischen Kommission rechtsverbindlich umfasst.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-57: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit rezdiviertem oder refraktärem Multiplem Myelom, die mindestens zwei Vortherapien, darunter Bortezomib und Lenalidomid, erhalten haben und unter der letzten Therapie eine Progression gezeigt haben.	erheblich

4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens

Für Arzneimittel, die zur Behandlung eines seltenen Leidens zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt.

Für diese Arzneimittel sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, im vorliegenden Abschnitt einzureichen. Dabei soll zunächst das Ausmaß des Zusatznutzens (in der Definition gemäß AM-NutzenV) begründet beschrieben werden. Unter Berücksichtigung des Ausmaßes des Zusatznutzens sollen Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, benannt werden. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Beschreiben Sie begründet das Ausmaß des Zusatznutzens unter Verwendung der Kategorisierung gemäß AM-NutzenV. Berücksichtigen Sie bei den Aussagen zum Zusatznutzen nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen. Verweisen Sie, wo notwendig, auf weitere Abschnitte von Modul 4.

Der Umsatz von Pomalidomid mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer hat in den vergangenen zwölf Kalendermonaten 50 Millionen Euro nicht überschritten. Nichtsdestotrotz hat sich Celgene als das verantwortliche pharmazeutische Unternehmen entschieden, ein über das geforderte Mindestmaß hinausgehendes Dossier einzureichen. Die Entscheidung wurde getroffen, da ein solches Vorgehen eine bessere Darlegung des aus Unternehmenssicht bestehenden erheblichen Zusatznutzens ermöglicht. Unsere rechtliche Einschätzung der Orphan Drug Regelungen im SGB V und der Bindungswirkung der EU Verordnung (EG) Nr. 141-2000 bleiben hiervon unberührt. Hierzu verweisen wir insbesondere auf die Darstellung unserer Rechtsauffassung in Abschnitt 4.4.2 sowie die dortige Darstellung der Studienergebnisse. Dort wird ausführlich begründet, warum für den Wirkstoff Pomalidomid bei indikationsgerechtem Einsatz ein erheblicher Zusatznutzen gesehen wird.

Geben Sie auf Basis des Ausmaßes des Zusatznutzens in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-58: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit rezidiviertem oder refraktärem Multiplem Myelom, die mindestens zwei Vortherapien, darunter Bortezomib und Lenalidomid, erhalten haben und unter der letzten Therapie eine Progression gezeigt haben.	Erheblich Aufgrund der an anderer Stelle in dem hier vorliegenden Dossier, insbesondere Abschnitt 4.4.2, gemachten Ausführungen wird für den Wirkstoff Pomalidomid bei indikationsgerechtem Einsatz ein erheblicher Zusatznutzen beansprucht.

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

Nicht anwendbar.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

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

Nicht anwendbar.

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

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

Nicht anwendbar.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁹, Molenberghs 2010²⁰). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006²¹) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006²²) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

¹⁹ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

²⁰ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

²¹ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

²² Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

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

In die Nutzenbewertung wurden die Endpunkte *Zeit bis zur Progression* und *Ansprechen* eingeschlossen, die in der Studie MM-003 anhand der Bladé Kriterien beurteilt wurden (Bladé et al., 1998). Unter den Parametern, die in diesen Kriterien zusammengefasst sind, nimmt das M-Protein eine zentrale Rolle ein. M-Proteine sind als Surrogatparameter für eine Nierenschädigung, die sich nachfolgend in einer Dialysepflicht manifestieren kann, allgemein anerkannt. Die große Akzeptanz zeigt sich darin, dass die Verlaufskontrolle einer Behandlung des Multiplen Myeloms in klinischen Leitlinien auf M-Proteinen beruht (National Comprehensive Cancer Network®, 2013, Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO) e.V., 2010).

Eine Validierung des Surrogatparameters *M-Protein* gemäß der hohen methodischen Standards in der Literatur ist nicht möglich, da sich keine klinischen Studien finden, die als Grundlage für die Untersuchung einer Kausalität zwischen dem M-Protein als Surrogatparameter und den genannten harten klinischen Endpunkten verwendet werden können. Dies ist maßgeblich darauf zurückzuführen, dass das Multiple Myelom mit einer Häufigkeit von weniger als fünf Erkrankten pro 10.000 Personen auftritt und dies somit eine faktische Limitierung in der Anzahl möglicher Studien und deren Umfang darstellt. Es ist deshalb auch unrealistisch anzunehmen, dass eine solche Validierung zukünftig erbracht werden kann.

M-Proteine setzen sich aus schweren und leichten Ketten zusammen. Letztere werden über die Niere ausgeschieden und können dabei sukzessive zu einer Schädigung der Tubuli führen (Stringer et al., 2011, Dimopoulos et al., 2008). Diese als *Cast-Nephropathie* bezeichneten pathologischen Prozesse können zu Nierenversagen und daraus resultierend zu einer Dialysepflicht führen. Bis zu 50 % der Patienten mit Multiplem Myelom weisen bei Erstdiagnose eine Nierenschädigung unterschiedlichen Schweregrades auf. Bei 10 % der Patienten handelt es sich um eine schwere Form, die unbehandelt innerhalb von Wochen zum Tode führen kann (Stringer et al., 2011). Nierenversagen stellt eine häufige Todesursache bei Patienten mit Multiplem Myelom dar (Herrera et al., 2004).

Bei Patienten, die mit Chemotherapie und *High cut-off*-Dialyse behandelt wurden, wurde ein linearer Zusammenhang zwischen der Reduktion freier Leichtketten im Serum und der Wahrscheinlichkeit einer Dialysefreiheit beobachtet (Hutchison et al., 2012). Die Studie von Oehrlein et al. zeigt darüber hinaus, dass ein Therapieansprechen (also Reduktion des M-Proteins) der Erkrankung und die damit einhergehende Krankheitskontrolle bei 50-60 % der rMM Patienten zu einer Verbesserung der Nierenfunktion führt (Oehrlein et al., 2012).

Trotz des Fehlens der methodisch gesicherten Validität des M-Proteins als Surrogatparameter erscheint es dennoch klinisch plausibel, dass M-Proteine über einen prognostischen Wert für das Auftreten von Nierenschädigungen und einer möglichen Dialysepflicht im weiteren Verlauf der Erkrankung verfügen. Weitere Belege für die Bedeutung des M-Proteins als Surrogat-

parameter für die Symptomlast von Patienten mit rrMM liefert die Studie MM-003, in der gezeigt wird, dass ein mindestens partielles Ansprechen (also die damit einhergehende Reduktion des M-Proteins) ein Indikator für eine Verbesserung der Lebensqualität und Funktionalität (EQ-5D Index Score, QLQ-C30 Allgemeiner Gesundheitszustand/Lebensqualität, Rollenfunktion, Physische und Soziale Funktion) sowie für eine Senkung der Symptomlast (QLQ-C30 Schmerz, Dyspnoe, Fatigue, Insomnie und QLQ-MY20 Krankheitssymptome) ist (Abschnitt 4.3.1.3.1.2).

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Datenquellen auf.

Studie	Datenquellen
MM-003	(Jacques und Celgene Corporation, 2013)

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie in Modul 4 angeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

1. AARONSON, N. K., AHMEDZAI, S., BERGMANN, B., BULLINGER, M., CULL, A., DUEZ, N. J., FILIBERTI, A., FLECHTNER, H., FLEISHMAN, S. B., DE HAES, J. C. J. M., KAASA, S., KLEE, M., OSOBA, D., RAZAVI, D., ROFE, P. B., SCHRAUB, S., SNEEUW, K., SULLIVAN, M. & TAKEDA, F. 1993. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *Journal of the National Cancer Institute*, 85, 365-76.
2. ANDERSON, K. C., KYLE, R. A., RAJKUMAR, S. V., STEWART, A. K., WEBER, D., RICHARDSON, P. & MYELOMA, A. F. P. O. C. E. I. M. 2008. Clinically relevant end points and new drug approvals for myeloma. *Leukemia*, 22, 231-9.
3. AVET-LOISEAU, H., LELEU, X., ROUSSEL, M., MOREAU, P., GUERIN-CHARBONNEL, C., CAILLOT, D., MARIT, G., BENBOUBKER, L., VOILLAT, L., MATHIOT, C., KOLB, B., MACRO, M., CAMPION, L., WETTERWALD, M., STOPPA, A. M., HULIN, C., FACON, T., ATTAL, M., MINVIELLE, S. & HAROUSSEAU, J. L. 2010. Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). *J Clin Oncol*, 28, 4630-4.
4. BACKSTROM, J. & CELGENE CORPORATION 2012. A Phase 1/2 Multi-Center, Randomized, Open Label, Dose-Escalation Study to Determine the Maximum Tolerated Dose, Safety, and Efficacy of CC-4047 Alone or in Combination with Low-Dose Dexamethasone in Patients with Relapsed and Refractory Multiple Myeloma

Who Have Received Prior Treatment that Includes Lenalidomide and Bortezomib. Celgene Corporation,.

5. BERENSON, J. R., LICHTENSTEIN, A., PORTER, L., DIMOPOULOS, M. A., BORDONI, R., GEORGE, S., LIPTON, A., KELLER, A., BALLESTER, O., KOVACS, M., BLACKLOCK, H., BELL, R., SIMEONE, J. F., REITSMA, D. J., HEFFERNAN, M., SEAMAN, J. & KNIGHT, R. D. 1998. Long-term pamidronate treatment of advanced multiple myeloma patients reduces skeletal events. Myeloma Aredia Study Group. *J Clin Oncol*, 16, 593-602.
6. BLADÉ, J., SAMSON, D., REECE, D., APPERLEY, J., BJÖRKSTRAND, B., GAHRTON, G., GERTZ, M., GIRALT, S., JAGANNATH, S. & VESOLE, D. 1998. Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT. European Group for Blood and Marrow Transplant. *Br J Haematol*, 102, 1115-23.
7. BÖHRER, S. & PÖNISCH, W. 2001. Multiples Myelom. In: MITROU, P. S. (ed.) *Non-Hodgkin-Lymphome*. Frankfurt/M.: Tumorzentrum Rhein-Main.
8. BREITKREUTZ, I., RAAB, M. S., VALLET, S., HIDESHIMA, T., RAJE, N., MITSIADES, C., CHAUHAN, D., OKAWA, Y., MUNSHI, N. C., RICHARDSON, P. G. & ANDERSON, K. C. 2008. Lenalidomide inhibits osteoclastogenesis, survival factors and bone-remodeling markers in multiple myeloma. *Leukemia*, 22, 1925-32.
9. BUNDESMINISTERIUM DER JUSTIZ 2010. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AMNutzenV).
10. BUNDESMINISTERIUM DER JUSTIZ 2013. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes v. 20. Dezember 1988, BGBl. I S. 2477). Berlin.
11. CELGENE EUROPE LIMITED 2013a. Adhärenz Lebensqualität (Pom+LD-Dex vs. HD-Dex, MM-003).
12. CELGENE EUROPE LIMITED 2013b. Gesamtüberleben: Adjustierung für Cross-Over (MM-003).
13. CELGENE EUROPE LIMITED 2013c. Interaktion_Sicherheit (MM-003).
14. CELGENE EUROPE LIMITED 2013d. Interaktion_Unerwünschte Ereignisse, Grad 3-4 (MM-003).
15. CELGENE EUROPE LIMITED 2013e. Interaktion_Wirksamkeit (MM-003).
16. CELGENE EUROPE LIMITED 2013f. Inzidenzraten Unerwünschte Ereignisse, Grad 3-4 (Pom+LD-Dex vs. HD-Dex, MM-003).

17. CELGENE EUROPE LIMITED 2013g. Lebensqualität je Behandlungszyklus (Pom+LD-Dex vs. HD-Dex, MM-003).
18. CELGENE EUROPE LIMITED 2013h. Minimally Clinically Important Differences (MM-003).
19. CELGENE EUROPE LIMITED 2013i. Mixed Model zur Veränderung der Lebensqualität (Pom+LD-Dex vs. HD-Dex, MM-003).
20. CELGENE EUROPE LIMITED 2013j. Progression als Determinante Unerwünschter Ereignisse (MM-003).
21. CELGENE EUROPE LIMITED 2013k. Progression und Ansprechen als Determinante der Lebensqualität (MM-003).
22. CELGENE EUROPE LIMITED 2013l. Verbesserung von Knochenschmerzen und ECOG PS.
23. CELGENE EUROPE LIMITED 2013m. Zeit bis zur Verschlechterung der Lebensqualität (Pom+LD-Dex vs. HD-Dex, MM-003).
24. COCKS, K., COHEN, D., WISLOFF, F., SEZER, O., LEE, S., HIPPE, E., GIMSING, P., TURESSON, I., HAJEK, R., SMITH, A., GRAHAM, L., PHILLIPS, A., STEAD, M., VELIKOVA, G. & BROWN, J. 2007. 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*, 43, 1670-8.
25. COSTA, L., BADIA, X., CHOW, E., LIPTON, A. & WARDLEY, A. 2008. Impact of skeletal complications on patients' quality of life, mobility, and functional independence. *Support Care Cancer*, 16, 879-89.
26. DELFORGE, M., BLADE, J., DIMOPOULOS, M. A., FACON, T., KROPFF, M., LUDWIG, H., PALUMBO, A., VAN DAMME, P., SAN-MIGUEL, J. F. & SONNEVELD, P. 2010. Treatment-related peripheral neuropathy in multiple myeloma: the challenge continues. *Lancet Oncol*, 11, 1086-95.
27. DEUTSCHE GESELLSCHAFT FÜR HÄMATOLOGIE UND MEDIZINISCHE ONKOLOGIE (DGHO) E.V. 2010. *Leitlinien Multiples Myelom* [Online]. onkopedia leitlinien. Available: <http://www.dgho-onkopedia.de/de/onkopedia/leitlinien/multiples-myelom/index.html> [Accessed 27.08.2013].
28. DEUTSCHE KREBSGESELLSCHAFT (DKG) E.V. 2011. *Lebensqualität des chronisch Kranken* [Online]. Available: http://www.krebsgesellschaft.de/lk_lebensqualitaet,124039.html [Accessed 26.04.2013 2013].
29. DEUTSCHER BUNDESTAG 2010. Beschlussempfehlung und Bericht des Ausschusses für Gesundheit (14. Ausschuss).

30. DIMOPOULOS, M. A., DELFORGE, M., HAJEK, R., KROPFF, M., PETRUCCI, M. T., LEWIS, P., NIXON, A., ZHANG, J., MEI, J. & PALUMBO, A. 2013. Lenalidomide, melphalan, and prednisone, followed by lenalidomide maintenance, improves health-related quality of life in newly diagnosed multiple myeloma patients aged 65 years or older: results of a randomized phase III trial. *Haematologica*, 98, 784-8.
31. DIMOPOULOS, M. A., KASTRITIS, E., ROSINOL, L., BLADE, J. & LUDWIG, H. 2008. Pathogenesis and treatment of renal failure in multiple myeloma. *Leukemia*, 22, 1485-93.
32. DIMOPOULOS, M. A., PALUMBO, A., HAJEK, R., KROPFF, M., PETRUCCI, M. T., LEWIS, P., MILLAR, S., ZHANG, J., MEI, J. & DELFORGE, M. 2011. Melphalan, Prednisone and Lenalidomide Followed by Lenalidomide Maintenance Displays Treatment Characteristics Favourable to Global Quality of Life in Newly Diagnosed Multiple Myeloma (NDMM) Patients \geq 65 Years. *ASH Annual Meeting Abstracts*.
33. DURIE, B. G., HAROUSSEAU, J. L., MIGUEL, J. S., BLADE, J., BARLOGIE, B., ANDERSON, K., GERTZ, M., DIMOPOULOS, M., WESTIN, J., SONNEVELD, P., LUDWIG, H., GAHRTON, G., BEKSAC, M., CROWLEY, J., BELCH, A., BOCCADARO, M., CAVO, M., TURESSON, I., JOSHUA, D., VESOLE, D., KYLE, R., ALEXANIAN, R., TRICOT, G., ATTAL, M., MERLINI, G., POWLES, R., RICHARDSON, P., SHIMIZU, K., TOSI, P., MORGAN, G. & RAJKUMAR, S. V. 2006. International uniform response criteria for multiple myeloma. *Leukemia*, 20, 1467-73.
34. EUROPEAN AGENCY FOR THE EVALUATION OF MEDICINAL PRODUCTS (EMA) - COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS (CPMP) 2001. Points to consider on application with 1. meta-analyses; 2. one pivotal study.
35. EUROPEAN MEDICINES AGENCY (EMA) 2013a. CHMP Assessment Report. Pomalidomide Celgene. Committee for Medicinal Products for Human Use (CHMP).
36. EUROPEAN MEDICINES AGENCY (EMA) 2013b. EMA/COMP Position on review of criteria for orphan designation of an orphan medicinal product submitted for marketing authorisation application. Pomalidomide Celgene. Orphan indication: Treatment of multiple myeloma. Committee for Orphan Medicinal Products (COMP).
37. EUROPEAN MEDICINES AGENCY (EMA) 2008. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man (CHMP/EWP/205/95 REV. 3). Methodological considerations for using progression-free survival (PFS) as primary endpoint in confirmatory trials for registration. Committee for Medicinal Products for Human Use (CHMP).
38. GAY, F. & PALUMBO, A. 2010. Management of disease- and treatment-related complications in patients with multiple myeloma. *Med Oncol*, 27 Suppl 1, S43-52.

39. GEMEINSAMER BUNDESAUSSCHUSS 2011. 5. Kapitel der Verfahrensordnung - Bewertung des Nutzens von Arzneimitteln nach § 35a SGB V. Berlin.
40. GREINER, W., WEIJNEN, T., NIEUWENHUIZEN, M., OPPE, S., BADIA, X., BUSSCHBACH, J., BUXTON, M., DOLAN, P., KIND, P., KRABBE, P., OHINMAA, A., PARKIN, D., ROSET, M., SINTONEN, H., TSUCHIYA, A. & DE CHARRO, F. 2003. A single European currency for EQ-5D health states. Results from a six-country study. *Eur J Health Econ*, 4, 222-31.
41. HAROUSSEAU, J. L., DIMOPOULOS, M. A., WANG, M., CORSO, A., CHEN, C., ATTAL, M., SPENCER, A., YU, Z., OLESNYCKYJ, M., ZELDIS, J. B., KNIGHT, R. D. & WEBER, D. M. 2010. Better quality of response to lenalidomide plus dexamethasone is associated with improved clinical outcomes in patients with relapsed or refractory multiple myeloma. *Haematologica*, 95, 1738-44.
42. HERRERA, G. A., JOSEPH, L., GU, X., HOUGH, A. & BARLOGIE, B. 2004. Renal Pathologic Spectrum in an Autopsy Series of Patients With Plasma Cell Dyscrasia. *Arch Pathol Lab Med*, 128, 875-79.
43. HUTCHISON, C. A., BATUMAN, V., BEHRENS, J., BRIDOUX, F., SIRAC, C., DISPENZIERI, A., HERRERA, G. A., LACHMANN, H., SANDERS, P. W., INTERNATIONAL, K. & MONOCLONAL GAMMOPATHY RESEARCH, G. 2012. The pathogenesis and diagnosis of acute kidney injury in multiple myeloma. *Nat Rev Nephrol*, 8, 43-51.
44. INSTITUT FÜR QUALITÄT UND WIRTSCHAFTLICHKEIT IM GESUNDHEITSWESEN (IQWiG) 2011a. Abschlussbericht N05-03C Version 1.0. Stammzelltransplantation bei Multiplem Myelom. Kurzfassung Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
45. INSTITUT FÜR QUALITÄT UND WIRTSCHAFTLICHKEIT IM GESUNDHEITSWESEN (IQWiG) 2011b. Allgemeine Methoden. Version 4.0 vom 23.09.2011. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
46. JACQUES, C. & CELGENE CORPORATION 2013. A Phase 3, Multicenter, Randomized, Open-label Study to Compare the Efficacy and Safety of Pomalidomide in Combination with Low-Dose Dexamethasone versus High-Dose Dexamethasone in Subjects with Relapsed and Refractory Multiple Myeloma. Celgene Corporation.
47. KAPOOR, P. & RAJKUMAR, S. V. 2011. Update on risk stratification and treatment of newly diagnosed multiple myeloma. *Int J Hematol*, 94, 310-20.
48. KLEIN, U., JAUCH, A., HIELSCHER, T., HILLEGASS, J., RAAB, M. S., SECKINGER, A., HOSE, D., HO, A. D., GOLDSCHMIDT, H. & NEBEN, K. 2011. Chromosomal aberrations +1q21 and del(17p13) predict survival in patients with recurrent multiple myeloma treated with lenalidomide and dexamethasone. *Cancer*, 117, 2136-44.

49. KOMMISSION DER EUROPÄISCHEN GEMEINSCHAFTEN 2009. Entscheidung der Kommission vom 8.10.2009 über die Ausweisung des Arzneimittels "Pomalidomid" als Arzneimittel für seltene Leiden gemäß Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates. Brüssel.
50. KVAM, A. K., FAYERS, P. & WISLOFF, F. 2010. What changes in health-related quality of life matter to multiple myeloma patients? A prospective study. *Eur J Haematol*, 84, 345-53.
51. KVAM, A. K., FAYERS, P. M. & WISLOFF, F. 2011. 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*, 87, 330-7.
52. LELEU, X. 2012. A Multicenter Randomized Open label Phase II study of Pomalidomide and Dexamethasone in relapse and refractory Multiple Myeloma patients who are progressive and did not achieve at least a partial response to Bortezomib and Lenalidomide – IFM 2009-02. Lille, France: Hopital Huriez - CHRU de Lille.
53. LELEU, X., ATTAL, M., ARNULF, B., MOREAU, P., TRAUILLÉ, C., MARIT, G., MATHIOT, C., PETILLON, M., MACRO, M., ROUSSEL, M., PEGOURIE, B., KOLB, B., STOPPA, A. M., HENNACHE, B., BRÉCHIGNAC, S., MEULEMAN, N., THIELEMANS, B., GARDERET, L., ROYER, B., HULIN, C., BENBOUBKER, L., DECAUX, O., ESCOFFRE-BARBE, M., MICHALLET, M., CAILLOT, D., FERMAND, J. P., AVET-LOISEAU, H. & FACON, T. 2013. Pomalidomide plus low-dose dexamethasone is active and well tolerated in bortezomib and lenalidomide-refractory multiple myeloma: Intergroupe Francophone du Myélome 2009-02. *Blood*, 121, 1968-75.
54. LONIAL, S., BAZ, R., BAHLIS, N. J., CHEN, C. I., ANDERSON, K. C., CHEN, M., ZAKI, M. & RICHARDSON, P. G. 2012. Improvement in Clinical Benefit Parameters with Pomalidomide (POM) in Combination with Low-Dose Dexamethasone (LoDEX) in Patients with Relapsed and Refractory Multiple Myeloma (RRMM): Results From a Phase 2 Study. *ASH Annual Meeting Abstracts*, 120, 4052.
55. MIGKOU, M., KASTRITIS, E., ROUSSOU, M., GKOTZAMANIDOU, M., GAVRIATOPOULOU, M., NIKITAS, N., MPARMPAROSSI, D., MATSOUKA, C., GIKA, D., TERPOS, E. & DIMOPOULOS, M. A. 2011. Short progression-free survival predicts for poor overall survival in older patients with multiple myeloma treated upfront with novel agent-based therapy. *Eur J Haematol*, 87, 323-9.
56. MOLS, F., OERLEMANS, S., VOS, A. H., KOSTER, A., VERELST, S., SONNEVELD, P. & VAN DE POLL-FRANSE, L. V. 2012. 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*, 89, 311-9.

57. NATIONAL COMPREHENSIVE CANCER NETWORK® 2013. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®). Multiple Myeloma. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®)*, 1.2013.
58. NIESVIZKY, R., RICHARDSON, P. G., RAJKUMAR, S. V., COLEMAN, M., ROSINOL, L., SONNEVELD, P., SCHUSTER, M. W., IRWIN, D., STADTMAUER, E. A., FACON, T., HAROUSSEAU, J. L., BORAL, A. L., ESSELTINE, D. L., ANDERSON, K. C. & BLADE, J. 2008. The relationship between quality of response and clinical benefit for patients treated on the bortezomib arm of the international, randomized, phase 3 APEX trial in relapsed multiple myeloma. *Br J Haematol*, 143, 46-53.
59. OEHRLEIN, K., LANGER, C., STURM, I., PONISCH, W., HAHN-AST, C., KUHN, S. & WEISEL, K. C. 2012. Successful treatment of patients with multiple myeloma and impaired renal function with lenalidomide: results of 4 German centers. *Clin Lymphoma Myeloma Leuk*, 12, 191-6.
60. OKEN, M. M., CREECH, R. H., TORMEY, D. C., HORTON, J., DAVIS, T. E., MCFADDEN, E. T. & CARBONE, P. P. 1982. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*, 5, 649-55.
61. RALSTON, S. H., GALLACHER, S. J., PATEL, U., CAMPBELL, J. & BOYLE, I. T. 1990. Cancer-Associated Hypercalcemia: Morbidity and Mortality - Clinical Experience in 126 Treated Patients. *Annals of Internal Medicine*, 112, 499-504.
62. REECE, D., SONG, K. W., FU, T., ROLAND, B., CHANG, H., HORSMAN, D. E., MANSOOR, A., CHEN, C., MASIH-KHAN, E., TRIEU, Y., BRUYERE, H., STEWART, D. A. & BAHLIS, N. J. 2009. Influence of cytogenetics in patients with relapsed or refractory multiple myeloma treated with lenalidomide plus dexamethasone: adverse effect of deletion 17p13. *Blood*, 114, 522-5.
63. RICHARDSON, P. G., DELFORGE, M., BEKSAC, M., WEN, P., JONGEN, J. L., SEZER, O., TERPOS, E., MUNSHI, N., PALUMBO, A., RAJKUMAR, S. V., HAROUSSEAU, J. L., MOREAU, P., AVET-LOISEAU, H., LEE, J. H., CAVO, M., MERLINI, G., VOORHEES, P., CHNG, W. J., MAZUMDER, A., USMANI, S., EINSELE, H., COMENZO, R., ORLOWSKI, R., VESOLE, D., LAHUERTA, J. J., NIESVIZKY, R., SIEGEL, D., MATEOS, M. V., DIMOPOULOS, M., LONIAL, S., JAGANNATH, S., BLADE, J., MIGUEL, J. S., MORGAN, G., ANDERSON, K. C., DURIE, B. G. & SONNEVELD, P. 2011. Management of treatment-emergent peripheral neuropathy in multiple myeloma. *Leukemia*, 26, 595-608.
64. RICHARDSON, P. G., SIEGEL, D., BAZ, R., KELLEY, S. L., MUNSHI, N. C., LAUBACH, J., SULLIVAN, D., ALSINA, M., SCHLOSSMAN, R. L., GHOBRIAL, I. M., DOSS, D., LOUGHNEY, N., MCBRIDE, L., BILOTTI, E., ANAND, P., NARDELLI, L., WEAR, S., LARKINS, G., CHEN, M., ZAKI, M. H., JACQUES, C. & ANDERSON, K. C. 2013. Phase I study of pomalidomide MTD, safety, and efficacy in patients with refractory multiple myeloma who have received lenalidomide and bortezomib. *Blood*, 121, 1961-67.

65. ROBINS, J. M. & TSIATIS, A. A. 1991. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Commun. Statist.-Theory Meth.*, 20, 2609-2631.
66. SAN MIGUEL, S., WEISEL, K., MOREAU, P., LACY, M., SONG, K., DELFORGE, M., KARLIN, L., GOLDSCHMIDT, H., BANOS, A., ORIOL, A., CHEN, C., CAVO, M., GARDERET, L., IVANOVA, V., MARTINEZ-LOPEZ, J., BELCH, A., PALUMBO, A., SCHEY, S., SONNEVELD, P., YU, X., STERNAS, L., JACQUES, C., ZAKI, M. & DIMOPOULOS, M. A. 2013. Survival advantage of pomalidomide plus low-dose dexamethasone in patients with refractory or relapsed/refractory multiple myeloma: MM-003 results of a randomised phase 3 trial. *Lancet Oncology (accepted for publication)*.
67. SCHICKERT, J. & SCHMITZ, A. 2012a. Orphan Drugs in der Nutzenbewertung nach § 35a SGB V. Teil 1: Die vereinfachte Nutzenbewertung von Orphan Drugs unterhalb der Umsatzschwelle. *pharmind*, 74, 416-420.
68. SCHICKERT, J. & SCHMITZ, A. 2012b. Orphan Drugs in der Nutzenbewertung nach § 35a SGB V. Teil 2: Die vollumfängliche Nutzenbewertung von Orphan Drugs oberhalb der Umsatzschwelle. *pharmind*, 74, 779-784.
69. SMITH, A., WISLOFF, F., SAMSON, D., FORUM, U. K. M., NORDIC MYELOMA STUDY, G. & BRITISH COMMITTEE FOR STANDARDS IN, H. 2005. Guidelines on the diagnosis and management of multiple myeloma 2005. *Br J Haematol*, 132, 410-51.
70. SONNEVELD, P., VERELST, S. G., LEWIS, P., GRAY-SCHOPFER, V., HUTCHINGS, A., NIXON, A. & PETRUCCI, M. T. 2013. Review of health-related quality of life data in multiple myeloma patients treated with novel agents. *Leukemia*, 1-11.
71. STRAKA, C., ODUNCU, F. S., FRANKE, D. & DIETZFELBINGER, H. 2012. Therapeutisches Konzept. In: STRAKA, C. & DIETZFELBINGER, H. (eds.) *MANUAL Multiples Myelom*. München: W. Zuckschwerdt Verlag.
72. STRINGER, S., BASNAYAKE, K., HUTCHISON, C. & COCKWELL, P. 2011. Recent advances in the pathogenesis and management of cast nephropathy (myeloma kidney). *Bone Marrow Res*, 2011, 493697.
73. U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES - FOOD AND DRUG ADMINISTRATION, CENTER FOR DRUG EVALUATION AND RESEARCH (CDER) & CENTER FOR BIOLOGICS EVALUATION AND RESEARCH (CBER) 2007. Guidance for Industry. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics
74. WILLHÖFT, C. & LIETZ, C. 2012. Die frühe Nutzenbewertung von Orphan Drugs nach § 35a SGB V *Arzneimittel und Recht*, 01/12, 19-23.

75. WISLOFF, F., EIKA, S., HIPPE, E., HJORTH, M., HOLMBERG, E., KAASA, S., PALVA, I. & WESTIN, J. 1996. Measurement of health-related quality of life in multiple myeloma. *British Journal of Haematology*, 92, 604-613.
76. WYRWICH, K. W., NIENABER, N. A., TIERNEY, W. M. & WOLINSKY, F. D. 1999. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Medical Care*, 37, 469-478.
77. YEH, H. S. & BERENSON, J. R. 2006. Treatment for myeloma bone disease. *Clin Cancer Res*, 12, 6279s-6284s.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ²³] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

²³ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Für die Zusatznutzenbewertung von Pomalidomid ist keine bibliographische Literaturrecherche erforderlich, da

- 1.) die erste Zulassung für den Wirkstoff Pomalidomid weltweit weniger als 12 Monate vor Dossiereinreichung erfolgte (USA: FDA-Zulassung am 08.02.2013).
- 2.) keine indirekten Vergleiche für die Nutzenbewertung des Wirkstoffs Pomalidomid herangezogen wurden.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. [clinicaltrials.gov](http://www.clinicaltrials.gov)), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Studienregister	Clinicaltrials
Internetadresse	www.clinicaltrials.gov
Datum der Suche	22.07.2013
Bezug	Studien zum zu bewertenden Arzneimittel
Suchstrategie	(Pomalidomide OR Actimid)
Treffer	51

Studienregister	International Clinical Trials Registry Search Portal (ICTRP)
Internetadresse	apps.who.int/trialsearch
Datum der Suche	22.07.2013
Bezug	Studien zum zu bewertenden Arzneimittel
Suchstrategie	(Pomalidomide OR Actimid)
Treffer	47

Studienregister	clinicaltrialsregister.eu
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	22.07.2013
Bezug	Studien zum zu bewertenden Arzneimittel
Suchstrategie	(Pomalidomide OR Actimid)
Treffer	9

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund

Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrundes.

Für die Zusatznutzenbewertung von Pomalidomid ist keine bibliographische Literaturrecherche erforderlich, da

- 1.) die erste Zulassung für den Wirkstoff Pomalidomid weltweit weniger als 12 Monate vor Dossiereinreichung erfolgte (USA: FDA-Zulassung am 08.02.2013).
- 2.) keine indirekten Vergleiche für die Nutzenbewertung des Wirkstoffs Pomalidomid herangezogen wurden.

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-59(Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponser-te Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

Nicht anwendbar, da keine abgebrochene Studie vorliegt.

Anhang 4-E: Liste der laufenden Studien

Benennen Sie in der nachfolgenden Tabelle die laufenden Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-60 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, Vergleichstherapie(n)						
MM-003	ja	ja	nein	ja [Jacques, C. & Celgene Corporation 2013]	ja Clinicaltrials.gov:d NCT01311687 Clinicaltrialsregister.eu:d EudraCT: 2010-019820-30 ICTRP:d NCT01311687 EUCTR2010-019820-30-BE	ja, zur Veröffentlichung in <i>Lancet Oncology</i> akzeptiert. (San Miguel et al., 2013)
2009-013319-36, 2009_28/0924	nein	ja	ja	ja [Leleu, X. 2012]	ja Clinicaltrials.gov: NCT01053949 (Leleu et al., 2013) Clinicaltrialsregister.eu:d EudraCT 2009-013319-36	ja (Leleu et al., 2013)
MCC-16705, PO-MM-PI-0039,	nein	ja	nein	nein	ja <u>Clinicaltrials.gov</u> : ^d NCT01432600 <u>ICTRP</u> : ^d NCT01432600	nein
HIC 1011007607	nein	ja	nein	nein	ja <u>Clinicaltrials.gov</u> : ^d NCT01319422 <u>ICTRP</u> : ^d NCT01319422	nein

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
MM-002	nein	ja	nein	ja [Backstrom, J. & Celgene Corporation 2012]	ja Clinicaltrials.gov : NCT00833833 (Richardson et al., 2013)	ja (Richardson et al., 2013)
12-128	nein	ja	nein	nein	ja Clinicaltrials.gov : ^d NCT01745588 ICTRP : ^d NCT01745588	nein
CC-4047-MM-007	nein	ja	nein	nein	ja Clinicaltrials.gov : ^d NCT01734928 ICTRP : ^d NCT01734928	nein

a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.
b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
d: Keine weiteren Zitate in den Registereinträgen vorhanden

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-61 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-61 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-61(Anhang): Studiendesign und -methodik für Studie MM-003

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Zielsetzung und Fragestellung: <ul style="list-style-type: none"> Vergleich der Wirksamkeit und der Sicherheit von Pom+LD-Dex und HD-Dex bei Studienteilnehmern mit refraktärem oder rezidiviertem und refraktärem MM Hypothesen: <ul style="list-style-type: none"> Nullhypothese: Die Überlebensfunktion für PFS ist zu jedem Zeitpunkt in der Pom+LD-Dex-Gruppe gleich jener der HD-Dex-Gruppe Alternativhypothese Die Überlebensfunktion für PFS ist zu jedem Zeitpunkt in der Pom+LD-Dex-Gruppe höher als jene der HD-Dex-Gruppe
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Allokation: randomisiert (Zuteilungsverhältnis 2: 1, Randomisierungseinheit: Studienteilnehmer) Verblindung: offen Studienhorizont: n. a., da ereignisgesteuert Design: 2-armig, parallel, aktiv-kontrolliert Studienorganisation: multizentrisch Phase: III
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Amendment 2 (06/2011) <ul style="list-style-type: none"> Einstufung von SPMs als SUE und Berichten über die gesamte Studiendauer Bestätigung der durch Prüfarzte erhobenen PD durch das IRAC bei allen Studienteilnehmern, um Verzerrungen zu minimieren Änderung des Einschlusskriterium Messbarer M-Proteinspiegel im Serum von

Item ^a	Charakteristikum	Studieninformation
		<p>1,0 g/dl auf 0.5 g/dl, da die einzuschließende Studienpopulation stark vorbehandelt war</p> <ul style="list-style-type: none"> • Änderung des Ausschlusskriteriums <i>Serum Bilirubin-Spiegel</i>, um einen Einschluss von Patienten mit erblicher Hyperbilirubinämie zu ermöglichen, da bei ihnen der höhere Bilirubinspiegel nicht durch Leberinsuffizienz verursacht ist. <p>Amendment 3 (11/11)</p> <ul style="list-style-type: none"> • Änderung des Ausschlusskriteriums <i>Keine maligne Vorerkrankung</i> außer Krankheitsfreiheit seit ≥ 5 J statt ≥ 3 J • Einschluss von Patienten mit PD während einer Therapie mit Alkylanzien, insofern ≥ 2 Zyklen gegeben wurden • Änderung des Ausschlusskriteriums <i>Allogene Knochenmarks- oder periphere Blutstammzelltransplantation</i> ≤ 12 Mo vor Beginn der Studienmedikation: Patienten, die eine immunsuppressive Therapie erhalten und diese nicht mind. 4 Wo (statt 12 Mo) vor Studienbeginn beenden, qualifizieren nicht für die Studie
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • ≥ 18 J bei Gabe des <i>informed consent</i> • Verständnis und Gabe eines <i>informed consent</i> • Fähigkeit vorgesehene Untersuchungen und andere im Studienprotokoll spezifizierten Prozeduren einzuhalten • Dokumentierte MM-Diagnose und messbarer M-Proteinspiegel im Serum ($\geq 0,5$ g/dl) oder Urin (≥ 200 mg/24 h) • ≥ 2 vorangegangene Myelom-Therapielinien • Refraktäres oder rezidiviertes und refraktäres MM, definiert als dokumentierte Krankheitsprogression während oder innerhalb 60 d nach Beenden der letzten Myelom-Therapie <ul style="list-style-type: none"> - Primär refraktär: Kein Ansprechen $>PD$ auf jegliche vorangegangene Myelom-Therapie - Rezidivierend und refraktär: Rezidiv nach Ansprechen $\geq SD$ für ≥ 2 Behandlungszyklen und PD während oder innerhalb 60 d nach Beenden der letzten Myelom-Therapie • ≥ 2 konsekutive Therapiezyklen (inkl.

Item ^a	Charakteristikum	Studieninformation
		<p>Lenalidomid und Bortezomib)</p> <ul style="list-style-type: none"> • Vorangegangene Therapie mit Alkylanzien: <ul style="list-style-type: none"> - als Teil einer Stammzelltransplantation oder - ≥ 6 konsekutive Zyklen einer Alkylanzien-basierten Therapie oder - PD während einer Behandlung mit Alkylanzien (Gabe von ≥ 2 Zyklen) • Versagen der letzten Lenalidomid-basierten Behandlung (PD während oder innerhalb 60 d nach Therapieende oder bei Ansprechen \geqPR Rezidiv innerhalb 6 Mo nach Therapieende) • Versagen der letzten Bortezomib-haltigen Therapie (PD innerhalb 60 d nach Therapieende, bei Ansprechen \geqPR Rezidiv innerhalb 6 Mo nach Therapieende oder Teilnehmer ohne Ansprechen \geqMR und mit Unverträglichkeit/Toxizität nach ≥ 2 Zyklen einer Bortezomibbasierten Therapie) • ECOG PS ≤ 2 • Frauen im gebärfähigen Alter nur bei: <ul style="list-style-type: none"> - gleichzeitiger Durchführung zweier zuverlässiger Kontrazeptionsmaßnahmen oder vollständiger heterosexueller Abstinenz für ≥ 28 d vor Studienbeginn, während der Studie und ≥ 28 d nach Ende der Studienmedikation und regelmäßige Durchführung von Schwangerschaftstests in diesem Zeitraum - Verzicht auf Stillen während und 28 d nach Studienteilnahme • Männern nur bei Bereitschaft <ul style="list-style-type: none"> - während und 28 d nach Studienteilnahme bei sexuellen Kontakten mit Frauen im gebärfähigen Alter Latex-Kondome zu verwenden - auf Samen- und Spermaspenden während der Pomalidomidtherapie und 28 d nach Ende der Studienmedikation zu verzichten • Verzicht auf Blutspende während der Studienmedikation und 28 d nach Ende der Studientherapie • Zusage, keine Studienmedikation weiterzugeben <p>Ausschlusskriterien:</p>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Absolute Anzahl Neutrophiler <1000/μl • Thrombozyten <75.000/μl für Teilnehmer, deren Plasmazellanteil im Knochenmark <50 % bzw. <30.000/μl für Teilnehmer, deren Plasmazellanteil im Knochenmark \geq50 % • Kreatinin-Clearance <45 ml/min berechnet aus der Cockcroft-Gault-Formel; falls Berechnung aus dem 24 h Urin \geq45 ml/min, ist ein Einschluss erlaubt • Korrigiertes Calcium im Serum >14 mg/dl • Hb <8 g/dl (vorherige Erythrozyten-Transfusion oder Gabe von rekombinantem humanen EPO erlaubt) • Serum SGOT/AST oder SGPT/ALT >das 3-fache des oberen Normwerts • Serum Bilirubin-Spiegel >2,0 mg/dl oder >das 3-fache des oberen Normwerts für Teilnehmer mit erblicher gutartiger Hyperbilirubinämie • Maligne Vorerkrankung (außer MM, Basalzell- oder Plattenepithelkarzinome der Haut, in situ Zervix- oder Brustkarzinome, zufälliger histologischer Befund von Prostata-Krebs mit TNM Status T1a oder T1b), sofern der Teilnehmer nicht seit \geq5 J krankheitsfrei war • Vorherige Therapie mit Pomalidomid • Überempfindlichkeit gegenüber Thalidomid, Lenalidomid oder Dexamethason • Resistenz auf HD-Dex in der letzten Behandlungslinie • Periphere Neuropathie \geqGrad 2 • Allogene Knochenmarks- oder periphere Blutstammzelltransplantation \leq12 Mo vor Beginn der Studienmedikation und kein Abbruch der immunsuppressiven Behandlung \geq4 Wochen vor Beginn der Studienmedikation und momentane immunsuppressive Behandlung erforderlich • Geplante oder erforderliche Stammzelltransplantation • Kongestive Herzinsuffizienz (NYHA Klasse III oder IV), Myokardinfarkt innerhalb 12 Mo vor Beginn der Studienmedikation oder instabile oder schwer einstellbare Angina pectoris (inkl. Prinzmetal Angina pectoris) • Eine der folgenden Behandlungen innerhalb 14 d vor Beginn der Studienmedikation:

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> - Plasmapherese - Aufwendige Operation - Strahlentherapie - Jegliche medikamentöse Myelomtherapie • Jegliche nicht durch den Sponsor genehmigte Anwendung von Prüfartzeimitteln innerhalb von 28 d oder 5 Halbwertszeiten vor Behandlungsbeginn • Dauerhafte Gabe von Steroiden oder Immunsuppressiva erfordernde Erkrankungen; Gabe von >10 mg Prednison oder äquivalentem Steroid/d innerhalb 3 Wo vor Studieneinschluss • Jegliche durch Studienteilnahme zu nicht akzeptablen Risiken führende Erkrankung (inkl. auffälliger Laborwerte) • Gastrointestinale Erkrankungen, die die Absorption von Pomalidomid signifikant verändern könnten • Keine Möglichkeit/Bereitschaft zur prophylaktischen Thrombosebehandlung • Jegliche schwerwiegende Erkrankung, auffällige Laborwerte oder psychische Erkrankung, die die Gabe des <i>informed consent</i> behindern • Schwangere oder stillende Frauen • Bekannte HIV - oder aktive, infektiöse Hepatitis A, B oder C Infektion
4b	Studienorganisation und Ort der Studiendurchführung	<p>Studienorganisation: Christian Jacques, MD (Celgene Corporation 86 Morris Avenue, Summit, NJ, USA)</p> <p>93 Studienzentren: Europa (58), Australien (10), UK (10), Kanada (10), Russland (4), USA (1)</p>
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>A: Pom+LD-Dex</p> <p>Teilnehmer ≤75 J:</p> <ul style="list-style-type: none"> • Pom (p. o.): 4 mg an d 1-21 eines 28 d Zyklus • LD-Dex (p. o.): 40 mg an d 1,8,15 und 22 eines 28 d Zyklus <p>Teilnehmer >75 J:</p> <ul style="list-style-type: none"> • Pom (p. o.): 4 mg an d 1-21 eines 28 d Zyklus • LD-Dex (p. o.): 20 mg an d 1,8,15 und 22 eines 28 d Zyklus <p>B: HD-Dex</p> <p>Teilnehmer ≤75 J:</p> <ul style="list-style-type: none"> • HD-Dex (p. o.): 40 mg an d 1-4, d 9-12

Item ^a	Charakteristikum	Studieninformation
		und d 17-20 eines 28 d Zyklus Teilnehmer >75 J: <ul style="list-style-type: none"> • HD-Dex (p. o.): 20 mg an d 1-4, d 9-12 und d 17-20 eines 28 d Zyklus
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium:</p> <ul style="list-style-type: none"> • PFS* <p>Sekundäre Endpunkte:</p> <ul style="list-style-type: none"> • OS* • Sicherheit • Ansprechrate gemäß den IMWG-Kriterien • Ansprechrate gemäß den EBMT-Kriterien • TTP • Zeit bis zum Ansprechen • Ansprehdauer • Klinisches Ansprechen <ul style="list-style-type: none"> - Erhöhung des Hb-Werts - Verbesserung von Knochenschmerzen - Verbesserung der Nierenfunktion - Verbesserung des ECOG PS • Lebensqualität <p>Erhebungsmethoden und -zeitpunkte</p> <ul style="list-style-type: none"> • ECOG PS: d 1 jedes Zyklus (inkl. Zyklus 1) und bei Behandlungsabbruch oder Progression • Lebensqualität: EORTC QLQ-MY20, EORTC QLQ-C30, EQ-5D; elektronische Übermittlung an die Studienteilnehmer an d 1 jedes Zyklus (inkl. Zyklus 1) und bei Behandlungsabbruch oder Progression • Ansprechen: d 1 jedes Zyklus ab Zyklus 2 und bei Therapieende, alle 28 d während des PFS <i>follow-up</i> • OS: alle 84 d, beginnend 28 d nach Studienende für 5 Jahre ab Randomisierung <p>Definitionen</p> <ul style="list-style-type: none"> • PFS: Zeit ab Randomisierung bis Krankheitsprogression gemäß den IMWG- und EBMT-Kriterien oder Tod • OS: Zeit ab Randomisierung bis zum Tod jeglicher Ursache • Ansprechen gemäß den IMWG- und EBMT-Kriterien • TTP: Zeit ab Randomisierung bis Zeitpunkt der ersten dokumentierten Krankheitsprogression

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Zeit bis zum Ansprechen: Zeit ab Randomisierung bis zum ersten dokumentierten Ansprechen (\geqPR) gemäß den IMWG- oder EMBT-Kriterien • Ansprehdauer: Zeit ab dem ersten dokumentierten Ansprechen (\geqPR) bis Krankheitsprogression • Klinisches Ansprechen: <ul style="list-style-type: none"> - Erhöhung des Hb-Wertes um ≥ 1 Kategorie ab Baseline in CTCAE Grad für Hb-Spiegel - Verbesserung der Knochenschmerzen um ≥ 1 Kategorie ab Baseline bzgl. der Antworten auf das EORTC MM Modul (QLQ-MY20), Frage 1 - Verbesserung der Nierenfunktion um ≥ 1 Kategorie ab Baseline, für Nierenfunktionen, die in normal, Grad 1 (Kreatinin-Clearance 60 ml/min-unterer Normalwert) oder Grad 2 (Kreatinin-Clearance 45-60 ml/min - Verbesserung des ECOG PS Status um ≥ 1 Kategorie ab Baseline
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderung der Zielkriterien nach Studienbeginn
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Geplante Aufnahme von 426 Teilnehmern (A: 284; B: 142)</p> <p>PFS:</p> <ul style="list-style-type: none"> • 242 Ereignisse (Progression oder Tod) nötig, um ein HR von 1,5 zwischen den beiden Studienarmen nachzuweisen (Zweiseitiges Signifikanzniveau=5 %, Teststärke=85 %) <p>OS:</p> <ul style="list-style-type: none"> • 212 Ereignisse (Tod) nötig um ein HR von 1,5 zwischen den beiden Studienarmen nachzuweisen (Zweiseitiges Signifikanzniveau=5 %, Teststärke=81 %)
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<ul style="list-style-type: none"> • Geplante Zwischenanalyse für PFS (<i>Futility-Analyse</i>) nach Krankheitsprogression oder Tod bei 50 % der geplanten Ereignisse (=121); Erwägen eines Studienabbruchs bei Unterschreiten der Futility Grenze • Geplante Zwischenanalyse für OS auf überragende Wirksamkeit zum Zeitpunkt der finalen PFS-Analyse oder bei 106 Todesfällen, Studienabbruch bei Über-

Item ^a	Charakteristikum	Studieninformation
		schreiten des O'Brien-Fleming-Grenzwertes für Überlegenheit
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	<ul style="list-style-type: none"> Die Randomisierung erfolgte mit einem validierten <i>interactive voice/web response system</i> (IVRS/IWRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	<p>Blockrandomisierung innerhalb der Strata</p> <p>Strata:</p> <ul style="list-style-type: none"> Alter (≤ 75 J vs. > 75 J) Krankheitspopulation (Refraktäre Erkrankung mit Progression innerhalb 60 Tagen auf Lenalidomid und Bortezomib vs. rezidierte und refraktäre Erkrankung mit Ansprechen \geq PR und Progression innerhalb von 6 Monaten vs. refraktäre/intolerante Erkrankung/Teilnehmer, mit Intoleranz/Toxizität nach ≥ 2 Bortezomibzyklen. Anzahl Vortherapien (2 vs. > 2)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Randomisierung erfolgte mit einem validierten IVRS/IWRS
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Aufnahme der Teilnehmer durch ausgewiesenes Studienpersonal an den Studienzentren
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>a) Nicht verblindet</p> <p>b) Nicht verblindet</p> <p>c) Verblindet</p> <p>Review aller Effektivitätsdaten durch ein verblindetes, unabhängiges Entscheidungskomitee (IRAC). Verblindete Festlegung des Ansprechens und der Progression</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<ul style="list-style-type: none"> PFS: Schätzung der Überlebensfunktionen für PFS durch Kaplan-Meier Methode; Vergleich der Überlebensfunktionen durch stratifizierten Log-Rank-Test Ansprechen gemäß den IMWG- und EBMT-Kriterien: Vergleich der Ansprech-

Item ^a	Charakteristikum	Studieninformation
		raten durch exakten Test <ul style="list-style-type: none"> • Zeit bis zum Ansprechen, klinisches Ansprechen: Vergleich durch Wilcoxon-Rangsummen-Test
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<ul style="list-style-type: none"> • Analysen für OS und PFS in den Strata • Zusätzliche Analysen für OS und PFS (Geschlecht, zytogenetische Kategorien bei Studieneinschluss, Ansprechen-Status am Ende des Zyklus 2)
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die <ol style="list-style-type: none"> a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden 	<ol style="list-style-type: none"> a) Interventionsgruppe: 302 Patienten, Kontrollgruppe: 153 Patienten b) Interventionsgruppe: 300 Patienten, Kontrollgruppe: 149 Patienten c) Interventionsgruppe: 302 Patienten, Kontrollgruppe: 153 Patienten
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	A: Pom+LD-Dex <ul style="list-style-type: none"> • PD (35,4 %) • UE (6,6 %) • Tod (6,3 %) • Rücknahme des <i>informed consent</i> (2,0 %) • Andere Gründe (4,3 %) B: HD-Dex <ul style="list-style-type: none"> • PD (49,0 %) • UE (5,9 %) • Tod (9,8 %) • Rücknahme des <i>informed consent</i> (3,3 %) • Andere Gründe (6,5 %)
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Studienaufnahme: 18.03.11 - 31.08.12 Nachbeobachtung: 07.09.12 (interim data cut-off)
14b	Informationen, warum die Studie endete oder beendet wurde	MM-003 wurde noch nicht beendet
a: nach CONSORT 2010.		

*MM-003 war auf PFS und OS gepowert (d. h. auch OS wurde confirmatorisch unter Einhaltung des globalen 5 % Signifikanzniveaus getestet)

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

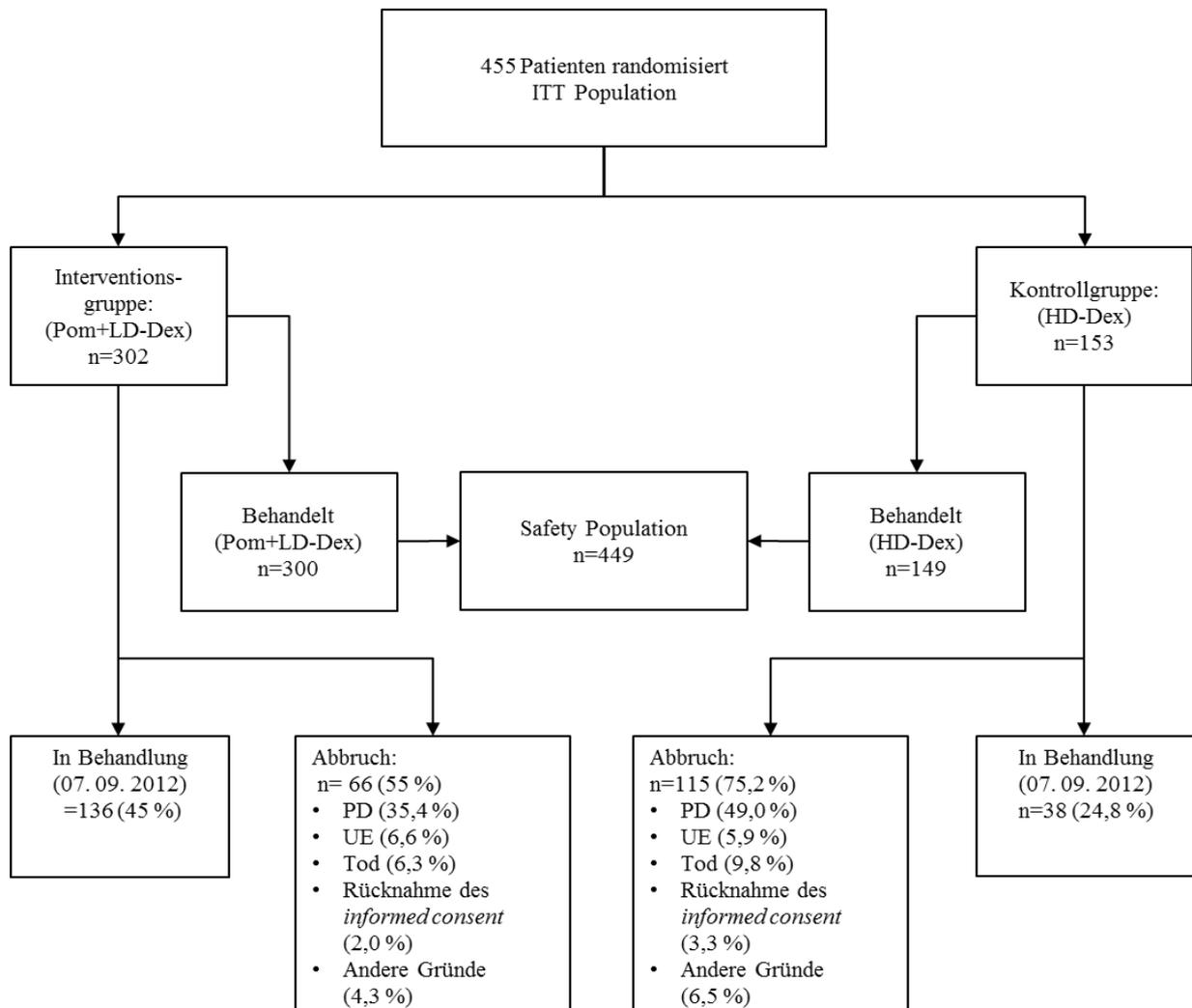


Abbildung 4-5(Anhang): Flussdiagramm der eingeschlossenen Studie - RCT

Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

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

Tabelle 4-62(Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MM-003

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

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

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

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

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

Studie: MM-003**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

Genaue Benennung der Quelle	Kürzel
Studienbericht	(A)
Studienprotokoll	(B)
Clinicaltrials.gov	(C)
Clinicaltrialsregister.eu	(D)
ICTRP	(E)

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

A: Kapitel/Abschnitt: *Investigational Plan: Overall Study Design and Plan: Description*

B: Kapitel/Abschnitt: *Overall Study Design: Study Design*

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

A: Kapitel/Abschnitt: *Treatments: Method of Assigning Patients to Treatment Groups*

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

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

A: Kapitel/Abschnitt: *Treatments: Method of Assigning Patients to Treatment Groups*

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

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

3. Verblindung von Patienten und Behandlern**Patient**

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

A: Kapitel/Abschnitt: *Treatments: Blinding*: Es handelt sich um eine Studie mit offenem Studiendesign.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

A: Kapitel/Abschnitt: *Treatments: Blinding*: Es handelt sich um eine Studie mit offenem Studiendesign.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht/unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen/Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

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

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zu Gunsten der zu evaluierenden Behandlung).

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

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

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

A: Ganzes Dokument

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine*Beachte die Hinweise zu Punkt A.4!*

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**

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

Aus dem Kontrollarm wechselte ein hoher Anteil Studienteilnehmer bei Krankheitsprogression in eine Begleitstudie und wurde mit Pomalidomid behandelt. Die Analyse der Daten erfolgte nach dem ITT-Prinzip, so dass sich dieser Effekt zuungunsten des Interventionsarms auswirkte.

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Studienteilnehmer der Kontrollgruppe hatten die Möglichkeit nach Krankheitsprogression in eine Begleitstudie zu wechseln und dort mit Pomalidomid behandelt zu werden. Insgesamt lag der Anteil Studienteilnehmer, die zu einer Pomalidomid-basierten Therapie gewechselt haben bei etwa 29 %, so dass hierdurch ein erhöhtes Verzerrungspotenzial nicht ausgeschlossen werden kann. Die Analyse der Daten erfolgte nach dem ITT-Prinzip, so dass sich dieser Effekt jedoch zuungunsten des Interventionsarms auswirkte.

Endpunkt: Progressionsfreies Überleben

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Discussion of the Study Design, Including the Choice of Control Groups*: Die behandelnden Ärzte und die Studienteilnehmer waren unverblindet, die erhobenen Daten wurden jedoch an ein verblindetes Komitee (IRAC, Independent Response Adjudication Committee) weitergeleitet. Die Zeitpunkte für eine Progression wurden von dem verblindeten IRAC festgelegt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders

wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zu Gunsten der zu evaluierenden Behandlung).

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

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

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

A: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Zeit bis zur Progression**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Discussion of the Study Design, Including the Choice of Control Groups*: Die behandelnden Ärzte und die Studienteilnehmer waren unverblindet, die erhobenen Daten wurden jedoch an ein verblindetes Komitee (IRAC) weitergeleitet. Die Zeitpunkte für eine Progression wurden von dem verblindeten IRAC festgelegt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders

wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zu Gunsten der zu evaluierenden Behandlung).

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

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

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

A: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gesamtansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Discussion of the Study Design, Including the Choice of Control Groups*: Die behandelnden Ärzte und die Studienteilnehmer waren unverblindet, die erhobenen Daten wurden jedoch an ein verblindetes Komitee (IRAC) weitergeleitet. Das Ansprechen wurde von dem verblindeten IRAC festgelegt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Objektives Ansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Discussion of the Study Design, Including the Choice of Control Groups*: Die behandelnden Ärzte und die Studienteilnehmer waren unverblindet, die erhobenen Daten wurden jedoch an ein verblindetes Komitee (IRAC) weitergeleitet. Das Ansprechen wurde von dem verblindeten IRAC festgelegt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Klinisches Ansprechen: Verbesserung des ECOG PS**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Klinisches Ansprechen: Verbesserung von Knochenschmerzen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Lebensqualität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Mind. 1 Unerwünschtes Ereignis**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Mind. 1 Schwerwiegendes Unerwünschtes Ereignis**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Mind. 1 zum Absetzen des Studienarzneimittels führendes Unerwünschtes Ereignis**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Tod ≤ 30 d nach letzter Dosis Studien-AM**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.

Endpunkt: Mind. 1 Unerwünschtes Ereignis \geq CTC Grad 3**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

A: Kapitel/Abschnitt: *Investigational plan*: MM-003 wurde mit offenem Studiendesign durchgeführt

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

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

A: Kapitel/Abschnitt: *Ganzes Dokument*

B: Kapitel/Abschnitt: *Ganzes Dokument*

C: Kapitel/Abschnitt: *Ganze Seite*

D: Kapitel/Abschnitt: *Ganze Seite*

E: Kapitel/Abschnitt: *Ganze Seite*

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

MM-003 wurde mit offenem Design durchgeführt. Ein Einfluss der Kenntnis der Gruppenzugehörigkeit auf das Ergebnis und eine dadurch resultierende Verzerrung kann nicht ausgeschlossen werden.
