

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Obinutuzumab (Gazyvaro[®])

Roche Pharma AG

Modul 4 A

Patienten mit follikulärem Lymphom (FL), die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	6
Abkürzungsverzeichnis.....	7
4 Modul 4 – allgemeine Informationen	9
4.1 Zusammenfassung der Inhalte von Modul 4.....	10
4.2 Methodik	15
4.2.1 Fragestellung	15
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	17
4.2.3 Informationsbeschaffung	17
4.2.3.1 Studien des pharmazeutischen Unternehmers	17
4.2.3.2 Bibliografische Literaturrecherche	17
4.2.3.3 Suche in Studienregistern	18
4.2.3.4 Selektion relevanter Studien	19
4.2.4 Bewertung der Aussagekraft der Nachweise.....	20
4.2.5 Informationssynthese und -analyse	21
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	21
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	22
4.2.5.3 Meta-Analysen.....	22
4.2.5.4 Sensitivitätsanalysen	23
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	24
4.2.5.6 Indirekte Vergleiche	25
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	27
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	27
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	27
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	27
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	28
4.3.1.1.3 Studien aus der Suche in Studienregistern	30
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	30
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	31
4.3.1.2.1 Studiendesign und Studienpopulationen	31
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	41
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	42
4.3.1.3.1 Endpunkte für die Patientenpopulation mit follikulärem Lymphom – RCT.....	43
4.3.1.3.1.1 Gesamtüberleben – RCT	43
4.3.1.3.1.2 Progressionsfreies Überleben (IRC) – RCT	47
4.3.1.3.1.3 Krankheitsfreies Überleben (IRC) – RCT	51
4.3.1.3.1.4 Gesundheitsbezogene Lebensqualität - FACT-Lym - RCT	57

4.3.1.3.1.5	Verträglichkeit – RCT	62
4.3.1.3.2	Subgruppenanalysen – RCT.....	75
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	75
4.3.2	Weitere Unterlagen.....	75
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	75
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	76
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	76
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	76
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	76
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	79
4.3.2.2	Nicht randomisierte vergleichende Studien.....	79
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	79
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	79
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	80
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	80
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	81
4.3.2.3	Weitere Untersuchungen.....	81
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	82
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	82
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	82
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	82
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	83
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	83
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	84
4.4.1	Beurteilung der Aussagekraft der Nachweise	84
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	84
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	91
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	91
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	91
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	91
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	92
4.5.4	Verwendung von Surrogatendpunkten.....	92
4.6	Liste der eingeschlossenen Studien.....	93
4.7	Referenzliste.....	94
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		98
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		100
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		101
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		102

Anhang 4-E : Methodik der eingeschlossenen Studien – RCT	103
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	106

Tabellenverzeichnis**Seite**

Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	28
Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	28
Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	30
Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	31
Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	32
Tabelle 4-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	33
Tabelle 4-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel – Patienten mit FL (Datenschnitt 01. September 2014 und 01. Mai 2015)	33
Tabelle 4-8: Datenschnitte und Wertigkeit der Analysen in GADOLIN.....	38
Tabelle 4-9: Anteil Patienten mit follikulärem Lymphom in GADOLIN	39
Tabelle 4-10: Anteil FL-Patienten in GADOLIN zu den Beobachtungszeitpunkten	40
Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	41
Tabelle 4-12: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	42
Tabelle 4-13: Operationalisierung von Gesamtüberleben.....	44
Tabelle 4-14: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel.....	44
Tabelle 4-15: Ergebnisse für das Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel.....	45
Tabelle 4-16: Operationalisierung von Progressionsfreies Überleben.....	47
Tabelle 4-17: Bewertung des Verzerrungspotenzials für PFS (IRC) in RCT mit dem zu bewertenden Arzneimittel	48
Tabelle 4-18: Ergebnisse für PFS (IRC) aus RCT mit dem zu bewertenden Arzneimittel	49
Tabelle 4-19: Operationalisierung von Krankheitsfreies Überleben.....	51
Tabelle 4-20: Bewertung des Verzerrungspotenzials für DFS und Rezidivrate (IRC) in RCT mit dem zu bewertenden Arzneimittel	52
Tabelle 4-21: Ergebnisse für Rezidivrate (IRC) aus RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-22: Zeit bis zum Rezidiv (IRC) aus RCT mit dem zu bewertenden Arzneimittel...	54
Tabelle 4-23: Landmark-Analyse für Rezidive (IRC) aus RCT mit dem zu bewertenden Arzneimittel.....	54
Tabelle 4-24: Ergebnisse für DFS (IRC) aus RCT mit dem zu bewertenden Arzneimittel.....	55
Tabelle 4-25: Operationalisierung von FACT-Lym.....	57

Tabelle 4-26: Bewertung des Verzerrungspotenzials für Gesundheitsbezogene Lebensqualität – FACT-Lym - in RCT mit dem zu bewertenden Arzneimittel	58
Tabelle 4-27: Anteil der Patienten mit Bewertung zu den Zeitpunkten Zyklus 5 Tag 1 und Follow-Up Visite 1 Monat 2 (Datenschnitt vom 01. September 2014 und 01. Mai 2015).....	59
Tabelle 4-28: Ergebnisse der Responderanalysen zur Verbesserung der Lebensqualität um die jeweilige MID aus RCT mit dem zu bewertenden Arzneimittel (01. September 2014)....	60
Tabelle 4-29: Ergebnisse der Responderanalysen zur Verbesserung der Lebensqualität um die jeweilige MID aus RCT mit dem zu bewertenden Arzneimittel (01. Mai 2015).....	61
Tabelle 4-30: Operationalisierung von Verträglichkeit	62
Tabelle 4-31: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-32: Wertigkeit der Analysen zur Darstellung der Verträglichkeit unter Berücksichtigung des Studiendesigns und der Verzerrung.....	65
Tabelle 4-33: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-34: Ergebnisse für IRR und UE, die in Zusammenhang mit IRR gesehen werden	67
Tabelle 4-35: Ergebnisse für hämatologische unerwünschte Ereignisse und Infektionen.....	68
Tabelle 4-36: Weitere ausgewählte UE aus RCT mit dem zu bewertenden Arzneimittel	71
Tabelle 4-37: Time-to-first-event Analyse (TTfE)	73
Tabelle 4-38: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	76
Tabelle 4-39: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	77
Tabelle 4-40: Operationalisierung von <Endpunkt xxx>.....	77
Tabelle 4-41: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	78
Tabelle 4-42: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	78
Tabelle 4-43: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	80
Tabelle 4-44: Operationalisierung von <Endpunkt xxx>.....	80
Tabelle 4-45: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	81
Tabelle 4-46: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	82
Tabelle 4-47: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	91
Tabelle 4-48 (Anhang): Studiendesign und -methodik für Studie	104
Tabelle 4-49 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GADOLIN	107

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	29
Abbildung 2: Studiendesign (7)	34
Abbildung 3: Kaplan-Meier-Kurve zum Endpunkt OS (Datenschnitt 01. Mai 2015) (12)	46
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	46
Abbildung 5: Kaplan-Meier-Kurve für den Endpunkt PFS (IRC) (Datenschnitt 01. Sept. 2014) (7).....	50
Abbildung 6: Kaplan-Meier-Kurve zum Endpunkt DFS (IRC) (Datenschnitt 01. September 2014) (18).....	55

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AEGT	Adverse event grouped term
AEPI	Adverse events of particular interest (unerwünschte Ereignisse von besonderem Interesse)
AESI	Adverse events of special interest (unerwünschte Ereignisse von speziellem Interesse)
ARR	Absolute Risikoreduktion
Benda	Bendamustin
CLL	Chronisch lymphatische Leukämie
CONSORT	Consolidated Standards of Reporting Trials
CR	Complete remission (vollständiges Ansprechen)
CSR	Clinical Study Report
DFS	Disease free survival (krankheitsfreies Überleben)
DIMDI	Deutsches Institut für Medizinische Dokumentation
DoR	Duration of response (Dauer des Ansprechens)
ECOG	Eastern Cooperative Oncology Group
EFS	Event free survival (ereignisfreies Überleben)
EG	Europäische Gemeinschaft
FACT	Functional Assessment of Cancer Therapy (gesundheitsbezogene Lebensqualität bei Krebspatienten)
FACT-G	Functional Assessment of Cancer Therapy general
FACT-Lym	Functional Assessment of Cancer Therapy – Lymphoma
FACT-Lym TOI	Functional Assessment of Cancer Therapy – Lymphoma Trial Outcome Index
FDA	Food and Drug Administration
FL	Follikuläres Lymphom
G-Benda	Induktionstherapie bestehend aus Obinutuzumab in Kombination mit Bendamustin, gefolgt von einer Obinutuzumab-Erhaltungstherapie
iNHL	Indolentes Non-Hodgkin-Lymphom
INN	International NonProprietary Name (Internationaler Freiname)
IRC	Independent Review Committee (unabhängiges Bewertungsgremium)
ITT	Intention to Treat
IVRS	Interactive Voice Response System

Abkürzung	Bedeutung
KI	Konfidenzintervall
LS	Least Squares (Methode der kleinsten Quadrate)
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference
MTC	Mixed Treatment Comparison
MW	Mittelwert
MWD	Mittelwertdifferenz
n.b.	Nicht berechenbar
n.e.	Nicht erreicht
n.z.	Nicht zutreffend
OR	Odds ratio (Quotenverhältnis)
ORR	Overall response (Gesamtansprechrage)
OS	Overall survival (Gesamtüberleben)
PFS	Progression free survival (progressionsfreies Überleben)
PR	Partial response (Partielles Ansprechen)
RCT	Randomized Controlled Trial
RR	Relatives Risiko
rRR	Reversiertes relatives Risiko
SE	Standard Error (Standardfehler)
SGB	Sozialgesetzbuch
SMQ	Standardised MedDRA Queries (standardisierte MedDRA Abfrage)
SPD	Sum of product diameters bzw. Summe der Produkte der Durchmesser
STD	Standardabweichung
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
TLS	Tumorlysesyndrom
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschte Ereignis/e
Verfo	Verfahrensordnung des Gemeinsamen Bundesausschusses
WHO	World Health Organization

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.

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

Obinutuzumab (Gazyvaro®) ist ein Arzneimittel zur Behandlung einer seltenen Erkrankung nach Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates (EU orphan designation number EU/3/15/1504) (1). Damit gilt für Obinutuzumab der medizinische Nutzen und Zusatznutzen mit der Zulassung als belegt (Absatz 1 Satz 10 §35a SGB V) (2).

Gemäß Absatz 1 Satz 1 Nummer 1 §12 des 5. Kapitels der Verfahrensordnung des Gemeinsamen Bundesausschusses (VerfO) ist im Dossier folgende Fragestellung zu untersuchen (3):

Wie groß ist unter Berücksichtigung patientenrelevanter Endpunkte zu Mortalität, Morbidität, Lebensqualität und Verträglichkeit das Ausmaß des Zusatznutzens von Obinutuzumab in Kombination mit Bendamustin als Induktionstherapie, gefolgt von einer Obinutuzumab-Erhaltungstherapie (G-Benda) für Patienten mit follikulärem Lymphom (FL), die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden im Vergleich zu der in der Zulassungsstudie herangezogenen Vergleichstherapie Bendamustin (Benda)?

Datenquellen

Zur Beantwortung der Fragestellung wurde die Zulassungsstudie GADOLIN (GAO4753g), eine randomisierte, kontrollierte Studie (RCT), herangezogen.

Ein-/Ausschlusskriterien für Studien

Nicht zutreffend.

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

Die Kriterien zur Bewertung der Aussagekraft folgen den Vorgaben der VerfO (3). Die Bewertung erfolgt auf der Basis der zur Verfügung stehenden Studienberichte zur Zulassungsstudie gegebenenfalls ergänzt durch für die Nutzenbewertung notwendige Zusatzanalysen.

Es erfolgt sowohl eine studienbezogene als auch endpunktbezogene Bewertung des Verzerrungspotentials anhand der vorgegebenen Bewertungsbögen in Anhang 4-F.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen***Mortalität*****v Gesamtüberleben (OS)**

G-Benda reduzierte das Sterberisiko im Vergleich zu Benda statistisch signifikant um 38 % (HR: 0,62 [0,39; 0,98]; $p=0,0379$). Bei einer medianen Beobachtungsdauer von 24,1 Monaten verstarben im G-Benda-Arm 18,3 % der Patienten und im Benda-Arm 28,1 % der Patienten. Insgesamt wurde der Median für den Endpunkt OS in GADOLIN bisher nicht erreicht.

Somit ist GADOLIN die erste randomisierte, kontrollierte Studie, die für den Endpunkt Gesamtüberleben (OS) einen signifikanten Vorteil gegenüber der Vergleichstherapie für Patienten mit Rituximab-refraktärem FL nachweisen konnte.

Auf Basis der vom G-BA geforderten bestverfügbaren Evidenz besteht nach Definition gemäß VerfO ein beträchtlicher Zusatznutzen für den patientenrelevanten Endpunkt Gesamtüberleben in der Domäne Mortalität.

Morbidität**v Progressionsfreies Überleben (PFS) (IRC)**

Durch G-Benda wurde das Risiko für ein PFS-Ereignis (Progression, Rezidiv oder Tod) im Vergleich zu Benda signifikant um 52 % reduziert (HR: 0,48; 95 % KI [0,34; 0,68]; $p < 0,0001$). Im Benda-Arm traten bei 54,2 % und im G-Benda-Arm bei 34,8 % der Patienten PFS-Ereignisse (IRC) auf.

Die mediane Dauer bis zu einem PFS-Ereignis (IRC) wurde für G-Benda nicht erreicht und betrug für Benda 13,8 Monate. Der Median für G-Benda liegt für diesen Datenschnitt nur für PFS (Investigator) vor, beträgt 29,2 Monate und wird im Follow-Up Datenschnitt vom 01. Mai 2015 durch das PFS (IRC) bestätigt. Dies bedeutet im Median einen Vorteil von 15,4 Monaten für G-Benda gegenüber Benda.

Aufgrund der statistisch signifikanten deutlichen Verzögerung eines Fortschreitens oder Wiedereintretens der Erkrankung durch G-Benda besteht ein beträchtlicher Zusatznutzen für den Endpunkt PFS (IRC) in der Domäne Morbidität.

v Krankheitsfreies Überleben (IRC)

Der Endpunkt krankheitsfreies Überleben operationalisiert, neben der Mortalität, in der Morbidität das patientenrelevante Wiedereintreten der Erkrankung durch ein Rezidiv bei Patienten, die zuvor unter der Therapie eine komplette Remission und damit völlige Freiheit jeglicher Krankheitszeichen erreicht haben. Die möglichst lange Aufrechterhaltung eines solchen Therapieerfolges ist neben dem Gesamtüberleben anerkanntes und angestrebtes Therapieziel

im Sinne des Patienten. Ein Effekt hinsichtlich der Rezidivrate, der Zeit bis zum Rezidiv und folglich krankheitsfreies Überleben ist als patientenrelevant zu betrachten.

Rezidive (Morbidität) wurden zum Zeitpunkt des Datenschnitts vom 01. September 2014 bei 6,3 % im G-Benda-Arm gegenüber 45,2 % im Benda-Arm beobachtet. Das Risiko ein Rezidiv zu erleiden wird durch G-Benda um 86 % reduziert. Im Benda-Arm rezidierten nach 6 Monaten ein 21,3 % und nach 12 Monaten 47,2 % aller Patienten, die zuvor eine komplette Remission erreicht hatten. Unter G-Benda wurden keine Frührezidive festgestellt (innerhalb von 6 Monaten nach CR) und nach 12 Monaten rezidierten 8,9 % der Patienten. Die mediane Dauer bis zu einem Rezidiv ist für G-Benda nicht erreicht und beträgt für Benda 13,0 Monate (95% KI [6,9, n.e.]). Die Hazard Ratio (G-Benda vs. Benda) beträgt 0,10 (95 % KI [0,02; 0,46]; $p = 0,0004$). In Abwesenheit eines Medians errechnet sich unter der Annahme stabiler Hazard Ratios (G-Benda vs. Benda 0,10 (95 % KI [0,02; 0,46]; $p = 0,0004$) ein geschätzter Median für die Rezidivfreiheit von G-Benda von 130,0 Monaten.

Gesamthaft besteht für den Endpunkt krankheitsfreies Überleben (Morbidität und Mortalität) ein signifikanter Vorteil für G-Benda (HR 0,10; 95 % KI [0,02; 0,44]; $p = 0,0002$). Die mediane Dauer des krankheitsfreien Überlebens beträgt für Benda 13,0 Monate, für G-Benda unter Annahme stabiler Hazard Ratios errechnet sich in Abwesenheit eines Medians eine Dauer von 94,3 Monaten.

Aufgrund der statistisch signifikanten Senkung der Rezidivrate, Verlängerung der Zeit bis zum Rezidiv, sowie der Reduktion des Risikos ein Rezidiv zu erleiden besteht aufgrund der langfristigen Abwesenheit aller Krankheitszeichen und Symptome basierend auf der Verfo ein beträchtlicher Zusatznutzen insgesamt für den Endpunkt Krankheitsfreies Überleben.

Gesundheitsbezogene Lebensqualität

v FACT-Lym

Die Gesundheitsbezogene Lebensqualität wird in der Zulassungsstudie GADOLIN systematisch beobachtet und mittels FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma), einem krankheitsspezifischen Messinstrument für Patienten mit Non-Hodgkin-Lymphom, erhoben.

Für den Endpunkt FACT-Lym ergaben sich keine statistisch signifikanten Unterschiede zwischen G-Benda und Benda. Dies entspricht gemäß Verfo keinem Zusatznutzen für G-Benda in der Endpunktkategorie Gesundheitsbezogene Lebensqualität. Die Ergebnisse zeigen jedoch, dass die vom Patienten berichtete Gesundheitsbezogene Lebensqualität bei deutlich besserer Wirksamkeit von G-Benda im Vergleich zu Benda nicht beeinträchtigt war.

Verträglichkeit

v Unerwünschte Ereignisse

Bei der vergleichenden Betrachtung der Sicherheitsprofile beider Arme ist eine Verzerrung zu Ungunsten des G-Benda-Arms festzuhalten. Diese ist begründet in der unterschiedlichen Behandlungsdauer, bedingt durch die zweijährige Erhaltungstherapie im G-Benda-Arm. Zudem wurde für die Mehrzahl der Patienten (150/168) im Benda-Arm keine systematische UE-Erfassung über die Induktionstherapie hinaus durchgeführt. Eine direkte Vergleichbarkeit beider Studienarme ist daher nur für die Induktionstherapie gegeben. Die Betrachtung der gesamten Studienphase wird aufgrund der Verzerrung durch eine Time-to-first-event Analyse ergänzt, die diese Verzerrung reduziert. Dennoch ist das Sicherheitsprofil von G-Benda gegenüber Benda aus statistisch-medizinischer Sicht insgesamt nicht nachteilig.

In GADOLIN ist die Inzidenz der unerwünschten Ereignisse jeglicher Art, UE $\geq$ Grad 3, UE Grad 3, UE Grad 4, UE Grad 5, SUE und Behandlungsabbruch aufgrund UE in beiden Therapiearmen sowohl in der Induktionstherapie als auch in der gesamten Studienphase vergleichbar.

Signifikante Unterschiede zu Ungunsten G-Benda bestanden in der Inzidenz von schwerwiegender und spät einsetzender Neutropenie, kardialen Ereignis und Hypotonie. Wenngleich schwerwiegende Neutropenien häufiger im G-Benda-Arm auftraten, führten diese nicht zu Unterschieden in den Inzidenzraten klinisch relevanter Neutropenie-assoziiierter Komplikationen, wie febrilen Neutropenien und Infektionen. Todesfälle aufgrund von Infektionen traten in beiden Behandlungsarmen in gleichem Umfang auf. Der Unterschied hinsichtlich kardialer Ereignisse zu Ungunsten von G-Benda erreichte statistische Signifikanz nur in der Betrachtung der gesamten Studienphase. Die kardialen UE im G-Benda-Arm waren bei 16/20 Patienten zum Zeitpunkt des aktuellsten Datenschnitts reversibel und keines zeigte einen fatalen Verlauf. Hypotonien traten ausschließlich im Rahmen von IRR auf, waren alle reversibel und fast ausschließlich UE Grad 1 oder 2 (107/110). Es trat kein Ereignis $>$ Grad 3 auf.

Signifikante Unterschiede zugunsten von G-Benda bestanden in der Inzidenz von Übelkeit, Erbrechen und ungewolltem Gewichtsverlust. Dieser Vorteil hinsichtlich gastrointestinaler Nebenwirkungen bestätigte sich auch in der Time-to-first-event Analyse. Diese wurde aufgrund des Ungleichgewichts bezüglich Behandlungsdauer und systematischer Dokumentation der UE zwischen den Behandlungsarmen durchgeführt. Sie zeigt zudem signifikante Vorteile für G-Benda für UE Grad 3 und 5, SUE, Thrombozytopenie, Infektionen, schwerwiegende Infektionen und Zweitmalignome. Ausschließlich für Hypotonien ergibt sich ein signifikanter Nachteil für G-Benda.

Die in der Effektivität überlegene Therapie mit G-Benda geht im Vergleich zu Benda in der Gesamtschau nicht mit einem Nachteil bezüglich Verträglichkeit einher. Dies entspricht gemäß VerFO keinem Zusatznutzen hinsichtlich des Endpunkts Verträglichkeit.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen***Zusammenfassung Mortalität***

Auf Basis der vom G-BA geforderten bestverfügbaren Evidenz besteht nach Definition gemäß VerfO ein beträchtlicher Zusatznutzen für den patientenrelevanten Endpunkt Gesamtüberleben in der Domäne Mortalität.

Zusammenfassung Morbidität

Aufgrund der statistisch signifikanten Verzögerung eines Fortschreitens oder Wiedereintretens der Erkrankung durch G-Benda sowie der statistisch signifikanten Verlängerung der Zeit ohne jegliche Krankheitszeichen mit Reduktion des Risikos für ein Rezidiv oder Tod um nahezu 90 % besteht insgesamt ein beträchtlicher Zusatznutzen in der Domäne Morbidität.

Zusammenfassung Gesundheitsbezogene Lebensqualität

Für den Endpunkt FACT-Lym ergaben sich keine statistisch signifikanten Unterschiede zwischen G-Benda und Benda. Dies entspricht gemäß VerfO keinem Zusatznutzen für G-Benda in der Endpunktkategorie Gesundheitsbezogene Lebensqualität.

Zusammenfassung Verträglichkeit

Die in der Effektivität überlegene Therapie mit G-Benda geht im Vergleich zu Benda in der Gesamtschau nicht mit einem Nachteil bezüglich Verträglichkeit einher. Dies entspricht gemäß VerfO keinem Zusatznutzen hinsichtlich des Endpunkts Verträglichkeit.

In der Gesamtschau stellen die unter G-Benda erzielten Therapieeffekte eine nachhaltige und bisher im Anwendungsgebiet nicht erreichte Verbesserung des therapielevanten Nutzens im Sinne einer Verlängerung der Überlebensdauer (statistisch signifikante Reduktion der Gesamtmortalität), einer langfristigen krankheits- und progressionsfreien Zeit (mit statistisch signifikanter Reduktion der Rezidiv- und Progressionsrate), vergleichbarer Gesundheitsbezogener Lebensqualität und nicht nachteiliger Verträglichkeit dar. Dies entspricht insgesamt einem beträchtlichen Zusatznutzen von G-Benda gegenüber Benda.

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

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

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.

Obinutuzumab (Gazyvaro[®]) ist ein Arzneimittel zur Behandlung einer seltenen Erkrankung nach Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates (EU orphan designation number EU/3/15/1504) (1). Damit gilt für Obinutuzumab der medizinische Nutzen und Zusatznutzen mit der Zulassung als belegt (Absatz 1 Satz 10 §35a SGB V) (2).

Gemäß Absatz 1 Satz 1 Nummer 1 §12 des 5. Kapitels der VerfO ist im Dossier folgende Fragestellung zu untersuchen (3):

Wie groß ist unter Berücksichtigung patientenrelevanter Endpunkte zu Mortalität, Morbidität, Gesundheitsbezogener Lebensqualität und Verträglichkeit das Ausmaß des Zusatznutzens von Obinutuzumab in Kombination mit Bendamustin als Induktionstherapie, gefolgt von einer Obinutuzumab-Erhaltungstherapie (G-Benda) für Patienten mit follikulärem Lymphom (FL), die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden im Vergleich zu der in der Zulassungsstudie herangezogenen Vergleichstherapie Bendamustin (Benda)? Die Fragestellung wird anhand der Zulassungsstudie diskutiert. Die Fragestellung konkretisiert sich weiterhin wie folgt:

Patientenpopulation

Gemäß Zulassung umfasst die für die Fragestellung betrachtete Zielpopulation Patienten mit follikulärem Lymphom (FL), die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden. Diese Patienten sind auch als Rituximab-refraktär definiert (4).

Intervention

Die zu bewertende Intervention ist Obinutuzumab (Gazyvaro[®]) in Kombination mit Bendamustin gefolgt von einer Erhaltungstherapie mit Obinutuzumab (G-Benda) gemäß den Angaben der Fachinformation (5). Obinutuzumab ist ein humanisierter, monoklonaler, Typ-II-Anti-CD20-Antikörper, welcher als Teil einer Chemo-Immuntherapie gegeben wird.

Vergleichstherapie

Als Vergleichstherapie wird für die vorliegende Fragestellung die Chemotherapie Bendamustin (Benda) gemäß internationaler Konsensusleitlinie herangezogen (6). Bendamustin ist eine alkylierende Substanz.

Patientenrelevante Endpunkte

Zur Beantwortung der vorliegenden Fragestellung werden folgende patientenrelevante Endpunkte in den Domänen Mortalität, Morbidität, Gesundheitsbezogene Lebensqualität und Verträglichkeit herangezogen:

Mortalität

- Gesamtüberleben

Morbidität

- Progressionsfreies Überleben
- Krankheitsfreies Überleben

Gesundheitsbezogene Lebensqualität

- FACT-Lym

Verträglichkeit

- Sicherheit

Weitere Details zur Auswahl, Validität und Patientenrelevanz der betrachteten Endpunkte sind in Kapitel 4.4.2 enthalten.

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. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

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.

Nicht zutreffend.

Im vorliegenden Dossier zur Nutzenbewertung von Obinutuzumab mit Orphan Status wird ausschließlich die Zulassungsstudie GADOLIN dargestellt.

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 erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ 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. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Nicht zutreffend.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmen ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

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

Nicht zutreffend.

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.

Nicht zutreffend.

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-F):

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 der behandelnden Personen
- 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 der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Kriterien zur Bewertung der Aussagekraft folgen den Vorgaben der Verfo (3). Die Bewertung erfolgt auf der Basis der zur Verfügung stehenden Studienberichte zur Zulassungsstudie gegebenenfalls ergänzt durch für die Nutzenbewertung notwendige Zusatzanalysen.

Es erfolgt sowohl eine studienbezogene als auch endpunktbezogene Bewertung des Verzerrungspotentials anhand der vorgegebenen Bewertungsbögen in Anhang 4-F.

Das Verzerrungspotential wurde entweder als "niedrig" oder "hoch" eingestuft.

Ein niedriges Verzerrungspotential liegt vor, wenn die Ergebnisse der Studie nicht relevant verzerrt sind. Von einem hohen Verzerrungspotential wird ausgegangen, wenn nicht ausgeschlossen werden kann, dass die Studienergebnisse in "unverzerrter" Darstellung zu anderen Schlussfolgerungen führen könnten.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

² Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.

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

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.

Nicht zutreffend.

Design und Methodik der in die vorliegende Nutzenbewertung eingeschlossenen Zulassungsstudie GADOLIN sind in Kapitel 4.3.1.2.1 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 patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. 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). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Die Beschreibung der in die vorliegende Nutzenbewertung eingeschlossenen patientenrelevanten Endpunkte erfolgt in Kapitel 4.4.2.

Angaben zur Operationalisierung und Validität eines Endpunkts erfolgen in den jeweiligen Unterkapiteln von Kapitel 4.3.1.3.1.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. 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 Effekten⁴ 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 sowohl anhand des Odds Ratios als auch des Relativen Risikos 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 verwendet⁵ 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 Heterogenität⁶ 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.

Nicht zutreffend.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. 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 Effekt-

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

⁵ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

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

maßes. 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.

Nicht zutreffend.

In GADOLIN wurden verschiedene Sensitivitätsanalysen für den Endpunkt Progressionsfreies Überleben (PFS) durchgeführt. Dabei wurde insbesondere der Einfluss spät aufgetretener Todesfälle, vorzeitiger Studienabbrüche, von Folgetherapien, fehlender IRC-Tumorbewertungen (IRC = Independent Review Committee bzw. unabhängiges Bewertungsgremium) oder eventueller Randomisierungsfehler auf die Ergebnisse der Hauptanalyse untersucht. Die Resultate der Sensitivitätsanalysen haben die Ergebnisse der Hauptanalyse für diesen Endpunkt bestätigt (7).

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

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen 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.

Nicht zutreffend.

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 möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹.

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vor-

⁷ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

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

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

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

¹¹ B. Schöttker, D. Lüthmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.

liegen (8-11). Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

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

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

Nicht zutreffend.

¹² Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

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

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*
- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, 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“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
GADOLIN	ja	ja	laufend	8 Jahre	G-Benda, Benda

Stand von Tabelle 4-1: 23.05.2016.

Geben Sie an, welchen Stand die Information in Tabelle 4-1 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die für die Nutzenbewertung relevante Studie GADOLIN ist in Kapitel 4.3.1.2 beschrieben.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-1 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
keine	entfällt aufgrund des Orphan-Status

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. 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 Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

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

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

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

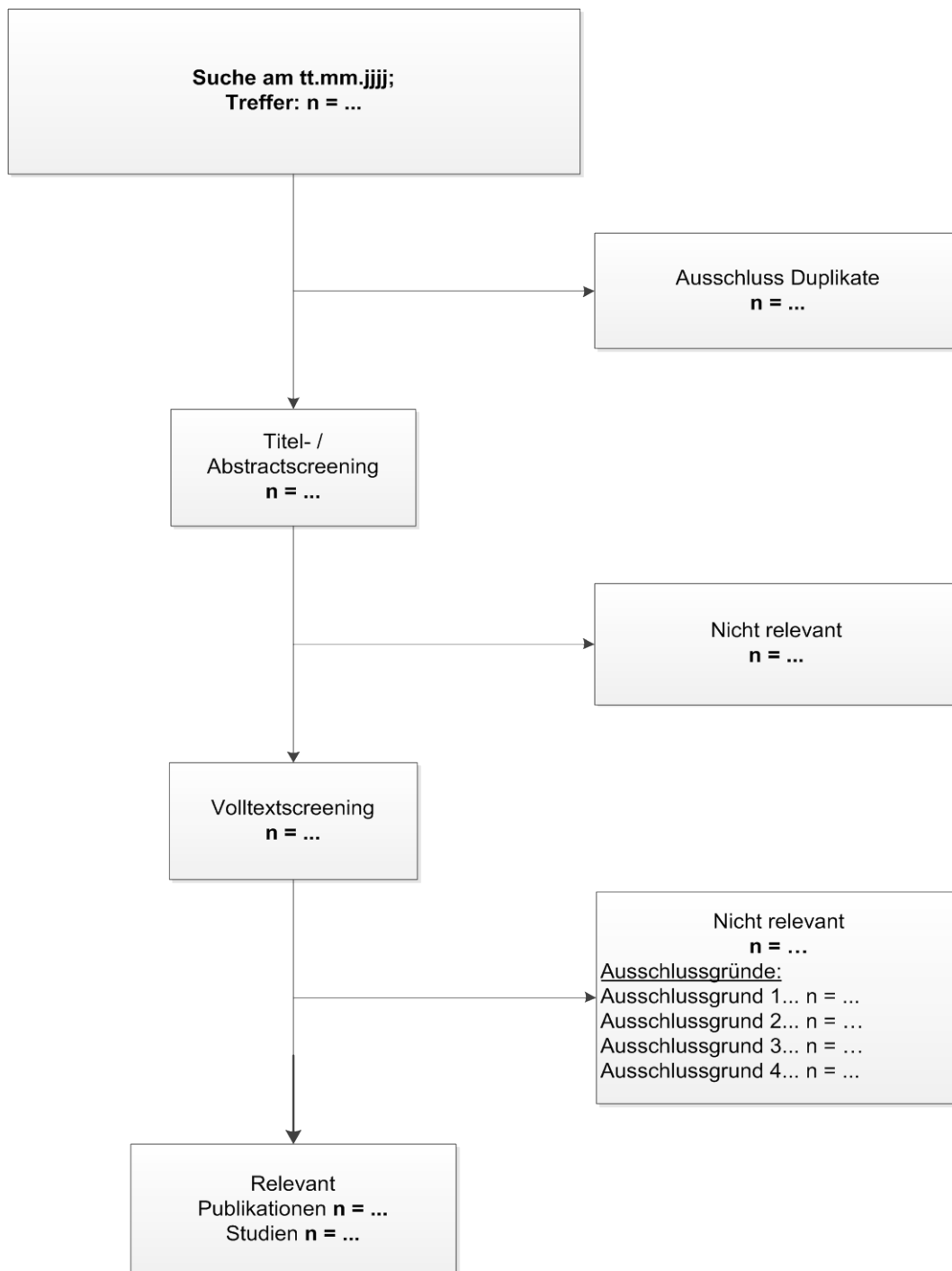


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Nicht zutreffend.

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-1) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[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 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate ^a)	Studie in Liste der Studien des pharma- zeutischen Unter- nehmers enthalten (ja/nein)	Studie durch biblio- grafische Literatur- recherche identifi- ziert (ja/nein)	Status (abge- schlossen/ abge- brochen/ laufend)
-	-	-	-	-
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-3 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Nicht zutreffend.

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 (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

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

jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Nicht zutreffend

Die Zulassungsstudie GADOLIN ist die für die Nutzenbewertung von Obinutuzumab mit Orphan-Status relevante Studie. Die folgende Tabelle wurde für eine bessere Lesbarkeit ausgefüllt.

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

Studie	Studienkategorie			verfügbare Quellen ^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						
aktivkontrolliert Vergleichstherapie(n)						
GADOLIN	ja	ja	nein	ja (7, 12)	ja [NCT01059630] [2009-015504-25] (13-16)	Ja [Abstract (17)]
a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden. b: Studie, für die der Unternehmer Sponsor war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

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-E zu hinterlegen.

Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studien- design	Population	Interventionen (Zahl der rando- misierten Patien- ten)	Studiendauer ^a	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patien- tenrelevante sekundäre Endpunkte
GADOLIN	RCT, offen, parallel	Erwachsene Patienten mit Rituximab- refraktärem indolenten Non-Hodgkin- Lymphom (iNHL)	<u>Gesamte Studien- population:</u> Datenschnitt 01. September 2014 G-Benda (N=194) Benda (N=202) Datenschnitt 01. Mai 2015 G-Benda (N=204) Benda (N=209) <u>Patienten mit FL:</u> Datenschnitt 01. September 2014 G-Benda (N=155) Benda (N=166) Datenschnitt 01. Mai 2015 G-Benda (N=164) Benda (N=171)	Screening-Phase: max. 28 Tage zur Überprüfung der Patienteneignung Behandlungsphase: max. 6 Behandlungs-Zyklen à 28 Tage Follow-Up Phase: bis Eintreten einer Progression/Rezidiv bzw. bis 25 Monate nach Ende der Induktionstherapie. Nachbeobachtung: bis Eintreten einer Progression/Rezidiv, Rücknahme der Einverständnis bzw. bis 24 Monate nach Ende der Follow-Up Phase. Eine Nachbeobachtung auf Überle- ben erfolgt bis max. 4 Jahre nach Einschluss des letz- ten Patienten. Nachbeobachtung auf Verträglichkeit Unerwünschte Ereignisse (UE) bis 28 Tage nach Behandlung UE Grad 3 und 4 bis 6 Monate nach Behandlung UE Infektionen Grad 3 und 4 bis 12 Monate nach Behandlung Ab Protokolländerung 8 (Oktober 2013) Angleichung der UE-Beobachtungsdauer (7): UE bis 28 Tage nach Follow-Up Phase UE Grad 3 und 4 bis 6 Monate nach Follow-Up Phase UE Infektionen Grad 3 und 4, die in Verbindung mit der Studienmedikation stehen, bis 24 Monate nach Follow-Up Phase	Ort: 82 Studienzentren in 14 Ländern: Europa (Belgien, Deutschland, Frank- reich, Italien, Nieder- lande, Österreich, Russland, Schweden, Schweiz, Spanien, Tschechische Repub- lik, Vereinigtes Kö- nigreich), Kanada, USA. Zeitraum: Randomisierung: April 2010 bis Januar 2015	Primärer Endpunkt: · Progressionsfreies Überle- ben (PFS) bewertet durch IRC Sekundäre Endpunkte: · PFS bewertet durch Prüfarzt, · Gesamtüberleben (OS), · Krankheitsfreies Überleben (DFS), · Verträglichkeit, · Gesundheitsbezogene Lebensqualität (FACT-Lym)

Benda: Bendamustin; DFS: Krankheitsfreies Überleben; FACT-Lym: Functional Assessment of Cancer Therapy – Lymphoma; G-Benda: G-Benda Induktionstherapie
gefolgt von einer Obinutuzumab Erhaltungstherapie; iNHL: indolentes Non-Hodgkin-Lymphom; IRC: Independent Review Committee; N: Anzahl der Patienten; OS:
Gesamtüberleben; PFS: progressionsfreies Überleben; RCT: Randomized Controlled Trial; UE: unerwünschte Ereignisse;
Quelle: CSR (7)

Tabelle 4-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	G-Benda	Benda	Vorbehandlung / Begleitbehandlung
GADOLIN	<u>Induktionstherapie:</u> Obinutuzumab: max. 6 Zyklen à 28 Tage Zyklus 1: Insgesamt 3.000 mg Jeweils 1.000 mg i.v., Tag 1, 8 und 15 Folgezyklen (2-6): 1.000 mg i.v. am Tag 1 <u>Bendamustin:</u> max. 6 Zyklen à 28 Tage Zyklus 1-6: 90 mg/m ² (KOF) i.v., Tag 1 und 2 <u>Erhaltungstherapie:</u> Obinutuzumab 1.000 mg i.v. alle 2 Monate für maximal 2 Jahre oder bis zur Progression	Bendamustin-Gabe: max. 6 Zyklen à 28 Tage Zyklus 1-6: 120 mg/m ² (KOF) i.v., Tag 1 und 2	Prä-Medikation zur Antikörpergabe: Acetaminophen/Paracetamol: (oral, 650-1000 mg) etwa 30 Minuten vor Beginn einer Obinutuzumab-Infusion Antihistaminikum: (z.B. Diphenhydramin 50-100 mg) etwa 30 Minuten vor Beginn einer Obinutuzumab-Infusion Kortikosteroide: vor der ersten Obinutuzumab- Infusion wird empfohlen prophylaktisch Korti- kosteroide (z.B. 100 mg i.v. Prednisolon oder Äquivalent) einzusetzen. Nach Ermessen des Prüfarztes können prophy- laktische Kortikosteroide auch vor weiteren Infusionen zur Verhinderung infusionsbeding- ter Reaktionen verabreicht werden. Sollten bei Patienten keine infusionsbedingten Reaktionen ≥ Grad 2 auftreten, kann nach Er- messen des Prüfarztes die Vorbehandlung bei den folgenden Infusionen weggelassen werden.
Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab-Erhaltungsthera- pie; i.v.: intravenös; KOF: Körperoberfläche; Wdh.: Wiederholung Quelle: CSR (7)			

Tabelle 4-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel – Patienten mit FL (Datenschnitt 01. September 2014 und 01. Mai 2015)

GADOLIN, Patienten mit FL Populationscharakteristika	Datenschnitt 01. September 2014		Datenschnitt 01. Mai 2015	
	G-Benda N = 155	Benda N = 166	G-Benda N = 164	Benda N = 171
Alter [Jahre] Mittelwert	61,8	62,5	61,8	62,4
Geschlecht [n (%)] männlich weiblich	85 (54,8) 70 (45,2)	95 (57,2) 71 (42,8)	91 (55,5) 73 (44,5)	98 (57,3) 73 (42,7)
Ethnische Herkunft [n (%)] kaukasisch andere unbekannt	139 (89,7) 7 (4,5) 9 (5,8)	143 (86,1) 8 (4,9) 15 (9,0)	144 (87,8) 8 (4,9) 12 (7,3)	148 (86,5) 8 (4,7) 15 (8,8)
Region [n (%)] West-Europa andere	72 (46,5) 83 (53,5)	78 (47,0) 88 (53,0)	78 (47,6) 86 (52,4)	79 (46,2) 92 (53,8)
Krankheitsstadium (lt. Ann Arbor)[n (%)] I II III IV Unbekannt Missing	9 (5,8) 15 (9,7) 31 (20,0) 90 (58,1) 10 (6,4) 0 (0,0)	9 (5,4) 19 (11,5) 45 (27,1) 82 (49,4) 10 (6,0) 1 (0,6)	9 (5,5) 16 (9,8) 33 (20,1) 96 (58,5) 10 (6,1) 0 (0,0)	9 (5,3) 20 (11,7) 45 (26,3) 86 (50,3) 10 (5,8) 1 (0,6)

GADOLIN, Patienten mit FL Populationscharakteristika	Datenschnitt 01. September 2014		Datenschnitt 01. Mai 2015	
	G-Benda N = 155	Benda N = 166	G-Benda N = 164	Benda N = 171
FLIPI [Kategorie (Anzahl Risikofaktoren)] [n (%)]				
Niedrig (0,1)	42 (27,1)	34 (20,5)	42 (25,6)	35 (20,5)
Intermediär (2)	47 (30,3)	58 (34,9)	51 (31,1)	60 (35,1)
Hoch (≥ 3)	60 (38,7)	67 (40,4)	64 (39,0)	69 (40,3)
Unbekannt	6 (3,9)	6 (3,6)	7 (4,3)	6 (3,5)
Missing	0 (0,0)	1 (0,6)	0 (0,0)	1 (0,6)
Anzahl Vorthérapien [n (%)]				
≤ 2	125 (80,6)	130 (78,3)	134 (81,7)	133 (77,8)
> 2	30 (19,4)	36 (21,7)	30 (18,3)	38 (22,2)
Refraktäre Patienten nach [n (%)]				
Rituximab-Monotherapie	25 (16,1)	39 (23,5)	25 (15,2)	42 (24,6)
Rituximab-Chemotherapie	130 (83,9)	127 (76,5)	139 (84,4)	129 (75,4)
Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab-Erhaltungstherapie; FL: Follikuläres Lymphom; FLIPI: FL International Prognostic Index; N: Anzahl der Patienten in Behandlungsarm; n: Anzahl der Patienten mit Ereignis Quelle: CSR, CSR update, Roche Nachauswertung (7, 12, 18)				

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

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

Design der GADOLIN-Studie

GADOLIN ist eine offene, randomisierte, kontrollierte, multinationale Studie, die in 82 Studienzentren in 14 Ländern durchgeführt wird. Die Therapie bestehend aus einer Induktionstherapie mit der Kombination Obinutuzumab + Bendamustin, gefolgt von einer Obinutuzumab-Erhaltungstherapie (G-Benda) und wird mit dem aktuellen Therapiestandard für diese Patientenpopulation – einer Monotherapie mit Bendamustin – verglichen (siehe Abbildung 2).

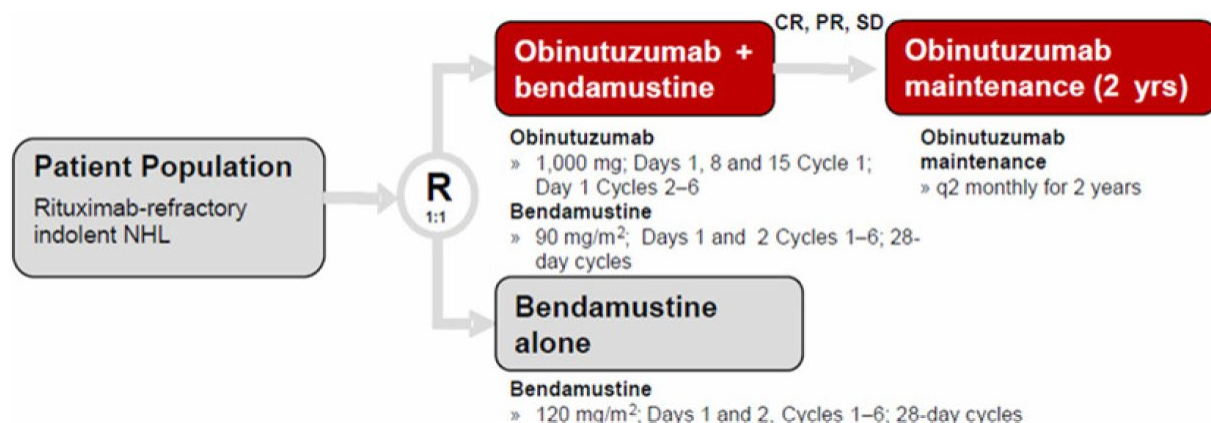


Abbildung 2: Studiendesign (7)

Studienziel: Primäres Ziel ist der Nachweis einer Überlegenheit von G-Benda gegenüber der derzeitigen Standardtherapie – Monotherapie mit Bendamustin – bei erwachsenen Patienten mit indolentem Non-Hodgkin-Lymphom (iNHL), die refraktär gegenüber einer Rituximab-haltigen Vortherapie sind.

Endpunkte: Der primäre Endpunkt der Studie ist das progressionsfreie Überleben (PFS), bewertet durch ein IRC, definiert als Zeitintervall von der Randomisierung bis zum Auftreten eines PFS-Ereignisses (Progression, Rezidiv oder Tod).

Die sekundären Endpunkte der Studie sind: PFS, bewertet durch Prüfarzt (Investigator); Gesamtüberleben (OS); Krankheitsfreies Überleben (DFS); Ereignisfreies Überleben (EFS); Gesamtansprechrate (ORR = overall response, gemessen als vollständiges Ansprechen (CR) + partielles Ansprechen (PR = partial response)); Dauer des Ansprechens (DoR = duration of response); Verträglichkeit und Gesundheitsbezogene Lebensqualität (FACT-G, FACT-Lym, FACT-TOI und EQ-5D).

Statistische Analysen von Time-to-event Endpunkten: Die Analyse des primären Endpunktes PFS (IRC) wurde unter Verwendung des zweiseitigen stratifizierten log-rank-Test zum Signifikanzniveau von 5 % durchgeführt. Die Hazard Ratios (HR) und zugehörige Konfidenzintervalle (KI) zur Quantifizierung des Unterschieds zwischen den Behandlungsgruppen wurden mit einem Cox-Modell stratifiziert nach Randomisierungsvariablen (indolente NHL, Art der Refraktärität, Anzahl Vortherapien) ermittelt. Die geographische Region wurde in der Analyse aufgrund fehlenden prognostischen Wertes nicht als Variable berücksichtigt. Die Überlebenszeitkurven und die daraus abgeleitete mediane Zeit bis zum Eintritt des Ereignisses sowie Überlebensraten zu definierten Zeitpunkten mit zugehörigen 95 % KI wurden mit der Kaplan-Meier-Methode ermittelt.

Ein- und Ausschlusskriterien: In die Studie wurden Patienten mit histologisch gesicherter Diagnose eines indolenten Non-Hodgkin-Lymphoms eingeschlossen, die auf Rituximab oder eine Rituximab-haltige Therapie kein Ansprechen oder Progression während oder bis zu 6 Monate nach der Behandlung zeigten.

Studiendauer: Der erste Patient wurde am 15. April 2010 randomisiert. Der letzte Patient wurde am 07. Januar 2015 in die Studie eingeschlossen. GADOLIN ist noch nicht beendet. Die finale Analyse bzgl. Gesamtüberleben und Verträglichkeit erfolgt am Ende der Studie gemäß Studienprotokoll nach 226 Todesfällen.

Datenschnitte:

Hintergrund

Studien mit geplanten sowie von Zulassungsbehörden zusätzlich geforderten Zwischenauswertungen, final bzw. konfirmatorischen Analysen und Follow-Up Datenschnitten stellen für die Bewertung der Studienergebnisse eine Herausforderung dar. Die unterschiedlichen Analysen (im Folgenden auch Datenschnitte genannt) sind für die Interpretation der Ergebnisse

meist nicht als gleichwertig zu betrachten und können je Endpunkt zu unterschiedlichen Zeitpunkten stattfinden.

- Final bzw. konfirmatorische Analysen
- Zwischenauswertungen (geplant und/oder ungeplant)
- Follow-Up Datenschnitte

Final bzw. konfirmatorische Analysen

Diese Analysen sind präspezifiziert und bilden die Grundlage der Studienplanung. Sie haben damit die höchste Aussagekraft. Die Studie ist geplant, um für definierte Endpunkte zu diesen Datenschnitten eine final bzw. konfirmatorische Aussage zu ermöglichen. Alle weiteren Analysen haben nicht die gleiche Wertigkeit. Finale bzw. konfirmatorische Analysen können zu mehreren Zeitpunkten durchgeführt werden, dann aber für unterschiedliche Endpunkte. Insbesondere bei Time-to-event Endpunkten kann es sein, dass die notwendige Anzahl von Ereignissen nicht zeitgleich für alle Endpunkte erreicht wird. Dies gilt z.B. dann, wenn der eine Endpunkt vom Verlauf der Erkrankung vor dem anderen liegt (Progress vor Tod).

Zwischenauswertungen

Geplant:

Mit einer adäquaten biometrischen Planung einer Zwischenauswertung sind die formalen statistischen Kriterien zur vorzeitigen Beendigung einer klinischen Studie erfüllt. Eine weitere Randbedingung besteht darin, dass medizinische Aspekte bei der Wahl des Zeitpunktes adäquat berücksichtigt wurden.

Ungeplant:

Nicht im Protokoll präspezifizierte Analysen sind möglichst zu vermeiden. Eine Ausnahme bilden Analysen, die durch behördliche Anforderung durchgeführt werden. Diese Analysen haben keinen konfirmatorischen Charakter und stellen keine Prüfung der medizinischen Fragestellung mit statistischen Mitteln dar. Sie haben vielmehr orientierenden Charakter. Mittels dieser Analysen sollen bestehende Trends im Rahmen des Zulassungsverfahrens überprüft werden, um die Unsicherheit im Hinblick auf die final bzw. konfirmatorische Analyse zu reduzieren.

Follow-Up Datenschnitte

Nach der final bzw. konfirmatorischen Analyse sind die weiteren späteren Datenschnitte aus statistischer Sicht als explorative ‚Follow-Up‘-Analysen zu betrachten, die keine inferenzstatistische Bedeutung haben. Dies gilt insbesondere in Fällen, in denen die Studien verblindete RCT sind, da die vorzeitige Beendigung der Studie zumeist auch Änderungen am Studiendesign mit sich bringt. Eine Ausnahme stellen die Datenschnitte zur Verträglichkeit dar, da hier die Akkumulation des Wissens, insbesondere auch bei seltenen Ereignissen große Bedeutung hat.

Datenschnitte in GADOLIN

Für GADOLIN ergibt sich folgende Situation (siehe Tabelle 4-8): Es waren insgesamt drei Zwischenauswertungen, wovon die dritte für die Endpunkte PFS, DFS/Rezidivrate und Gesundheitsbezogene Lebensqualität die final bzw. konfirmatorische wurde, geplant (7, 12).

Die erste geplante Zwischenauswertung untersuchte ausschließlich die Verträglichkeit und wurde zum Zeitpunkt, als 20 Patienten rekrutiert waren und mindestens einen Zyklus vollendet hatten, durchgeführt (18. April 2011).

Die zweite geplante Zwischenauswertung wurde durchgeführt, als 100 Prüfarzt berichtete PFS Ereignisse dokumentiert waren (29. Juli 2013). Zu dem Zeitpunkt der zweite Zwischenbewertung hatte die Studie das Potential aus Sicherheitsgründen oder aus Gründen der Sinnhaftigkeit basierend auf IRC bewertetes PFS frühzeitig zu stoppen.

Die dritte geplante Zwischenauswertung wurde durchgeführt, als 170 IRC bewertete PFS Ereignisse dokumentiert waren (01. September 2014). Da der primäre Endpunkt PFS zum Zeitpunkt dieser Zwischenauswertung die präspezifizierte Signifikanzgrenze für den p-Wert unterschritt, empfahl das IDMC die Studie komplett auszuwerten. Die dritte Zwischenauswertung mit Datenschnitt am 01. September 2014 ist damit als die primäre Analyse mit final bzw. konfirmatorischer Wertigkeit für alle Endpunkte (PFS, DFS/Rezidivrate, Gesundheitsbezogene Lebensqualität), nicht jedoch für Gesamtüberleben und Verträglichkeit, anzusehen.

Die finale Analyse bzgl. Gesamtüberleben und Verträglichkeit erfolgt am Ende der Studie gemäß Studienprotokoll nach 226 Todesfällen.

Eine ungeplante Zwischenauswertung wurde aufgrund behördlicher Anforderungen (gefordert für Tag-90 innerhalb des FDA Zulassungsprozesses) am 01. Mai 2015 durchgeführt. Sie ist damit ein Follow-Up Datenschnitt für die bereits zum dritten Datenschnitt als final bzw. konfirmatorisch ausgewerteten Endpunkte PFS, DFS/Rezidivrate, Gesundheitsbezogene Lebensqualität. Für Gesamtüberleben und Verträglichkeit hat dieser Datenschnitt jedoch aufgrund der Akkumulation des Wissens im Sinne der bestverfügbaren Evidenz, eine besondere Bedeutung und wird im Nutzendossier für diese Endpunkte primär berücksichtigt.

Tabelle 4-8: Datenschnitte und Wertigkeit der Analysen in GADOLIN

Auswertungen	Anzahl Ereignisse (präspezifiziert)		Datenschnitt		Wertigkeit der Wirksamkeitsanalysen			
	Inv/ IRC	OS			PFS IRC	DFS/ Rezidiv-rate	Lebens-qualität	Gesamt-über-leben
Präspezifizierte ZA Verträglichkeit (bei 20 randomisierten Patienten)	n.a.		1	18. Apr 2011				
Präspezifizierte ZA Futility PFS IRC oder Verträglichkeit	100 (Inv)		2	29. Jul 2013	ZA			
Präspezifizierte ZA Wirksamkeit und Futility à Primärer Endpunkt unterschritt Signifikanzgrenze	170 (IRC)		3	01. Sep 2014	final bzw. konfirmatorisch	final bzw. konfirmatorisch	final bzw. konfirmatorisch	ZA
ungeplante ZA von der Behörde gefordert (90 Tage FDA)			4	01. Mai 2015	FU	FU	FU	ZA*
In der Zukunft anstehende Auswertungen:								
Präspezifizierte finale Auswertung, Studienende Wirksamkeit à OS und Verträglichkeit		226			FU	FU	FU	final bzw. konfirmatorisch
FU: Follow-Up, ZA: Zwischenauswertung, Die präspezifizierte finale Auswertung für die Wirksamkeitsendpunkte zum Zeitpunkt, an dem 260 IRC PFS Ereignisse dokumentiert sind, entfällt, da zum Zeitpunkt der dritten Zwischenauswertung die präspezifizierte Signifikanzgrenze für den p-Wert von PFS IRC unterschritten wurde und die dritte Zwischenauswertung damit als die final bzw. konfirmatorische Analyse anzusehen ist. *78 Ereignisse (Todesfälle) zum Zeitpunkt des Datenschnitts								

Fazit: Der dritte Datenschnitt vom 01. September 2014 ist der final bzw. konfirmatorische für PFS, DFS/Rezidivrate sowie Gesundheitsbezogene Lebensqualität und ist damit für diese Endpunkte der entscheidende für die Erstellung des Nutzendossiers. Für Gesamtüberleben und Verträglichkeit hat der vierte Datenschnitt vom 01. Mai 2015 ungeachtet dessen, dass dies keine präspezifizierte Analyse war, aufgrund der Akkumulation des Wissens im Sinne der bestverfügbaren Evidenz eine besondere Bedeutung und wird daher im Nutzendossier primär für diese Endpunkte berücksichtigt. Die finale konfirmatorische Analyse für das Gesamtüberleben steht noch aus.

Randomisierung: Die Patienten wurden im Verhältnis 1:1 den zwei Studienarmen zugeteilt, stratifiziert nach den Faktoren *indolentes NHL* (Subtyp follikuläre Lymphome oder sonstige), *Art der Refraktärität* (Rituximab-Monotherapie vs. Rituximab + Chemotherapie), *Anzahl der Vorthapien* (≤ 2 oder > 2) und *geographische Region* (Osteuropa und Naher Osten / Westeuropa / Nord- und Zentral-Amerika und Karibik).

Studienpopulation: Zum Zeitpunkt des Datenschnitts vom 01. September 2014 wurden insgesamt 396 Patienten in zwei Studienarme randomisiert (202 Patienten zu Benda und 194 Patienten zu G-Benda). Mit Datenschnitt vom 01. Mai 2015 waren 209 Patienten im Benda- und 204 Patienten im G-Benda-Arm. Die Verteilung nach Stratifikationsmerkmal Subtyp iNHL ist für beide vorliegenden Datenschnitte der Tabelle 4-9 zu entnehmen.

Die im vorliegenden Nutzendossier betrachtete Kombinationstherapie aus Obinutuzumab und Bendamustin ist für die Indikation FL zugelassen (5). Entsprechend wird nachfolgend ausschließlich die zulassungskonforme Studienpopulation bestehend aus Patienten mit der Diagnose FL betrachtet. Nachfolgende Angaben bezüglich ITT-Population beziehen sich auf diese FL-Population (ITT FL).

Die demographischen und krankheitsbezogenen Charakteristika in beiden Studienarmen sind vergleichbar (siehe Tabelle 4-7).

Tabelle 4-9: Anteil Patienten mit follikulärem Lymphom in GADOLIN

Patientenpopulation in GADOLIN	G-Benda (n (%))	Benda (n (%))
Datenschnitt 01. September 2014		
Randomisiert	194 (100 %)	202 (100 %)
Follikuläres Lymphom	155 (79,9 %)	166 (82,2 %)
kein follikuläres Lymphom	39 (20,1 %)	36 (17,8 %)
Datenschnitt 01. Mai 2015		
Randomisiert	204 (100 %)	209 (100 %)
Follikuläres Lymphom	164 (80,4 %)	171 (81,8 %)
kein follikuläres Lymphom	40 (19,6 %)	38 (18,2 %)
n: Anzahl der Patienten Quelle: CSR und CSR update (7, 12)		

Die Studie GADOLIN wurde mit dem dritten Datenschnitt vom 01. September 2014, der als der final bzw. konfirmatorische Datenschnitt für alle Endpunkte der Morbidität und Gesundheitsbezogene Lebensqualität gilt, vorzeitig beendet. In der folgenden Tabelle 4-10 sind die Zahlen der Patienten abgebildet, welche die verschiedenen Beobachtungszeitpunkte bis zum Zeitpunkt der relevanten Datenschnitte vom 01. September 2014 sowie 01. Mai 2015 erreichten.

Tabelle 4-10: Anteil FL-Patienten in GADOLIN zu den Beobachtungszeitpunkten

GADOLIN, Patienten mit FL zum Zeitpunkt	Datenschnitt 01. September 2014		Datenschnitt 01. Mai 2015	
	G-Benda N = 155	Benda N = 166	G-Benda N = 164	Benda N = 171
Gesamtbeobachtungsdauer in Jahren	278,3	280,7	352,6	347,1
Mittlere Beobachtungsdauer in Monaten [min; max]	21,6 [0,4;48,5]	20,3 [0,0;50,0]	25,8 [0,4;55,2]	24,4 [0,0;58,0]
> 0 Monate	155 (100 %)	166 (100 %)	164 (100 %)	171 (100 %)
> 6 Monate	130 (83,9 %)	140 (84,3 %)	143 (87,2 %)	156 (91,2 %)
> 9 Monate	125 (80,6 %)	128 (77,1 %)	135 (82,3 %)	140 (81,9 %)
> 1,0 Jahr	113 (72,9 %)	114 (68,7 %)	130 (79,3 %)	125 (73,1 %)
> 1,5 Jahre	94 (60,6 %)	92 (55,4 %)	113 (68,9 %)	107 (62,6 %)
> 2,0 Jahre	63 (40,6 %)	64 (38,6 %)	83 (50,6 %)	86 (50,3 %)
> 2,5 Jahre	42 (27,1 %)	38 (22,9 %)	67 (40,9 %)	60 (35,1 %)
> 3,0 Jahre	19 (12,3 %)	22 (13,3 %)	44 (26,8 %)	39 (22,8 %)
> 3,5 Jahre	9 (5,8 %)	8 (4,8 %)	25 (15,2 %)	20 (11,7 %)
> 4,0 Jahre	1 (0,6 %)	1 (0,6 %)	14 (8,5 %)	9 (5,3 %)
> 4,5 Jahre	0 (0,0 %)	0 (0,0 %)	3 (1,8 %)	2 (1,2 %)
> 5,0 Jahre	0 (0,0 %)	0 (0,0 %)	0 (0,0 %)	0 (0,0 %)
Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab-Erhaltungstherapie; FL: Follikuläres Lymphom; N: Anzahl der Patienten in Behandlungsarm Quelle: CSR update, Nachauswertung Roche (12, 18)				

Interventionen

Benda (Monotherapie)

Die Patienten in diesem Studienarm erhalten Bendamustin in einer Dosierung von 120 mg/m² Körperoberfläche (KOF) als intravenöse Infusion über 60 Minuten an den Tagen 1 und 2 eines 28-tägigen Zyklus. Die Behandlung umfasst maximal 6 Zyklen.

G-Benda

Induktionstherapie (Kombinationstherapie):

Im ersten Behandlungszyklus werden 1000 mg Obinutuzumab als intravenöse Infusion an den Tagen 1, 8 und 15 verabreicht. In den Folgezyklen 2 bis 6 erfolgt die intravenöse Infusion von 1000 mg Obinutuzumab jeweils am Tag 1.

Kombiniert wird Obinutuzumab mit Bendamustin in einer Dosierung von 90 mg/m² KOF als intravenöse Infusion über 60 Minuten an den Tagen 1 und 2 eines 28-tägigen Zyklus. Die Behandlung umfasst maximal 6 Zyklen.

Erhaltungstherapie (Monotherapie):

Nach Beendigung des 6. Zyklus erfolgt für Patienten, die auf die Induktionstherapie angesprochen haben, eine Obinutuzumab-Erhaltungstherapie (ohne Bendamustin) alle 2 Monate bis zum Progress, jedoch für maximal 2 Jahre.

Mit der Protokolländerung 10 vom 07. Mai 2015 wurde aufgrund der Überlegenheit der Kombinationstherapie ein Cross-over von Patienten aus dem Benda-Arm in den G-Benda-Arm ermöglicht. Dies betrifft Patienten, die sich zu diesem Zeitpunkt noch in der Induktionstherapie befinden oder diese weniger als 3 Monate ($\pm$ vier Wochen) zuvor beendet haben.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
GADOLIN	ja	ja	nein	nein	ja	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei GADOLIN handelt es sich um eine offene randomisierte, multinationale klinische Studie. Die Zuteilung eines Patienten zu einer Behandlungsgruppe erfolgte nach einer Überprüfung der Ein- und Ausschlusskriterien durch ein zentrales IVRS (Interactive Voice Response System).

Es erfolgte eine stratifizierte Randomisierung mit den folgenden Stratifizierungsfaktoren: *indolentes NHL* (Subtyp follikuläre Lymphome oder sonstige), *Art der Refraktärität* (Rituximab-Monotherapie vs. Rituximab + Chemotherapie), *Anzahl der Vortherapien* (≤ 2 oder > 2) und *geographische Region* (Osteuropa und Naher Osten / Westeuropa / Nord- und Zentralamerika und Karibik).

Durch die zentrale Vergabe der Behandlungsgruppe blieb die zugewiesene Behandlungsgruppe dem Prüfarzt bis zum Zeitpunkt der tatsächlichen Randomisierung unbekannt.

Eine Verblindung des Prüfarztes oder der Patienten wurde aufgrund der erheblichen Belastung für den Patienten durch weitere Infusionen (mehrstündige Infusion des Placebo-Antikörpers und der Placebo-Prä-Medikation) nicht durchgeführt. Zudem ist aufgrund des bekannten Nebenwirkungsprofils von sowohl Verum als auch der Prä-Medikation (Korti-

kosteroid, Antihistaminikum) von einer Selbstentblindung des Prüfarztes auszugehen. Zur Beurteilung der Tumorprogression wurde neben der unverblindeten Erhebung durch den Prüfarzt eine verblindete Beurteilung durch ein unabhängiges Bewertungsgremium etabliert.

Durch die fehlende Verblindung besteht grundsätzlich ein Verzerrungspotential, das je Endpunkt unterschiedlich ist und sich deshalb für die dargestellten Ergebnisse dort niederschlägt. Auch wird die fehlende Verblindung durch weitere Maßnahmen wie z.B. der verblindeten Bewertung des progressionsfreien Überlebens durch das IRC oder Sensitivitätsanalysen auf Endpunktebene teilweise kompensiert. Das Verzerrungspotential ist aus diesem Grund endpunktspezifisch zu betrachten.

Die Berichterstattung der Studiendaten im Studienbericht folgte der im Studienprotokoll und Analyseplan a priori festgelegten Auswertungsstrategie. Andere Verzerrungsquellen auf Studienebene konnten nicht identifiziert werden. Die Patienten in den beiden betrachteten Studienarmen zeigten keine Unterschiede hinsichtlich ihrer Baseline-Charakteristika und Beobachtungsdauern.

Die Bewertung des Zusatznutzens im vorliegenden Dossier basiert auf dem Vergleich einer Teilpopulation der GADOLIN. Basis sind die Patienten mit FL. Diese Beschränkung auf einen Teil der Studienpopulation stellt keine Verletzung des ITT-Prinzips dar, da durch den Stratifizierungsfaktor „Subtyp des indolenten Non-Hodgkin-Lymphoms“ die Gleichverteilung der FL-Patienten auf die Behandlungsarme sichergestellt war. Gleichzeitig wird die Einschränkung der Studienpopulation nicht aus studienimmanenten Gründen vorgenommen, sondern aufgrund einer Indikationsstellung der Zulassungsbehörde. Ein Einfluss auf das Verzerrungspotential wird daher ausgeschlossen.

Von einer relevanten Verzerrung aller Studienergebnisse ist somit nicht auszugehen. Aus diesem Grund wird das Verzerrungspotential der GADOLIN als *niedrig* eingestuft.

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 4-12: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Gesamt- überleben	Progressionsfreies Überleben (IRC)	Krankheitsfreies Überleben	Gesundheitsbezogene Lebens- qualität FACT-Lym	Verträglichkeit
GADOLIN	ja	ja	ja	ja	ja

4.3.1.3.1 Endpunkte für die Patientenpopulation mit follikulärem Lymphom – RCT

4.3.1.3.1.1 Gesamtüberleben – RCT

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

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

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen 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). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

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

Tabelle 4-13: Operationalisierung von Gesamtüberleben

Studie	Operationalisierung
GADOLIN	Der Endpunkt Gesamtüberleben (OS) ist definiert als Zeit (in Monaten) zwischen dem Tag der Randomisierung bis zum Tod (unabhängig von der Todesursache). <u>Erhebung:</u> Patienten, welche die Studienmedikation ohne vorherigen Progress abbrechen (z.B. aufgrund von Toxizität), verbleiben in der Studie und werden bezüglich OS in der Nachbeobachtungsphase beobachtet, ungeachtet dessen, ob eine neue Therapie begonnen wird. Patienten mit Progression/Rezidiv gehen direkt in die Nachbeobachtungsphase über. Die Patienten werden in der Follow-Up Phase alle zwei Monate, danach in der Nachbeobachtungsphase zweimal jährlich für mindestens 4 Jahre nach Einschluss des letzten Patienten in die Studie (bis zum Eintreten von 226 Todesfällen) oder bis zum Tod des letzten Patienten in der Studie bezüglich des Versterbens nachbeobachtet. <u>Zensurierung:</u> Patienten, für die das Ereignis Tod nicht dokumentiert wurde, werden zum Zeitpunkt der Analyse mit dem letzten bekannten Datum, an dem sie nachweislich noch lebten, zensiert. Für die Auswertung werden die Datenschnitte vom 01. September 2014 sowie 01. Mai 2015 dargestellt. Die Analysepopulation ist die ITT FL.

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

Tabelle 4-14: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
GADOLIN	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene liegt ein niedriges Verzerrungspotential vor.

Die Erhebung des Gesamtüberlebens erfolgte mit Kenntnis der Studientherapie.

Die Analyse des Gesamtüberlebens für die FL-Population folgte dem ITT-Prinzip (Subtyp des indolenten Non-Hodgkin-Lymphoms war Stratifizierungsfaktor) mit einer adäquaten Zensierung der noch lebenden Patienten zum letzten Beobachtungszeitpunkt.

Die im Studienbericht dargestellten Ergebnisse folgten den im Analyseplan spezifizierten Analysen. Ein Wechsel zwischen den beiden Therapiearmen G-Benda und Benda innerhalb der Studie war zum Zeitpunkt beider Datenschnitte nicht erlaubt. Mit Protokolländerung Nummer 10 war der Wechsel für Patienten des Benda-Arms in den G-Benda-Arm erlaubt. Der zweite Datenschnitt erfolgte vor Protokolländerung 10 und somit bestehen keine verzerrungsrelevanten Aspekte.

Sonstige verzerrungsrelevante Aspekte konnten nicht identifiziert werden.

Insgesamt wird deshalb das Verzerrungspotential für den Endpunkt Gesamtüberleben als *niedrig* bewertet.

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 4-15: Ergebnisse für das Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel

Gesamtüberleben					
Datenschnitt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Datenschnitt (01. Sep 2014)	25 / 155 (16,1)	36 / 166 (21,7)	n.e.	n.e. [39,8; n.e.]	0,71 [0,43; 1,19] 0,1976
Datenschnitt (01. Mai 2015)	30 / 164 (18,3)	48 / 171 (28,1)	n.e.	n.e. [42,2; n.e.]	0,62 [0,39; 0,98] 0,0379
Population: ITT FL; (%): Anteil der Patienten mit Ereignis in %; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.e. = nicht erreicht Statistik: Median basierend auf Kaplan-Meier Methode; 95 % KI: Brookmeyer und Crowley Methode; p-Wert: stratifizierter log-rank Test; Stratifizierungsfaktoren: Anzahl vorheriger Therapien, Art des Therapieversagens Quelle: CSR und CSR update (7, 12)					

Im Datenschnitt vom 01. September 2014 war eine Tendenz zugunsten von G-Benda gegenüber Benda bezüglich des Endpunkts OS (HR: 0,71 [0,43; 1,19]; p=0,1976) nachweisbar. Im Datenschnitt vom 01. Mai 2015 bestand ein statistisch signifikanter Unterschied zugunsten von G-Benda gegenüber Benda (HR: 0,62 [0,39; 0,98]; p=0,0379) (siehe Tabelle 4-15).

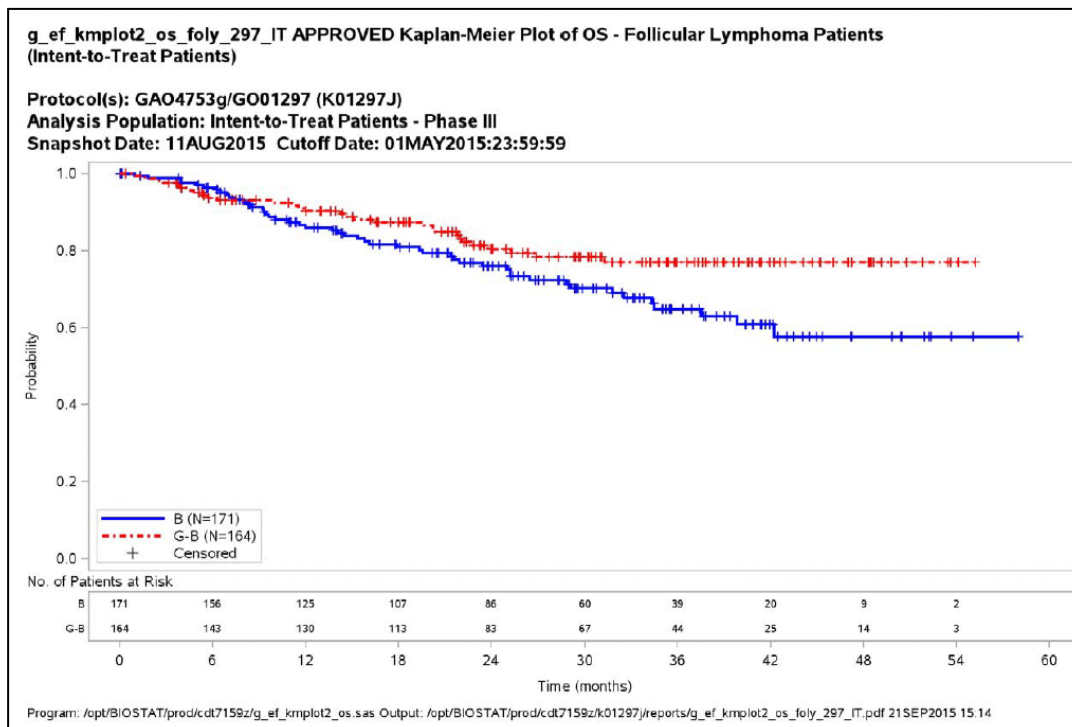


Abbildung 3: Kaplan-Meier-Kurve zum Endpunkt OS (Datenschnitt 01. Mai 2015) (12)

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht zutreffend. Dies gilt für alle in der vorliegenden Nutzenbewertung dargestellten Endpunkte.

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

4.3.1.3.1.2 Progressionsfreies Überleben (IRC) – RCT

Tabelle 4-16: Operationalisierung von Progressionsfreies Überleben

Studie	Operationalisierung
GADOLIN	Progressionsfreies Überleben (PFS), beurteilt durch ein unabhängiges Bewertungsgremium (IRC), stellt in GADOLIN den primären Endpunkt dar. PFS definiert als die Zeit von der Randomisierung bis zum Eintreten eines PFS-Ereignisses. PFS-Ereignisse sind das Fortschreiten der Erkrankung (Nachweis einer Progression oder Rezidiv) gemäß NHL-Responsekriterien (19) oder Tod (unabhängig von der Todesursache), je nachdem, welches Ereignis früher eintritt. Eine Progression (bei partieller Remission) bzw. Rezidiv (nach kompletter Remission) liegt vor, wenn mindestens eine der folgenden Bedingungen erfüllt ist: • Bildung neuer Läsionen (längster Durchmesser > 1,5 cm) während oder zum Ende der Behandlung, selbst wenn sich andere Läsionen zurückbilden. • Zunahme von mindestens 50 % in der Summe der Produkte der Durchmesser (SPD) zuvor identifizierter Lymphknoten oder anderer zuvor identifizierter Läsionen (z.B. in Milz oder Leber). Um als Progress gewertet zu werden, muss ein Lymphknoten mit einer kurzen Achse < 1 cm sich um mindestens 50 % vergrößern und sich auf eine Größe von 1,5 x 1,5 cm ausdehnen oder sich auf > 1,5 cm in der langen Achse vergrößern. • Zunahme von mindestens 50 % des längsten Durchmessers eines zuvor identifizierten Lymphknoten mit kürzestem Durchmesser > 1 cm. • Zunahme oder erstmaliges Auftreten von extralymphatischen Manifestationen (maligner Pleura-/Perikarderguss oder Aszites, Hepato-/Splenomegalie, Lymphangitis der Haut oder Lunge, Knochenläsionen, zystische Läsionen). <u>Erhebung:</u> • Baseline Tumorerhebung maximal 4 Wochen vor Randomisierung. • Während der Induktionstherapie erfolgt eine vollständige Erhebung des Tumorstatus (d.h. vollständige körperliche Untersuchung sowie bildgebende Verfahren) innerhalb von 14 Tagen vor Tag 1 von Zyklus 4. Bei Verdacht auf eine Progression auf der Basis eines Laborwerts oder klinischen Befundes können ungeplante Untersuchungen zur Abklärung des Tumorstatus durchgeführt werden. • Vollständige Erhebung des Tumorstatus bei Ende der Induktionstherapie innerhalb von 4-6 Wochen • In der Follow-Up Phase wird der Tumorstatus alle drei Monate nach dem letzten Zyklus der Induktionstherapie bzw. sechs Behandlungszyklen bis zur finalen Follow-Up Visite (28 Tage nach Monat 24) erhoben. • Weiterbeobachtung von Patienten mit vorzeitigem Therapieabbruch bis zum Eintreten eines PFS-Ereignisses. • In der Nachbeobachtungsphase ab Monat 25 erfolgt eine halbjährliche Erhebung des Tumorstatus bis einschließlich Monat 49. <u>Zensierung:</u> • Zensierung von Patienten ohne PFS-Ereignis zum Zeitpunkt der letzten IRC-Tumorbewertung. Falls keine IRC-Tumorbewertung für einen Patienten vorliegt, wird dieser zum Zeitpunkt der Randomisierung plus ein Tag zensiert. <u>PFS (IRC):</u> • Bewertung des Tumorstatus durch ein unabhängiges Bewertungsgremium (IRC) bestehend aus mindestens drei verblindeten Experten (ein Onkologe und zwei Radiologen) gemäß einer IRC-Charter (7). • Die Bewertung stützt sich auf folgende Informationen aus Pflichtvisiten: Anamnese zu Baseline inkl. aller Angaben zu Vorbehandlungen, Ergebnisberichte von körperlichen Untersuchungen der Patienten sowie radiologische Berichte von Ergebnissen der Computertomographien (CT) vor und nach Behandlung. Für die Auswertung werden die Datenschnitte vom 01. September 2014 sowie 01. Mai 2015 dargestellt. Der final bzw. konfirmatorische Datenschnitt ist der Datenschnitt vom 01. September 2014. Die Analysepopulation ist die ITT FL.

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

Tabelle 4-17: Bewertung des Verzerrungspotenzials für PFS (IRC) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
GADOLIN	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene liegt ein *niedriges* Verzerrungspotential vor.

Der Endpunkt wurde durch das IRC verblindet bewertet. Die Mitglieder des unabhängigen Bewertungsgremiums (IRC) waren gegenüber der verabreichten Therapie und der Beurteilung des Prüfarztes verblindet. Die CT-Aufnahmen wurden verblindet und unabhängig von den Befunden der Prüfarzte bewertet.

Es besteht ein mögliches Verzerrungspotential bzgl. Diagnostik und Progressionsbeurteilung, da die Beurteilung des Onkologen neben der Beurteilung der verblindeten CT-Aufnahmen auch auf Basis der Befunde des Prüfarztes erfolgen konnte. Der Einfluss auf das Verzerrungspotential wird wegen dieses sehr geringen subjektiven Anteils mit *niedrig* beurteilt.

Die Analyse des progressionsfreien Überlebens folgte für die FL-Population dem ITT-Prinzip (Subtyp des indolenten Non-Hodgkin-Lymphoms war Stratifizierungsfaktor) mit einer adäquaten Zensierung der Patienten ohne PFS-Ereignis.

Die im Studienbericht dargestellten Ergebnisse folgten den im Analyseplan spezifizierten Analysen.

Weitere verzerrungsrelevante Aspekte konnten nicht identifiziert werden.

Insgesamt wird deshalb das Verzerrungspotential für den Endpunkt PFS (IRC) als *niedrig* bewertet.

Tabelle 4-18: Ergebnisse für PFS (IRC) aus RCT mit dem zu bewertenden Arzneimittel

PFS (IRC)					
Datenschnitt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Final bzw. konfirmatorisch (01. Sep 2014)	54 / 155 (34,8)	90 / 166 (54,2)	n.e. [22,5; n.e.]	13,8 [11,4; 16,2]	0,48 [0,34; 0,68] < 0,0001
Follow-Up (01. Mai 2015)	67 / 164 (40,9)	108 / 171 (63,2)	29,2 [20,5; n.e.]	13,8 [11,5; 15,8]	0,47 [0,34; 0,64] < 0,0001
Population: ITT FL; (%): Anteil der Patienten mit Ereignis in %; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; IRC: Independent Review Committee; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; n.e.: nicht erreicht Statistik: Median basierend auf Kaplan-Meier Methode; 95 % KI: Brookmeyer und Crowley Methode; p-Wert: stratifizierter log-rank Test; Stratifizierungsfaktoren: Anzahl vorheriger Therapien, Art des Therapieversagens Quelle: CSR und CSR update (7, 12)					

In der Analyse des primären Endpunkts PFS (IRC) zum Zeitpunkt des final bzw. konfirmatorischen Datenschnittes vom 01. September 2014 zeigte sich ein statistisch signifikanter Unterschied zugunsten von G-Benda gegenüber Benda (HR: 0,48 [0,34; 0,68]; p<0,0001).

Mit diesem Ergebnis ist der primäre Endpunkt nach Studienprotokoll erreicht und der Datenschnitt vom 01. September 2014 gilt für alle Endpunkte außerhalb von Gesamtüberleben und Verträglichkeit als final bzw. konfirmatorisch.

Der Median für G-Benda liegt zum final bzw. konfirmatorischen Datenschnitt vom 01. September 2014 nur für PFS (Investigator) vor und beträgt 29,2 Monate; dieser Wert bestätigt sich im Follow-Up Datenschnitt vom 01. Mai 2015 durch das PFS (IRC) und bedeutet im Median einen Vorteil von 15,4 Monaten für G-Benda gegenüber Benda (12).

Die Follow-Up Analyse zum Zeitpunkt des Datenschnittes vom 01. Mai 2015 bestätigt den statistisch signifikanten Unterschied zugunsten von G-Benda gegenüber Benda bezüglich PFS (IRC) (HR: 0,47 [0,34; 0,64]; p<0,0001).

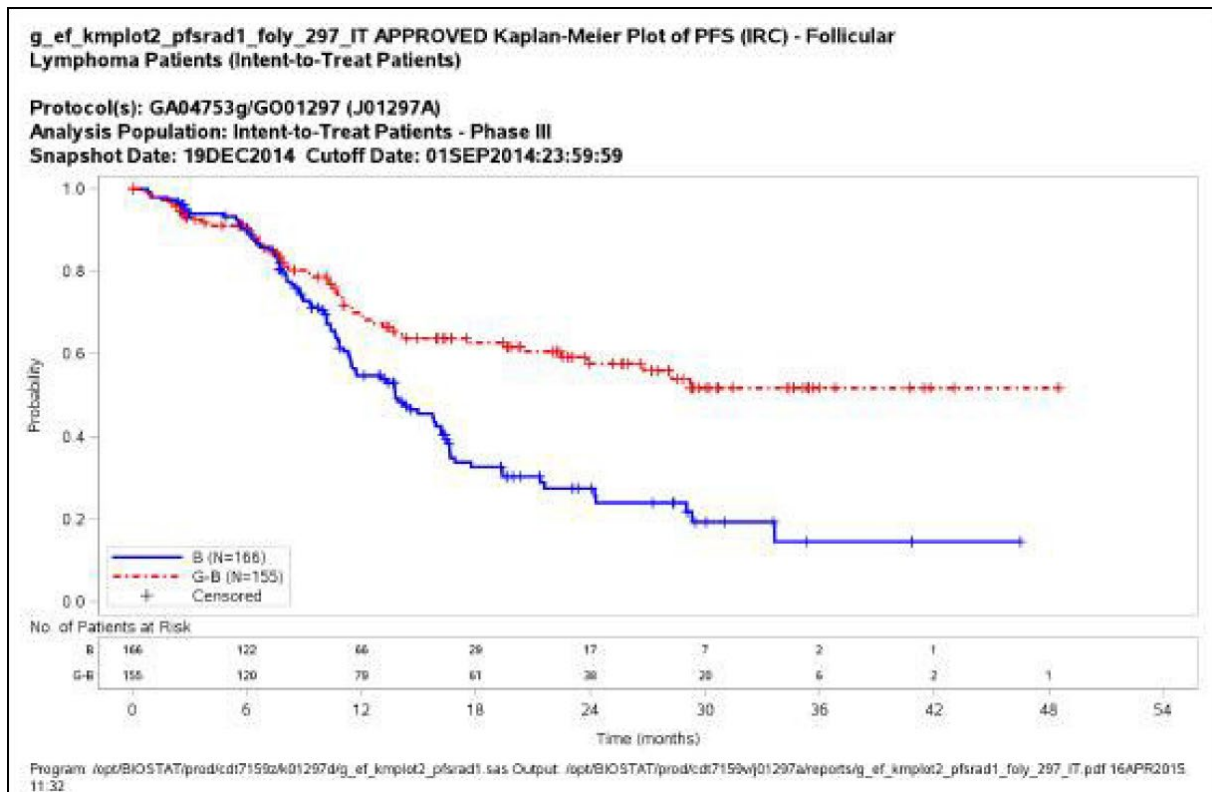


Abbildung 5: Kaplan-Meier-Kurve für den Endpunkt PFS (IRC) (Datenschnitt 01. Sept. 2014) (7)

In GADOLIN wurden verschiedene Sensitivitätsanalysen für den Endpunkt PFS durchgeführt. Dabei wird insbesondere der Einfluss spät auftretender Todesfälle, vorzeitiger Studienabbrüche, von Folgetherapien, fehlender IRC-Tumorbewertungen oder eventueller Randomisierungsfehler auf die Ergebnisse der Hauptanalyse untersucht. Die Resultate der Sensitivitätsanalysen haben die Ergebnisse der Hauptanalyse für diesen Endpunkt bestätigt (7).

4.3.1.3.1.3 Krankheitsfreies Überleben (IRC) – RCT

Tabelle 4-19: Operationalisierung von Krankheitsfreies Überleben

Studie	Operationalisierung
GADOLIN	Krankheitsfreies Überleben (DFS = disease-free survival) ist in GADOLIN definiert als die Zeit vom ersten Erreichen einer kompletten Remission (CR) bis zum Eintreten eines DFS-Ereignisses bestimmt durch IRC. DFS-Ereignisse sind das Wiederauftreten der Erkrankung (Nachweis eines Rezidivs, bestimmt durchs IRC) oder Tod (unabhängig von der Todesursache), je nachdem, welches Ereignis früher eintritt. Die komplette Remission (CR) ist ein Ergebnis der Tumorstuserhebung, beurteilt durch ein unabhängiges Bewertungsgremium (IRC). Gemäß der im Protokoll festgelegten Responsekriterien liegt eine CR dann vor, wenn <i>alle</i> der folgenden Kriterien erfüllt sind (19): · Vollständiger Rückgang aller zuvor bestehenden klinischen Manifestationen des Lymphoms und krankheitsbezogener Symptome. · Hepato- und/oder Splenomegalie, sofern vor Beginn der Therapie vorliegend, werden sowohl in der körperlichen Untersuchung als auch in der radiologischen Bildgebung als normal groß befundet. Eine noduläre Infiltration der Leber und/oder Milz ist nicht feststellbar. · Fehlender Nachweis einer Knochenmarkinfiltration durch das Lymphom, wenn vor Beginn der Therapie eine Beteiligung dokumentiert wurde. Die Bedingungen für das Vorliegen eines Rezidivs folgen den NHL-Responsekriterien (19): · Lymphknoten werden als pathologisch gewertet, wenn diese in der Längsachse 1,5 cm überschreiten (unabhängig von der Querachse) oder in der Längsachse bei 1,1-1,5 cm gleichzeitig in der Querachse 1 cm überschreiten. · Bildung neuer Läsionen · Auftreten von extralymphatischen Manifestationen (maligner Pleura-/Perikarderguss oder Aszites, Hepatomegalie, Splenomegalie oder normal große Milz mit Läsionen, Lymphangitis der Haut oder Lunge, Knochenläsionen, zystische Läsionen). <u>Zensierung:</u> · Patienten ohne IRC-dokumentiertes Rezidiv nach Erreichen einer kompletten Remission werden zum Zeitpunkt der letzten IRC-Tumorbewertung zensiert. <u>Erhebung:</u> · Während der Induktionstherapie erfolgt eine vollständige Erhebung des Tumorstatus (d. h. vollständige körperliche Untersuchung sowie bildgebende Verfahren) am Tag 1 von Zyklus 4. · Vollständige Erhebung des Tumorstatus bei Ende der Induktion innerhalb von 4-6 Wochen. · In der Follow-Up Phase wird der Tumorstatus alle drei Monate nach dem letzten Zyklus der Induktionstherapie bzw. sechs Behandlungszyklen bis zur finalen Follow-Up Visite (28 Tage nach Monat 24) erhoben. · In der Nachbeobachtungsphase ab Monat 25 erfolgt eine halbjährliche Erhebung des Tumorstatus bis einschließlich Monat 49. · Bei Verdacht auf ein Rezidiv auf Basis von Symptomatik oder eines klinischen Befundes werden ungeplante Untersuchungen (innerhalb von vier Wochen) zur Abklärung des Tumorstatus durchgeführt. <u>Komplette Remission (CR) nach IRC:</u> · Bewertung des IRC-Tumorstatus bestehend aus drei verblindeten Experten (ein Onkologe und zwei Radiologen) gemäß einer IRC-Charter (7). Für die Auswertung werden die Datenschnitte vom 01. September 2014 sowie 01. Mai 2015 dargestellt. Der final bzw. konfirmatorische Datenschnitt ist der Datenschnitt vom 01. September 2014. Die Analysepopulation setzt sich aus FL-Patienten mit CR während der Studie zusammen.

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

Tabelle 4-20: Bewertung des Verzerrungspotenzials für DFS und Rezidivrate (IRC) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
GADOLIN	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene liegt ein niedriges Verzerrungspotential vor.

Der Endpunkt wurde durch das IRC verblindet bewertet. Die Mitglieder des unabhängigen Bewertungsgremiums (IRC) waren gegenüber der verabreichten Therapie und der Beurteilung des Prüfarztes verblindet. Die CT-Aufnahmen wurden verblindet und unabhängig von den Befunden der Prüfarzte bewertet.

Es besteht ein mögliches Verzerrungspotential bzgl. Diagnostik und Progressionsbeurteilung, da die Beurteilung des Onkologen auch auf Basis der Befunde des Prüfarztes erfolgte. Der Einfluss auf das Verzerrungspotential wird wegen dieses sehr geringen subjektiven Anteils mit niedrig beurteilt.

Die Analyse zu krankheitsfreiem Überleben und der Rezidivrate erfolgt in der Subgruppe der Patienten, die während der Studie eine komplette Remission erreicht hatten. Es wurde nicht hiernach randomisiert. Insofern besteht grundsätzlich die Möglichkeit, dass die Patientenkollektive der beiden betrachteten Behandlungsarme unterschiedlich sein könnten. Allerdings konnten keine nennenswerten Unterschiede in Demographie oder Tumorlast zwischen den Behandlungsarmen festgestellt werden (18).

Die im Studienbericht dargestellten Ergebnisse folgten den im Analyseplan spezifizierten Analysen.

Weitere verzerrungsrelevante Aspekte konnten nicht identifiziert werden.

Insgesamt wird deshalb das Verzerrungspotential für den Endpunkt DFS/Rezidivrate (IRC) als *hoch* bewertet.

Krankheitsfreies Überleben (DFS) ist ein zusammengesetzter Endpunkt aus den Ereignissen Rezidiv (Morbidität) und Tod (Mortalität) bei Patienten, die eine komplette Remission erreicht haben. Im Folgenden wird der Anteil Morbidität im patientenrelevanten Endpunkt Rezidivrate und Zeit bis zum Rezidiv für GADOLIN separat neben DFS gesamthaft dargestellt (siehe Tabelle 4-21, Tabelle 4-22, Tabelle 4-23 und Tabelle 4-24).

Tabelle 4-21: Ergebnisse für Rezidivrate (IRC) aus RCT mit dem zu bewertenden Arzneimittel

Rezidivrate (IRC)					
Endpunkt	Patienten mit Ereignis n / N (%)		RR [95% KI] rRR [95% KI] p-Wert	OR [95% KI]	ARR in % [95% KI]
	G-Benda	Benda	G-Benda vs. Benda		
Final bzw. konfirmatorisch (01. Sep 2014)	2 / 32 (6,3)	14 / 31 (45,2)	0,14 [0,03; 0,56] n.z. 0,0055	0,08 [0,02; 0,40]	-38,9 % [-58,3; -19,5]
Follow-Up (01. Mai 2015)	3 / 35 (8,6)	15 / 33 (45,5)	0,19 [0,06; 0,59] n.z. 0,0043	0,11 [0,03; 0,44]	-36,9 % [-56,2; -17,5]
Population: FL-Patienten mit kompletter Remission (CR) während der Studie; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; IRC: Independent Review Committee; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten mit kompletter Remission; n.b.: nicht berechenbar; OR: Odds Ratio Statistik: p-Wert wurde mittels Wald Test bestimmt; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (die Berechnung und Darstellung des reversierten relativen Risikos erfolgte, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg betrachten zu können, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation des pU basiert in allen Fällen auf dem relativen Risiko. Quelle: Nachauswertung Roche (18)					

Zum Zeitpunkt beider Datenschnitte zeigte sich für die Rezidivrate ein signifikanter Unterschied zwischen beiden Behandlungsarmen.

In der final bzw. konfirmatorischen Analyse zum Zeitpunkt des Datenschnittes vom 01. September 2014 war die Rezidivrate im G-Benda-Arm um 38,9 % (95 % KI: [-58,3; -19,5]) geringer als im Benda-Arm. Das Risiko ein Rezidiv zu erleiden wird durch G-Benda um 86 % reduziert. Die Ergebnisse bedeuten eine mehr als 7-fache Reduktion der Häufigkeit der Ereignisse.

Die Follow-Up Analyse zum Zeitpunkt des Datenschnittes vom 01. Mai 2015 bestätigt den statistisch signifikanten Unterschied zugunsten von G-Benda gegenüber Benda (RR: 0,19; 95 % KI: [0,06; 0,59]; p = 0,0043).

Im Nachfolgenden wird zusätzlich zur Rezidivrate die Zeit bis zum Rezidiv dargestellt (Tabelle 4-22).

Tabelle 4-22: Zeit bis zum Rezidiv (IRC) aus RCT mit dem zu bewertenden Arzneimittel

Rezidiv (IRC)					
Datenschnitt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Final bzw. konfirmatorisch (01. Sep. 2014)	2 / 32 (6,3)	14 / 31 (45,2)	n.e.	13,0 [6,9; n.e.]	0,10 [0,02; 0,46] 0,0004
Follow-Up (01. Mai 2015)	3 / 35 (8,6)	15 / 33 (45,5)	n.e.	13,0 [8,2; n.e.]	0,15 [0,04; 0,52] 0,0007
Population: FL-Patienten mit CR während der Studie; (%): Anteil der Patienten mit Ereignis in %; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; IRC: Independent Review Committee; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; n.e.: nicht erreicht Statistik: Median basierend auf Kaplan-Meier Methode; 95 % KI: Brookmeyer und Crowley Methode; p-Wert: stratifizierter log-rank Test; Stratifizierungsfaktoren: Anzahl vorheriger Therapien, Art des Therapieversagens Quelle: CSR, CSR update und Nachauswertung Roche (7, 12, 18)					

Auch die Zeit bis zum Rezidiv zeigt einen statistisch signifikanten Vorteil für G-Benda gegenüber Benda (HR 0,10; 95 % KI [0,02; 0,46]; p = 0,0004). Der Median wurde im G-Benda-Arm bisher nicht erreicht. Aus Berechnungen auf Grundlage des Hazard Ratios und dem Median von Benda unter der Annahme, dass die Harzards über die Zeit konstant sind, ergibt sich ein geschätzter Median für G-Benda von 130 Monaten zum final bzw. konfirmatorischen Datenschnitt.

Nachfolgend wird zusätzlich eine Landmarkanalyse zu den Zeitpunkten 6 Monate und 12 Monate nach CR dargestellt. Für spätere Zeitpunkte ist aufgrund der noch laufenden Studie die Zahl der Patienten, die den Analysezeitpunkt bereits erreicht haben (siehe Tabelle 4-10), nicht ausreichend.

Tabelle 4-23: Landmark-Analyse für Rezidive (IRC) aus RCT mit dem zu bewertenden Arzneimittel

Rezidive – Landmark-Analyse								
Datenschnitt	6 Monate nach Erreichen einer CR				12 Monate nach Erreichen einer CR			
	G-Benda		Benda		G-Benda		Benda	
	Patienten noch unter Risiko N	Ereignis-freier Anteil in % [95 % KI]	Patienten noch unter Risiko N	Ereignis-freier Anteil in % [95 % KI]	Patienten noch unter Risiko N	Ereignis-freier Anteil in % [95 % KI]	Patienten noch unter Risiko N	Ereignis-freier Anteil in % [95 % KI]
Final bzw. konfirmatorisch (01. Sep. 2014)	24	100 [n.b.]	19	78,7 [58,6;89,9]	20	91,1 [68,8;97,7]	12	52,8 [31,7;70,2]
Follow-Up (01. Mai 2015)	29	100 [n.b.]	21	80,8 [62,1;90,9]	21	91,7 [70,6;97,9]	13	52,8 [32,5;69,6]
Population: FL-Patienten mit kompletter Remission (CR) während der Studie; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; N: Anzahl der Patienten; n.b.: nicht berechenbar Statistik: basierend auf Kaplan-Meier Methode; 95% KI: Brookmeyer und Crowley Methode Quelle: Nachauswertung Roche (18)								

Zum Zeitpunkt 6 Monate nach Erreichen einer CR waren *alle* Patienten im G-Benda-Arm rezidivfrei, im Benda-Arm 78,7 % zum Zeitpunkt des final bzw. konfirmatorischen Datenschnitts vom 01. September 2014. 12 Monate nach Erreichen einer CR waren noch 91,7 % der mit G-Benda behandelten Patienten rezidivfrei. Im Benda-Arm waren zu diesem Zeitpunkt nahezu die Hälfte rezidiert (52,8 % ereignisfrei) (siehe Tabelle 4-23). Der Datenschnitt vom 01. Mai 2015 bestätigt dieses Ergebnis.

Das krankheitsfreie Überleben (definiert als die Zeit zwischen CR und den Ereignissen Rezidiv oder Tod) ist in nachfolgender Tabelle 4-24 dargestellt.

Tabelle 4-24: Ergebnisse für DFS (IRC) aus RCT mit dem zu bewertenden Arzneimittel

Krankheitsfreies Überleben (DFS)					
Datenschnitt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Final bzw. konfirmatorisch (01. Sep 2014)	2 / 32 (6,3)	15 / 31 (48,4)	n.e.	13,0 [6,9; n.e.]	0,10 [0,02; 0,44] 0,0002
Follow-Up (01. Mai 2015)	3 / 35 (8,6)	16 / 33 (48,5)	n.e.	13,0 [8,2; n.e.]	0,14 [0,04; 0,48] 0,0004

Population: FL-Patienten mit CR während der Studie; (%): Anteil der Patienten mit Ereignis in %; Benda: Bendamustine; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; IRC: Independent Review Committee; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; n.e.: nicht erreicht

Statistik: Median basierend auf Kaplan-Meier Methode; 95 % KI: Brookmeyer und Crowley Methode; p-Wert: stratifizierter log-rank Test; Stratifizierungsfaktoren: Anzahl vorheriger Therapien, Art des Therapieversagens

Quelle: CSR, CSR update und Nachauswertung Roche (7, 12, 18)

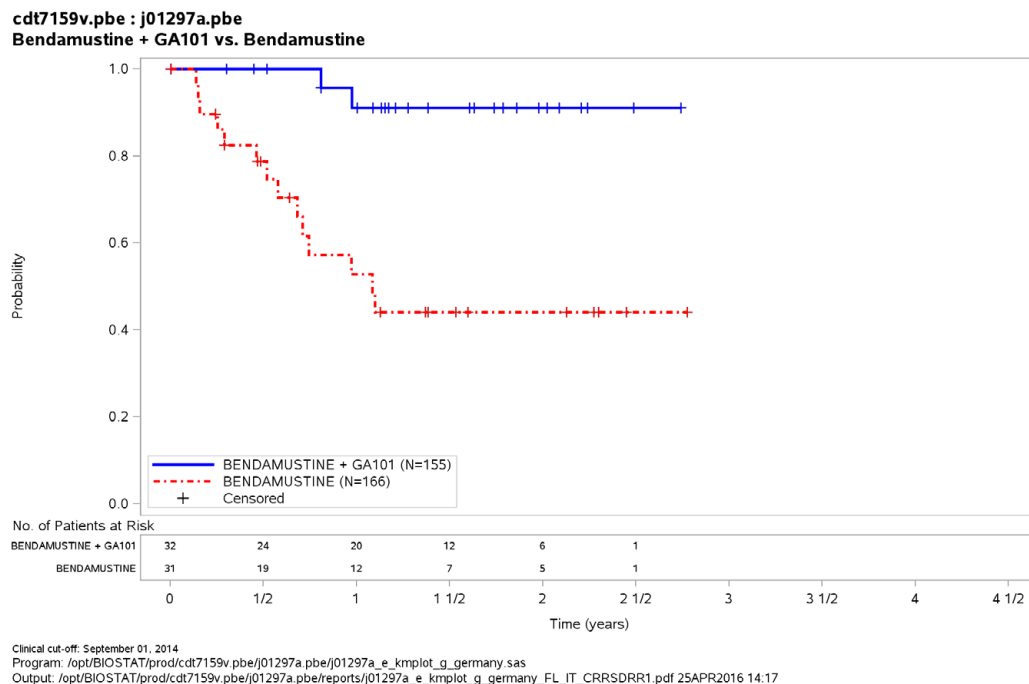


Abbildung 6: Kaplan-Meier-Kurve zum Endpunkt DFS (IRC) (Datenschnitt 01. September 2014) (18)

Es zeigte sich in der final bzw. konfirmatorischen Analyse vom 01. September 2014 ein statistisch signifikanter Unterschied zugunsten von G-Benda gegenüber Benda bezüglich des Endpunkts Krankheitsfreies Überleben (HR: 0,10 [0,02; 0,44]; $p=0,0002$), welcher mit dem Follow-Up Datenschnitt vom 1. Mai 2015 bestätigt wurde (HR: 0,14 [0,04; 0,48]; $p=0,0004$). Im Benda-Arm betrug der Median zum Zeitpunkt des final bzw. konfirmatorischen Datenschnitts 13,0 Monate (95 % KI [6,9;n.e.]). Im G-Benda-Arm war der Median zum Zeitpunkt des Datenschnitts noch nicht erreicht. Aus Berechnungen auf Grundlage des Hazard Ratios und dem Median von Benda unter der Annahme, dass die Harzards über die Zeit konstant sind, ergibt sich ein geschätzter Median für G-Benda von 130 Monaten zum final bzw. konfirmatorischen Datenschnitt. Zum Zeitpunkt des Follow-Up Datenschnitts betrug der Median im Benda-Arm 13,0 Monate (95 % KI [8,2;n.e.] und war im G-Benda-Arm nicht erreicht wurde aber wie beschrieben auf 94,3 Monate geschätzt.

Zum Zeitpunkt beider Datenschnitte war bei einem Patienten Tod das DFS-Ereignis im Benda-Arm (7, 12). Dieser Patient verstarb mehr als 12 Monate nach Erreichen der CR. Bei allen anderen Patienten lag ein Rezidiv vor.

Für das krankheitsfreie Überleben konnte sowohl für den Anteil Morbidität (Rezidivrate, Zeit bis zum Rezidiv) als auch gesamthaft ein statistisch signifikanter Vorteil für G-Benda gegenüber Benda gezeigt werden. In Abwesenheit eines Median konnte in der Landmarkanalyse kein Frührezidiv für G-Benda (21,3 % für Benda) und nach 12 Monaten eine Rezidivrate unter 10 % für G-Benda gegenüber nahezu 50 % für Benda nachgewiesen werden.

4.3.1.3.1.4 Gesundheitsbezogene Lebensqualität - FACT-Lym - RCT

Tabelle 4-25: Operationalisierung von FACT-Lym

Studie	Operationalisierung
GADOLIN	Zur Erhebung der Gesundheitsbezogenen Lebensqualität wird der FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma), ein Messinstrument für Lymphom-Patienten eingesetzt. Der FACT-Lym wurde 2013 in einer Studie an Patienten mit Non-Hodgkin-Lymphom validiert (20). Der FACT-Lym beinhaltet neben dem generischen Kernfragebogen FACT-G (FACT-General), der die wichtigsten Dimensionen der Gesundheitsbezogenen Lebensqualität bei Krebspatienten erhebt, zusätzliche Lymphom-spezifische Aspekte in der LymS-Subskala. Die Subskalen des FACT-Lym umfassen: · Körperliches Wohlbefinden (PWB) · Soziales/familiäres Wohlbefinden (SWB) · Emotionales Wohlbefinden (EWB) · Funktionelles Wohlbefinden (FWB) · Lymphom-spezifische Aspekte (FACT-LymS) Es lassen sich die folgenden drei etablierten und validierten Scores berechnen: · FACT-G (FACT-G=PWB+SWB+EWB+FWB) · FACT-Lym Trial Outcome Index (FACT-Lym TOI=PWB+FWB+LymS) · FACT-Lym Gesamtscore (FACT-Lym Total=LymS+FACT-G). Höhere Werte bedeuten eine bessere Lebensqualität. <u>Erhebung:</u> Der FACT-Lym wird während der Induktionstherapie jeweils am Tag 1 von Zyklus 1, 3 und 5 vor der Gabe der Studienmedikation ausgefüllt. In der Follow-Up Phase (nach dem letzten Zyklus bzw. 6 Behandlungszyklen) wird der Fragebogen alle zwei Monate sowie ca. 28 Tage nach Monat 24 (bzw. bei der Visite nach Behandlungsabbruch) erhoben. In der Nachbeobachtungsphase erfolgt die Erhebung alle 6 Monate. <u>Auswertung:</u> Die Scores FACT-G, FACT-Lym, LymS und FACT-Lym TOI werden an den Zeitpunkten Zyklus 5 Tag 1 und Follow-Up Visite 1 Monat 2 wie folgt ausgewertet: · Responderanalysen: Verbesserung der Lebensqualität um folgende MID (MID = minimal important difference) - o FACT-Lym Lymphoma Subscale: 5 (20, 21) - o FACT-G: 7 (22, 23) - o FACT-Lym TOI: 6 - o FACT-Lym Total: 7 · Auswertungen mittels adjustierter ANCOVA (Stratifikationsfaktoren Anzahl vorheriger Therapien, Art des Therapieversagens) Für die Auswertung werden die Datenschnitte vom 01. September 2014 sowie 01. Mai 2015 dargestellt. Der final bzw. konfirmatorische Datenschnitt ist der Datenschnitt vom 01. September 2014. Die Analysepopulation setzt sich aus FL-Patienten mit vorhandener Bewertung, der PRO-Population, zusammen.

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

Tabelle 4-26: Bewertung des Verzerrungspotenzials für Gesundheitsbezogene Lebensqualität – FACT-Lym - in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
GADOLIN	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene liegt ein niedriges Verzerrungspotential vor.

Die Erhebung dieses patientenberichteten Endpunkts erfolgte mit Kenntnis der Studientherapie.

Die Analyse der Gesundheitsbezogenen Lebensqualität folgt dem ITT-Prinzip.

Die im Studienbericht dargestellten Ergebnisse folgen der im statistischen Analyseplan spezifizierten Analyse.

Weitere verzerrungsrelevante Aspekte konnten nicht identifiziert werden.

Insgesamt wird das Verzerrungspotential für die Skalen des FACT-Lym als *hoch* bewertet, da eine relevante Verzerrung durch das offene Studiendesign nicht ausgeschlossen werden kann.

Dargestellte Zeitpunkte

Die Gesundheitsbezogene Lebensqualität wird im Folgenden zum Zeitpunkt Zyklus 5 Tag 1 (C5d1) sowie zum Zeitpunkt der Follow-Up Visite 1 Monat 2 (FU1M2) nach Ende der Induktionstherapie dargestellt. Der Zeitpunkt Zyklus 5 Tag 1 wurde gewählt, da dies der bestmögliche Zeitpunkt ist, um die Gesundheitsbezogene Lebensqualität während der Induktionstherapie (Chemo(-immun)therapie) darzustellen. Der Zeitpunkt FU1M2 wurde gewählt, um die Gesundheitsbezogene Lebensqualität zu einem Zeitpunkt nach Erholung von den Auswirkungen der Chemo(-immun)therapie zu reflektieren.

Rücklaufquote - Anteil der Patienten mit Bewertung

Die Gesamtzahl der Patienten, die die Zeitpunkte der Erhebung (C5d1, FU1M2) innerhalb der Studie bereits erreicht haben, gehen gesamthaft in die Analyse ein. Im Folgenden werden die Patienten mit auswertbarem Fragebogen als „Patienten mit Bewertung“ bezeichnet. Die Rücklaufquote liegt lediglich für den Zeitpunkt Zyklus 5 Tag 1 für beide Therapiearme über 70 %. Für den Zeitpunkt Follow-Up Visite 1 Monat 2 fällt die Rücklaufquote für den Benda-Arm unter 70 % (siehe Tabelle 4-27).

Tabelle 4-27: Anteil der Patienten mit Bewertung zu den Zeitpunkten Zyklus 5 Tag 1 und Follow-Up Visite 1 Monat 2 (Datenschnitt vom 01. September 2014 und 01. Mai 2015)

FACT-Lym - Anteil der Patienten mit Bewertung									
			Final bzw. confirmatorischer DS (01. September 2014)				Follow-Up DS (01. Mai 2015)		
			G-Benda N=155		Benda N=166		G-Benda N=164		Benda N=171
Zyklus 5 Tag 1	Patienten, die C5d1 erreicht haben* (n)		135		144		145		157
			m	m/n %	m	m/n %	m	m/n %	m
	Patienten mit Bewertung	FACT-Lym Lymphoma Subscale	114	84,4	105	72,9	122	84,1	114
		FACT-G	110	81,5	103	71,5	118	81,4	112
		FACT-Lym TOI	116	85,9	106	73,6	124	85,5	115
		FACT-Lym Total	109	80,7	104	72,2	117	80,7	113
Follow Up Visite 1 Monat 2	Patienten, die FU1M2 erreicht haben* (n)		122		121		135		135
			m	m/n %	m	m/n %	m	m/n %	m
	Patienten mit Bewertung	FACT-Lym Lymphoma Subscale	99	81,1	71	58,7	109	80,7	82
		FACT-G	96	78,7	71	58,7	106	78,5	82
		FACT-Lym TOI	101	82,8	72	59,5	111	82,2	83
		FACT-Lym Total	95	77,9	70	57,9	105	77,8	81

Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; DS: Datenschnitt; FACT-Lym: Functional Assessment of Cancer Therapy - Lymphoma; FACT-G: FACT general; FACT-Lym TOI: FACT Lym Trial Outcome Index; n: Anzahl auswertbarer Patienten; N: Anzahl der Patienten in FL ITT; m= Anzahl Patienten mit Bewertung

*Patienten ohne Studienabbruch bzw. ohne Ereignisse Progress oder Tod vor dem angegebenen Zeitpunkt, gingen in die Analyse ein.

Quelle: Nachauswertung Roche (18)

Responderanalysen (Verbesserung der Gesundheitsbezogenen Lebensqualität um MID)

Die Responderanalysen für die Verbesserung der Gesundheitsbezogenen Lebensqualität um eine definierte MID auf Patientenebene zeigten für keinen der untersuchten FACT-Lym-Scores zu keinem Zeitpunkt (C5d1, FU1M2) weder zum final bzw. konfirmatorischen noch zum Follow-Up Datenschnitt einen statistisch signifikanten Unterschied zwischen G-Benda und Benda (siehe Tabelle 4-28 und Tabelle 4-29).

Tabelle 4-28: Ergebnisse der Responderanalysen zur Verbesserung der Lebensqualität um die jeweilige MID aus RCT mit dem zu bewertenden Arzneimittel (01. September 2014)

Responderanalysen (Verbesserung der Gesundheitsbezogenen Lebensqualität um MID) – Final bzw. konfirmatorischer Datenschnitt (01. September 2014)							
Zeitpunkt	Endpunkt	MID	Patienten mit Ereignis n / N (%)		RR [95% KI] rRR [95% KI] p-Wert	OR [95% KI]	ARR in % [95% KI]
			G-Benda N=135	Benda N=144	G-Benda vs. Benda		
Zyklus 5 Tag 1	FACT-Lym Lymphoma Subscale	5	28 / 114 (24,6)	30 / 105 (28,6)	0,86 [0,55; 1,34] n.z. 0,5020	0,81 [0,45; 1,48]	-4,0 [-15,7; 7,7]
	FACT-G	7	21 / 110 (19,1)	23 / 103 (22,3)	0,85 [0,50; 1,45] n.z. 0,5600	0,82 [0,42; 1,59]	-3,2 [-14,1; 7,7]
	FACT-Lym TOI	6	40 / 116 (34,5)	28 / 106 (26,4)	1,31 [0,87; 1,96] 0,77 [0,51; 1,15] 0,1969	1,47 [0,82; 2,61]	8,1 [-4,0; 20,1]
	FACT-Lym Total	7	30 / 109 (27,5)	29 / 104 (27,9)	0,99 [0,64; 1,52] n.z. 0,9530	0,98 [0,54; 1,79]	-0,4 [-12,4; 11,7]
			G-Benda N=122	Benda N=121	G-Benda vs. Benda		
Follow Up Visite 1 Monat 2	FACT-Lym Lymphoma Subscale	5	29 / 99 (29,3)	22 / 71 (31,0)	0,95 [0,60; 1,50] n.z. 0,8119	0,92 [0,48; 1,79]	-1,7 [-15,7; 12,3]
	FACT-G	7	24 / 96 (25,0)	19 / 71 (26,8)	0,93 [0,56; 1,57] n.z. 0,7967	0,91 [0,45; 1,84]	-1,8 [-15,2; 11,7]
	FACT-Lym TOI	6	39 / 101 (38,6)	23 / 72 (31,9)	1,21 [0,80; 1,83] 0,83 [0,55; 1,26] 0,3731	1,34 [0,71; 2,53]	6,7 [-7,7; 21,0]
	FACT-Lym Total	7	31 / 95 (32,6)	23 / 70 (32,9)	0,99 [0,64; 1,55] n.z. 0,9756	0,99 [0,51; 1,91]	-0,2 [-14,7; 14,3]

Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; FACT-Lym: Functional Assessment of Cancer Therapy - Lymphoma; FACT-G: FACT general; FACT-Lym TOI: FACT Lym Trial Outcome Index; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; OR: Odds Ratio; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Statistik: p-Wert mittels Wald Test bestimmt; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des reversierten relativen Risikos, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation basiert in allen Fällen auf dem RR.)
Datenschnitt: 01. September 2014; Quelle: Nachauswertung Roche (18)

Tabelle 4-29: Ergebnisse der Responderanalysen zur Verbesserung der Lebensqualität um die jeweilige MID aus RCT mit dem zu bewertenden Arzneimittel (01. Mai 2015)

Responderanalysen (Verbesserung der Gesundheitsbezogenen Lebensqualität um MID) - Follow-Up Datenschnitt (01. Mai 2015)							
Zeitpunkt	Endpunkt	MID	Patienten mit Ereignis n / N (%)		RR [95% KI] rRR [95% KI] p-Wert	OR [95% KI]	ARR in % [95% KI]
			G-Benda N=145	Benda N=157	G-Benda vs. Benda		
Zyklus 5 Tag 1	FACT-Lym Lymphoma Subscale	5	31 / 122 (25,4)	31 / 114 (27,2)	0,93 [0,61; 1,43] n.z. 0,7558	0,91 [0,51; 1,63]	-1,8 [-13,0; 9,5]
	FACT-G	7	22 / 118 (18,6)	28 / 112 (25,0)	0,75 [0,45; 1,22] n.z. 0,2454	0,69 [0,37; 1,29]	-6,4 [-17,0; 4,3]
	FACT-Lym TOI	6	42 / 124 (33,9)	31 / 115 (27,0)	1,26 [0,85; 1,85] 0,80 [0,54; 1,17] 0,2495	1,39 [0,80; 2,42]	6,9 [-4,7; 18,5]
	FACT-Lym Total	7	33 / 117 (28,2)	33 / 113 (29,2)	0,97 [0,64; 1,45] n.z. 0,8671	0,95 [0,54; 1,69]	-1,0 [-12,7; 10,7]
			G-Benda N=135	Benda N=135	G-Benda vs. Benda		
Follow Up Visite 1 Monat 2	FACT-Lym Lymphoma Subscale	5	33 / 109 (30,3)	25 / 82 (30,5)	0,99 [0,64; 1,53] n.z. 0,9748	0,99 [0,53; 1,85]	-0,2 [-13,4; 13,0]
	FACT-G	7	28 / 106 (26,4)	23 / 82 (28,0)	0,94 [0,59; 1,51] n.z. 0,8025	0,92 [0,48; 1,76]	-1,6 [-14,5; 11,2]
	FACT-Lym TOI	6	43 / 111 (38,7)	26 / 83 (31,3)	1,24 [0,83; 1,84] 0,81 [0,54; 1,20] 0,2922	1,39 [0,76; 2,53]	7,4 [-6,1; 20,9]
	FACT-Lym Total	7	35 / 105 (33,3)	27 / 81 (33,3)	1,00 [0,66; 1,51] n.z. 1,0000	1,00 [0,54; 1,85]	0,0 [-13,7; 13,7]
Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; FACT Lym: Functional Assessment of Cancer Therapy - Lymphoma; FACT-G: FACT general; FACT-Lym TOI: FACT Lym Trial Outcome Index; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; OR: Odds Ratio; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Statistik: p-Wert mittels Wald Test bestimmt; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des reversierten relativen Risikos, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation basiert in auf dem RR.) Datenschnitt: 01. Mai 2015; Quelle: Nachauswertung Roche (18)							

Ergänzend zu den Responderanalysen wurden ANCOVAs berechnet. In der ANCOVA zeigte sich für keinen der untersuchten FACT-Lym-Scores zu keinem Zeitpunkt (C5d1, FU1M2) weder zum final bzw. konfirmatorischen noch zum Follow-Up Datenschnitt ein statistisch signifikanter Unterschied zwischen G-Benda und Benda (18).

4.3.1.3.1.5 Verträglichkeit – RCT

Tabelle 4-30: Operationalisierung von Verträglichkeit

Studie	Operationalisierung
GADOLIN	<u>Analyse-Population</u>: Die Analyse-Population ist die FL Safety-Population. Diese schließt alle Patienten ein, die mindestens eine Dosis der Studienmedikation erhalten haben. Die Patienten werden einer Behandlungsgruppe nach der tatsächlich erhaltenen Studienmedikation zugeteilt: Patienten, die unabhängig von der zugeteilten Medikation mindestens eine Dosis von Obinutuzumab erhielten, wurden im G-Benda-Arm ausgewertet. Insgesamt wurden 335 Patienten in die FL ITT-Population randomisiert (G-Benda: N = 164; Benda: N = 171). Drei Patienten im Benda-Arm erhielten keine Studienmedikation. Es erhielten 164 Patienten G-Benda und 168 Patienten Benda und entsprechend setzt sich die FL Safety-Population zusammen. <u>Unerwünschte Ereignisse</u> (UE) wurden gemäß Medical Dictionary for Regulatory Activities (MedDRA Version: 18) bzw. Roche INN (International non-proprietary Name) Drug Terms and Procedures Dictionary for medications and treatments kodiert. Die Daten zu den UE werden nach Inzidenz und NCI-CTCAE-Grad (NCI-CTCAE = National Cancer Institute - Common Terminology Criteria for Adverse Events), Version 4.0 dargestellt. <u>Schwerwiegende unerwünschte Ereignisse</u> (SUE) werden nach den ICH-Kriterien definiert. Ein SUE ist jedes aufgetretene Ereignis, das eine signifikante Bedrohung, eine Kontraindikation, eine Nebenwirkung oder eine Vorsichtsmaßnahme sein könnte. <u>Therapieabbruch wegen UE</u> wird mit einem permanenten Abbruch der Studienmedikation für folgende UE festgelegt: IRR Grad 4; IRR Grad 3 bei erneuter Gabe; Neutropenie/Anämie Grad 3 oder 4, die nicht zu ≤ Grad 2 abklingt und die Behandlung um mindestens 4 Wochen verzögert; Thrombozytopenie, die nicht zu ≤ Grad 1 abklingt (oder < 20 % Änderung zum Ausgangswert bei Patienten mit Knochenmarkbeteiligung zu Studienbeginn) und die Behandlung um mindestens 4 Wochen verzögert; wiederholt aufgetretene Zytopenie Grad 4 oder nicht-hämatologische Toxizität Grad 3 oder 4 trotz Absetzen von Bendamustin; Hepatitis B Reaktivierung trotz angemessener antiviraler Therapie; Progression der Erkrankung. <u>UE von besonderem Interesse</u> (AEPI = Adverse Events of Particular Interest) werden auf Basis der Wirkungsweise von Obinutuzumab und der Notwendigkeit der Aufdeckung des Verträglichkeitsprofils definiert (Infusionsbedingte Reaktionen (IRR), Neutropenie, Thrombozytopenie, Infektion, kardiale UE, Zweitmalignome, gastrointestinale (GI) Perforation, Hepatitis-B-Reaktivierung). <u>UE von speziellem Interesse</u> (AESI = Adverse Events of Special Interest) werden prospektiv im Studienprotokoll definiert und umfassen alle UE, bei denen eine gesonderte Überwachung und Verträglichkeitsanalyse notwendig ist (Tumor-Lyse-Syndrom (TLS), SUE Infektionen, SUE Neutropenie, SUE IRR).
Details zur Operationalisierung der dargestellten UE	
UE	Definition/Sicherheitsanalyse
Infusionsbedingte Reaktionen (IRR)	Behandlungsbedingtes UE, welches während oder innerhalb von 24 Stunden nach Beendigung der Infusion auftrat. Der Prüfarzt kreuzte ein Kästchen im eCRF für alle entsprechend aufgetretenen UE an. Es wurde empfohlen, ein zusätzliches UE „IRR“ zu melden. IRR beinhalten sowohl UE mit Preferred Term „Infusion Related Reaction“ (Infusionsbedingte Reaktion) als auch berichtete Anzeichen und Symptome infusionsbedingter Reaktionen.
Schwerwiegende infusionsbedingte Reaktionen	Behandlungsbedingtes UE, welches während oder innerhalb von 24 Stunden nach Beendigung der Infusion auftrat und schwerwiegend war. Der Prüfarzt kreuzte ein Kästchen im eCRF für alle entsprechend aufgetretenen SUE an.
Neutropenien	Roche Standard AEGT “Neutropenia and associated complications” (Neutropenien und assoziierte Komplikationen)

Spät einsetzende Neutropenien	Neutropenie ($< 1,0 \times 10^9/l$), die 24 Tage nach Ende der Behandlung oder später auftrat und der Patient bei der letzten Visite vor Behandlungsabschluss oder Behandlungsstopp eine absolute Neutrophilenzahl im Normbereich ($\geq 1,0 \times 10^9/l$) hatte.
Verlängerte Neutropenien	Neutropenie ($< 1,0 \times 10^9/l$), die zwischen 24 und 41 Tage nach Ende der Gabe der Studienmedikation auftrat und die auch bei der letzten Visite vor Behandlungsabbruch oder Behandlungsstopp vorhanden war.
Schwerwiegende Neutropenien	Roche Standard AEGT “Neutropenia and associated complications” (Neutropenien und assoziierte Komplikationen), Neutropenien, die schwerwiegend waren
Infektionen und Infestationen	Systemorganklasse “Infections and Infestations” (Infektionen und Infestationen)
Schwerwiegende Infektionen	Systemorganklasse “Infections and Infestations” (Infektionen und Infestationen), Infektionen, die schwerwiegend waren
Thrombozytopenien	SMQ “Haematopoietic Thrombocytopenia” (Hämatopoetische Thrombozytopenie)
Akute Thrombozytopenien	Thrombozytopenie, die während oder innerhalb von 24 Stunden nach der Infusion auftrat
Gastrointestinale Perforation	SMQ “Gastrointestinal perforation” (Gastrointestinale Perforation)
Kardiale Ereignisse	Systemorganklasse „Cardiac Disorders“ (Kardiale Ereignisse)
Zweite maligne Erkrankung	Systemorganklasse „Neoplasms benign, malignant and unspecified (including cysts and polyps)” (benigne, maligne und nicht spezifizierte (inklusive Zysten und Polypen) Neoplasien), auftretend ab 6 Monate nach der ersten Gabe der Studienmedikation
Hepatitis B Reaktivierung	Anstieg der HBV DNS nach Studienbeginn (≥ 100 IU/ml) an zwei aufeinanderfolgenden Visiten
Tumorlyse-Syndrom	Preferred Term „Tumor Lysis Syndrome“ (Tumorlyse-Syndrom)
Übelkeit	Preferred Term „Nausea“ (Übelkeit)
Erbrechen	Preferred Term „Emesis“ (Erbrechen)
Gewichtsverlust	Preferred Term „Weight Loss“ (Gewichtsverlust)
Hypotonie	Preferred Term „Hypotension“ (Hypotonie)
AEGT = adverse event group terms; SMQ = Standardised MedDRA Queries	

Die ausgewählten UE werden mittels standardisierter MedDRA-Abfragen (Queries), z.B. (SMQ) analysiert, da diese mit einem Set an UE konsistent sind, die besonders im Interesse von globalen regulatorischen Behörden stehen. Sind keine SMQ verfügbar, werden Gruppen von MedDRA Adverse Event Grouped Terms (AEGTs) verwendet. Alle UE nach Studieneintritt werden bis zu 28 Tage nach Beendigung der Studienmedikation oder zum Zeitpunkt des Behandlungsabbruchs dokumentiert. Alle UE Grad 3 und 4 werden bis zu 6 Monate nach Beendigung der Studienmedikation dokumentiert. Alle Infektionen Grad 3 und 4 (mit der Behandlung zusammenhängend) werden bis zu 24 Monate nach Beendigung der Studienmedikation dokumentiert. Mit der Behandlung zusammenhängende SUE wurden ohne zeitliche Begrenzung dokumentiert.

Ergebnisdarstellung erfolgt getrennt nach Induktionstherapie und gesamter Studienphase, um dem vorliegenden Studiendesign Rechnung zu tragen (nach der Induktionstherapie erfolgte im Benda-Arm eine Weiterbeobachtung ohne weitere Therapie bzw. im G-Benda-Arm eine Erhaltungstherapie mit Obinutuzumab Monotherapie für maximal zwei Jahre). Zusätzlich zu den Inzidenzen werden Time-to-first-event Analysen dargestellt.

Bis Version 8 (Oktober 2013) des Studienprotokolls fand im Benda-Arm keine systematische Erfassung von UE im Anschluss an die Induktionstherapie statt (150 Patienten mit Dokumentation bis Tag 28 nach Induktionstherapie; 18 Patienten mit Dokumentation im Follow-Up) (siehe Anhang 4-F).

Für die Datenauswertung wird ausschließlich der aktuellste Datenschnitt vom 01. Mai 2015 berücksichtigt.

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

Tabelle 4-31: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
GADOLIN	niedrig	nein	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene liegt ein niedriges Verzerrungspotential vor.

Die Erhebung von unerwünschten Ereignissen erfolgte durch den Prüfarzt mit Kenntnis der Studienbehandlung. Eine Beeinflussung der Studienergebnisse zur Verträglichkeit durch die Kenntnis der Studienbehandlung erscheint möglich aufgrund der subjektiven Komponente hinsichtlich der Beurteilung des Schweregrades einiger der erfassten UE.

Die Analyse der Verträglichkeit erfolgte adäquat auf der Basis der Patienten mit mindestens einer Studienmedikationsgabe gemäß dem "as-treated"-Prinzip, d.h. die Patienten wurden gemäß der tatsächlich erhaltenen Prüfmedikation ausgewertet. Die im Studienbericht dargestellten Ergebnisse folgen der im statistischen Analyseplan spezifizierten Analysen.

Folgende Aspekte könnten zu einer Verzerrung der Ergebnisse in der gesamten Studienphase beigetragen haben:

- Unterschiedliche Behandlungsdauer zwischen den Studienarmen (Benda: 6 Monate, entsprechend 6 Zyklen Induktionstherapie mit Benda; G-Benda: 2,5 Jahre (30 Monate), entsprechend 6 Zyklen Induktionstherapie mit G-Benda, gefolgt von einer Obinutuzumab-Erhaltungstherapie für Patienten mit Ansprechen auf die Induktion für maximal 2 Jahre);
- bis Version 8 des Studienprotokolls (Oktober 2013) keine systematische Erfassung von UE im Anschluss an die Induktionstherapie im Benda-Arm (150 Patienten mit Dokumentation bis Tag 28 nach Induktionstherapie; 18 Patienten mit Dokumentation im Follow-Up).

Die oben genannten Aspekte könnten zu einer Ergebnisverzerrung geführt haben, da weniger Daten zu UE im Benda-Arm in der gesamten Studienphase erfasst wurden (Underreporting). Dies wird zu einer Verzerrung zu Ungunsten des G-Benda-Arms in der gesamten Studienphase geführt haben.

Weitere verzerrungsrelevante Aspekte konnten nicht identifiziert werden.

Insgesamt wird deshalb das Verzerrungspotential für die Endpunkte zur Verträglichkeit als *hoch* bewertet.

Aufgrund unterschiedlicher Therapiedauer (6 Monate im Benda-Arm gegenüber maximal 30 Monaten im G-Benda-Arm) wird in der Darstellung der Ereignisse im Folgenden zwischen den Behandlungsphasen unterschieden. Die Induktionstherapie umfasst 6 Monate Therapie in beiden Behandlungsarmen mit G-Benda oder Benda. Die gesamte Studiendauer umfasst im G-Benda-Arm eine Erhaltungstherapie für Patienten mit Ansprechen auf die Induktionstherapie maximal zwei Jahre Behandlung mit Obinutuzumab als Monotherapie; im Benda-Arm ausschließlich eine Beobachtung (ab Version 8 des Studienprotokolls im Oktober 2013; 18 Patienten mit systematischer Dokumentation über Tag 28 nach Therapieende hinaus; 150 Patienten nur bis Tag 28 nach Induktionstherapie). Eine direkte Vergleichbarkeit ist somit primär nur für die Induktionstherapie mit gleicher Behandlungsdauer in beiden Armen gegeben. Unter Berücksichtigung des vorliegenden Studiendesigns, wird die Inzidenz aller UE über die gesamte Studiendauer hinweg dargestellt. Der unterschiedlichen Behandlungsdauer wird durch die Darstellung der Time-to-first-event Analyse (TTfE) Rechnung getragen. Die nachfolgende Tabelle 4-32 veranschaulicht die Wertigkeit der Analysen.

Tabelle 4-32: Wertigkeit der Analysen zur Darstellung der Verträglichkeit unter Berücksichtigung des Studiendesigns und der Verzerrung

	Inzidenz unerwünschter Ereignisse	TTfE
Vergleich der Induktionstherapie G-Benda versus Benda (6 Monate)	Direkte Vergleichbarkeit von aktiver Behandlung gegeben	Vollständige Abbildung mit Reduktion der Verzerrung
Vergleich der gesamten Studienphase	Vollständige Abbildung mit hoher Verzerrung zu Ungunsten G-Benda	

Tabelle 4-33: Ergebnisse für Verträglichkeit aus RCT mit dem zu bewertenden Arzneimittel

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
Patienten mit UE	Induktionstherapie	162 / 164 (98,8)	165 / 168 (98,2)	1,01 [0,98; 1,03] 0,99 [0,97; 1,02] 0,6713	1,47 [0,24; 8,93]	0,57 [-2,05; 3,18]
	gesamte Studiendauer	162 / 164 (98,8)	166 / 168 (98,8)	1,00 [0,98; 1,02] n.z. 0,9807	0,98 [0,14; 7,01]	-0,03 [-2,38; 2,32]
Patienten mit UE ≥ Grad 3	Induktionstherapie	85 / 164 (51,8)	86 / 168 (51,2)	1,01 [0,82; 1,25] 0,99 [0,80; 1,22] 0,9073	1,03 [0,67; 1,58]	0,64 [-10,11; 11,39]
	gesamte Studienphase	111 / 164 (67,7)	105 / 168 (62,5)	1,08 [0,92; 1,27] 0,92 [0,79; 1,08] 0,3224	1,26 [0,80; 1,98]	5,18 [-5,06; 15,42]
Patienten mit UE Grad 3	Induktionstherapie	75 / 164 (45,7)	80 / 168 (47,6)	0,96 [0,76; 1,21] n.z. 0,7305	0,93 [0,60; 1,43]	-1,89 [-12,62; 8,84]
	gesamte Studienphase	101 / 164 (61,6)	98 / 168 (58,3)	1,06 [0,89; 1,26] 0,95 [0,79; 1,13] 0,5455	1,15 [0,74; 1,78]	3,25 [-7,28; 13,79]

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
Patienten mit UE Grad 4	Induktionstherapie	40 / 164 (24,4)	34 / 168 (20,2)	1,21 [0,81; 1,80] 0,83 [0,55; 1,24] 0,3646	1,27 [0,76; 2,13]	4,15 [-4,8; 13,1]
	gesamte Studien- phase	54 / 164 (32,9)	40 / 168 (23,8)	1,38 [0,98; 1,96] 0,72 [0,51; 1,02] 0,0676	1,57 [0,97; 2,54]	9,12 [-0,54; 18,77]
Patienten mit UE Grad 5	Induktionstherapie	2 / 164 (1,2)	3 / 168 (1,8)	0,68 [0,12; 4,03] n.z. 0,6739	0,68 [0,11; 4,12]	-0,57 [-3,18; 2,05]
	gesamte Studien- phase	10 / 164 (6,1)	11 / 168 (6,5)	0,93 [0,41; 2,13] n.z. 0,8663	0,93 [0,38; 2,25]	-0,45 [-5,68; 4,78]
Patienten mit SUE	Induktionstherapie	44 / 164 (26,8)	37 / 168 (22,0)	1,22 [0,83; 1,78] 0,82 [0,56; 1,20] 0,3094	1,30 [0,79; 2,15]	4,81 [-4,43; 14,04]
	gesamte Studien- phase	64 / 164 (39,0)	58 / 168 (34,5)	1,13 [0,85; 1,50] 0,88 [0,67; 1,17] 0,3957	1,21 [0,78; 1,90]	4,50 [-5,86; 14,87]
Patienten mit Behandlungsab- bruch wegen UE	Induktionstherapie	20 / 164 (12,2)	30 / 168 (17,9)	0,68 [0,40; 1,15] n.z. 0,1532	0,64 [0,35; 1,18]	-5,66 [-13,32; 1,99]
	gesamte Studien- phase	29 / 164 (17,7)	30 / 168 (17,9)	0,99 [0,62; 1,57] n.z. 0,9669	0,99 [0,56; 1,73]	-0,17 [-8,40; 8,05]
Population: Safety FL; 2. Datenschnitt: 01. Mai 2015; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; OR: Odds Ratio; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des rRR, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend) Quelle: Nachauswertung Roche (18)						

In GADOLIN ist die Inzidenz der unerwünschten Ereignisse jeglicher Art, UE $\geq$ Grad 3, UE Grad 3, UE Grad 4, UE Grad 5, SUE und Behandlungsabbrüche aufgrund UE bis zum Zeitpunkt des aktuellsten Datenschnitts vom 01. Mai 2015 in beiden Therapiearmen sowohl in der Induktionstherapie als auch in der gesamten Studienphase vergleichbar (siehe Tabelle 4-33).

Über die Dauer der gesamten Studie wurden 10 Todesfälle (6,1 %) im G-Benda-Arm und 11 Todesfälle (6,5 %) im Benda-Arm als Grad 5 UE klassifiziert (siehe Tabelle 4-33). Es liegt kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen vor (RR: 0,93; 95 % KI: [0,41; 2,13]; p = 0,8663). Zwei von zehn UE Grad 5 im G-Benda-Arm und vier von elf UE Grad 5 im Benda-Arm wurden als in Zusammenhang mit der Studienmedikation stehend bewertet. Bei den zwei Ereignissen im G-Benda-Arm handelte es sich um myeloische Systemerkrankungen (ein myelodysplastisches Syndrom und eine akute myeloische Leukämie). Im Benda-Arm wurden zwei Fälle von akuter myeloischer Leukämie als UE Grad 5 dokumentiert (7, 12). Das myelodysplastische Syndrom und die akute myeloische Leukämie sind in den vorliegenden Fällen aufgrund der Latenz am ehesten mit der Vortherapie assoziiert (24). Nach dem aktuellen Stand der Forschung werden sie nicht in einen Zusammenhang

mit einer Anti-CD20-mAK-Therapie gestellt (25, 26). Bei den zwei weiteren mit der Studienmedikation in Verbindung gebrachten Todesfällen im Benda-Arm handelte es sich um Grad 5 Infektionen. Bei beiden führte eine *Pneumocystis jirovecii*-Pneumonie zum Tod.

Sowohl bei Betrachtung der Induktionstherapie (G-Benda: 12,2 % vs. Benda: 17,9 %; RR: 0,68; 95 % KI: [0,40; 1,15]; $p = 0,1532$) als auch der gesamten Studienphase (G-Benda: 17,7 % vs. Benda: 17,9 %; RR: 0,99; 95 % KI: [0,62; 1,57]; $p = 0,9669$) zeigen sich hinsichtlich der Therapieabbrüche aufgrund eines UE keine Unterschiede in beiden Behandlungsarmen (siehe Tabelle 4-33). Im Benda-Arm waren Thrombozytopenie (46,7 %) und Neutropenie (13,3 %) der Hauptgrund für den Behandlungsabbruch aufgrund UE. Die häufigsten Ursachen im G-Benda-Arm waren Neutropenien (24,1 %), Thrombozytopenien (20,7 %) und infusionsbedingte Reaktionen (IRR) (13,8 %) (7, 12).

In der folgenden Tabelle 4-34 werden IRR, das Tumorlysesyndrom sowie die mit IRR assoziierte Hypotonie dargestellt.

Tabelle 4-34: Ergebnisse für IRR und UE, die in Zusammenhang mit IRR gesehen werden

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
IRR	Induktionstherapie	109 / 164 (66,5)	111 / 168 (66,1)	1,01 [0,86; 1,17] 0,99 [0,85; 1,16] 0,9398	1,02 [0,65; 1,60]	0,39 [-9,78; 10,56]
	gesamte Studienphase	110 / 164 (67,1)	111 / 168 (66,1)	1,02 [0,87; 1,18] 0,99 [0,85; 1,15] 0,8466	1,05 [0,66; 1,65]	1,00 [-9,15; 11,15]
Schwerwiegende IRR	Induktionstherapie	8 / 164 (4,9)	2 / 168 (1,2)	4,10 [0,88; 19,01] 0,24 [0,05; 1,13] 0,0716	4,26 [0,89; 20,35]	3,69 [0,01; 7,37]
	gesamte Studienphase	8 / 164 (4,9)	2 / 168 (1,2)	4,10 [0,88; 19,01] 0,24 [0,05; 1,13] 0,0716	4,26 [0,89; 20,35]	3,69 [0,01; 7,37]
Tumorlyse-Syndrom	Induktionstherapie	1 / 164 (0,6)	2 / 168 (1,2)	0,51 [0,05; 5,59] n.z. 0,5834	0,51 [0,05; 5,67]	-0,58 [-2,61; 1,45]
	gesamte Studienphase	1 / 164 (0,6)	2 / 168 (1,2)	0,51 [0,05; 5,59] n.z. 0,5834	0,51 [0,05; 5,67]	-0,58 [-2,61; 1,45]
Hypotonie	Induktionstherapie	19 / 164 (11,6)	3 / 168 (1,8)	6,49 [1,96; 21,51] 0,15 [0,05; 0,51] 0,0022	7,21 [2,09; 24,85]	9,80 [4,51; 15,09]
	gesamte Studienphase	21 / 164 (12,8)	3 / 168 (1,8)	7,17 [2,18; 23,58] 0,14 [0,04; 0,46] 0,0012	8,08 [2,36; 27,64]	11,02 [5,53; 16,51]
Population: Safety FL; 2. Datenschnitt: 01. Mai 2015; (%) = Anteil der Patienten mit Ereignis in %; ARR = Absolute Risikoreduktion; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; OR: Odds Ratio; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des rRR, um beobachtete relative Effekte im Fall von $RR > 1$ gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation des pU basiert in allen Fällen auf dem RR) Quelle: Nachauswertung Roche (18)						

Für IRR, schwerwiegende IRR und TLS bestand sowohl in der Induktionstherapie als auch über die gesamte Studienphase kein signifikanter Unterschied zwischen den Behandlungsarmen (siehe Tabelle 4-34).

Keine IRR führte zu einem Todesfall. Alle Grad 3, Grad 4 und schwerwiegende IRR waren reversibel (18).

Bei einem Patienten im Benda-Arm führte das TLS zu einem fatalen Ereignis. Das TLS unter G-Benda war ein nicht schwerwiegendes reversibles Grad 3 Ereignis. Die im Vergleich zur Therapie der chronisch lymphatischen Leukämie (CLL) mit Obinutuzumab deutlich geringere Inzidenz des TLS beim follikulären Lymphom lässt sich durch die geringere Menge zirkulierender B-Lymphozyten erklären.

Eine Hypotonie trat sowohl in der Induktionstherapie als auch über die gesamte Studienphase statistisch signifikant häufiger unter G-Benda gegenüber Benda auf (siehe Tabelle 4-34). Alle Hypotonie-Ereignisse im G-Benda-Arm stehen im Zusammenhang mit IRR und waren reversibel. Die Anzahl der Grad 3 Hypotonien im G-Benda-Arm war ähnlich der im Benda-Arm (drei versus zwei Ereignisse). Alle übrigen Hypotonien waren als Grad 1 und Grad 2 UE nicht als schwer bzw. schwerwiegend zu betrachten.

Nachfolgend werden in Tabelle 4-35 hämatologische unerwünschte Ereignisse dargestellt.

Tabelle 4-35: Ergebnisse für hämatologische unerwünschte Ereignisse und Infektionen

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
Neutropenie	Induktionstherapie	50 / 164 (30,5)	46 / 168 (27,4)	1,11 [0,79; 1,56] 0,90 [0,64; 1,26] 0,5328	1,16 [0,72; 1,87]	3,11 [-6,65; 12,86]
	gesamte Studienphase	62 / 164 (37,8)	48 / 168 (28,6)	1,32 [0,97; 1,80] 0,76 [0,55; 1,03] 0,0760	1,52 [0,96; 2,41]	9,23 [-0,85; 19,32]
Schwerwiegende Neutropenie	Induktionstherapie	12 / 164 (7,3)	4 / 168 (2,4)	3,07 [1,01; 9,33] 0,33 [0,11; 0,99] 0,0476	3,24 [1,02; 10,25]	4,94 [0,33; 9,54]
	gesamte Studienphase	15 / 164 (9,1)	5 / 168 (3,0)	3,07 [1,14; 8,26] 0,33 [0,12; 0,87] 0,0261	3,28 [1,16; 9,25]	6,17 [1,06; 11,28]
Verlängerte Neutropenie***	gesamte Studienphase	3 / 164 (1,8)	0 / 168 (0,0)	[0,0; 1,460]* 0,1194**	n.b.	n.b.
Spät einsetzende Neutropenie***	gesamte Studienphase	12 / 164 (7,3)	0 / 168 (0,0)	[0,0; 0,275]* 0,0002**	n.b.	n.b.

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
Infektionen	Induktionstherapie	72 / 164 (43,9)	83 / 168 (49,4)	0,89 [0,71; 1,12] n.z. 0,3164	0,80 [0,52; 1,23]	-5,50 [-16,22; 5,21]
	gesamte Studienphase	107 / 164 (65,2)	100 / 168 (59,5)	1,10 [0,93; 1,30] 0,91 [0,77; 1,08] 0,2827	1,28 [0,82; 1,99]	5,72 [-4,68; 16,12]
Schwerwiegende Infektionen	Induktionstherapie	11 / 164 (6,7)	22 / 168 (13,1)	0,51 [0,26; 1,02] n.z. 0,0578	0,48 [0,22; 1,02]	-6,39 [-12,77; -0,01]
	gesamte Studienphase	22 / 164 (13,4)	31 / 168 (18,5)	0,73 [0,44; 1,20] n.z. 0,2134	0,68 [0,38; 1,24]	-5,04 [-12,89; 2,81]
Thrombozytopenie	Induktionstherapie	24 / 164 (14,6)	40 / 168 (23,8)	0,61 [0,39; 0,97] n.z. 0,0373	0,55 [0,31; 0,96]	-9,18 [-17,59; -0,76]
	gesamte Studienphase	25 / 164 (15,2)	40 / 168 (23,8)	0,64 [0,41; 1,01] n.z. 0,0526	0,58 [0,33; 1,00]	-8,57 [-17,04; -0,10]
Akute Thrombozytopenie	Induktionstherapie	1 / 164 (0,6)	0 / 168 (0,0)	[0,0; 14,328]* 0,4940**	n.b.	n.b.
	gesamte Studienphase	1 / 164 (0,6)	0 / 168 (0,0)	[0,0; 14,328]* 0,4940**	n.b.	n.b.
Hepatitis-B-Reaktivierung***	gesamte Studienphase	1 / 164 (0,6)	2 / 168 (1,2)	0,51 [0,05; 5,59] n.z. 0,5834	0,51 [0,05; 5,67]	-0,58 [-2,61; 1,45]

Population: Safety FL; 2. Datenschnitt: 01. Mai 2015; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; OR: Odds Ratio; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des rRR, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation basiert in allen Fällen auf dem RR.)

* exaktes Konfidenzintervall des RR; ** exakter p-Wert mittels Fisher's Test bestimmt, ***Fälle von verlängerten Neutropenien, spät einsetzenden Neutropenien und Hepatitis-B-Reaktivierung werden auf Basis der Visiten nach der Induktionstherapie diagnostiziert. Eine klare Zuordnung zur Induktionstherapie kann aus diesem Grund nicht getroffen werden, alle aufgetretenen UE sind in der Gesamtstudienphase aufgeführt.

Quelle: Nachauswertung Roche (18)

Für Neutropenien, Infektionen, schwerwiegende Infektionen, akute Thrombozytopenien, Hepatitis-B-Reaktivierung liegt sowohl in der Induktionstherapie als auch über die gesamte Studienphase kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen vor (siehe Tabelle 4-35).

Schwerwiegende Neutropenien traten statistisch signifikant häufiger im G-Benda-Arm als im Benda-Arm auf. Bei 3 Patienten (1,8 %) trat unter G-Benda eine verlängerte Neutropenie und bei 12 Patienten (7,3 %) eine spät einsetzende Neutropenie auf (siehe Tabelle 4-35). In diesem Zusammenhang ist zu berücksichtigen, dass nur für 18 Patienten im Benda-Arm eine systematische Dokumentation der UE über die Induktionstherapie hinaus erfolgte. Als reiner Laborparameter hat die Neutropenie keinen unmittelbaren Einfluss auf das Wohlbefinden des Patienten. Sie ist dann medizinisch relevant, wenn gleichzeitig ein Infekt vorliegt. Für den Endpunkt febrile Neutropenien lässt sich für die Gesamtstudienphase kein Unterschied zwischen beiden Therapiearmen feststellen (G-Benda: 5,5 % vs. Benda: 3,0 %; RR: 1,84; 95 % KI: [0,63; 5,39]; $p = 0,2632$) (12).

Für die Endpunkte Infektionen und schwerwiegende Infektionen zeigen sich keine statistisch signifikanten Unterschiede zwischen den Therapiearmen (siehe Tabelle 4-35). Neben einem numerischen Vorteil bei schwerwiegenden Infektionen zugunsten von G-Benda, lassen sich niedrigere Inzidenzen bei Infektionen $\geq$ Grad 3 erkennen. In der gesamten FL Safety-Population wurden diese im G-Benda-Arm bei 10,7 % und im Benda-Arm bei 22,7 % diagnostiziert (18). Im Benda-Arm verliefen fünf Infektionen tödlich. Drei Patienten hatten eine Pneumonie und bei zwei Patienten führte eine Sepsis zum Tod. Unter G-Benda hatten drei Patienten Infektionen mit fatalem Verlauf (zweimal Sepsis und ein Multiorganversagen in Folge einer septisch verlaufenden Gastroenteritis).

In der Induktionstherapie bestand für Thrombozytopenien ein statistisch signifikanter Vorteil für G-Benda (RR: 0,61; 95 % KI: [0,39; 0,97]; $p = 0,0373$) (siehe Tabelle 4-35). Es waren keine Blutungen (Grad 3-5) mit einer Thrombozytopenie assoziiert. Bei einem Patienten im G-Benda-Arm (0,6 %) wurde eine akute Thrombozytopenie diagnostiziert, die ohne Behandlung reversibel war.

In Tabelle 4-36 werden die Ergebnisse für die UE von besonderem Interesse, gastrointestinale Perforation, kardiale Ereignisse und Zweitmalignome, dargestellt. Des Weiteren wurden aufgrund der statistisch signifikanten Unterschiede und Häufigkeit die gastrointestinalen UE Übelkeit, Erbrechen und Gewichtsverlust ausgewählt.

Tabelle 4-36: Weitere ausgewählte UE aus RCT mit dem zu bewertenden Arzneimittel

Endpunkt		Patienten mit Ereignis n/N (%)		G-Benda vs. Benda		
		G-Benda	Benda	RR [95 % KI] rRR [95 % KI] p-Wert	OR [95 % KI]	ARR in % [95 % KI]
Gastrointestinale Perforationen	Induktionstherapie	1 / 164 (0,6)	0 / 168 (0,0)	[0,0; 14,328]* 0,4940**	n.b.	n.b.
	gesamte Studienphase	1 / 164 (0,6)	0 / 168 (0,0)	[0,0; 14,328]* 0,4940**	n.b.	n.b.
Übelkeit	Induktionstherapie	78 / 164 (47,6)	106 / 168 (63,1)	0,75 [0,62; 0,92] n.z. 0,0051	0,53 [0,34; 0,82]	-15,53 [-26,10; -4,97]
	gesamte Studienphase	84 / 164 (51,2)	106 / 168 (63,1)	0,81 [0,67; 0,98] n.z. 0,0305	0,61 [0,40; 0,95]	-11,88 [-22,45; -1,30]
Erbrechen	Induktionstherapie	31 / 164 (18,9)	48 / 168 (28,6)	0,66 [0,44; 0,98] n.z. 0,0414	0,58 [0,35; 0,97]	-9,67 [-18,76; -0,58]
	gesamte Studienphase	35 / 164 (21,3)	48 / 168 (28,6)	0,75 [0,51; 1,09] n.z. 0,1312	0,68 [0,41; 1,12]	-7,23 [-16,50; 2,04]
Gewichtsverlust	Induktionstherapie	6 / 164 (3,7)	18 / 168 (10,7)	0,34 [0,14; 0,84] n.z. 0,0191	0,32 [0,12; 0,82]	-7,06 [-12,54; -1,57]
	gesamte Studienphase	7 / 164 (4,3)	18 / 168 (10,7)	0,40 [0,17; 0,93] n.z. 0,0330	0,37 [0,15; 0,92]	-6,45 [-12,05; -0,84]
Kardiale Ereignisse	Induktionstherapie	13 / 164 (7,9)	7 / 168 (4,2)	1,90 [0,78; 4,65] 0,53 [0,22; 1,28] 0,1582	1,98 [0,77; 5,10]	3,76 [-1,36; 8,88]
	gesamte Studienphase	20 / 164 (12,2)	9 / 168 (5,4)	2,28 [1,07; 4,85] 0,44 [0,21; 0,94] 0,0331	2,45 [1,08; 5,56]	6,84 [0,78; 12,89]
Zweitmalignom	Induktionstherapie	1 / 164 (0,6)	0 / 168 (0,0)	[0,0; 14,328]* 0,4940**	n.b.	n.b.
	gesamte Studienphase	12 / 164 (7,3)	9 / 168 (5,4)	1,37 [0,59; 3,15] 0,73 [0,32; 1,69] 0,4654	1,39 [0,57; 3,40]	1,96 [-3,28; 7,20]

Population: Safety FL; 2. Datenschnitt: 01. Mai 2015; (%): Anteil der Patienten mit Ereignis in %; ARR: Absolute Risikoreduktion; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; OR: Odds Ratio; RR: Relatives Risiko G-Benda vs. Benda; rRR: reversiertes relatives Risiko Benda vs. G-Benda (Berechnung und Darstellung des rRR, um beobachtete relative Effekte im Fall von RR > 1 gleichgerichtet über alle Endpunkte hinweg zu betrachten, ansonsten n.z.: nicht zutreffend, die inhaltliche und statistische Interpretation basiert in allen Fällen auf dem RR.)
 * exaktes Konfidenzintervall des RR; ** exakter p-Wert mittels Fisher's Test bestimmt
 Quelle: Nachauswertung Roche (18)

Für gastrointestinale Perforationen und Zweitmalignome bestand sowohl in der Induktionstherapie als auch über die gesamte Studienphase kein signifikanter Unterschied zwischen den Behandlungsarmen.

Bei einem Patienten (0,6 %) mit vorbestehender Divertikulose wurde im Rahmen einer Divertikulitis eine gedeckte gastrointestinale Perforation im G-Benda-Arm während der Erhaltungstherapie diagnostiziert, welche durch eine konservative Therapie mit Antibiotika komplikationslos ausheilte (siehe Tabelle 4-36).

Übelkeit und Erbrechen traten in der Induktionstherapie signifikant häufiger im Benda-Arm im Vergleich zum G-Benda-Arm auf (siehe Tabelle 4-36). Die statistische Signifikanz hinsichtlich Erbrechens verliert sich aufgrund von vier Patienten in der Erhaltungstherapie bzw. Follow-Up Phase im G-Benda-Arm.

Ungewollter Gewichtsverlust wurde signifikant häufiger bei Patienten im Benda-Arm als im G-Benda-Arm dokumentiert (siehe Tabelle 4-36).

Kardiale UE traten in der Induktionstherapie vergleichbar häufig in beiden Behandlungsarmen auf (siehe Tabelle 4-36). Unter Berücksichtigung der gesamten Studienphase erreichte dieser Unterschied mit Ereignissen bei sieben Patienten im G-Benda-Arm Signifikanzniveau. Schwerwiegende kardiale UE wurden bei 3,7 % der Patienten im G-Benda-Arm und 1,2 % der Patienten im Benda-Arm dokumentiert; dieser Unterschied ist nicht statistisch signifikant (RR: 3,0; 95 % KI: [0,61; 14,65]; $p = 0,1746$) (18). Kein kardiales Ereignis zeigte einen fatalen Verlauf. Vier der kardialen Ereignisse waren bis zum Datenschnitt vom 01. Mai 2015 noch nicht als abgeklungen dokumentiert. Es handelt sich um eine Sinusbradykardie Grad 1, ein Vorhofflimmern Grad 2, eine Herzinsuffizienz Grad 2 und eine Herzinsuffizienz Grad 4. Letztere stellte eine akute Verschlechterung einer vorbestehenden Herzinsuffizienz dar. Im Therapiealltag stellen kardiale Vorerkrankungen keine Kontraindikation einer Obinutuzumab-Therapie dar, sind jedoch im Rahmen von möglichen Infusionsreaktionen präventiv zu berücksichtigen (5).

Aufgrund der unterschiedlichen Behandlungsdauer (6 Monate im Benda-Arm gegenüber maximal 30 Monaten im G-Benda-Arm) und abweichender systematischer Beobachtungsdauer (150 Patienten mit Dokumentation bis Tag 28 nach Induktionstherapie; 18 Patienten mit Dokumentation im Follow-Up im Benda-Arm) beider Studienarme wurde zusätzlich eine Time-to-first-event Analyse (siehe Tabelle 4-37) der unerwünschten Ereignisse durchgeführt (siehe auch Tabelle 4-32).

Tabelle 4-37: Time-to-first-event Analyse (TTfE)

Endpunkt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Patienten mit UE	162 / 164 (98,8)	166 / 168 (98,8)	0,07 [0,03; 0,10]	0,10 [0,07; 0,10]	1,04 [0,84; 1,29] 0,6998
Patienten mit UE $\geq$ Grad 3	111 / 164 (67,7)	105 / 168 (62,5)	5,65 [4,60; 11,86]	5,13 [3,91; 8,41]	0,76 [0,58; 1,01] 0,0608
Patienten mit UE Grad 3	101 / 164 (61,6)	98 / 168 (58,3)	11,14 [5,26; 17,15]	6,77 [4,63; 8,67]	0,69 [0,51; 0,93] 0,0135
Patienten mit UE Grad 4	54 / 164 (32,9)	40 / 168 (23,8)	n.e.	12,22 [9,82; 20,44]	0,95 [0,61; 1,47] 0,8112
Patienten mit UE Grad 5	10 / 164 (6,1)	11 / 168 (6,5)	n.e.	18,69 [12,81; 24,54]	0,11 [0,03; 0,34] 0,0001
Patienten mit SUE	64 / 164 (39,0)	58 / 168 (34,5)	42,51 [26,58; n.e.]	11,96 [8,67; 18,69]	0,60 [0,40; 0,89] 0,0097
Patienten mit Behandlungsab- bruch wegen UE	29 / 164 (17,7)	30 / 168 (17,9)	n.e.	n.e.	0,64 [0,36; 1,12] 0,1125
IRR	110 / 164 (67,1)	111 / 168 (66,1)	0,10 [0,07; 0,99]	0,95 [0,10; 1,08]	1,09 [0,84; 1,42] 0,4808
Schwerwiegende IRR	8 / 164 (4,9)	2 / 168 (1,2)	n.e.	n.e.	4,14 [0,88; 19,51] 0,0500
Tumorlyse-Syndrom	1 / 164 (0,6)	2 / 168 (1,2)	n.e.	n.e.	0,51 [0,05; 5,67] 0,5803
Neutropenie	62 / 164 (37,8)	48 / 168 (28,6)	n.e.	n.e. [12,81; n.e.]	1,01 [0,68; 1,49] 0,9709
Schwerwiegende Neutropenie	15 / 164 (9,1)	5 / 168 (3,0)	n.e.	n.e. [12,81; n.e.]	2,14 [0,74; 6,15] 0,1497
Verlängerte Neutropenie*	3 / 164 (1,8)	0 / 168 (0,0)	n.b.	n.b.	n.b.
Spät einsetzende Neutropenie*	12 / 164 (7,3)	0 / 168 (0,0)	n.b.	n.b.	n.b.
Thrombozytopenie	25 / 164 (15,2)	40 / 168 (23,8)	n.e.	n.e.	0,61 [0,37; 1,01] 0,0497
Akute Thrombozytopenie	1 / 164 (0,6)	0 / 168 (0,0)	n.b.	n.b.	n.b.
Infektionen	107 / 164 (65,2)	100 / 168 (59,5)	8,51 [5,29; 11,83]	5,65 [4,44; 7,43]	0,68 [0,51; 0,91] 0,0088
Schwerwiegende Infektionen	22 / 164 (13,4)	31 / 168 (18,5)	n.e.	17,08 [9,95; 18,69]	0,22 [0,11; 0,44] 0,0000

Endpunkt	Patienten mit Ereignis n / N (%)		Median in Monaten [95 % KI]		Hazard Ratio [95 % KI] p-Wert
	G-Benda	Benda	G-Benda	Benda	G-Benda vs. Benda
Hepatitis B Reaktivierung*	1 / 164 (0,6)	0 / 168 (1,2)	n.b.	n.b.	n.b.
Gastrointestinale Perforationen	1 / 164 (0,6)	0 / 168 (0,0)	n.b.	n.b.	n.b.
Kardiale Ereignisse	20 / 164 (12,2)	9 / 168 (5,4)	n.e.	22,24 [n.e.; n.e.]	1,16 [0,49; 2,78] 0,7328
Zweitmalignom	12 / 164 (7,3)	9 / 168 (5,4)	n.e.	20,47 [20,44; 24,54]	0,08 [0,03; 0,19] 0,0000
Übelkeit	84 / 164 (51,2)	106 / 168 (63,1)	8,38 [1,87; n.e.]	1,05 [0,23; 2,56]	0,68 [0,51; 0,92] 0,0076
Erbrechen	35 / 164 (21,3)	48 / 168 (28,6)	n.e.	n.e.	0,63 [0,40; 0,99] 0,0432
Gewichtsverlust	7 / 164 (4,3)	18 / 168 (10,7)	n.e.	n.e.	0,32 [0,13; 0,80] 0,0104
Hypotonie	21 / 164 (12,8)	3 / 168 (1,8)	n.e.	n.e.	6,17 [1,82; 20,94] 0,0007
Population: Safety FL; 2. Datenschnitt: 01. Mai 2015; (%): Anteil der Patienten mit Ereignis in %; Benda: Bendamustin; G-Benda: G-Benda Induktionstherapie gefolgt von einer Obinutuzumab Erhaltungstherapie; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten; n.b.: nicht berechenbar; n.e.: nicht erreicht Statistik: Median basierend auf Kaplan-Meier Methode; 95 % KI: Brookmeyer und Crowley Methode; p-Wert: log-rank Test *Fälle von verlängerten Neutropenien, spät einsetzenden Neutropenien und Hepatitis-B-Reaktivierung werden auf Basis der Visiten nach der Induktionstherapie diagnostiziert. Der exakte Beginn des Ereignisses ist damit nicht erfasst, daher kann eine TTfE Analyse nicht durchgeführt werden Quelle: Nachauswertung Roche (18)					

Die Time-to-first-event Analyse zeigt einen statistisch signifikanten Unterschied zugunsten von G-Benda bei der Betrachtung der SUE, der UE Grad 3 und 5 sowie der UE Thrombozytopenie, Infektionen, schwerwiegende Infektionen, Zweitmalignome, Übelkeit, Erbrechen und Gewichtsverlust.

Ein signifikanter Nachteil in dieser Betrachtung von G-Benda gegenüber Benda zeigt sich hinsichtlich schwerwiegender IRR und Hypotonie; diese beiden UE waren in allen Fällen reversibel (siehe Erläuterung unterhalb der Tabelle 4-34).

Verlängerte Neutropenie, spät einsetzende Neutropenie und Hepatitis-B-Reaktivierung wurden auf Basis von Visiten nach Ende der Induktionstherapie diagnostiziert. Im Benda-Arm ist lediglich für 18 Patienten für diese Ereignisse aufgrund der Version 8 des Protokolls eine Dokumentation vorhanden.

4.3.1.3.2 Subgruppenanalysen – RCT

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

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

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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.

Nicht zutreffend.

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.

Nicht zutreffend

Eine Zusammenfassung der Ergebnisse ist in Kapitel 4.4.2 zu finden.

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).** Benennen Sie

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

Nicht zutreffend.

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.**

Nicht zutreffend.

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 4-38: 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

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 4-39: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

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

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-40: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

Tabelle 4-41: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-42: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen 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.

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

Nicht zutreffend.

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

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

Nicht zutreffend.

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-F. 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-F.

Tabelle 4-43: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

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 4-44: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-45: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

Nicht zutreffend.

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).** Benennen Sie

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

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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

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

Nicht zutreffend.

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

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

Tabelle 4-46: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

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

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

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

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.2.1.3.2.***

Nicht zutreffend.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

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

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Obinutuzumab (Gazyvaro[®]) ist ein Arzneimittel zur Behandlung einer seltenen Erkrankung nach Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates (EU orphan designation number EU/3/15/1504) (1). Damit gilt für Obinutuzumab der medizinische Nutzen und Zusatznutzen mit der Zulassung als belegt (Absatz 1 Satz 10 §35a SGB V) (2).

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. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. 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.

Obinutuzumab (Gazyvaro[®]) ist ein Arzneimittel zur Behandlung eines seltenen Leidens (Orphan Drug) entsprechend Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und

des Rates vom 16. Dezember 1999, dessen Umsatz in 12 Kalendermonaten den Betrag von 50 Mio. € nicht übersteigt (1, 2). Gemäß § 35a Abs. 1 Satz 10 SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Es ist das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen nachzuweisen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht. Das Ausmaß des Zusatznutzens ist auf Grundlage der Zulassung und der die Zulassung begründenden Studien zu bestimmen. Es bedarf keiner Ausführungen zu einer zweckmäßigen Vergleichstherapie.

Anwendungsgebiet

Das im vorliegenden Nutzendossier betrachtete Anwendungsgebiet von Obinutuzumab umfasst Patienten mit FL, die auf eine Behandlung mit Rituximab oder einem Rituximabhaltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden (5). In GADOLIN wurden Rituximab-refraktäre Patienten mit indolentem NHL randomisiert, von denen ca. 80 % an einem FL erkrankt waren.

Zulassungsstudie GADOLIN

Patienten mit einem FL, die nach der initialen Chemo-Immuntherapie eine mehrjährige Remission erreichen, können auch in der nächsten Therapielinie mit einem PFS von etwa bis zu 4,5 Jahren rechnen (27, 28). Hingegen werden bei den Rituximab-refraktären Patienten mit den derzeit eingeschränkten Therapieoptionen lediglich Remissionen mit einer Dauer von weniger als einem Jahr erreicht (29-33). Es bestand daher ein großer Bedarf speziell für dieses Patientenkollektiv, solide Evidenz zu generieren, um einen neuen, effektiveren und verträglichen Therapiestandard zu etablieren.

GADOLIN ist die erste randomisierte kontrollierte Studie, die ausschließlich Rituximab-refraktäre Patienten mit indolentem NHL einschloss. Mit einer für die Indikation außerordentlich großen Studienpopulation von insgesamt 413 Patienten (davon 335 FL-Patienten), kann eine signifikante Überlegenheit von G-Benda gegenüber Benda bezüglich des primären Studienendpunkts progressionsfreies Überleben (IRC) und auch sekundärer Endpunkte demonstriert werden. Dies wird erreicht bei einem zwischen den Studienarmen vergleichbaren Sicherheitsprofil. Zudem zeigt sich in der Auswertung des Datenschnitts vom 01. Mai 2015 ein statistisch signifikanter Unterschied zugunsten von G-Benda gegenüber Benda bezüglich des Endpunkts Gesamtüberleben.

Patientenrelevante Endpunkte

Mortalität

v Gesamtüberleben

Die Mortalität in GADOLIN wurde durch die Erfassung des Gesamtüberlebens gemessen.

Das FL gehört zur Gruppe der niedrig malignen indolenten NHL und gilt bisher nur in frühen Stadien mit Standardtherapien als potenziell heilbar (4). Allerdings werden FL in über 80 % der Fälle erst im fortgeschrittenen Stadium diagnostiziert (4), weshalb auch bei der Therapie von FL-Patienten ein verlängertes Überleben bei gleichzeitigem Erhalt der Gesundheitsbezogenen Lebensqualität angestrebt wird (34).

Ein fehlendes oder nur kurzes Ansprechen auf eine Therapie geht mit einem kürzeren Gesamtüberleben einher (35, 36). Da die in GADOLIN untersuchte Patientenpopulation dieses Merkmal aufweist, handelt es sich um eine Population mit ungünstiger Prognose. Für die Gesamtpopulation von GADOLIN ist die finale Analyse bzgl. Gesamtüberleben nach 226 aufgetretenen Todesfällen vorgesehen.

Zum Zeitpunkt des für den Endpunkt Gesamtüberleben relevanten Datenschnitts vom 01. Mai 2015 ereigneten sich 23,3 % Todesfälle in der FL-Population von GADOLIN (mediane Beobachtungsdauer G-Benda 24,1 Monate bzw. Benda 24,1 Monate). Bei diesem Datenschnitt zeigte sich für die FL-Patienten bereits ein statistisch signifikanter Unterschied zugunsten von G-Benda gegenüber Benda (HR 0,62; 95 % KI [0,39; 0,98]; $p = 0,0379$). G-Benda reduzierte das Sterberisiko statistisch signifikant um 38 % im Vergleich zu Benda. Im G-Benda-Arm verstarben 18,3 % und im Benda-Arm 28,1 % der Patienten (12).

Somit ist GADOLIN die erste randomisierte, kontrollierte Studie, die für den Endpunkt Gesamtüberleben (OS) einen signifikanten Vorteil gegenüber der Vergleichstherapie für Patienten mit Rituximab-refraktärem FL nachweisen konnte.

Fazit:

Der statistisch signifikante Unterschied zwischen den Behandlungsarmen im Datenschnitt vom 01. Mai 2015 zeigt einen Vorteil für G-Benda. Auf Basis der bestverfügbaren Evidenz besteht nach Definition gemäß VerfO ein beträchtlicher Zusatznutzen für den patientenrelevanten Endpunkt Gesamtüberleben in der Domäne Mortalität.

Morbidität

v Progressionsfreies Überleben (IRC)

Der Endpunkt PFS operationalisiert die Mortalität und das Fortschreiten oder Wiedereintreten der Erkrankung (Progression oder Rezidiv) eines FL-Patienten. Das Erreichen eines möglichst langen krankheits- oder progressionsfreien Überlebens ist neben dem Gesamtüberleben ein Ziel bei der Therapie des FL. Zur Beurteilung des Einzelereignisses Tod können die Ergebnisse zum Mortalitätsendpunkt Gesamtüberleben herangezogen werden. Aus diesem Grund wird an dieser Stelle nicht auf die Komponente Gesamtüberleben eingegangen.

Das progressionsfreie Überleben stellt laut EMA einen patientenrelevanten Endpunkt dar („Prolonged PFS/DFS as such, however, is considered to be of benefit to the patient“) und ist primärer Endpunkt der Studie (37).

Der G-BA hat bei der frühen Nutzenbewertung von Idelalisib für das Anwendungsgebiet „Erstlinientherapie von CLL-Patienten mit 17p-Deletion oder TP53-Mutation, die für eine Chemoimmuntherapie ungeeignet sind“ den Zusatznutzen basierend auf PFS und ORR abgeleitet (38). Voraussetzung für die Betrachtung der deutlichen Effekte bzgl. PFS und ORR und die Ableitung des Zusatznutzens war hierbei zunächst die Anerkennung der Endpunkte PFS und ORR als patientenrelevant. Zudem hat der G-BA PFS und ORR ebenfalls für das Anwendungsgebiet „Patienten mit rezidivierender CLL, für die eine Chemotherapie nicht angezeigt ist“ als patientenrelevant anerkannt: Die deutlichen Effekte bzgl. PFS und ORR stützten den Zusatznutzen für den Endpunkt Gesamtüberleben, für den die Daten als unreif betrachtet wurden (38). Zur Wahrung der Verfahrenskonsistenz im Therapiefeld iNHL ist PFS auch für das hier betrachtete Anwendungsgebiet als patientenrelevant anzusehen. Die Operationalisierung von PFS in GADOLIN ist vergleichbar zur Operationalisierung in der relevanten Idelalisib-Studie (38).

In GADOLIN wurde PFS sowohl durch den Prüfarzt als auch durch ein unabhängiges, verblindetes IRC beurteilt. Das durch das IRC bestimmte PFS stellt aufgrund seiner Verblindung ein objektives Maß dar. Bis zum Zeitpunkt des Datenschnitts vom 01. September 2014 wurden 144 PFS-Ereignisse vom IRC beobachtet: 34,8 % im G-Benda-Arm vs. 54,2 % im Benda-Arm (7).

Das Risiko für ein PFS-Ereignis wird signifikant durch G-Benda um 52 % reduziert (HR 0,48; 95 % KI [0,34; 0,68]; $p < 0,0001$). Der statistisch signifikante Vorteil von G-Benda gegenüber Benda wurde im Datenschnitt vom 01. Mai 2015 bestätigt (HR: 0,47; 95 % KI [0,34; 0,64]; $p < 0,0001$) (12).

Die mediane Dauer bis zu einem PFS-Ereignis (IRC) wurde für G-Benda nicht erreicht und betrug für Benda 13,8 Monate (siehe Tabelle 4-18). Der Median für G-Benda liegt für diesen Datenschnitt nur für PFS (Investigator) vor, beträgt 29,2 Monate und wird im Follow-Up Datenschnitt vom 01. Mai 2015 durch das PFS (IRC) bestätigt (12). Dies bedeutet im Median einen Vorteil von 15,4 Monaten für G-Benda gegenüber Benda.

Die Ergebnisse der einzelnen Datenschnitte zeigen konsistente Behandlungseffekte. Unter Berücksichtigung der Schwere der Erkrankung und des therapeutischen Ziels der Behandlung rezidivierter bzw. refraktärer FL-Patienten stellt die erzielte Halbierung des Risikos für Progression oder Tod (PFS-Ereignis) unter G-Benda einen bisher nicht erreichten deutlichen therapeutischen Effekt in der Zielpopulation dar. Somit eignet sich PFS auch in dieser Untergruppe der iNHL als patientenrelevanter Endpunkt und wird zur Ableitung des Ausmaßes des Zusatznutzens herangezogen.

Fazit:

Aufgrund der statistisch signifikanten deutlichen Verzögerung eines Fortschreitens oder Wiedereintretens der Erkrankung durch G-Benda besteht ein beträchtlicher Zusatznutzen für den Endpunkt PFS (IRC) in der Domäne Morbidität.

v Krankheitsfreies Überleben (IRC)

Der Endpunkt krankheitsfreies Überleben operationalisiert, neben der Mortalität, in der Morbidität das patientenrelevante Wiedereintreten der Erkrankung durch ein Rezidiv bei Patienten, die zuvor unter der Therapie eine komplette Remission und damit völlige Freiheit jeglicher Krankheitszeichen erreicht haben. Die möglichst lange Aufrechterhaltung eines solchen Therapieerfolges ist neben dem Gesamtüberleben anerkanntes und angestrebtes Therapieziel im Sinne des Patienten (39). Ein Effekt hinsichtlich der Rezidivrate, der Zeit bis zum Rezidiv und folglich krankheitsfreies Überleben ist als patientenrelevant zu betrachten.

Rezidive (Morbidität) wurden zum Zeitpunkt des Datenschnitts vom 01. September 2014 bei 6,3 % im G-Benda-Arm gegenüber 45,2 % im Benda-Arm beobachtet (7). Das Risiko ein Rezidiv zu erleiden wird durch G-Benda um 86 % reduziert. Im Benda-Arm rezidierten nach 6 Monaten ein 21,3 % und nach 12 Monaten 47,2 % aller Patienten, die zuvor eine komplette Remission erreicht hatten. Unter G-Benda wurden keine Frührezidive festgestellt (innerhalb von 6 Monaten nach CR) und nach 12 Monaten rezidierten 8,9 % der Patienten. Die mediane Dauer bis zu einem Rezidiv ist für G-Benda nicht erreicht und beträgt für Benda 13,0 Monate (95% KI [6,9, n.e.]) (7, 18). In Abwesenheit eines Medians errechnet sich unter der Annahme stabiler Hazard Ratios (G-Benda vs. Benda 0,10 (95 % KI [0,02; 0,46]; $p = 0,0004$) ein geschätzter Median für die Rezidivfreiheit von G-Benda von 130,0 Monaten.

Gesamthaft besteht für den Endpunkt krankheitsfreies Überleben (Morbidität und Mortalität) ein signifikanter Vorteil für G-Benda (HR 0,1; 95 % KI [0,02; 0,44]; $p = 0,0002$). Die mediane Dauer des krankheitsfreien Überlebens beträgt für Benda 13,0 Monate, für G-Benda unter Annahme stabiler Hazard Ratios errechnet sich in Abwesenheit eines Medians eine Dauer von 94,3 Monaten.

Fazit:

Aufgrund der statistisch signifikanten Senkung der Rezidivrate, Verlängerung der Zeit bis zum Rezidiv, sowie der Reduktion des Risikos ein Rezidiv zu erleiden besteht aufgrund der langfristigen Abwesenheit aller Krankheitszeichen und Symptome ein beträchtlicher Zusatznutzen insgesamt für den Endpunkt Krankheitsfreies Überleben.

Zusammenfassung Morbidität

Aufgrund der statistisch signifikanten Verzögerung eines Fortschreitens oder Wiedereintretens der Erkrankung durch G-Benda um 15,4 Monate sowie der statistisch signifikanten Verlängerung der Zeit ohne jegliche Krankheitszeichen mit Reduktion des Risikos für ein Rezidiv um nahezu 90 % besteht insgesamt ein beträchtlicher Zusatznutzen in der Domäne Morbidität.

Gesundheitsbezogene Lebensqualität

v FACT-Lym

Gesundheitsbezogene Lebensqualität, erhoben mittels allgemeiner oder spezifischer Patientenfragebögen ist stets patientenrelevant, weil die Beeinträchtigung des Patienten allgemein,

funktionell, körperlich, sozial oder in der eigenen Rolle, direkt vom Patienten berichtet, messbar wird. Gerade bei Patienten mit hämatologischen Erkrankungen wird die Lebensqualität zusätzlich durch Krankheitssymptome, sowie durch die Therapie selbst (meist Chemotherapie und/oder Bestrahlung) stark beeinflusst. Therapieziel kann daher neben einer angestrebten Verbesserung der Gesundheitsbezogenen Lebensqualität auch deren *Erhalt* trotz effektiverer Interventionen bei gleichzeitig besserer Wirksamkeit sein.

In GADOLIN wird die Gesundheitsbezogene Lebensqualität systematisch beobachtet und mittels FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma), einem krankheitsspezifischen Messinstrument für Patienten mit Non-Hodgkin-Lymphom, erhoben. Zum Zeitpunkt der beiden Datenschnitte zeigten sich für keine der untersuchten FACT-Lym-Scores statistisch signifikante Unterschiede zwischen G-Benda und Benda.

Mit diesen Daten konnte der Erhalt der Gesundheitsbezogenen Lebensqualität über den dargestellten Erhebungszeitraum hinweg gezeigt werden. Diese Ergebnisse weisen darauf hin, dass auch in dieser Population die Therapie mit G-Benda neben signifikant stärkerer Wirksamkeit nicht zu einer Beeinträchtigung der Gesundheitsbezogenen Lebensqualität führt und sich somit von Benda bezüglich Auswirkungen auf die Gesundheitsbezogene Lebensqualität nicht unterscheidet.

Fazit:

Für den Endpunkt FACT-Lym ergaben sich keine statistisch signifikanten Unterschiede zwischen G-Benda und Benda. Dies entspricht gemäß VerFO keinem Zusatznutzen für G-Benda in der Endpunktkategorie Gesundheitsbezogene Lebensqualität. Die Ergebnisse zeigen jedoch, dass die vom Patienten berichtete Gesundheitsbezogene Lebensqualität bei deutlich besserer Wirksamkeit von G-Benda im Vergleich zu Benda nicht beeinträchtigt war.

Verträglichkeit

v Unerwünschte Ereignisse

Bei der vergleichenden Betrachtung der Sicherheitsprofile beider Arme ist eine Verzerrung zu Ungunsten des G-Benda-Arms festzuhalten. Diese ist begründet in der unterschiedlichen Behandlungsdauer, bedingt durch die zweijährige Erhaltungstherapie im G-Benda-Arm. Zudem wurde für die Mehrzahl der Patienten (150/168) im Benda-Arm keine systematische UE-Erfassung über die Induktionstherapie hinaus durchgeführt. Eine direkte Vergleichbarkeit beider Studienarme ist daher nur für die Induktionstherapie gegeben. Die Betrachtung der gesamten Studienphase wird aufgrund der Verzerrung durch eine Time-to-first-event Analyse ergänzt, die diese Verzerrung reduziert. Dennoch ist das Sicherheitsprofil von G-Benda gegenüber Benda aus statistisch-medizinischer Sicht insgesamt nicht nachteilig.

In GADOLIN ist die Inzidenz der unerwünschten Ereignisse jeglicher Art, UE $\geq$ Grad 3, UE Grad 3, UE Grad 4, UE Grad 5, SUE und Behandlungsabbruch aufgrund UE in beiden Therapiearmen sowohl in der Induktionstherapie als auch in der gesamten Studienphase vergleichbar.

Signifikante Unterschiede zu Ungunsten G-Benda bestanden in der Inzidenz von schwerwiegender und spät einsetzender Neutropenie, kardialen Ereignis und Hypotonie. Wenngleich

schwerwiegende Neutropenien häufiger im G-Benda-Arm auftraten, führten diese nicht zu Unterschieden in den Inzidenzraten klinisch relevanter Neutropenie-assoziiierter Komplikationen, wie febrilen Neutropenien und Infektionen. Todesfälle aufgrund von Infektionen traten in beiden Behandlungsarmen in gleichem Umfang auf. Der Unterschied hinsichtlich kardialer Ereignisse zu Ungunsten von G-Benda erreichte statistische Signifikanz nur in der Betrachtung der gesamten Studienphase. Die kardialen UE im G-Benda-Arm waren bei 16/20 Patienten zum Zeitpunkt des aktuellsten Datenschnitts reversibel und keines zeigte einen fatalen Verlauf. Hypotonien traten ausschließlich im Rahmen von IRR auf, waren alle reversibel und fast ausschließlich UE Grad 1 oder 2 (22/25). Es trat kein Ereignis > Grad 3 auf.

Signifikante Unterschiede zugunsten von G-Benda bestanden in der Inzidenz von Übelkeit, Erbrechen und ungewolltem Gewichtsverlust. Dieser Vorteil hinsichtlich gastrointestinaler Nebenwirkungen bestätigte sich auch in der Time-to-first-event Analyse. Diese wurde aufgrund des Ungleichgewichts bezüglich Behandlungsdauer und systematischer Dokumentation der UE zwischen den Behandlungsarmen durchgeführt. Sie zeigt zudem signifikante Vorteile für G-Benda für UE Grad 3 und 5, SUE, Thrombozytopenie, Infektionen, schwerwiegende Infektionen und Zweitmalignome. Ausschließlich für Hypotonien ergibt sich ein signifikanter Nachteil für G-Benda.

Fazit:

Die in der Effektivität überlegene Therapie mit G-Benda geht im Vergleich zu Benda in der Gesamtschau nicht mit einem Nachteil bezüglich Verträglichkeit einher. Dies entspricht gemäß Verfo keinem Zusatznutzen hinsichtlich des Endpunkts Verträglichkeit.

Gesamtfazit zum Ausmaß des Zusatznutzens von G-Benda gegenüber Benda

Für Patienten mit follikulärem Lymphom, die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden, konnte in GADOLIN gezeigt werden, dass eine Therapie mit G-Benda sowohl das Gesamtüberleben als auch das progressions- und das krankheitsfreie Überleben im Vergleich zu Benda statistisch signifikant verlängert.

Die Gesundheitsbezogene Lebensqualität der Patienten unter G-Benda bleibt trotz Kombinationstherapie in der Induktionstherapie und einer Erhaltungstherapie bei deutlich besserer Wirksamkeit erhalten und vergleichbar mit dem Benda-Arm.

Die Verträglichkeit von G-Benda war trotz deutlich längerer Behandlungsdauer und Kombinationstherapie in der Induktionsphase gegenüber Benda insgesamt nicht nachteilig.

In der Gesamtschau stellen die unter G-Benda erzielten Therapieeffekte eine nachhaltige und bisher im Anwendungsgebiet nicht erreichte Verbesserung des therapielevanten Nutzens im Sinne einer Verlängerung der Überlebensdauer (statistisch signifikante Reduktion der Gesamtmortalität), einer langfristigen krankheits- und progressionsfreien Zeit (mit statistisch signifikanter Reduktion der Rezidiv- und Progressionsrate), vergleichbarer Gesundheitsbezogener Lebensqualität und nicht nachteiliger Verträglichkeit dar. Dies entspricht insgesamt einem beträchtlichen Zusatznutzen von G-Benda gegenüber Benda.

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 4-47: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit follikulärem Lymphom (FL), die auf eine Behandlung mit Rituximab oder einem Rituximab-haltigen Regime nicht angesprochen haben oder während bzw. bis zu 6 Monate nach der Behandlung progredient wurden	Beträchtlich

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 zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte 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 zutreffend.

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

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

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹³, Molenberghs 2010¹⁴). 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 2006¹⁵) 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 2006¹⁶) 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:

¹³ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁴ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

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

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

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

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

Nicht zutreffend.

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 Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Studie	Titel	Datenquellen
GADOLIN	An open-label, multicenter, randomized, Phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin lymphoma	Studienregistereintrag [Clinicaltrials.gov (13); ICTRP WHO (14); EU CTR (15); PharmNet.Bund (16)] Studienbericht e [CSR (7); Update CSR (12); Nachauswertung (18)] Publikationen [Abstract (17)]

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. EUROPEAN PARLIAMENT. REGULATION (EC) No 141/2000 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 1999 on orphan medicinal products. 2000 22.01.2000. In: Official Journal of the European Communities. [L18/1-L/5].
2. Bundesministerium der Justiz und für Verbraucherschutz. Das Fünfte Buch Sozialgesetzbuch – Gesetzliche Krankenversicherung – (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477, 2482), das durch Artikel 2 Absatz 12 des Gesetzes vom 17. Februar 2016 (BGBl. I S. 203) geändert worden ist. 2016 [cited 14.04.2016]. Available from: http://www.gesetze-im-internet.de/bundesrecht/sgb_5/gesamt.pdf.
3. Gemeinsamer Bundesausschuss G-BA. Verfahrensordnung des Gemeinsamen Bundesausschusses in der Fassung vom Dezember 2008, zuletzt geändert am 18. Dezember 2014, veröffentlicht im Bundesanzeiger BAnz AT 15.04.2015 B2, in Kraft getreten am 16. April 2015. 2015.
4. Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V. (DGHO), Ch. Buske, M. Herold, W. Willenbacher, M. Dreyling. DGHO-Leitlinie: Follikuläres Lymphom, Stand Mai 2012. 2012 [cited 21.04.2016]. Available from: <https://www.onkopedia.com/de/onkopedia/guidelines/follikulaeres-lymphom/@@view/html/index.html>.
5. Roche. Fachinformation Gazyvaro® (Obinutuzumab) 1000 mg Konzentrat zur Herstellung einer Infusionslösung; Fachinformation Stand 2016. 2016 [cited 2016].
6. B. D. Cheson, C. M. Wendtner, A. Pieper, et al. Optimal use of bendamustine in chronic lymphocytic leukemia, non-Hodgkin lymphomas, and multiple myeloma: treatment recommendations from an international consensus panel. 2010 Feb. In: Clin Lymphoma Myeloma Leuk. 2010/03/13. [21-7]. DOI: 10.3816/CLML.2010.n.002.
7. F. Hoffmann-La Roche Ltd. Primary Clinical Study Report – GAO4753g/GO01297 – An open-label, multicenter, randomized, Phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin's lymphoma. Clinical Study Report – 1051204. 2015
8. G. Gartlehner, C. G. Moore. Direct versus indirect comparisons: a summary of the evidence. 2008 Spring. In: Int J Technol Assess Health Care. 2008/04/11. [170-7]. DOI: 10.1017/S0266462308080240.

9. A. Sutton, A. E. Ades, N. Cooper, K. Abrams. Use of indirect and mixed treatment comparisons for technology assessment. 2008. In: Pharmacoeconomics. 2008/09/05. [753-67]. DOI: 2696 [pii].
10. F. Song, Y. K. Loke, T. Walsh, A. M. Glenny, A. J. Eastwood, D. G. Altman. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. 2009. In: BMJ. 2009/04/07. [b1147].
11. G. Salanti, V. Marinho, J. P. T. Higgins. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. 2009 Aug. In: J Clin Epidemiol. 2009/01/23. [857-64]. DOI: 10.1016/j.jclinepi.2008.10.001.
12. F. Hoffmann-La Roche Ltd. Update Clinical Study Report – GAO4753g/GO01297 - An open-label, multicenter, randomized, Phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin lymphoma. Report No. 1067639, December 2015. 2015
13. Genentech, Roche Pharma AG. A Study to Investigate the Efficacy and Safety of Bendamustine Compared With Bendamustine+RO5072759 (GA101) in Patients With Rituximab-Refractory, Indolent Non-Hodgkin's Lymphoma (GADOLIN), last verified November 2015. 2015 8. November 2015 [cited 28.04.2016]. Available from: <http://ClinicalTrials.gov/show/NCT01059630>.
14. Genentech. ICTRP NCT01059630; An OpenLabel, Multicenter, Randomized, Phase III Study to Investigate the Efficacy and Safety of Bendamustine Compared With Bendamustine+RO5072759 (GA101) in Patients With RituximabRefractory, Indolent NonHodgkin's Lymphoma; last refreshed 31. August 2015 2015 [updated 31.05.2015; cited 2016 28.04.2016]. Available from: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01059630>.
15. Genentech, Roche Pharma AG. EudraCT_2009-015504-25; An Open Label, Multicenter, randomized, phase III study to investigate the efficacy and safety of Bendamustine compared with Bendamustine + RO5072759 (GA101) in patients with Rituximab-refractory, indolent Non-Hodgkin's Lymphoma 2009 [cited 2016 28.04.2016]. EudraCT_2009-015504-25]. Available from: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number%3A2009-015504-25.
16. Genentech. An Open Label, Multicenter, randomized, phase III study to investigate the efficacy and safety of Bendamustine compared with Bendamustine + RO5072759 (GA101) in patients with Rituximab-refractory, indolent Non-Hodgkin's Lymphoma; Auszug PharmNetBund - Clinical Trials 2010 [28.04.2016].
17. L. H. Sehn, N. S. Chua, J. Mayer, et al. GADOLIN: Primary results from a phase III study of obinutuzumab plus bendamustine compared with bendamustine alone in patients with rituximab-refractory indolent non-Hodgkin lymphoma. (LBA8502). 2015 May 20 [cited 27.04.2016]. In: Journal of Clinical Oncology [Internet]. Available from:

http://meeting.ascopubs.org/cgi/content/abstract/33/15_suppl/LBA8502?sid=8da1cef1-4f1e-415d-82df-69f3804fbf1a.

18. Roche Pharma AG. Nachauswertung zu GADOLIN. 2016
19. B. D. Cheson, B. Pfistner, M. E. Juweid, et al. Revised response criteria for malignant lymphoma. 2007 Feb 10. In: J Clin Oncol. [579-86]. DOI: 10.1200/JCO.2006.09.2403.
20. Fay J. Hlubocky, Kimberly Webster, John Cashy, Jennifer Beaumont, David Cella. The Development and Validation of a Measure of Health-Related Quality of Life for Non-Hodgkin's Lymphoma: The Functional Assessment of Cancer Therapy—Lymphoma (FACT-Lym). 2013. In: Lymphoma. [1-9]. DOI: 10.1155/2013/147176.
21. David Cella, Webster Kimberly, John Cashy, et al. Development and Initial Validation of The FACT-Lymphoma: A Health-related Quality of Life Questionnaire for People with Non-Hodgkin's Lymphoma; Final Report January 31, 2005. 2005 [cited 18.09.2015].
22. K. J. Yost, C. A. Thompson, D. T. Eton, et al. The Functional Assessment of Cancer Therapy - General (FACT-G) is valid for monitoring quality of life in patients with non-Hodgkin lymphoma. 2013 Feb. In: Leuk Lymphoma. [290-7]. DOI: 10.3109/10428194.2012.711830.
23. N. D. Wagner-Johnston, A. K. Gopal, B. S. Kahl, et al. Patient-reported outcomes data from a phase 2 study of idelalisib in patients with refractory indolent B-cell non-Hodgkin lymphoma (iNHL). 2014. In: Journal of clinical oncology.
24. S. Bhatia. Therapy-related myelodysplasia and acute myeloid leukemia. 2013 Dec. In: Semin Oncol. [666-75]. DOI: 10.1053/j.seminoncol.2013.09.013.
25. P. McLaughlin, E. Estey, A. Glassman, et al. Myelodysplasia and acute myeloid leukemia following therapy for indolent lymphoma with fludarabine, mitoxantrone, and dexamethasone (FND) plus rituximab and interferon alpha. 2005 Jun 15. In: Blood. [4573-5]. DOI: 10.1182/blood-2004-08-3035.
26. J. W. Friedberg. Secondary malignancies after therapy of indolent non-Hodgkin's lymphoma. 2008 Mar. In: Haematologica. [336-8]. DOI: 10.3324/haematol.12585.
27. M. J. Rummel, N. Niederle, G. Maschmeyer, et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. 2013 Apr 6. In: Lancet. 2013/02/26. [1203-10]. DOI: 10.1016/S0140-6736(12)61763-2.
28. M. H. van Oers, M. Van Glabbeke, L. Giurgea, et al. Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. 2010 Jun 10. In: J Clin Oncol. [2853-8]. DOI: 10.1200/JCO.2009.26.5827.
29. M. S. Czuczman, L. Fayad, V. Delwail, et al. Ofatumumab monotherapy in rituximab-refractory follicular lymphoma: results from a multicenter study. 2012 Apr 19. In: Blood. 2012/03/06. [3698-704]. DOI: 10.1182/blood-2011-09-378323.
30. J. W. Friedberg, P. Cohen, L. Chen, et al. Bendamustine in patients with rituximab-refractory indolent and transformed non-Hodgkin's lymphoma: results from a phase II

- multicenter, single-agent study. 2008 Jan 10. In: J Clin Oncol. [204-10]. DOI: 10.1200/JCO.2007.12.5070.
31. A. K. Gopal, B. S. Kahl, S. de Vos, et al. PI3Kdelta inhibition by idelalisib in patients with relapsed indolent lymphoma. 2014 Mar 13. In: N Engl J Med. [1008-18]. DOI: 10.1056/NEJMoa1314583.
 32. B. S. Kahl, N. L. Bartlett, J. P. Leonard, et al. Bendamustine is effective therapy in patients with rituximab-refractory, indolent B-cell non-Hodgkin lymphoma: results from a Multicenter Study. 2010 Jan 1. In: Cancer. 2009/11/06. [106-14]. DOI: 10.1002/cncr.24714.
 33. T. E. Witzig, L. I. Gordon, F. Cabanillas, et al. Randomized controlled trial of yttrium-90-labeled ibritumomab tiuxetan radioimmunotherapy versus rituximab immunotherapy for patients with relapsed or refractory low-grade, follicular, or transformed B-cell non-Hodgkin's lymphoma. 2002 May 15. In: J Clin Oncol. [2453-63].
 34. C. McNamara, J. Davies, M. Dyer, et al. Guidelines on the investigation and management of follicular lymphoma. 2012 Feb. In: Br J Haematol. [446-67]. DOI: 10.1111/j.1365-2141.2011.08969.x.
 35. C. Casulo, M. Byrtek, K. L. Dawson, et al. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. 2015 Aug 10. In: J Clin Oncol. [2516-22]. DOI: 10.1200/JCO.2014.59.7534.
 36. S. Montoto, A. Lopez-Guillermo, A. Ferrer, et al. Survival after progression in patients with follicular lymphoma: analysis of prognostic factors. 2002 Apr. In: Ann Oncol. [523-30].
 37. EMA European Medicines Agency. Guideline on the evaluation products in man; 13 December 2012; EMA/CHMP/205/95/Rev.4; Oncology Working Party. 2013 01.07.2013 [cited 28.11.2014]. [cited 28.11.2014].
 38. Gemeinsamer Bundesausschuss G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Idelalisib. 2015 [cited 04.04.2016]. [cited 04.04.2016].
 39. Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V. (DGHO). Gesundheitspolitische Schriftenreihe der DGHO, Band 2: Nutzenbewertung von Arzneimitteln der Onkologie und Hämatologie 2013. Eine Standortbestimmung. Gesundheitspolitische Schriftenreihe der DGHO ed2013 August 2013.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. 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. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁷ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht zutreffend.

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. 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. 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

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht zutreffend.

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht zutreffend.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht zutreffend.

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: 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 4-48 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 4-48 inklusive eines Flow-Charts für den Patientenfluss.

Ausführliche Angaben zur in die vorliegende Nutzenbewertung eingeschlossenen GADOLIN-Studie können den zugehörigen Studienberichten entnommen werden.

Tabelle 4-48 (Anhang): Studiendesign und -methodik für Studie

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	
4b	Studienorganisation und Ort der Studiendurchführung	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	
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	
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	
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	
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	

Item ^a	Charakteristikum	Studieninformation
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?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

Ausführliche Angaben zur in die vorliegende Nutzenbewertung eingeschlossenen GADOLIN-Studie können den zugehörigen Studienberichten entnommen werden.

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte 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:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. 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 sollen die Bögen studienbezogen ausgefüllt werden. 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. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

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 4-49 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GADOLIN

Studie: GADOLIN**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

Genaue Benennung der Quelle	Referenz
Primary Clinical Study Report – GAO4753g/GO01297 – An open-label, multicenter, randomized, phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin's lymphoma. Clinical Study Report – 1051204.	(7)
Clinical Study Protocol - An open-label, multicenter, randomized, phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin's lymphoma. Protocol Number GAO4753g	
Statistical Analysis Plan - An open-label, multicenter, randomized, phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin's lymphoma. Protocol number - GAO4753g/GO01297.	
Update Clinical Study Report – GAO4753g/GO01297 - An open-label, multicenter, randomized, Phase III study to investigate the efficacy and safety of bendamustine compared with bendamustine + RO5072759 (GA101) in patients with rituximab-refractory, indolent non-Hodgkin lymphoma. Report No. 1067639,	(12)
Nachauswertungen für das vorliegende Dossier	(18)

A Verzerrungsaspekte auf Studienebene:**Einstufung als randomisierte Studie**

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Nach Überprüfung der Eignung der Patienten in einer Screening-Phase erfolgt eine randomisierte Zuteilung der Patienten zu einem der zwei Behandlungsarme (Benda oder G-Benda) im Verhältnis 1:1 mittels eines IVRS (interactive voice response system).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

Es wurde eine zentrale Randomisierung im Verhältnis 1:1 mittels eines IVRS mit Stratifizierung nach folgenden Faktoren durchgeführt:

- 1) Subtyp des iNHL (follikulär vs. andere);
- 2) Refraktärität (refraktär gegenüber einer Rituximab-Monotherapie vs. refraktär gegenüber einer Kombinationstherapie aus Rituximab + Chemotherapie);
- 3) Anzahl der vorangegangenen Therapien (≤ 2 vs. > 2);
- 4) geographische Region (Osteuropa und Naher Osten / Westeuropa / Nord und Zentral Amerika und Karibik).

Die Randomisierung erfolgt unter Verwendung eines hierarchischen dynamischen Randomisierungsschemas mit drei Ebenen: auf der ersten Ebene wurde eine vergleichbare Anzahl Patienten zwischen den Studienarmen sichergestellt; auf der zweiten Ebene wurde eine vergleichbare Anzahl Patienten zwischen den acht o. g. Stratifizierungsgruppen gewährleistet. Wenn ein Gleichgewicht auf den ersten beiden Ebenen erreicht wurde, wurden die Studienarme nach geographischer Region ausbalanciert.

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ ja ☐ unklar ☐ nein

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

2.

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

☒ ja ☐ unklar ☐ nein

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

Die Randomisierung wurde zentral unter Verwendung eines IVRS durchgeführt. Der Prüfarzt hatte somit keine Kenntnis über die Randomisierungsliste und keinen Zugriff darauf.

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

☐ ja ☐ unklar ☐ nein

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

3. Verblindung von Patienten und behandelnden Personen**Patient:**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde unverblindet durchgeführt. Sowohl Patienten als auch Prüfarzte sind unverblindet gegenüber der zugewiesenen Intervention.

behandelnde bzw. weiterbehandelnde Personen:

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde unverblindet durchgeführt. Sowohl Patienten als auch Prüfarzte sind unverblindet gegenüber der zugewiesenen Intervention.

Eine Verblindung des Prüfarztes oder der Patienten wurde aufgrund der erheblichen Belastung für den Patienten durch weitere Infusionen (Antikörper-Placebo, Placebo-Prä-Medikation) nicht durchge-

führt. Zudem ist aufgrund des bekannten Nebenwirkungsprofils von sowohl Verum und Vergleichstherapie als auch der Prä-Medikation von einer Selbstentblindung des Prüfarztes auszugehen. Für die Beurteilung der Progression wurde neben der unverblindeten Erhebung durch den Prüfarzt eine verblindete Beurteilung durch ein unabhängiges Bewertungsgremium (IRC) durchgeführt.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

Es gibt keine Anhaltspunkte für eine ergebnisgesteuerte Darstellung der Studiendaten im Studienbericht. Die im Studienbericht dargestellten Analysen der im Dossier betrachteten Endpunkte folgen der im statistischen Analyseplan bzw. Studienprotokoll a priori festgelegten Auswertungsstrategie.

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☐ ja ☒ nein

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

Die Bewertung des Zusatznutzens im vorliegenden Dossier basiert auf dem Vergleich einer Teilpopulation der GADOLIN, der Patienten mit FL. Basis sind die der Patienten mit follikulärem Non-Hodgkin-Lymphom. Diese Beschränkung auf einen Teil der Studienpopulation stellt keine Verletzung des ITT-Prinzips dar, da durch den Stratifizierungsfaktor „Subtyp des indolenten Non-Hodgkin-Lymphoms“ die Gleichverteilung der FL-Patienten auf die Behandlungsarme sichergestellt war, ergibt sich durch die Reduktion auch keine Imbalance zwischen den Behandlungsarmen. Gleichzeitig wird die Einschränkung der Studienpopulation nicht aus studienimmanenten Gründen vorgenommen, sondern aufgrund einer Indikationsstellung der Zulassungsbehörde. Ein Einfluss auf das Verzerrungspotential wird daher ausgeschlossen.

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Die Randomisierung wurde zentral unter Verwendung eines IVRS durchgeführt. Es handelte sich um eine offene Studie. Durch die fehlende Verblindung besteht grundsätzlich ein Verzerrungspotential, das je Endpunkt unterschiedlich ist und sich deshalb für die dargestellten Ergebnisse dort niederschlägt. Auch wird die fehlende Verblindung durch weitere Maßnahmen wie z.B. die verblindete Bewertung des progressionsfreien Überlebens oder Sensitivitätsanalysen auf Endpunktebene teilweise kompensiert. Das Verzerrungspotential ist aus diesem Grund endpunktspezifisch zu betrachten. In den Studienunterlagen finden sich weder Hinweise auf ergebnisabhängige Berichterstattung der relevanten Endpunkte noch Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial beeinflussen. Daher wird das Verzerrungspotenzial auf Studienebene als „niedrig“ eingestuft.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**Endpunkt: Gesamtüberleben****1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde unverblindet durchgeführt. Sowohl Patienten als auch Prüfarzte sind unverblindet gegenüber der zugewiesenen Intervention.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Analyse des Gesamtüberlebens folgt dem ITT-Prinzip (alle randomisierten FL-Patienten gingen in die Analyse ein) mit adäquaten Zensierungsverfahren.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung. Die im Studienbericht dargestellten Ergebnisse folgen den im statistischen Analyseplan spezifizierten Analysen. Die Ergebnisse des Datenschnitts vom 01. Mai 2015 stellen durch die längere Beobachtungsdauer belastbarere Daten in Form der best verfügbaren Evidenz dar.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ ja ☐ nein

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

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu Verzerrungen führen können.

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

☒ niedrig ☐ hoch

Begründung für die Einstufung

Es konnten keine verzerrungsrelevanten Punkte identifiziert werden, daher wird das Verzerrungspotenzial als „niedrig“ eingestuft.

Endpunkt: PFS (IRC)**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gutachter des unabhängigen Bewertungsgremiums (IRC) sind gegenüber der verabreichten Therapie und der Bewertung des Prüfarztes verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Analyse des Endpunkts folgt dem ITT-Prinzip (alle randomisierten FL-Patienten gingen in die Analyse ein) mit adäquaten Zensierungsverfahren.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung. Die im Studienbericht dargestellten Ergebnisse folgen den im statistischen Analyseplan spezifizierten Analysen.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ ja ☐ nein

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

Der Endpunkt wurde durch das IRC verblindet bewertet. Die Mitglieder des unabhängigen Bewertungsgremiums (IRC) waren gegenüber der verabreichten Therapie und der Beurteilung des Prüfarztes verblindet. Die CT-Aufnahmen wurden verblindet und unabhängig von den Befunden der Prüfarzte bewertet.

Es besteht ein mögliches Verzerrungspotential bzgl. Diagnostik und Progressionsbeurteilung, da die Beurteilung des Onkologen neben der Beurteilung der verblindeten CT-Aufnahmen auch auf Basis der potentiell verzerrten Befunde des Prüfarztes erfolgen konnte. Der Einfluss auf das Verzerrungspotential wird wegen dieses sehr geringen subjektiven Anteils mit niedrig beurteilt.

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Es konnten keine verzerrungsrelevanten Punkte identifiziert werden, daher wird das Verzerrungspotenzial für diesen Endpunkt als „niedrig“ eingestuft.

Endpunkt: DFS (IRC)**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gutachter des unabhängigen Bewertungsgremiums (IRC) sind gegenüber der verabreichten Therapie und der Bewertung des Prüfarztes verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☒ nein

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

Die Analyse zur krankheitsfreien Zeit und der dazugehörigen Rezidivrate erfolgt in der Subgruppe der Patienten, die während der Studie eine komplette Remission als beste Response erreicht hatten. Es wurde nicht hiernach randomisiert.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung. Die im Studienbericht dargestellten Ergebnisse folgen den im statistischen Analyseplan spezifizierten Analysen.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ ja ☐ nein

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

Der Endpunkt wurde durch das IRC verblindet bewertet. Die Mitglieder des unabhängigen Bewertungsgremiums (IRC) waren gegenüber der verabreichten Therapie und der Beurteilung des Prüfarztes verblindet. Die CT-Aufnahmen wurden verblindet und unabhängig von den Befunden der Prüfarzte bewertet.

Es besteht ein mögliches Verzerrungspotential bzgl. Diagnostik und Rezidivbeurteilung, da die Beurteilung des Onkologen neben der Beurteilung der verblindeten CT-Aufnahmen auch auf Basis der potentiell verzerrten Befunde des Prüfarztes erfolgen konnte. Der Einfluss auf das Verzerrungspotential wird wegen dieses sehr geringen subjektiven Anteils mit niedrig beurteilt.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Die Analyse des Endpunktes folgte nicht dem ITT-Prinzip. Insgesamt wird daher das Verzerrungspotenzial für diesen Endpunkt als „hoch“ eingestuft.

Endpunkt: FACT-Lym**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde unverblindet durchgeführt. Sowohl Patienten als auch Prüfer sind unverblindet gegenüber der erhaltenen Intervention.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Analyse folgt dem ITT-Prinzip.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung. Die im Studienbericht dargestellten Ergebnisse folgen den im statistischen Analyseplan spezifizierten Analysen.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ ja ☐ nein

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

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu Verzerrungen führen können.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Aufgrund des offenen Studiendesigns kann eine relevante Verzerrung nicht ausgeschlossen werden. Insgesamt wird daher das Verzerrungspotenzial für diesen Endpunkt als „hoch“ eingestuft.

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

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde unverblindet durchgeführt. Sowohl Patienten als auch Prüfarzte sind unverblindet gegenüber der erhaltenen Intervention.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Analyse der Verträglichkeit erfolgt standardmäßig auf der Population aller Patienten mit mindestens einer Prüfmedikationsgabe. Diese werden nicht „as randomized“ sondern „as treated“ ausgewertet.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung. Die im Studienbericht dargestellten Ergebnisse folgen den im statistischen Analyseplan spezifizierten Analysen.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☐ ja ☒ nein

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

Folgende Aspekte könnten zu einer Verzerrung der Ergebnisse in der gesamten Studienphase beigetragen haben:

- Unterschiedliche Behandlungsdauer zwischen den beiden Studienarmen (Benda-Arm: 6 Monate, entsprechend 6 Zyklen Induktionstherapie mit Bendamustin; G-Benda-Arm: 2,5 Jahre (30 Monate), entsprechend 6 Zyklen Induktionstherapie mit Obinutuzumab + Bendamustin gefolgt von einer Obinutuzumab-Monotherapie für bis zu 2 Jahre);
- bis Version 8 des Studienprotokolls fand im Benda-Arm keine systematische Erfassung von UE im Anschluss an die Induktionstherapie statt (150 Patienten mit Dokumentation bis Tag 28 nach Induktionstherapie; 18 Patienten mit Dokumentation im Follow-Up) (das heißt von Randomisierung April 2010 bis Oktober 2013).

Die oben genannten Aspekte könnten zu einer Ergebnisverzerrung geführt haben, da u. U. weniger Daten zu UE im Benda-Arm in der gesamten Studienphase erfasst wurden. Dies könnte wiederum zu einer Verzerrung zu Ungunsten des Obinutuzumab-Bendamustin-Studienarms in der gesamten Studienphase geführt haben.

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Aufgrund der sonstigen Aspekte und der damit verbundenen Verzerrung in der gesamten Studienphase zu Ungunsten von G-Benda wird daher das Verzerrungspotenzial für diesen Endpunkt als „hoch“ eingestuft.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

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

A Verzerrungsaspekte 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 Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgt). 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:

1.

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

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

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

☐ **nein:** Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

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

☐ **ja:** Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.☐ **nein:** Die Gruppenzuteilung erfolgt nicht verdeckt.

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

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgt ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen 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 behandelnden Personen**Patient:**☐ **ja:** Die Patienten sind verblindet.☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet sind.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. 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 behandelnden Personen nicht verblindet sind.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

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 Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht

zwangsläufig zu einer Verzerrung 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 dies nicht zur Verzerrung anderer Endpunkte führt.

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

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

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

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

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

☐ **ja**

☐ **nein**

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

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

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

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

☐ **hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene 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: _____

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 erfolgt.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ☐ **ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-Follow-Up Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-Follow-Up Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
 - Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-Follow-Up Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).
- ☐ **unklar:** Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-Follow-Up Patienten nicht einschätzbar.

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- ☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- ☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ☐ **ja**
- ☐ **nein**

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgt, 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.

Begründung für die Einstufung:
