

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Ruxolitinib (Jakavi®)

Novartis Pharma GmbH

Modul 4 A

*Behandlung von krankheitsbedingter Splenomegalie oder Symptomen bei
Erwachsenen mit primärer Myelofibrose (auch bekannt als chronische
idiopathische Myelofibrose), Post-Polycythaemia-vera-Myelofibrose oder Post-
Essentieller-Thrombozythämie-Myelofibrose*

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

Inhaltsverzeichnis

	Seite
Abbildungsverzeichnis	8
Abkürzungsverzeichnis.....	10
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik	20
4.2.1 Fragestellung	20
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	21
4.2.3 Informationsbeschaffung	23
4.2.3.1 Studien des pharmazeutischen Unternehmers	23
4.2.3.2 Bibliografische Literaturrecherche	24
4.2.3.3 Suche in Studienregistern	25
4.2.3.4 Selektion relevanter Studien	26
4.2.4 Bewertung der Aussagekraft der Nachweise.....	26
4.2.5 Informationssynthese und -analyse	28
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	28
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	28
4.2.5.3 Meta-Analysen.....	36
4.2.5.4 Sensitivitätsanalysen	37
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	37
4.2.5.6 Indirekte Vergleiche	40
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	42
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	42
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	42
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	42
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	44
4.3.1.1.3 Studien aus der Suche in Studienregistern	45
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	47
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	48
4.3.1.2.1 Studiendesign und Studienpopulationen	48
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	62
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	64
4.3.1.3.1 Endpunkte – RCT.....	65
4.3.1.3.1.1 Reduktion der Milzgröße – RCT	65
4.3.1.3.1.2 Linderung der Krankheitssymptomatik – RCT	79
4.3.1.3.1.3 Gesamtüberleben – RCT	91
4.3.1.3.1.4 Leukämiefreies Überleben – RCT	99
4.3.1.3.1.5 Progressionsfreies Überleben – RCT	106
4.3.1.3.1.6 Veränderung des ECOG-Performance-Status – RCT	112
4.3.1.3.1.7 Gesundheitsbezogene Lebensqualität – RCT	120

4.3.1.3.1.8	Verträglichkeit – RCT	134
4.3.1.3.2	Subgruppenanalysen – RCT.....	149
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	162
4.3.2	Weitere Unterlagen.....	166
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	166
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	166
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	166
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	167
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	167
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	169
4.3.2.2	Nicht randomisierte vergleichende Studien	170
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	170
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	170
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	171
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	171
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	172
4.3.2.3	Weitere Untersuchungen.....	173
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	173
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	173
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	173
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	173
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	174
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	174
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	176
4.4.1	Beurteilung der Aussagekraft der Nachweise	176
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	178
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	183
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	184
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	184
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	184
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	184
4.5.4	Verwendung von Surrogatendpunkten	184
4.6	Liste der eingeschlossenen Studien.....	186
4.7	Referenzliste.....	189
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		197
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		202
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		205
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		206

Anhang 4-E : Methodik der eingeschlossenen Studien – RCT	217
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	241

Tabellenverzeichnis

	Seite
Tabelle 4-1: Ein- und Ausschlusskriterien für die Auswahl von Studien für die Nutzenbewertung von Ruxolitinib	15
Tabelle 4-2: Ein- und Ausschlusskriterien für die Auswahl von Studien für die Nutzenbewertung von Ruxolitinib	23
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	43
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	44
Tabelle 4-5: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	46
Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	48
Tabelle 4-7: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	49
Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	51
Tabelle 4-9: Auswahl häufiger Begleitmedikamente mit möglichem Bezug zur Behandlung von Myelofibrose-Symptomen in der Studie COMFORT-I.....	52
Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	62
Tabelle 4-12: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-13: Operationalisierung von „Reduktion der Milzgröße“	66
Tabelle 4-14: Bewertung des Verzerrungspotenzials für den Endpunkt „Reduktion der Milzgröße“ in RCT mit dem zu bewertenden Arzneimittel	68
Tabelle 4-15: Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel	70
Tabelle 4-16: Operationalisierung von „Linderung der Krankheitssymptomatik“	80
Tabelle 4-17: Bewertung des Verzerrungspotenzials für den Endpunkt „Linderung der Krankheitssymptomatik“ in RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-18: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-19: Operationalisierung von „Gesamtüberleben“	92
Tabelle 4-20: Bewertung des Verzerrungspotenzials für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel.....	92
Tabelle 4-21: Ergebnisse für den Endpunkt „Gesamtüberleben“ aus RCT mit dem zu bewertenden Arzneimittel	94

Tabelle 4-22: Operationalisierung von „Leukämiefreies Überleben“	100
Tabelle 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt „Leukämiefreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-24: Ergebnisse für den Endpunkt „Leukämiefreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel.....	102
Tabelle 4-25: Operationalisierung von „Progressionsfreies Überleben“	107
Tabelle 4-26: Bewertung des Verzerrungspotenzials für den Endpunkt „Progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-27: Ergebnisse für den Endpunkt „Progressionsfreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel.....	109
Tabelle 4-28: Operationalisierung von „Veränderung des ECOG-Performance-Status“	113
Tabelle 4-29: Bewertung des Verzerrungspotenzials für den Endpunkt „Veränderung des ECOG-Performance-Status“ in RCT mit dem zu bewertenden Arzneimittel.....	114
Tabelle 4-30: Ergebnisse für den Endpunkt „Veränderung des ECOG-Performance-Status“ aus RCT mit dem zu bewertenden Arzneimittel.....	115
Tabelle 4-31: Operationalisierung von „Gesundheitsbezogene Lebensqualität“	121
Tabelle 4-32: Bewertung des Verzerrungspotenzials für den Endpunkt „Gesundheitsbezogene Lebensqualität“ in RCT mit dem zu bewertenden Arzneimittel	122
Tabelle 4-33: Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel: EORTC-QLQ-C30.....	123
Tabelle 4-34: Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel: FACT-Lym	128
Tabelle 4-35: Klinische Relevanzschwellen für die Skalen des Fragebogens FACT-Lym ..	132
Tabelle 4-36: Operationalisierung von „Verträglichkeit“	135
Tabelle 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt „Verträglichkeit“ in RCT mit dem zu bewertenden Arzneimittel	136
Tabelle 4-38: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel	138
Tabelle 4-39: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse (Inzidenz $\geq 10\%$)	138
Tabelle 4-40: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse adjustiert nach Patientenjahren	140
Tabelle 4-41: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: unerwünschte Ereignisse von besonderem Interesse unter der Behandlung mit Ruxolitinib im Zeitverlauf in der Studie COMFORT-II	141
Tabelle 4-42: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse unter der Behandlung mit Ruxolitinib im Zeitverlauf in der Studie COMFORT-I	142
Tabelle 4-43: Subgruppenanalysen für „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen“ aus der Studie COMFORT-II	149

Tabelle 4-44: Subgruppenanalysen für „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen“ aus der Studie COMFORT-I.....	152
Tabelle 4-45: Subgruppenanalysen für „Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen“ aus der Studie COMFORT-I.....	156
Tabelle 4-46: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	167
Tabelle 4-47: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	167
Tabelle 4-48: Operationalisierung von <Endpunkt xxx>.....	168
Tabelle 4-49: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	168
Tabelle 4-50: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	169
Tabelle 4-51: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	171
Tabelle 4-52: Operationalisierung von <Endpunkt xxx>.....	171
Tabelle 4-53: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	172
Tabelle 4-54: Operationalisierung von <Endpunkt xxx> - weitere Untersuchungen	174
Tabelle 4-55: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	183
Tabelle 4-56: Bibliographische Literaturrecherche in MEDLINE	198
Tabelle 4-57: Bibliographische Literaturrecherche in EMBASE	199
Tabelle 4-58: Bibliographische Literaturrecherche in der Cochrane-Library: Cochrane Database of Systematic Reviews.....	200
Tabelle 4-59: Bibliographische Literaturrecherche in der Cochrane-Library: Cochrane Central Register of Controlled Trials	201
Tabelle 4-60: Suche im Studienregister ClinicalTrials.gov	202
Tabelle 4-61: Suche im Studienregister Klinische Prüfungen PharmNet.Bund	203
Tabelle 4-62: Suche im Studienregister EU Clinical Trials Register.....	203
Tabelle 4-63: Suche im ICTRP Search Portal.....	203
Tabelle 4-64 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister ClinicalTrials.gov.....	206
Tabelle 4-65 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister Klinische Prüfungen PharmNet.Bund	209
Tabelle 4-66 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister EU Clinical Trials Register.....	210
Tabelle 4-67 (Anhang): Liste der ausgeschlossenen Studien aus dem ICTRP Search Portal	212
Tabelle 4-68 (Anhang): Studiendesign und -methodik für Studie COMFORT-II (CINC424A2352).....	218

Tabelle 4-69 (Anhang): Studiendesign und -methodik für Studie COMFORT-I (INCB 18424-351)	230
Tabelle 4-70 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie COMFORT-II	242
Tabelle 4-71 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie COMFORT-I.....	266

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Zusammenhang zwischen Reduktion der Milzlänge und Gesamtüberleben..	31
Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	45
Abbildung 4-3: Design der Zulassungsstudie COMFORT-II	55
Abbildung 4-4: Design der Zulassungsstudie COMFORT-I	58
Abbildung 4-5: Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel: Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn	73
Abbildung 4-6: Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel: Dauer des Ansprechens.....	74
Abbildung 4-7: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel: Veränderung des TSS gegenüber Studienbeginn.....	87
Abbildung 4-8: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel: Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn	88
Abbildung 4-9: Ergebnisse für den Endpunkt „Gesamtüberleben“ aus RCT mit dem zu bewertenden Arzneimittel	95
Abbildung 4-10: Ergebnisse für den Endpunkt „Leukämiefreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel.....	103
Abbildung 4-11: Ergebnisse für den Endpunkt „Progressionsfreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel	110
Abbildung 4-12: Ergebnisse für den Endpunkt „Veränderung des ECOG-Performance-Status“ aus RCT mit dem zu bewertenden Arzneimittel.....	117
Abbildung 4-13: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: Hämoglobin-Konzentration im Studienverlauf	143
Abbildung 4-14: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: Thrombozytenzahl im Studienverlauf	144
Abbildung 4-15: Flow-Chart zum Patientenfluss in der COMFORT-II-Studie (primäre Datenanalyse vom 4. Januar 2011)	228
Abbildung 4-16: Flow-Chart zum Patientenfluss in der COMFORT-II-Studie (3-Jahres-Update vom 1. Dezember 2012)	229
Abbildung 4-17: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (primäre Datenanalyse vom 2. November 2010)	239
Abbildung 4-18: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (2-Jahres-Update vom 1. März 2012)	240
Abbildung 4-19: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (3-Jahres-Update vom 25. Januar 2013)	240

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALT	Alaninaminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANC	Absolute Neutrophilenzahl
ASH	American Society of Hematology
BAT	Best Available Therapy
bid	Bis in die (lat., zweimal täglich)
BSC	Best Supportive Care
CMH	Cochran-Mantel-Haenszel
COMFORT	Controlled Myelofibrosis Study with Oral JAK-Inhibitor Therapy
CONSORT	Consolidated Standards of Reporting Trials
CSR	Clinical Study Report
CT	Computertomographie
CTCAE	Common Terminology Criteria for Adverse Events
DIMDI	Deutsches Institut für Medizinische Dokumentation
DSMB	Data Safety Monitoring Board
ECOG	Eastern Cooperative Oncology Group
EG	Europäische Gemeinschaft
EHA	European Hematology Association
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EORTC-QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
EPAR	European Public Assessment Report
EU-CTR	EU Clinical Trials Register
EWB	Emotional Well-Being
FACT-Lym	Functional Assessment of Cancer Therapy – Lymphoma
FDA	Food and Drug Administration
FWB	Functional Well-Being
G-BA	Gemeinsamer Bundesausschuss
G-CSF	Granulocyte colony-stimulating factor
HR	Hazard ratio

ICTRP	International Clinical Trials Registry
IPSS	International Prognostic Scoring System
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive Voice Response System
IWG-MRT	International Working Group for Myelofibrosis Research and Treatment
JAK	Januskinase
k. A.	Keine Angabe
KI	Konfidenzintervall
LFS	Leukemia-free survival
LSM	Least Square Means
LymS	Lymphoma Subscale
MedDRA	Medical Dictionary for Regulatory Activities
MFSAF	Myelofibrosis Symptom Assessment Form
MID	Minimal important difference
MRT	Magnetresonanztomographie
MTC	Mixed Treatment Comparison
NSAID	Non-steroidal anti-inflammatory drug
OR	Odds ratio
OS	Overall survival
PET-MF	Post-Essentielle-Thrombozythämie-Myelofibrose
PFS	Progression-free survival
PGIC	Patient Global Impression of Change
PMF	Primäre Myelofibrose
PP	Per protocol
PPV-MF	Post-Polycythaemia-vera-Myelofibrose
PT	Preferred Term
PWB	Physical Well-Being
RCT	Randomized Controlled Trial
RR	Risk ratio
SD	Standardabweichung
SGB	Sozialgesetzbuch

SMQ	Standardized MedDRA Query
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
SWB	Social/Family Well-Being
TOI	Trial Outcome Index
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TSS	Total Symptom Score
UE	Unerwünschtes Ereignis
ULN	Upper limit of normal
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

Ziel der vorliegenden Untersuchung ist es, Ausmaß und Wahrscheinlichkeit des medizinischen Zusatznutzens von Ruxolitinib (Jakavi®) im Vergleich zur aktuell bestmöglichen, patientenindividuell optimierten, unterstützenden Behandlung (Best Supportive Care, BSC) bei der Behandlung von krankheitsbedingter Splenomegalie oder Symptomen bei Erwachsenen mit primärer Myelofibrose (PMF), Post-Polycythaemia-vera-Myelofibrose (PPV-MF) oder Post-Essentieller-Thrombozythämie-Myelofibrose (PET-MF) auf Grundlage der besten verfügbaren Evidenz zu ermitteln. Die Bewertung soll hinsichtlich der patientenrelevanten Endpunkte *Mortalität* (Gesamtüberleben), *Morbidität* (Reduktion der Milzgröße, Linderung der Krankheitssymptomatik, leukämiefreies Überleben, progressionsfreies Überleben, Veränderung des ECOG-Performance-Status), *gesundheitsbezogene Lebensqualität* (EORTC-QLQ-C30, FACT-Lym) und *Verträglichkeit* erfolgen.

Datenquellen

Die Bewertung wurde auf Grundlage randomisierter klinischer Studien vorgenommen. Hierzu wurden alle Studien zu Ruxolitinib im vorliegenden Anwendungsgebiet gelistet, die an die Zulassungsbehörden übermittelt wurden bzw. an denen Novartis finanziell beteiligt ist oder war. Um einen möglichst vollständigen Studienpool für die Nutzenbewertung von Ruxolitinib zu generieren, wurde außerdem eine systematische Literaturrecherche in den Datenbanken MEDLINE, EMBASE und der Cochrane Library durchgeführt. Zusätzlich wurde in den Suchregistern ClinicalTrials.gov, EU Clinical Trials Register (EU-CTR) und Klinische Prüfungen PharmNet.Bund sowie über das International Clinical Trials Registry Platform (ICTRP) Search Portal nach laufenden, abgeschlossenen und abgebrochenen Studien mit Ruxolitinib im Indikationsgebiet „Myelofibrose“ gesucht. Die Auswahl der relevanten Studien wurde anhand von prädefinierten Selektionskriterien durch zwei Reviewer unabhängig voneinander vorgenommen.

Ein-/Ausschlusskriterien für Studien

Die Selektion von Studien für die Nutzenbewertung von Ruxolitinib basierte auf vorab definierten Ein- und Ausschlusskriterien. Eine Übersicht über diese Kriterien ist in Tabelle 4-1 dargestellt.

Tabelle 4-1: Ein- und Ausschlusskriterien für die Auswahl von Studien für die Nutzenbewertung von Ruxolitinib

Einschlusskriterien		
E1	Patientenpopulation	Erwachsene Patienten mit krankheitsbedingter Splenomegalie oder Symptomen bei primärer Myelofibrose (PMF), Post-Polycythaemia-vera-Myelofibrose (PPV-MF) oder Post-Essentieller-Thrombozythämie-Myelofibrose (PET-MF)
E2	Intervention	Ruxolitinib (Jakavi®)
E3	Vergleichstherapie	Best Supportive Care (BSC)
E4	Endpunkte	Mindestens einer der folgenden Endpunkte: • Gesamtüberleben • Reduktion der Milzgröße • Linderung der Krankheitssymptomatik • Leukämiefreies Überleben • Progressionsfreies Überleben • Veränderung des ECOG-Performance-Status • Gesundheitsbezogene Lebensqualität (EORTC-QLQ-C30, FACT-Lym) • Verträglichkeit
E5	Studientyp	Randomisierte kontrollierte Studie (RCT)
Ausschlusskriterien		
A1	Studien, für die weder eine Volltext-Publikation noch ein Studienbericht oder ein ausführlicher Ergebnisbericht in einem Studienregister verfügbar sind oder für die noch keine Ergebnisse vorliegen	
A2	Review, Übersichtsartikel, Kommentar, Survey, Note, Editorial, Letter, Erratum	
A3	Tierexperimentelle Studien, In-vitro-Untersuchungen	

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

Da sie die höchste Ergebnissicherheit aufweisen, wurden ausschließlich randomisierte kontrollierte Studien in die vorliegende Nutzenbewertung eingeschlossen (Evidenzstufe 1b). Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise wurde außerdem das Verzerrungspotential der relevanten Studien beurteilt. Entsprechend den Vorgaben der Verfahrensordnung des Gemeinsamen Bundesausschusses (G-BA) wurden dabei sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Aspekte berücksichtigt. Ein niedriges Verzerrungspotenzial lag vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt waren. Unter einer relevanten Verzerrung wurde verstanden, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden. Eine Zusammenfassung der Ergebnisse mittels Meta-Analyse war wegen unzureichender Vergleichbarkeit der Studien nicht sinnvoll.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die Bewertung beruht auf den beiden randomisierten kontrollierten Phase-III-Studien COMFORT-I und COMFORT-II, in denen Wirksamkeit und Sicherheit von Ruxolitinib bei Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) untersucht wurden. In COMFORT-II wurde die Ruxolitinib-Behandlung mit der nach Ermessen des Prüfarztes besten verfügbaren Therapie (Best Available Therapy, BAT) verglichen, wobei auch eine Kombination mehrerer Therapieoptionen, ein Therapiewechsel sowie eine Nichtbehandlung möglich waren. Den Patienten der COMFORT-I-Studienpopulation stand aufgrund von Refraktärität, Resistenz oder Intoleranz gegenüber anderen Therapien nur die „Nichtbehandlung“ als beste verfügbare Therapie zur Verfügung. Der Vergleich in dieser Studie erfolgte gegenüber Placebo, wobei das Studienprotokoll den Einsatz weiterer unterstützender Begleitmedikationen erlaubte. Damit entsprechen die Kontrollinterventionen beider Studien der zweckmäßigen Vergleichstherapie Best Supportive Care.

Beide Studien belegen eine drastische **Reduktion der Milzgröße** unter der Therapie mit Ruxolitinib. Sowohl in COMFORT-I als auch in COMFORT-II war der Anteil der Patienten, die nach 24 Wochen bzw. 48 Wochen eine klinisch relevante Milzvolumenreduktion um $\geq 35\%$ erreichten, im Ruxolitinib-Arm signifikant höher als im jeweiligen Kontrollarm (41,9% vs. 0,7% bzw. 28,5% vs. 0%; $p < 0,0001$ für beide Vergleiche). Gegenüber dem Ausgangswert konnte unter der Therapie mit Ruxolitinib jeweils eine mediane Abnahme des Milzvolumens um circa 30% erreicht werden, wohingegen es in den Komparator-Armen sogar zu einer Zunahme des Milzvolumens kam. Eine Milzvolumenreduktion um $\geq 35\%$ wurde bei der Mehrzahl der Patienten bereits nach 12 Wochen nachgewiesen und blieb über einen Zeitraum von drei Jahren erhalten. Der Nutzen der Behandlung zeigte sich dabei in allen vorab definierten Subgruppen, unabhängig vom Alter der Patienten, ihrem Geschlecht, dem Myelofibrose-Subtyp, der Risikogruppe oder dem Milzvolumen zu Studienbeginn.

Die Myelofibrose-assoziierten Symptome bzw. die **Linderung der Krankheitssymptomatik** unter der Therapie wurde in der Studie COMFORT-I mit Hilfe des MFSAF-Fragebogens erfasst, der die charakteristischen Symptome (Bauchbeschwerden, Schmerzen linksseitig unter den Rippen, Völlegefühl, Nachtschweiß, Juckreiz sowie Muskel- und Knochenschmerzen) in einem Summenscore, dem Total Symptom Score (TSS), zusammenfasst. Der Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen war unter der Therapie mit Ruxolitinib signifikant höher als unter Placebo (45,9% vs. 5,3%; $p < 0,0001$). In der Ruxolitinib-Gruppe konnte bei der Mehrzahl der Patienten eine deutliche Verbesserung des TSS nachgewiesen werden, wobei der Rückgang im Median bei 56,2% lag. Einige Patienten waren sogar vollständig symptomfrei. Unter der Placebo-Behandlung kam es hingegen beim überwiegenden Teil der Patienten zu einer mitunter dramatischen

Verschlechterung der Symptomatik. Die Verbesserung der Myelofibrose-Symptome unter der Therapie mit Ruxolitinib zeigte sich nicht nur beim Gesamtscore sondern auch in Bezug auf alle sechs untersuchten Einzelsymptome sowie beim ebenfalls abgefragten Item „Inaktivität“ ($p < 0,0001$ für alle Vergleiche).

Nach 3-jähriger Nachbeobachtungszeit konnte außerdem eine deutliche Verbesserung des **Gesamtüberlebens** unter der Therapie mit Ruxolitinib nachgewiesen werden. So zeigte die ITT-Auswertung von COMFORT-II eine signifikante Reduktion des Mortalitätsrisikos im Ruxolitinib-Arm der Studie (HR = 0,48 [0,28-0,85]; $p = 0,009$). Auch die Ergebnisse der COMFORT-I-Studie weisen auf einen positiven Trend in Bezug auf ein verlängertes Gesamtüberleben hin und bestätigen somit den Überlebensvorteil für Patienten, die von Beginn der Studie an mit Ruxolitinib behandelt wurden (HR = 0,69 [0,46-1,03]; $p = 0,067$). Die Tatsache, dass ein mortalitätssenkender Therapieeffekt von Ruxolitinib trotz des in beiden Studien erlaubten Cross-overs nachweisbar war, hebt dessen Bedeutung noch weiter hervor, da der tatsächliche Überlebensvorteil durch Ruxolitinib aufgrund der Auswertung nach dem ITT-Prinzip wahrscheinlich unterschätzt wird.

Die Auswertung standardisierter Fragebögen zur Erfassung der **gesundheitsbezogenen Lebensqualität** zeigte darüber hinaus, dass die Reduktion der Milzgröße sowie der konstitutionellen Symptome mit einer für die Patienten spürbaren Linderung der Erkrankung einherging. Dabei wurde unter anderem der generische QLQ-C30-Fragebogen der EORTC eingesetzt, der spezifisch für die Erfassung der Lebensqualität von Krebspatienten entwickelt und validiert wurde. In beiden Studien konnte unter der Behandlung mit Ruxolitinib in Bezug auf nahezu alle Subskalen und Symptome des Fragebogens eine deutliche Verbesserung im Vergleich zu Studienbeginn erreicht werden. In der jeweiligen Kontrollgruppe kam es dagegen – wenn überhaupt – nur zu einer minimalen Verbesserung der Scores, meist wurde sogar eine Verschlechterung der Werte beobachtet. In der COMFORT-II-Studie wurde zusätzlich der krankheitsspezifische Fragebogen FACT-Lym eingesetzt, um die typischen Myelofibrose-Symptome zu erfassen. Auch hier erfuhren Patienten unter der Therapie mit Ruxolitinib über alle Bereiche hinweg eine deutliche Verbesserung gegenüber Studienbeginn, wohingegen die Werte unter der BAT-Kontrollbehandlung unverändert blieben oder sich teilweise sogar verschlechterten.

Dementsprechend konnten unter der Therapie mit Ruxolitinib in beiden Studien mehr Patienten ihren **ECOG-Performance-Status** erhalten oder verbessern als unter der jeweiligen Kontrollintervention. Umgekehrt war der Anteil der Patienten, bei denen im Studienverlauf eine Verschlechterung des ECOG-Status dokumentiert wurde, unter BAT bzw. Placebo deutlich höher.

Die Sicherheitsanalyse der beiden Studien COMFORT-I und COMFORT-II belegt eine insgesamt gute **Verträglichkeit** von Ruxolitinib. So war die Häufigkeit unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse sowie von Ereignissen, die zum Abbruch der Therapie führten, jeweils zwischen den Behandlungsarmen vergleichbar. Lediglich der Anteil der Patienten mit einem Grad-3/4-Ereignis war in der Ruxolitinib-Gruppe in COMFORT-II höher als in der Kontrollgruppe. Umgekehrt war in COMFORT-I der Anteil der Patienten mit einem schwerwiegenden unerwünschten Ereignis unter Ruxolitinib niedriger als unter Placebo. Entsprechend dem Wirkmechanismus von Ruxolitinib als JAK-Inhibitor waren Anämien und Thrombozytopenien die häufigsten Nebenwirkungen. Sie waren jedoch in beiden Studien durch Dosisreduktionen oder symptomorientierte Supportivmaßnahmen gut kontrollierbar und führten nur selten zum Abbruch der Therapie. Auch nach längerem Follow-up ergaben sich keine neuen oder unerwarteten Sicherheitsprobleme. Im Gegenteil, die Inzidenz der unerwünschten Ereignisse ging bei längerer Behandlung sogar zurück.

Ruxolitinib stellt damit eine wichtige neue Therapieoption dar, die zu einer klinisch relevanten, lang anhaltenden Reduktion der Splenomegalie, einer deutlichen Verminderung der belastenden Krankheitssymptomatik sowie einer Verbesserung der Lebensqualität führt und das Gesamtüberleben von Patienten mit Myelofibrose verlängert.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Vor Einführung von Ruxolitinib waren keine ausreichenden Therapiemöglichkeiten für die Behandlung der Myelofibrose vorhanden. Ruxolitinib ist das erste und einzige Medikament, dessen Nutzen bei der Behandlung von Myelofibrose-Patienten nach den Standards der evidenzbasierten Medizin belegt ist und das deshalb eine arzneimittelrechtliche Zulassung im vorliegenden Anwendungsgebiet besitzt.

Die Nutzenbewertung von Ruxolitinib beruht auf den beiden randomisierten kontrollierten Studien COMFORT-I und COMFORT-II, in denen die Wirksamkeit von Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie Best Supportive Care bei Patienten mit primärer oder sekundärer Myelofibrose untersucht wurde. Die Aussagekraft der Nachweise zur Abschätzung des Zusatznutzens von Ruxolitinib ist aufgrund der hohen Evidenzstufe der beiden Studien (Evidenzstufe Ib), der guten Studienqualität sowie der Validität der herangezogenen Endpunkte als „hoch“ einzuschätzen. Eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens ist deshalb mit hoher Ergebnissicherheit möglich.

Wie die Ergebnisse der beiden Studien COMFORT-I und COMFORT-II zeigen, kommt es unter der Therapie mit Ruxolitinib zu einer klinisch bedeutenden, lang anhaltenden Reduktion des pathologisch vergrößerten Milzvolumens, einer drastischen Verminderung der belastenden Krankheitssymptomatik hinsichtlich aller untersuchten Symptome und einer signifikanten Senkung des Mortalitätsrisikos. Damit konnte bei Patienten mit Myelofibrose erstmals eine Abschwächung schwerwiegender Symptome der Erkrankung sowie eine Verlängerung des Gesamtüberlebens durch eine medikamentöse Therapieoption erreicht werden. Zugleich ging der therapeutische Effekt mit einer nachhaltigen Verbesserung der gesundheitsbezogenen Lebensqualität und des ECOG-Performance-Status einher. Wie die Auswertung von a priori geplanten Subgruppenanalysen zeigt, die jeweils für den primären Endpunkt der beiden Studien durchgeführt wurden, war der Vorteil der Ruxolitinib-Therapie dabei über alle vorab definierten Subgruppen hinweg konsistent.

Der Zusatznutzen von Ruxolitinib gegenüber Best Supportive Care ergibt sich somit aus der Abschwächung schwerwiegender Symptome der Erkrankung sowie einer deutlichen Verlängerung der Überlebensdauer bei guter Verträglichkeit. Hinzu kommt eine für die Patientinnen und Patienten spürbare Linderung der Erkrankung, die sich in der Verbesserung der Lebensqualität und des allgemeinen Gesundheitszustandes widerspiegelt. Insbesondere auch unter Berücksichtigung des Schweregrades der Erkrankung und des Fehlens alternativer Therapiemöglichkeiten ist der Zusatznutzen von Ruxolitinib damit für alle Patienten im zugelassenen Anwendungsgebiet als **beträchtlich** einzustufen.

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.

Ziel der vorliegenden Untersuchung ist es, Ausmaß und Wahrscheinlichkeit des medizinischen Zusatznutzens von Ruxolitinib (Jakavi®) im Vergleich zur aktuell bestmöglichen, patientenindividuell optimierten, unterstützenden Behandlung (Best Supportive Care, BSC) bei der Behandlung von krankheitsbedingter Splenomegalie oder Symptomen bei Erwachsenen mit primärer Myelofibrose (PMF), Post-Polycythaemia-vera-Myelofibrose (PPV-MF) oder Post-Essentieller-Thrombozythämie-Myelofibrose (PET-MF) auf Grundlage der besten verfügbaren Evidenz zu ermitteln. Die Bewertung soll hinsichtlich der patientenrelevanten Endpunkte *Mortalität* (Gesamtüberleben), *Morbidität* (Reduktion der Milzgröße, Linderung der Krankheitssymptomatik, leukämiefreies Überleben, progressionsfreies Überleben, Veränderung des ECOG-Performance-Status), *gesundheitsbezogene Lebensqualität* (EORTC-QLQ-C30, FACT-Lym) und *Verträglichkeit* erfolgen.

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.

Nach der oben formulierten Fragestellung zur Bewertung des medizinischen Nutzens und Zusatznutzens von Ruxolitinib (Jakavi®) wurden folgende Kriterien für den Einschluss von Studien in die Nutzenbewertung berücksichtigt:

Patientenpopulation:

Die Patientenpopulation soll erwachsene Patienten mit krankheitsbedingter Splenomegalie oder Symptomen bei primärer Myelofibrose (PMF), Post-Polycythaemia-vera-Myelofibrose (PPV-MF) oder Post-Essentieller-Thrombozythämie-Myelofibrose (PET-MF) einschließen. Da Studien, die für die Nutzenbewertung herangezogen werden, für die gemäß Zulassung in Frage kommenden Patienten aussagekräftig sein müssen, entspricht die Patientenpopulation der in Deutschland geltenden Zulassungsindikation (1).

Intervention:

Die Intervention ist – entsprechend der Zulassung – Ruxolitinib (Jakavi®) in einer Anfangsdosierung von 15 mg 2x täglich bei Patienten mit einer Thrombozytenzahl zwischen 100.000/µl und 200.000/µl bzw. 20 mg 2x täglich bei Patienten mit einer Thrombozytenzahl über 200.000/µl (1).

Vergleichsintervention:

Als Vergleichstherapie wird nach den Vorgaben des Gemeinsamen Bundesausschusses Best Supportive Care (BSC) definiert (2). Als Best Supportive Care wird dabei die Therapie verstanden, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität gewährleistet.

Endpunkte:

Als Zielgröße zum Nachweis des Zusatznutzens sollen patientenrelevante Endpunkte herangezogen werden, die die Verfahrensordnung des Gemeinsamen Bundesausschusses mit Morbidität, Mortalität und gesundheitsbezogene Lebensqualität angibt. Für das vorliegende Dossier werden deshalb folgende Endpunkte definiert:

Mortalität

- Gesamtüberleben

Morbidität

- Reduktion der Milzgröße
- Linderung der Krankheitssymptomatik
- Leukämiefreies Überleben
- Progressionsfreies Überleben
- Veränderung des ECOG-Performance-Status

Gesundheitsbezogene Lebensqualität

- European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC-QLQ-C30)
- Functional Assessment of Cancer Therapy – Lymphoma (FACT-Lym)

Verträglichkeit (therapiebedingte Morbidität)

Auf Patientenrelevanz und Validität der entsprechenden Endpunkte wird in Abschnitt 4.2.5.2 näher eingegangen.

Studientypen:

Für die Beantwortung der oben formulierten Fragestellung ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien (randomized controlled trials, RCTs) möglich und praktisch durchführbar. Da sie die höchste Ergebnissicherheit aufweisen, sollen nur RCTs in die Nutzenbewertung einbezogen werden (3). Studien, für die weder eine Volltext-Publikation noch ein Studienbericht oder ein ausführlicher Ergebnisbericht aus einem Studienregister verfügbar sind, werden aus der Bewertung ausgeschlossen, da bei Fehlen umfassender Informationen Verzerrungspotenzial und Ergebnissicherheit der Studie nicht bewertet werden können. Studien, für die noch keine Ergebnisse vorliegen, sowie tierexperimentelle Studien und In-vitro-Untersuchungen bleiben ebenfalls unberücksichtigt.

Studiendauer:

Es werden keine Einschränkungen bezüglich der Studiendauer vorgenommen.

Eine Übersicht über die Ein- und Ausschlusskriterien für die Auswahl von Studien für die Nutzenbewertung von Ruxolitinib ist in Tabelle 4-2 dargestellt.

Tabelle 4-2: Ein- und Ausschlusskriterien für die Auswahl von Studien für die Nutzenbewertung von Ruxolitinib

Einschlusskriterien		
E1	Patientenpopulation	Erwachsene Patienten mit krankheitsbedingter Splenomegalie oder Symptomen bei primärer Myelofibrose (PMF), Post-Polycythaemia-vera-Myelofibrose (PPV-MF) oder Post-Essentieller-Thrombozythämie-Myelofibrose (PET-MF)
E2	Intervention	Ruxolitinib (Jakavi®)
E3	Vergleichstherapie	Best Supportive Care (BSC)
E4	Endpunkte	Mindestens einer der folgenden Endpunkte: • Gesamtüberleben • Reduktion der Milzgröße • Linderung der Krankheitssymptomatik • Leukämiefreies Überleben • Progressionsfreies Überleben • Veränderung des ECOG-Performance-Status • Gesundheitsbezogene Lebensqualität (EORTC-QLQ-C30, FACT-Lym) • Verträglichkeit
E5	Studientyp	Randomisierte kontrollierte Studie (RCT)
Ausschlusskriterien		
A1	Studien, für die weder eine Volltext-Publikation noch ein Studienbericht oder ein ausführlicher Ergebnisbericht in einem Studienregister verfügbar sind oder für die noch keine Ergebnisse vorliegen	
A2	Review, Übersichtsartikel, Kommentar, Survey, Note, Editorial, Letter, Erratum	
A3	Tierexperimentelle Studien, In-vitro-Untersuchungen	

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.

Die bibliografische Literaturrecherche für die im vorliegenden Dossier formulierte Fragestellung wurde in den Datenbanken MEDLINE und EMBASE, der Cochrane Database of Systematic Reviews und dem Cochrane Central Register of Controlled Trials durchgeführt. Die bibliografische Literaturrecherche zielte primär darauf ab, alle Publikationen zu Ruxolitinib im Indikationsbereich „Myelofibrose“ zu erfassen. Für jede Datenbank wurde eine eigene adaptierte Suchstrategie verwendet. Die vollständigen Suchstrategien sind im Anhang 4-A dokumentiert. Die Suche wurde dabei so sensitiv wie möglich gehalten. So

wurden keinerlei Einschränkungen bezüglich des Zeitpunktes der Veröffentlichung und der Sprache getroffen. Allerdings wurde der Studientyp in den Datenbanken EMBASE und MEDLINE mit Hilfe des Filters nach Wong et al. auf „randomisierte kontrollierte Studie“ beschränkt (4), um das Verzerrungspotential auf Studienebene so weit wie möglich niedrig zu halten. Die Suche wurde am 16. Januar 2014 durchgeführt. Eine Aktualisierung fand am 24. Februar 2014 statt.

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 Unternehmer 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.

Die Recherche erfolgte in den Studienregistern ClinicalTrials.gov, EU Clinical Trials Register (EU-CTR) und Klinische Prüfungen PharmNet.Bund sowie über das International Clinical Trials Registry Platform (ICTRP) Search Portal. Um die Sensitivität der Suche so hoch wie möglich zu halten, wurden die Schlagwörter auf „myelofibrosis“ und “ruxolitinib” bzw. entsprechende Alternativbegriffe beschränkt. Die Suchstrategie und die vollständigen Rechercheergebnisse sind in Anhang 4-B dokumentiert. Die Suche im Studienregister ClinicalTrials.gov sowie dem ICTRP Search Portal erfolgte am 16. Januar 2014, die Suche in den Registern Klinische Prüfungen PharmNet.Bund und EU-CTR wurde am 14. Januar 2014 durchgeführt. Eine Aktualisierung fand am 24. Februar 2014 statt.

4.2.3.4 Selektion relevanter Studien

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

Die Liste der Treffer wurde auf Grundlage der prädefinierten Selektionskriterien von zwei Reviewern unabhängig voneinander bewertet. Publikationen, die eines der Ausschlusskriterien erfüllten bzw. eines der Einschlusskriterien nicht erfüllten, wurden ausgeschlossen. Die Volltexte der potentiell relevanten Studien wurden erneut anhand der prädefinierten Kriterien hinsichtlich ihrer Relevanz bewertet. Etwaige Diskrepanzen zwischen den beiden Reviewern wurden diskutiert und im Konsens gelöst.

4.2.4 Bewertung der Aussagekraft der Nachweise

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

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise wurde das Verzerrungspotential der Studien sowohl unter allgemeinen (endpunktübergreifenden) als auch endpunktspezifischen Aspekten beurteilt. Dazu wurde vornehmlich der Studienbericht (clinical study report, CSR) und verfügbare Volltextpublikationen der eingeschlossenen Studien herangezogen. Ergänzend wurden weitere Dokumente wie zum Beispiel der European

Public Assessment Report (EPAR) gesichtet. Details zum Verzerrungspotenzial sind den Bewertungsbögen in Anhang 4-F zu entnehmen. Die eingesetzte Methodik weicht somit nicht von der oben beschriebenen Methodik ab.

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.

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

Die Beschreibung des Designs und der Methodik der RCTs erfolgte anhand der durch das CONSORT-Statement vorgegebenen Standards. Ausführliche Informationen zu den Items 2b bis 14 wurden jeweils dem Studienprotokoll bzw. den relevanten Publikationen entnommen und sind zusammen mit dem CONSORT-Flow-Chart in Anhang 4-E dokumentiert.

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,

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.

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.

Patientencharakteristika

Im Abschnitt 4.3.1.2.1 sind die Patientencharakteristika der Studienpopulationen innerhalb der einzelnen Behandlungsgruppen dargestellt, da diese in den verschiedenen Armen einer Studie ausgeglichen sein müssen und das untersuchte Patientenkollektiv repräsentativ für Patienten mit Myelofibrose sein sollte. Die Beschreibung umfasst dabei demographische Charakteristika wie Alter und Geschlecht der Patienten, ihre Hautfarbe sowie krankheits-spezifische Daten zur Erfassung der Krankheitsschwere. Dazu wurde der Myelofibrose-Subtyp (PMF, PPV-MF, PET-MF) des Patienten angegeben und das krankheitsspezifische Risiko mit Hilfe des IPSS-Risikoprofils quantifiziert. Außerdem wurde das Ausmaß der Splenomegalie, das Vorhandensein einer *JAK2*^{V617F}-Mutation, verschiedene krankheits-relevante Blutwerte (Hämoglobin, Thrombozyten, Anteil der Blasten im peripheren Blut) sowie eventuelle Vorbehandlungen dokumentiert.

Endpunkte

Als Zielgröße zum Nachweis des Zusatznutzens sollen patientenrelevante Endpunkte herangezogen werden, die die Verfahrensordnung des Gemeinsamen Bundesausschuss mit Mortalität, Morbidität und Lebensqualität angibt. Dabei wurden folgende Endpunkte als unmittelbar patientenrelevant für die vorliegende Nutzenbewertung eingestuft:

Mortalität

- Gesamtüberleben

Morbidität

- Reduktion der Milzgröße
- Linderung der Krankheitssymptomatik
- Leukämiefreies Überleben
- Progressionsfreies Überleben
- Veränderung des ECOG-Performance-Status

Gesundheitsbezogene Lebensqualität

- European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC-QLQ-C30)
- Functional Assessment of Cancer Therapy – Lymphoma (FACT-Lym)

Verträglichkeit (therapiebedingte Morbidität)

Die Patientenrelevanz der oben genannten Endpunkte begründet sich dabei wie folgt:

Gesamtüberleben

Das Gesamtüberleben (overall survival, OS) ist definiert als der Zeitraum von Studienbeginn bis zum Tod des Patienten ungeachtet der zugrundeliegenden Todesursache und gilt in der Onkologie unbestritten als patientenrelevanter Endpunkt (5). Auch in einem kürzlich veröffentlichten Rapid Report des IQWiG zur Bewertung von Surrogatendpunkten in der Onkologie wird das Gesamtüberleben als herausragendes Therapieziel beschrieben (6). Aufgrund der unbeeinflussbaren Definition des gemessenen Ereignisses wird das Gesamtüberleben häufig als der zuverlässigste Endpunkt in der Onkologie betrachtet.

Reduktion der Milzgröße

Die Vergrößerung der Milz (Splenomegalie) ist ein charakteristisches Merkmal der Myelofibrose und betrifft die große Mehrheit der Patienten. So betrug der Anteil der PMF-Patienten mit tastbarer Splenomegalie in zwei großen Datenbankauswertungen 79% bzw. 89% (7;8). Dementsprechend ist das Vorhandensein der Splenomegalie eines der Kriterien zur Diagnose einer primären Myelofibrose sowie einer Post-Polycythaemia-vera-Myelofibrose bzw. Post-Essentiellen-Thrombozythämie-Myelofibrose (9;10). Die Milzvergrößerung korreliert mit der Krankheitslast und den Allgemeinsymptomen (11). Patienten, die an einer Splenomegalie leiden, werden durch ein zusätzliches Gewicht von bis zu 10 kg belastet. Die massiv vergrößerte Milz ist durch die Bauchdecke nach außen sichtbar und drückt gleichzeitig

auf andere innere Organe wie Magen, Darm und Lunge (12). Dabei kommt es häufig zu weiteren klinisch relevanten Symptomen wie Bauchschmerzen, Völlegefühl, vorzeitigem Sättigungsgefühl und Pfortaderhochdruck (13). Bei progressivem Verlauf besteht außerdem das Risiko eines Milzinfarkts (12). Der Zusammenhang zwischen der Milzvergrößerung und dem Auftreten weiterer klinischer Symptome wurde kürzlich in einer retrospektiven Studie an Myelofibrose-Patienten systematisch untersucht. Gegenüber Patienten ohne Splenomegalie litt ein signifikant höherer Anteil der Patienten mit Splenomegalie unter Schmerzen und Beschwerden im linken subkostalen Bereich (85% versus 37%), allgemeinen Bauchbeschwerden und Schmerzen (67% versus 50%), vorzeitigem Sättigungsgefühl (79% versus 47%) und Nachtschweiß (65% versus 50%) (14).

Die Vergrößerung der Milz ist somit mit einer Vielzahl belastender, klinisch relevanter Symptome assoziiert. Umgekehrt geht die Reduktion der Milzgröße mit einer Verminderung der Symptomlast sowie einer Verbesserung der gesundheitsbezogenen Lebensqualität einher (15;16). Darüber hinaus wird die klinische Relevanz der Milzgrößenreduktion durch den Zusammenhang mit einem verlängerten Gesamtüberleben untermauert. So ergab eine Untersuchung an Patienten, die im Rahmen der Phase-I/II-Studie INCB18424-251 mit Ruxolitinib behandelt wurden, dass die Reduktion der Milzlänge mit einer deutlichen Erhöhung der Überlebenswahrscheinlichkeit verbunden ist (Abbildung 4-1) (17). Folglich ist eine Reduktion des pathologisch vergrößerten Milzvolumens unmittelbar patientenrelevant.

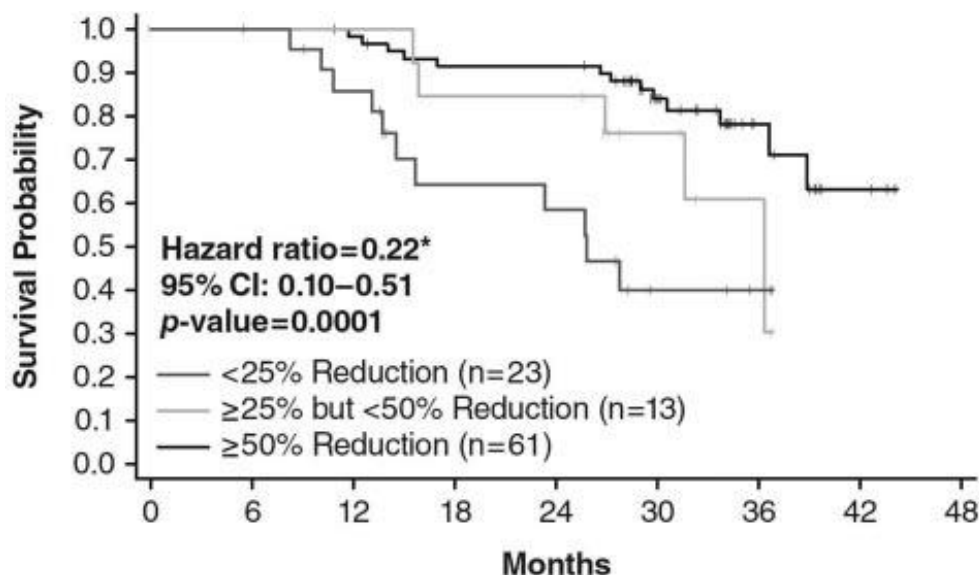


Abbildung 4-1: Zusammenhang zwischen Reduktion der Milzlänge und Gesamtüberleben

Kaplan-Meier-Kurven für das Gesamtüberleben von Patienten unter der Behandlung mit Ruxolitinib in Abhängigkeit vom Ausmaß der Milzlängenreduktion in der initialen Phase-I/II-Studie INCB18424-251. Aus Verstovsek S et al. Blood 2012 Aug 9;120(6):1202-9. * Vergleich zwischen Patienten mit einer Reduktion der Milzlänge um < 25% und Patienten mit einer Reduktion der Milzlänge um ≥ 50%.

Entsprechend seiner Patientenrelevanz wurde die initiale Bewertung des Zusatznutzens von Ruxolitinib im vorliegenden Anwendungsgebiet vom Gemeinsamen Bundesausschuss auf dem Endpunkt „Reduktion der Milzgröße“ abgestellt (18;19). Darüber hinaus bestätigt auch die Europäische Arzneimittel-Agentur (EMA) im European Public Assessment Report (EPAR) zu Ruxolitinib, dass eine Milzvolumenreduktion nach klinischen Erfahrungen mit einer Verminderung der abdominellen Symptome verbunden ist und insbesondere ein Rückgang des Milzvolumens um $\geq 35\%$, was einer Reduktion der Milzlänge um 50% entspricht, klinisch bedeutsam ist (20).

Linderung der Krankheitssymptomatik

Symptome sind die vom Patienten wahrgenommenen Krankheitszeichen. Die Symptome der Myelofibrose stellen für Patienten eine Krankheitslast dar und bedeuten sowohl eine Verschlechterung des Gesundheitszustands als auch eine Beeinträchtigung der Alltagsfunktion und der Lebensqualität (21). Die Linderung der Krankheitssymptomatik ist folglich ein patientenrelevanter Endpunkt.

Die durch die Myelofibrose-Symptome bedingte Krankheitslast wurde in der COMFORT-I-Studie mittels des modifizierten MFSAF v2.0 (Myelofibrosis Symptom Assessment Form Version 2.0)-Fragebogens quantitativ erfasst (22-24). Der MFSAF v2.0 ist ein speziell zur Erfassung der Symptome der Myelofibrose auf Basis des MFSAF weiterentwickelter Fragebogen, über den die folgenden sechs charakteristischen Symptome der Myelofibrose erfasst werden: Nachtschweiß, Juckreiz, Bauchbeschwerden, Schmerzen linksseitig unter den Rippen, Völlegefühl sowie Knochen- und/oder Muskelschmerzen (25). Der Fragebogen quantifiziert die Ausprägung der Symptome jeweils in Bezug auf die höchste Intensität innerhalb der letzten 24 Stunden auf einer numerischen Skala (von 0-10). Aus dem numerischen Einzelscore für jedes Symptom ergibt sich durch Addition ein Gesamtscore (Total Symptom Score, TSS), der die Krankheitslast quantitativ erfasst.

Die Entwicklung des Fragebogens erfolgte in enger Abstimmung mit der amerikanischen Zulassungsbehörde Food and Drug Administration (FDA) und basierend auf den FDA-Leitlinien zu patientenbezogenen Endpunkten (26). Die Ergebnisse des MFSAF v2.0-Fragebogens bei der Erfassung der Myelofibrose-Symptomatik korrelieren mit denen anderer validierter Erhebungsinstrumente wie dem QLQ-C30-Fragebogen der EORTC (European Organisation for Research and Treatment of Cancer) (27), dem ECOG (Eastern Cooperative Oncology Group)-Performance-Status (28) und der PGIC (Patient Global Impression of Change)-Bewertung (29). Dabei wird eine Verbesserung im TSS um $\geq 50\%$ von den Patienten als bedeutend eingestuft: 91,2% der Patienten beschrieben ihre Verfassung im PGIC als „viel“ oder „sehr viel“ besser. Darüber hinaus zeigten Patienten mit einer Reduktion des TSS um

≥ 50% eine statistisch signifikante Verbesserung in allen Subskalen des QLQ-C30-Fragebogens (16). Eine detaillierte Beschreibung der Validierung und der Validität des MFSAF v2.0 ist im beigefügten Evidenz-Dossier dokumentiert (30).

Leukämiefreies Überleben

In beiden für die vorliegende Nutzenbewertung herangezogenen Studien wurde das leukämiefreie Überleben als Endpunkt untersucht. Dieses war definiert als der Zeitraum von der Randomisierung bis zum Auftreten einer leukämischen Transformation (definiert als ein Anstieg der Blasten im peripheren Blut auf ≥ 20% für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf ≥ 20%) oder bis zum Tod ungeachtet der zugrundeliegenden Todesursache (23;31).

Das 10-Jahresrisiko für eine Transformation der Myelofibrose in eine akute myeloische Leukämie (AML) beträgt 8-23% (32-35). Diese ist nur äußerst schwer behandelbar und führt gewöhnlich innerhalb weniger Monate zum Tod (36): In einer Studie von 91 Patienten mit leukämischer Transformation war der Verlauf bei einer medianen Überlebenszeit von 2,6 Monaten in 98% der Fälle tödlich (34). Die Transformation in eine AML geht nicht nur mit einer außerordentlich ungünstigen Überlebensprognose sondern auch mit weiteren klinischen Symptomen einher, deren Auftreten bzw. Ausbleiben patientenrelevant sind. So treten nach dem Übergang in eine AML neben den auch für die Myelofibrose typischen klinischen Symptomen wie Nachtschweiß, Appetitlosigkeit oder Knochenschmerzen verstärkt Infektionen und Blutungen auf, die den Gesundheitszustand des Patienten weiter beeinträchtigen (37). Die Verhinderung der leukämischen Transformation stellt deshalb ein wichtiges Therapieziel bei der Behandlung der Myelofibrose dar und ist damit in höchstem Maße patientenrelevant.

Progressionsfreies Überleben

Das progressionsfreie Überleben wurde in der COMFORT-II-Studie als Endpunkt untersucht. Es war definiert als der Zeitraum von der Randomisierung bis zum Auftreten eines der folgenden Ereignisse: der Zunahme des Milzvolumens um > 25% im Vergleich zum niedrigsten Wert während der Studie, dem Nachweis einer leukämischen Transformation (definiert als ein Anstieg der Blasten im peripheren Blut auf ≥ 20% für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf ≥ 20%), der Durchführung einer Milzbestrahlung bzw. Splenektomie oder dem Tod ungeachtet der zugrundeliegenden Todesursache (31).

Entsprechend der klinischen Bedeutung der Splenomegalie und der damit einhergehenden Symptomatik sowie der Patientenrelevanz der leukämischen Transformation, wird die Erkrankungsprogression nicht nur durch den Übergang in eine AML sondern vor allem auch über das (Wieder-)Auftreten oder das Voranschreiten einer Splenomegalie bzw. die Notwendigkeit zur Durchführung einer Splenektomie oder Milzbestrahlung definiert. Basierend auf der Patientenrelevanz der Einzelkriterien ist der Endpunkt „Progressionsfreies Überleben“ damit ebenfalls patientenrelevant.

Veränderung des ECOG-Performance-Status

Der ECOG (Eastern Cooperative Oncology Group)-Performance-Status ist ein häufig eingesetztes, etabliertes Messinstrument, das den physischen Zustand von Krebspatienten und deren Einschränkung bei Aktivitäten des alltäglichen Lebens mit Hilfe einer Skala von 0 (keine Einschränkung) bis 5 (Tod) beschreibt (28). Der ECOG-Performance-Status bzw. dessen Veränderung unter der Therapie ist also ein direktes Maß für den Gesundheitszustand des Patienten sowie seine Fähigkeit, ein selbstbestimmtes Leben zu führen, und ist daher in höchstem Maße patientenrelevant. Auch der Gemeinsame Bundesausschuss bekräftigt in der Nutzenbewertung zu Ruxolitinib vom 17. Dezember 2012, dass die Betrachtung des ECOG-Performance-Status sowie dessen Veränderung im Studienverlauf sinnvoll ist (38).

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität ist eine wichtige Zielgröße bei der medizinischen Versorgung krebskranker Patienten. Sie spiegelt direkt die Auswirkungen der Erkrankung bzw. deren Linderung auf das Befinden des Patienten und seinen Alltag wider und ist damit unmittelbar patientenrelevant. Auch der Erhalt der Lebensqualität stellt gerade in der palliativen Situation, die auf Dauer fast immer mit einer Verschlechterung der Lebensqualität einhergeht, einen patientenrelevanten Endpunkt dar.

Der Fragebogen QLQ-C30 der EORTC (European Organisation for Research and Treatment of Cancer) ist spezifisch für die Erfassung der Lebensqualität von Krebspatienten entwickelt und validiert (27). Mit dem Fragebogen werden die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität (allgemeiner Gesundheitszustand/Lebensqualität, körperliche Funktion, Rollenfunktion, kognitive Funktion, emotionale Funktion und soziale Funktion), die häufigsten von Tumorpatienten berichteten Symptome (Fatigue, Übelkeit/ Erbrechen, Schmerzen, Dyspnoe, Schlafstörungen, Appetitlosigkeit, Obstipation und Diarrhö) sowie finanzielle Schwierigkeiten erfasst (27). Osoba und Kollegen beschreiben außerdem die Ableitung einer klinisch relevanten Differenz (minimal important difference, MID) durch den Fragebogen (39). Dementsprechend wird eine Veränderung der Skala von 5-10 Punkten vom

Patienten als „gering“ empfunden, eine Veränderung um 10-20 Punkte als „moderat“ und eine Veränderung von mehr als 20 Punkten als „sehr groß“. Bei dem Fragebogen QLQ-C30 der EORTC handelt es sich somit um ein etabliertes und validiertes Messinstrument, das die Lebensqualität adäquat erfasst und klinisch relevante Änderungen darzustellen vermag (27;39).

Da der generische QLQ-C30-Fragebogen nicht dazu geeignet ist, spezifische Symptome von Myelofibrose-Patienten zu erfassen, wurde in der Studie COMFORT-II zusätzlich der Fragebogen FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma) eingesetzt, der ursprünglich dazu entwickelt wurde, die Lebensqualität von Patienten mit Non-Hodgkin-Lymphom zu erfassen. Neben dem generischen Kernfragebogen FACT-G (FACT-General), der die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität wie physisches Wohlbefinden (PWB), soziales/familiäres Wohlbefinden (SWB), emotionales Wohlbefinden (EWB) und funktionelles Wohlbefinden (FWB) abfragt, werden im LymS (Lymphoma Subscale) krankheitsspezifische Probleme wie die auch für die Myelofibrose typischen Symptome Fieber, Nachtschweiß, Juckreiz und Gewichtsverlust erfasst (40;41). In einer Übersicht von Webster und Kollegen wird für den Fragebogen FACT-G ein klinisch relevanter Schwellenwert (MID) zur Beurteilung des Therapieerfolgs präsentiert (42). Arbeiten von Carter et al. und Hlubocky et al. beschreiben außerdem die Ableitung einer MID für die Skalen LymS und FACT-TOI sowie den FACT-Lym Total Score (41;43). Auch bei dem Fragebogen FACT-Lym handelt es sich somit um ein etabliertes und validiertes Messinstrument, das die Lebensqualität der Patienten adäquat erfasst und klinisch relevante Änderungen darzustellen vermag.

Verträglichkeit

Unerwünschte Ereignisse haben einen direkten Einfluss auf den Gesundheitszustand des Patienten und sind damit als Endpunkt unmittelbar patientenrelevant. Sie reflektieren vor allem in Form der therapieassoziierten, unerwünschten Ereignisse die Verträglichkeit der Therapie (therapiebedingte Morbidität), wobei Klassifikation, Schwere und Häufigkeit der Ereignisse Kriterien zur Differenzierung darstellen. Ergänzend zu den Wirksamkeitsendpunkten dient die Dokumentation unerwünschter Arzneimittelwirkungen der Nutzen-Risiko-Bewertung. Die Erfassung der unerwünschten Ereignisse in den in die Nutzenbewertung eingeschlossenen Studien erfolgte gemäß internationalen Standards und ist somit validiert.

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.

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

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 Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

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

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

Da keine Meta-Analysen durchgeführt wurden, entfällt auch die Durchführung von Sensitivitätsanalysen zur Untersuchung der Stabilität der Ergebnisse bezüglich methodischer Faktoren im Rahmen meta-analytischer Untersuchungen. Stattdessen wurden verschiedene Auswertungen zur Überprüfung der Robustheit der Ergebnisse auf der Ebene der Einzelstudien vorgenommen. So wurden für die Endpunkte „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ “ bzw. „Anteil der Patienten mit einer Verbesserung des Total Symptom Score um $\geq 50\%$ “ entsprechende Sensitivitätsanalysen durchgeführt, im Rahmen derer auch Patienten mit fehlendem Baseline-Wert berücksichtigt wurden. Patienten ohne Baseline-Wert wurden dabei als Non-Responder gewertet. Weitere Sensitivitätsanalysen sind in den Studienberichten der beiden in die Nutzenbewertung eingeschlossenen Studien dargestellt.

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.

Um Erkenntnisse über die Konsistenz des Therapieeffektes von Ruxolitinib zu erhalten, wurde jeweils für den Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ “ (primärer Endpunkt der beiden eingeschlossenen Studien) eine Analyse des Effekts in präspezifizierten Subgruppen durchgeführt.

Die Analyse des Zielkriteriums erfolgte dabei getrennt in folgenden a priori definierten Subgruppen:

COMFORT-II:

- Geschlecht (männlich, weiblich)
- Alter (≤ 65 Jahre, > 65 Jahre)
- Hautfarbe (Weiße, Nicht-Weiße)
- Myelofibrose-Typ (PMF, PPV-MF, PET-MF)

- Risikogruppe (intermediär-2, hoch)
- Milzvolumen zu Studienbeginn ($\leq$ Median, $>$ Median)
- Thrombozytenzahl zu Studienbeginn ($\leq 200.000/\mu\text{l}$, $> 200.000/\mu\text{l}$)
- Vorbehandlung mit Hydroxycarbamid (ja, nein)

COMFORT-I:

- Geschlecht (männlich, weiblich)
- Alter (≤ 65 Jahre, > 65 Jahre)
- Myelofibrose-Typ (PMF, PPV-MF, PET-MF)
- Vorbehandlung mit Hydroxycarbamid in den letzten 3 Monaten vor Studienbeginn (ja, nein)
- Vorliegen der $JAK2^{V617F}$ -Mutation zu Studienbeginn (ja, nein)
- Milzvolumen zu Studienbeginn ($\leq$ Median, $>$ Median)
- tastbare Milzlänge zu Studienbeginn (≤ 10 cm, > 10 cm)
- Land (USA vs. Kanada und Australien)
- Hautfarbe (Weiße, Nicht-Weiße)

Im Rahmen der Subgruppenanalyse wurden neben den Ansprechraten in den jeweiligen Behandlungsgruppen auch die Differenz der Ansprechraten zwischen den Gruppen, das relative Risiko sowie das Odds Ratio angegeben und die zugehörigen 95%-Konfidenzintervalle berechnet. Mögliche Effektmodifikationen wurden anhand von Interaktionstermen untersucht. Dazu wurden die den Analysen zugrundeliegenden Modelle im Rahmen der Subgruppenanalyse um die entsprechende Subgruppenvariable und den Interaktionsterm $\text{Behandlung} \times \text{Subgruppenvariable}$ erweitert. Entsprechend den Empfehlungen des IQWiG wurde ein p-Wert von $< 0,05$ als Beleg für eine Effektmodifikation bzw. ein p-Wert zwischen $0,05$ und $< 0,2$ als Hinweis auf eine Effektmodifikation gewertet (3). Lag mindestens ein Hinweis auf unterschiedliche Effekte in den Subgruppen vor oder waren keine Interaktionstests durchführbar, so wurden neben den einzelnen Subgruppenergebnissen auch die zugehörigen p-Werte berichtet.

Da Post-hoc-Analysen von nicht präspezifizierten Subgruppen aus verschiedenen methodischen Gründen als problematisch anzusehen sind, wurden im vorliegenden Dossier keine gesonderten Post-hoc-Analysen durchgeführt (44;45). So können die Ergebnisse von post hoc durchgeführten Subgruppenanalysen nicht als methodisch korrekte Prüfung einer Hypothese betrachtet und interpretiert werden (3). Generell sind Subgruppenanalysen kritisch zu beurteilen, da die zufällige Verteilung durch die Randomisierung in Subgruppen bei fehlender Stratifizierung nach dem jeweiligen Subgruppenmerkmal oftmals nicht mehr gewährleistet ist und die statistische Power aufgrund geringerer Patientenzahlen innerhalb der einzelnen

Subgruppen häufig nicht ausreicht, um Unterschiede zwischen den Behandlungsgruppen interferenzstatistisch aufzudecken. Dies gilt insbesondere für die Durchführung von nicht präspezifizierten Subgruppenanalysen, da diese Faktoren bereits bei der Planung einer Studie berücksichtigt werden müssten. Umgekehrt können aufgrund der steigenden Zahl an statistischen Tests zufällig signifikante Ergebnisse entstehen (Problem des multiplen Testens) (3;44;46-48). Post-hoc-Analysen von Subgruppen sind daher nicht dazu geeignet, definitive Schlüsse über das Ausmaß des Behandlungseffektes in einer bestimmten Subgruppe zu ziehen, sondern dienen vielmehr der Hypothesengenerierung (45;46;49).

Dennoch werden der Vollständigkeit wegen für den Endpunkt „Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen“ Post-hoc-Analysen nach verschiedenen Subgruppenmerkmalen im Dossier präsentiert, da diese im Studienbericht der COMFORT-I-Studie dargestellt sind. Bei der Interpretation der Ergebnisse ist jedoch zu beachten, dass diese Auswertungen als Post-hoc-Analysen rein explorativen Charakter aufweisen.

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

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.

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, vorliegen.

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 Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Es wurden keine indirekten Vergleiche durchgeführt.

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

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

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 4-3: 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
CINC424A2352 COMFORT-II A Randomized Study of the JAK Inhibitor INC424 Tablets Compared to Best Available Therapy in Subjects with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia Myelofibrosis (PET-MF)	ja	ja	Laufend, aber nicht mehr rekrutierend	Seit Juli 2009, voraussichtlich bis Januar 2015	1. Ruxolitinib (15 mg bid oder 20 mg bid) 2. BAT (nach Ermessen des Prüfarztes)
INCB 18424-351 COMFORT-I A Randomized, Double-blind, Placebo-controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF)	ja	nein	Laufend, aber nicht mehr rekrutierend	Seit August 2009, voraussichtlich bis Januar 2015	1. Ruxolitinib (15 mg bid oder 20 mg bid) 2. Placebo
Abkürzungen: BAT, Best Available Therapy; bid, bis in die (lat., zweimal täglich)					

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

Stand: 26. Februar 2014

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

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

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie

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.

Im Rahmen der bibliographischen Literaturrecherche wurden folgende Datenbanken durchsucht: MEDLINE, EMBASE, die Cochrane Database of Systematic Reviews sowie das Cochrane Central Register of Clinical Trials. Die Suche wurde am 16. Januar 2014 durchgeführt. Eine Aktualisierung erfolgte am 24. Februar 2014. Die dabei verwendeten Suchstrategien sind in Anhang 4-A dargestellt. In der Recherche wurden insgesamt 226 Treffer identifiziert (MEDLINE: 34, EMBASE: 177, Cochrane Library: 15). Nach Ausschluss von Duplikaten blieben noch 174 Literaturstellen übrig. Die Auswahl relevanter Studien aus diesem Pool erfolgte durch zwei unabhängige Reviewer anhand der in Abschnitt 4.2.2 beschriebenen Kriterien und führte zum Ausschluss von 141 weiteren Treffern. Die verbleibenden 33 Literaturstellen dokumentieren alle Ergebnisse der beiden Studien COMFORT-I und COMFORT-II, die bereits in Abschnitt 4.3.1.1.1 aufgelistet wurden und der Bewertung des Zusatznutzens von Ruxolitinib im vorliegenden Dossier zugrunde liegen. Der Ablauf der Recherche ist im Flussdiagramm in Abbildung 4-2 dokumentiert.

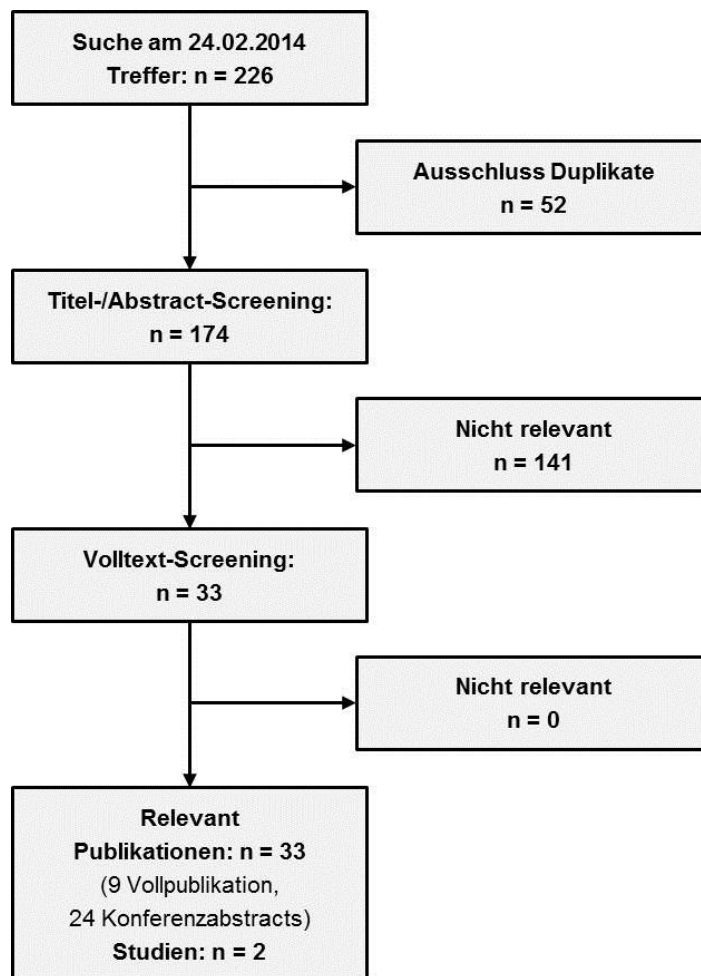


Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-3) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. 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 4-5: 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 pharmazeutischen Unternehmens enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein)	Status (abgeschlossen / abgebrochen / laufend)
CINC424A2352 COMFORT-II A Randomized Study of the JAK Inhibitor INC424 Tablets Compared to Best Available Therapy in Subjects with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia Myelofibrosis (PET-MF)	ClinicalTrials.gov: NCT00934544 (50) EU-CTR: EudraCT 2009-009858-24 (51) ICTRP: NCT00934544 (52)	ja	ja	Laufend, aber nicht mehr rekrutierend
INCB 18424-351 COMFORT-I A Randomized, Double-blind, Placebo-controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF)	ClinicalTrials.gov: NCT00952289 (53) ICTRP: NCT00952289 (54)	ja	ja	Laufend, aber nicht mehr rekrutierend
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. Abkürzungen: EU-CTR, EU Clinical Trials Register; ICTRP, International Clinical Trials Registry Platform				

Geben Sie an, welchen Stand die Information in Tabelle 4-5 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 Dokuments maßgeblichen Zeitpunkt liegen.

Die Recherche im Studienregister ClinicalTrials.gov sowie im ICTRP Search Portal wurde am 16. Januar 2014 durchgeführt, die Suche in den Registern Klinische Prüfungen PharmNet.Bund und EU-CTR erfolgte am 14. Januar 2014. Eine Aktualisierung fand am

24. Februar 2014 statt. Die dabei verwendeten Suchstrategien sind in Anhang 4-B dargestellt. Die Suche ergab insgesamt 82 Treffer (davon 32 Treffer im Register ClinicalTrials.gov, 4 Treffer im Register Klinische Prüfungen PharmNet.Bund, 13 Treffer im EU-CTR und 33 Treffer im ICTRP Search Portal). Nach Ausschluss der nicht relevanten Studien blieben nur zwei Studien übrig: die COMFORT-I-Studie und die COMFORT-II-Studie, die bereits in Abschnitt 4.3.1.1.1 aufgelistet wurden und der Bewertung des Zusatznutzens von Ruxolitinib im vorliegenden Dossier zugrunde liegen. Eine Liste der ausgeschlossenen Studien findet sich in Anhang 4-D.

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.

Tabelle 4-6: 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])	Register-eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
COMFORT-I	ja	nein	ja	ja (23;24;55;56)	ja (53;54)	ja (16;22;57-60)
COMFORT-II	ja	ja	nein	ja (31;61;62)	ja (50-52)	ja (63-65)
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 4-7: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
COMFORT-II	Randomisierte kontrollierte Phase-III-Studie offen parallel	Erwachsene Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) mit Intermediär-2- oder Hochrisikoprofil, die nicht für eine allogene Stamm- zelltransplantation in Betracht kamen	Ruxolitinib (15 mg bid bzw. 20 mg bid) (N = 146) BAT nach Ermessen des Prüfarztes (Kombinationen, Therapiewechsel, Nichtbehandlung erlaubt) (N = 73)	Primäre Datenanalyse nachdem alle Patienten 48 Wochen behandelt worden waren Weiterbehandlung ursprünglich bis zu Woche 144 geplant; durch Amendment 7 wurde die Studiendauer auf 5 Jahre verlängert Bei Zunahme des Milzvolumens um $\geq 25\%$ bzw. Splenektomie konnten Patienten unter Weiterführung bzw. Neuaufnahme der Ruxolitinib-Therapie in die Extensionsphase der Studie wechseln Aufgrund der Überlegenheit von Ruxolitinib in der Primäranalyse konnten ab Februar 2011 alle Patienten unabhängig von diesen Progressions- kriterien in die Extensions- phase der Studie wechseln (Amendment 5)	56 Zentren in 9 europäischen Ländern: Belgien, Deutschland, Frankreich, Großbritannien, Italien, Niederlande, Österreich, Schweden, Spanien Seit Juli 2009 bis voraussichtlich Januar 2015	<u>Primärer Endpunkt:</u> Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen <u>Sekundäre Endpunkte:</u> Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen Dauer des Ansprechens (Milzvolumenreduktion) Zeit bis zum ersten Ansprechen (Milzvolumenreduktion) Progressionsfreies Überleben Leukämiefreies Überleben Gesamtüberleben <u>Explorative Endpunkte:</u> Gesundheitsbezogene Lebensqualität (EORTC- QLQ-C30, FACT-Lym) und ECOG-Performance-Status (jeweils Veränderung gegenüber Studienbeginn)

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

COMFORT-I	Randomisierte kontrollierte Phase-III-Studie doppelblind Placebo-kontrolliert parallel	Erwachsene Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) mit Intermediär-2- oder Hochrisikoprofil, die refraktär, resistent oder intolerant gegenüber anderen Therapieoptionen waren oder aus anderen Gründen nicht für eine Behandlung mit diesen in Betracht kamen	Ruxolitinib (15 mg bid bzw. 20 mg bid) (N = 155) Placebo (N = 154)	Primäre Datenanalyse nachdem alle Patienten 24 Wochen und 50% der Patienten 36 Wochen behandelt worden waren Weiterbehandlung ursprünglich bis zu Woche 144 geplant; durch Amendment 4 wurde die Studiendauer auf 5 Jahre verlängert Bei Zunahme des Milzvolumens um $\geq 25\%$ konnten Patienten zur offenen Ruxolitinib-Therapie wechseln (vor Woche 24 musste zusätzlich eine definierte Symptomverschlechterung vorliegen) Nach der Primäranalyse im November 2010 wurde die Studie entblindet und alle Patienten konnten unabhängig vom Milzvolumen in den Ruxolitinib-Arm wechseln	68 Zentren in den USA, 6 Zentren in Kanada, 15 Zentren in Australien Seit August 2009 bis voraussichtlich Januar 2015	<u>Primärer Endpunkt:</u> Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen <u>Sekundäre Endpunkte:</u> Dauer des Ansprechens (Milzvolumenreduktion) Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen Prozentuale/absolute Veränderung des TSS gegenüber Studienbeginn nach 24 Wochen Gesamtüberleben <u>Explorative Endpunkte:</u> Leukämiefreies Überleben Gesundheitsbezogene Lebensqualität (EORTC-QLQ-C30) und ECOG-Performance-Status (jeweils Veränderung gegenüber Studienbeginn)
Abkürzungen: BAT, Best Available Therapy; bid, bis in die (lat., zweimal täglich); ECOG, Eastern Cooperative Oncology Group; EORTC-QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose						

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

Studie	<Gruppe 1>	<Gruppe 2>	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
COMFORT-II	Ruxolitinib entsprechend der Thrombozytenzahl (N = 146): > 200.000/ μ l: Anfangsdosis 20 mg bid > 100.000/ μ l und $\leq$ 200.000/ μ l: Anfangsdosis 15 mg bid Dosisanapassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit erlaubt Höchstsdosis: 25 mg bid	BAT nach Ermessen des Prüfarztes; Kombinationen, Therapiewechsel, Nichtbehandlung erlaubt (N = 73) Keine Medikation (N = 24) Medikation (N = 49)* - Hydroxycarbamid (N = 34) - Anagrelid (N = 4) - Cytarabin (N = 2) - Glucocorticoide (N = 12) - Epoetin- α (N = 5) - Thalidomid (N = 3) - Lenalidomid (N = 2) - Mercaptopurin (N = 3) - Thioguanin (N = 1) - Danazol (N = 3) - Interferon- α (N = 3) - Melphalan (N = 2) - Acetylsalicylsäure (N = 2) - Colchicin (N = 1)	
COMFORT-I	Ruxolitinib entsprechend der Thrombozytenzahl (N = 155): > 200.000/ μ l: Anfangsdosis 20 mg bid > 100.000/ μ l und $\leq$ 200.000/ μ l: Anfangsdosis 15 mg bid Dosisanapassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit erlaubt Höchstsdosis: 25 mg bid	Placebo (N = 154) (2x tägl. je 3 oder 4 Tabletten) Auch im Placebo-Arm wurden „Dosisanapassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten	
* manche Patienten erhielten mehrere Medikamente Abkürzungen: BAT, Best Available Therapy; bid, bis in die (lat., zweimal täglich)			

Tabelle 4-9: Auswahl häufiger Begleitmedikamente mit möglichem Bezug zur Behandlung von Myelofibrose-Symptomen in der Studie COMFORT-I

Begleitmedikationen *	Ruxolitinib (N = 155) n, %	Placebo (N = 151) n, %
Opioide	40 (25,8)	61 (40,4)
Morphium	10 (6,5)	13 (8,6)
Oxycodonhydrochlorid	14 (9,0)	28 (18,5)
Antirheumatika (nicht-steroidal)	58 (37,4)	56 (37,1)
Ibuprofen	33 (21,3)	30 (19,9)
Andere Analgetika/Antipyretika	91 (58,7)	101 (66,9)
Paracetamol	84 (54,2)	92 (60,9)
Acetylsalicylsäure	4 (2,6)	8 (5,3)
Antihistaminika (systemische)	38 (24,5)	45 (29,8)
Corticosteroide (systemische)	33 (21,3)	31 (20,5)
Antiemetika	19 (12,3)	22 (14,6)
Propulsiva	6 (3,9)	10 (6,6)
Motilitätshemmer	16 (10,3)	21 (13,9)
Ulku­stherapeutika	61 (39,4)	63 (41,7)
Omeprazol	23 (14,8)	22 (14,6)
Pantoprazol	12 (7,7)	8 (5,3)
Antazida	10 (6,5)	7 (4,6)
Laxantien	27 (17,4)	26 (17,2)
Antithrombotika	88 (56,8)	83 (55,0)
Acetylsalicylsäure	74 (47,7)	69 (45,7)
Warfarin	10 (6,5)	5 (3,3)
Antivirale Medikamente	7 (4,5)	9 (6,0)
Antibiotika		
Penicilline	23 (14,8)	8 (5,3)
Andere Beta-Lactam-Antibiotika	23 (14,8)	18 (11,9)
Makrolide/Lincosamide/Streptogramine	20 (12,9)	17 (11,3)
Quinolone	29 (18,7)	20 (13,2)
Andere Antibiotika	17 (11,0)	17 (11,3)
Antimykotika (systemische)	4 (2,6)	9 (6,0)
Eisenpräparate	15 (9,7)	15 (9,9)
Antidepressiva	37 (23,9)	28 (18,5)
Anxiolytika	24 (15,5)	25 (16,6)
Hypnotika/Sedativa	34 (21,9)	25 (16,6)
* Begleitmedikationen, die bei $\geq 5\%$ der Patienten in einem der beiden Studienarme eingesetzt wurden, ausgewählt nach Relevanz		

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

Charakteristikum	COMFORT-II		COMFORT-I	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Alter (in Jahren)				
Mittelwert ($\pm$ SD)	65,1 $\pm$ 9,7	65,2 $\pm$ 10,3	66,7 $\pm$ 8,8	68,7 $\pm$ 8,7
Median	67	66	66	70
Spannweite	35 – 83	35 – 85	43 – 91	40 – 86
Alter nach Kategorie (n, %)				
< 65 Jahre	69 (47,3)	36 (49,3)	70 (45,2)	52 (33,8)
$\geq$ 65 Jahre	77 (52,7)	37 (50,7)	85 (54,8)	102 (66,2)
Geschlecht (n, %)				
männlich	83 (56,8)	42 (57,5)	79 (51,0)	88 (57,1)*
weiblich	63 (43,2)	31 (42,5)	76 (49,0)	65 (42,2)*
Hautfarbe (n, %)				
Weiß	118 (80,8)	67 (91,8)	138 (89,0)	139 (90,3)
Sonstige	0 (0,0)	1 (1,4)	17 (11,0)	15 (9,7)
unbekannt [†]	28 (19,2)	5 (6,8)	0 (0,0)	0 (0,0)
Myelofibrose-Subtyp (n, %)				
PMF	77 (52,7)	39 (53,4)	70 (45,2)	84 (54,5)*
PPV-MF	48 (32,9)	20 (27,4)	50 (32,3)	47 (30,5)*
PET-MF	21 (14,4)	14 (19,2)	35 (22,6)	22 (14,3)*
IPSS-Risikogruppe (n, %)				
Hochrisiko	88 (60,3)	43 (58,9)	90 (58,1)	99 (64,3)
Intermediär-Risiko 2	58 (39,7)	29 (39,7)	64 (41,3)	54 (35,1)
unbekannt	0 (0,0)	1 (1,4)	1 (0,6)	1 (0,6)
ECOG-Performance-Status (n, %)				
0	58 (39,7)	26 (35,6)	47 (31,1)	38 (25,5)
1	77 (52,7)	37 (50,7)	87 (57,6)	82 (55,0)
2	10 (6,8)	9 (12,3)	14 (9,3)	25 (16,8)
3	1 (0,7)	1 (1,4)	3 (2,0)	4 (2,7)
fehlt	0 (0,0)	0 (0,0)	4 (2,6)	5 (3,2)
Milzvolumen (in cm³)				
n	144	72	155	153
Mittelwert ($\pm$ SD)	2662 $\pm$ 1351	2631 $\pm$ 1405	2746 $\pm$ 1247	2798 $\pm$ 1389
Median	2408	2318	2598	2566
Spannweite	451 – 7765	728 – 7701	478 – 7462	521 – 8881

Charakteristikum	COMFORT-II		COMFORT-I	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Milzgröße (cm unter dem Rippenbogen)				
n	146	72	155	153
Mittelwert (± SD)	14,9 ± 6,5	15,8 ± 6,7	16,1 ± 5,7	16,4 ± 6,3
Median	14	15	16	16
Spannweite	5 – 30	5 – 37	0 – 33	5 – 34
Milzgröße nach Kategorie (n, %)				
< 10 cm (n, %)	47 (32,2)	17 (23,3)	32 (20,6)	27 (17,5)
≥ 10 cm (n, %)	99 (67,8)	55 (75,3)	123 (79,4)	126 (81,8)
JAK2^{V617F}-Mutation (n, %)				
ja	110 (75,3)	49 (67,1)	113 (72,9)	123 (79,9)
nein	35 (24,0)	20 (27,4)	40 (25,8)	27 (17,5)
unbekannt	1 (0,7)	4 (5,5)	2 (1,3)	4 (2,6)
Vorbehandlung mit Hydroxycarbamid (n, %)	110 (75,3)	50 (68,5)	104 (67,1)	87 (56,5)
Vorausgegangene Milzbestrahlung (n, %)	0 (0,0)	4 (5,5)	1 (0,6)	0 (0,0)
Hämoglobin (g/l)				
Mittelwert (± SD)	k. A.	k. A.	108,3 ± 20,1	106,8 ± 21,2
Median			105	105
Spannweite			66 – 170	35 – 173
Hämoglobin < 10 g/dl (n, %)	66 (45,2)	38 (52,1)	k. A.	k. A.
Thrombozyten (10⁹/l)				
Mittelwert (± SD)	k. A.	k. A.	321,1 ± 201,9	279,5 ± 150,8
Median			262	238
Spannweite			81 – 984	100 – 887
Blasten im peripheren Blut ≥ 1% (n, %)	111 (76,0)	54 (74,0)	k. A.	k. A.
* Daten von einem Patienten fehlen, da die Unterlagen beim Umzug des Prüfzentrums verloren gingen † Daten von mehreren Patienten aus Frankreich fehlen aus Datenschutzgründen Abkürzungen: BAT, Best Available Therapy; ECOG, Eastern Cooperative Oncology Group; IPSS, International Prognostic Scoring System; k. A., keine Angabe; PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose; SD, Standardabweichung				

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.

COMFORT-II

Die Phase-III-Studie COMFORT-II (**C**Ontrolled **M**yl**o**Fibrosis study with **O**Ral JAK inhibitor **T**herapy) ist eine randomisierte, offene Multicenterstudie (31;62-64). In der Studie wurden Wirksamkeit und Sicherheit der Behandlung mit Ruxolitinib gegenüber der nach Ermessen des Prüfarztes besten verfügbaren Therapie (Best Available Therapy, BAT) bei Patienten mit Myelofibrose untersucht. Dabei wurden erwachsene Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) der IPSS-Risikostufen intermediär-2 bzw. hoch eingeschlossen, die nicht für eine allogene Stammzelltransplantation in Betracht kamen. Insgesamt wurden 219 Patienten im Verhältnis 2:1 stratifiziert nach Risikogruppe randomisiert und so der Therapie mit Ruxolitinib bzw. BAT zugewiesen (Abbildung 4-3).

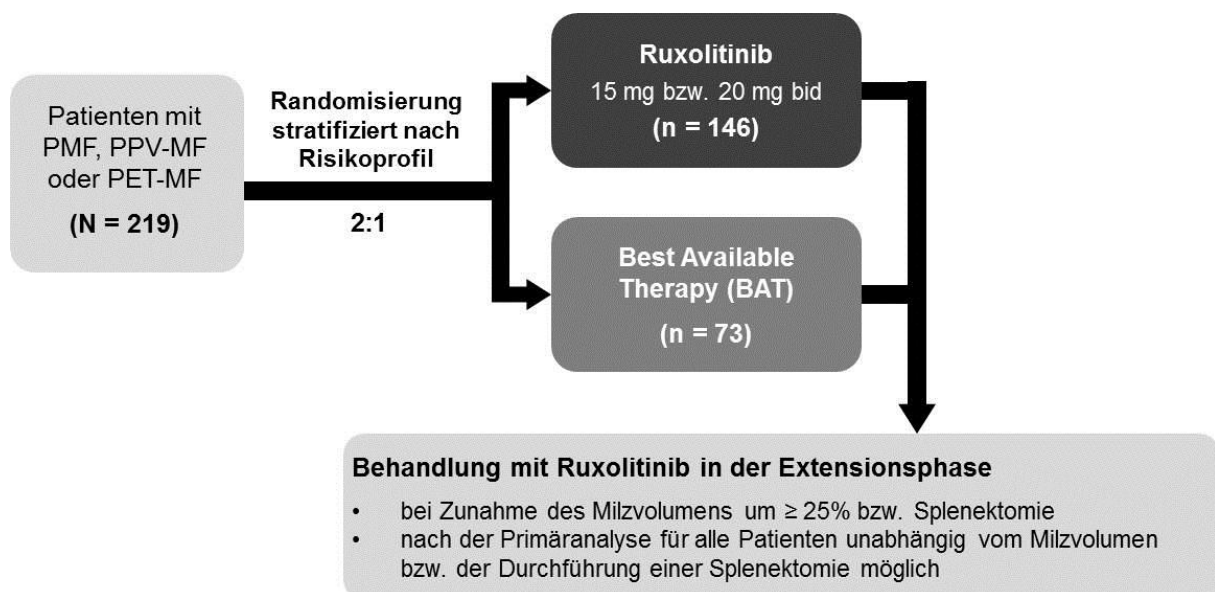


Abbildung 4-3: Design der Zulassungsstudie COMFORT-II

Bei der Studie handelt es sich um eine randomisierte, aktivkontrollierte Phase-III-Studie. Die Dosierung von Ruxolitinib betrug 15 mg bid bzw. 20 mg bid in Abhängigkeit von der zu Studienbeginn gemessenen Thrombozytenzahl. Im Vergleichsarm erhielten die Patienten die nach Ermessen des behandelnden Arztes beste verfügbare Therapie (BAT). BAT, Best Available Therapy; bid, bis in die (lat., zweimal täglich); PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose.

Der primäre Endpunkt der Studie war der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen. Sekundäre Endpunkte umfassten den Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen, die Dauer des Ansprechens sowie die Zeit bis zum ersten Ansprechen in Bezug auf die Milzvolumenreduktion, progressionsfreies Überleben (progression-free survival, PFS), leukämiefreies Überleben (leukemia-free survival, LFS) und Gesamtüberleben (overall survival, OS). Das Milzvolumen wurde zu Studienbeginn und dann alle 12 Wochen mittels Magnetresonanztomographie (MRT) oder Computertomographie (CT) erfasst. Die Auswertung der Aufnahmen erfolgte dabei durch einen unabhängigen zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war.

Im Interventionsarm erhielten die Patienten in Abhängigkeit von der zu Studienbeginn gemessenen Thrombozytenzahl Ruxolitinib in einer Dosierung von 15 mg bid bzw. 20 mg bid (zweimal täglich). Dosisanpassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit waren gemäß Studienprotokoll erlaubt, eine Höchstdosis von 25 mg bid (50 mg/Tag) durfte jedoch nicht überschritten werden. Im Vergleichsarm konnte der jeweilige Prüfarzt die nach seinem Ermessen beste verfügbare Therapie (BAT) auswählen, wobei auch eine Kombination mehrerer Therapieoptionen, ein Therapiewechsel sowie eine Nichtbehandlung möglich waren. Insgesamt erhielten 32,9% der Patienten keine medikamentöse Therapie („Nichtbehandlung“), die übrigen Patienten wurden vorwiegend mit Hydroxycarbamid (46,6%) und/oder Glucocorticoiden (16,4%) behandelt. Bei Zunahme des Milzvolumens um $\geq 25\%$ oder der Durchführung einer Splenektomie war es Patienten aus beiden Studienarmen gestattet, die Studie unter Weiterführung bzw. Neuaufnahme der Ruxolitinib-Therapie in der Extensionsphase fortzusetzen. Aufgrund der deutlichen Überlegenheit von Ruxolitinib in der Primäranalyse konnten ab Februar 2011 dann alle Patienten unabhängig von den erwähnten Progressionskriterien zur Behandlung mit Ruxolitinib in die Extensionsphase wechseln (Amendment 5 des Studienprotokolls). Die Weiterbehandlung war ursprünglich bis zu Woche 144 geplant. Um Langzeitdaten zur Wirksamkeit und Sicherheit der Ruxolitinib-Therapie erheben zu können, wurde die Studiendauer mit Amendment 7 des Studienprotokolls jedoch auf 5 Jahre verlängert. Die Studie wird voraussichtlich bis Januar 2015 fortgesetzt (62).

Insgesamt wurden 219 Patienten zwischen Juli 2009 und Januar 2010 auf die beiden Studienarme randomisiert: 146 Patienten wurden der Ruxolitinib-Gruppe zugeteilt und 73 Patienten der BAT-Gruppe („ITT-Population“) (31;64). Alle Patienten erhielten mindestens eine Dosis der Studienmedikation, so dass die Sicherheitspopulation der ITT-Population entspricht (31). Der Patientenfluss ist in Anhang 4-E mittels CONSORT-Flow-Chart dargestellt (Abbildung 4-15 und Abbildung 4-16).

Die Basischarakteristika der Patienten in den beiden Studienarmen zu Studienbeginn waren gut ausgeglichen (Tabelle 4-10) (20;31;64). Das mediane Alter der Patienten betrug 67 Jahre im Ruxolitinib-Arm und 66 Jahre im BAT-Arm. Der Anteil der männlichen Patienten war mit durchschnittlich 57% geringfügig höher als der der weiblichen Patienten. Etwa die Hälfte der Patienten hatte eine primäre Myelofibrose, bei einem Drittel war eine PPV-MF und bei den übrigen eine PET-MF diagnostiziert worden. Aufgrund der nach IPSS-Risikoscore stratifizierten Randomisierung war die Verteilung der Patienten auf die Risikogruppen ausgewogen. Rund 60% der Patienten in jedem Studienarm gehörten dabei zur Hochrisikogruppe. Auch Parameter zur Erfassung der Krankheitsschwere wie das Milzvolumen bzw. die Milzgröße, verschiedene krankheitsrelevante Blutwerte (Hämoglobin, Thrombozyten) sowie der Anteil der Patienten mit $JAK2^{V617F}$ -Mutation waren in beiden Gruppen vergleichbar.

Die primäre Datenanalyse wurde am 4. Januar 2011 durchgeführt, nachdem alle Patienten mindestens 48 Wochen lang behandelt worden waren (bzw. die Behandlung vorzeitig abgebrochen hatten) (31). Zu diesem Zeitpunkt wurden noch 62,3% der Patienten im Ruxolitinib-Arm mit der ursprünglich zugewiesenen Studienmedikation behandelt, während dies nur bei 42,5% der Patienten im BAT-Arm der Fall war. Die mediane Behandlungsdauer betrug 51,4 Wochen unter der Therapie mit Ruxolitinib im Vergleich zu 45,1 Wochen unter der Therapie mit BAT. Ursache für den Abbruch der Therapie in der BAT-Gruppe war in den meisten Fällen das Zurückziehen der Einwilligungserklärung des Patienten, nämlich bei 12,3% der Patienten im Vergleich zu 1,4% der Patienten in der Ruxolitinib-Gruppe (Abbildung 4-15). Ein 3-Jahres-Update erfolgte am 1. Dezember 2012, nachdem alle Patienten mindestens 144 Wochen lang behandelt worden waren (62). Zu diesem Zeitpunkt wurden noch 45,2% der Patienten im Ruxolitinib-Arm weiterhin mit Ruxolitinib behandelt. Von den Patienten der BAT-Gruppe erhielt kein Patient mehr die ursprünglich zugewiesene Studienmedikation. Insgesamt waren 61,6% der Patienten aus dem BAT-Arm der Studie zur Behandlung mit Ruxolitinib in die Extensionsphase der Studie gewechselt, 38,4% der BAT-Patienten hatten die Studie vollständig beendet (Abbildung 4-16). Dementsprechend betrug die mediane Behandlungsdauer 136,3 Wochen unter der Therapie mit Ruxolitinib im Vergleich zu 45,1 Wochen unter der Therapie mit BAT.

COMFORT-I

Die Phase-III-Studie COMFORT-I (**C**Ontrolled **M**yl**o**Fibrosis study with **O**Ral JAK inhibitor **T**herapy) ist eine randomisierte, doppelblinde Multicenterstudie, in der Wirksamkeit und Sicherheit der Behandlung mit Ruxolitinib gegenüber Placebo untersucht wurden (22-24;55;56;59). In die Studie wurden erwachsene Patienten mit primärer oder sekundärer

Myelofibrose (PMF, PPV-MF, PET-MF) der IPSS-Risikostufen intermediär-2 bzw. hoch eingeschlossen. Darüber hinaus mussten die Patienten refraktär, resistent oder intolerant gegenüber anderen Therapieoptionen sein oder – nach Einschätzung des behandelnden Arztes – aus anderen Gründen nicht für eine Behandlung mit diesen in Betracht kommen. Insgesamt wurden 309 Patienten im Verhältnis 1:1 randomisiert und so der Therapie mit Ruxolitinib bzw. Placebo zugewiesen (Abbildung 4-4).

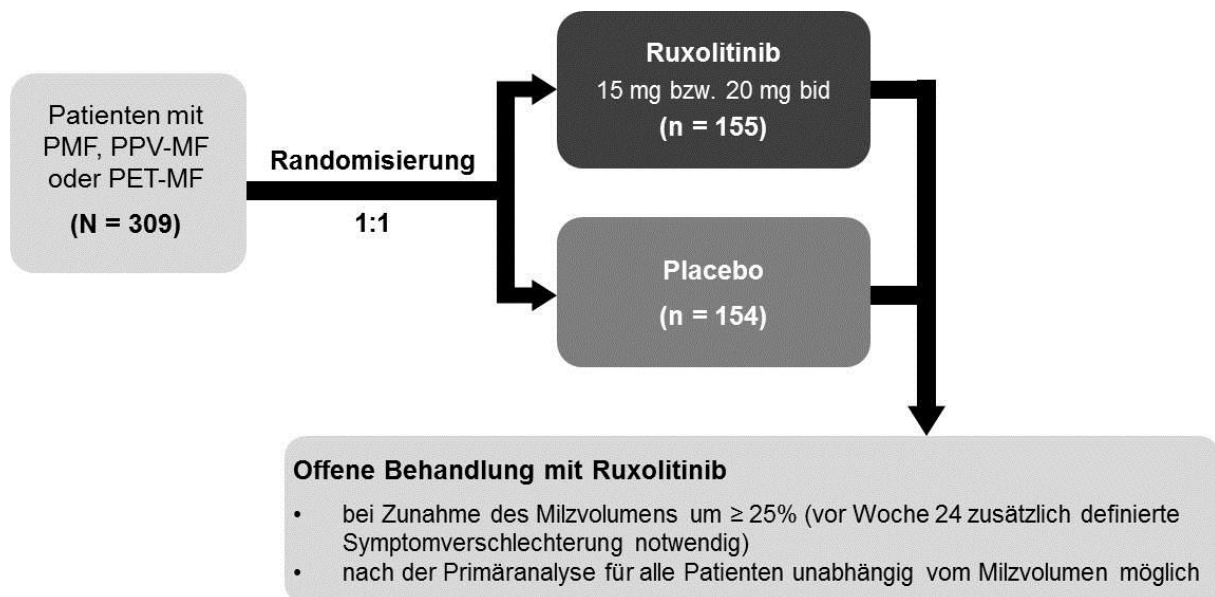


Abbildung 4-4: Design der Zulassungsstudie COMFORT-I

Bei der Studie handelt es sich um eine randomisierte, doppelblinde, placebo-kontrollierte Phase-III-Studie. Die Dosierung von Ruxolitinib betrug 15 mg bid bzw. 20 mg bid in Abhängigkeit von der zu Studienbeginn gemessenen Thrombozytenzahl. Bid, bis in die (lat., zweimal täglich); PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose.

Der primäre Endpunkt der Studie war der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen. Sekundäre Endpunkte umfassten die Dauer des Ansprechens in Bezug auf die Milzvolumenreduktion, den Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) des MFSAF-Fragebogens nach 24 Wochen, die prozentuale sowie absolute Veränderung des TSS gegenüber Studienbeginn nach 24 Wochen und das Gesamtüberleben (overall survival, OS). Das Milzvolumen wurde zu Studienbeginn und dann alle 12 Wochen mittels Magnetresonanztomographie (MRT) oder Computertomographie (CT) erhoben und durch einen unabhängigen zentralen Reviewer ausgewertet. Die Erfassung der Krankheitssymptome mit Hilfe des MFSAF-Fragebogens erfolgte täglich bis zu Woche 24.

Im Interventionsarm erhielten die Patienten in Abhängigkeit von der zu Studienbeginn gemessenen Thrombozytenzahl Ruxolitinib in einer Dosierung von 15 mg bid bzw. 20 mg bid (zweimal täglich). Dosisanpassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit waren gemäß Studienprotokoll erlaubt, eine Höchstdosis von 25 mg bid (50 mg/Tag) durfte jedoch nicht überschritten werden. Auch im Placebo-Arm wurden entsprechende „Dosisanpassungen“ durchgeführt, um die Verblindung von Patienten und Behandlern zu gewährleisten. Bei Zunahme des Milzvolumens um $\geq 25\%$ war es Patienten aus der Placebo-Gruppe gestattet, zur offenen Ruxolitinib-Therapie zu wechseln. Vor Woche 24 musste zusätzlich auch eine definierte Symptomverschlechterung vorliegen. Nach der primären Datenanalyse im November 2010 wurde die Studie entblindet und alle Patienten konnten unabhängig vom Milzvolumen in den Ruxolitinib-Arm der Studie wechseln. Die Weiterbehandlung war ursprünglich bis zu Woche 144 geplant. Um Langzeitdaten zur Wirksamkeit und Sicherheit der Ruxolitinib-Therapie erheben zu können, wurde die Studiendauer mit Amendment 4 des Studienprotokolls jedoch auf 5 Jahre verlängert (24). Die Studie wird voraussichtlich im Januar 2015 beendet.

Insgesamt wurden 309 Patienten zwischen August 2009 und April 2010 auf die beiden Studienarme randomisiert: 155 Patienten wurden der Ruxolitinib-Gruppe zugeteilt und 154 Patienten der Placebo-Gruppe („ITT-Population“) (23). Nur 307 der ursprünglich randomisierten Patienten erhielten mindestens eine Dosis der Studienmedikation: 155 Patienten erhielten Ruxolitinib und 152 Patienten Placebo. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (23). Der Patientenfluss ist in Anhang 4-E mittels CONSORT-Flow-Chart dargestellt (Abbildung 4-17, Abbildung 4-18, Abbildung 4-19).

Die Basischarakteristika der Patienten in den beiden Studienarmen zu Studienbeginn waren im Wesentlichen vergleichbar (Tabelle 4-10) (20;22;23). Etwa die Hälfte der Patienten hatte eine primäre Myelofibrose, ein Drittel eine PPV-MF und die übrigen Patienten eine PET-MF. Das mediane Alter der Patienten im Placebo-Arm war mit 70 Jahren geringfügig höher als im Ruxolitinib-Arm (66 Jahre). Auch der Anteil der Patienten mit Hochrisikoprofil sowie der Anteil der Patienten mit $JAK2^{V617F}$ -Mutation war in der Placebo-Gruppe etwas höher als in der Ruxolitinib-Gruppe (64,3% vs. 58,1% bzw. 79,9% vs. 72,9%). Dagegen waren Patienten in der Ruxolitinib-Gruppe etwas häufiger mit Hydroxycarbamid vorbehandelt (67,1% vs. 56,5%). Parameter zur Erfassung der Krankheitsschwere wie das Milzvolumen bzw. die Milzgröße und verschiedene krankheitsrelevante Blutwerte (Hämoglobin, Thrombozyten) waren zwischen den Gruppen gut ausgeglichen.

Die primäre Datenanalyse wurde am 2. November 2010 durchgeführt, nachdem alle Patienten mindestens 24 Wochen und 50% der Patienten 36 Wochen lang behandelt worden waren (bzw. die Behandlung vorzeitig abgebrochen hatten) (23). Zu diesem Zeitpunkt wurden noch 86,5% der Patienten im Ruxolitinib-Arm mit der ursprünglich zugewiesenen Studienmedikation behandelt, während dies nur bei 51,7% der Patienten im Placebo-Arm der Fall war. Die mediane Behandlungsdauer betrug 33,7 Wochen unter der Therapie mit Ruxolitinib und 30,1 Wochen unter der Therapie mit Placebo. Ursache für den Abbruch der Therapie in der Placebo-Gruppe war in den meisten Fällen die Progression der Erkrankung, nämlich bei 7,9% der Patienten im Vergleich zu 1,9% der Patienten in der Ruxolitinib-Gruppe (Abbildung 4-17). Ein 3-Jahres-Update erfolgte am 25. Januar 2013, nachdem alle Patienten mindestens 144 Wochen lang behandelt worden waren (24). Zu diesem Zeitpunkt wurden noch 49,7% der Patienten im Ruxolitinib-Arm weiterhin mit Ruxolitinib behandelt. In der Placebo-Gruppe hatten dagegen alle Patienten die Studie entweder abgebrochen (26,5%) oder waren in den Ruxolitinib-Arm der Studie gewechselt (73,5%) (Abbildung 4-19). Dementsprechend betrug die mediane Behandlungsdauer 144,6 Wochen unter der Therapie mit Ruxolitinib im Vergleich zu 37,1 Wochen unter der Therapie mit Placebo.

Einschluss der Studien in die Nutzenbewertung

Sowohl die COMFORT-II-Studie als auch die COMFORT-I-Studie entspricht den in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien, die bei der Auswahl von Studien für die Nutzenbewertung von Ruxolitinib heranzuziehen sind. Nach diesen Kriterien soll der klinische Nutzen der Therapie mit Ruxolitinib bei der Behandlung von krankheitsbedingter Splenomegalie oder Symptomen bei Erwachsenen mit primärer Myelofibrose, Post-Polycythaemia-vera-Myelofibrose oder Post-Essentieller-Thrombozythämie-Myelofibrose gezeigt werden. Als Vergleichstherapie ist die nach Ermessen des behandelnden Arztes für den Patienten aktuell bestmögliche, unterstützende Behandlung (BSC) heranzuziehen.

In die beiden Studien COMFORT-I und COMFORT-II wurden ausschließlich Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) eingeschlossen. Entsprechend den Einschlusskriterien mussten die Patienten außerdem eine tastbare Milzlänge von mindestens 5 cm unter dem Rippenbogen aufweisen, was mit der Definition einer Splenomegalie übereinstimmt (23;31). Tatsächlich lag das mediane Milzvolumen studien- und gruppenübergreifend bei etwa 2500 cm³. Dies entspricht einer Zunahme des Milzvolumens auf mehr als das 10-fache der normalen Milzgröße von circa 200 cm³ (23;31;66-68). Die Erfassung der Basischarakteristika in COMFORT-II zeigt außerdem, dass 67,6% der Patienten zu Studienbeginn konstitutionelle Symptome wie Fieber, Nachtschweiß oder Gewichtsverlust aufwiesen (31). Auch in der COMFORT-I-Studie wurden bei der

Mehrzahl der Patienten für die Erkrankung typische Symptome dokumentiert (23). Das Patientenkollektiv beider Studien umfasst also Patienten mit primärer oder sekundärer Myelofibrose mit krankheitsbedingter Splenomegalie und/oder Symptomen und entspricht damit der in Abschnitt 4.2.2 definierten Patientenpopulation, welche wiederum durch die in Deutschland geltende Zulassungsindikation bestimmt wird. Auch der Gemeinsame Bundesausschuss bestätigt auf Seite 51 der Nutzenbewertung zu Ruxolitinib vom 17. Dezember 2012, dass die in beiden Studien untersuchten Patientenpopulationen die gemäß Fachinformation zugelassene Zielpopulation repräsentieren (38).

In der COMFORT-II-Studie wurde die Ruxolitinib-Behandlung mit der aktuell besten verfügbaren Therapie (BAT) verglichen. Für die Auswahl dieser Therapie war der jeweilige Prüfarzt zuständig, wobei auch eine Kombination mehrerer Therapieoptionen, ein Therapiewechsel sowie eine Nichtbehandlung möglich waren (Tabelle 4-8). Die Kontrollintervention der COMFORT-II-Studie entspricht damit der in Abschnitt 4.2.2 definierten Vergleichstherapie Best Supportive Care. Auch der Gemeinsame Bundesausschuss stellt auf Seite 52 der Nutzenbewertung zu Ruxolitinib fest, dass die Studie COMFORT-II für die Beurteilung des Therapieeffektes gegenüber der gängigen Standardtherapie im Indikationsgebiet geeignet ist (38). Während Patienten in COMFORT-II lediglich nicht für eine allogene Stammzelltransplantation in Frage kommen durften, mussten Patienten der COMFORT-I-Studie gegenüber anderen Therapien refraktär, resistent oder intolerant sein oder aus anderen Gründen nicht für eine Behandlung mit diesen in Frage kommen (23;31). Das Patientenkollektiv von COMFORT-I repräsentiert damit den Teil der Zielpopulation, dem in der Kontrollgruppe nur die „Nichtbehandlung“ als beste verfügbare Therapie zur Verfügung steht, weshalb der Vergleich in dieser Studie gegenüber Placebo vorgenommen wurde. Da die „Nichtbehandlung“ im vorliegenden Anwendungsgebiet Bestandteil der besten verfügbaren Therapie ist, erfüllt auch die Studie COMFORT-I die in Abschnitt 4.2.2 definierten Einschlusskriterien bezüglich der zu untersuchenden Vergleichstherapie. Tatsächlich erhielten im BAT-Arm der COMFORT-II-Studie ebenfalls 32,9% der Patienten keine medikamentöse Therapie („Nichtbehandlung“) (31). Darüber hinaus war der Einsatz weiterer Begleitmedikationen in der COMFORT-I-Studie entsprechend dem Studienprotokoll erlaubt. Diese waren zwar hinsichtlich der zu untersuchenden antineoplastischen Therapien, nicht aber in Bezug auf unterstützende Maßnahmen und Arzneimittel eingeschränkt, die für die Sicherheit und das Wohlbefinden des Patienten erforderlich waren (23). Die Liste der im Studienverlauf tatsächlich verabreichten Begleitmedikationen enthält beispielsweise Opioide, nicht-steroidale Antirheumatika (non-steroidal anti-inflammatory drugs, NSAIDs), andere Analgetika bzw. Antipyretika, Antihistaminika und Glucocorticoide, Medikamente zur Behandlung von Infektionserkrankungen sowie verschiedene Magen-Darm-Therapeutika (Tabelle 4-9). Auch Antidepressiva, Sedativa/Hypnotika und Anxiolytika kamen zum Einsatz

(23). In beiden Studien erhielten die Patienten unter der Vergleichstherapie somit eine patientenindividuell unterstützende Behandlung im Sinne von Best Supportive Care, wobei sich Unterschiede in der Zusammensetzung dieser Behandlung durch das differierende Patientenkollektiv ergeben. Der Übersichtlichkeit wegen wird die in den Studien verwendete Bezeichnung der jeweiligen Vergleichstherapie (BAT bzw. Placebo) im Folgenden beibehalten.

Zusammenfassend lässt sich also festhalten, dass die beiden Studien COMFORT-I und COMFORT-II entsprechend den in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien für die Nutzenbewertung von Ruxolitinib heranzuziehen sind. Die COMFORT-I-Studie wurde in den USA, Kanada und Australien durchgeführt, während sich die Studienzentren für COMFORT-II ausschließlich in westeuropäischen Ländern befanden (Belgien, Deutschland, Frankreich, Großbritannien, Italien, Niederlande, Österreich, Schweden, Spanien). Beide Studien schlossen außerdem überwiegend Patienten weißer Hautfarbe ein (siehe Tabelle 4-10). Die Ergebnisse können damit uneingeschränkt auf den deutschen Versorgungskontext übertragen werden.

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			
COMFORT-II	ja	ja	nein	nein	ja	ja	niedrig
COMFORT-I	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Von einem relevanten Verzerrungspotenzial ist nicht auszugehen, da eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da alle wichtigen Endpunkte objektiv erfasst werden konnten und damit stabil gegenüber möglichen Verzerrungen sind. So wurde der Endpunkt „Reduktion der Milzgröße“ durch klar definierte Parameter erfasst, die nicht von der subjektiven Beurteilung des Prüfers oder des Patienten beeinflusst werden. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann. Auch die Endpunkte „Gesamtüberleben“, „Leukämiefreies Überleben“ und „Progressionsfreies Überleben“ sind unabhängig von subjektiven Einflussfaktoren. Insgesamt ist das Verzerrungspotenzial für die COMFORT-II-Studie daher als „niedrig“ einzustufen. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für die COMFORT-I-Studie als „niedrig“ eingestuft, da eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte, sowohl Patient als auch Behandler verblindet waren und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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	Reduktion der Milzgröße	Linderung der Krankheits-symptomatik	Gesamtüberleben	Leukämiefreies Überleben	Progressionsfreies Überleben	Veränderung des ECOG-Performance-Status	Gesundheitsbezogene Lebensqualität	Verträglichkeit
COMFORT-II	ja	nein	ja	ja	ja	ja	ja	ja
COMFORT-I	ja	ja	ja	ja	nein	ja	ja	ja

Im Folgenden sind die Endpunkte Reduktion der Milzgröße (in den Operationalisierungen Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen, prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 bzw. 48 Wochen, Zeit bis zum ersten Ansprechen, Dauer des Ansprechens) sowie Linderung der Krankheitssymptomatik (mit den Operationalisierungen Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen, prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen) dargestellt. Als weitere Zielkriterien werden Gesamtüberleben, leukämiefreies Überleben, progressionsfreies Überleben, die Veränderung des ECOG-Performance-Status, gesundheitsbezogene Lebensqualität (Veränderung des EORTC-QLQ-C30 bzw. FACT-Lym gegenüber Studienbeginn) und Verträglichkeit untersucht.

4.3.1.3.1 Endpunkte – RCT

4.3.1.3.1.1 Reduktion der Milzgröße – 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 „Reduktion der Milzgröße“

Studie	Operationalisierung
COMFORT-II	Der Endpunkt „Reduktion der Milzgröße“ wurde durch vier Operationalisierungen erfasst: 1. <u>Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen</u> Der Rückgang des Milzvolumens wurde mittels MRT oder CT gemessen und der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen im Vergleich zu Studienbeginn angegeben. Patienten ohne Milzvolumen-Bestimmung nach 24 bzw. 48 Wochen wurden dabei als Non-Responder klassifiziert. Nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn wurden in die Analyse einbezogen („ITT-Population, evaluierbare Patienten“). Im Rahmen einer Sensitivitätsanalyse wurde der Anteil der Patienten mit einer Reduktion des Milzvolumens zum jeweiligen Zeitpunkt zusätzlich unter Einschluss aller Patienten kalkuliert. Patienten ohne auswertbare Milzvolumen-Messung zu Studienbeginn wurden dabei als Non-Responder gewertet („ITT-Population“). 2. <u>Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 bzw. 48 Wochen</u> Der Rückgang des Milzvolumens wurde mittels MRT oder CT gemessen und die prozentuale Veränderung des Milzvolumens nach 24 bzw. 48 Wochen im Vergleich zum Baseline-Wert zu Studienbeginn angegeben. 3. <u>Dauer des Ansprechens</u> Die Dauer des Ansprechens in Bezug auf die Milzvolumenreduktion war definiert als die Zeit vom ersten Erreichen einer Milzvolumenreduktion von $\geq 35\%$ bis zum Verlust dieses Ansprechens bei gleichzeitiger Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie. Falls kein Ereignis auftrat, erfolgte eine Zensurierung zum Zeitpunkt der letzten Milzuntersuchung. 4. <u>Zeit bis zum ersten Ansprechen</u> Die Zeit bis zum ersten Ansprechen in Bezug auf die Milzvolumenreduktion war definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Erreichen einer Milzvolumenreduktion um $\geq 35\%$. Die radiologischen Untersuchungen wurden durch die lokalen Studienärzte durchgeführt. Die Auswertung der Aufnahmen erfolgte durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war. Die Analyse des Milzvolumens durch MRT bzw. CT erfolgte zum Zeitpunkt der Baseline-Untersuchung, nach 12, 24, 36 und 48 Wochen sowie danach alle 12 Wochen. Ab November 2012 wurden radiologische Milzvolumenuntersuchungen nur noch alle 24 Wochen durchgeführt.
COMFORT-I	Der Endpunkt „Reduktion der Milzgröße“ wurde durch vier Operationalisierungen erfasst: 1. <u>Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen</u> Der Rückgang des Milzvolumens wurde mittels MRT oder CT gemessen und der Anteil der Patienten mit einer Milzvolumenreduktion um $\geq 35\%$ nach 24 Wochen im Vergleich zu Studienbeginn angegeben. Patienten ohne Milzvolumen-Bestimmung nach 24 Wochen wurden als Non-Responder klassifiziert. Nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn wurden in die Analyse einbezogen („ITT-Population, evaluierbare Patienten“). Im Rahmen einer Sensitivitätsanalyse wurde der Anteil der

Patienten mit einer Reduktion des Milzvolumens zum jeweiligen Zeitpunkt zusätzlich unter Einschluss aller Patienten kalkuliert. Patienten ohne auswertbare Milzvolumen-Messung zu Studienbeginn wurden dabei als Non-Responder gewertet („ITT-Population“).

2. Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 Wochen

Der Rückgang des Milzvolumens wurde mittels MRT oder CT gemessen und die prozentuale Veränderung des Milzvolumens nach 24 Wochen im Vergleich zum Baseline-Wert zu Studienbeginn angegeben.

3. Dauer des Ansprechens

Die Dauer des Ansprechens in Bezug auf die Milzvolumenreduktion war definiert als die Zeit vom ersten Erreichen einer Milzvolumenreduktion von $\geq 35\%$ bis zum Verlust dieses Ansprechens bei gleichzeitiger Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie. Falls kein Ereignis auftrat, erfolgte eine Zensurierung zum Zeitpunkt der letzten Milzuntersuchung.

4. Zeit bis zum ersten Ansprechen

Die Zeit bis zum ersten Ansprechen in Bezug auf die Milzvolumenreduktion war definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Erreichen einer Milzvolumenreduktion um $\geq 35\%$.

Die radiologischen Untersuchungen wurden durch die lokalen Studienärzte durchgeführt. Die Auswertung der Aufnahmen erfolgte durch einen unabhängigen, zentralen Reviewer.

Die Analyse des Milzvolumens durch MRT bzw. CT erfolgte zum Zeitpunkt der Baseline-Untersuchung, nach 12, 24, 36, 48, 60 und 72 Wochen sowie danach alle 24 Wochen.

Abkürzungen: CT, Computertomographie; ITT, Intention to treat; MRT, Magnetresonanztomographie

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 den Endpunkt „Reduktion der Milzgröße“ 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
Operationalisierung: Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen						
COMFORT-II	niedrig	ja*	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 bzw. 48 Wochen						
COMFORT-II	niedrig	ja*	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Dauer des Ansprechens						
COMFORT-II	niedrig	ja*	NA [†]	ja	ja	niedrig
COMFORT-I	niedrig	ja	NA [†]	ja	ja	niedrig
Operationalisierung: Zeit bis zum ersten Ansprechen						
COMFORT-II	niedrig	ja*	NA [†]	ja	ja	niedrig
COMFORT-I	niedrig	ja	NA [†]	ja	ja	niedrig
* Die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens erfolgte durch einen unabhängigen, zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war. [†] Bei der vorliegenden Untersuchung handelt es sich um einen Sonderfall, bei dem eine Auswertung nach dem ITT-Prinzip nicht sinnvoll wäre. Die Frage nach der adäquaten Umsetzung des ITT-Prinzips ist somit nicht zutreffend. Abkürzungen: NA, not applicable						

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Milzvolumen durch klar definierte Parameter erfasst wurde, die nicht von der subjektiven Beurteilung des Prüfers

beeinflusst werden und somit stabil gegenüber möglichen Verzerrungen sind. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

Stellen Sie die Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ 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 den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel

Reduktion der Milzgröße	COMFORT-II		COMFORT-I	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Anteil der Patienten mit Milzvolumenreduktion ≥ 35%: ITT-Population, evaluierbare Patienten				
Nach 24 Wochen				
N (gesamt)	144	72	155	153
Ansprechen (n, %)	46 (31,9)	0 (0,0)	65 (41,9)	1 (0,7)
Differenz vs. BAT/Placebo [95%-KI]	31,9 [17,7 – 45,3]		41,3 [30,7 – 51,0]	
RR vs. BAT/Placebo [95%-KI] [†]	46,8 [2,9 – 749,1]		64,2 [9,0 – 456,6]	
p-Wert	< 0,0001 [*]		< 0,0001 ^{**}	
Nach 48 Wochen				
N (gesamt)	144	72		
Ansprechen (n, %)	41 (28,5)	0 (0,0)		
Differenz vs. BAT/Placebo [95%-KI]	28,5 [14,2 – 42,0]		-	
RR vs. BAT/Placebo [95%-KI] [†]	41,8 [2,6 – 669,7]			
p-Wert	< 0,0001 [*]			
Anteil der Patienten mit Milzvolumenreduktion ≥ 35%: ITT-Population (Sensitivitätsanalyse)				
Nach 24 Wochen				
N (gesamt)	146	73	155	154
Ansprechen (n, %)	46 (31,5)	0 (0,0)	65 (41,9)	1 (0,6)
Differenz vs. BAT/Placebo [95%-KI]	31,5 [17,4 – 44,8]		41,3 [30,9 – 51,2]	
RR vs. BAT/Placebo [95%-KI] [†]	46,8 [2,9 – 749,1]		64,6 [9,1 – 459,6]	
p-Wert	< 0,0001 [*]		< 0,0001 ^{**}	
Nach 48 Wochen				
N (gesamt)	146	73		
Ansprechen (n, %)	41 (28,1)	0 (0,0)		
Differenz vs. BAT/Placebo [95%-KI]	28,1 [13,9 – 41,5]		-	
RR vs. BAT/Placebo [95%-KI] [†]	41,8 [2,6 – 669,8]			
p-Wert	< 0,0001 ^{**}			

[†] Bei Auftreten von Nullzellen wurde zur Berechnung des relativen Risikos ein Korrekturwert von 0,5 zu jeder Zellohäufigkeit der entsprechenden Vierfeldertafel addiert. Auf die Darstellung des Odds Ratios wurde aufgrund der geringen Zahl der Ereignisse in der Kontrollgruppe verzichtet.

^{*} Berechnung des p-Werts mit dem exakten Cochran-Mantel-Haenszel (CMH)-Test

^{**} Berechnung des p-Werts mit dem exakten Fisher-Test

Reduktion der Milzgröße	COMFORT-II		COMFORT-I	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn				
Nach 24 Wochen				
N (gesamt)	125	45	139	106
Mittelwert (± SD)	-29,23 ± 19,48	2,70 ± 14,12	-31,6 ± 18,92	8,1 ± 15,31
Median	-27,5	5,1	-33,0	8,5
Spannweite	-71,2 – 36,9	-33,3 – 29,7	-75,9 – 25,1	-46,4 – 48,8
Differenz vs. BAT/Placebo [95%-KI]	-31,7 [-38,0; -25,5] [†]		-39,7 [-44,1; -35,3] [†]	
p-Wert	< 0,0001 [*]		< 0,0001 [*]	
Nach 48 Wochen				
N (gesamt)	98	34		
Mittelwert (± SD)	-30,05 ± 22,12	7,34 ± 12,99		
Median	-28,3	8,5	-	
Spannweite	-78,6 – 14,7	-20,0 – 32,9		
Differenz vs. BAT/Placebo [95%-KI]	-36,7 [-44,7; -28,8] [†]			
p-Wert	< 0,0001 [*]			
[†] Berechnung der LSM (Least Square Means)-Differenz mit zugehörigem 95%-Konfidenzintervall auf Basis einer Kovarianzanalyse mit dem Milzvolumen zu Studienbeginn als Kovariate				
[*] Berechnung des p-Werts mit Hilfe der oben genannten Kovarianzanalyse				
Dauer des Ansprechens (anhaltende Milzvolumenreduktion ≥ 35%)*				
N (gesamt)	75	1	91	k. A.
Zahl der Ereignisse (n, %)	31 (41,3)	0 (0,0)	34 (37,4)	
Zahl der Zensierungen (n, %)	44 (58,7)	1 (100)	57 (62,6)	
Mediane Dauer [95%-KI]	NA	NA	135,0 [107,6 – NA]	
Kaplan-Meier-Schätzer				
Nach 24 Wochen [95%-KI]	0,86 [0,76 – 0,92]	NA	k. A.	k. A.
Nach 48 Wochen [95%-KI]	0,73 [0,61 – 0,82]	NA	0,76 [0,66 – 0,84]	k. A.
Nach 72 Wochen [95%-KI]	0,68 [0,56 – 0,78]	NA	k. A.	k. A.
Nach 96 Wochen [95%-KI]	0,61 [0,49 – 0,72]	NA	0,67 [0,55 – 0,76]	k. A.
Nach 120 Wochen [95%-KI]	0,56 [0,43 – 0,67]	NA	k. A.	k. A.
Nach 144 Wochen [95%-KI]	0,50 [0,36 – 0,63]	NA	0,27 [0,02 – 0,63]	k. A.
[*] Schätzung der Dauer des Ansprechens mit Hilfe der Kaplan-Meier-Methode				

Reduktion der Milzgröße	COMFORT-II		COMFORT-I	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Zeit bis zum ersten Ansprechen (Milzvolumenreduktion $\geq 35\%$)*				
N (gesamt)	75	1	93	k. A.
Mediane Zeit zum ersten Ansprechen	12,43	15,43	12,7	
[95%-KI]	[12,14 – 23,57]	[NA – NA]	[12,3 – 13,1]	
Kaplan-Meier-Schätzer				
Nach 12 Wochen	0,21	0,0	0,20	k. A.
[95%-KI]	[0,13 – 0,31]	[NA – NA]	[0,14 – 0,30]	
Nach 24 Wochen	0,61	1,0	0,70	k. A.
[95%-KI]	[0,49 – 0,71]	[NA – NA]	[0,60 – 0,79]	
Nach 48 Wochen	0,89	1,0	0,87	k. A.
[95%-KI]	[0,80 – 0,95]	[NA – NA]	[0,79 – 0,93]	
Nach 72 Wochen	0,93	1,0	k. A.	k. A.
[95%-KI]	[0,85 – 0,97]	[NA – NA]		
Nach 96 Wochen	0,95	1,0	k. A.	k. A.
[95%-KI]	[0,86 – 0,98]	[NA – NA]		
Nach 144 Wochen	1,00	1,0	k. A.	k. A.
[95%-KI]	[NA – NA]	[NA – NA]		
* Schätzung der Zeit bis zum ersten Ansprechen mit Hilfe der Kaplan-Meier-Methode				
Abkürzungen: BAT, Best Available Therapy; ITT, Intention to treat; k. A., keine Angabe; KI, Konfidenzintervall; LSM, Least Square Means; NA, not applicable; RR, Risk Ratio; SD, Standardabweichung				

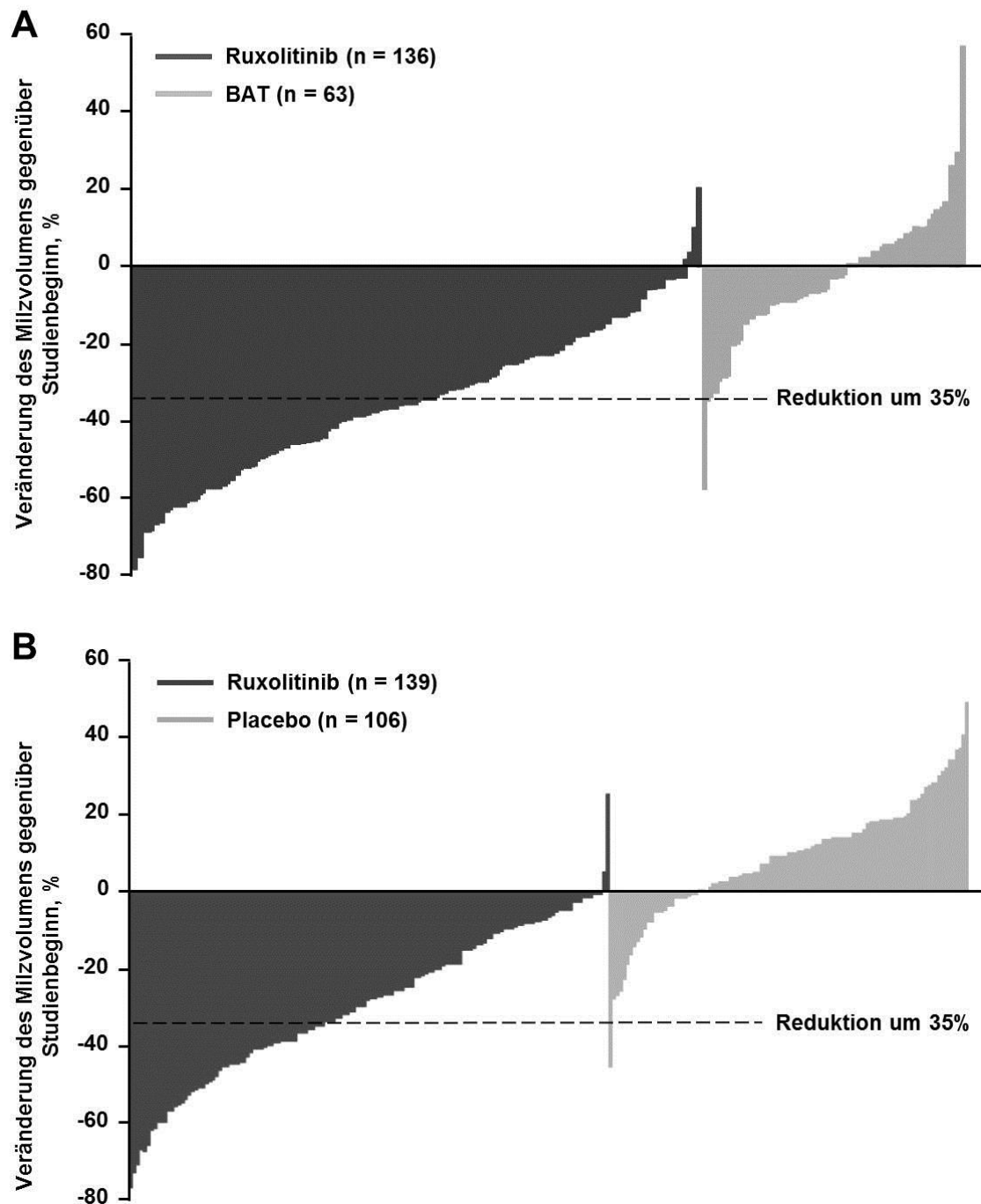


Abbildung 4-5: Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel: Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn

(A) Größte prozentuale Veränderung des Milzvolumens im Vergleich zum Ausgangswert in der Studie COMFORT-II innerhalb der ersten 48 Wochen nach Studienbeginn. Nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn sowie mindestens einer weiteren Messung wurden in die Analyse einbezogen. (B) Größte prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn in der Studie COMFORT-I zu Studienwoche 24. Dargestellt sind nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn sowie zu Studienwoche 24. Jede Linie entspricht einem Patienten. BAT, Best Available Therapy.

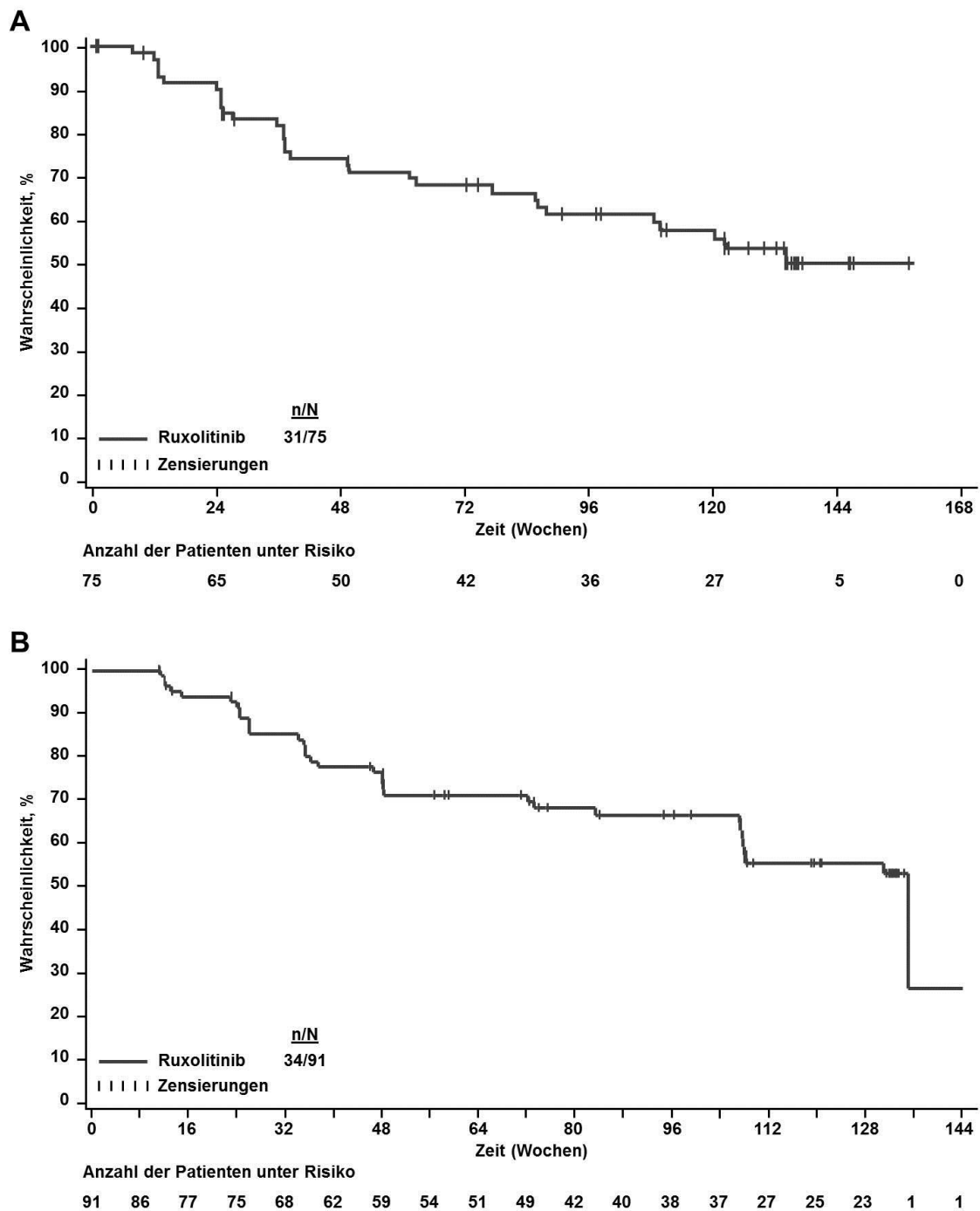


Abbildung 4-6: Ergebnisse für den Endpunkt „Reduktion der Milzgröße“ aus RCT mit dem zu bewertenden Arzneimittel: Dauer des Ansprechens

Kaplan-Meier-Kurven für die Dauer des Ansprechens in Bezug auf die Milzvolumenreduktion unter der Behandlung mit Ruxolitinib in den beiden Studien COMFORT-II (**A**) und COMFORT-I (**B**) nach einem medianen Follow-up von 151 Wochen bzw. 149 Wochen. In die Analyse wurden alle Patienten eingeschlossen, die zu irgendeinem Zeitpunkt im Studienverlauf eine Milzvolumenreduktion von $\geq 35\%$ erreicht hatten. Auf der X-Achse ist die Zeit seit dem ersten Erreichen eines solchen Ansprechens dargestellt. n/N, Zahl der Ereignisse/Zahl der Patienten.

Reduktion des Milzvolumens

Der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen bzw. 48 Wochen war der primäre Endpunkt der beiden Studien COMFORT-I und COMFORT-II (22;23;31;64). Der Rückgang des Milzvolumens im Vergleich zur Baseline wurde mittels Magnetresonanztomographie (MRT) oder Computertomographie (CT) gemessen. Die Auswertung der Aufnahmen erfolgte durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war.

Wie die Primäranalyse der **COMFORT-II**-Studie im Januar 2011 zeigte, erreichten 28,5% der Patienten im Ruxolitinib-Arm eine Milzvolumenreduktion um $\geq 35\%$ nach 48 Wochen, während dies für keinen der Patienten in der Kontrollgruppe der Fall war ($p < 0,0001$) (Tabelle 4-15) (31;64). Die Studie erreichte damit ihren primären Wirksamkeitsendpunkt. Der Vorteil der Ruxolitinib-Therapie war dabei bereits nach 24 Wochen zu erkennen. So hatten zu diesem Zeitpunkt 31,9% der Patienten eine Milzvolumenreduktion um $\geq 35\%$ erreicht im Vergleich zu 0% unter der BAT-Kontrollbehandlung ($p < 0,0001$) (31;64). Auch die Sensitivitätsanalyse, in die auch Patienten ohne auswertbare Milzvolumen-Messung zu Studienbeginn einbezogen wurden, bestätigte die Überlegenheit von Ruxolitinib mit signifikant höheren Ansprechraten nach 24 Wochen und 48 Wochen (Tabelle 4-15) (31). Bei der Erhebung des primären Endpunkts betrug die mediane Milzvolumenreduktion unter der Therapie mit Ruxolitinib 28,3%, während das Milzvolumen im Kontrollarm sogar um 8,5% zunahm ($p < 0,0001$) (Tabelle 4-15) (31). Insgesamt erreichten 97,1% der Patienten im Ruxolitinib-Arm bis zu diesem Zeitpunkt bei mindestens einer Messung eine Milzvolumenreduktion unter den Ausgangswert zu Studienbeginn (Abbildung 4-5A). Unter der BAT-Kontrolltherapie war dies dagegen nur bei 55,6% der Patienten der Fall, wobei der primäre Endpunkt, die Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen, von keinem Patienten erreicht wurde. Lediglich bei einem Patienten war ein entsprechendes Ansprechen zum Zeitpunkt der 12-Wochen-Evaluation zu verzeichnen.

Auch die **COMFORT-I**-Studie erreichte ihren primären Endpunkt und bestätigt damit die Überlegenheit von Ruxolitinib bei der Behandlung der krankheitsbedingten Splenomegalie bei Patienten mit Myelofibrose. So war der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen unter der Therapie mit Ruxolitinib signifikant höher als unter Placebo (41,9% vs. 0,7%; $p < 0,0001$) (Tabelle 4-15) (22;23). Auch die Auswertung der Sensitivitätsanalyse, in der auch Patienten ohne auswertbare Milzvolumen-Messung zu Studienbeginn berücksichtigt wurden, ergab einen klaren Vorteil für die Ruxolitinib-Therapie (Tabelle 4-15) (23). Nach 24 Wochen betrug die mediane Milzvolumenreduktion unter der Therapie mit Ruxolitinib 33%, während das Milzvolumen im Kontrollarm um median 8,5% zunahm ($p < 0,0001$) (Tabelle 4-15) (23). Insgesamt konnte in der

Ruxolitinib-Gruppe nach 24 Wochen bei nahezu allen Patienten ein Rückgang des Milzvolumens nachgewiesen werden (Abbildung 4-5B) (22;23). Lediglich bei zwei Patienten war eine moderate Zunahme des Milzvolumens zu verzeichnen. Beide Patienten wurden jedoch aufgrund von Thrombozytopenien mit einer stark reduzierten Ruxolitinib-Dosis von 5 mg bid behandelt, die deutlich unter den empfohlenen Dosen von 15 mg bid bzw. 20 mg bid liegt (23). Im Gegensatz dazu kam es bei der Mehrzahl der Patienten unter Placebo zu einer deutlichen Zunahme des Milzvolumens. Der einzige Patient, bei dem zu diesem Zeitpunkt ein Rückgang des Milzvolumens um $\geq 35\%$ gemessen wurde, verstarb bereits wenige Tage nach dieser Untersuchung an den Folgen einer Krankheitsprogression. Vermutlich handelte es sich bei der gemessenen Milzvolumenreduktion nicht um ein echtes klinisches Ansprechen, sondern um einen Milzinfarkt, der auch die anschließende Verschlechterung des Gesundheitszustands des Patienten erklären würde (23).

Dauer des Ansprechens

Die Dauer des Ansprechens in Bezug auf die Milzvolumenreduktion wurde in beiden Studien als sekundärer Wirksamkeitsendpunkt untersucht (22;23;31;64). Sie war definiert als die Zeit vom ersten Erreichen einer Milzvolumenreduktion von $\geq 35\%$ bis zum Verlust dieses Ansprechens bei gleichzeitiger Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie (24;62).

Wie das 3-Jahres-Update der **COMFORT-II-Studie** zeigte, blieb eine bereits erreichte Reduktion des Milzvolumens bei den meisten Patienten über diesen Zeitraum erhalten (62;63). Basierend auf Kaplan-Meier-Schätzungen lag die Wahrscheinlichkeit, unter der Behandlung mit Ruxolitinib eine Milzvolumenreduktion um $\geq 35\%$ für mindestens 48 bzw. 144 Wochen beizubehalten, bei 73% (95%-KI: 61-82%) bzw. 50% (95%-KI: 36-63%) (Tabelle 4-15). Die mediane Ansprechdauer wurde bislang noch nicht erreicht. Da unter der BAT-Behandlung nur bei einem einzigen Teilnehmer eine Milzvolumenreduktion von $\geq 35\%$ gemessen wurde, war die Berechnung einer Ansprechdauer im Kontrollarm der Studie nicht möglich (31;62).

Auch in **COMFORT-I** wurde eine unter der Therapie mit Ruxolitinib erreichte Milzvolumenreduktion um $\geq 35\%$ von der Mehrzahl der Patienten dauerhaft beibehalten (24). So lag die Wahrscheinlichkeit eines Ansprechens für 48 Wochen bzw. 96 Wochen bei 76% (95%-KI: 66-84%) bzw. 67% (95%-KI: 55-76%), die mediane Ansprechdauer betrug 135 Wochen (Tabelle 4-15). Die Wahrscheinlichkeit, unter der Behandlung mit Ruxolitinib eine Milzvolumenreduktion um $\geq 35\%$ für mindestens 144 Wochen beizubehalten, lag zwar nur noch bei 27% (95%-KI: 2-63%), allerdings ist die Schätzung aufgrund der zahlreichen Zensurierungen nach Woche 128 wenig zuverlässig (24). Wie in der COMFORT-II-Studie war

auch in COMFORT-I eine Berechnung der Ansprechdauer im Kontrollarm nicht sinnvoll, da nur bei einem einzelnen Patienten eine Reduktion des Milzvolumens von $\geq 35\%$ dokumentiert wurde und dieser bereits wenige Tage nach der Messung verstarb.

Zeit bis zum ersten Ansprechen

In der **COMFORT-II**-Studie lag die mediane Zeit bis zum ersten Ansprechen unter der Therapie mit Ruxolitinib bei 12,4 Wochen und fiel damit zeitlich mit der ersten Milzvolumen-Evaluation nach 12 Wochen zusammen (Tabelle 4-15) (31;62;64). Bei 44 der insgesamt 75 Patienten (58,7%), die zu irgendeinem Zeitpunkt unter Ruxolitinib eine Reduktion des Milzvolumens um $\geq 35\%$ erreicht hatten, wurde ein solches Ansprechen bereits zum Zeitpunkt der 12-Wochen-Untersuchung nachgewiesen. In der BAT-Kontrollgruppe hatte nur ein Patient überhaupt ein solches Ansprechen erreicht und zwar nach 15,4 Wochen.

Auch in **COMFORT-I** hatten nach 12 Wochen bereits 61 der insgesamt 93 Patienten (65,6%) eine Milzvolumenreduktion von $\geq 35\%$ erreicht. Basierend auf einer Kaplan-Meier-Analyse betrug die mediane Zeit bis zum ersten Ansprechen 12,7 Wochen (Tabelle 4-15) (23;24), fiel also wie in COMFORT-II mit dem Zeitpunkt der ersten Milzuntersuchung zusammen.

Zusammenfassung

Die beiden Studien COMFORT-I und COMFORT-II belegen damit eine deutliche Reduktion der krankheitsbedingten Splenomegalie bei Patienten mit Myelofibrose unter der Therapie mit Ruxolitinib. Sowohl nach 24 Wochen (COMFORT-I) als auch nach 48 Wochen (COMFORT-II) war der Anteil der Patienten, die eine klinisch relevante Reduktion des Milzvolumens von mindestens 35% aufwiesen, im Ruxolitinib-Arm der jeweiligen Studie signifikant höher als im Kontrollarm. Im Vergleich zum Ausgangswert konnte in beiden Studien unter Ruxolitinib eine mediane Abnahme des Milzvolumens um circa 30% erreicht werden, wohingegen es in den Komparator-Armen sogar zu einer Zunahme des Milzvolumens kam. Eine Milzvolumenreduktion um $\geq 35\%$ konnte bei der Mehrzahl der Patienten bereits nach 12 Wochen nachgewiesen werden und blieb bei den meisten Patienten über einen Zeitraum von drei Jahren erhalten. Der Nutzen von Ruxolitinib zeigte sich dabei in allen vorab definierten Subgruppen, unabhängig vom Alter der Patienten, ihrem Geschlecht, dem Myelofibrose-Subtyp, der Risikogruppe oder dem Milzvolumen zu Studienbeginn (siehe auch Abschnitt 4.3.1.3.2).

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

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 Linderung der Krankheitssymptomatik – 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 4-16: Operationalisierung von „Linderung der Krankheitssymptomatik“

Studie	Operationalisierung
COMFORT-I	Der Endpunkt „Linderung der Krankheitssymptomatik“ wurde durch zwei Operationalisierungen erfasst: <u>1. Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen</u> Die Myelofibrose-assoziierten Symptome wurden mit Hilfe des MFSAF (Myelofibrosis Symptom Assessment Form)-Fragebogens erfasst und der Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) um $\geq 50\%$ im Vergleich zu Studienbeginn angegeben. Patienten ohne TSS-Bestimmung nach 24 Wochen wurden dabei als Non-Responder klassifiziert. Nur Patienten mit TSS-Bestimmung zu Studienbeginn wurden in die Analyse einbezogen („ITT-Population, evaluierbare Patienten“). Im Rahmen einer Sensitivitätsanalyse wurde der Anteil der Patienten mit einer Verbesserung des TSS zusätzlich unter Einschluss aller Patienten kalkuliert. Patienten ohne TSS-Bestimmung zu Studienbeginn wurden dabei ebenfalls als Non-Responder gewertet („ITT-Population“). <u>2. Prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen</u> Die Myelofibrose-assoziierten Symptome wurden mit Hilfe des MFSAF-Fragebogens erfasst und die prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome nach 24 Wochen im Vergleich zu Studienbeginn angegeben. Der hier verwendete Fragebogen MFSAF v2.0 besteht aus 7 Items, über die die charakteristischen Symptome der Myelofibrose abgefragt werden: - Bauchbeschwerden - Schmerzen linksseitig unter den Rippen - Völlegefühl - Nachtschweiß - Juckreiz - Muskel- und Knochenschmerzen - Inaktivität* Diese Symptome sowie die daraus resultierende Einschränkung der Aktivität werden mit 0 (nicht vorhanden) bis 10 (schlimmstmöglich) bewertet. Aus dem numerischen Score der Einzelsymptome wird durch Addition ein Gesamtscore (Total Symptom Score, TSS) berechnet, der aus 6 Einzelsymptomen (ohne das Item „Inaktivität“) besteht und daher einen Wert zwischen 0 und 60 annehmen kann. Die Erfassung der Symptome durch den MFSAF-Fragebogen erfolgte mit Hilfe eines elektronischen Patiententagebuchs und wurde täglich bis zu Woche 24 durchgeführt. Die Patienten begannen mit dem Ausfüllen des Patiententagebuchs bereits 7 Tage vor der Einnahme der ersten Studienmedikation. Der Basiswert zu Studienbeginn wurde als Mittelwert aus den TSS-Werten dieser 7 Tage bestimmt. Dementsprechend wurde der TSS nach 24 Wochen als Mittelwert aus den TSS-Werten der letzten 28 Tage vor der 24 Wochen-Untersuchung berechnet. * Das Item „Inaktivität“ wurde in Absprache mit der FDA nicht als Symptom gewertet, da die Inaktivität <i>per se</i> kein Symptom darstellt sondern Folge der eigentlichen Symptome ist. Es ist deshalb nicht Bestandteil des Total Symptom Score. Abkürzungen: FDA, Food and Drug Administration; ITT, Intention to treat; MFSAF, Myelofibrosis Symptom Assessment Form; TSS, Total Symptom Score

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 den Endpunkt „Linderung der Krankheitssymptomatik“ 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
Operationalisierung: Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen						
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen						
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Linderung der Krankheitssymptomatik“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Die Linderung der Krankheitssymptomatik wurde nur in der COMFORT-I-Studie, nicht aber in der COMFORT-II-Studie als Endpunkt erhoben. Die folgende Darstellung beschränkt sich daher auf die Daten der **COMFORT-I**-Studie.

Tabelle 4-18: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel

Linderung der Krankheitssymptomatik	COMFORT-I	
	Ruxolitinib (N= 155)	Placebo (N = 154)
Anteil der Patienten mit einer Verbesserung des TSS um ≥ 50% nach 24 Wochen: ITT-Population, evaluierbare Patienten		
N (gesamt)	148	152
Ansprechen (n, %)	68 (45,9)	8 (5,3)
Differenz vs. Placebo [95%-KI]	40,7 [31,2 – 50,1]	
RR vs. Placebo [95%-KI]	8,7 [4,3 – 17,5]	
OR vs. Placebo [95%-KI] [†]	15,28 [6,93 – 33,66]	
p-Wert [*]	< 0,0001	
Anteil der Patienten mit einer Verbesserung des TSS um ≥ 50% nach 24 Wochen: ITT-Population (Sensitivitätsanalyse)		
N (gesamt)	155	154
Ansprechen (n, %)	68 (43,9)	8 (5,2)
Differenz vs. Placebo [95%-KI]	38,7 [29,5 – 47,9]	
RR vs. Placebo [95%-KI]	8,4 [4,2 – 17,0]	
OR vs. Placebo [95%-KI] [†]	14,90 [6,78 – 32,74]	
p-Wert [*]	< 0,0001	
[†] Berechnung des Odds Ratios mittels logistischer Regression		
[*] Berechnung des p-Werts mittels Chi ² -Test		
Prozentuale/absolute Veränderung des TSS gegenüber Studienbeginn nach 24 Wochen		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	18,0 ± 10,9	16,5 ± 11,5
Median	17,8	14,8
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-8,6 ± 10,0	3,2 ± 9,4
Median	-6,9	2,0
Differenz vs. Placebo [95%-KI] [†]	-11,3 [-13,6; -8,95]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	129	103
Mittelwert (± SD)	-46,1 ± 48,6	41,8 ± 99,3
Median	-56,2	14,6
Differenz vs. Placebo [95%-KI] [†]	-85,6 [-105; -66,4]	
p-Wert [*]	< 0,0001	

Linderung der Krankheitssymptomatik	COMFORT-I	
	Ruxolitinib (N= 155)	Placebo (N = 154)
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Bauchbeschwerden“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	3,7 ± 2,2	3,4 ± 2,4
Median	3,7	3,3
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,8 ± 2,2	0,6 ± 2,1
Median	-1,6	0,3
Differenz vs. Placebo [95%-KI] [†]	-2,20 [-2,17; -1,69]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	124	100
Mittelwert (± SD)	-29,4 ± 125,8	55,7 ± 181,0
Median	-55,6	7,7
Differenz vs. Placebo [95%-KI] [†]	-78,4 [-117; -39,6]	
p-Wert [*]	< 0,0001	
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Schmerzen linksseitig unter den Rippen“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	2,7 ± 2,3	2,5 ± 2,5
Median	2,3	1,6
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,4 ± 2,1	0,4 ± 2,3
Median	-1,0	0,0
Differenz vs. Placebo [95%-KI] [†]	-1,73 [-2,23; -1,23]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	109	86
Mittelwert (± SD)	-46,9 ± 111,5	62,8 ± 224,6
Median	-75,0	-7,7
Differenz vs. Placebo [95%-KI] [†]	-107 [-155; -59,5]	
p-Wert [*]	< 0,0001	

Linderung der Krankheitssymptomatik	COMFORT-I	
	Ruxolitinib (N= 155)	Placebo (N = 154)
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Völlegefühl“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	3,7 ± 2,2	3,3 ± 2,2
Median	4,0	3,0
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,8 ± 2,2	0,7 ± 2,3
Median	-1,7	0,6
Differenz vs. Placebo [95%-KI] [†]	-2,36 [-2,89; -1,83]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	122	96
Mittelwert (± SD)	-43,0 ± 70,9	70,3 ± 309,2
Median	-62,0	18,9
Differenz vs. Placebo [95%-KI] [†]	-106 [-162; -50,2]	
p-Wert [*]	< 0,0001	
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Nachtschweiß“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	2,7 ± 2,5	2,4 ± 2,1
Median	2,4	2,1
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,3 ± 2,2	0,4 ± 1,9
Median	-1,0	0,4
Differenz vs. Placebo [95%-KI] [†]	-1,59 [-2,05; -1,13]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	103	90
Mittelwert (± SD)	-42,1 ± 95,5	51,0 ± 139,0
Median	-75,4	14,9
Differenz vs. Placebo [95%-KI] [†]	-83,9 [-117; -50,8]	
p-Wert [*]	< 0,0001	

Linderung der Krankheitssymptomatik	COMFORT-I	
	Ruxolitinib (N= 155)	Placebo (N = 154)
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Juckreiz“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	2,2 ± 2,1	2,3 ± 2,5
Median	1,7	1,4
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,3 ± 2,3	0,4 ± 2,4
Median	-1,0	0,0
Differenz vs. Placebo [95%-KI] [†]	-1,74 [-2,23; -1,26]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	98	77
Mittelwert (± SD)	-42,8 ± 111,0	111,9 ± 573,5
Median	-85,3	-3,7
Differenz vs. Placebo [95%-KI] [†]	-158 [-273; -43,1]	
p-Wert [*]	< 0,0001	
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Muskel- und Knochenschmerzen“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	3,0 ± 2,5	2,7 ± 2,4
Median	2,6	2,1
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,0 ± 2,1	0,7 ± 1,9
Median	-0,6	0,5
Differenz vs. Placebo [95%-KI] [†]	-1,60 [-2,09; -1,12]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	108	89
Mittelwert (± SD)	-21,2 ± 125,1	37,7 ± 107,0
Median	-46,6	12,3
Differenz vs. Placebo [95%-KI] [†]	-54,1 [-86,8; -21,5]	
p-Wert [*]	< 0,0001	

Linderung der Krankheitssymptomatik	COMFORT-I	
	Ruxolitinib (N= 155)	Placebo (N = 154)
Prozentuale/absolute Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen: „Inaktivität“		
Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	3,8 ± 2,6	3,7 ± 2,4
Median	3,7	3,7
Absolute Veränderung gegenüber Studienbeginn		
N (gesamt)	131	105
Mittelwert (± SD)	-1,2 ± 1,9	0,6 ± 2,3
Median	-0,8	0,6
Differenz vs. Placebo [95%-KI] [†]	-1,78 [-2,28; -1,27]	
p-Wert [*]	< 0,0001	
Prozentuale Veränderung gegenüber Studienbeginn		
N (gesamt) [‡]	119	97
Mittelwert (± SD)	-32,1 ± 69,4	60,0 ± 238,7
Median	-37,1	15,2
Differenz vs. Placebo [95%-KI] [†]	-88,9 [-133; -44,6]	
p-Wert [*]	< 0,0001	
[†] Berechnung der LSM-Differenz mit zugehörigem 95%-Konfidenzintervall auf Basis einer Kovarianzanalyse mit dem Baseline-Wert als Kovariate		
[*] Berechnung des p-Werts mittels Wilcoxon-Rangsummentest		
[‡] Patienten mit 0 als Ausgangswert wurden nicht in die Analyse eingeschlossen; eine gesonderte Beschreibung der Ausgangswerte für dieses Patientenkollektiv findet sich im Studienbericht der COMFORT-I-Studie (23)		
Abkürzungen: ITT, Intention to treat; KI, Konfidenzintervall; LSM, Least Square Means; MFSAF, Myelofibrosis Symptom Assessment Form; OR, Odds Ratio; RR, Risk Ratio; SD, Standardabweichung; TSS, Total Symptom Score		

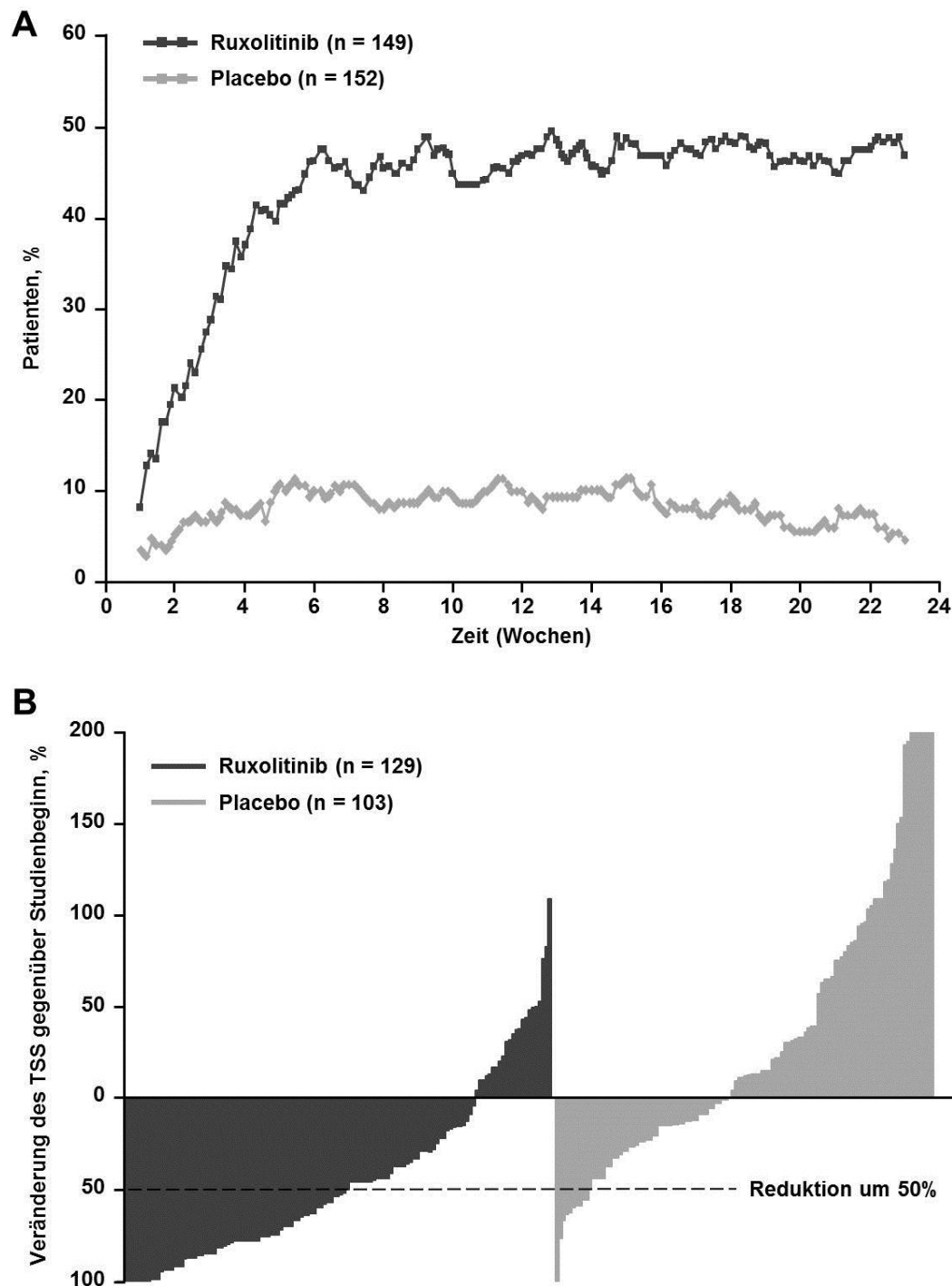


Abbildung 4-7: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel: Veränderung des TSS gegenüber Studienbeginn

(A) Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ gegenüber Studienbeginn im Zeitverlauf bis Studienwoche 24 in der Studie COMFORT-I. Der Gesamtscore für den jeweiligen Tag wurde als Mittelwert aus den Werten der vorangegangenen 7 Tage berechnet. In die Analyse wurden nur Patienten mit TSS-Bestimmung zu Studienbeginn einbezogen. **(B)** Größte prozentuale Veränderung des TSS gegenüber Studienbeginn nach 24 Wochen in der Studie COMFORT-I. Dargestellt sind nur Patienten mit TSS-Bestimmung zu Studienbeginn sowie zu Studienwoche 24. Eine Verschlechterung des TSS um $\geq 200\%$ ist in der Abbildung als 200% dargestellt. MFSAF, Myelofibrosis Symptom Assessment Form; TSS, Total Symptom Score.

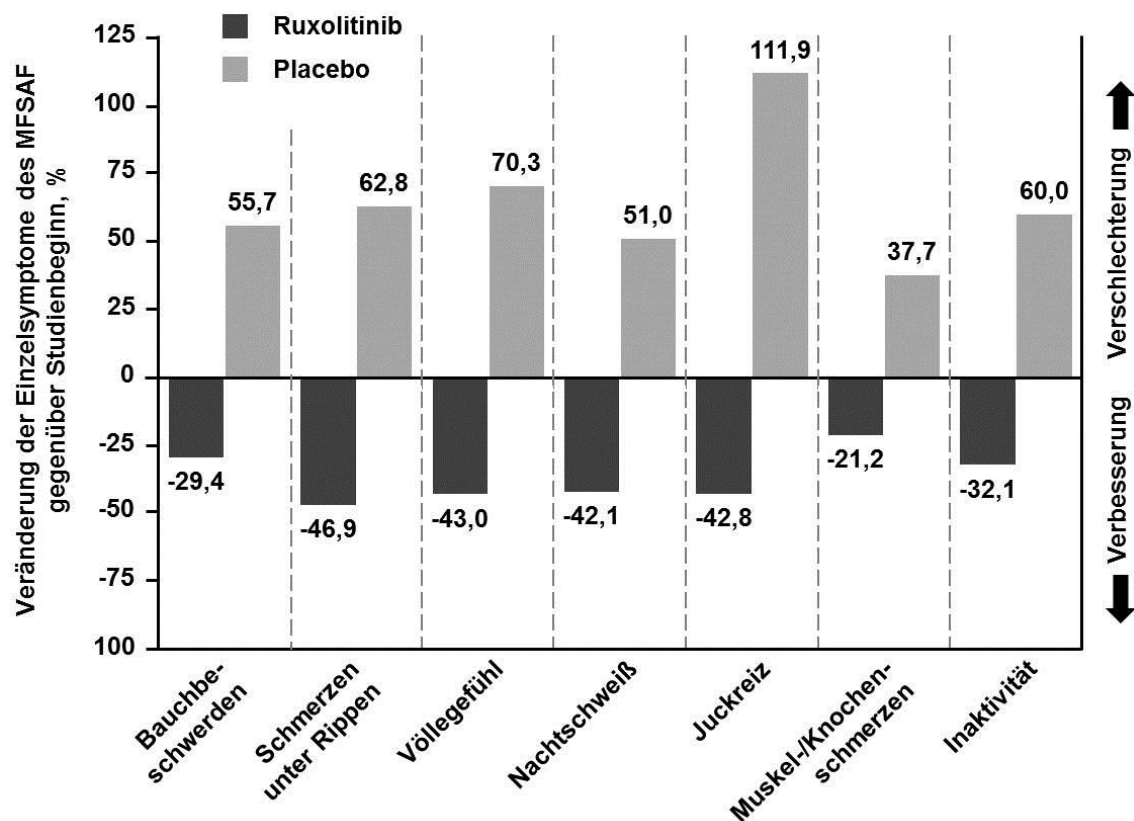


Abbildung 4-8: Ergebnisse für den Endpunkt „Linderung der Krankheitssymptomatik“ aus RCT mit dem zu bewertenden Arzneimittel: Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn

Prozentuale Veränderung der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen. In die Analyse wurden nur Patienten mit MFSAF-Bewertung zu Studienbeginn sowie zu Studienwoche 24 eingeschlossen, Patienten mit 0 als Ausgangswert wurden ebenfalls nicht berücksichtigt. MFSAF, Myelofibrosis Symptom Assessment Form.

Die Myelofibrose-assoziierten Symptome bzw. die Linderung der Krankheitssymptomatik unter der Therapie wurde in der **COMFORT-I**-Studie mit Hilfe des modifizierten MFSAF v2.0 (Myelofibrosis Symptom Assessment Form)-Fragebogens erfasst (22;23). Dieser besteht aus sieben Items, über die die folgenden, für die Erkrankung charakteristischen Symptome abgefragt werden: Bauchbeschwerden, Schmerzen linksseitig unter den Rippen, Völlegefühl, Nachtschweiß, Juckreiz, Muskel- und Knochenschmerzen und Inaktivität. Diese Symptome sowie die daraus resultierende Einschränkung der Aktivität für den Patienten werden mit 0 (nicht vorhanden) bis 10 (schlimmstmöglich) bewertet. Aus dem numerischen Score der Einzelsymptome wird durch Addition ein Gesamtscore (Total Symptom Score, TSS) berechnet, der sich aus sechs der sieben krankheitsspezifischen Symptome des MFSAF zusammensetzt. Das Einzel-Item „Inaktivität“ wurde auf Vorschlag der FDA nicht als

Symptom gewertet, da die Inaktivität eine Folge der eigentlichen Symptome ist (30). Die Erfassung der Symptome durch den MFSAF-Fragebogen erfolgte täglich mit Hilfe eines elektronischen Patiententagebuchs (22;23).

Wie die Primäranalyse der COMFORT-I-Studie im November 2010 zeigte, war der Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen unter der Therapie mit Ruxolitinib signifikant höher als unter Placebo (45,9% vs. 5,3%; $p < 0,0001$) (Tabelle 4-18) (22;23). Auch die Auswertung der Sensitivitätsanalyse, in der auch Patienten ohne TSS-Bestimmung zu Studienbeginn berücksichtigt wurden, ergab einen klaren Vorteil für die Ruxolitinib-Therapie (43,9% vs. 5,2%; $p < 0,0001$) (Tabelle 4-18) (23). Die meisten Patienten, die nach 24 Wochen eine Verbesserung des TSS um $\geq 50\%$ zeigten, erreichten ein solches Ansprechen bereits innerhalb der ersten vier Wochen nach Therapiebeginn (Abbildung 4-7A) (22;23). Nach 24 Wochen betrug die mediane Verbesserung des TSS unter der Therapie mit Ruxolitinib 56,2%, während es im Kontrollarm zu einer medianen Verschlechterung um 14,6% kam ($p < 0,0001$) (Tabelle 4-18) (22;23). Insgesamt konnte im Ruxolitinib-Arm nach 24 Wochen bei der Mehrzahl der Patienten eine deutliche Verbesserung des TSS nachgewiesen werden, einige Patienten waren sogar vollständig symptomfrei. Unter der Placebo-Behandlung kam es hingegen beim überwiegenden Teil der Patienten zu einer mitunter dramatischen Verschlechterung der Symptomatik (Abbildung 4-7B) (22;23). Die Verbesserung der Myelofibrose-Symptome unter der Therapie mit Ruxolitinib war dabei nicht nur in der Reduktion des Gesamtscores, sondern auch in Bezug auf alle sechs untersuchten Einzelsymptome sowie beim ebenfalls abgefragten Item „Inaktivität“ ersichtlich (Abbildung 4-8) (22;23). So zeigten die Ergebnisse einer Post-hoc-Analyse, dass es in der Ruxolitinib-Gruppe bei jedem der Einzelsymptome zu einer Verbesserung kam, während unter der Placebo-Kontrolltherapie jeweils eine Verschlechterung zu verzeichnen war ($p < 0,0001$ für alle Vergleiche) (Tabelle 4-18) (23). Zusammenfassend belegen die Daten der COMFORT-I-Studie damit eine ebenso schnelle wie deutliche Reduktion der schwerwiegenden, mit der Erkrankung assoziierten Symptome unter der Behandlung mit Ruxolitinib.

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.

Da die Linderung der Krankheitssymptomatik nur in einer der beiden in die Nutzenbewertung eingeschlossenen Studien untersucht wurde (COMFORT-I), war eine Zusammenfassung der Ergebnisse mit Hilfe einer Meta-Analyse nicht möglich.

4.3.1.3.1.3 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 4-19: Operationalisierung von „Gesamtüberleben“

Studie	Operationalisierung
COMFORT-II	Das Gesamtüberleben (overall survival, OS) war definiert als die Zeit von der Randomisierung bis zum Tod des Patienten ungeachtet der zugrundeliegenden Todesursache. Falls kein Ereignis auftrat oder bei fehlender Follow-up-Information erfolgte eine Zensierung zum letzten Zeitpunkt, an dem der Patient nachweislich am Leben war (Zeitpunkt des letzten Kontakts). Daten zum Gesamtüberleben wurden alle 3 Monate erhoben.
COMFORT-I	Das Gesamtüberleben war definiert als die Zeit von der Randomisierung bis zum Tod des Patienten ungeachtet der zugrundeliegenden Todesursache. Falls kein Ereignis auftrat oder bei fehlender Follow-up-Information erfolgte eine Zensierung zum letzten Zeitpunkt, an dem der Patient nachweislich am Leben war (Zeitpunkt des letzten Kontakts). Daten zum Gesamtüberleben wurden alle 6 Monate erhoben.
Abkürzungen: OS, overall survival	

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 den Endpunkt „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
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesamtüberleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im

Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da der Endpunkt „Gesamtüberleben“ unabhängig von subjektiven Einflussfaktoren ist. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesamtüberleben“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Tabelle 4-21: Ergebnisse für den Endpunkt „Gesamtüberleben“ aus RCT mit dem zu bewertenden Arzneimittel

Gesamtüberleben	COMFORT-II		COMFORT-I	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Datenschnitt: Primäranalyse [†]				
Gesamtüberleben (Wochen)				
Zahl der Ereignisse (n, %)	6 (4,1)	4 (5,5)	10 (6,5)	14 (9,1)
Zahl der Zensierungen (n, %)	140 (95,9)	69 (94,5)	145 (93,5)	140 (90,9)
Medianes Gesamtüberleben [95%-KI]	NA	NA	NA	NA
Hazard Ratio	0,70		0,67	
[95%-KI]	[0,20 – 2,49]		[0,30 – 1,50]	
p-Wert	0,58*		0,327**	
Datenschnitt: 3-Jahres-Update [‡]				
Gesamtüberleben (Wochen)				
Zahl der Ereignisse (n, %)	29 (19,9)	22 (30,1)	42 (27,1)	54 (35,1)
Zahl der Zensierungen (n, %)	117 (80,1)	51 (69,9)	113 (72,9)	100 (64,9)
Medianes Gesamtüberleben [95%-KI]	NA	NA	NA	NA
Hazard Ratio	0,48		0,69	
[95%-KI]	[0,28 – 0,85]		[0,46 – 1,03]	
p-Wert	0,009*		0,067**	
Kaplan-Meier-Schätzer				
Nach 24 Wochen	k. A.	k. A.	0,96	0,97
[95%-KI]			[0,92 – 0,98]	[0,92 – 0,99]
Nach 48 Wochen	0,96	0,94	0,92	0,84
[95%-KI]	[0,91 – 0,98]	[0,84 – 0,98]	[0,87 – 0,95]	[0,77 – 0,89]
Nach 96 Wochen	0,88	0,84	0,81	0,71
[95%-KI]	[0,81 – 0,93]	[0,72 – 0,92]	[0,73 – 0,86]	[0,63 – 0,78]
Nach 144 Wochen	0,81	0,61	0,74	0,61
[95%-KI]	[0,74 – 0,87]	[0,47 – 0,73]	[0,66 – 0,81]	[0,53 – 0,69]
[†] nach einem medianen Follow-up von 51 Wochen (COMFORT-II) bzw. 32 Wochen (COMFORT-I)				
[‡] nach einem medianen Follow-up von 151 Wochen (COMFORT-II) bzw. 149 Wochen (COMFORT-I)				
* Berechnung des p-Werts mit Hilfe eines Logrank-Tests stratifiziert nach Risikogruppe				
** Berechnung des p-Werts mittels Logrank-Test				
Abkürzungen: BAT, Best Available Therapy; k. A., keine Angabe; KI, Konfidenzintervall; NA, not applicable				

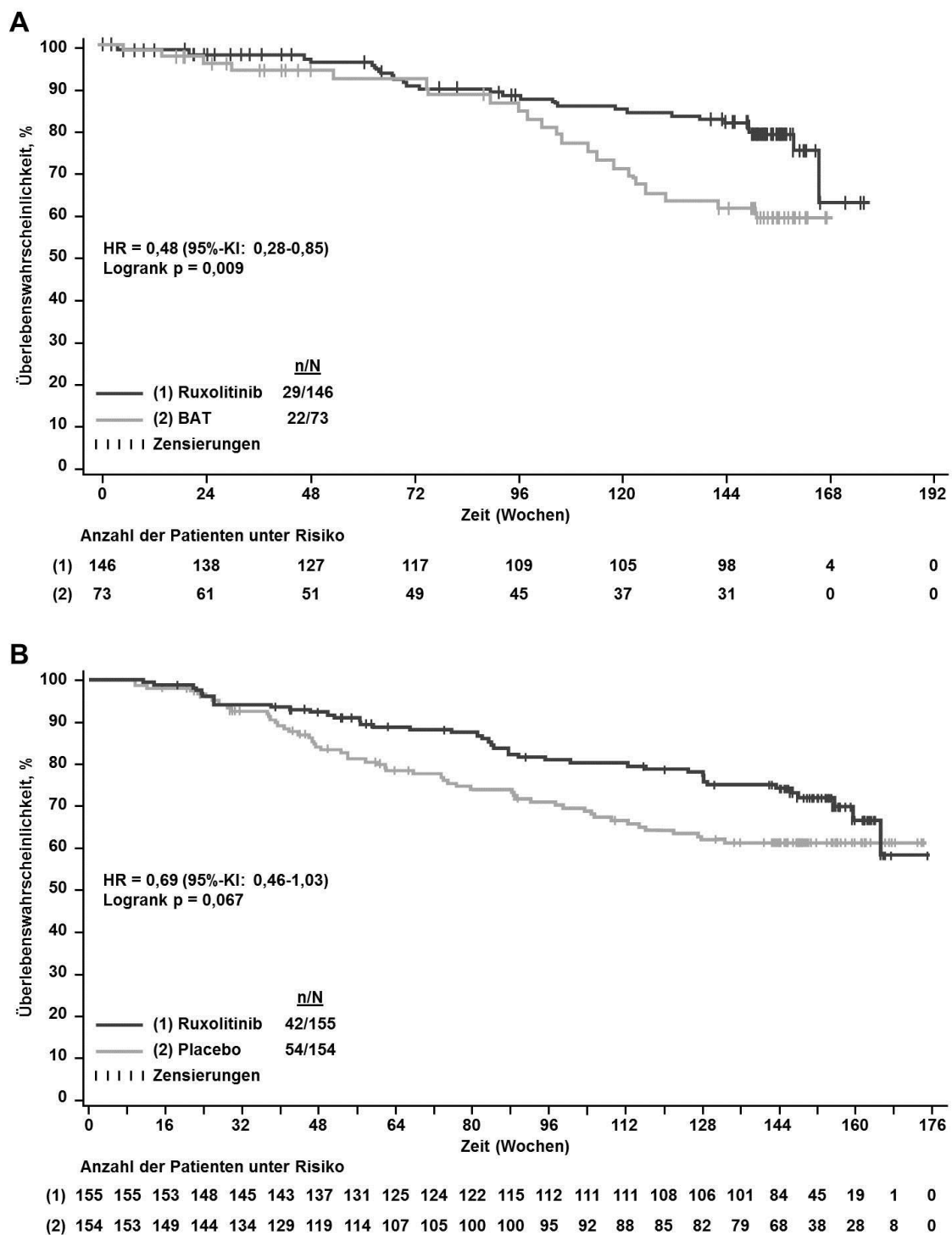


Abbildung 4-9: Ergebnisse für den Endpunkt „Gesamtüberleben“ aus RCT mit dem zu bewertenden Arzneimittel

(A) Kaplan-Meier-Kurven für das Gesamtüberleben in der Studie COMFORT-II unter der Behandlung mit Ruxolitinib bzw. BAT nach einem medianen Follow-up von 151 Wochen. Auswertung gemäß ITT. (B) Kaplan-Meier-Kurven für das Gesamtüberleben in der Studie COMFORT-I unter der Behandlung mit Ruxolitinib bzw. Placebo nach einem medianen Follow-up von 149 Wochen. Auswertung gemäß ITT. BAT, Best Available Therapy; ITT, Intention to treat; n/N, Zahl der Ereignisse/Zahl der Patienten.

Das Studiendesign der Studien COMFORT-I und COMFORT-II war nicht primär auf den langfristigen Vergleich von Überlebensparametern zwischen dem Ruxolitinib-Arm und dem Kontrollarm der jeweiligen Studie ausgerichtet (22;23;31;64). So war es den Patienten sowohl in COMFORT-I als auch in COMFORT-II aus ethischen Gründen erlaubt, bei Vorliegen bestimmter Progressionskriterien auf die Behandlung mit Ruxolitinib umzustellen. Aufgrund der deutlichen Überlegenheit von Ruxolitinib konnten außerdem in beiden Studien alle Patienten nach der primären Datenanalyse zur Ruxolitinib-Therapie wechseln (23;31). Bei der Interpretation der Ergebnisse ist jedoch zu beachten, dass die Auswertung internationalen Standards entsprechend nach dem Intention-to-treat-Prinzip erfolgte (23;24;31;62). Da hierbei auch die Effekte der Ruxolitinib-Therapie in den Vergleichsarm der Studie eingehen, zeigen Patienten unter der Vergleichstherapie voraussichtlich ein besseres Gesamtüberleben als ohne den Therapiewechsel zu erwarten gewesen wäre (69). Somit ist sowohl in COMFORT-I als auch in COMFORT-II von einer Unterschätzung des mortalitätssenkenden Effekts von Ruxolitinib auszugehen.

Da die mediane Nachbeobachtungszeit bei der Primäranalyse der **COMFORT-II**-Studie lediglich 51 Wochen betrug, war die Zahl der Ereignisse zu diesem Zeitpunkt noch sehr gering. Insgesamt verstarben nur 10 Patienten, 6 Patienten (4,1%) im Ruxolitinib-Arm und 4 Patienten (5,5%) im BAT-Arm der Studie (Tabelle 4-21) (31;64). Damit war nach dieser kurzen Behandlungszeit noch kein unmittelbarer Überlebensvorteil für einen der beiden Studienarme ersichtlich (HR = 0,70 [0,20-2,49]; p = 0,58). Die Daten der 3-Jahres-Analyse, die zum ursprünglich geplanten Studienende nach 151 Wochen medianer Nachbeobachtungszeit durchgeführt wurde, zeigten jedoch eine deutliche Senkung des Mortalitätsrisikos unter der Therapie mit Ruxolitinib. Während zu diesem Zeitpunkt in der BAT-Kontrollgruppe bereits 30,1% der Patienten verstorben waren, lag der Anteil der Todesfälle in der Ruxolitinib-Gruppe bei nur 19,9% (Tabelle 4-21) (62;63). Dies entspricht einer signifikanten Reduktion der Mortalitätsrate unter der Therapie mit Ruxolitinib um 52% (HR = 0,48 [0,28-0,85]; p = 0,009) (Abbildung 4-9A). Dementsprechend betrug die Überlebenswahrscheinlichkeit nach 144 Wochen im Ruxolitinib-Arm 81%, wohingegen sie unter der BAT-Behandlung nur bei 61% lag. Das mediane Gesamtüberleben wurde bislang in keinem der beiden Studienarme erreicht (Tabelle 4-21) (62;63).

Da in der COMFORT-II-Studie vor Amendment 3 des Studienprotokolls im Februar 2010 bei vorzeitigem Abbruch der Behandlung keine weitere Erhebung von Überlebensdaten geplant war, lag für insgesamt 28 Patienten, die die Studie bis zu diesem Zeitpunkt in beiden Gruppen beendet hatten, keine Follow-up-Information in Bezug auf das Gesamtüberleben vor (31;70). Für die neueste Auswertung zum 3,5-Jahres-Update, die beim diesjährigen Kongress der European Hematology Association (EHA) im Juni 2014 präsentiert werden soll, lagen

zusätzlich Informationen zum Überleben für weitere 15 Patienten vor (5 Patienten aus der Ruxolitinib-Gruppe, 10 Patienten aus der BAT-Gruppe) (70). Bei Einschluss dieser Patienten in die Analyse zeigte sich ebenfalls eine statistisch signifikante Senkung des Mortalitätsrisikos unter der Therapie mit Ruxolitinib (HR = 0,58 [0,36-0,93]; $p = 0,02$). Insgesamt waren zum Zeitpunkt der Auswertung bereits 41,1% der Patienten im BAT-Arm verstorben, während der Anteil der Todesfälle in der Ruxolitinib-Gruppe bei nur 27,4% lag (70).

Auch die Ergebnisse der **COMFORT-I**-Studie bestätigen den Überlebensvorteil für Patienten unter der Therapie mit Ruxolitinib. Zwar war auch hier die Gesamtzahl der beobachteten Todesfälle zum Zeitpunkt der Primäranalyse nach 32 Wochen medianem Follow-up zu gering für einen statistisch validen Vergleich zwischen den Studienarmen, bei der Auswertung der 3-Jahres-Analyse zeichnete sich jedoch ein deutlicher Trend bezüglich eines verlängerten Gesamtüberlebens unter der Therapie mit Ruxolitinib ab (Tabelle 4-21) (22-24). So konnte das Mortalitätsrisiko gegenüber Placebo um 31% gesenkt werden (HR = 0,69 [0,46-1,03]; $p = 0,067$). Die Wirksamkeitsunterschiede zwischen den beiden Therapien waren dabei bereits 32 Wochen nach Therapiebeginn zu beobachten und zeigten sich in einer frühen und lang anhaltenden Separierung der Kaplan-Meier-Kurven (Abbildung 4-9B) (24;71). Lediglich nach Woche 144 näherten sich die Kurven – möglicherweise bedingt durch die zahlreichen Zensurierungen zu diesem Zeitpunkt – wieder aneinander an.

Die beiden Studien COMFORT-I und COMFORT-II zeigen damit eine deutliche Reduktion des Mortalitätsrisikos unter der Therapie mit Ruxolitinib bei Patienten mit primärer oder sekundärer Myelofibrose und bestätigen so die Ergebnisse früherer Analysen, die bereits im Jahr 2012 auf dem Kongress der American Society of Hematology (ASH) präsentiert wurden (72;73). Dieses Ergebnis ist insbesondere deshalb bemerkenswert, da der mortalitätssenkende Therapieeffekt von Ruxolitinib aufgrund des in beiden Studien erlaubten Cross-overs eher unterschätzt wird und der tatsächliche Überlebensvorteil durch Ruxolitinib somit noch höher sein dürfte.

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

4.3.1.3.1.4 Leukämiefreies Überleben – 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 4-22: Operationalisierung von „Leukämiefreies Überleben“

Studie	Operationalisierung
COMFORT-II	Das leukämiefreie Überleben (leukemia-free survival, LFS) war definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Auftreten eines der folgenden Ereignisse: - Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ - Tod durch jedwede Ursache Falls kein Ereignis auftrat erfolgte eine Zensierung zum Zeitpunkt der letzten Untersuchung. Daten zur leukämischen Transformation sowie zum Gesamtüberleben wurden alle 3 Monate erhoben.
COMFORT-I	Das leukämiefreie Überleben war definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Auftreten eines der folgenden Ereignisse: - Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ - Tod durch jedwede Ursache Falls kein Ereignis auftrat erfolgte eine Zensierung zum Zeitpunkt der letzten Untersuchung. Daten zur leukämischen Transformation sowie zum Gesamtüberleben wurden alle 6 Monate erhoben.
Abkürzungen: LFS, leukemia-free survival	

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 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt „Leukämiefreies Überleben“ 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
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Leukämiefreies Überleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Auftreten einer leukämischen Transformation durch klar definierte Laborparameter erfasst wurde und somit wie die Erhebung des Gesamtüberlebens unabhängig von subjektiven Einflussfaktoren ist. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Leukämiefreies Überleben“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Tabelle 4-24: Ergebnisse für den Endpunkt „Leukämiefreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel

Leukämiefreies Überleben	COMFORT-II		COMFORT-I	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Datenschnitt: Primäranalyse [†]				
LFS (Wochen)				
Zahl der Ereignisse (n, %)	6 (4,1)	4 (5,5)	11 (7,1)	14 (9,1)
Zahl der Zensierungen (n, %)	140 (95,9)	69 (94,5)	144 (92,9)	140 (90,9)
Medianes LFS [95%-KI]	NA	NA	NA	NA
Hazard Ratio	0,65		k. A.	
[95%-KI]	[0,18 – 2,31]			
p-Wert	0,507*		0,447**	
Datenschnitt: 3-Jahres-Update [‡]				
LFS (Wochen)				
Zahl der Ereignisse (n, %)	32 (21,9)	23 (31,5)	45 (29,0)	58 (37,7)
Zahl der Zensierungen (n, %)	114 (78,1)	50 (68,5)	110 (71,0)	96 (62,3)
Medianes LFS [95%-KI]	NA	NA	NA	NA
Hazard Ratio	0,51		0,69	
[95%-KI]	[0,30 – 0,88]		[0,46 – 1,01]	
p-Wert	0,014*		0,057**	
Kaplan-Meier-Schätzer				
Nach 24 Wochen	0,98	0,97	k. A.	k. A.
[95%-KI]	[0,94 – 0,99]	[0,89 – 0,99]		
Nach 48 Wochen	0,95	0,94	0,90	0,84
[95%-KI]	[0,90 – 0,98]	[0,84 – 0,98]	[0,84 – 0,94]	[0,77 – 0,89]
Nach 96 Wochen	0,87	0,82	0,80	0,70
[95%-KI]	[0,80 – 0,92]	[0,70 – 0,90]	[0,73 – 0,86]	[0,62 – 0,77]
Nach 144 Wochen	0,78	0,59	0,73	0,58
[95%-KI]	[0,70 – 0,85]	[0,45 – 0,71]	[0,65 – 0,79]	[0,50 – 0,66]
[†] nach einem medianen Follow-up von 51 Wochen (COMFORT-II) bzw. 32 Wochen (COMFORT-I)				
[‡] nach einem medianen Follow-up von 151 Wochen (COMFORT-II) bzw. 149 Wochen (COMFORT-I)				
* Berechnung des p-Werts mit Hilfe eines Logrank-Tests stratifiziert nach Risikogruppe				
** Berechnung des p-Werts mittels Logrank-Test				
Abkürzungen: BAT, Best Available Therapy; k. A., keine Angabe; KI, Konfidenzintervall; LFS, leukemia-free survival; NA, not applicable				

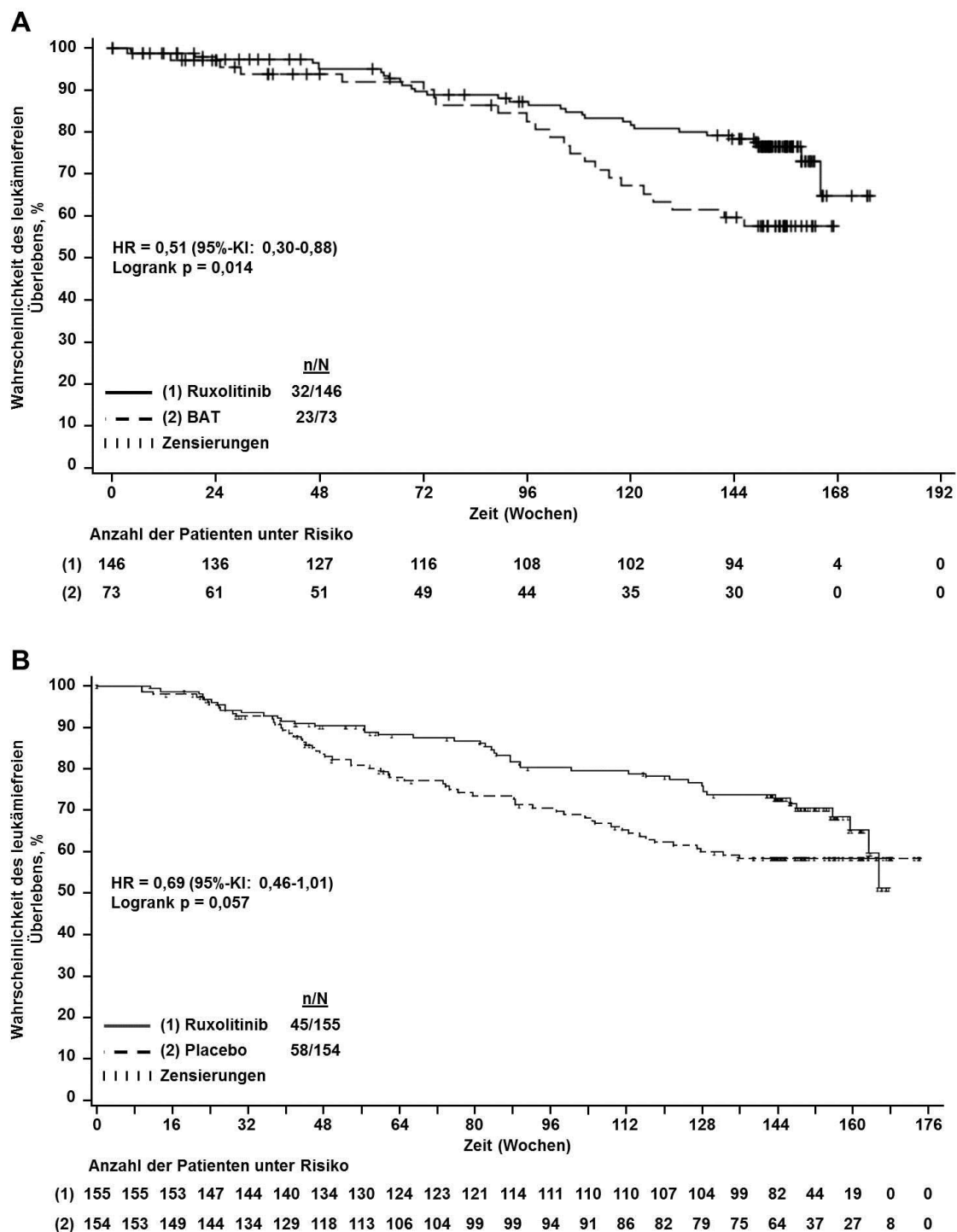


Abbildung 4-10: Ergebnisse für den Endpunkt „Leukämiefreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel

(A) Kaplan-Meier-Kurven für das leukämiefreie Überleben in der Studie COMFORT-II unter der Behandlung mit Ruxolitinib bzw. BAT nach einem medianen Follow-up von 151 Wochen. Auswertung gemäß ITT. (B) Kaplan-Meier-Kurven für das leukämiefreie Überleben in der Studie COMFORT-I unter der Behandlung mit Ruxolitinib bzw. Placebo nach einem medianen Follow-up von 149 Wochen. Auswertung gemäß ITT. BAT, Best Available Therapy; ITT, Intention to treat; n/N, Zahl der Ereignisse/Zahl der Patienten.

In beiden Studien COMFORT-I und COMFORT-II wurde das leukämiefreie Überleben als Endpunkt untersucht (23;31). Dieses war definiert als der Zeitraum von der Randomisierung bis zum Auftreten einer leukämischen Transformation (definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$) oder bis zum Tod ungeachtet der zugrundeliegenden Todesursache.

Die mediane Nachbeobachtungszeit bei der Primäranalyse der **COMFORT-II**-Studie betrug lediglich 51 Wochen. Die Zahl der Ereignisse in diesem Zeitraum war dementsprechend gering. Insgesamt wurden nur 10 Ereignisse dokumentiert, nämlich 6 (4,1%) im Ruxolitinib-Arm und 4 (5,5%) im BAT-Arm der Studie (HR = 0,65 [0,18-2,31]; p = 0,507) (Tabelle 4-24) (31;64). Bei allen Ereignissen handelte es sich um Todesfälle, der Übergang in eine akute myeloische Leukämie wurde bei keinem der Patienten diagnostiziert (31). Zum Zeitpunkt der 3-Jahres-Analyse wurde dagegen bereits bei 31,5% der Patienten im BAT-Arm der Studie ein Ereignis erfasst, während dies nur bei 21,9% der Patienten unter der Therapie mit Ruxolitinib der Fall war (Tabelle 4-24) (62). Die Zahl der Patienten mit einer leukämischen Transformation als Ereignis betrug dabei 5 (3,4%) in der Ruxolitinib-Gruppe bzw. 4 (5,5%) in der BAT-Gruppe (62;63). Somit ergibt sich unter der Therapie mit Ruxolitinib eine signifikante Reduktion des Transformations- bzw. Mortalitätsrisikos um 49% (HR = 0,51 [0,30-0,88]; p = 0,014) (Abbildung 4-10A). Die Wahrscheinlichkeit des leukämiefreien Überlebens nach 144 Wochen betrug 78% im Ruxolitinib-Arm, wohingegen sie unter der BAT-Behandlung nur bei 59% lag (62).

Auch die Ergebnisse der **COMFORT-I**-Studie bestätigen die Überlegenheit der Ruxolitinib-Therapie in Bezug auf das leukämiefreie Überleben. Zwar war auch hier die Zahl der Ereignisse zum Zeitpunkt der Primäranalyse nach 32 Wochen medianem Follow-up zu gering für einen statistisch validen Vergleich zwischen den Studienarmen, bei der Auswertung der 3-Jahres-Analyse zeigte sich jedoch ein deutlicher Trend hinsichtlich eines verlängerten leukämiefreien Überlebens unter der Therapie mit Ruxolitinib (Tabelle 4-24) (23;24). Zu diesem Zeitpunkt wurde bei 37,7% der Patienten im Placebo-Arm der Studie ein Ereignis dokumentiert, während nur 29,0% der Patienten im Ruxolitinib-Arm betroffen waren (24). Dabei wurde bei jeweils 7 Patienten pro Gruppe eine akute myeloische Leukämie diagnostiziert. Das Risiko, eine leukämische Transformation zu erleiden oder zu versterben, konnte unter der Therapie mit Ruxolitinib somit um 31% gesenkt werden (HR = 0,69 [0,46-1,01]; p = 0,057) (Abbildung 4-10B). Dementsprechend betrug die Wahrscheinlichkeit des leukämiefreien Überlebens nach 144 Wochen 73% in der Ruxolitinib-Gruppe, wohingegen sie unter der Placebo-Behandlung nur bei 58% lag (24).

Die beiden Studien COMFORT-I und COMFORT-II belegen damit eine deutliche Verlängerung des leukämiefreien Überlebens unter der Therapie mit Ruxolitinib. Da die Mehrzahl der Ereignisse Todesfälle darstellten, ist die Überlegenheit gegenüber der Vergleichsintervention vornehmlich auf den Überlebensvorteil unter der Ruxolitinib-Behandlung zurückzuführen (siehe Abschnitt 4.3.1.3.1.3).

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

4.3.1.3.1.5 Progressionsfreies Überleben – 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 4-25: Operationalisierung von „Progressionsfreies Überleben“

Studie	Operationalisierung
COMFORT-II	Das progressionsfreie Überleben (progression-free survival, PFS) war definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Auftreten eines der folgenden Ereignisse: - Zunahme des Milzvolumens um > 25% im Vergleich zum niedrigsten Wert während der Studie - Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ - Milzbestrahlung, Splenektomie oder Tod Falls kein Ereignis auftrat erfolgte eine Zensierung zum Zeitpunkt der letzten Untersuchung. Daten zur leukämischen Transformation sowie zum Gesamtüberleben wurden alle 3 Monate erhoben.
Abkürzungen: PFS, progression-free survival	

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 den Endpunkt „Progressionsfreies Überleben“ 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
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Progressionsfreies Überleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten

und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Auftreten einer Krankheitsprogression durch klar definierte Laborparameter bzw. die Auswertung von MRT- bzw. CT-Aufnahmen erfasst wurde und somit wie die Erhebung des Gesamtüberlebens unabhängig von subjektiven Einflussfaktoren ist. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Das progressionsfreie Überleben wurde nur in der COMFORT-II-Studie, nicht aber in der COMFORT-I-Studie als Endpunkt erhoben. Die folgende Darstellung beschränkt sich daher auf die Daten der **COMFORT-II**-Studie.

Tabelle 4-27: Ergebnisse für den Endpunkt „Progressionsfreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel

Progressionsfreies Überleben	COMFORT-II	
	Ruxolitinib (N= 146)	BAT (N = 73)
Datenschnitt: Primäranalyse [†]		
PFS (Wochen)		
Zahl der Ereignisse (n, %)	44 (30,1)	19 (26,0)
Zahl der Zensierungen (n, %)	102 (69,9)	54 (74,0)
Medianes PFS [95%-KI]	60,4 [60,1 – NA]	60,1 [49,4 – NA]
Hazard Ratio	0,81	
[95%-KI]	[0,47 – 1,39]	
p-Wert	0,457*	
Datenschnitt: 3-Jahres-Update [‡]		
PFS (Wochen)		
Zahl der Ereignisse (n, %)	83 (56,8)	35 (47,9)
Zahl der Zensierungen (n, %)	63 (43,2)	38 (52,1)
Medianes PFS [95%-KI]	84,1 [64,1 – 120,1]	72,3 [59,3 – 98,0]
Hazard Ratio	0,75	
[95%-KI]	[0,50 – 1,12]	
p-Wert	0,158*	
Kaplan-Meier-Schätzer		
Nach 24 Wochen	0,97	0,87
[95%-KI]	[0,92 – 0,99]	[0,76 – 0,93]
Nach 48 Wochen	0,75	0,73
[95%-KI]	[0,66 – 0,81]	[0,59 – 0,83]
Nach 96 Wochen	0,47	0,38
[95%-KI]	[0,38 – 0,56]	[0,24 – 0,53]
Nach 144 Wochen	0,33	0,24
[95%-KI]	[0,25 – 0,42]	[0,12 – 0,38]
[†] nach einem medianen Follow-up von 51 Wochen		
[‡] nach einem medianen Follow-up von 151 Wochen		
* Berechnung des p-Werts mit Hilfe eines Logrank-Tests stratifiziert nach Risikogruppe		
Abkürzungen: BAT, Best Available Therapy; KI, Konfidenzintervall; NA, not applicable; PFS, progression-free survival		

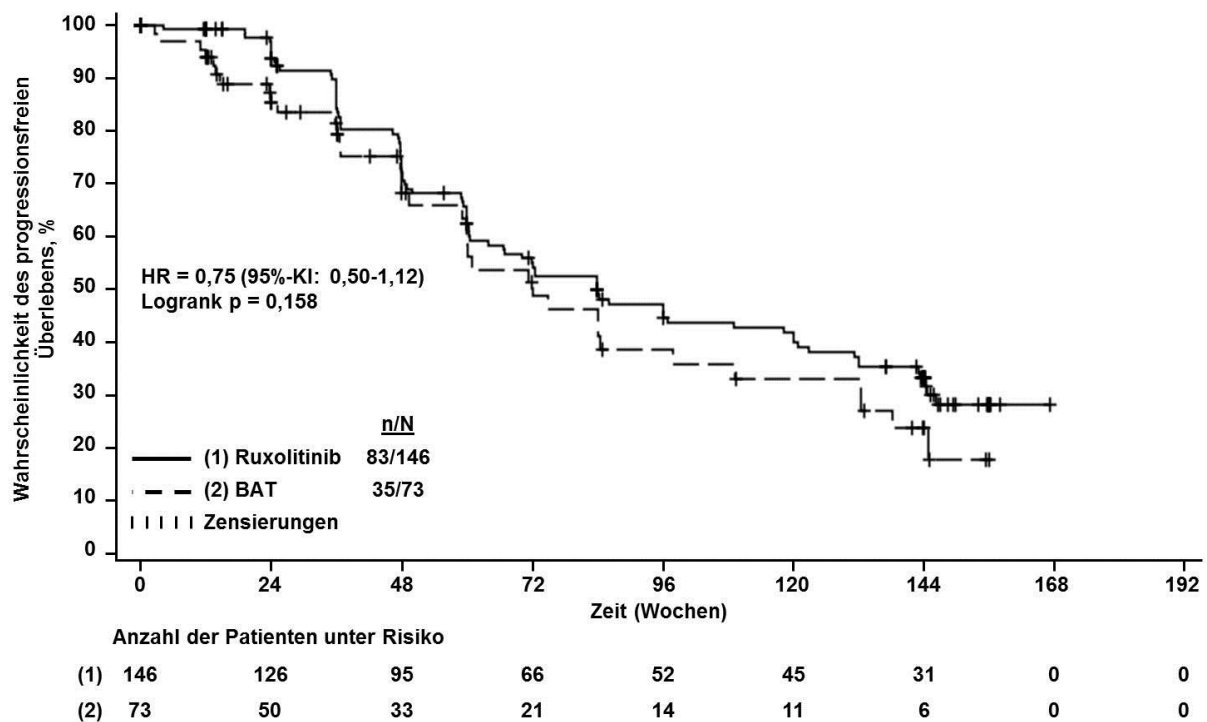


Abbildung 4-11: Ergebnisse für den Endpunkt „Progressionsfreies Überleben“ aus RCT mit dem zu bewertenden Arzneimittel

Kaplan-Meier-Kurven für das progressionsfreie Überleben in der Studie COMFORT-II unter der Behandlung mit Ruxolitinib bzw. BAT nach einem medianen Follow-up von 151 Wochen. Auswertung gemäß ITT. BAT, Best Available Therapy; ITT, Intention to treat; n/N, Zahl der Ereignisse/Zahl der Patienten.

In der **COMFORT-II**-Studie wurde das progressionsfreie Überleben als sekundärer Endpunkt erhoben (31;64). Dieses war definiert als der Zeitraum von der Randomisierung bis zum Auftreten eines der folgenden Ereignisse: der Zunahme des Milzvolumens um > 25% im Vergleich zum niedrigsten Wert während der Studie, dem Nachweis einer leukämischen Transformation (definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$), der Durchführung einer Milzbestrahlung bzw. Splenektomie oder dem Tod ungeachtet der zugrundeliegenden Todesursache (31).

Bei der Primäranalyse im Januar 2011 war bezüglich des progressionsfreien Überlebens kein Unterschied zwischen den beiden Behandlungsgruppen zu beobachten. Die Zahl der Patienten mit einem Progressionsereignis betrug 44 (30,1%) unter der Therapie mit Ruxolitinib im Vergleich zu 19 (26,0%) unter der Kontrolltherapie (Tabelle 4-27) (31). Auch zum Zeitpunkt der 3-Jahres-Analyse war die Zahl der Patienten mit einer Krankheitsprogression in den beiden Studienarmen vergleichbar. So wurden 83 (56,8%) Progressionsereignisse in der

Ruxolitinib-Gruppe und 35 (47,9%) Ereignisse in der BAT-Gruppe dokumentiert (Tabelle 4-27) (62). Der Anteil der Patienten mit einer Zunahme des Milzvolumens als Ereignis war dabei mit 50,7% in der Ruxolitinib-Gruppe etwas höher als in der BAT-Gruppe (35,6%). Dabei ist jedoch zu beachten, dass 20 Patienten im Ruxolitinib-Arm zwar eine Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie aufwiesen, bei ihnen aber immer noch eine Reduktion des Milzvolumens um $\geq 35\%$ gegenüber Studienbeginn vorlag, ein Wert, der in der Kontrollgruppe während der gesamten Studiendauer nur von einem einzigen Patienten erreicht wurde (62). Nach 144 Wochen betrug der geschätzte Anteil an Patienten ohne Krankheitsprogression 33% unter der Therapie mit Ruxolitinib im Vergleich zu 24% unter BAT (HR = 0,75 [0,50-1,12]; p = 0,158) (Abbildung 4-11) (62).

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.

Da das progressionsfreie Überleben nur in einer der beiden in die Nutzenbewertung eingeschlossenen Studien untersucht wurde (COMFORT-II), war eine Zusammenfassung der Ergebnisse mit Hilfe einer Meta-Analyse nicht möglich.

4.3.1.3.1.6 Veränderung des ECOG-Performance-Status – 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 4-28: Operationalisierung von „Veränderung des ECOG-Performance-Status“

Studie	Operationalisierung														
COMFORT-II	Der Endpunkt „Veränderung des ECOG-Performance-Status“ wurde durch zwei Operationalisierungen erfasst: <u>Veränderung des ECOG-Performance-Status gegenüber Studienbeginn</u> Deskriptive Darstellung des ECOG-Performance-Status zu allen Erhebungszeitpunkten sowie dessen Veränderung gegenüber Studienbeginn. <u>Zeit bis zur Verschlechterung des ECOG-Performance-Status</u> Die Zeit bis zur Verschlechterung des ECOG-Performance-Status war definiert als die Zeit von der Randomisierung bis zur ersten dokumentierten Zunahme des ECOG-Performance-Status von 0 oder 1 auf ≥ 2 bzw. bis zur ersten Zunahme des ECOG-Performance-Status von 2 auf ≥ 3. Falls kein Ereignis auftrat, erfolgte eine Zensurierung zum Zeitpunkt der letzten Untersuchung. Bei Patienten, die in den Ruxolitinib-Arm wechselten bzw. eine weitere antineoplastische Therapie erhielten, erfolgte eine Zensurierung zum Zeitpunkt der letzten Untersuchung vor Beginn der neuen Therapie. Der ECOG-Performance-Status beschreibt den körperlichen Zustand von Tumorpatienten mit Hilfe einer Skala von 0 bis 5. Die Bedeutung der einzelnen Stufen ist in der folgenden Tabelle dargestellt: <table border="1"> <thead> <tr> <th>Stufe</th><th>Beschreibung</th></tr> </thead> <tbody> <tr> <td>0</td><td>normale, uneingeschränkte Aktivität wie vor der Erkrankung</td></tr> <tr> <td>1</td><td>Einschränkung bei körperlicher Anstrengung, gehfähig, leichte körperliche Arbeit oder Arbeit im Sitzen möglich</td></tr> <tr> <td>2</td><td>gehfähig, Selbstversorgung möglich, aber nicht arbeitsfähig; kann mehr als 50% der Wachzeit aufstehen</td></tr> <tr> <td>3</td><td>nur begrenzte Selbstversorgung möglich; 50% oder mehr der Wachzeit an Bett oder Stuhl gebunden</td></tr> <tr> <td>4</td><td>völlig pflegebedürftig, keinerlei Selbstversorgung möglich; völlig an Bett oder Stuhl gebunden</td></tr> <tr> <td>5</td><td>Tod</td></tr> </tbody> </table> Der ECOG-Performance-Status wurde zu Studienbeginn, nach 4, 8, 16, 24, 48 und 72 Wochen sowie danach alle 24 Wochen erhoben.	Stufe	Beschreibung	0	normale, uneingeschränkte Aktivität wie vor der Erkrankung	1	Einschränkung bei körperlicher Anstrengung, gehfähig, leichte körperliche Arbeit oder Arbeit im Sitzen möglich	2	gehfähig, Selbstversorgung möglich, aber nicht arbeitsfähig; kann mehr als 50% der Wachzeit aufstehen	3	nur begrenzte Selbstversorgung möglich; 50% oder mehr der Wachzeit an Bett oder Stuhl gebunden	4	völlig pflegebedürftig, keinerlei Selbstversorgung möglich; völlig an Bett oder Stuhl gebunden	5	Tod
Stufe	Beschreibung														
0	normale, uneingeschränkte Aktivität wie vor der Erkrankung														
1	Einschränkung bei körperlicher Anstrengung, gehfähig, leichte körperliche Arbeit oder Arbeit im Sitzen möglich														
2	gehfähig, Selbstversorgung möglich, aber nicht arbeitsfähig; kann mehr als 50% der Wachzeit aufstehen														
3	nur begrenzte Selbstversorgung möglich; 50% oder mehr der Wachzeit an Bett oder Stuhl gebunden														
4	völlig pflegebedürftig, keinerlei Selbstversorgung möglich; völlig an Bett oder Stuhl gebunden														
5	Tod														
COMFORT-I	Der Endpunkt „Veränderung des ECOG-Performance-Status“ wurde durch eine Operationalisierung erfasst: <u>Veränderung des ECOG-Performance-Status gegenüber Studienbeginn</u> Deskriptive Darstellung des ECOG-Performance-Status zu allen Erhebungszeitpunkten sowie dessen Veränderung gegenüber Studienbeginn. Der ECOG-Performance-Status wurde zu Studienbeginn, nach 4, 8, 12, 16, 24, 48 und 72 Wochen sowie danach alle 24 Wochen erhoben.														
Abkürzungen: ECOG, Eastern Cooperative Oncology Group															

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 4-29: Bewertung des Verzerrungspotenzials für den Endpunkt „Veränderung des ECOG-Performance-Status“ 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
Operationalisierung: Veränderung des ECOG-Performance-Status gegenüber Studienbeginn						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Zeit bis zur Verschlechterung des ECOG-Performance-Status						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Veränderung des ECOG-Performance-Status“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, bei der Analyse der entsprechenden Zielgrößen kann jedoch von einer unverzerrten Bewertung ausgegangen werden, da die Erhebung des ECOG-Status anhand objektiver Kriterien (z. B. „gehfähig“, „Selbstversorgung möglich“) durch den Prüfarzt erfolgte. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Veränderung des ECOG-Performance-Status“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren,

die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

Stellen Sie die Ergebnisse für den Endpunkt „Veränderung des ECOG-Performance-Status“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-30: Ergebnisse für den Endpunkt „Veränderung des ECOG-Performance-Status“ aus RCT mit dem zu bewertenden Arzneimittel

Veränderung des ECOG-Performance-Status	COMFORT-II		COMFORT-I [†]	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 151)
ECOG-Performance-Status*				
Zu Studienbeginn (n, %)				
0	58 (39,7)	26 (35,6)	47 (30,3)	38 (25,2)
1	77 (52,7)	37 (50,7)	87 (56,1)	80 (53,0)
2	10 (6,8)	9 (12,3)	14 (9,0)	25 (16,6)
3	1 (0,7)	1 (1,4)	3 (1,9)	4 (2,6)
fehlt	0 (0,0)	0 (0,0)	4 (2,6)	4 (2,6)
Nach 4 Wochen (n, %)				
0	71 (48,6)	25 (34,2)	55 (35,5)	28 (18,5)
1	57 (39,0)	29 (39,7)	79 (51,0)	87 (57,6)
2	5 (3,4)	8 (11,0)	12 (7,7)	27 (17,9)
3	2 (1,4)	1 (1,4)	2 (1,3)	1 (0,7)
fehlt	11 (7,5)	10 (13,7)	7 (4,5)	8 (5,3)
Nach 8 Wochen (n, %)				
0	72 (49,3)	25 (34,2)	65 (41,9)	31 (20,5)
1	51 (34,9)	27 (37,0)	71 (45,8)	75 (49,7)
2	6 (4,1)	6 (8,2)	12 (7,7)	31 (20,5)
3	0 (0,0)	0 (0,0)	1 (0,6)	4 (2,6)
fehlt	17 (11,6)	14 (19,2)	6 (3,9)	10 (6,6)
Nach 12 Wochen (n, %)				
0			58 (37,4)	28 (18,5)
1	-	-	74 (47,7)	76 (50,3)
2			12 (7,7)	26 (17,2)
3			1 (0,6)	6 (4,0)
fehlt			10 (6,5)	15 (9,9)

Veränderung des ECOG- Performance-Status	COMFORT-II		COMFORT-I [†]	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 151)
Nach 16 Wochen (n, %)				
0	65 (44,5)	16 (21,9)	60 (38,7)	22 (14,6)
1	53 (36,3)	26 (35,6)	69 (44,5)	70 (46,4)
2	6 (4,1)	5 (6,8)	12 (7,7)	28 (18,5)
3	0 (0,0)	1 (1,4)	2 (1,3)	3 (2,0)
fehlt	21 (14,4)	25 (34,2)	12 (7,7)	28 (18,5)
Nach 24 Wochen (n, %)				
0	61 (41,8)	13 (17,8)	59 (38,1)	24 (15,9)
1	41 (28,1)	22 (30,1)	73 (47,1)	59 (39,1)
2	2 (1,4)	3 (4,1)	6 (3,9)	22 (14,6)
3	0 (0,0)	0 (0,0)	2 (1,3)	4 (2,6)
fehlt	42 (28,8)	35 (47,9)	15 (9,7)	41 (27,2)
Nach 48 Wochen ** (n, %)				
0	52 (35,6)	19 (26,0)		
1	27 (18,5)	6 (8,2)	-	-
2	3 (2,1)	4 (5,5)		
3	1 (0,7)	1 (1,4)		
fehlt	63 (43,2)	43 (58,9)		
Zeit bis zur Verschlechterung des ECOG-Performance-Status				
Zahl der Ereignisse (n, %)	13 (8,9)	16 (21,9)	-	
Zahl der Zensierungen (n, %)	133 (91,1)	57 (78,1)		
Mediane Zeit bis zur Verschlechterung (Wochen)	NA	55,7		
[95%-KI]		[49,6 – 55,7]		
Hazard Ratio		0,32		
[95%-KI]		[0,16 – 0,67]	-	
Kaplan-Meier-Schätzer				
Nach 24 Wochen	0,93	0,89		
[95%-KI]	[0,87 – 0,96]	[0,78 – 0,95]		
Nach 48 Wochen	0,91	0,74		
[95%-KI]	[0,84 – 0,95]	[0,59 – 0,85]		
[†] Die Zahl der Patienten mit fehlender ECOG-Bestimmung sowie die Prozentangaben in Bezug auf die Gesamtpopulation sind im Studienbericht nicht angegeben (23). Die entsprechenden Werte stellen Ergebnisse eigener Berechnungen dar, die in Absprache mit der zuständigen Statistikabteilung durchgeführt wurden.				
* Nur ECOG-Scores ≤ 3 sind angegeben.				
** In der COMFORT-I-Studie wurde der ECOG-Performance-Status regelmäßig bis zu Woche 144 bestimmt (24). Da nach der Primäranalyse alle Patienten innerhalb kurzer Zeit aus dem Placebo-Arm in den Ruxolitinib-Arm wechselten und Patienten unter Placebo bei der Auswertung nur bis zum Cross-over in den Ruxolitinib-Arm berücksichtigt wurden, ist die Analyse nach Woche 24 für einen Vergleich zwischen den beiden Studienarmen wenig aussagekräftig. Auf die Darstellung der erhobenen Daten wurde deshalb verzichtet.				
Abkürzungen: BAT, Best Available Therapy; ECOG, Eastern Cooperative Oncology Group; KI, Konfidenzintervall; NA, not applicable				

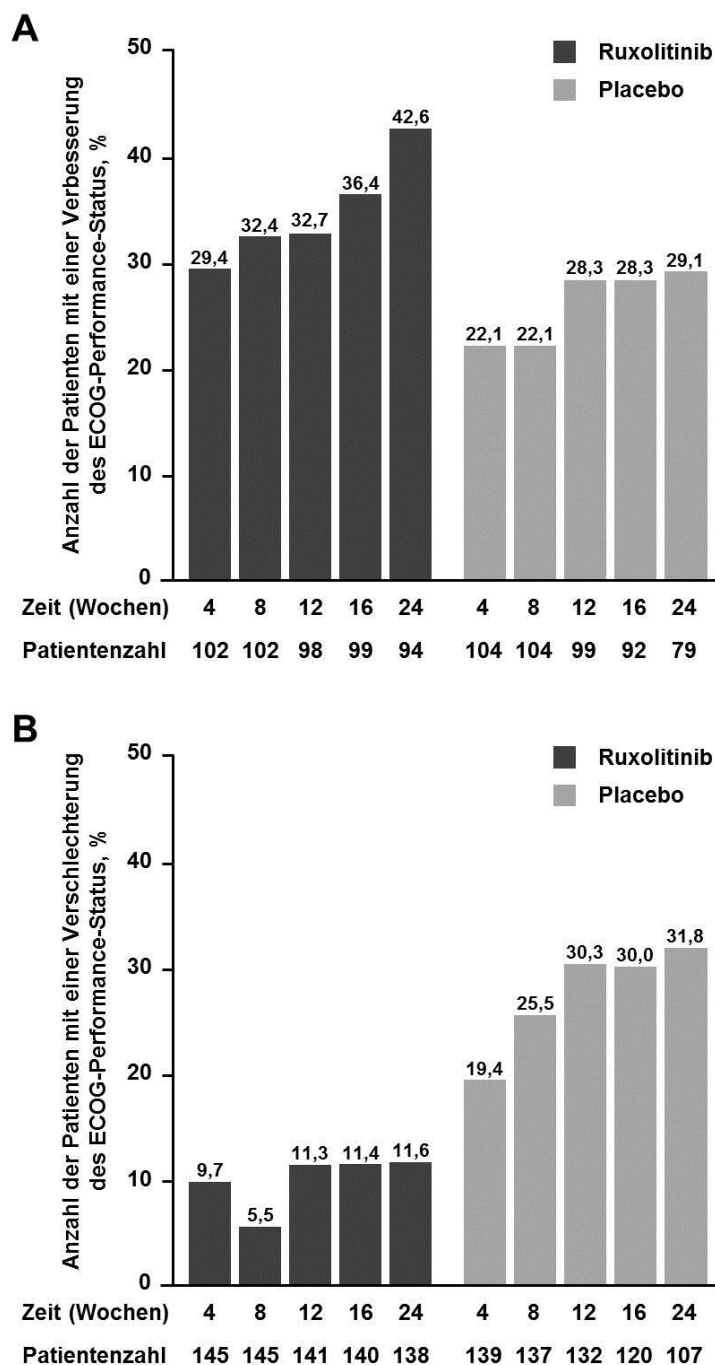


Abbildung 4-12: Ergebnisse für den Endpunkt „Veränderung des ECOG-Performance-Status“ aus RCT mit dem zu bewertenden Arzneimittel

(A) Anteil der Patienten mit einer Verbesserung des ECOG-Status gegenüber Studienbeginn unter Ruxolitinib bzw. Placebo in COMFORT-I. Die Patientenzahl umfasst alle Patienten mit einem ECOG-Status von > 0 zu Studienbeginn, für die zum jeweiligen Zeitpunkt eine ECOG-Bestimmung vorlag. **(B)** Anteil der Patienten mit einer Verschlechterung des ECOG-Status gegenüber Studienbeginn unter Ruxolitinib bzw. Placebo in der Studie COMFORT-I. Die Patientenzahl umfasst alle Patienten mit einem ECOG-Status von < 5 zu Studienbeginn, für die zum jeweiligen Zeitpunkt eine ECOG-Bestimmung vorlag. Achtung: Das Diagramm in Abbildung 4-11B wurde korrigiert, da es im Studienbericht fehlerhaft dargestellt ist und nicht mit den Angaben in der zugehörigen Tabelle 14.2.18.1 übereinstimmt. Eine Verbesserung bzw. Verschlechterung des ECOG-Status war definiert als eine Abnahme bzw. Zunahme des Baseline-Scores um ≥ 1 . ECOG, Eastern Cooperative Oncology Group.

Um den Gesundheitszustand der Patienten im Verlauf der Behandlung zu charakterisieren, wurde in beiden Studien die Veränderung des ECOG (Eastern Cooperative Oncology Group)-Performance-Status analysiert, der den physischen Zustand von Krebspatienten und deren Einschränkung bei Aktivitäten des alltäglichen Lebens mit Hilfe einer Skala von 0 (keine Einschränkung) bis 5 (Tod) beschreibt (28).

In der Studie **COMFORT-II** war der ECOG-Performance-Status zu Studienbeginn zwischen den beiden Gruppen vergleichbar (Tabelle 4-30) (31). Mehr als 35% der Patienten in den beiden Studienarmen hatten einen ECOG-Performance-Status von 0, die Hälfte hatte einen ECOG-Status von 1. Zu den späteren Analyse-Zeitpunkten stieg jedoch der Anteil der Patienten mit einem ECOG-Status von 0 bzw. 1 in der Ruxolitinib-Gruppe im Vergleich zu dem in der BAT-Gruppe deutlich an. So hatten nach 48 Wochen 35,6% bzw. 18,5% der Patienten im Ruxolitinib-Arm einen ECOG-Status von 0 bzw. 1, während dies nur bei 26,0% bzw. 8,2% der Patienten im BAT-Arm der Fall war (Tabelle 4-30) (31). Von denjenigen Patienten, die zu Studienbeginn einen ECOG-Status von 0 hatten, verschlechterten sich insgesamt 60% der Patienten unter der BAT-Therapie aber nur 34,5% der Patienten unter der Behandlung mit Ruxolitinib (31). Umgekehrt konnten mehr Patienten unter der Ruxolitinib-Therapie als unter der BAT-Kontrollbehandlung den besten Wert von 0 im Verlauf der Studie halten (31). Dementsprechend war auch das Risiko, eine Verschlechterung des ECOG-Performance-Status zu erleiden, in der Ruxolitinib-Gruppe geringer als in der BAT-Gruppe (HR = 0,32 [0,16-0,67]) (Tabelle 4-30) (31).

Auch in der Studie **COMFORT-I** war der Anteil der Patienten, die zum Zeitpunkt der Primäranalyse einen ECOG-Performance-Status von 0 oder 1 hatten, im Ruxolitinib-Arm der Studie höher als im Placebo-Arm (Tabelle 4-30) (23). Da der Performance-Status zu Studienbeginn zwischen den beiden Studienarmen nicht ganz ausgeglichen war, wurde hier zusätzlich die Veränderung im Studienverlauf erfasst (Abbildung 4-12) (23). In der Ruxolitinib-Gruppe zeigte sich nach 24 Wochen bei 42,6% der Patienten eine Verbesserung des ECOG-Performance-Status gegenüber Studienbeginn, während dies nur bei 29,1% der Patienten im Placebo-Arm der Fall war. Dagegen war der Anteil der Patienten, bei denen nach 24 Wochen eine Verschlechterung des ECOG-Performance-Status dokumentiert wurde, in der Placebo-Gruppe mit 31,8% im Vergleich zu 11,6% in der Ruxolitinib-Gruppe deutlich höher (23).

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

4.3.1.3.1.7 Gesundheitsbezogene Lebensqualität – 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 4-31: Operationalisierung von „Gesundheitsbezogene Lebensqualität“

Studie	Operationalisierung
COMFORT-II	Der Endpunkt „Gesundheitsbezogene Lebensqualität“ wurde durch zwei Operationalisierungen erfasst: 1. <u>Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn</u> Deskriptive Darstellung der Skalenwerte des QLQ-C30-Fragebogens sowie deren Veränderung gegenüber Studienbeginn zu allen Erhebungszeitpunkten. Mit dem generischen QLQ-C30-Fragebogen der EORTC werden über 30 Fragen die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität erfasst. Diese umfassen zwei Globalbeurteilungen (allgemeiner Gesundheitszustand/Lebensqualität), fünf Funktionsskalen (körperliche Funktion, Rollenfunktion, kognitive Funktion, emotionale Funktion und soziale Funktion), drei Symptomskalen (Fatigue, Übelkeit/Erbrechen, Schmerzen) und einige Einzel-Items, die die häufigsten von Tumorpatienten berichteten Symptome (Dyspnoe, Schlafstörungen, Appetitlosigkeit, Obstipation, Diarrhö) sowie finanzielle Schwierigkeiten erfassen. Zur besseren Vergleichbarkeit werden die Zahlenwerte des Fragebogens jeweils in lineare Werte auf einer Skala von 0 bis 100 transformiert. Für die Globalbeurteilungen und die Funktionsskalen sind höhere Werte mit einem besseren Zustand bzw. einem besseren Befinden des Patienten gleichzusetzen, während höhere Werte bei den Symptomskalen eine stärker ausgeprägte Symptomatik anzeigen. 2. <u>Veränderung des FACT-Lym gegenüber Studienbeginn</u> Deskriptive Darstellung der Summenscores des FACT-Lym-Fragebogens sowie deren Veränderung gegenüber Studienbeginn zu allen Erhebungszeitpunkten. Der FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma) besteht aus dem generischen Kernfragebogen FACT-G (FACT-General), der die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität wie physisches Wohlbefinden (PWB), soziales/familiäres Wohlbefinden (SWB), emotionales Wohlbefinden (EWB) und funktionelles Wohlbefinden (FWB) abfragt, und der Subskala LymS, die krankheitsspezifische Probleme erfasst. Folgende Auswertungen können durchgeführt werden, wobei höhere Werte mit einer besseren Lebensqualität gleichzusetzen sind: - FACT-G Total Score (Range 0-108): Kernfragebogen FACT-G (27 Items) - LymS (Range 0-60): krankheitsspezifische Subskala (15 Items) - FACT-Lym Total Score (Range 0-168): FACT-G Total Score + LymS - FACT-TOI (Range 0-116): PWB + FWB + LymS Daten zur Lebensqualität wurden zu Studienbeginn, nach 8, 16, 24 und 48 Wochen erhoben.
COMFORT-I	Der Endpunkt „Gesundheitsbezogene Lebensqualität“ wurde durch eine Operationalisierung erfasst: 1. <u>Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn</u> Deskriptive Darstellung der Skalenwerte des QLQ-C30-Fragebogens sowie deren Veränderung gegenüber Studienbeginn zu allen Erhebungszeitpunkten. Daten zur Lebensqualität wurden zu Studienbeginn, nach 4, 8, 12, 16, 24, 48 und 72 Wochen sowie danach alle 24 Wochen erhoben.
Abkürzungen: EORTC-QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EWB, Emotional Well-Being; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; FWB, Functional Well-Being; LymS, Lymphoma Subscale; PWB, Physical Well-Being; SWB, Social/Family Well-Being; TOI, Trial Outcome Index	

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.

Tabelle 4-32: Bewertung des Verzerrungspotenzials für den Endpunkt „Gesundheitsbezogene Lebensqualität“ 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
Operationalisierung: Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn						
COMFORT-II	niedrig	nein	ja	ja	ja	hoch
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Veränderung des FACT-Lym gegenüber Studienbeginn						
COMFORT-II	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesundheitsbezogene Lebensqualität“ als „hoch“ eingestuft, da eine relevante Verzerrung der Ergebnisse aufgrund des offenen Studiendesigns nicht ausgeschlossen werden kann. Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesundheitsbezogene Lebensqualität“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Tabelle 4-33: Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel: EORTC-QLQ-C30

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Allgemeiner Gesundheitszustand/ Lebensqualität (Range: 0-100)				
N (gesamt)	133	66	136	104
Mittelwert zu Studienbeginn (± SD)	55,5 ± 20,8	50,1 ± 23,3	52,6 ± 23,1	54,5 ± 22,6
Nach 24 Wochen				
N (gesamt)	95	37	136	104
Abs. Veränderung zu Beginn (± SD)	8,4 ± 18,57	5,2 ± 23,76	12,3 ± 25,4	-3,4 ± 21,5
Differenz vs. BAT/Placebo [95%-KI]	k. A.		14,62 [9,38; 19,86] [†]	
p-Wert	k. A.		< 0,0001*	
Nach 48 Wochen				
N (gesamt)	66	27	-	
Abs. Veränderung zu Beginn (± SD)	9,1 ± 24,23	3,4 ± 24,27		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Körperliche Funktion (Range: 0-100)				
N (gesamt)	136	67	137	107
Mittelwert zu Studienbeginn (± SD)	69,5 ± 22,3	65,1 ± 24,0	69,5 ± 21,1	67,6 ± 20,5
Nach 24 Wochen				
N (gesamt)	98	39	137	107
Abs. Veränderung zu Beginn (± SD)	3,7 ± 14,80	-1,2 ± 16,59	9,3 ± 18,5	-3,5 ± 17,4
Differenz vs. BAT/Placebo [95%-KI]	k. A.		13,40 [9,08; 17,71] [†]	
p-Wert	k. A.		< 0,0001*	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	4,1 ± 16,79	-3,6 ± 19,16		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Rollenfunktion (Range: 0-100)				
N (gesamt)	136	67	137	105
Mittelwert zu Studienbeginn (± SD)	67,0 ± 30,1	64,4 ± 32,9	64,6 ± 29,2	65,1 ± 30,6
Nach 24 Wochen				
N (gesamt)	96	39	137	105
Abs. Veränderung zu Beginn (± SD)	5,4 ± 27,52	-6,0 ± 27,96	6,4 ± 29,7	-11,1 ± 30,9
Differenz vs. BAT/Placebo [95%-KI]	k. A.		17,34 [10,42; 24,26] [†]	
p-Wert	k. A.		< 0,0001*	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	9,9 ± 28,05	-5,4 ± 26,47		
p-Wert	k. A.			

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Emotionale Funktion (Range: 0-100)				
N (gesamt)	133	66	137	105
Mittelwert zu Studienbeginn (± SD)	75,3 ± 22,3	70,2 ± 26,1	73,7 ± 23,5	76,3 ± 23,5
Nach 24 Wochen				
N (gesamt)	95	37	137	105
Abs. Veränderung zu Beginn (± SD)	4,6 ± 19,74	-0,2 ± 20,27	8,3 ± 20,7	-1,2 ± 23,0
Differenz vs. BAT/Placebo [95%-KI]	k. A.		8,19 [3,40; 12,99] [†]	
p-Wert	k. A.		0,0009*	
Nach 48 Wochen				
N (gesamt)	67	27	-	
Abs. Veränderung zu Beginn (± SD)	5,2 ± 19,62	0,1 ± 21,56		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Kognitive Funktion (Range: 0-100)				
N (gesamt)	133	66	136	103
Mittelwert zu Studienbeginn (± SD)	80,2 ± 21,7	75,3 ± 26,7	81,1 ± 19,1	79,3 ± 20,1
Nach 24 Wochen				
N (gesamt)	95	37	136	103
Abs. Veränderung zu Beginn (± SD)	0,2 ± 19,82	0,0 ± 18,43	2,5 ± 19,0	-1,1 ± 21,0
Differenz vs. BAT/Placebo [95%-KI]	k. A.		4,42 [-0,16; 9,01] [†]	
p-Wert	k. A.		0,0587*	
Nach 48 Wochen				
N (gesamt)	67	27	-	
Abs. Veränderung zu Beginn (± SD)	0,7 ± 19,56	-1,9 ± 19,25		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Soziale Funktion (Range: 0-100)				
N (gesamt)	132	66	135	106
Mittelwert zu Studienbeginn (± SD)	80,4 ± 25,1	73,5 ± 30,1	68,0 ± 26,6	65,9 ± 29,8
Nach 24 Wochen				
N (gesamt)	94	37	135	106
Abs. Veränderung zu Beginn (± SD)	4,4 ± 22,56	4,1 ± 32,73	9,4 ± 25,5	-9,0 ± 28,1
Differenz vs. BAT/Placebo [95%-KI]	k. A.		19,29 [13,22; 25,35] [†]	
p-Wert	k. A.		< 0,0001*	
Nach 48 Wochen				
N (gesamt)	67	27	-	
Abs. Veränderung zu Beginn (± SD)	1,2 ± 18,65	10,5 ± 29,65		
p-Wert	k. A.			

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Fatigue (Range: 0-100)				
N (gesamt)	136	67	137	107
Mittelwert zu Studienbeginn (± SD)	44,4 ± 27,6	51,2 ± 28,7	50,7 ± 24,82	52,2 ± 26,76
Nach 24 Wochen				
N (gesamt)	98	39	137	107
Abs. Veränderung zu Beginn (± SD)	-8,1 ± 23,26	-1,0 ± 23,57	-14,8 ± 27,22	1,8 ± 24,71
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-17,3 [-23,1; -11,5] [†]	
p-Wert	k. A.		< 0,0001 [*]	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	-12,8 ± 28,57	0,4 ± 25,39		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Übelkeit/Erbrechen (Range: 0-100)				
N (gesamt)	135	67	136	105
Mittelwert zu Studienbeginn (± SD)	5,6 ± 12,1	6,7 ± 18,1	8,7 ± 13,90	9,7 ± 16,47
Nach 24 Wochen				
N (gesamt)	97	39	136	105
Abs. Veränderung zu Beginn (± SD)	-1,0 ± 12,22	-0,9 ± 15,27	-5,9 ± 14,42	-0,0 ± 18,92
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-6,59 [-9,79; -3,40] [†]	
p-Wert	k. A.		< 0,0001 [*]	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	-1,7 ± 13,14	2,4 ± 17,40		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Schmerzen (Range: 0-100)				
N (gesamt)	135	67	137	104
Mittelwert zu Studienbeginn (± SD)	23,6 ± 29,3	30,6 ± 30,8	28,5 ± 25,58	30,9 ± 28,48
Nach 24 Wochen				
N (gesamt)	98	39	137	104
Abs. Veränderung zu Beginn (± SD)	-5,4 ± 24,50	3,0 ± 28,06	-6,7 ± 29,32	8,3 ± 27,47
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-16,5 [-22,6; -10,3] [†]	
p-Wert	k. A.		< 0,0001 [*]	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	-1,9 ± 28,80	3,0 ± 17,60		
p-Wert	k. A.			

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Dyspnoe (Range: 0-100)				
N (gesamt)	136	67	136	105
Mittelwert zu Studienbeginn (± SD)	34,3 ± 30,9	38,8 ± 36,5	37,2 ± 28,14	35,2 ± 31,97
Nach 24 Wochen				
N (gesamt)	97	38	136	105
Abs. Veränderung zu Beginn (± SD)	-1,7 ± 28,61	5,3 ± 31,51	-11,0 ± 28,42	1,0 ± 27,53
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-11,0 [-17,1; -4,87] [†]	
p-Wert	k. A.		0,0005*	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	-6,3 ± 30,40	4,8 ± 34,80		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Schlafstörungen (Range: 0-100)				
N (gesamt)	135	67	137	105
Mittelwert zu Studienbeginn (± SD)	38,0 ± 34,6	24,9 ± 31,4	38,4 ± 32,06	38,7 ± 31,06
Nach 24 Wochen				
N (gesamt)	97	39	137	105
Abs. Veränderung zu Beginn (± SD)	-9,3 ± 31,08	6,0 ± 28,48	-12,9 ± 29,50	-2,2 ± 32,12
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-10,8 [-17,3; -4,33] [†]	
p-Wert	k. A.		0,0012*	
Nach 48 Wochen				
N (gesamt)	68	28	-	
Abs. Veränderung zu Beginn (± SD)	-12,3 ± 35,46	6,0 ± 24,09		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Appetitlosigkeit (Range: 0-100)				
N (gesamt)	136	67	135	107
Mittelwert zu Studienbeginn (± SD)	16,7 ± 26,0	21,4 ± 32,2	32,8 ± 32,33	30,8 ± 28,86
Nach 24 Wochen				
N (gesamt)	98	39	135	107
Abs. Veränderung zu Beginn (± SD)	-10,9 ± 24,77	0,9 ± 34,61	-21,0 ± 32,78	0,6 ± 33,96
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-20,2 [-26,8; -13,7] [†]	
p-Wert	k. A.		< 0,0001*	
Nach 48 Wochen				
N (gesamt)	69	28	-	
Abs. Veränderung zu Beginn (± SD)	-8,2 ± 27,06	9,5 ± 33,77		
p-Wert	k. A.			

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Obstipation (Range: 0-100)				
N (gesamt)	133	67	135	107
Mittelwert zu Studienbeginn (± SD)	10,5 ± 22,2	7,0 ± 16,0	14,8 ± 22,18	14,6 ± 25,15
Nach 24 Wochen				
N (gesamt)	95	38	135	107
Abs. Veränderung zu Beginn (± SD)	2,5 ± 21,88	1,8 ± 13,30	-2,5 ± 28,69	-0,9 ± 22,64
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-1,43 [-7,12; 4,26] [†]	
p-Wert	k. A.		0,6207*	
Nach 48 Wochen				
N (gesamt)	66	27	-	
Abs. Veränderung zu Beginn (± SD)	-0,5 ± 26,47	2,5 ± 12,83		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Diarrhö (Range: 0-100)				
N (gesamt)	133	66	137	105
Mittelwert zu Studienbeginn (± SD)	17,8 ± 26,8	14,6 ± 26,2	20,2 ± 24,37	17,1 ± 24,51
Nach 24 Wochen				
N (gesamt)	95	37	137	105
Abs. Veränderung zu Beginn (± SD)	-5,3 ± 25,87	2,7 ± 31,80	-5,8 ± 24,21	6,0 ± 27,65
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-10,3 [-16,0; -4,54] [†]	
p-Wert	k. A.		0,0005*	
Nach 48 Wochen				
N (gesamt)	65	27	-	
Abs. Veränderung zu Beginn (± SD)	-4,1 ± 23,20	-6,2 ± 24,52		
p-Wert	k. A.			
Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn: Finanzielle Schwierigkeiten (Range: 0-100)				
N (gesamt)	132	65	136	105
Mittelwert zu Studienbeginn (± SD)	10,9 ± 23,1	12,8 ± 24,07	16,7 ± 26,61	19,0 ± 28,07
Nach 24 Wochen				
N (gesamt)	94	36	136	105
Abs. Veränderung zu Beginn (± SD)	0,0 ± 19,55	-1,9 ± 17,72	-3,9 ± 20,71	3,2 ± 25,55
Differenz vs. BAT/Placebo [95%-KI]	k. A.		-8,03 [-13,2; -2,82] [†]	
p-Wert	k. A.		0,0026*	
Nach 48 Wochen				
N (gesamt)	67	26	-	
Abs. Veränderung zu Beginn (± SD)	0,0 ± 17,41	-1,3 ± 19,96		
p-Wert	k. A.			

Gesundheitsbezogene Lebensqualität	COMFORT-II		COMFORT-I [‡]	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 154)
[‡] In der COMFORT-I-Studie wurden Daten zur gesundheitsbezogenen Lebensqualität regelmäßig bis zu Woche 144 erhoben (24). Da nach der Primäranalyse innerhalb kurzer Zeit alle Patienten aus dem Placebo-Arm in den Ruxolitinib-Arm der Studie wechselten und Patienten im Placebo-Arm bei der Auswertung nur bis zum Cross-over in den Ruxolitinib-Arm berücksichtigt wurden, ist die Analyse nach Woche 24 für einen Vergleich zwischen den beiden Studienarmen wenig aussagekräftig. Auf eine entsprechende Darstellung der erhobenen Daten wurde aus diesem Grund verzichtet. [†] Berechnung der LSM-Differenz mit zugehörigem 95%-Konfidenzintervall auf Basis einer Kovarianzanalyse mit dem Baseline-Wert als Kovariate [*] Berechnung des p-Werts mit Hilfe der oben genannten Kovarianzanalyse Abkürzungen: BAT, Best Available Therapy; EORTC-QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; k. A., keine Angabe; LSM, Least Square Means; SD, Standardabweichung				

Tabelle 4-34: Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel: FACT-Lym

Gesundheitsbezogene Lebensqualität	COMFORT-II	
	Ruxolitinib (N= 146)	BAT (N = 73)
Veränderung des FACT-Lym gegenüber Studienbeginn: FACT-Lym Total Score (Range: 0-168)		
Studienbeginn		
N (gesamt)	141	70
Mittelwert (± SD)	116,2 ± 24,89	113,5 ± 23,29
Nach 8 Wochen		
N (gesamt)	116	50
Absolute Veränderung zu Studienbeginn (± SD)	12,1 ± 22,1	-0,1 ± 19,9
Nach 16 Wochen		
N (gesamt)	116	42
Absolute Veränderung zu Studienbeginn (± SD)	10,5 ± 22,0	-3,3 ± 13,4
Nach 24 Wochen		
N (gesamt)	100	40
Absolute Veränderung zu Studienbeginn (± SD)	11,9 ± 20,3	-0,8 ± 17,1
Nach 48 Wochen		
N (gesamt)	70	29
Absolute Veränderung zu Studienbeginn (± SD)	11,3 ± 25,0	-0,9 ± 19,6

Gesundheitsbezogene Lebensqualität	COMFORT-II	
	Ruxolitinib (N= 146)	BAT (N = 73)
Veränderung des FACT-Lym gegenüber Studienbeginn: FACT-TOI (Range: 0-116)		
Studienbeginn		
N (gesamt)	141	70
Mittelwert ($\pm$ SD)	78,6 $\pm$ 19,84	76,0 $\pm$ 19,84
Nach 8 Wochen		
N (gesamt)	116	50
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	10,1 $\pm$ 17,0	0,0 $\pm$ 14,4
Nach 16 Wochen		
N (gesamt)	116	42
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	8,7 $\pm$ 15,9	-2,7 $\pm$ 10,0
Nach 24 Wochen		
N (gesamt)	100	40
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	9,0 $\pm$ 14,2	-0,4 $\pm$ 13,8
Nach 48 Wochen		
N (gesamt)	70	29
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	9,1 $\pm$ 18,8	-0,9 $\pm$ 16,1
Veränderung des FACT-Lym gegenüber Studienbeginn: FACT-G Total Score (Range: 0-108)		
Studienbeginn		
N (gesamt)	141	70
Mittelwert ($\pm$ SD)	73,5 $\pm$ 17,19	72,0 $\pm$ 14,06
Nach 8 Wochen		
N (gesamt)	116	50
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	6,3 $\pm$ 15,0	-0,1 $\pm$ 12,5
Nach 16 Wochen		
N (gesamt)	116	42
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	5,3 $\pm$ 16,4	-2,9 $\pm$ 10,4
Nach 24 Wochen		
N (gesamt)	100	40
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	6,3 $\pm$ 17,2	-1,1 $\pm$ 11,0
Nach 48 Wochen		
N (gesamt)	68	29
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	8,9 $\pm$ 15,9	0,1 $\pm$ 13,8

Gesundheitsbezogene Lebensqualität	COMFORT-II	
	Ruxolitinib (N= 146)	BAT (N = 73)
Veränderung des FACT-Lym gegenüber Studienbeginn: LymS (Range: 0-60)		
Studienbeginn		
N (gesamt)	139	69
Mittelwert ($\pm$ SD)	43,4 $\pm$ 9,67	42,1 $\pm$ 10,41
Nach 8 Wochen		
N (gesamt)	112	49
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	5,9 $\pm$ 7,3	1,2 $\pm$ 6,3
Nach 16 Wochen		
N (gesamt)	113	42
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	5,7 $\pm$ 7,3	-0,4 $\pm$ 6,0
Nach 24 Wochen		
N (gesamt)	99	40
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	6,1 $\pm$ 6,7	0,3 $\pm$ 8,4
Nach 48 Wochen		
N (gesamt)	67	28
Absolute Veränderung zu Studienbeginn ($\pm$ SD)	6,0 $\pm$ 7,6	0,7 $\pm$ 9,1
Abkürzungen: BAT, Best Available Therapy; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; LymS, Lymphoma Subscale; SD, Standardabweichung; TOL, Trial Outcome Index		

Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn

Mit dem QLQ-C30-Fragebogen der EORTC (European Organisation für Research and Treatment of Cancer) werden die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität abgefragt (27). Er umfasst zwei Globalbeurteilungen (allgemeiner Gesundheitszustand/Lebensqualität), fünf Funktionsskalen (körperliche Funktion, Rollenfunktion, kognitive Funktion, emotionale Funktion und soziale Funktion), drei Symptomskalen (Fatigue, Übelkeit/Erbrechen, Schmerzen) sowie einige Einzel-Items, die die häufigsten von Tumorpatienten berichteten Symptome wie Dyspnoe, Schlafstörungen, Appetitlosigkeit, Obstipation und Diarrhö erfassen. Die einzelnen Skalen bzw. Einzel-Items können jeweils Werte von 0 bis 100 annehmen. Für die Globalbeurteilungen und die Funktionsskalen sind höhere Werte mit einer besseren Lebensqualität des Patienten gleichzusetzen, während höhere Werte bei den Symptomskalen eine stärker ausgeprägte Symptomatik anzeigen. Dabei wurde ein Unterschied von mindestens 10 Punkten als Schwellenwert für die klinische Relevanz beschrieben (31;39;65).

Im Verlauf der **COMFORT-II**-Studie konnte unter der Behandlung mit Ruxolitinib in Bezug auf nahezu alle Subskalen und Symptome des QLQ-C30-Fragebogens eine deutliche Verbesserung im Vergleich zu Studienbeginn erreicht werden (Tabelle 4-33) (31). In der Kontrollgruppe kam es dagegen – wenn überhaupt – nur zu einer minimalen Verbesserung der Scores, meist wurde sogar eine Verschlechterung der Werte beobachtet (31). Eine klinisch relevante Differenz von ≥ 10 Punkten zwischen den Behandlungsgruppen konnte sowohl in Bezug auf die Globalskala „allgemeiner Gesundheitszustand/Lebensqualität“ als auch die Subskala „Rollenfunktion“ mindestens einmal in der Zeit bis zur Primäranalyse nachgewiesen werden. Auch bei den Symptomen Fatigue, Schmerzen, Dyspnoe, Schlafstörungen, Appetitverlust und Diarrhö wurde die Schwelle der klinischen Relevanz zu mindestens einem Erhebungszeitpunkt überschritten (31;64). Dabei waren die QLQ-C30-Scores zu Studienbeginn in allen Dimensionen zwischen den beiden Studienarmen vergleichbar. Lediglich beim Einzel-Item „Schlafstörungen“ war ein relevanter Unterschied zwischen den beiden Gruppen zu beobachten (31).

Auch in der **COMFORT-I**-Studie erfuhren Patienten unter der Therapie mit Ruxolitinib über alle Bereiche hinweg eine deutliche Verbesserung gegenüber Studienbeginn, wohingegen es unter der Placebo-Behandlung meist sogar zu einer Verschlechterung der Anfangswerte kam (Tabelle 4-33) (23). Nach 24 Wochen hatten sich die QLQ-C30-Scores in der Ruxolitinib-Gruppe sowohl bei der Globalbeurteilung „allgemeiner Gesundheitszustand/Lebensqualität“ als auch bei vier der fünf Funktionsskalen statistisch signifikant gegenüber Placebo verbessert (jeweils $p < 0,001$) (22;23). Auch in Bezug auf die Symptome konnte – mit Ausnahme der Obstipation – eine signifikante Überlegenheit von Ruxolitinib gegenüber Placebo gezeigt werden (jeweils $p < 0,001$) (23). Die klinisch relevante Differenz von mindestens 10 Punkten zwischen den Behandlungsgruppen wurde dabei sowohl bei der Globalbeurteilung und den Funktionsskalen „körperliche Funktion“, „Rollenfunktion“ und „soziale Funktion“ als auch bei der Mehrzahl der Symptome überschritten. Die Verbesserung der Lebensqualitätsparameter sowie der Symptome blieb dabei über einen Zeitraum von mindestens drei Jahren stabil (24).

Veränderung des FACT-Lym gegenüber Studienbeginn

Um auch spezifische Symptome von Myelofibrose-Patienten zu erfassen, wurde in der Studie **COMFORT-II** zusätzlich der Fragebogen FACT-Lym (Functional Assessment of Cancer Therapy – Lymphoma) eingesetzt. Neben dem generischen Kernfragebogen FACT-G (FACT-General), der die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität wie physisches Wohlbefinden (PWB), soziales/familiäres Wohlbefinden (SWB), emotionales Wohlbefinden (EWB) und funktionelles Wohlbefinden (FWB) abfragt, werden über die

Subskala LymS (Lymphoma Subscale) des Fragebogens krankheitsspezifische Probleme wie die auch für die Myelofibrose typischen Symptome Fieber, Nachtschweiß, Juckreiz und Gewichtsverlust erfasst (40;41). Bei der Auswertung der Skalen des FACT-Lym sind höhere Werte mit einer besseren Lebensqualität gleichzusetzen. Die klinisch relevante Differenz (minimal important difference, MID) für die jeweiligen Skalen ist in Tabelle 4-35 angegeben.

Tabelle 4-35: Klinische Relevanzschwellen für die Skalen des Fragebogens FACT-Lym

Skala	Zusammensetzung	Range	MID	Referenz
FACT-G Total Score	Generischer Kernfragebogen 4 Bereiche: PWB, SWB, EWB, FWB	0-108	3-7	(42)
LymS	Krankheitsspezifische Subskala	0-60	2,9-5,4	(43)
			3-5	(41)
FACT-Lym Total Score	FACT-G Total Score + LymS	0-168	6,5-11,2	(43)
FACT-TOI	PWB + FWB + LymS	0-116	5,5-11	(43)
Abkürzungen: EWB, Emotional Well-Being; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; FACT-G, FACT-General; FWB, Functional Well-Being; LymS, Lymphoma Subscale; MID, minimal important difference; PWB, Physical Well-Being; SWB, Social/Family Well-Being; TOI, Trial Outcome Index				

Wie die Auswertung des Fragebogens FACT-Lym zeigte, kam es unter der Behandlung mit Ruxolitinib nicht nur zu einer Verbesserung der allgemeinen Lebensqualität, sondern auch zu einer Reduktion der spezifischen Myelofibrose-Symptome (Tabelle 4-34) (31;64). So verbesserten sich in der Ruxolitinib-Gruppe im Vergleich zu Studienbeginn sowohl der generische FACT-G Total Score als auch die krankheitsspezifische Subskala LymS, in der die für die Erkrankung typischen Symptome erfasst werden. Auch bei den zusammengesetzten Skalen FACT-Lym Total Score und FACT-TOI war eine deutliche Verbesserung gegenüber Studienbeginn zu verzeichnen. Im Gegensatz dazu blieben die Werte aller FACT-Lym-Skalen unter der BAT-Kontrollbehandlung unverändert, teilweise mussten die Patienten sogar eine Verschlechterung hinnehmen (31;64). Die Unterschiede zwischen den Behandlungsgruppen zeigten sich durchgehend ab der ersten Erhebung nach Woche 8 bis zur letzten Messung nach 48 Wochen (31). Die klinische Relevanzschwelle wurde dabei über alle Skalen hinweg zu mindestens einem Messzeitpunkt überschritten (41-43). Bei der krankheitsspezifischen Subskala LymS lag die Differenz zwischen den beiden Gruppen zu allen Zeitpunkten über der oberen Grenze der beschriebenen MID (41;43).

Zusammenfassung

Die beiden Studien COMFORT-I und COMFORT-II belegen damit eine deutliche, lang anhaltende Verbesserung der gesundheitsbezogenen Lebensqualität sowie eine drastische

Reduktion der Myelofibrose-assoziierten Symptome unter der Therapie mit Ruxolitinib. Dagegen war die zweckmäßige Vergleichstherapie nicht dazu geeignet, die für die Erkrankung typische Symptomatik zu kontrollieren und die Lebensqualität der betroffenen Patienten zu verbessern (65).

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

4.3.1.3.1.8 Verträglichkeit – 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 4-36: Operationalisierung von „Verträglichkeit“

Studie	Operationalisierung
COMFORT-II	Der Endpunkt umfasste folgende Operationalisierungen • Unerwünschte Ereignisse (UEs) • Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4) • Schwerwiegende unerwünschte Ereignisse (SUEs) • Unerwünschte Ereignisse, die zum Therapieabbruch führten • Häufige unerwünschte Ereignisse* Die Patienten wurden kontinuierlich hinsichtlich UEs bzw. SUEs bis 28 Tage nach der letzten Behandlung überwacht. Unerwünschte Ereignisse wurden mit Hilfe des Medical Dictionary for Regulatory Activities (MedDRA, Version 12.0) kodierte und nach CTCAE (Common Terminology Criteria for Adverse Events, Version 3.0) in Bezug auf ihren Schweregrad klassifiziert. Alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten, wurden in die Analyse eingeschlossen (Sicherheitspopulation).
COMFORT-I	Der Endpunkt umfasste folgende Operationalisierungen • Unerwünschte Ereignisse (UEs) • Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4) • Schwerwiegende unerwünschte Ereignisse (SUEs) • Unerwünschte Ereignisse, die zum Therapieabbruch führten • Häufige unerwünschte Ereignisse* Die Patienten wurden kontinuierlich hinsichtlich UEs bzw. SUEs bis 28 Tage nach der letzten Behandlung überwacht. Unerwünschte Ereignisse wurden mit Hilfe des Medical Dictionary for Regulatory Activities (MedDRA, Version 12.0) kodierte und nach CTCAE (Common Terminology Criteria for Adverse Events, Version 3.0) in Bezug auf ihren Schweregrad klassifiziert. Alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten, wurden in die Analyse eingeschlossen (Sicherheitspopulation).
* Ereignisse, die in einer der beiden Studien bei $\geq 10\%$ der Patienten unter der Behandlung mit Ruxolitinib aufgetreten sind, operationalisiert über MedDRA PT und ausgewählt nach Relevanz Abkürzungen: CTCAE, Common Terminology Criteria for Adverse Events; MedDRA, Medical Dictionary for Regulatory Activities; PT, Preferred Term; SUE, schwerwiegendes unerwünschtes Ereignis; UE, unerwünschtes Ereignis	

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 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt „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
Operationalisierung: Unerwünschte Ereignisse						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Schwerwiegende unerwünschte Ereignisse						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Unerwünschte Ereignisse, die zum Therapieabbruch führten						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig
Operationalisierung: Häufige unerwünschte Ereignisse						
COMFORT-II	niedrig	nein	ja	ja	ja	niedrig
COMFORT-I	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

COMFORT-II

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen

oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74). Siehe auch Tabelle 4-70 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

COMFORT-I

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Siehe auch Tabelle 4-71 in Anhang 4-F für eine detaillierte Bewertung des Verzerrungspotenzials.

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

Tabelle 4-38: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel

Verträglichkeit	COMFORT-II*		COMFORT-I†	
	Ruxolitinib (N= 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 151)
Unerwünschte Ereignisse				
Patienten mit Ereignis (n, %)	145 (99,3)	66 (90,4)	151 (97,4)	148 (98,0)
Differenz vs. BAT/Placebo [95%-KI]	8,9 [1,0 – 16,8]		-0,6 [-4,6 – 3,4]	
RR vs. BAT/Placebo [95%-KI]	1,1 [1,0 – 1,2]		1,0 [1,0 – 1,0]	
OR vs. BAT/Placebo [95%-KI]	15,4 [1,9 – 127,5]		0,8 [0,2 – 3,5]	
Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)				
Patienten mit Ereignis (n, %)	63 (43,2)	18 (24,7)	73 (47,1)	67 (44,4)
Differenz vs. BAT/Placebo [95%-KI]	18,5 [4,7 – 32,3]		2,7 [-9,1 – 14,5]	
RR vs. BAT/Placebo [95%-KI]	1,8 [1,1 – 2,7]		1,1 [0,8 – 1,4]	
OR vs. BAT/Placebo [95%-KI]	2,3 [1,2 – 4,3]		1,1 [0,7 – 1,8]	
Schwerwiegende unerwünschte Ereignisse				
Patienten mit Ereignis (n, %)	44 (30,1)	21 (28,8)	43 (27,7)	53 (35,1)
Differenz vs. BAT/Placebo [95%-KI]	1,4 [-12,4 – 15,2]		-7,4 [-18,4 – 3,7]	
RR vs. BAT/Placebo [95%-KI]	1,0 [0,7 – 1,6]		0,8 [0,6 – 1,1]	
OR vs. BAT/Placebo [95%-KI]	1,1 [0,6 – 2,0]		0,7 [0,4 – 1,2]	
Unerwünschte Ereignisse, die zum Therapieabbruch führten				
Patienten mit Ereignis (n, %)	12 (8,2)	6 (8,2)	17 (11,0)	16 (10,6)
Differenz vs. BAT/Placebo [95%-KI]	0,0 [-8,7 – 8,7]		0,4 [-7,2 – 8,0]	
RR vs. BAT/Placebo [95%-KI]	1,0 [0,4 – 2,6]		1,0 [0,5 – 2,0]	
OR vs. BAT/Placebo [95%-KI]	1,0 [0,4 – 2,8]		1,0 [0,5 – 2,1]	
* nach einer medianen Expositionszeit von 51,4 Wochen bzw. 45,1 Wochen im Ruxolitinib- bzw. BAT-Arm				
† nach einer medianen Expositionszeit von 33,7 Wochen bzw. 30,1 Wochen im Ruxolitinib- bzw. Placebo-Arm				
Abkürzungen: BAT, Best Available Therapy; CTCAE, Common Terminology Criteria for Adverse Events; KI, Konfidenzintervall; OR, Odds Ratio; RR, Risk Ratio				

Tabelle 4-39: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse (Inzidenz ≥ 10%)

Verträglichkeit	COMFORT-II*		COMFORT-I†	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 151)
Hämatologische UEs, alle Grade (n, %)*				
Anämie	60 (41,1)	9 (12,3)	48 (31,0)	21 (13,9)
Thrombozytopenie	65 (44,5)	7 (9,6)	53 (34,2)	14 (9,3)
Hämatologische UEs, Grad 3/4 (n, %)				
Anämie	16 (11,0)	3 (4,1)	24 (15,5)	7 (4,6)
Thrombozytopenie	11 (7,5)	3 (4,1)	13 (8,4)	3 (2,0)

Verträglichkeit	COMFORT-II*		COMFORT-I†	
	Ruxolitinib (N = 146)	BAT (N = 73)	Ruxolitinib (N = 155)	Placebo (N = 151)
Nicht-hämatologische UEs, alle Grade (n, %)[‡]				
Diarrhö	34 (23,3)	9 (12,3)	36 (23,2)	32 (21,2)
Übelkeit	19 (13,0)	5 (6,8)	23 (14,8)	29 (19,2)
Erbrechen	13 (8,9)	1 (1,4)	19 (12,3)	15 (9,9)
Obstipation	11 (7,5)	4 (5,5)	20 (12,9)	18 (11,9)
Schwindel	11 (7,5)	4 (5,5)	23 (14,8)	10 (6,6)
Kopfschmerzen	15 (10,3)	3 (4,1)	23 (14,8)	8 (5,3)
Fatigue	18 (12,3)	6 (8,2)	39 (25,2)	51 (33,8)
Insomnie	9 (6,2)	5 (6,8)	18 (11,6)	15 (9,9)
Asthenie	26 (17,8)	7 (9,6)	8 (5,2)	12 (7,9)
Dyspnoe	23 (15,8)	13 (17,8)	27 (17,4)	26 (17,2)
Husten	20 (13,7)	11 (15,1)	15 (9,7)	13 (8,6)
Pyrexie	20 (13,7)	7 (9,6)	17 (11,0)	11 (7,3)
Arthralgie	18 (12,3)	5 (6,8)	17 (11,0)	13 (8,6)
Extremitätenschmerz	17 (11,6)	3 (4,1)	19 (12,3)	15 (9,9)
Abdominalschmerz	16 (11,0)	10 (13,7)	16 (10,3)	62 (41,1)
Peripheres Ödem	32 (21,9)	19 (26,0)	29 (18,7)	34 (22,5)
Nicht-hämatologische UEs, Grad 3/4 (n, %)				
Diarrhö	2 (1,4)	0 (0,0)	3 (1,9)	0 (0,0)
Übelkeit	1 (0,7)	0 (0,0)	0 (0,0)	1 (0,7)
Erbrechen	0 (0,0)	0 (0,0)	1 (0,6)	1 (0,7)
Obstipation	0 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
Schwindel	0 (0,0)	0 (0,0)	1 (0,6)	0 (0,0)
Kopfschmerzen	2 (1,4)	0 (0,0)	0 (0,0)	0 (0,0)
Fatigue	1 (0,7)	0 (0,0)	8 (5,2)	10 (6,6)
Insomnie	0 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
Asthenie	2 (1,4)	1 (1,4)	1 (0,6)	4 (2,6)
Dyspnoe	1 (0,7)	3 (4,1)	2 (1,3)	6 (4,0)
Husten	0 (0,0)	1 (1,4)	0 (0,0)	0 (0,0)
Pyrexie	3 (2,1)	0 (0,0)	1 (0,6)	1 (0,7)
Arthralgie	1 (0,7)	0 (0,0)	3 (1,9)	1 (0,7)
Extremitätenschmerz	1 (0,7)	0 (0,0)	2 (1,3)	0 (0,0)
Abdominalschmerz	5 (3,4)	2 (2,7)	4 (2,6)	17 (11,3)
Peripheres Ödem	0 (0,0)	0 (0,0)	0 (0,0)	2 (1,3)
* nach einer medianen Expositionszeit von 51,4 Wochen bzw. 45,1 Wochen im Ruxolitinib- bzw. BAT-Arm † nach einer medianen Expositionszeit von 33,7 Wochen bzw. 30,1 Wochen im Ruxolitinib- bzw. Placebo-Arm ‡ Ereignisse, die in einer der beiden Studien bei $\geq 10\%$ der Patienten unter der Behandlung mit Ruxolitinib aufgetreten sind, operationalisiert über MedDRA PT und ausgewählt nach Relevanz Abkürzungen: BAT, Best Available Therapy; MedDRA, Medical Dictionary for Regulatory Activities; PT, Preferred Term; UE, unerwünschtes Ereignis				

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

Tabelle 4-40: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse adjustiert nach Patientenjahren

Verträglichkeit	COMFORT-II			COMFORT-I		
	Ruxolitinib* (N = 146)	BAT (N = 73)	Ruxolitinib (Cross-over) (N = 45)	Ruxolitinib* (N = 155)	Placebo (N = 151)	Ruxolitinib (Cross-over) (N = 111)
Anzahl nicht-hämatologischer UEs pro 100 Patientenjahre, alle Grade[†]						
Exposition (Patientenjahre)	304,87	66,98	44,95	330,58	98,85	171,61
Diarrhö	15,4	19,4	20,0	21,8	41,2	16,3
Übelkeit	8,9	10,5	8,9	12,5	32,3	9,1
Erbrechen	7,2	1,5	6,7	8,8	17,7	4,0
Obstipation	6,2	6,0	2,2	14,0	20,1	7,7
Schwindel	5,2	7,5	8,9	14,3	10,0	9,4
Kopfschmerzen	6,9	6,0	15,6	10,8	9,9	10,1
Fatigue	11,2	11,9	13,3	27,2	67,6	19,5
Insomnie	3,9	10,5	8,9	9,5	14,9	5,2
Asthenie	11,5	13,4	17,8	5,3	11,8	4,0
Dyspnoe	11,2	22,4	20,0	14,9	30,0	9,9
Husten	10,5	17,9	11,1	12,9	13,2	9,9
Pyrexie	11,5	10,5	13,3	11,8	12,0	9,1
Arthralgie	8,9	10,5	13,3	10,3	15,2	8,6
Extremitätenschmerz	7,2	6,0	20,0	12,8	16,5	10,0
Abdominalschmerz	6,6	19,4	6,7	9,5	82,2	9,4
Peripheres Ödem	17,4	31,4	17,8	17,7	40,9	15,8
* unter der Behandlung mit Ruxolitinib in der randomisierten Behandlungsphase sowie der Extensionsphase der Studie						
† häufige unerwünschte Ereignisse, operationalisiert über MedDRA PT und ausgewählt nach Relevanz						
Abkürzungen: BAT, Best Available Therapy; MedDRA, Medical Dictionary for Regulatory Activities; PT, Preferred Term; UE, unerwünschtes Ereignis						

Tabelle 4-41: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: unerwünschte Ereignisse von besonderem Interesse unter der Behandlung mit Ruxolitinib im Zeitverlauf in der Studie COMFORT-II

Verträglichkeit	Zeit unter der Behandlung mit Ruxolitinib (Wochen)*						
	0-24 (N = 146)	24-48 (N = 134)	48-72 (N = 116)	72-96 (N = 101)	96-120 (N = 93)	120-144 (N = 81)	144-168 (N = 72)
Hämatologische UEs, alle Grade (%)**							
Anämie [†]	34,9	12,7	8,6	13,9	8,6	7,4	8,3
Thrombozytopenie [‡]	43,2	22,4	15,5	12,9	10,8	12,3	2,8
Nicht-hämatologische UEs, alle Grade (%)**							
Blutungen	17,1	14,2	9,5	11,9	7,5	9,9	6,9
Epistaxis	6,8	1,5	0,9	4,0	0	1,2	1,4
Hämatom	5,5	4,5	3,4	1,0	0	2,5	1,4
Infektionen	50,0	35,1	37,9	25,7	43,0	33,3	25,0
Bronchitis	3,4	6,7	8,6	3,0	10,8	4,9	4,2
Gastroenteritis	5,5	3,0	0,9	1,0	2,2	1,2	0
Nasopharyngitis	13,7	5,2	7,8	4,0	10,8	3,7	4,2
Harnwegsinfektion	4,8	2,2	5,2	4,0	5,4	3,7	2,8
Gewichtszunahme	8,2	8,2	5,2	5,0	2,2	0	0
* in der randomisierten Behandlungsphase sowie der Extensionsphase der Studie ** Ereignisse von besonderem Interesse, die in Bezug auf einen MedDRA PT bei $\geq 5\%$ der Patienten in einem der Zeitintervalle aufgetreten sind, operationalisiert über Standardized MedDRA Queries (SMQ) [†] schließt die Preferred Terms „Anämie“ und „Normochromische normozytische Anämie“ ein (der SMQ-Begriff ist „Erythropenie“) [‡] schließt die Preferred Terms „Thrombozytopenie“ und „erniedrigte Thrombozytenzahl“ ein Abkürzungen: MedDRA, Medical Dictionary for Regulatory Activities; PT, Preferred Term; SMQ, Standardized MedDRA Query; UE, unerwünschtes Ereignis							

Tabelle 4-42: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: häufige unerwünschte Ereignisse unter der Behandlung mit Ruxolitinib im Zeitverlauf in der Studie COMFORT-I

Verträglichkeit	Zeit unter der Behandlung mit Ruxolitinib*			
	< 12 Monate (N = 155)	12-23 Monate (N = 130)	24-35 Monate (N = 103)	≥ 36 Monate (N = 82)
Nicht-hämatologische UEs, alle Grade (%)[†]				
Diarrhö	27,8	6,7	10,8	3,9
Übelkeit	16,6	6,8	5,1	5,9
Erbrechen	13,7	2,8	2,4	5,5
Obstipation	14,5	8,6	10,1	9,0
Ekchymose	21,2	10,4	5,7	0
Schwindel	18,1	10,4	3,0	3,5
Kopfschmerzen	16,6	5,1	2,7	0
Fatigue	29,0	15,2	15,3	7,7
Insomnie	13,8	5,7	3,7	0
Dyspnoe	19,2	10,2	2,9	3,3
Husten	13,1	13,3	4,0	6,0
Pyrexie	13,5	7,3	8,5	2,9
Arthralgie	11,8	5,8	6,6	6,3
Extremitätenschmerz	18,0	6,2	4,2	3,3
Abdominalschmerz	13,8	5,7	3,6	0
Peripheres Ödem	21,3	8,4	12,6	0
Infektion der oberen Atemwege	7,7	11,1	4,0	3,2
* in der randomisierten Behandlungsphase sowie der Extensionsphase				
[†] Ereignisse, die bei ≥ 10% der Patienten in einem der Zeitintervalle erstmals aufgetreten sind, operationalisiert über MedDRA PT und ausgewählt nach Relevanz				
Abkürzungen: MedDRA, Medical Dictionary for Regulatory Activities; PT, Preferred Term; UE, unerwünschtes Ereignis				

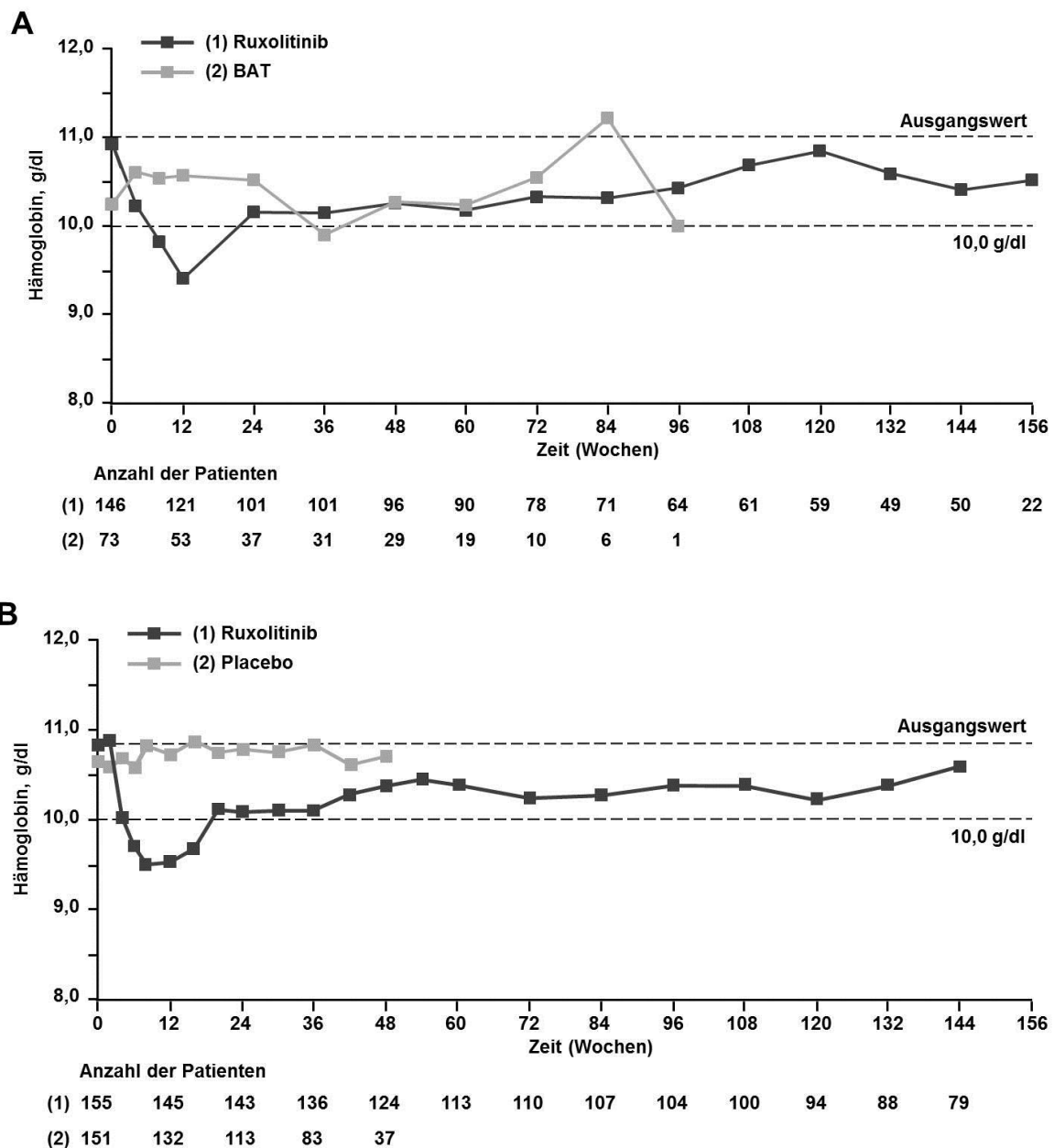


Abbildung 4-13: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: Hämoglobin-Konzentration im Studienverlauf

(A) Mittlere Hämoglobinkonzentration im Zeitverlauf in der Studie COMFORT-II unter der Behandlung mit Ruxolitinib bzw. BAT nach einem medianen Follow-up von 151 Wochen. (B) Mittlere Hämoglobinkonzentration im Zeitverlauf in der Studie COMFORT-I unter der Behandlung mit Ruxolitinib bzw. Placebo nach einem medianen Follow-up von 149 Wochen. BAT, Best Available Therapy; n/N, Zahl der Ereignisse/Zahl der Patienten.

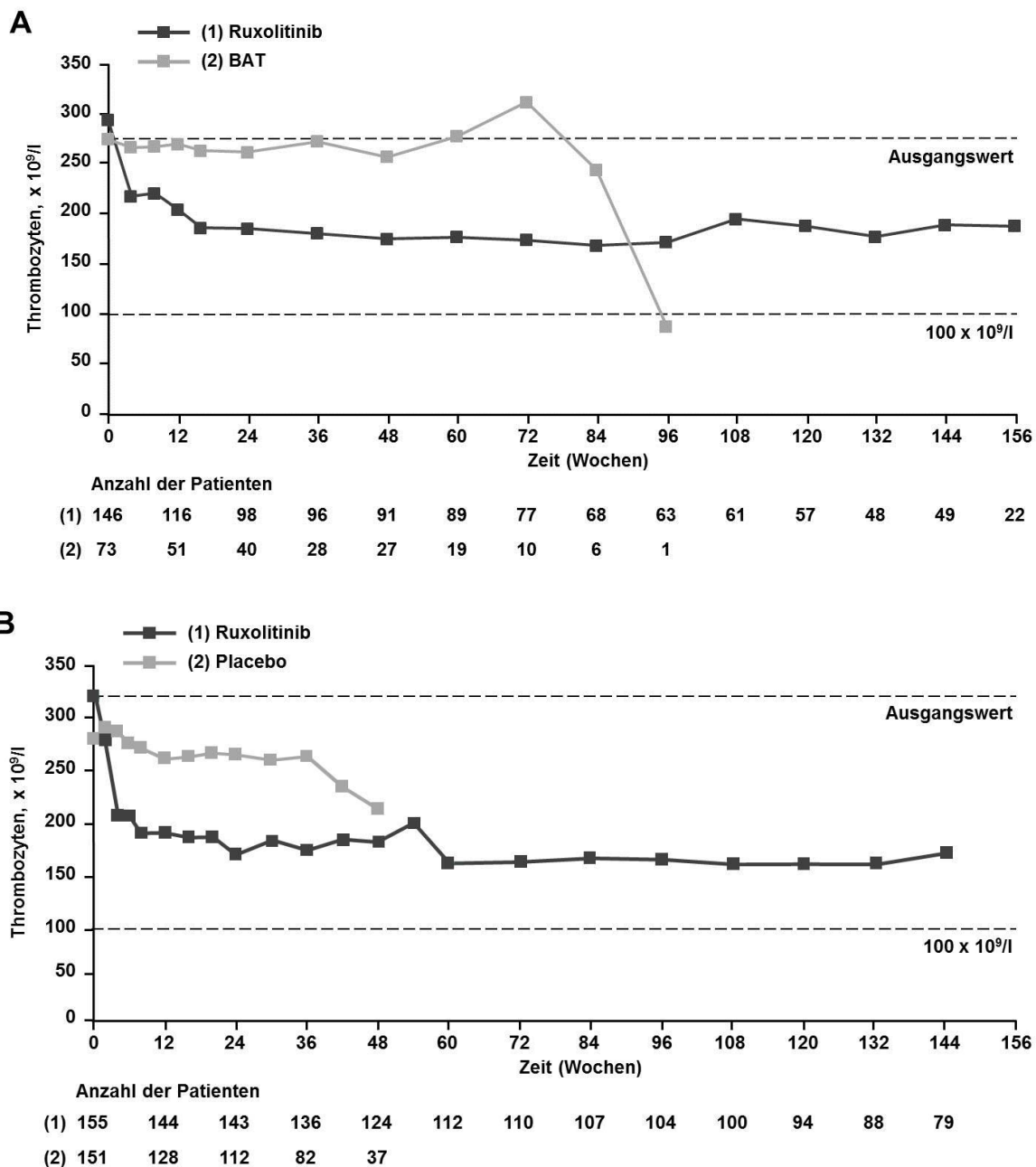


Abbildung 4-14: Ergebnisse für den Endpunkt „Verträglichkeit“ aus RCT mit dem zu bewertenden Arzneimittel: Thrombozytenzahl im Studienverlauf

(A) Mittlere Thrombozytenzahl im Zeitverlauf in der Studie COMFORT-II unter der Behandlung mit Ruxolitinib bzw. BAT nach einem medianen Follow-up von 151 Wochen. (B) Mittlere Thrombozytenzahl im Zeitverlauf in der Studie COMFORT-I unter der Behandlung mit Ruxolitinib bzw. Placebo nach einem medianen Follow-up von 149 Wochen. BAT, Best Available Therapy; n/N, Zahl der Ereignisse/Zahl der Patienten.

Die Evaluation der unerwünschten Ereignisse in den beiden Studien COMFORT-I und COMFORT-II erfolgte jeweils mit Hilfe der Sicherheitspopulation (Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten) (23;31). Dabei wurden alle unerwünschten Ereignisse, die unter der Therapie mit der Studienmedikation bzw. innerhalb von 28 Tagen nach der Einnahme der letzten Dosis auftraten, mit Hilfe der MedDRA (Medical Dictionary for Regulatory Activities)-Terminologie kodiert und unabhängig von ihrem vermuteten Kausalzusammenhang mit der Behandlung dargestellt. Die beschriebenen unerwünschten Ereignisse setzen sich somit sowohl aus den Nebenwirkungen der Therapie als auch aus den Symptomen der Grunderkrankung sowie unabhängigen Ereignissen zusammen.

Unerwünschte Ereignisse unterschiedlichen Schweregrads

Wie die Daten der Primäranalyse der **COMFORT-II**-Studie zeigen, wurde bei fast allen Patienten in den beiden Studienarmen über ein unerwünschtes Ereignis während der Behandlung mit der Studienmedikation berichtet (Tabelle 4-38) (31;64). Die Häufigkeit unerwünschter Ereignisse war dabei unter der Therapie mit Ruxolitinib bzw. BAT vergleichbar. Auch die Häufigkeit schwerwiegender unerwünschter Ereignisse war mit jeweils circa 30% zwischen den beiden Behandlungsarmen ausgeglichen. Lediglich der Anteil der Patienten mit einem Grad-3- oder Grad-4-Ereignis war mit 43,2% in der Ruxolitinib-Gruppe im Vergleich zu 24,7% in der BAT-Gruppe höher. Die unerwünschten Ereignisse waren im Allgemeinen jedoch durch entsprechende symptomorientierte Supportivmaßnahmen sowie Dosisanpassungen gut kontrollierbar und führten nicht zu einem vermehrten Abbruch der Therapie im Ruxolitinib-Arm der Studie. So betrug der Anteil der Patienten, die die Therapie aufgrund von unerwünschten Ereignissen vorzeitig beendeten, in beiden Studienarmen jeweils 8,2% (31;64).

Auch die Primäranalyse der **COMFORT-I**-Studie bestätigt die gute Verträglichkeit von Ruxolitinib bei der Behandlung von Myelofibrose-Patienten (Tabelle 4-38) (22;23). So war die Häufigkeit unerwünschter Ereignissen, schwerer unerwünschter Ereignisse (CTCAE Grad 3 bzw. 4) bzw. unerwünschter Ereignissen, die zum Therapieabbruch führten, zwischen dem Ruxolitinib-Arm und dem Placebo-Arm der Studie vergleichbar. Der Anteil der Patienten mit einem schwerwiegenden unerwünschten Ereignis war mit 27,7% in der Ruxolitinib-Gruppe sogar etwas niedriger als in der Placebo-Gruppe (35,1%) (22;23).

Häufige unerwünschte Ereignisse (Inzidenz $\geq$ 10%)

Die Mehrzahl der unerwünschten Ereignisse in den beiden Studien **COMFORT-I** und **COMFORT-II** war von milder bis moderater Natur und trat studienübergreifend in den beiden Behandlungsgruppen jeweils vergleichbar häufig auf (Tabelle 4-39) (22;23;31;38;64).

Nicht-hämatologische Ereignisse von höherem Schweregrad (Grad 3 oder 4) wurden kaum beobachtet. Während es unter Ruxolitinib häufiger zu Durchfall, Schwindel und Kopfschmerzen kam, wurden periphere Ödeme, Abdominalschmerz und Dyspnoe vermehrt in den beiden Komparator-Armen beobachtet (22;64).

In beiden Studien waren Anämie und Thrombozytopenie die häufigsten unerwünschten Ereignisse unter der Therapie mit Ruxolitinib (Tabelle 4-39) (22;23;31;64). Diese Ereignisse, die als Folge der Knochenmarkfibrose und der damit einhergehenden gestörten Blutbildung auftreten, sind auch für die unbehandelte Erkrankung typisch, traten jedoch unter der Behandlung mit Ruxolitinib gehäuft auf (22;64;75;76). Da JAK2-Kinasen eine wichtige Rolle bei der Hämatopoese spielen, sind entsprechende Nebenwirkungen aufgrund des Wirkmechanismus von Ruxolitinib als JAK-Inhibitor zu erwarten (58;75). Sowohl Anämien als auch Thrombozytopenien waren jedoch in beiden Studien durch Dosisreduktionen und Bluttransfusionen gut zu beherrschen und führten bei weniger als 1% der Patienten zu einer vorzeitigen Beendigung der Studienteilnahme (22;23;31;64). Auch die Inzidenz von Blutungen infolge einer Thrombozytopenie war gering. So wurden in der COMFORT-I-Studie nur bei 11% der Patienten unter der Ruxolitinib-Therapie Blutungsereignisse (operationalisiert durch die MedDRA Preferred Terms „Blutung“, „Hämorrhagie“ und „Epistaxis“) dokumentiert, wobei dies auch bei 8,6% der Patienten im Placebo-Arm der Fall war (23). Ekchymosen wurden mit 18,7% in der Ruxolitinib-Gruppe im Vergleich zu 9,3% in der Placebo-Gruppe zwar etwas häufiger beobachtet, in der COMFORT-II-Studie konnte in Bezug auf die Inzidenz von Ekchymosen jedoch kein relevanter Gruppenunterschied festgestellt werden (1,4% vs. 0%) (23;31). Da noch keine in der Praxis etablierten Dosiswerte vorlagen, waren in beiden Studien bei Auftreten von Thrombozytopenien bzw. Neutropenien im Studienprotokoll Dosisreduktionen oder Therapieunterbrechungen ausdrücklich vorgesehen, die dementsprechend unter der Therapie mit Ruxolitinib häufiger zu beobachten waren als im jeweiligen Vergleichsarm (23;31).

Langzeitdaten zur Sicherheit und Verträglichkeit

Wie die neuesten Auswertungen der 3-Jahres-Analysen zeigen, ergaben sich auch nach längerem Follow-up keine neuen oder unerwarteten Sicherheitsprobleme (24;62;63;71). Im Gegenteil, die Inzidenz der unerwünschten Ereignisse ging nach längerer Behandlung mit Ruxolitinib sogar zurück. So traten Anämien und Thrombozytopenien sowohl in **COMFORT-I** als auch in **COMFORT-II** vor allem zu Beginn der Therapie auf, im weiteren Studienverlauf kam es dann zu einer Stabilisierung der Hämoglobinwerte sowie der Thrombozytenzahl (22;59;63;64;71). In der COMFORT-II-Studie sanken die mittleren Hämoglobinwerte von einem Ausgangswert von 10,9 g/dl innerhalb der ersten 12 Wochen auf

einen Nadir von 9,4 g/dl ab, um sich dann auf einem Niveau von circa 10,4 g/dl ähnlich dem der Patienten im BAT-Arm der Studie einzupendeln (Abbildung 4-13A) (62;63). In der Studie COMFORT-I war ein ähnlicher Kurvenverlauf zu beobachten: hier sanken die mittleren Hämoglobinwerten innerhalb der ersten 8 bis 12 Therapiewochen um etwa 10-12% ab, stabilisierten sich aber nach etwa 24 Wochen auf einem Wert nur leicht unterhalb des Ausgangswerts (Abbildung 4-13B) (24;59). Dementsprechend war der Anteil der Patienten, die ein Erythrozytenkonzentrat benötigten, in den ersten 24 Wochen unter der Therapie mit Ruxolitinib leicht erhöht, näherte sich dann aber dem Wert im Vergleichsarm der jeweiligen Studie an (59;63). Auch die mittlere Thrombozytenzahl stabilisierte sich in beiden Studien nach einem anfänglichen Abfall in den ersten Studienwochen im weiteren Behandlungsverlauf im unteren Normalbereich (Abbildung 4-14A, Abbildung 4-14B) (24;63;71). Darüber hinaus zeigte sich nach 3-jährigem Follow-up, dass sowohl die mittlere Hämoglobinkonzentration als auch die Thrombozytenzahl über diesen Zeitraum stabil blieben (24;62;63).

Auch in Bezug auf nicht-hämatologische unerwünschte Ereignisse ergaben sich mit längerem Follow-up keine Hinweise auf ein verändertes Sicherheitsprofil (24;62;63;71). Da die absolute Expositionszeit der Patienten unter Ruxolitinib aufgrund der Möglichkeit zum Cross-over in den Ruxolitinib-Arm sehr viel höher war als unter der jeweiligen Vergleichstherapie, war ein direkter Vergleich der unerwünschten Ereignisse zwischen den Behandlungsarmen zum Zeitpunkt der 3-Jahres-Analysen nicht mehr sinnvoll. Stattdessen wurde die Zahl der unerwünschten Ereignisse adjustiert nach Patientenjahren ausgewertet (Tabelle 4-40) (24;62;63). Bezogen auf die Therapiedauer trat die Mehrzahl der unerwünschten Ereignisse unter der Therapie mit Ruxolitinib seltener auf als unter der jeweiligen Kontrolltherapie. In Übereinstimmung damit nahm die Inzidenz neu auftretender unerwünschter Ereignisse mit der Dauer der Ruxolitinib-Therapie ab (Tabelle 4-41, Tabelle 4-42) (24;63;71).

Zusammenfassung

Insgesamt bestätigen die beiden Studien COMFORT-I und COMFORT-II die gute Verträglichkeit von Ruxolitinib, die bereits aus der initialen Phase-I/II-Studie bekannt war (77). Obwohl die mediane Expositionszeit unter der Behandlung mit Ruxolitinib bereits zum Zeitpunkt der Primäranalyse in beiden Studien circa 10% über der der entsprechenden Kontrolltherapie lag, war die Häufigkeit unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse sowie von Ereignissen, die zum Abbruch der Therapie führten, zwischen den Behandlungsarmen vergleichbar. Lediglich der Anteil der Patienten mit einem Grad-3- oder Grad-4-Ereignis war in COMFORT-II in der Ruxolitinib-Gruppe höher als in der Kontrollgruppe. Umgekehrt war in COMFORT-I der Anteil der Patienten mit einem schwerwiegenden unerwünschten Ereignis unter Ruxolitinib niedriger als unter Placebo. Auch

nach 3-jährigem Follow-up und einer Expositionsdauer von mehr als 850 Patientenjahren ergaben sich keine neuen oder unerwarteten Sicherheitsprobleme. Im Gegenteil, die Inzidenz der unerwünschten Ereignisse ging nach längerer Behandlung mit Ruxolitinib sogar zurück. Entsprechend dem Wirkmechanismus von Ruxolitinib als JAK-Inhibitor waren Anämien und Thrombozytopenien die häufigsten unerwünschten Ereignisse unter der Therapie mit Ruxolitinib, sie waren jedoch in beiden Studien durch Dosisreduktionen und Bluttransfusionen gut kontrollierbar.

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.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Studien mittels Meta-Analyse nicht sinnvoll.

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.

Tabelle 4-43: Subgruppenanalysen für „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen“ aus der Studie COMFORT-II

Reduktion des Milzvolumens um ≥ 35% nach 48 Wochen	Ruxolitinib (N= 144)	BAT (N = 72)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Geschlecht“			
Männer	N = 81	N = 42	NA
Ansprechen (n, %)	20 (24,7)	0 (0,0)	
Differenz vs. BAT [95%-KI]	24,7 [13,5 – 35,9]		
RR vs. BAT [95%-KI] [†]	21,5 [1,3 – 346,9]		
p-Wert [*]	0,0001		
Frauen	N = 63	N = 30	
Ansprechen (n, %)	21 (33,3)	0 (0,0)	
Differenz vs. BAT [95%-KI]	33,3 [19,2 – 47,4]		
RR vs. BAT [95%-KI] [†]	20,8 [1,3 – 332,7]		
p-Wert [*]	0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 48 Wochen	Ruxolitinib (N= 144)	BAT (N = 72)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Alter“			
≤ 65 Jahre	N = 68	N = 36	NA
Ansprechen (n, %)	20 (29,4)	0 (0,0)	
Differenz vs. BAT [95%-KI]	29,4 [16,5 – 42,4]		
RR vs. BAT [95%-KI] [†]	22,0 [1,4 – 353,3]		
p-Wert [*]	0,0001		
> 65 Jahre	N = 76	N = 36	
Ansprechen (n, %)	21 (27,6)	0 (0,0)	
Differenz vs. BAT [95%-KI]	27,6 [15,5 – 39,7]		
RR vs. BAT [95%-KI] [†]	20,7 [1,3 – 331,8]		
p-Wert [*]	0,0001		
Subgruppenanalyse „Hautfarbe“			
Weiße ^{**}	N = 116	N = 67	NA
Ansprechen (n, %)	33 (28,4)	0 (0,0)	
Differenz vs. BAT [95%-KI]	28,4 [19,1 – 37,8]		
RR vs. BAT [95%-KI] [†]	38,9 [2,4 – 625,4]		
p-Wert [*]	< 0,0001		
Nicht-Weiße ^{**}	N = 0	N = 1	
Ansprechen (n, %)	0	0 (0,0)	
Differenz vs. BAT [95%-KI]	NA		
RR vs. BAT [95%-KI] [†]	NA		
p-Wert [*]	NA		
Subgruppenanalyse „Risikogruppe“			
Hochrisiko	N = 70	N = 35	NA
Ansprechen (n, %)	17 (24,3)	0 (0,0)	
Differenz vs. BAT [95%-KI]	24,3 [12,1 – 36,5]		
RR vs. BAT [95%-KI] [†]	17,7 [1,1 – 286,8]		
p-Wert [*]	0,0006		
Intermediär-Risiko 2	N = 74	N = 37	
Ansprechen (n, %)	24 (32,4)	0 (0,0)	
Differenz vs. BAT [95%-KI]	32,4 [19,7 – 45,1]		
RR vs. BAT [95%-KI] [†]	24,8 [1,6 – 397,3]		
p-Wert [*]	< 0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 48 Wochen	Ruxolitinib (N= 144)	BAT (N = 72)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Myelofibrose-Typ“			
PMF	N = 76	N = 38	NA
Ansprechen (n, %)	14 (18,4)	0 (0,0)	
Differenz vs. BAT [95%-KI]	18,4 [7,7 – 29,1]		
RR vs. BAT [95%-KI] [†]	14,7 [0,9 – 239,8]		
p-Wert [*]	0,0044		
PPV-MF	N = 48	N = 20	
Ansprechen (n, %)	20 (41,7)	0 (0,0)	
Differenz vs. BAT [95%-KI]	41,7 [24,2 – 59,2]		
RR vs. BAT [95%-KI] [†]	17,6 [1,1 – 277,2]		
p-Wert [*]	0,0003		
PET-MF	N = 20	N = 14	
Ansprechen (n, %)	7 (35,0)	0 (0,0)	
Differenz vs. BAT [95%-KI]	35,0 [8,0 – 62,0]		
RR vs. BAT [95%-KI] [†]	10,7 [0,7 – 173,6]		
p-Wert [*]	0,0262		
Subgruppenanalyse „Thrombozytenzahl zu Studienbeginn“			
> 200.000/µl	N = 88	N = 47	NA
Ansprechen (n, %)	31 (35,2)	0 (0,0)	
Differenz vs. BAT [95%-KI]	35,2 [23,6 – 46,8]		
RR vs. BAT [95%-KI] [†]	34,0 [2,1 – 543,1]		
p-Wert [*]	< 0,0001		
≤ 200.000/µl	N = 56	N = 25	
Ansprechen (n, %)	10 (17,9)	0 (0,0)	
Differenz vs. BAT [95%-KI]	17,9 [4,9 – 30,8]		
RR vs. BAT [95%-KI] [†]	9,6 [0,6 – 157,4]		
p-Wert [*]	0,0270		
Subgruppenanalyse „Milzvolumen zu Studienbeginn“			
≤ Median	N = 72	N = 36	NA
Ansprechen (n, %)	18 (25,0)	0 (0,0)	
Differenz vs. BAT [95%-KI]	25,0 [12,9 – 37,1]		
RR vs. BAT [95%-KI] [†]	18,8 [1,2 – 302,6]		
p-Wert [*]	0,0006		
> Median	N = 72	N = 36	
Ansprechen (n, %)	23 (31,9)	0 (0,0)	
Differenz vs. BAT [95%-KI]	31,9 [19,1 – 44,8]		
RR vs. BAT [95%-KI] [†]	23,8 [1,5 – 381,4]		
p-Wert [*]	< 0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 48 Wochen	Ruxolitinib (N= 144)	BAT (N = 72)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Vorbehandlung mit Hydroxycarbamid“			
Ja	N = 108	N = 49	NA
Ansprechen (n, %)	28 (25,9)	0 (0,0)	
Differenz vs. BAT [95%-KI]	25,9 [16,2 – 35,7]		
RR vs. BAT [95%-KI] [†]	26,1 [1,6 – 419,8]		
p-Wert [*]	< 0,0001		
Nein	N = 36	N = 23	
Ansprechen (n, %)	13 (36,1)	0 (0,0)	
Differenz vs. BAT [95%-KI]	36,1 [16,9 – 55,4]		
RR vs. BAT [95%-KI] [†]	17,5 [1,1 – 281,1]		
p-Wert [*]	0,0008		

[†] Bei Auftreten von Nullzellen wurde zur Berechnung des relativen Risikos ein Korrekturwert von 0,5 zu jeder Zelhäufigkeit der entsprechenden Vierfeldertafel addiert. Auf die Darstellung des Odds Ratios wurde aufgrund der geringen Zahl der Ereignisse in der Kontrollgruppe verzichtet.

^{*} Berechnung des p-Werts mit dem exakten Fisher-Test

[‡] Berechnung mittels logistischer Regression mit dem Milzvolumen zu Studienbeginn, der Behandlung und dem Subgruppenmerkmal als Kovariaten und dem Interaktionsterm Behandlung*Subgruppenmerkmal

^{**} Daten von mehreren Patienten aus Frankreich fehlen aufgrund von Datenschutzgründen

Abkürzungen: BAT, Best Available Therapy; KI, Konfidenzintervall; NA, not applicable; PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose; RR, Risk Ratio

Tabelle 4-44: Subgruppenanalysen für „Anteil der Patienten mit einer Reduktion des Milzvolumens um ≥ 35% nach 24 Wochen“ aus der Studie COMFORT-I

Reduktion des Milzvolumens um ≥ 35% nach 24 Wochen	Ruxolitinib (N= 155)	Placebo (N = 153)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Geschlecht“			
Männer	N = 79	N = 87	NA
Ansprechen (n, %)	20 (25,3)	1 (1,1)	
Differenz vs. Placebo [95%-KI]	24,2 [13,1 – 35,2]		
RR vs. Placebo [95%-KI] [†]	22,0 [3,0 – 160,4]		
p-Wert [*]	< 0,0001		
Frauen	N = 76	N = 65	NA
Ansprechen (n, %)	45 (59,2)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	59,2 [46,7 – 71,7]		
RR vs. Placebo [95%-KI] [†]	78,0 [4,9 – 1.241,8]		
p-Wert [*]	< 0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 24 Wochen	Ruxolitinib (N= 155)	Placebo (N = 153)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Alter“			
≤ 65 Jahre	N = 70	N = 52	NA
Ansprechen (n, %)	32 (45,7)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	45,7 [32,4 – 59,1]		
RR vs. Placebo [95%-KI] [†]	48,5 [3,0 – 774,6]		
p-Wert [*]	< 0,0001		
> 65 Jahre	N = 85	N = 101	
Ansprechen (n, %)	33 (38,8)	1 (1,0)	
Differenz vs. Placebo [95%-KI]	37,8 [26,2 – 49,5]		
RR vs. Placebo [95%-KI] [†]	39,2 [5,5 – 280,7]		
p-Wert [*]	< 0,0001		
Subgruppenanalyse „Hautfarbe“			
Weiß	N = 138	N = 138	NA
Ansprechen (n, %)	56 (40,6)	1 (0,7)	
Differenz vs. Placebo [95%-KI]	39,9 [30,8 – 48,9]		
RR vs. Placebo [95%-KI] [†]	56,0 [7,9 – 398,9]		
p-Wert [*]	< 0,0001		
Nicht-Weiß	N = 17	N = 14	
Ansprechen (n, %)	9 (52,9)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	52,9 [22,7 – 83,2]		
RR vs. Placebo [95%-KI] [†]	15,8 [1,0 – 250,2]		
p-Wert [*]	0,0013		
Subgruppenanalyse „Land“			
USA	N = 115	N = 121	NA
Ansprechen (n, %)	49 (42,6)	1 (0,8)	
Differenz vs. Placebo [95%-KI]	41,8 [31,8 – 51,8]		
RR vs. Placebo [95%-KI] [†]	51, 6 [7,2 – 367,3]		
p-Wert [*]	< 0,0001		
Kanada/Australien	N = 40	N = 32	
Ansprechen (n, %)	16 (40,0)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	40,0 [22,0 – 58,0]		
RR vs. Placebo [95%-KI] [†]	26,6 [1,7 – 426,4]		
p-Wert [*]	< 0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 24 Wochen	Ruxolitinib (N= 155)	Placebo (N = 153)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Myelofibrose-Typ“			
PMF	N = 70	N = 83	NA
Ansprechen (n, %)	27 (38,6)	1 (1,2)	
Differenz vs. Placebo [95%-KI]	37,4 [24,4 – 50,3]		
RR vs. Placebo [95%-KI] [†]	32,0 [4,5 – 229,7]		
p-Wert [*]	< 0,0001		
PPV-MF	N = 50	N = 47	
Ansprechen (n, %)	25 (50,0)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	50,0 [34,1 – 65,9]		
RR vs. Placebo [95%-KI] [†]	48,0 [3,0 – 766,8]		
p-Wert [*]	< 0,0001		
PET-MF	N = 35	N = 22	
Ansprechen (n, %)	13 (37,1)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	37,1 [17,4 – 56,9]		
RR vs. Placebo [95%-KI] [†]	17,3 [1,1 – 276,3]		
p-Wert [*]	0,0008		
Subgruppenanalyse „Milzvolumen zu Studienbeginn“			
≤ Median	N = 77	N = 77	NA
Ansprechen (n, %)	30 (39,0)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	39,0 [26,8 – 51,2]		
RR vs. Placebo [95%-KI] [†]	61,0 [3,8 – 980,1]		
p-Wert [*]	< 0,0001		
> Median	N = 78	N = 76	
Ansprechen (n, %)	35 (44,9)	1 (1,3)	
Differenz vs. Placebo [95%-KI]	43,6 [30,9 – 56,2]		
RR vs. Placebo [95%-KI] [†]	34,1 [4,8 – 242,7]		
p-Wert [*]	< 0,0001		
Subgruppenanalyse „Tastbare Milzlänge zu Studienbeginn“			
≤ 10 cm	N = 32	N = 26	NA
Ansprechen (n, %)	9 (28,1)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	28,1 [9,1 – 47,2]		
RR vs. Placebo [95%-KI] [†]	15,5 [0,9 – 255,1]		
p-Wert [*]	0,0029		
> 10cm	N = 123	N = 126	
Ansprechen (n, %)	56 (45,5)	1 (0,8)	
Differenz vs. Placebo [95%-KI]	44,7 [35,0 – 54,5]		
RR vs. Placebo [95%-KI] [†]	57,4 [8,1 – 408,0]		
p-Wert [*]	< 0,0001		

Reduktion des Milzvolumens um ≥ 35% nach 24 Wochen	Ruxolitinib (N= 155)	Placebo (N = 153)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Vorbehandlung mit Hydroxycarbamid“			
Ja	N = 52	N = 51	NA
Ansprechen (n, %)	29 (55,8)	1 (2,0)	
Differenz vs. Placebo [95%-KI]	53,8 [37,8 – 69,8]		
RR vs. Placebo [95%-KI] [†]	28,4 [4,0 – 201,1]		
p-Wert [*]	< 0,0001		
Nein	N = 103	N = 102	
Ansprechen (n, %)	36 (35,0)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	35,0 [24,8 – 45,1]		
RR vs. Placebo [95%-KI] [†]	72,3 [4,5 – 1.162,4]		
p-Wert [*]	< 0,0001		
Subgruppenanalyse „Vorliegen der JAK2 ^{V617F} -Mutation“			
Ja	N = 113	N = 122	NA
Ansprechen (n, %)	54 (47,8)	1 (0,8)	
Differenz vs. Placebo [95%-KI]	47,0 [36,8 – 57,2]		
RR vs. Placebo [95%-KI] [†]	58,3 [8,2 – 414,5]		
p-Wert [*]	< 0,0001		
Nein	N = 40	N = 27	
Ansprechen (n, %)	11 (27,5)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	27,5 [10,6 – 44,4]		
RR vs. Placebo [95%-KI] [†]	15,7 [1,0 – 255,8]		
p-Wert [*]	0,0021		
[†] Bei Auftreten von Nullzellen wurde zur Berechnung des relativen Risikos ein Korrekturwert von 0,5 zu jeder Zelhäufigkeit der entsprechenden Vierfeldertafel addiert. Auf die Darstellung des Odds Ratios wurde aufgrund der geringen Zahl der Ereignisse in der Kontrollgruppe verzichtet.			
[*] Berechnung des p-Werts mit dem exakten Fisher-Test			
[‡] Berechnung mittels logistischer Regression mit dem Milzvolumen zu Studienbeginn, der Behandlung und dem Subgruppenmerkmal als Kovariaten und dem Interaktionsterm Behandlung*Subgruppenmerkmal			
Abkürzungen: KI, Konfidenzintervall; NA, not applicable; PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose; RR, Risk Ratio			

Tabelle 4-45: Subgruppenanalysen für „Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen“ aus der Studie COMFORT-I

Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen	Ruxolitinib (N= 148)	Placebo (N= 152)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Geschlecht“			
Männer	N = 76	N = 87	0,8173
Ansprechen (n, %)	35 (46,1)	5 (5,7)	
Differenz vs. Placebo [95%-KI]	40,3 [26,8 – 53,8]		
RR vs. Placebo [95%-KI]	8,0 [3,3 – 19,4]		
OR vs. Placebo [95%-KI]	15,11 [5,41 – 42,22]		
Frauen	N = 72	N = 64	
Ansprechen (n, %)	33 (45,8)	3 (4,7)	
Differenz vs. Placebo [95%-KI]	41,1 [27,0 – 55,2]		
RR vs. Placebo [95%-KI]	9,8 [3,1 – 30,4]		
OR vs. Placebo [95%-KI]	17,25 [4,95 – 60,14]		
Subgruppenanalyse „Alter“			
≤ 65 Jahre	N = 69	N = 52	0,9815
Ansprechen (n, %)	33 (47,8)	3 (5,8)	
Differenz vs. Placebo [95%-KI]	42,1 [27,0 – 57,1]		
RR vs. Placebo [95%-KI]	8,3 [2,7 – 25,6]		
OR vs. Placebo [95%-KI]	14,98 [4,26 – 52,70]		
> 65 Jahre	N = 79	N = 100	
Ansprechen (n, %)	35 (44,3)	5 (5,0)	
Differenz vs. Placebo [95%-KI]	39,3 [26,4 – 52,2]		
RR vs. Placebo [95%-KI]	8,9 [3,6 – 21,6]		
OR vs. Placebo [95%-KI]	15,61 [5,68 – 42,87]		
Subgruppenanalyse „Hautfarbe“			
Weiß	N = 132	N = 137	NA
Ansprechen (n, %)	62 (47,0)	8 (5,8)	
Differenz vs. Placebo [95%-KI]	41,1 [31,0 – 51,2]		
RR vs. Placebo [95%-KI]	8,0 [4,0 – 16,1]		
OR vs. Placebo [95%-KI]	14,35 [6,50 – 31,70]		
p-Wert [*]	< 0,0001		
Nicht-Weiß	N = 16	N = 14	
Ansprechen (n, %)	6 (37,5)	0 (0,0)	
Differenz vs. Placebo [95%-KI]	37,5 [7,1 – 67,9]		
RR vs. Placebo [95%-KI] [†]	11,5 [0,7 – 187,0]		
p-Wert [*]	0,0185		

Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen	Ruxolitinib (N= 148)	Placebo (N = 152)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Land“			
USA	N = 111	N = 120	0,7519
Ansprechen (n, %)	53 (47,7)	7 (5,8)	
Differenz vs. Placebo [95%-KI]	41,9 [30,9 – 53,0]		
RR vs. Placebo [95%-KI]	8,2 [3,9 – 17,2]		
OR vs. Placebo [95%-KI]	14,90 [6,35 – 34,93]		
Kanada/Australien	N = 37	N = 32	
Ansprechen (n, %)	15 (40,5)	1 (3,1)	
Differenz vs. Placebo [95%-KI]	37,4 [17,6 – 57,3]		
RR vs. Placebo [95%-KI]	13,0 [1,8 – 92,9]		
OR vs. Placebo [95%-KI]	23,11 [2,79 – 191,28]		
Subgruppenanalyse „Risikogruppe“			
Hochrisiko	N = 83	N = 98	0,9337
Ansprechen (n, %)	38 (45,8)	5 (5,1)	
Differenz vs. Placebo [95%-KI]	40,7 [28,0 – 53,4]		
RR vs. Placebo [95%-KI]	9,0 [3,7 – 21,8]		
OR vs. Placebo [95%-KI]	16,30 [5,96 – 44,55]		
Intermediär-Risiko 2	N = 64	N = 53	
Ansprechen (n, %)	30 (46,9)	3 (5,7)	
Differenz vs. Placebo [95%-KI]	41,2 [25,8 – 56,7]		
RR vs. Placebo [95%-KI]	8,3 [2,7 – 25,6]		
OR vs. Placebo [95%-KI]	14,66 [4,14 – 51,91]		
Subgruppenanalyse „Myelofibrose-Typ“			
PMF	N = 66	N = 82	0,5280
Ansprechen (n, %)	27 (40,9)	2 (2,4)	
Differenz vs. Placebo [95%-KI]	38,5 [24,8 – 52,2]		
RR vs. Placebo [95%-KI]	16,8 [4,1 – 68,0]		
OR vs. Placebo [95%-KI]	29,20 [6,55 – 130,25]		
PPV-MF	N = 48	N = 47	
Ansprechen (n, %)	26 (54,2)	5 (10,6)	
Differenz vs. Placebo [95%-KI]	43,5 [24,8 – 62,3]		
RR vs. Placebo [95%-KI]	5,1 [2,1 – 12,1]		
OR vs. Placebo [95%-KI]	10,17 [3,39 – 30,55]		
PET-MF	N = 34	N = 22	
Ansprechen (n, %)	15 (44,1)	1 (4,5)	
Differenz vs. Placebo [95%-KI]	39,6 [17,0 – 62,1]		
RR vs. Placebo [95%-KI]	9,7 [1,4 – 68,4]		
OR vs. Placebo [95%-KI]	14,44 [1,71 – 122,10]		

Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen	Ruxolitinib (N= 148)	Placebo (N = 152)	Interaktionstest (p-Wert) [‡]
Subgruppenanalyse „Milzvolumen zu Studienbeginn“			
$\leq$ Median	N = 73	N = 76	0,6480
Ansprechen (n, %)	30 (41,1)	4 (5,3)	
Differenz vs. Placebo [95%-KI]	35,8 [22,1 – 49,5]		
RR vs. Placebo [95%-KI]	7,8 [2,9 – 21,1]		
OR vs. Placebo [95%-KI]	12,52 [4,11 – 38,12]		
> Median	N = 75	N = 76	
Ansprechen (n, %)	38 (50,7)	4 (5,3)	
Differenz vs. Placebo [95%-KI]	45,4 [31,7 – 59,1]		
RR vs. Placebo [95%-KI]	9,6 [3,6 – 25,6]		
OR vs. Placebo [95%-KI]	19,37 [6,34 – 59,17]		
Subgruppenanalyse „Tastbare Milzlänge zu Studienbeginn“			
≤ 10 cm	N = 30	N = 26	0,0123
Ansprechen (n, %)	11 (36,7)	4 (15,4)	
Differenz vs. Placebo [95%-KI]	21,3 [-4,4 – 47,0]		
RR vs. Placebo [95%-KI]	2,4 [0,9 – 6,6]		
OR vs. Placebo [95%-KI]	2,95 [0,75 – 11,57]		
p-Wert*	0,1290		
> 10cm	N = 118	N = 125	
Ansprechen (n, %)	57 (48,3)	4 (3,2)	
Differenz vs. Placebo [95%-KI]	45,1 [34,8 – 55,5]		
RR vs. Placebo [95%-KI]	15,1 [5,7 – 40,3]		
OR vs. Placebo [95%-KI]	28,41 [9,84 – 82,02]		
p-Wert*	< 0,0001		
Subgruppenanalyse „Vorbehandlung mit Hydroxycarbamid“			
Ja	N = 46	N = 51	0,9208
Ansprechen (n, %)	22 (47,8)	3 (5,9)	
Differenz vs. Placebo [95%-KI]	41,9 [24,1 – 59,8]		
RR vs. Placebo [95%-KI]	8,1 [2,6 – 25,4]		
OR vs. Placebo [95%-KI]	14,69 [3,99 – 54,06]		
Nein	N = 102	N = 101	
Ansprechen (n, %)	46 (45,1)	5 (5,0)	
Differenz vs. Placebo [95%-KI]	40,1 [28,6 – 51,7]		
RR vs. Placebo [95%-KI]	9,1 [3,8 – 22,0]		
OR vs. Placebo [95%-KI]	16,18 [6,04 – 43,33]		

Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen	Ruxolitinib (N= 148)	Placebo (N = 152)	Interaktionstest (p-Wert) [‡]	
Subgruppenanalyse „Vorliegen der JAK2 ^{V617F} -Mutation“				
Ja	N = 107	N = 121	0,1341	
Ansprechen (n, %)	56 (52,3)	6 (5,0)		
Differenz vs. Placebo [95%-KI]	47,4 [36,3 – 58,5]			
RR vs. Placebo [95%-KI]	10,6 [4,7 – 23,5]			
OR vs. Placebo [95%-KI]	21,60 [8,70 – 53,63]			
p-Wert [*]	< 0,0001			
Nein	N = 39	N = 27		
Ansprechen (n, %)	11 (28,2)	2 (7,4)		
Differenz vs. Placebo [95%-KI]	20,8 [0,4 – 41,2]			
RR vs. Placebo [95%-KI]	3,8 [0,9 – 15,8]			
OR vs. Placebo [95%-KI]	4,43 [0,87 – 22,61]			
p-Wert [*]	0,0575			
Subgruppenanalyse „Ruxolitinib-Anfangsdosis“				
≤ 15 mg bid	N = 51	N = 58	0,2877	
Ansprechen (n, %)	16 (31,4)	3 (5,2)		
Differenz vs. Placebo [95%-KI]	26,2 [10,4 – 42,0]			
RR vs. Placebo [95%-KI]	6,1 [1,9 – 16,9]			
OR vs. Placebo [95%-KI]	8,42 [2,28 – 31,04]			
20 mg bid	N = 97	N = 91		
Ansprechen (n, %)	52 (53,6)	5 (5,5)		
Differenz vs. Placebo [95%-KI]	48,1 [36,1 – 60,2]			
RR vs. Placebo [95%-KI]	9,8 [4,1 – 23,3]			
OR vs. Placebo [95%-KI]	21,17 [7,79 – 57,52]			
[†] Bei Auftreten von Nullzellen wurde zur Berechnung des relativen Risikos ein Korrekturwert von 0,5 zu jeder Zelhäufigkeit der entsprechenden Vierfeldertafel addiert. Auf die Darstellung des Odds Ratios wurde aufgrund der geringen Zahl der Ereignisse in der Kontrollgruppe verzichtet.				
[*] Berechnung des p-Werts mit dem exakten Fisher-Test				
[‡] Berechnung mittels logistischer Regression mit dem Milzvolumen zu Studienbeginn, der Behandlung und dem Subgruppenmerkmal als Kovariaten und dem Interaktionsterm Behandlung*Subgruppenmerkmal				
Abkürzungen: Bid, bis in die (lat., zweimal täglich); KI, Konfidenzintervall; NA, not applicable; OR, Odds Ratio; PET-MF, Post-Essentielle-Thrombozythämie-Myelofibrose; PMF, Primäre Myelofibrose; PPV-MF, Post-Polycythaemia-vera-Myelofibrose; RR, Risk Ratio				

Um Erkenntnisse  ber die Konsistenz des Therapieeffektes von Ruxolitinib zu erhalten, wurde jeweils f r den Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ “ (prim rer Endpunkt der beiden eingeschlossenen Studien) eine Analyse des Effekts in pr spezifizierten Subgruppen durchgef hrt (Tabelle 4-43, Tabelle 4-44) (23;31). So wurden Subgruppenanalysen zu folgenden Merkmalen dargestellt: Geschlecht (m nnlich,

weiblich), Alter (< 65 Jahre, $\geq$ 65 Jahre), Hautfarbe (Weiße, Nicht-Weiße), Myelofibrose-Typ (PMF, PPV-MF, PET-MF), Milzvolumen zu Studienbeginn (< Median, $\geq$ Median) und Vorbehandlung mit Hydroxycarbamid (ja, nein). In COMFORT-II wurde die Wirksamkeit der Ruxolitinib-Therapie außerdem zusätzlich getrennt nach Risikogruppe (intermediär-2, hoch) und der Thrombozytenzahl zu Studienbeginn (< 200.000/ μ l, $\geq$ 200.000/ μ l) ausgewertet, in COMFORT-I erfolgten weitere Subgruppenanalysen nach Land (USA, Kanada/Australien), der Milzlänge zu Studienbeginn (< 10 cm, $\geq$ 10 cm) und dem Vorliegen der *JAK2*^{V617F}-Mutation zu Studienbeginn (ja, nein) (23;31). Die in COMFORT-I ursprünglich geplante Analyse nach Ethnie und dem Einsatz von CYP3A4-Induktoren wurde aufgrund der geringen Fallzahlen nicht dargestellt. So waren lediglich 4 Patienten lateinamerikanischer Abstammung und nur ein Patient wurde neben der Studienmedikation auch mit einem Induktor des Isoenzym CYP3A4 behandelt (23).

Wie die im Rahmen der beiden Studien durchgeführten Subgruppenanalysen zeigten, war der Vorteil der Ruxolitinib-Therapie gegenüber der jeweiligen Kontrollbehandlung über alle vorab definierten Subkollektive hinweg konsistent (Tabelle 4-43, Tabelle 4-44) (23;31). Dies galt sowohl unabhängig vom Alter oder Geschlecht der Patienten als auch unabhängig von ihrer geographischen Herkunft oder ethnischen Zugehörigkeit. Auch krankheitsspezifische Faktoren wie der Myelofibrose-Subtyp, das Vorliegen der *JAK2*^{V617F}-Mutation sowie die Krankheitsschwere (dargestellt durch das Milzvolumen sowie die Risikogruppe zu Studienbeginn) hatten keinen Einfluss auf die Wirksamkeit der Ruxolitinib-Therapie. Ein Interaktionstest konnte für keine der Subgruppenanalysen angegeben werden, da eine Anpassung des logistischen Regressionsmodells in keinem Fall möglich war. Grund dafür ist das überwiegende Ausbleiben von Ereignissen im BAT-/Placebo-Arm. Die Überlegenheit von Ruxolitinib zeigte sich jedoch in allen Subgruppen mit signifikant höheren Ansprechraten im Vergleich zur jeweiligen Kontrollgruppe (23;31).

Wie bereits ausführlich in Abschnitt 4.2.5.5 diskutiert, sind Post-hoc-Analysen von nicht präspezifizierten Gruppen aus methodischen Gründen problematisch und somit nicht dazu geeignet, definitive Schlüsse über das Ausmaß des Behandlungseffektes in einer bestimmten Subgruppe im Rahmen der Nutzenbewertung zu ziehen (44-46;49). Aus diesem Grund wurden im vorliegenden Dossier keine gesonderten Post-hoc-Analysen durchgeführt. Da im Studienbericht der COMFORT-I-Studie für den sekundären Endpunkt „Anteil der Patienten mit einer Verbesserung des TSS um $\geq$ 50% nach 24 Wochen“ Post-hoc-Analysen nach verschiedenen Subgruppenmerkmalen berichtet wurden (23), werden diese der Vollständigkeit wegen im Folgenden dennoch dargestellt. Bei der Interpretation der Ergebnisse ist jedoch zu beachten, dass diese Post-hoc-Auswertungen rein explorativen Charakter aufweisen.

Es wurden Subgruppenanalysen zu den folgenden Merkmalen durchgeführt (Tabelle 4-45): Geschlecht (männlich, weiblich), Alter (< 65 Jahre, ≥ 65 Jahre), Hautfarbe (Weiße, Nicht-Weiße), Land (USA, Kanada/Australien), Myelofibrose-Typ (PMF, PPV-MF, PET-MF), Milzlänge zu Studienbeginn (< 10 cm, ≥ 10 cm), Milzvolumen zu Studienbeginn ($< \text{Median}$, $\geq \text{Median}$), Vorbehandlung mit Hydroxycarbamid (ja, nein), Vorliegen der $JAK2^{V617F}$ -Mutation zu Studienbeginn (ja, nein), Risikogruppe (intermediär-2, hoch) und Ruxolitinib-Anfangsdosis (≤ 15 mg bid, 20 mg bid) (23). Wie bei den Subgruppenanalysen zum primären Endpunkt zeigte sich der Vorteil der Ruxolitinib-Therapie über alle Subgruppen hinweg mit durchgehend höheren Ansprechraten gegenüber Placebo (Tabelle 4-45) (23). Aus den durchgeführten Interaktionstests ergab sich ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal „Milzlänge zu Studienbeginn“ ($p = 0,0123$) bzw. ein Hinweis auf eine Effektmodifikation durch das Merkmal „Vorliegen der $JAK2^{V617F}$ -Mutation“ ($p = 0,1341$) (Tabelle 4-45), allerdings waren auch hier die Ansprechraten in allen Subgruppen unter der Therapie mit Ruxolitinib deutlich höher als unter Placebo (23). Die Tatsache, dass bei Patienten mit einer Milzlänge von < 10 cm zu Studienbeginn bzw. Patienten ohne $JAK2^{V617F}$ -Mutation keine statistische Signifikanz erreicht wurde, ist sehr wahrscheinlich durch die deutlich geringere Patientenzahl in den beiden Subgruppen bedingt (56 vs. 243 bzw. 66 vs. 228). Darüber hinaus zeigte sich sowohl für Patienten mit einer Milzlänge von < 10 cm zu Studienbeginn als auch für Patienten ohne $JAK2^{V617F}$ -Mutation in Bezug auf den primären Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen“ ein statistisch signifikanter Vorteil für die Behandlung mit Ruxolitinib (Tabelle 4-44) (23). Eine eventuelle Effektmodifikation durch die Milzgröße konnte außerdem im Interaktionstest für das Merkmal „Milzvolumen zu Studienbeginn“ nicht bestätigt werden. Zusammenfassend lässt sich also festhalten, dass der Behandlungseffekt zugunsten von Ruxolitinib bei allen untersuchten Endpunkten über alle Subkollektive konsistent war.

Da es sich bei den beiden in die Nutzenbewertung eingeschlossenen Studien COMFORT-I und COMFORT-II um verschiedene Studientypen (placebo- bzw. aktiv-kontrolliert) mit unterschiedlichen Patientenpopulationen (Patienten ohne weitere Therapieoptionen bzw. Patienten, die lediglich nicht für eine allogene Stammzelltransplantation in Frage kamen) handelt, ist eine Zusammenfassung der Daten mittels Meta-Analyse nicht sinnvoll.

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Die Bewertung beruht auf den beiden randomisierten kontrollierten Studien COMFORT-I und COMFORT-II, in denen Wirksamkeit und Sicherheit von Ruxolitinib bei Patienten mit primärer oder sekundärer Myelofibrose untersucht wurden. In COMFORT-II wurde die Ruxolitinib-Behandlung mit der nach Ermessen des Prüfarztes besten verfügbaren Therapie (BAT) verglichen, wobei auch eine Nichtbehandlung möglich war. Den Patienten der COMFORT-I-Studienpopulation stand aufgrund von Refraktärität, Resistenz oder Intoleranz gegenüber anderen Therapien nur die „Nichtbehandlung“ als beste verfügbare Therapie zur Verfügung. Der Vergleich in dieser Studie erfolgte gegenüber Placebo, wobei das Studienprotokoll den Einsatz weiterer unterstützender Begleitmedikationen erlaubte. Damit entsprechen die Kontrollinterventionen beider Studien der zweckmäßigen Vergleichstherapie Best Supportive Care (BSC). Die Ergebnisse der beiden Studien lassen sich für die patientenrelevanten Endpunkte wie folgt zusammenfassen:

Reduktion der Milzgröße

Der Anteil der Patienten mit einer Milzvolumenreduktion um $\geq 35\%$ nach 24 Wochen (COMFORT-I) bzw. 48 Wochen (COMFORT-II) war der primäre Endpunkt der beiden Studien. Sowohl in **COMFORT-I** als auch in **COMFORT-II** wurde dieses Ziel unter Ruxolitinib signifikant häufiger erreicht als im jeweiligen Kontrollarm (41,9% vs. 0,7% bzw. 28,5% vs. 0%; $p < 0,0001$ für beide Vergleiche). Gegenüber dem Ausgangswert konnte in beiden Studien unter Ruxolitinib eine mediane Abnahme des Milzvolumens um circa 30% nachgewiesen werden, wohingegen es in den Komparator-Armen sogar zu einer Zunahme des Milzvolumens um median 8,5% kam. Eine klinisch relevante Reduktion des Milzvolumens um $\geq 35\%$ wurde bei der Mehrzahl der Patienten bereits nach 12 Wochen festgestellt und blieb über einen Zeitraum von drei Jahren erhalten. Die beiden Studien belegen damit eine drastische Reduktion der krankheitsbedingten Splenomegalie bei Patienten mit Myelofibrose unter der Therapie mit Ruxolitinib. Der Nutzen der Behandlung zeigte sich dabei in allen vorab definierten Subgruppen, unabhängig vom Alter der Patienten, ihrem Geschlecht, dem Myelofibrose-Subtyp, der Risikogruppe oder dem Milzvolumen zu Studienbeginn.

Linderung der Krankheitssymptomatik

Die Myelofibrose-assoziierten Symptome bzw. die Linderung der Krankheitssymptomatik unter der Therapie wurde in der Studie **COMFORT-I** mit Hilfe des MFSAF-Fragebogens erfasst, der die charakteristischen Symptome (Bauchbeschwerden, Schmerzen linksseitig unter den Rippen, Völlegefühl, Nachtschweiß, Juckreiz sowie Muskel- und Knochenschmerzen) in einem Summenscore, dem Total Symptome Score (TSS), zusammenfasst. Der Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen war unter der Therapie mit Ruxolitinib signifikant höher als unter Placebo (45,9% vs. 5,3%; $p < 0,0001$). In der Ruxolitinib-Gruppe konnte bei der Mehrzahl der Patienten eine deutliche Verbesserung des TSS nachgewiesen werden, wobei der Rückgang im Median bei 56,2% lag. Einige Patienten waren sogar vollständig symptomfrei. Unter der Placebo-Behandlung kam es hingegen beim überwiegenden Teil der Patienten zu einer mitunter dramatischen Verschlechterung der Symptomatik. Die Verbesserung der Myelofibrose-Symptome unter der Therapie mit Ruxolitinib zeigte sich dabei nicht nur beim Gesamtscore sondern auch in Bezug auf alle sechs untersuchten Einzelsymptome sowie beim ebenfalls abgefragten Item „Inaktivität“ ($p < 0,0001$ für alle Vergleiche).

Gesamtüberleben

Nach 3-jährigem Follow-up konnte außerdem eine deutliche Verbesserung des Gesamtüberlebens unter der Therapie mit Ruxolitinib nachgewiesen werden. So zeigte die ITT-Auswertung von **COMFORT-II** eine signifikante Reduktion des Mortalitätsrisikos im Ruxolitinib-Arm der Studie (HR = 0,48 [0,28-0,85]; $p = 0,009$). Auch die Ergebnisse der **COMFORT-I**-Studie weisen auf einen positiven Trend in Bezug auf ein verlängertes Gesamtüberleben hin und bestätigen somit die Überlegenheit der Ruxolitinib-Therapie (HR = 0,69 [0,46-1,03]; $p = 0,067$). Dieses Ergebnis ist insbesondere deshalb bemerkenswert, da der tatsächliche Überlebensvorteil durch Ruxolitinib aufgrund des in beiden Studien erlaubten Cross-overs eher unterschätzt wird.

Leukämiefreies Überleben

Auch in Bezug auf das leukämiefreie Überleben ergab die Auswertung der 3-Jahres-Analyse der beiden Studien einen klaren Vorteil für die Ruxolitinib-Therapie. So zeigte sich in der **COMFORT-II**-Studie eine signifikante Reduktion des Transformations- bzw. Mortalitätsrisikos um 49% (HR = 0,51 [0,30-0,88]; $p = 0,014$). In **COMFORT-I** konnte das Risiko, eine leukämische Transformation zu erleiden oder zu versterben, unter der Behandlung mit Ruxolitinib ebenfalls deutlich gesenkt werden (HR = 0,69 [0,46-1,01]; $p = 0,057$). Da die Mehrzahl der Ereignisse Todesfälle darstellten, ist die Überlegenheit gegenüber der

Vergleichsintervention vornehmlich auf den Überlebensvorteil unter der Ruxolitinib-Therapie zurückzuführen.

Progressionsfreies Überleben

Für das progressionsfreie Überleben, das in der Studie **COMFORT-II** als sekundärer Endpunkt erhoben wurde, ergaben sich bislang keine signifikanten Unterschiede zwischen den Behandlungsgruppen. Nach 144 Wochen betrug die Wahrscheinlichkeit des progressionsfreien Überlebens 33% in der Ruxolitinib-Gruppe, während sie unter der BAT-Kontrollbehandlung bei 24% lag (HR = 0,75 [0,50-1,12]; p = 0,158).

Veränderung des ECOG-Performance-Status

Um den Gesundheitszustand der Patienten im Verlauf der Behandlung zu erfassen, wurde in beiden Studien die Veränderung des ECOG-Performance-Status analysiert, der den physischen Zustand von Krebspatienten und deren Einschränkungen bei Aktivitäten des alltäglichen Lebens beschreibt. Sowohl in **COMFORT-I** als auch in **COMFORT-II** konnten unter der Therapie mit Ruxolitinib mehr Patienten ihren ECOG-Performance-Status erhalten oder verbessern als unter der jeweiligen Kontrollintervention. Umgekehrt war der Anteil der Patienten, bei denen im Studienverlauf eine Verschlechterung des ECOG-Status dokumentiert wurde, unter BAT bzw. Placebo deutlich höher.

Gesundheitsbezogene Lebensqualität

Die Auswertung standardisierter Fragebögen zur Erfassung der gesundheitsbezogenen Lebensqualität zeigte zudem, dass die Reduktion der Milzgröße sowie der konstitutionellen Symptome mit einer für die Patienten spürbaren Linderung der Erkrankung einherging.

In beiden Studien wurde der generische QLQ-C30-Fragebogen der EORTC eingesetzt, der spezifisch für die Erfassung der Lebensqualität von Krebspatienten entwickelt und validiert wurde. Im Verlauf der **COMFORT-II**-Studie konnte unter der Behandlung mit Ruxolitinib in Bezug auf nahezu alle Subskalen und Symptome des Fragebogens eine deutliche Verbesserung im Vergleich zu Studienbeginn erreicht werden. In der Kontrollgruppe wurde dagegen – wenn überhaupt – nur eine minimale Verbesserung der Werte beobachtet, meist kam es sogar zu einer Verschlechterung. Auch in der **COMFORT-I**-Studie hatten sich die QLQ-C30-Scores nach 24 Wochen in der Ruxolitinib-Gruppe sowohl bei der Globalbeurteilung als auch bei vier der fünf Funktionsskalen statistisch signifikant gegenüber Placebo verbessert (jeweils $p < 0,001$). In Bezug auf die Symptome konnte – mit Ausnahme

der Obstipation – ebenfalls eine signifikante Überlegenheit von Ruxolitinib gegenüber Placebo gezeigt werden (jeweils $p < 0,001$).

In der **COMFORT-II**-Studie wurde zusätzlich der krankheitsspezifische Fragebogen FACT-Lym eingesetzt. In der Ruxolitinib-Gruppe verbesserte sich im Vergleich zu Studienbeginn neben dem generischen FACT-G Total Score und den zusammengesetzten Skalen auch die krankheitsspezifische Subskala LymS, in der die für die Myelofibrose typischen Symptome erfasst werden. Im Gegensatz dazu blieben die Werte aller Skalen unter der BAT-Therapie unverändert oder verschlechterten sich sogar.

Verträglichkeit

Insgesamt bestätigen die beiden Studien **COMFORT-I** und **COMFORT-II** die gute Verträglichkeit von Ruxolitinib, die bereits aus der initialen Phase-I/II-Studie bekannt war. Obwohl die mediane Expositionszeit unter der Behandlung mit Ruxolitinib bereits zum Zeitpunkt der Primäranalyse in beiden Studien circa 10% über der der entsprechenden Kontrolltherapie lag, war die Häufigkeit unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse sowie von Ereignissen, die zum Abbruch der Therapie führten, zwischen den Behandlungsarmen vergleichbar. Lediglich der Anteil der Patienten mit einem Grad-3- oder Grad-4-Ereignis war in COMFORT-II in der Ruxolitinib-Gruppe höher als in der Kontrollgruppe. Umgekehrt war in COMFORT-I der Anteil der Patienten mit einem schwerwiegenden unerwünschten Ereignis unter Ruxolitinib niedriger als unter Placebo. Entsprechend dem Wirkmechanismus von Ruxolitinib als JAK-Inhibitor waren Anämien und Thrombozytopenien die häufigsten unerwünschten Ereignisse unter der Therapie mit Ruxolitinib. Sie waren jedoch in beiden Studien durch Dosisreduktionen oder symptomorientierte Supportivmaßnahmen gut kontrollierbar und führten nur selten zum Abbruch der Therapie. Auch nach längerem Follow-up ergaben sich keine neuen oder unerwarteten Sicherheitsprobleme. Im Gegenteil, die Inzidenz der unerwünschten Ereignisse ging nach längerer Behandlung mit Ruxolitinib sogar zurück.

Zusammenfassung

Ruxolitinib stellt somit eine wichtige neue Therapieoption dar, die zu einer klinisch relevanten, lang anhaltenden Reduktion der Splenomegalie, einer deutlichen Verminderung der belastenden Krankheitssymptomatik sowie einer Verbesserung der Lebensqualität führt und das Gesamtüberleben von Patienten mit Myelofibrose verlängert. Zusammenfassend ergibt sich aus den Ergebnissen randomisierter kontrollierter Studien für Patienten mit primärer oder sekundärer Myelofibrose damit ein therapeutisch bedeutender, beträchtlicher Zusatznutzen von Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie Best Supportive Care.

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 4-46: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

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

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

Tabelle 4-47: 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 4-48: 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 4-49: 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 4-50: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend.

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

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

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 4-51: 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 4-52: 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 4-53: 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 4-54: 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 4.3.1.3.2.***

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.

Die Nutzenbewertung von Ruxolitinib beruht auf den beiden großen, multizentrischen Phase-III-Studien COMFORT-I und COMFORT-II, in denen die Wirksamkeit von Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie BSC (Best Supportive Care) bei Patienten mit primärer oder sekundärer Myelofibrose untersucht wurde. Die Vollständigkeit des Studienpools wurde durch die geforderte systematische Recherche in den Datenbanken MEDLINE, EMBASE und der Cochrane-Library bestätigt (siehe Abschnitt 4.3.1.1.2). Auch die Suche in den Studienregistern ClinicalTrials.gov, EU Clinical Trials Register und Klinische Prüfungen PharmNet.Bund sowie über das ICTRP Search Portal ergab keine weiteren laufenden oder abgeschlossenen RCTs mit Ruxolitinib im vorliegenden Anwendungsgebiet (siehe Abschnitt 4.3.1.1.3).

Studienqualität

Beide Studien weisen eine hohe Ergebnissicherheit auf. Die Erzeugung der Randomisierungssequenz war jeweils adäquat, die Gruppenzuteilung erfolgte verdeckt und es lagen weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vor. Während in der COMFORT-I-Studie sowohl Patienten als auch Behandler verblindet waren, konnte eine Verblindung in der COMFORT-II-Studie aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (BAT) nicht durchgeführt werden. Das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da die patientenrelevanten Endpunkte objektiv erfasst werden konnten und die Ergebnisse damit stabil gegenüber möglichen Verzerrungen sind. So wurde der Endpunkt „Reduktion der Milzgröße“ durch klar definierte Parameter erfasst, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen unabhängigen, zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann. Auch die Endpunkte „Progressionsfreies Überleben“, „Leukämiefreies Überleben“ und „Gesamtüberleben“ sind unabhängig von subjektiven Einflussfaktoren. Insgesamt ist das Verzerrungspotenzial der Ergebnisse der Studien COMFORT-I und COMFORT-II daher als „niedrig“ einzustufen.

Validität der Endpunkte

Wie ausführlich in Abschnitt 4.2.5.2 diskutiert, sind alle Endpunkte in den herangezogenen Operationalisierungen unmittelbar patientenrelevant. Alle Auswertungen in Bezug auf den Rückgang der krankheitsbedingten Splenomegalie wurden in den relevanten Studien mit Hilfe standardisierter Verfahren erhoben und stellen somit valide Endpunkte dar. Da sie auf klar definierten Messparametern beruhen, ist zudem trotz der fehlenden Verblindung in der COMFORT-II-Studie von einer unverzerrten, validen Erhebung auszugehen. Dies gilt auch für die Endpunkte „Progressionsfreies Überleben“, „Leukämiefreies Überleben“ und „Gesamtüberleben“, die ebenfalls unabhängig von der subjektiven Beurteilung durch den Prüfarzt sind. Endpunkte zur gesundheitsbezogenen Lebensqualität sowie zur Erfassung der Krankheitssymptomatik wurden mit Hilfe validierter Fragebögen erhoben und sind ebenfalls in hohem Maße patientenrelevant. Während in der COMFORT-I-Studie sowohl Patienten als auch Behandler verblindet waren und somit von einer unverzerrten Bewertung auszugehen ist, kann eine eventuelle Verzerrung der Ergebnisse zur gesundheitsbezogenen Lebensqualität in der COMFORT-II-Studie aufgrund des offenen Studiendesigns nicht ausgeschlossen werden. Es ist jedoch zu beachten, dass die Verblindung einer Studie, in der die Wirksamkeit von Ruxolitinib im Vergleich zur besten verfügbaren Therapie untersucht werden soll, wegen des Einsatzes verschiedener Einzeltherapien im Kontrollarm mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten generell nicht durchführbar ist. Lediglich für den Teil der Zielpopulation, dem in der Kontrollgruppe nur die „Nichtbehandlung“ (Placebo) als beste verfügbare Therapie zur Verfügung steht, können Daten mit Hilfe einer verblindeten Studie erhoben werden, was im Rahmen der COMFORT-I-Studie realisiert wurde. Die vorgelegten Daten stellen somit die allgemein bestmögliche Evidenz dar, die zum Nachweis des Zusatznutzens von Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie im vorliegenden Indikationsgebiet generiert werden kann. Darüber hinaus zeigte sich eine hohe Konsistenz in Bezug auf das Ausmaß des klinischen Ansprechens, sowohl für die verschiedenen Wirksamkeitsendpunkte innerhalb der jeweiligen Studie als auch zwischen den beiden Studien. Insgesamt ist die Validität der herangezogenen Endpunkte deshalb als „hoch“ zu bewerten.

Evidenzstufe

Bei beiden in die vorliegende Nutzenbewertung eingeschlossenen Studien handelt es sich um randomisierte kontrollierte Vergleichsstudien, die somit dem höchsten Evidenzgrad entsprechen (Evidenzstufe Ib). Da sie die höchste Ergebnissicherheit aufweisen, sollen – sofern verfügbar – vorzugsweise randomisierte kontrollierte Studien in die Nutzenbewertung einbezogen werden (3).

Fazit

Unter Berücksichtigung der hohen Evidenzstufe und der guten Studienqualität sowie der Validität der herangezogenen Endpunkte ist die Aussagekraft der Nachweise zur Abschätzung des Zusatznutzens von Ruxolitinib im vorliegenden Anwendungsgebiet als „hoch“ einzuschätzen. Eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens auf Grundlage der erbrachten Nachweise ist deshalb mit hoher Ergebnissicherheit möglich.

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.

Vor Einführung von Ruxolitinib gab es keine zugelassenen medikamentösen Therapieoptionen für die Behandlung der Myelofibrose. Bei Patienten, die nicht für eine Stammzelltransplantation in Betracht kamen, versuchte man, einzelne Krankheitserscheinungen wie die Splenomegalie oder die konstitutionellen Symptome durch die palliative Therapie mit ver-

schiedenen „off-label“ eingesetzten Medikamenten zu lindern (36;78;79). Der Einsatz dieser Substanzen war jedoch häufig aufgrund gravierender Nebenwirkungen limitiert und brachte oft nur bei einem kleinen Teil der Patienten vorübergehende Besserung (12;13;36;79;80). Zur Behandlung massiver, therapieresistenter Splenomegalien blieb als letzte therapeutische Option häufig nur die Splenektomie bzw. Bestrahlung der Milz (12;13;81). Allerdings sind beide Eingriffe mit erheblichen Komplikationsraten behaftet, die zu einer therapiebedingten Mortalität von 5-10% führen (13;36;81). Vor der Zulassung von Ruxolitinib waren somit keine ausreichenden Therapiemöglichkeiten für die Behandlung der Myelofibrose vorhanden. Wie eine systematische Literaturrecherche ergab, liegt für keines der bis dahin eingesetzten Medikamente eine randomisierte kontrollierte Studie oder eine prospektive Kohortenstudie vor, die dessen Wirksamkeit bei der Therapie der Myelofibrose belegen könnte (82;83).

Ruxolitinib ist das erste und einzige Medikament, dessen Nutzen bei der Behandlung von Myelofibrose-Patienten nach den Standards der evidenzbasierten Medizin belegt ist und das infolgedessen eine arzneimittelrechtliche Zulassung im vorliegenden Anwendungsgebiet besitzt (1;84). Alleine aus dieser Tatsache folgt bereits die therapeutische Bedeutung von Ruxolitinib für die Behandlung von Patienten mit primärer oder sekundärer Myelofibrose.

Im Folgenden soll anhand der Kriterien der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) belegt werden, dass sich zum gegenwärtigen Stand der Erkenntnisse ein **beträchtlicher Zusatznutzen** für Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie Best Supportive Care (BSC) bei der Behandlung der krankheitsbedingten Splenomegalie sowie der Symptome bei Erwachsenen mit PMF, PPV-MF oder PET-MF ergibt.

Gemäß der AM-NutzenV liegt ein „beträchtlicher Zusatznutzen“ vor, wenn gegenüber der zweckmäßigen Vergleichstherapie eine bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens erreicht wird, insbesondere

- eine Abschwächung schwerwiegender Symptome
- eine moderate Verlängerung der Lebensdauer
- eine spürbare Linderung der Erkrankung
- eine relevante Vermeidung schwerwiegender Nebenwirkungen oder
- eine bedeutsame Vermeidung anderer Nebenwirkungen

Die Nutzenbewertung von Ruxolitinib beruht auf den beiden randomisierten kontrollierten Phase-III-Studien COMFORT-I und COMFORT-II, in denen die Wirksamkeit von Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie BSC bei Patienten mit primärer oder sekundärer Myelofibrose untersucht wurde. Wie bereits in Abschnitt 4.4.1 dargestellt, ist die Aussagekraft der Nachweise zur Abschätzung des Zusatznutzens von Ruxolitinib

aufgrund der hohen Evidenzstufe der beiden Studien (Evidenzstufe Ib), der guten Studienqualität sowie der Validität der herangezogenen Endpunkte als „hoch“ einzuschätzen. Eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens ist deshalb mit hoher Ergebnissicherheit möglich. Darüber hinaus können die Ergebnisse beider Studien uneingeschränkt auf den deutschen Versorgungskontext übertragen werden, da überwiegend Patienten weißer Hautfarbe eingeschlossen wurden und davon auszugehen ist, dass die medizinischen Versorgungsstandards der Länder, in denen sich Studienzentren befanden (Westeuropa, USA, Kanada und Australien), mit denen in Deutschland vergleichbar sind. Auf Grundlage der erbrachten Nachweise begründet sich der Zusatznutzen für Ruxolitinib dabei wie folgt:

Abschwächung schwerwiegender Symptome

Die Ergebnisse der beiden Studien COMFORT-I und COMFORT-II belegen für Patienten mit Myelofibrose unter der Behandlung mit Ruxolitinib gegenüber der zweckmäßigen Vergleichstherapie BSC erstmalig eine signifikante Abschwächung schwerwiegender Symptome der Erkrankung.

So konnte in beiden Studien eine ebenso schnelle wie lang anhaltende klinisch relevante **Reduktion des pathologisch vergrößerten Milzvolumens** nachgewiesen werden. Während die Ansprechraten in den jeweiligen Komparator-Armen nur bei 0,7% bzw. 0% lagen, erreichten 41,9% der Patienten (COMFORT-I, nach 24 Wochen) bzw. 28,5% der Patienten (COMFORT-II, nach 48 Wochen) unter Ruxolitinib eine Reduktion des Milzvolumens um $\geq 35\%$. Insgesamt wurde bei nahezu allen Patienten im Ruxolitinib-Arm der beiden Studien eine Milzvolumenreduktion unter den Ausgangswert zu Studienbeginn beobachtet. Dagegen kam es bei der Mehrzahl der Patienten unter der Kontrolltherapie sogar zu einer weiteren Zunahme der Milzgröße. Der Rückgang der krankheitsbedingten Splenomegalie unter der Therapie mit Ruxolitinib blieb dabei bei den meisten Patienten über einen Zeitraum von mindestens drei Jahren stabil. Wie die Ergebnisse von a priori geplanten Subgruppenanalysen zeigen, war der therapeutische Effekt dabei über alle Subgruppen hinweg nachweisbar.

Darüber hinaus belegen beide Studien eine drastische **Verminderung der belastenden Krankheitssymptomatik** unter der Behandlung mit Ruxolitinib. In COMFORT-I wurden die Myelofibrose-assoziierten Symptome mit Hilfe des MFSAF-Fragebogens erfasst, der die charakteristischen Symptome (Bauchbeschwerden, Schmerzen linksseitig unter den Rippen, Völlegefühl, Nachtschweiß, Juckreiz sowie Muskel- und Knochenschmerzen) als Summenscore (TSS) zusammenfasst. In der Ruxolitinib-Gruppe konnte bei der Mehrzahl der Patienten eine deutliche Verbesserung des TSS nachgewiesen werden, während es unter der Placebo-Behandlung meist zu einer mitunter dramatischen Verschlechterung der Symptomatik kam. Der Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen war

unter der Therapie mit Ruxolitinib signifikant höher als unter Placebo (45,9% vs. 5,3%; $p < 0,0001$). Auch in Bezug auf die Einzelsymptome sowie beim ebenfalls abgefragten Item „Inaktivität“ zeigte sich eine statistisch signifikante Verbesserung in der Ruxolitinib-Gruppe ($p < 0,0001$ für alle Vergleiche). Anders als in COMFORT-I wurden in COMFORT-II spezifische Symptome der Erkrankung wie Fieber, Nachtschweiß, Juckreiz und Gewichtsverlust mit Hilfe des FACT-Lym-Fragebogens erhoben und in der krankheitsspezifischen Subskala LymS zusammengefasst. Auch hier konnte für Patienten unter der Therapie mit Ruxolitinib eine deutliche Verbesserung gegenüber Studienbeginn nachgewiesen werden. Dabei wurde die klinisch relevante Differenz von mindestens 5,4 Punkten zwischen den Behandlungsgruppen zu allen Erhebungszeitpunkten überschritten.

Wie bereits ausführlich in Abschnitt 4.2.5.2 dargestellt, werden sowohl die Splenomegalie und die damit verbundenen Beschwerden als auch die konstitutionellen Symptome von den Betroffenen als äußerst belastend empfunden, was zu einer gravierenden Einschränkung der gesundheitsbezogenen Lebensqualität führt (38;85;86). Mit der Reduktion der Milzgröße und der Linderung der charakteristischen Myelofibrose-Symptomatik kommt es unter der Therapie mit Ruxolitinib somit zu einer Abschwächung schwerwiegender Symptome der Erkrankung. Wie die Analyse des ECOG-Performance-Status im Studienverlauf zeigte, ging die Linderung dieser Symptome durch die Behandlung mit Ruxolitinib zugleich mit einer Verbesserung des allgemeinen Gesundheitszustandes der Patienten einher. Gemäß den Kriterien der AM-NutzenV ist der Zusatznutzen von Ruxolitinib gegenüber Best Supportive Care damit in der Nutzendimension „Morbidity“ als **beträchtlich** einzustufen.

Moderate Verlängerung der Lebensdauer

Darüber hinaus konnte unter der Therapie mit Ruxolitinib im Vergleich zu Best Supportive Care eine **Verbesserung des Gesamtüberlebens** erreicht werden. So belegt die ITT-Auswertung von COMFORT-II eine signifikante Reduktion des Mortalitätsrisikos im Ruxolitinib-Arm der Studie (HR = 0,48 [0,28-0,85]; $p = 0,009$). Auch die Ergebnisse der COMFORT-I-Studie bestätigen einen Überlebensvorteil für Patienten unter Ruxolitinib (HR = 0,69 [0,46-1,03]; $p = 0,067$). Dieses Ergebnis ist insbesondere bemerkenswert, da der Überlebensvorteil durch Ruxolitinib aufgrund des in beiden Studien erlaubten Cross-overs eher unterschätzt wird. Da in beiden Studien bislang weniger als 50% der Patienten verstorben sind, lassen sich keine Aussagen zum medianen Gesamtüberleben bzw. zu dessen Verlängerung unter der Therapie mit Ruxolitinib treffen. Aufgrund der signifikanten Senkung des Mortalitätsrisikos unter Ruxolitinib liegt jedoch eine mindestens moderate Verlängerung der Lebensdauer vor. Gemäß den Vorgaben der AM-NutzenV ist der Zusatznutzen von Ruxolitinib gegenüber BSC daher auch in der Nutzendimension „Mortality“ **beträchtlich**.

Spürbare Linderung der Erkrankung

Die Erfassung der gesundheitsbezogenen Lebensqualität mit Hilfe standardisierter, validierter Fragebögen zeigte außerdem, dass Patienten unter Ruxolitinib über alle Bereiche hinweg eine **Verbesserung des subjektiven Befindens** erfuhren. So konnte in beiden Studien in der Ruxolitinib-Gruppe in Bezug auf nahezu alle Subskalen und Symptome des QLQ-C30-Fragebogens der EORTC eine deutliche Verbesserung im Vergleich zu Studienbeginn erreicht werden. In der jeweiligen Kontrollgruppe kam es dagegen – wenn überhaupt – nur zu einer minimalen Verbesserung der Scores, meist wurde sogar eine Verschlechterung der Werte beobachtet. In der COMFORT-II-Studie wurde zusätzlich der krankheitsspezifische Fragebogen FACT-Lym eingesetzt. Auch hier erfuhren Patienten unter der Therapie mit Ruxolitinib über alle Bereiche hinweg eine deutliche Verbesserung gegenüber Studienbeginn, wohingegen die Werte unter der BAT-Kontrollbehandlung unverändert blieben oder sich teilweise sogar verschlechterten. Mit der Verbesserung der gesundheitsbezogenen Lebensqualität kommt es unter der Therapie mit Ruxolitinib somit zu einer für die Patientinnen und Patienten spürbaren Linderung der Erkrankung. Entsprechend den Kriterien der AM-NutzenV ergibt sich gegenüber der Vergleichstherapie Best Supportive Care somit auch hinsichtlich der Nutzendimension „Lebensqualität“ ein **beträchtlicher Zusatznutzen** von Ruxolitinib.

Vermeidung von Nebenwirkungen

Wie die Ergebnisse der beiden Studien COMFORT-I und COMFORT-II zeigen, wird das Ausmaß des Zusatznutzens nicht durch ein ungünstiges Sicherheitsprofil gegenüber BSC als der zweckmäßigen Vergleichstherapie eingeschränkt. Aus den verfügbaren Ergebnissen ergeben sich keine Anhaltspunkte für einen höheren Schaden durch Ruxolitinib. Insgesamt bestätigte die Sicherheitsanalyse der beiden Studien die **gute Verträglichkeit** von Ruxolitinib, die bereits aus der initialen Phase-I/II-Studie bekannt war. Entsprechend dem Wirkmechanismus von Ruxolitinib als JAK-Inhibitor waren Anämien und Thrombozytopenien die häufigsten Nebenwirkungen. Sie waren jedoch in beiden Studien gut kontrollierbar und führten nur selten zum Abbruch der Therapie. Auch nach längerem Follow-up ergaben sich keine neuen oder unerwarteten Sicherheitsprobleme. Im Gegenteil, die Inzidenz der unerwünschten Ereignisse ging im Verlauf der Behandlung sogar zurück. Dementsprechend stellte der Gemeinsame Bundesausschuss in den tragenden Gründen zum Beschluss vom 7. März 2013 über die erste Nutzenbewertung von Ruxolitinib fest, dass die Nebenwirkungen als für die Patienten bedeutend, aber überwiegend als kontrollierbar und behandelbar einzustufen sind und deshalb **nicht zu einer Herabstufung des Ausmaßes des Zusatznutzens** führen (19).

Fazit

Der Zusatznutzen von Ruxolitinib gegenüber Best Supportive Care ergibt sich aus dem klinisch bedeutenden, lang anhaltenden Rückgang der pathologischen Milzvergrößerung, der drastischen Reduktion der belastenden Krankheitssymptomatik hinsichtlich aller untersuchten Symptome, der deutlichen Senkung des Mortalitätsrisikos sowie der nachhaltigen Verbesserung des allgemeinen Gesundheitszustandes und der gesundheitsbezogenen Lebensqualität. Insbesondere auch unter Berücksichtigung des Schweregrades der Erkrankung und des Fehlens alternativer Therapiemöglichkeiten ist der Zusatznutzen von Ruxolitinib damit für alle Patienten im zugelassenen Anwendungsgebiet als **beträchtlich** einzustufen.

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 4-55: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit krankheitsbedingter Splenomegalie oder Symptomen bei primärer Myelofibrose (auch bekannt als chronische idiopathische Myelofibrose), Post-Polycythaemia-vera-Myelofibrose oder Post-Essentieller-Thrombozythämie-Myelofibrose	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 nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

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

Nicht 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,

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

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:

- 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.

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.

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).

COMFORT-II (CINC424A2352)

Registereinträge:

- ClinicalTrials.gov: NCT00934544
- EU Clinical Trials Register: EudraCT 2009-009858-24
- ICTRP Search Portal: NCT00934544

Studienberichte:

- Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Oct 17.
- Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (Addendum1). 2011 Nov 8.
- Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jun 19.

Publikationen:

- Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. N Engl J Med 2012 Mar 1;366(9):787-98.
- Cervantes F, Vannucchi AM, Kiladjian JJ, Al-Ali HK, Sirulnik A, Stalbovskaya V, et al. Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3

study comparing ruxolitinib with best available therapy for myelofibrosis. Blood 2013 Dec 12;122(25):4047-53.

- Harrison CN, Mesa RA, Kiladjian JJ, Al-Ali HK, Gisslinger H, Knoops L, et al. Health-related quality of life and symptoms in patients with myelofibrosis treated with ruxolitinib versus best available therapy. Br J Haematol 2013 Jul;162(2):229-39.
- European Medicines Agency. European Public Assessment Report Jakavi. Procedure No.: EMEA/H/C/002464. 2012 Oct 4.

COMFORT-I (INCB 18424-351)

Registereinträge:

- ClinicalTrials.gov: NCT00952289
- ICTRP Search Portal: NCT00952289

Studienberichte:

- Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Apr 27.
- Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 1). 2011 May 4.
- Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 2). 2011 Oct 31.
- Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jul 24.

Publikationen:

- Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *N Engl J Med* 2012 Mar 1;366(9):799-807.
- Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Efficacy, safety and survival with ruxolitinib treatment in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. *Haematologica* 2013 Dec;98(12):1865-71.
- Mesa RA, Gotlib J, Gupta V, Catalano JV, Deininger MW, Shields AL, et al. Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: a randomized, double-blind, placebo-controlled trial. *J Clin Oncol* 2013 Apr 1;31(10):1285-92.
- Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. The clinical benefit of ruxolitinib across patient subgroups: analysis of a placebo-controlled, Phase III study in patients with myelofibrosis. *Br J Haematol* 2013 May;161(4):508-16.
- Verstovsek S, Gotlib J, Gupta V, Atallah E, Mascarenhas J, Quintas-Cardama A, et al. Management of cytopenias in patients with myelofibrosis treated with ruxolitinib and effect of dose modifications on efficacy outcomes. *Onco Targets Ther* 2013;7:13-21.
- Mesa RA, Shields A, Hare T, Erickson-Viitanen S, Sun W, Sarlis NJ, et al. Progressive burden of myelofibrosis in untreated patients: assessment of patient-reported outcomes in patients randomized to placebo in the COMFORT-I study. *Leuk Res* 2013 Aug;37(8):911-6.
- European Medicines Agency. European Public Assessment Report Jakavi. Procedure No.: EMEA/H/C/002464. 2012 Oct 4.

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) Novartis Pharma GmbH. Fachinformation Jakavi. 2013 Nov.
- (2) Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß §8 Abs. 1 AM-NutzenV. Beratungsanforderung 2013-B-128 Ruxolitinib. 2014 Mar 30.
- (3) Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Version 4.1. www.iqwig.de. 2013 Nov 28.
- (4) Wong SS, 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 Oct;94(4):451-5.
- (5) European Medicines Agency. Committee for Medicinal Products for Human Use. Guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/205/95/Rev.4. 2012 Dec 13.
- (6) Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Rapid Report: Aussagekraft von Surrogatendpunkten in der Onkologie. www.iqwig.de. 2011 Nov 21.
- (7) Gangat N, Caramazza D, Vaidya R, George G, Begna K, Schwager S, et al. DIPSS plus: a refined Dynamic International Prognostic Scoring System for primary myelofibrosis that incorporates prognostic information from karyotype, platelet count, and transfusion status. J Clin Oncol 2011 Feb 1;29(4):392-7.
- (8) Cervantes F, Dupriez B, Pereira A, Passamonti F, Reilly JT, Morra E, et al. New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. Blood 2009 Mar 26;113(13):2895-901.
- (9) Barosi G, Mesa RA, Thiele J, Cervantes F, Campbell PJ, Verstovsek S, et al. Proposed criteria for the diagnosis of post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the International Working Group for Myelofibrosis Research and Treatment. Leukemia 2008 Feb;22(2):437-8.
- (10) Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 2009 Jul 30;114(5):937-51.
- (11) Cervantes F. How I treat splenomegaly in myelofibrosis. Blood Cancer J 2011 Oct;1(10):e37.

- (12) Mesa RA. How I treat symptomatic splenomegaly in patients with myelofibrosis. *Blood* 2009 May 28;113(22):5394-400.
- (13) Vannucchi AM. Management of myelofibrosis. *Hematology Am Soc Hematol Educ Program* 2011;2011:222-30.
- (14) Mitra D, Kaye JA, Piecoro LT, Brown J, Reith K, Mughal TI, et al. Symptom burden and splenomegaly in patients with myelofibrosis in the United States: a retrospective medical record review. *Cancer Med* 2013 Dec;2(6):889-98.
- (15) Mesa RA, Cortes JE, Cervantes F, Milligan D, Masszi T, Mishchenko E, et al. Symptom Burden and Health-Related Quality of Life (HRQoL) in Patients with Myelofibrosis (MF) treated with Fedratinib (SAR302503) in a Phase III Study (JAKARTA). *Blood* 2013 Nov 15;122:4061.
- (16) Mesa RA, Gotlib J, Gupta V, Catalano JV, Deininger MW, Shields AL, et al. Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: a randomized, double-blind, placebo-controlled trial. *J Clin Oncol* 2013 Apr 1;31(10):1285-92.
- (17) Verstovsek S, Kantarjian HM, Estrov Z, Cortes JE, Thomas DA, Kadia T, et al. Long-term outcomes of 107 patients with myelofibrosis receiving JAK1/JAK2 inhibitor ruxolitinib: survival advantage in comparison to matched historical controls. *Blood* 2012 Aug 9;120(6):1202-9.
- (18) Gemeinsamer Bundesausschuss. 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 - Ruxolitinib. Verfügbar unter: http://www.g-ba.de/downloads/39-261-1665/2013-03-07_AM-RL-XII_Ruxolitinib_BAnz.pdf. 2013 Mar 7.
- (19) Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Ruxolitinib. Verfügbar unter: http://www.g-ba.de/downloads/40-268-2215/2013-03-07_AM-RL-XII_Ruxolitinib_TrG.pdf. 2013 Mar 7.
- (20) European Medicines Agency. European Public Assessment Report Jakavi. Procedure No.: EMEA/H/C/002464. 2012 Oct 4.
- (21) Mesa RA, Niblack J, Wadleigh M, Verstovsek S, Camoriano J, Barnes S, et al. The burden of fatigue and quality of life in myeloproliferative disorders (MPDs): an international Internet-based survey of 1179 MPD patients. *Cancer* 2007 Jan 1;109(1):68-76.
- (22) Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *N Engl J Med* 2012 Mar 1;366(9):799-807.

- (23) Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Apr 27.
- (24) Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jul 24.
- (25) Mesa RA, Schwager S, Radia D, Cheville A, Hussein K, Niblack J, et al. The Myelofibrosis Symptom Assessment Form (MFSAF): an evidence-based brief inventory to measure quality of life and symptomatic response to treatment in myelofibrosis. *Leuk Res* 2009 Sep;33(9):1199-203.
- (26) US Department of Health and Human Services, Food and Drug Administration (FDA). Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
- (27) Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993 Mar 3;85(5):365-76.
- (28) Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982 Dec;5(6):649-55.
- (29) Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. *Qual Life Res* 2002 May;11(3):207-21.
- (30) Incyte Corporation. Evidenzdossier "Myelofibrosis Symptom Assessment Form" (MFSAF v2.0), Evidence Dossier Ruxolitinib (INCB018424), Patient Reported Outcome Measurement Strategy. 2011 Mar 17.
- (31) Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Oct 17.
- (32) Cervantes F, Tassies D, Salgado C, Rovira M, Pereira A, Rozman C. Acute transformation in nonleukemic chronic myeloproliferative disorders: actuarial probability and main characteristics in a series of 218 patients. *Acta Haematol* 1991;85(3):124-7.

- (33) Dupriez B, Morel P, Demory JL, Lai JL, Simon M, Plantier I, et al. Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood* 1996 Aug 1;88(3):1013-8.
- (34) Mesa RA, Li CY, Ketterling RP, Schroeder GS, Knudson RA, Tefferi A. Leukemic transformation in myelofibrosis with myeloid metaplasia: a single-institution experience with 91 cases. *Blood* 2005 Feb 1;105(3):973-7.
- (35) Tefferi A. Pathogenesis of myelofibrosis with myeloid metaplasia. *J Clin Oncol* 2005 Nov 20;23(33):8520-30.
- (36) Barbui T, Barosi G, Birgegard G, Cervantes F, Finazzi G, Griesshammer M, et al. Philadelphia-negative classical myeloproliferative neoplasms: critical concepts and management recommendations from European LeukemiaNet. *J Clin Oncol* 2011 Feb 20;29(6):761-70.
- (37) Lowenberg B, Downing JR, Burnett A. Acute myeloid leukemia. *N Engl J Med* 1999 Sep 30;341(14):1051-62.
- (38) Gemeinsamer Bundesausschuss. Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO. Wirkstoff: Ruxolitinib. Verfügbar unter: <http://www.g-ba.de/informationen/nutzenbewertung/38/#tab/nutzenbewertung>. 2012 Dec 17.
- (39) Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998 Jan;16(1):139-44.
- (40) Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993 Mar;11(3):570-9.
- (41) Hlubocky FJ, Webster K, Cashy J, Beaumont J, Cella D. 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). *Lymphoma*, vol. 2013, Article ID 147176, 9 pages. doi:10.1155/2013/147176; 2013.
- (42) Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes* 2003;1:79.
- (43) Carter GC, Liepa AM, Zimmermann AH, Morschhauser F. Validation of the Functional Assessment of Cancer Therapy - Lymphoma (FACT-Lym) in patients with relapsed/refractory mantle cell lymphoma. *ASH Annual Meeting Abstracts*; 2008 Dec 7.
- (44) Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 2005 Jan 8;365(9454):176-86.

- (45) Bestehorn M, Tunder R. AMNOG - Erste Erfahrungen und mögliche Auswirkungen auf die Klinische Forschung. *Pharmacocon Ger Res Artic*, doi: 10.1007/s40275-013-0002-1; 2013 Aug 14.
- (46) Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000 Mar 25;355(9209):1064-9.
- (47) Cook DI, GebSKI VJ, Keech AC. Subgroup analysis in clinical trials. *Med J Aust* 2004 Mar 15;180(6):289-91.
- (48) Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992 Jan 1;116(1):78-84.
- (49) Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ* 2012;344:e1553.
- (50) ClinicalTrials.gov. Controlled Myelofibrosis Study With Oral Janus-associated Kinase (JAK) Inhibitor Treatment-II: The COMFORT-II Trial. <http://clinicaltrials.gov/show/NCT00934544>. 2014 Jan 16.
- (51) EU Clinical Trials Register. COntrolled MyeloFibrosis Study with ORal JAK Inhibitor Treatment-II. <https://www.clinicaltrialsregister.eu/ctr-search/trial/2009-009858-24/DE>. 2014 Jan 10.
- (52) International Clinical Trials Registry Platform Search Portal. Controlled Myelofibrosis Study With Oral Janus-associated Kinase (JAK) Inhibitor Treatment-II: The COMFORT-II Trial. <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00934544>. 2014 Jan 10.
- (53) ClinicalTrials.gov. COntrolled MyeloFibrosis Study With ORal JAK Inhibitor Treatment: The COMFORT-I Trial. <http://clinicaltrials.gov/show/NCT00952289>. 2014 Jan 16.
- (54) International Clinical Trials Registry Platform Search Portal. COntrolled MyeloFibrosis Study With ORal JAK Inhibitor Treatment: The COMFORT-I Trial. <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00952289>. 2014 Jan 10.
- (55) Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 1). 2011 May 4.
- (56) Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-

Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 2). 2011 Oct 31.

- (57) Mesa RA, Shields A, Hare T, Erickson-Viitanen S, Sun W, Sarlis NJ, et al. Progressive burden of myelofibrosis in untreated patients: assessment of patient-reported outcomes in patients randomized to placebo in the COMFORT-I study. *Leuk Res* 2013 Aug;37(8):911-6.
- (58) Verstovsek S, Gotlib J, Gupta V, Atallah E, Mascarenhas J, Quintas-Cardama A, et al. Management of cytopenias in patients with myelofibrosis treated with ruxolitinib and effect of dose modifications on efficacy outcomes. *Onco Targets Ther* 2013;7:13-21.
- (59) Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Efficacy, safety and survival with ruxolitinib treatment in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. *Haematologica* 2013 Dec;98(12):1865-71.
- (60) Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. The clinical benefit of ruxolitinib across patient subgroups: analysis of a placebo-controlled, Phase III study in patients with myelofibrosis. *Br J Haematol* 2013 May;161(4):508-16.
- (61) Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (Addendum1). 2011 Nov 8.
- (62) Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jun 19.
- (63) Cervantes F, Vannucchi AM, Kiladjian JJ, Al-Ali HK, Sirulnik A, Stalbovskaya V, et al. Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3 study comparing ruxolitinib with best available therapy for myelofibrosis. *Blood* 2013 Dec 12;122(25):4047-53.
- (64) Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. *N Engl J Med* 2012 Mar 1;366(9):787-98.
- (65) Harrison CN, Mesa RA, Kiladjian JJ, Al-Ali HK, Gisslinger H, Knoops L, et al. Health-related quality of life and symptoms in patients with myelofibrosis treated with ruxolitinib versus best available therapy. *Br J Haematol* 2013 Jul;162(2):229-39.

- (66) Prassopoulos P, Daskalogiannaki M, Raissaki M, Hatjidakis A, Gourtsoyiannis N. Determination of normal splenic volume on computed tomography in relation to age, gender and body habitus. *Eur Radiol* 1997;7(2):246-8.
- (67) Mazonakis M, Damilakis J, Maris T, Prassopoulos P, Gourtsoyiannis N. Estimation of spleen volume using MR imaging and a random marking technique. *Eur Radiol* 2000;10(12):1899-903.
- (68) Harris A, Kamishima T, Hao HY, Kato F, Omatsu T, Onodera Y, et al. Splenic volume measurements on computed tomography utilizing automatically contouring software and its relationship with age, gender, and anthropometric parameters. *Eur J Radiol* 2010 Jul;75(1):e97-101.
- (69) Aidelsburger P, Wasem J. Kosten-Nutzen-Bewertungen von onkologischen Therapien. Gutachten für die Deutsche Krebsgesellschaft e.V. 2008 Jul.
- (70) Harrison C, Niederwieser D, Vannucchi AM, Kiladjian JJ, Barbui T, Gisslinger H, et al. Results from a 3.5-Year Update of COMFORT-II, A Phase 3 Study Comparing Ruxolitinib (RUX) With Best Available Therapy (BAT) for the Treatment of Myelofibrosis. Abstract eingereicht zur Präsentation auf dem Kongress der European Hematology Association (EHA) im Juni 2014.
- (71) Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Long-term outcomes of Ruxolotinib therapy in patients with myelofibrosis: 3-year update from COMFORT-I. *ASH Annual Meeting Abstracts*; 2013 Dec.
- (72) Cervantes F, Kiladjian JJ, Niederwieser D, Sirulnik A, Stalbovskaya V, McQuitty M, et al. Long-Term Efficacy, Safety, and Survival Findings From COMFORT-II, a Phase 3 Study Comparing Ruxolitinib With Best Available Therapy for the Treatment of Myelofibrosis. *ASH Annual Meeting Abstracts*; 2012 Dec.
- (73) Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Long-Term Outcome of Ruxolitinib Treatment in Patients With Myelofibrosis: Durable Reductions in Spleen Volume, Improvements in Quality of Life, and Overall Survival Advantage in COMFORT-I. *ASH Annual Meeting Abstracts*; 2012 Dec.
- (74) Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Eribulin - Nutzenbewertung gemäß § 35a SGB V. www.g-ba.de. 2012 Jan 30.
- (75) Harrison C, Mesa R, Ross D, Mead A, Keohane C, Gotlib J, et al. Practical management of patients with myelofibrosis receiving ruxolitinib. *Expert Rev Hematol* 2013 Oct;6(5):511-23.
- (76) Mughal TI, Vaddi K, Sarlis NJ, Verstovsek S. Myelofibrosis-associated complications: pathogenesis, clinical manifestations, and effects on outcomes. *Int J Gen Med* 2014;7:89-101.
- (77) Verstovsek S, Kantarjian H, Mesa RA, Pardananani AD, Cortes-Franco J, Thomas DA, et al. Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *N Engl J Med* 2010 Sep 16;363(12):1117-27.

- (78) Pardanani A, Vannucchi AM, Passamonti F, Cervantes F, Barbui T, Tefferi A. JAK inhibitor therapy for myelofibrosis: critical assessment of value and limitations. *Leukemia* 2011 Feb;25(2):218-25.
- (79) Tefferi A, Vainchenker W. Myeloproliferative neoplasms: molecular pathophysiology, essential clinical understanding, and treatment strategies. *J Clin Oncol* 2011 Feb 10;29(5):573-82.
- (80) Abdel-Wahab OI, Levine RL. Primary myelofibrosis: update on definition, pathogenesis, and treatment. *Annu Rev Med* 2009;60:233-45.
- (81) Cervantes F, Martinez-Trillos A. Myelofibrosis: an update on current pharmacotherapy and future directions. *Expert Opin Pharmacother* 2013 May;14(7):873-84.
- (82) Novartis Pharma GmbH: Dossier zur Nutzenbewertung gemäß § 35a SGB V, Ruxolitinib (JAKAVI®), Modul 4. 2012 Sep 7.
- (83) Boston Healthcare Associates International GmbH. Systematisches Review zur Behandlung der Splenomegalie und Symptomatik bei Myelofibrose. Manuskript in Vorbereitung. 2014 Apr 4.
- (84) Deutsches Institut für Medizinische Dokumentation und Information (DIMDI): Datenbankabfrage (AJ29) am 21.03.2014.
- (85) Tefferi A. Essential thrombocythemia, polycythemia vera, and myelofibrosis: current management and the prospect of targeted therapy. *Am J Hematol* 2008 Jun;83(6):491-7.
- (86) Tefferi A. Primary myelofibrosis: 2012 update on diagnosis, risk stratification, and management. *Am J Hematol* 2011 Dec;86(12):1017-26.

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

Im Folgenden sind die Suchstrategien für Abschnitt 4.3.1.1.2 „Studien aus der bibliographischen Literaturrecherche“ dokumentiert.

Tabelle 4-56: Bibliographische Literaturrecherche in MEDLINE

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations /MEDLINE Daily Update	
Suchoberfläche	Ovid	
Datum der Suche	24.02.2014	
Zeitsegment	1946 to February Week 2 2014 / February 21, 2014 / February 21, 2014	
Suchfilter	Filter für randomisierte kontrollierte Studien nach 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.	
#	Suchbegriffe	Ergebnis
1	ruxolitinib*.mp.	147
2	jakafi*.mp.	15
3	jakavi*.mp.	3
4	incb18424*.mp.	3
5	incb-18424*.mp.	4
6	"incb 18424".mp.	4
7	incb018424*.mp.	126
8	incb-018424*.mp.	2
9	"incb 018424".mp.	2
10	inc424*.mp.	4
11	inc-424*.mp.	0
12	"inc 424".mp.	0
13	1092939-17-7.rm.	0
14	or/1-13	201
15	exp Primary Myelofibrosis/	4775
16	myelofibros*.mp.	6272
17	osteomyelofibros*.mp.	138
18	myelo-fibros*.mp.	5
19	myelosclero*.mp.	195
20	(myeloid* adj2 metaplasia*).mp.	1153
21	(bone adj2 marrow* adj2 fibros*).mp.	524
22	or/15-21	6720
23	randomized controlled trial.pt.	363619

24	randomized.mp.	546582
25	placebo.mp.	155035
26	or/23-25	594912
27	and/14,22,26	34
28	remove duplicates from 27	34

Tabelle 4-57: Bibliographische Literaturrecherche in EMBASE

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	24.02.2014	
Zeitsegment	1974 to 2014 February 21	
Suchfilter	Filter für randomisierte kontrollierte Studien nach 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.	
#	Suchbegriffe	Ergebnis
1	exp ruxolitinib/	693
2	ruxolitinib*.mp.	703
3	jakafi*.mp.	62
4	jakavi*.mp.	13
5	incb18424*.mp.	21
6	incb-18424*.mp.	60
7	"incb 18424".mp.	60
8	incb018424*.mp.	113
9	incb-018424*.mp.	186
10	"incb 018424".mp.	186
11	inc424*.mp.	21
12	inc-424*.mp.	15
13	"inc 424".mp.	15
14	1092939-17-7.rn.	479
15	or/1-14	832
16	exp myelofibrosis/	6538
17	exp myeloid metaplasia/	3193
18	myelofibros*.mp.	8670
19	osteomyelofibros*.mp.	251
20	myelo-fibros*.mp.	10
21	myelosclero*.mp.	220

22	(myeloid* adj2 metaplasia*).mp.	3619
23	(bone adj2 marrow* adj2 fibros*).mp.	730
24	or/16-23	10092
25	random:.tw.	900461
26	placebo:.mp.	335601
27	double-blind:.tw.	151302
28	or/25-27	1116638
29	and/15,24,28	177

Tabelle 4-58: Bibliographische Literaturrecherche in der Cochrane-Library: Cochrane Database of Systematic Reviews

Datenbank		COCHRANE Database of Systematic Reviews
Suchoberfläche		Ovid
Datum der Suche		24.02.2014
Zeitsegment		2005 to December 2013
#	Suchbegriffe	Ergebnis
1	ruxolitinib*.mp.	1
2	jakafi*.mp.	0
3	jakavi*.mp.	0
4	incb18424*.mp.	0
5	incb-18424*.mp.	0
6	"incb 18424".mp.	0
7	incb018424*.mp.	1
8	incb-018424*.mp.	0
9	"incb 018424".mp.	0
10	inc424*.mp.	0
11	inc-424*.mp.	0
12	"inc 424".mp.	0
13	or/1-12	1

Tabelle 4-59: Bibliographische Literaturrecherche in der Cochrane-Library: Cochrane Central Register of Controlled Trials

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	24.02.2014	
Zeitsegment	January 2014	
#	Suchbegriffe	Ergebnis
1	ruxolitinib*.mp.	11
2	jakafi*.mp.	0
3	jakavi*.mp.	0
4	incb18424*.mp.	0
5	incb-18424*.mp.	0
6	"incb 18424".mp.	0
7	incb018424*.mp.	6
8	incb-018424*.mp.	0
9	"incb 018424".mp.	0
10	inc424*.mp.	0
11	inc-424*.mp.	0
12	"inc 424".mp.	0
13	or/1-12	14

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

Im Folgenden sind die Suchstrategien für Abschnitt 4.3.1.1.3 „Studien aus der Suche in Studienregistern“ dokumentiert.

Tabelle 4-60: Suche im Studienregister ClinicalTrials.gov

Studienregister	ClinicalTrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	24.02.2014
Suchstrategie	Search Terms= (ruxolitinib OR jakafi OR jakavi OR incb18424 OR incb-18424 OR incb018424 OR incb-018424 OR inc424 OR inc-424) AND (myelofibrosis OR myelofibroses OR osteomyelofibrosis OR osteomyelofibroses OR myelosclerosis OR myeloscleroses OR myeloid-metaplasia OR bone-marrow-fibrosis OR bone-marrow-fibroses OR PMF OR PPV-MF OR PET-MF)
Treffer	32

Tabelle 4-61: Suche im Studienregister Klinische Prüfungen PharmNet.Bund

Studienregister	Klinische Prüfungen PharmNet.Bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	24.02.2014
Suchstrategie	Suche nach: ruxolitinib? In: active substance
Treffer	4

Tabelle 4-62: Suche im Studienregister EU Clinical Trials Register

Studienregister	EU Clinical Trials Register
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	24.02.2014
Suchstrategie	Search Terms= (ruxolitinib OR jakafi OR jakavi OR incb18424 OR incb-18424 OR incb018424 OR incb-018424 OR inc424 OR inc-424) AND (myelofibrosis OR myelofibroses OR osteomyelofibrosis OR osteomyelofibroses OR myelosclerosis OR myeloscleroses OR myeloid-metaplasia OR bone-marrow-fibrosis OR bone-marrow-fibroses OR PMF OR PPV-MF OR PET-MF)
Treffer	13

Tabelle 4-63: Suche im ICTRP Search Portal

Studienregister	ICTRP Search Portal
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	24.02.2014
Suchstrategie	Search Terms= Ruxolitinib AND myelofibrosis OR jakafi AND myelofibrosis OR jakavi AND myelofibrosis OR incb18424 AND myelofibrosis OR incb-18424 AND myelofibrosis OR incb018424 AND myelofibrosis OR incb-018424 AND myelofibrosis OR inc424 OR inc-424 AND myelofibrosis OR Ruxolitinib AND myelofibroses OR jakafi AND myelofibroses OR jakavi AND myelofibroses OR incb18424 AND myelofibroses OR incb-18424 AND myelofibroses OR incb018424 AND myelofibroses OR incb-018424 AND myelofibroses OR inc424 OR inc-424 AND myelofibroses
Treffer	76 Treffer zu 33 Studien

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

ClinicalTrials.gov

Tabelle 4-64 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister ClinicalTrials.gov

Nr.	Studien-ID	Zitat	Ausschlussgrund
1	NCT01732445	Ruxolitinib Phosphate and Danazol in Treating Anemia in Patients With Myelofibrosis. http://ClinicalTrials.gov/show/NCT01732445 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
2	NCT01787487	Evaluation of RUX and AZA Combination as a Therapy For Patients With Myelofibrosis And Myelodysplastic Syndrome/ Myeloproliferative Neoplasm. http://clinicaltrials.gov/show/NCT01787487 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
3	NCT01790295	Ruxolitinib Prior to Transplant in Patients With Myelofibrosis. http://clinicaltrials.gov/show/NCT01790295 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
4	NCT01392443	Asian Phase II Study of INC424 in Patients With Primary Myelofibrosis (MF), Post-PV MF or Post-ET MF. http://clinicaltrials.gov/show/NCT01392443 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
5	NCT01433445	Panobinostat and Ruxolitinib in Primary Myelofibrosis, Post-polycythemia Vera-myelofibrosis or Post-essential Thrombocythemia-myelofibrosis. http://clinicaltrials.gov/show/NCT01433445 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
6	NCT01340651	Study of Ruxolitinib (INCB018424) Sustained Release Formulation in Myelofibrosis Patients. http://clinicaltrials.gov/show/NCT01340651 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
7	NCT01445769	Alternative Dosing Strategy of Ruxolitinib in Patients With Myelofibrosis. http://clinicaltrials.gov/show/NCT01445769 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
8	NCT01348490	Ruxolitinib (INCB018424) in Subjects With Primary Myelofibrosis, Post Essential Thrombocythemia-myelofibrosis and Post Polycythemia Vera-myelofibrosis. http://clinicaltrials.gov/show/NCT01348490 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
9	NCT00509899	Open Label Ruxolitinib (INCB018424) in Patients With Myelofibrosis and Post Polycythemia Vera/Essential Thrombocythemia Myelofibrosis. http://clinicaltrials.gov/show/NCT00509899 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
10	NCT01969838	Momelotinib Versus Ruxolitinib in Subjects With Myelofibrosis. http://clinicaltrials.gov/show/NCT01969838 . Abgerufen am 16.01.2014.	Einschlusskriterium E3 nicht erfüllt
11	NCT01693601	Panobinostat and Ruxolitinib In MyElofibrosis (PRIME Trial). http://clinicaltrials.gov/show/NCT01693601 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
12	NCT01375140	Ruxolitinib and Lenalidomide for Patients With Myelofibrosis. http://clinicaltrials.gov/show/NCT01375140 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
13	NCT01644110	Ruxolitinib and Pomalidomide Combination Therapy in Patients With Primary and Secondary MF (POMINC). http://clinicaltrials.gov/show/NCT01644110 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
14	NCT01317875	Study of the JAK Inhibitor Ruxolitinib Administered Orally to Patients With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). http://clinicaltrials.gov/show/NCT01317875 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
15	NCT01787552	A Phase Ib/II Dose-finding Study to Assess the Safety and Efficacy of LDE225 + INC424 in Patients With MF. http://clinicaltrials.gov/show/NCT01787552 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
16	NCT01558739	Exploratory Phase II Study of INC424 Patients With Primary Myelofibrosis (PMF) or Post Polycythaemia Myelofibrosis (PPV MF) or Post Essential Thrombocythaemia Myelofibrosis (PET-MF) (MACS2030). http://clinicaltrials.gov/show/NCT01558739 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
17	NCT01369498	Efficacy and Safety of GS-6624 in Adults With Primary, Post Polycythemia Vera or Post Essential Thrombocythemia Myelofibrosis. http://clinicaltrials.gov/show/NCT01369498 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 nicht erfüllt
18	NCT01493414	INC424 for Patients With Myelofibrosis, Post Polycythemia Myelofibrosis or Post-essential Thrombocythemia Myelofibrosis. http://clinicaltrials.gov/show/NCT01493414 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
19	NCT01795677	JAK2 Inhibitors RUXOLITINIB in Patients With Myelofibrosis. http://clinicaltrials.gov/show/NCT01795677 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
20	NCT01981850	A Phase 2 Study Of PRM-151 In Subjects With Myelofibrosis. http://clinicaltrials.gov/show/NCT01981850 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
21	NCT01523171	Phase II, Open Label, Single Arm Study of SAR302503 In Myelofibrosis Patients Previously Treated With Ruxolitinib (JAKARTA2). http://clinicaltrials.gov/show/NCT01523171 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
22	NCT01730248	A Study to Find the Maximum Tolerated Dose of the Experimental Combination of the Drugs INC424 and BKM120 in Patients With Primary or Secondary Myelofibrosis. http://clinicaltrials.gov/show/NCT01730248 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
23	NCT01950780	An Open-Label Pilot Study to Evaluate the Efficacy of Ruxolitinib in Moderate to Severe Alopecia Areata. http://clinicaltrials.gov/show/NCT01950780 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
24	NCT01164163	INCB18424 in Treating Young Patients With Relapsed or Refractory Solid Tumor, Leukemia, or Myeloproliferative Disease. http://clinicaltrials.gov/show/NCT01164163 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
25	NCT01776723	A Sequential Two-Stage Dose Escalation Study to Evaluate the Safety and Efficacy of Ruxolitinib. http://clinicaltrials.gov/show/NCT01776723 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
26	NCT01702064	Ruxolitinib in Combination With Nilotinib in Chronic Myeloid Leukemia (CML) Patients. http://clinicaltrials.gov/show/NCT01702064 . Abgerufen am 16.01.2014.	Einschlusskriterium E1, E2 und E5 nicht erfüllt
27	NCT02015208	Ruxolitinib in the Treatment of Chronic Lymphocytic Leukemia. http://clinicaltrials.gov/show/NCT02015208 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
28	NCT01594723	A Study of LY2784544 in Participants With Myeloproliferative Neoplasms. http://clinicaltrials.gov/show/NCT01594723 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
29	NCT02066532	Phase I/II Trial of Ruxolitinib in Combination With Trastuzumab in Metastatic HER2 Positive Breast Cancer. http://clinicaltrials.gov/show/NCT02066532 . Abgerufen am 24.02.2014.	Einschlusskriterium E1, E2 und E5 nicht erfüllt
30	NCT02055781	Oral Pacritinib Versus Best Available Therapy to Treat Myelofibrosis With Thrombocytopenia. http://clinicaltrials.gov/show/NCT02055781 . Abgerufen am 24.02.2014.	Einschlusskriterium E2 nicht erfüllt

Klinische Prüfungen PharmNet.Bund

Tabelle 4-65 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister Klinische Prüfungen PharmNet.Bund

Nr.	Studien-ID	Zitat	Ausschlussgrund
1	EudraCT 2012-004023-20	A Phase Ib/II, open-label, multi-center, dose-finding study to assess the safety and efficacy of the oral combination of LDE225 and INC424 (Ruxolitinib) in patients with myelofibrosis. http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/ . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
2	EudraCT 2012-002431-29	A Phase-Ib/II Study of Ruxolitinib and Pomalidomide Combination Therapy in Patients with Primary and Secondary Myelofibrosis. http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/ . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
3	EudraCT 2010-024473-39	An open-label, multicenter, expanded access study of INC424 for patients with primary myelofibrosis (PMF) of post polycythemia vera myelofibrosis (PPV MF) or post-essential thrombocythemia myelofibrosis (PET-MF). http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/ . Abgerufen am 14.01.2014.	Einschlusskriterium E5 nicht erfüllt
4	EudraCT 2010-020807-57	Randomized, open label, multicenter phase III study of Efficacy and Safety in Polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK iNhibitor INC424 tablets verSus bEst available care (The RESPONSE Trial). http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/ . Abgerufen am 14.01.2014.	Einschlusskriterium E1 nicht erfüllt

EU Clinical Trials Register

Tabelle 4-66 (Anhang): Liste der ausgeschlossenen Studien aus dem Studienregister EU Clinical Trials Register

Nr.	Studien-ID	Zitat	Ausschlussgrund
1	EudraCT 2010-024473-39	An open-label, multicenter, expanded access study of INC424 for patients with primary myelofibrosis (PMF) of post polycythemia vera myelofibrosis (PPV MF) or post-essential thrombocythemia myelofibrosis (PET-MF). https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024473-39 . Abgerufen am 14.01.2014.	Einschlusskriterium E5 nicht erfüllt
2	EudraCT 2010-023055-29	A Phase Ib, open-label, dose-finding study of the JAK inhibitor INC424 tablets administered orally to patients with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF) and baseline platelet counts $\geq 50 \times 10^9/L$ and $< 100 \times 10^9/L$. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-023055-29 . Abgerufen am 14.01.2014.	Einschlusskriterium E5 nicht erfüllt
3	EudraCT 2012-004239-21	A Randomized Controlled Phase 3 Study of Oral Pacritinib versus Best Available Therapy in Patients with Primary Myelofibrosis, Post-Polycythemia Vera Myelofibrosis, or Post-Essential Thrombocythemia Myelofibrosis. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004239-21 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
4	EudraCT 2012-004023-20	A Phase Ib/II, open-label, multi-center, dose-finding study to assess the safety and efficacy of the oral combination of LDE225 and INC424 (Ruxolitinib) in patients with myelofibrosis. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004023-20 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
5	EudraCT 2011-001012-56	A Phase 2 Study of LY2784544 in Patients with Myeloproliferative Neoplasms. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-001012-56 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
6	EudraCT 2011-005066-38	A UK open-label, multicentre, exploratory Phase II study of INC424 for patients with primary myelofibrosis (PMF) or post polycythemia myelofibrosis (PPV MF) or post-essential thrombocythemia myelofibrosis (PET-MF). https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-005066-38 . Abgerufen am 14.01.2014.	Einschlusskriterium E5 nicht erfüllt
7	EudraCT 2012-002431-29	A Phase-Ib/II Study of Ruxolitinib and Pomalidomide Combination Therapy in Patients with Primary and Secondary Myelofibrosis. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002431-29 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
8	EudraCT 2013-003295-12	Combination Therapy Interferon Alpha + JAK1-2 Inhibitor in The Ph-Negative Chronic Myeloid Neoplasms. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-003295-12 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
9	EudraCT 2011-005226-21	A Phase II, Multicenter, Open Label, Single Arm Study of SAR302503 in Subjects Previously Treated with Ruxolitinib and with a Current Diagnosis of Intermediate or High-Risk Primary Myelofibrosis, Post-Polycythemia Vera Myelofibrosis, or Post-Essential Thrombocythemia Myelofibrosis. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-005226-21 . Abgerufen am 14.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
10	EudraCT 2011-005279-18	A RandoMised study of best Available therapy versus JAK Inhibition in patients with high risk Polycythaemia Vera or Essential Thrombocythaemia who are resistant or intolerant to HydroxyCarbamide. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-005279-18 . Abgerufen am 14.01.2014.	Einschlusskriterium E1 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
11	EudraCT 2012-002253-30	A phase 2 study of Ruxolitinib in patients with splanchnic vein thrombosis associated with myeloproliferative neoplasm. https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002253-30 . Abgerufen am 24.02.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
12	EudraCT 2013-003583-31	Randomized, open label, multicenter phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with polycythemia vera who are hydroxyurea resistant or intolerant (Response 2). https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-003583-31 . Abgerufen am 24.02.2014.	Einschlusskriterium E1 nicht erfüllt

ICTRP Search Portal

Tabelle 4-67 (Anhang): Liste der ausgeschlossenen Studien aus dem ICTRP Search Portal

Nr.	Studien-ID	Zitat	Ausschlussgrund
1	NCT01981850	A Phase 2 Study Of PRM-151 In Subjects With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01981850 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
2	EUCTR2013-003295-12-DK	Combination Therapy Interferon Alpha + JAK1-2 Inhibitor in The Ph-Negative Chronic Myeloid Neoplasms. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2013-003295-12-DK . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
3	NCT01969838	Momelotinib Versus Ruxolitinib in Subjects With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01969838 . Abgerufen am 16.01.2014.	Einschlusskriterium E3 nicht erfüllt
4	NCT01877005	A Phase II Study of Oral JAK1/JAK2 Inhibitor INC424 in Adult Patients With Relapsed/Refractory Classical Hodgkin's Lymphoma (HIJAK). http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01877005 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
5	NCT01790295	Ruxolitinib Prior to Transplant in Patients With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01790295 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
6	NCT01787552	A Phase Ib/II Dose-finding Study to Assess the Safety and Efficacy of LDE225 + INC424 in Patients With MF. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01787552 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
7	NCT01787487	Evaluation of RUX and AZA Combination as a Therapy For Patients With Myelofibrosis And Myelodysplastic Syndrome/ Myeloproliferative Neoplasm. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01787487 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
8	NCT01795677	JAK2 Inhibitors RUXOLITINIB in Patients With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01795677 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
9	NCT01732445	Ruxolitinib Phosphate and Danazol in Treating Anemia in Patients With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01732445 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
10	NCT01730248	A Study to Find the Maximum Tolerated Dose of the Experimental Combination of the Drugs INC424 and BKM120 in Patients With Primary or Secondary Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01730248 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
11	NCT01693601	Panobinostat and Ruxolitinib In MyElofibrosis (PRIME Trial) PRIME. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01693601 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
12	NCT01644110	Ruxolitinib and Pomalidomide Combination Therapy in Patients With Primary and Secondary MF (POMINC). http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01644110 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
13	EUCTR2011-005066-38-GB	Safety and efficacy study of INC424 in patients with myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005066-38-GB . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
14	EUCTR2011-005279-18-GB	A study of best available therapy versus Ruxolitinib in patients with high risk Polycythaemia Vera or Essential Thrombocythaemia who are resistant or intolerant to HydroxyCarbamide (standard therapy). http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005279-18-GB . Abgerufen am 16.01.2014.	Einschlusskriterium E1 nicht erfüllt
15	EUCTR2011-005226-21-GB	A Phase II, Multicenter, Open Label, Single Arm Study of SAR302503 in Subjects Previously Treated with Ruxolitinib and with a Current Diagnosis of Intermediate or High-Risk Primary Myelofibrosis, Post-Polycythemia Vera Myelofibrosis, or Post-Essential Thrombocythemia Myelofibrosis - JAKARTA2. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005226-21-GB . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
16	NCT01445769	Alternative Dosing Strategy of Ruxolitinib in Patients With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01445769 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
17	NCT01392443	Asian Phase II Study of INC424 in Patients With Primary Myelofibrosis (MF), Post-PV MF or Post-ET MF. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01392443 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
18	NCT01433445	Panobinostat and Ruxolitinib in Primary Myelofibrosis, Post-polycythemia Vera-myelofibrosis or Post-essential Thrombocythemia-myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01433445 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
19	NCT01375140	Ruxolitinib and Lenalidomide for Patients With Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01375140 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
20	NCT01369498	Efficacy and Safety of GS-6624 in Adults With Primary, Post Polycythemia Vera or Post Essential Thrombocythemia Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01369498 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 nicht erfüllt
21	EUCTR2010-024473-39-AT	Safety and efficacy study of INC424 in patients with myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-024473-39-AT . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
22	NCT01348490	Ruxolitinib (INCB018424) in Subjects With Primary Myelofibrosis, Post Essential Thrombocythemia-myelofibrosis and Post Polycythemia Vera-myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01348490 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
23	NCT01340651	Study of INCB018424 Sustained Release Formulation in Myelofibrosis Patients. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01340651 . Abgerufen am 16.01.2014.	Einschlusskriterium E2 und E5 nicht erfüllt
24	NCT01317875	Study of the JAK Inhibitor Ruxolitinib Administered Orally to Patients With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01317875 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt

Nr.	Studien-ID	Zitat	Ausschlussgrund
25	EUCTR2010-023055-29-AT	A Phase Ib, open-label, dose-finding study of the JAK inhibitor INC424 tablets administered orally to patients with Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF) or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF) and baseline platelet counts = 50 x10 ⁹ /L and <100 x10 ⁹ /L. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-023055-29-AT . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
26	EUCTR2010-020807-57-BE	Randomized, open label, multicenter phase III study of Efficacy and Safety in POLycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK iNhibitor INC424 tablets verSus bEst available care (The RESPONSE Trial). http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020807-57-BE . Abgerufen am 16.01.2014.	Einschlusskriterium E1 nicht erfüllt
27	NCT01243944	Study of Efficacy and Safety in Polycythemia Vera Subjects Who Are Resistant to or Intolerant of Hydroxyurea: JAK Inhibitor INC424 (INCB018424) Tablets Versus Best Available Care: (The RESPONSE Trial). http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01243944 . Abgerufen am 16.01.2014.	Einschlusskriterium E1 nicht erfüllt
28	NCT00509899	Open Label Ruxolitinib (INCB018424) in Patients With Myelofibrosis and Post Polycythemia Vera/Essential Thrombocythemia Myelofibrosis. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00509899 . Abgerufen am 16.01.2014.	Einschlusskriterium E5 nicht erfüllt
29	NCT02049450	Study of Efficacy and Safety of INC424 in Regularly Transfused Patients With Thalassemia. http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT02049450 . Abgerufen am 24.02.2014.	Einschlusskriterium E1 und E5 nicht erfüllt
30	EUCTR2013-003583-31-ES	Ruxolitinib efficacy and safety in patients with HU resistant or intolerant polycythemia vera vs best available therapy. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2013-003583-31-ES . Abgerufen am 24.02.2014.	Einschlusskriterium E1 nicht erfüllt
31	EUCTR2012-002253-30-IT	A phase 2 study of Ruxolitinib in patients with splanchnic vein thrombosis associated with myeloproliferative neoplasm. http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2012-002253-30-IT . Abgerufen am 24.02.2014.	Einschlusskriterium E1 und E5 nicht erfüllt

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 4-68 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 4-68 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-68 (Anhang): Studiendesign und -methodik für Studie COMFORT-II (CINC424A2352)

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	<u>Primäres Studienziel:</u> Untersuchung der Wirksamkeit, Sicherheit und Verträglichkeit von Ruxolitinib bei zweimal täglicher Gabe im Vergleich zu BAT bei Patienten mit PMF, PPV-MF und PET-MF <u>Sekundäres Studienziel:</u> Untersuchung der Populations-Pharmakokinetik von Ruxolitinib
Methoden		
3	Studiendesign	Randomisierte kontrollierte Phase-III-Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, offene, multizentrische Parallelgruppenstudie Randomisierung der 219 Patienten im Verhältnis 2:1, stratifiziert nach Risikoprofil (intermediär-2 bzw. hoch): - Interventionsgruppe (146 Patienten): Ruxolitinib 15 mg 2x täglich bzw. 20 mg 2x täglich (entsprechend der Thrombozytenzahl zu Studienbeginn) - Kontrollgruppe (73 Patienten): Best Available Therapy (BAT) nach Ermessen des behandelnden Prüfarztes (Kombinationen, Therapiewechsel oder Nichtbehandlung erlaubt) <u>Randomisierte Behandlungsphase:</u> Behandlung bis zum Ende der Studie bzw. bis zur Progression der Erkrankung, definiert als - Zunahme des Milzvolumens um $\geq 25\%$ - Milzbestrahlung oder Splenektomie - Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ - Tod <u>Extensionsphase:</u> Bei Zunahme des Milzvolumens um $\geq 25\%$ bzw. der Durchführung einer Splenektomie konnten Patienten aus beiden Studienarmen unter Weiterführung bzw. Neuaufnahme der Ruxolitinib-Therapie in die Extensionsphase der Studie wechseln Aufgrund der Überlegenheit von Ruxolitinib in der Primäranalyse konnten ab Februar 2011 alle Patienten unabhängig von den erwähnten Progressionskriterien in die Extensionsphase der Studie wechseln (Amendment 5 des Studienprotokolls) Weiterbehandlung ursprünglich bis zu Woche 144 geplant; durch Amendment 7 (Mai 2012) wurde die Studiendauer auf 5 Jahre verlängert <u>Primäre Datenanalyse:</u> Nachdem alle Patienten 48 Wochen behandelt worden waren

Item ^a	Charakteristikum	Studieninformation
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Das Studienprotokoll wurde insgesamt siebenmal geändert. Folgende wesentliche Änderungen ergaben sich: <u>Amendment 3 vom 25. Februar 2010:</u> - nach Abschluss der Rekrutierungsphase im Januar 2010 - auch bei vorzeitigem Abbruch der Behandlung sollen für alle Patienten Daten in Bezug auf ihr Überleben, eine leukämische Transformation sowie die Einnahme neuer antineoplastischer Medikamente erhoben werden <u>Amendment 5 vom 2. Februar 2011:</u> - nach Schluss der Datenbank für die Primäranalyse am 4. Januar 2011 - aufgrund der klaren Überlegenheit von Ruxolitinib bei der primären Datenanalyse wird allen Patienten gestattet, zur Behandlung mit Ruxolitinib in die Extensionsphase der Studie wechseln <u>Amendment 7 vom 21. Mai 2012:</u> - Verlängerung der Studiendauer von 144 Wochen auf 5 Jahre, um Langzeitdaten zur Wirksamkeit und Sicherheit der Therapie mit Ruxolitinib erheben zu können
4	Probanden / Patienten	Erwachsene Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) mit Intermediär-2- oder Hochrisikoprofil, die nicht für eine allogene Stammzelltransplantation in Betracht kamen
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> 1. Erwachsene Patienten (≥ 18 Jahre alt) 2. Diagnose einer PMF, PPV-MF oder PET-MF nach den WHO-Kriterien, unabhängig vom JAK2-Mutationsstatus 3. Intermediär-2- bzw. Hochrisikoprofil entsprechend den IWG-MRT-Kriterien, d.h. Vorhandensein von 2 bzw. ≥ 3 der folgenden Risikofaktoren - Alter ≥ 65 Jahre - konstitutionelle Symptome (Fieber, Gewichtsverlust, Nachtschweiß) - Hämoglobin $< 10\text{g/dl}$ - Leukozyten $> 25 \times 10^9/\text{l}$ - Blasten $\geq 1\%$ 4. Blasten im peripheren Blut $< 10\%$ 5. Eastern Cooperative Oncology Group (ECOG)-Status ≤ 3 6. Milz tastbar mit ≥ 5 cm unter dem Rippenbogen 7. Stabiles Therapieregime für ≥ 2 Wochen vor dem Screening bzw. ≥ 4 Wochen vor der Baseline-Untersuchung 8. Keine Vorbehandlung mit einem JAK-Inhibitor <u>Ausschlusskriterien:</u> 1. Lebenserwartung weniger als 6 Monate 2. Schwangere bzw. stillende Patientinnen 3. Patienten im fortpflanzungsfähigen Alter, die keine Vorsichtsmaßnahmen treffen wollen, um die Zeugung eines

Item ^a	Charakteristikum	Studieninformation
		Kindes bzw. eine Schwangerschaft zu verhindern 4. Inadäquate Knochenmarkfunktion, definiert als - absolute Neutrophilenzahl (ANC) $\leq 1000/\mu\text{l}$ - Thrombozytenzahl $< 100.000/\mu\text{l}$ ohne Gabe von Wachstumsfaktoren bzw. Transfusionen 5. Vorgeschichte einer Thrombozytenzahl $< 50.000/\mu\text{l}$ bzw. einer ANC $< 500/\mu\text{l}$, außer wenn diese während der Behandlung einer myeloproliferativen Erkrankung oder während einer Chemotherapie auftrat 6. Inadäquate Leber- oder Nierenfunktion: - direktes Bilirubin $> 2,0 \times \text{ULN}$ (upper limit of normal) - Alaninaminotransferase (ALT) $> 2,5 \times \text{ULN}$ - Kreatinin $> 2,0 \text{ mg/dl}$ 7. Behandlungbedürftige, klinisch signifikante bakterielle, fungale, parasitäre oder virale Infektion 8. Patienten dürfen derzeit keine Möglichkeit für eine allogene Stammzelltransplantation haben (wegen fehlender Patienteneignung bzw. mangelnder Spenderverfügbarkeit) 9. Vorgeschichte einer malignen Erkrankung in den letzten 5 Jahren mit Ausnahme eines behandelten Plattenepithel- bzw. Basalzellkarzinoms im frühen Stadium 10. Kardiale Erkrankungen, die nach Einschätzung des Prüfarztes die Sicherheit des Patienten oder seine Compliance beeinträchtigen können 11. Unkontrollierte oder instabile Angina pectoris 12. Behandlung mit anderen Medikamenten, die sich noch in der klinischen Prüfung befinden, innerhalb von 14 Tagen (bzw. 6 Halbwertszeiten) vor Studienbeginn 13. Patienten, deren Myelofibrose-Therapieregime innerhalb von 2 Wochen vor Beginn des Screenings modifiziert wurde 14. Patienten mit Milzbestrahlung innerhalb von 12 Monaten vor Beginn des Screenings 15. Patienten mit Vorhofflimmern (auch paroxysmal) 16. Behandlung mit hämatopoetischen Wachstumsfaktoren bzw. Agonisten an deren Rezeptoren (z.B. Erythropoetin, G-CSF, Romiplostim, Eltrombopag) innerhalb von 2 Wochen vor dem Screening bzw. 4 Wochen vor der Baseline-Untersuchung 17. Patienten mit Myokardinfarkt oder akutem Koronarsyndrom (innerhalb der letzten 6 Monate) 18. Ruhepuls > 110 Schläge/Minute 19. Systolischer Blutdruck $> 160 \text{ mmHg}$ 20. Diastolischer Blutdruck $> 100 \text{ mmHg}$ 21. Unfähigkeit zur Einwilligungserklärung 22. Alkohol- oder Drogenabhängigkeit, die die Fähigkeit, den Vorgaben des Studienprotokolls zu folgen, beeinträchtigt 23. Patienten mit irgendeiner Erkrankung, die nach Einschätzung des Prüfarztes die Sicherheit des Patienten oder seine Compliance beeinträchtigen kann

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Die Studie wurde in 56 Zentren in 9 europäischen Ländern durchgeführt. In folgenden Ländern befanden sich Studienzentren: Belgien, Deutschland, Frankreich, Großbritannien, Italien, Niederlande, Österreich, Schweden, Spanien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<u>Interventionsgruppe (N = 146):</u> Ruxolitinib zur oralen Einnahme in Abhängigkeit von der Thrombozytenzahl zu Studienbeginn: - Thrombozytenzahl > 200.000/μl: Anfangsdosis 20 mg 2x täglich (je vier 5 mg-Tabletten) - Thrombozytenzahl > 100.000/μl und $\leq$ 200.000/μl: Anfangsdosis 15 mg 2x täglich (je drei 5 mg-Tabletten) Dosisanpassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit gemäß Studienprotokoll erlaubt Höchst-dosis: 25 mg 2x täglich (50 mg/Tag) <u>Kontrollgruppe (N = 73):</u> Best Available Therapy (BAT): die nach dem Ermessen des jeweiligen Prüfarztes beste verfügbare Therapie, wobei auch ein Therapiewechsel, eine Kombination verschiedener Therapieoptionen oder eine Nichtbehandlung möglich waren Folgende Medikamente kamen zum Einsatz: - Hydroxycarbamid - Anagrelid - Cytarabin - Prednison - Methylprednison - Prednisolon - Epoetin-α - Thalidomid - Lenalidomid - Mercaptopurin - Thioguanin - Danazol - Pegasys - Interferon-α - Melphalan - Acetylsalicylsäure - Colchicin
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.	<u>Das primäre Zielkriterium war:</u> 1. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq$ 35% nach 48 Wochen - Rückgang des Milzvolumens im Vergleich zur Baseline gemessen mittels Magnetresonanztomographie (MRT) bzw. Computertomographie (CT)

Item ^a	Charakteristikum	Studieninformation
	Angaben zur Validierung von Erhebungsinstrumenten	- radiologische Untersuchung durch lokale Studienärzte; Auswertung der Aufnahmen durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war <u>Die wichtigsten sekundären Zielkriterien waren:</u> 1. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen - Rückgang des Milzvolumens im Vergleich zur Baseline gemessen mittels Magnetresonanztomographie (MRT) bzw. Computertomographie (CT) - radiologische Untersuchung durch lokale Studienärzte; Auswertung der Aufnahmen durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war 2. Dauer des Ansprechens (Milzvolumenreduktion) - Zeit vom ersten Erreichen einer Milzvolumenreduktion von $\geq 35\%$ bis zum Verlust dieses Ansprechens bei gleichzeitiger Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie - falls kein Ereignis auftrat, erfolgte eine Zensurierung zum Zeitpunkt der letzten Milzuntersuchung 3. Zeit bis zum ersten Ansprechen (Milzvolumenreduktion) - definiert als die Zeit von der Randomisierung bis zum ersten dokumentierten Erreichen einer Milzvolumenreduktion um $\geq 35\%$ 4. Progressionsfreies Überleben (progression-free survival, PFS) - Zeit von der Randomisierung bis zum ersten dokumentierten Auftreten eines der folgenden Ereignisse: • Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie • Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ • Milzbestrahlung, Splenektomie oder Tod 5. Leukämiefreies Überleben (leukemia-free survival, LFS) - Zeit von der Randomisierung bis zum ersten dokumentierten Auftreten eines der folgenden Ereignisse: • Leukämische Transformation definiert als ein Anstieg der Blasten im peripheren Blut auf $\geq 20\%$ für mindestens 8 Wochen bzw. ein Anstieg der Blasten im Knochenmark auf $\geq 20\%$ • Tod durch jedwede Ursache 6. Gesamtüberleben (overall survival, OS) - Zeit von der Randomisierung bis zum Tod durch jedwede Ursache

Item ^a	Charakteristikum	Studieninformation
		<u>Erhebungszeitpunkte</u> Analyse des Milzvolumens durch MRT bzw. CT: Baseline, nach 12, 24, 36 und 48 Wochen sowie danach alle 12 Wochen; ab November 2012 nur noch alle 24 Wochen Erhebung der Daten zur leukämischen Transformation, zur Einnahme neuer antineoplastischer Medikamente sowie zum Gesamtüberleben alle 3 Monate
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	Geplante Rekrutierungszahl: 150 Zahl der tatsächlich rekrutierten Patienten: 219 (nach einer Randomisierung im Verhältnis 2:1 erhielten 146 Patienten Ruxolitinib und 73 Patienten BAT)
7a	Wie wurden die Fallzahlen bestimmt?	Primärer Endpunkt der Studie war der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen. Unter Annahme einer Ansprechrate von 35% in der Interventionsgruppe bzw. 10% in der Kontrollgruppe ergab eine Fallzahl von 150 Patienten (100 bzw. 50 Patienten in der Interventions- bzw. Kontrollgruppe) im Chi ² -Test bei einem Signifikanzniveau von 0,05 eine Power von 90%.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte mit Hilfe eines IVRS (Interactive Voice Response System).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte im Verhältnis 2:1 in 6er-Blöcken und war stratifiziert nach Risikoprofil (intermediär-2 bzw. hoch).
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	Vor Therapiebeginn wurde jedem Patienten mit Hilfe des IVRS eine Randomisierungsnummer zugeteilt, durch die der Patient einem der beiden Studienarme zugewiesen wurde.
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?	Die Randomisierung erfolgte mit Hilfe eines IVRS-Systems. Nach Verifizierung aller verfügbaren Anforderungen kontaktierte das Studienpersonal des Zentrums das IVRS, durch das der Patient einem der beiden Studienarme zugewiesen wurde.
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?	Da es sich um eine offene Studie handelt, war sowohl Ärzten als auch Patienten die Zuteilung der Studienmedikation bekannt. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Bestimmung des Milzvolumens durch einen zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Es wurden folgende Analysepopulationen festgelegt: 1. Intention-to-treat (ITT)-Population - Auswertung aller randomisierten Patienten entsprechend dem Intention-to-treat (ITT)-Prinzip 2. Per-Protocol (PP)-Population - Auswertung aller Patienten der ITT-Population, bei denen es zu keiner relevanten Protokollverletzung gekommen war - relevante Protokollverletzungen waren z. B.: • keine bestätigte Diagnose einer PMF, PPV-MF oder PET-MF • Vorbehandlung mit einem JAK-Inhibitor • Milzbestrahlung innerhalb von 12 Monaten vor Beginn des Screenings • Einsatz von Hydroxycarbamid, Interferon, Thalidomid, Busulfan, Lenalidomid oder Anagrelid in der Interventionsgruppe • Einnahme der falschen Studienmedikation • keine Milzvolumen-Messung nach der Baseline-Untersuchung vorhanden 3. Sicherheitspopulation - Auswertung aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten - Patienten der BAT-Gruppe, die keine Behandlung erhielten, wurden ebenfalls in die Auswertung eingeschlossen <u>Statistische Methoden zur Bewertung relevanter Zielkriterien:</u> 1. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen - Vergleich der Ansprechraten mit Hilfe eines exakten Cochran-Mantel-Haenszel-Tests stratifiziert nach Risikoprofil mit einem zweiseitigen Signifikanzniveau von 0,05 - Berechnung der exakten Konfidenzintervalle basierend auf der Binominalverteilung - nur Patienten mit auswertbarer Milzvolumen-Messung

Item ^a	Charakteristikum	Studieninformation
		zu Studienbeginn wurden in die Analyse einbezogen - Patienten ohne Milzvolumen-Bestimmung nach 48 Wochen wurden als Non-Responder klassifiziert 2. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen - Vergleich der Ansprechraten mit Hilfe eines exakten Cochran-Mantel-Haenszel-Tests stratifiziert nach Risikoprofil mit einem zweiseitigen Signifikanzniveau von 0,05 - Berechnung der exakten Konfidenzintervalle basierend auf der Binominalverteilung - nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn wurden in die Analyse einbezogen - Patienten ohne Milzvolumen-Bestimmung nach 24 Wochen wurden als Non-Responder klassifiziert - eine statistische Auswertung dieses Endpunktes sollte nur dann erfolgen, wenn zuvor bei der Analyse des primären Zielkriteriums ein signifikanter Unterschied zwischen den beiden Gruppen nachgewiesen worden war (hierarchische Testprozedur) 3. Dauer des Ansprechens (Milzvolumenreduktion) - Schätzung der Dauer des Ansprechens mit Hilfe der Kaplan-Meier-Methode (nur Responder) 4. Zeit bis zum ersten Ansprechen (Milzvolumenreduktion) - Schätzung der Zeit bis zum ersten Ansprechen mit Hilfe der Kaplan-Meier-Methode (nur Responder) 5. Progressionsfreies Überleben (progression-free survival, PFS) - Schätzung des progressionsfreien Überlebens mit Hilfe der Kaplan-Meier-Methode - Vergleich der Kaplan-Meier-Kurven zwischen den Gruppen mittels eines stratifizierten Logrank-Tests und Berechnung der Hazard Ratios und zugehörigen 95%-Konfidenzintervalle mit einem stratifizierten Cox-Regressionsmodell 6. Leukämiefreies Überleben (leukemia-free survival, LFS) - Schätzung des leukämiefreien Überlebens mit Hilfe der Kaplan-Meier-Methode - Vergleich der Kaplan-Meier-Kurven zwischen den Gruppen mittels eines stratifizierten Logrank-Tests und Berechnung der Hazard Ratios und zugehörigen 95%-Konfidenzintervalle mit einem stratifizierten Cox-Regressionsmodell 7. Gesamtüberleben (overall survival, OS) - Schätzung des Gesamtüberlebens mit Hilfe der Kaplan-Meier-Methode - Vergleich der Kaplan-Meier-Kurven zwischen den Gruppen mittels eines stratifizierten Logrank-Tests und Berechnung der Hazard Ratios und zugehörigen 95%-Konfidenzintervalle mit einem stratifizierten Cox-Regressionsmodell

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<u>Subgruppenanalysen:</u> Zur Überprüfung der Konsistenz des Behandlungseffektes wurden für den primären Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen“ Subgruppenanalysen durchgeführt. Dabei wurden jeweils die Ansprechraten angegeben und die Odds Ratios sowie die zugehörigen 95%-Konfidenzintervalle mit Hilfe eines logistischen Regressionsmodells berechnet. Die Analyse des Zielkriteriums erfolgte dabei getrennt in folgenden a priori definierten Subgruppen: - Geschlecht (männlich, weiblich) - Alter (≤ 65 Jahre, > 65 Jahre) - Hautfarbe (Weiße, Nicht-Weiße) - Myelofibrose-Typ (PMF, PPV-MF, PET-MF) - Risikogruppe (intermediär-2, hoch) - Milzvolumen zu Studienbeginn ($\leq$ Median, $>$ Median) - Thrombozytenzahl zu Studienbeginn ($\leq 200.000/\mu\text{l}$, $> 200.000/\mu\text{l}$) - Vorbehandlung mit Hydroxycarbamid (ja, nein) <u>Sensitivitätsanalysen:</u> Zur Einschätzung der Robustheit der Ergebnisse wurden verschiedene Sensitivitätsanalysen durchgeführt. So wurde z. B. der primäre Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen“ auch unter Berücksichtigung von Patienten mit fehlendem Baseline-Wert ausgewertet. Die Schätzung des Gesamtüberlebens und des leukämiefreien Überlebens erfolgte zusätzlich auch unter Zensurierung der Patienten am Ende der randomisierten Studienphase. <u>Adjustierung für multiples Testen:</u> Die Analyse des primären Zielkriteriums „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 48 Wochen“ sowie des wichtigsten sekundären Zielkriteriums „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen“ erfolgte mit Hilfe einer hierarchischen Testabfolge. So wurde der wichtigste sekundäre Endpunkt nur bei Nachweis eines signifikanten Unterschieds in Bezug auf den primären Endpunkt ausgewertet.
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	a) Anzahl der randomisierten Patienten - Ruxolitinib: n = 146 - BAT: n = 73 b) Anzahl der tatsächlich behandelten Patienten: - Ruxolitinib: n = 146 - BAT: n = 73

Item ^a	Charakteristikum	Studieninformation
	haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	c) Anzahl der bei der Analyse des primären Zielkriteriums berücksichtigten Patienten: - Ruxolitinib: n = 144 - BAT: n = 72 2 Ruxolitinib-Patienten und 1 BAT-Patient wurden aufgrund fehlender bzw. nicht auswertbarer Milzvolumen-Messung zu Studienbeginn nicht in die Analyse des primären Zielkriteriums eingeschlossen
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Die primäre Datenanalyse fand am 4. Januar 2011 statt, nachdem alle Patienten 48 Wochen behandelt worden waren. Ein 3-Jahres-Update erfolgte am 1. Dezember 2012, nachdem alle Patienten 144 Wochen behandelt worden waren. Anzahl der Patienten, die die Therapie in der randomisierten Phase zum Zeitpunkt der jeweiligen Analyse beendet hatten: <u>Primäranalyse vom 4. Januar 2011:</u> 1. Ruxolitinib: 55 von 146 (37,7%) - Eintritt in die Extensionsphase (laut Studienprotokoll nach Krankheitsprogression erlaubt): n = 29 - unerwünschte Ereignisse: n = 12 - Rückzug der Einwilligungserklärung: n = 2 - Krankheitsprogression: n = 1 - sonstige Gründe: n = 11 2. BAT: 42 von 73 (57,5%) - Eintritt in die Extensionsphase (laut Studienprotokoll nach Krankheitsprogression erlaubt): n = 18 - unerwünschte Ereignisse: n = 4 - Rückzug der Einwilligungserklärung: n = 9 - Krankheitsprogression: n = 3 - sonstige Gründe: n = 8 <u>3-Jahres-Update vom 1. Dezember 2012:</u> 1. Ruxolitinib: 80 von 146 (54,8%) - unerwünschte Ereignisse: n = 24 - Rückzug der Einwilligungserklärung: n = 9 - Krankheitsprogression: n = 22 - sonstige Gründe: n = 25 2. BAT: 73 von 73 (100%) - Cross-over in den Ruxolitinib-Arm: n = 45 - Abbruch der Studienbehandlung: n = 28 • unerwünschte Ereignisse: n = 5 • Rückzug der Einwilligungserklärung: n = 9 • Krankheitsprogression: n = 4 • sonstige Gründe: n = 10
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme	Der Einschreibungszeitraum war zwischen 1. Juli 2009 und 22. Januar 2010. Die primäre Datenanalyse fand am 4. Januar

Item ^a	Charakteristikum	Studieninformation
	der Probanden / Patienten und der Nachbeobachtung	2011 statt. Zu diesem Zeitpunkt waren alle Patienten 48 Wochen behandelt worden, das mediane Follow-up betrug 51 Wochen. Ein 3-Jahres-Update erfolgte am 1. Dezember 2012. Zu diesem Zeitpunkt waren alle Patienten 144 Wochen behandelt worden, das mediane Follow-up betrug 151 Wochen. Die geplante Studiendauer beträgt 5 Jahre.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie läuft noch und wird voraussichtlich im Januar 2015 beendet.
a: nach CONSORT 2010.		

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

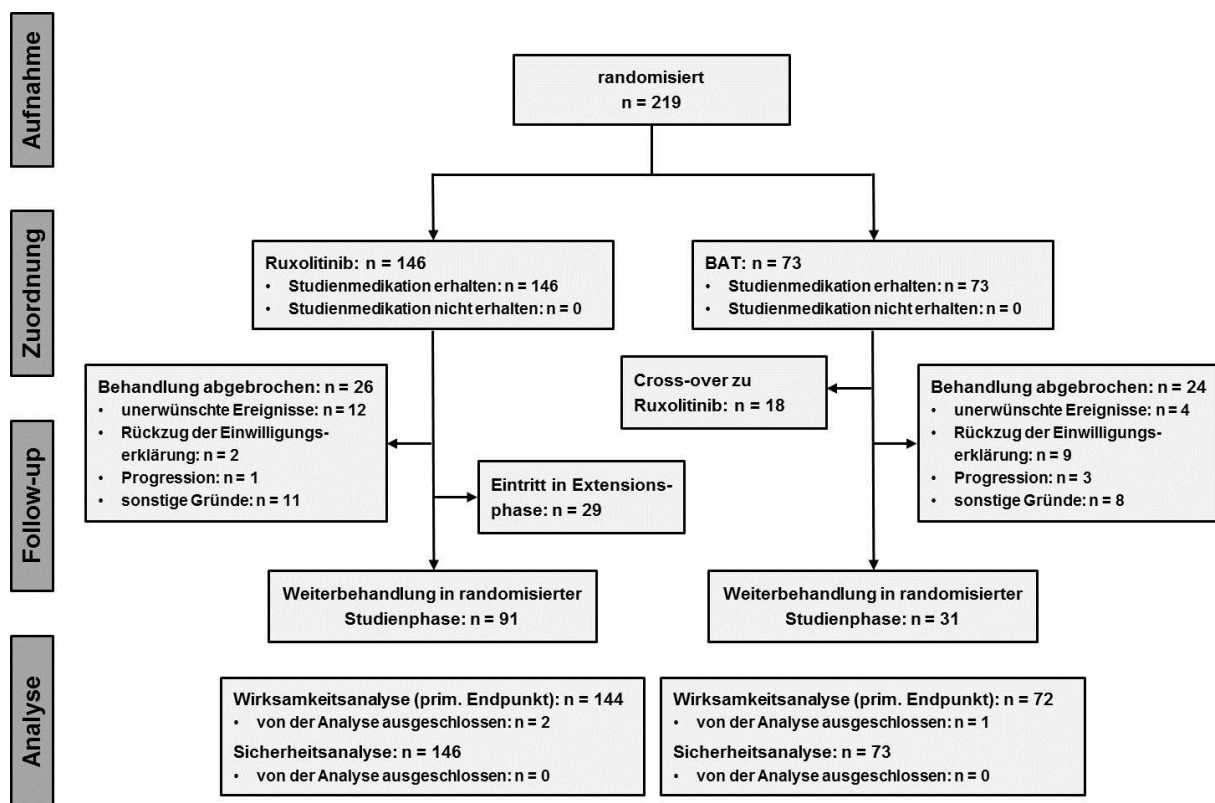


Abbildung 4-15: Flow-Chart zum Patientenfluss in der COMFORT-II-Studie (primäre Datenanalyse vom 4. Januar 2011)

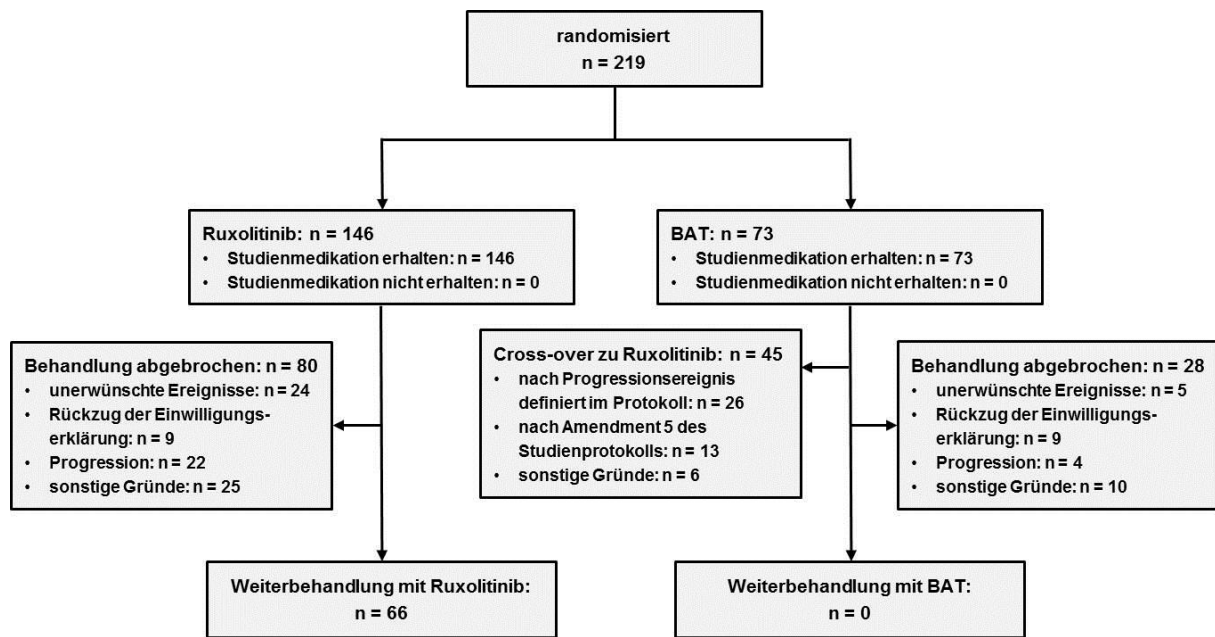


Abbildung 4-16: Flow-Chart zum Patientenfluss in der COMFORT-II-Studie (3-Jahres-Update vom 1. Dezember 2012)

Tabelle 4-69 (Anhang): Studiendesign und -methodik für Studie COMFORT-I (INCB 18424-351)

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	<u>Primäres Studienziel:</u> - Untersuchung der Wirksamkeit, Sicherheit und Verträglichkeit von Ruxolitinib bei zweimal täglicher Gabe im Vergleich zu Placebo bei Patienten mit PMF, PPV-MF und PET-MF <u>Sekundäre Studienziele:</u> - Untersuchung der Auswirkung von Ruxolitinib auf die MF-Symptomatik aus Sicht der Patienten - Untersuchung der Populations-Pharmakokinetik von Ruxolitinib
Methoden		
3	Studiendesign	Randomisierte kontrollierte Phase-III-Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, placebo-kontrollierte, multizentrische Parallelgruppenstudie Randomisierung der 309 Patienten im Verhältnis 1:1: - Interventionsgruppe (155 Patienten): Ruxolitinib 15 mg 2x täglich bzw. 20 mg 2x täglich (entsprechend der Thrombozytenzahl zu Studienbeginn) - Kontrollgruppe (154 Patienten): Placebo <u>Cross-over:</u> Bei Zunahme des Milzvolumens um $\geq 25\%$ konnten Patienten in der Placebo-Gruppe zur offenen Ruxolitinib-Therapie wechseln. Vor Woche 24 musste zusätzlich auch eine definierte Symptomverschlechterung vorliegen. Patienten in der Ruxolitinib-Gruppe schieden vor Woche 24 in diesem Fall aus der Studie aus, danach konnten sie unter Fortsetzung der Ruxolitinib-Therapie in der Studie verbleiben. Nach der primären Datenanalyse im November 2010 wurde die Studie entblindet und alle Patienten konnten unabhängig vom Milzvolumen in den Ruxolitinib-Arm der Studie wechseln. Weiterbehandlung ursprünglich bis zu Woche 144 geplant; durch Amendment 4 (September 2012) wurde die Studiendauer auf 5 Jahre verlängert <u>Primäre Datenanalyse:</u> Nachdem alle Patienten 24 Wochen und 50% der Patienten 36 Wochen behandelt worden waren
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Das Studienprotokoll wurde insgesamt viermal geändert. Bis auf Amendment 4 erfolgten alle Änderungen vor Studienbeginn. <u>Amendment 4 vom 24. September 2012:</u> - Verlängerung der Studiendauer von 144 Wochen auf 5 Jahre, um Langzeitdaten zur Wirksamkeit und Sicherheit der Therapie mit Ruxolitinib erheben zu können

Item ^a	Charakteristikum	Studieninformation
4	Probanden / Patienten	Erwachsene Patienten mit primärer oder sekundärer Myelofibrose (PMF, PPV-MF, PET-MF) mit Intermediär-2- oder Hochrisikoprofil, die refraktär, resistent oder intolerant gegenüber anderen Therapieoptionen waren oder aus anderen Gründen nicht für eine Behandlung mit diesen in Betracht kamen
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> 1. Erwachsene Patienten (≥ 18 Jahre alt) 2. Diagnose einer PMF, PPV-MF oder PET-MF nach den WHO-Kriterien, unabhängig vom JAK2-Mutationsstatus 3. Intermediär-2- bzw. Hochrisikoprofil entsprechend den IWG-MRT-Kriterien, d.h. Vorhandensein von 2 bzw. ≥ 3 der folgenden Risikofaktoren - Alter ≥ 65 Jahre - konstitutionelle Symptome (Gewichtsverlust, Fieber, Nachtschweiß) - Hämoglobin $< 10\text{g/dl}$ - Leukozyten $> 25 \times 10^9/\text{l}$ - Blasten $\geq 1\%$ 4. Patienten, die refraktär, resistent oder intolerant gegenüber anderen Therapieoptionen sind oder nach Einschätzung des Prüfarztes aus anderen Gründen nicht für eine Behandlung mit diesen in Betracht kommen 5. Eastern Cooperative Oncology Group (ECOG)-Status ≤ 3 6. Milz tastbar mit ≥ 5 cm unter dem Rippenbogen 7. Blasten im peripheren Blut $< 10\%$ 8. CD34+ Zellen $> 20 \times 10^6/\text{l}$ 9. Keine Vorbehandlung mit einem JAK-Inhibitor <u>Ausschlusskriterien:</u> 1. Lebenserwartung weniger als 6 Monate 2. Schwangere bzw. stillende Patientinnen 3. Patienten im fortpflanzungsfähigen Alter, die keine entsprechenden Vorsichtsmaßnahmen treffen wollen, um die Zeugung eines Kindes bzw. eine Schwangerschaft zu verhindern 4. Inadäquate Knochenmarkfunktion, definiert als - absolute Neutrophilenzahl (ANC) $\leq 1000/\mu\text{l}$ - Thrombozytenzahl $< 100.000/\mu\text{l}$ ohne Gabe von Wachstumsfaktoren bzw. Transfusionen 5. Vorgeschichte einer Thrombozytenzahl $< 50.000/\mu\text{l}$ bzw. einer ANC $< 500/\mu\text{l}$, außer wenn diese während der Behandlung einer myeloproliferativen Erkrankung oder während einer Chemotherapie auftrat 6. Inadäquate Leber- oder Nierenfunktion: - direktes Bilirubin $> 2,0 \times \text{ULN}$ (upper limit of normal) - Alaninaminotransferase (ALT) $> 2,5 \times \text{ULN}$ - Kreatinin $> 2,0 \text{ mg/dl}$ 7. Behandlungsbedürftige, klinisch signifikante bakterielle,

Item ^a	Charakteristikum	Studieninformation
		fungale, parasitäre oder virale Infektion 8. Vorgeschichte einer malignen Erkrankung in den letzten 5 Jahren mit Ausnahme eines behandelten Plattenepithel- bzw. Basalzellkarzinoms im frühen Stadium 9. Kardiale Erkrankungen, die nach Einschätzung des Prüfarztes die Sicherheit des Patienten oder seine Compliance beeinträchtigen können 10. Patienten mit unkontrollierter oder instabiler Angina pectoris 11. Patienten mit Vorhofflimmern (auch paroxysmal) 12. Patienten mit Myokardinfarkt oder akutem Koronarsyndrom (innerhalb der letzten 6 Monate) 13. Ruhepuls > 110 Schläge/Minute 14. Systolischer Blutdruck > 160 mmHg 15. Diastolischer Blutdruck > 100 mmHg 16. Patienten mit Milzbestrahlung innerhalb von 12 Monaten vor Studienbeginn 17. Zeitgleiche Behandlung mit folgenden Medikamenten: - andere Medikamente, die sich noch in der klinischen Prüfung befinden - Induktoren des Isoenzym CYP3A4, Rifampicin oder Johanniskraut - Hydroxycarbamid, Interferon, Thalidomid, Busulfan, Lenalidomid oder Anagrelid - Hämatopoetische Wachstumsfaktoren bzw. Agonisten an deren Rezeptoren (z.B. Erythropoetin, G-CSF, Romiplostim, Eltrombopag) 18. Unfähigkeit zur Einwilligungserklärung 19. Alkohol- oder Drogenabhängigkeit, die die Fähigkeit, den Vorgaben des Studienprotokolls zu folgen, beeinträchtigt 20. Patienten mit irgendeiner Erkrankung, die nach Einschätzung des Prüfarztes die Sicherheit des Patienten oder seine Compliance beeinträchtigen kann
4b	Studienorganisation und Ort der Studiendurchführung	Die Studie wurde in 68 Zentren in den USA, 6 Zentren in Kanada und 15 Zentren in Australien durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<u>Interventionsgruppe (N = 155):</u> Ruxolitinib zur oralen Einnahme in Abhängigkeit von der Thrombozytenzahl zu Studienbeginn: - Thrombozytenzahl > 200.000/µl: Anfangsdosis 20 mg 2x täglich (je vier 5 mg-Tabletten) - Thrombozytenzahl > 100.000/µl und ≤ 200.000/µl: Anfangsdosis 15 mg 2x täglich (je drei 5 mg-Tabletten) Dosisanpassungen bzw. Therapieunterbrechungen aufgrund fehlender Wirksamkeit oder Unverträglichkeit gemäß Studienprotokoll erlaubt Höchstosis: 25 mg 2x täglich (50 mg/Tag) <u>Kontrollgruppe (N = 154):</u> Placebo 2x täglich je 3 oder 4 Tabletten

Item ^a	Charakteristikum	Studieninformation
		Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<u>Das primäre Zielkriterium war:</u> 1. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen - Rückgang des Milzvolumens im Vergleich zur Baseline gemessen mittels Magnetresonanztomographie (MRT) bzw. Computertomographie (CT) - radiologische Untersuchung durch lokale Studienärzte; Auswertung der Aufnahmen durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war <u>Die wichtigsten sekundären Zielkriterien waren:</u> 1. Dauer des Ansprechens (Milzvolumenreduktion) - Zeit vom ersten Erreichen einer Milzvolumenreduktion von $\geq 35\%$ bis zum Verlust dieses Ansprechens bei gleichzeitiger Zunahme des Milzvolumens um $> 25\%$ im Vergleich zum niedrigsten Wert während der Studie - falls kein Ereignis auftrat, erfolgte eine Zensurierung zum Zeitpunkt der letzten Milzuntersuchung 2. Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) um $\geq 50\%$ nach 24 Wochen - ermittelt mit Hilfe des modifizierten MFSAF v2.0 (Myelofibrosis Symptom Assessment Form) in Form eines elektronischen Patiententagebuchs 3. Prozentuale/absolute Veränderung des Total Symptom Score (TSS) gegenüber Studienbeginn nach 24 Wochen - ermittelt mit Hilfe des modifizierten MFSAF v2.0 (Myelofibrosis Symptom Assessment Form) in Form eines elektronischen Patiententagebuchs 4. Gesamtüberleben (overall survival, OS) - Zeit von der Randomisierung bis zum Tod durch jedwede Ursache <u>Erhebungszeitpunkte</u> Analyse des Milzvolumens durch MRT bzw. CT: Baseline, nach 12, 24, 36, 48, 60 und 72 Wochen sowie danach alle 24 Wochen Erfassung der Myelofibrose-Symptome mit Hilfe des MFSAF-Tagebuchs: täglich bis zu Woche 24 Erhebung der Daten zur leukämischen Transformation sowie zum Gesamtüberleben auch nach dem Ausscheiden aus der Studie alle 6 Monate
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	Geplante Rekrutierungszahl: 240 Zahl der tatsächlich rekrutierten Patienten: 309

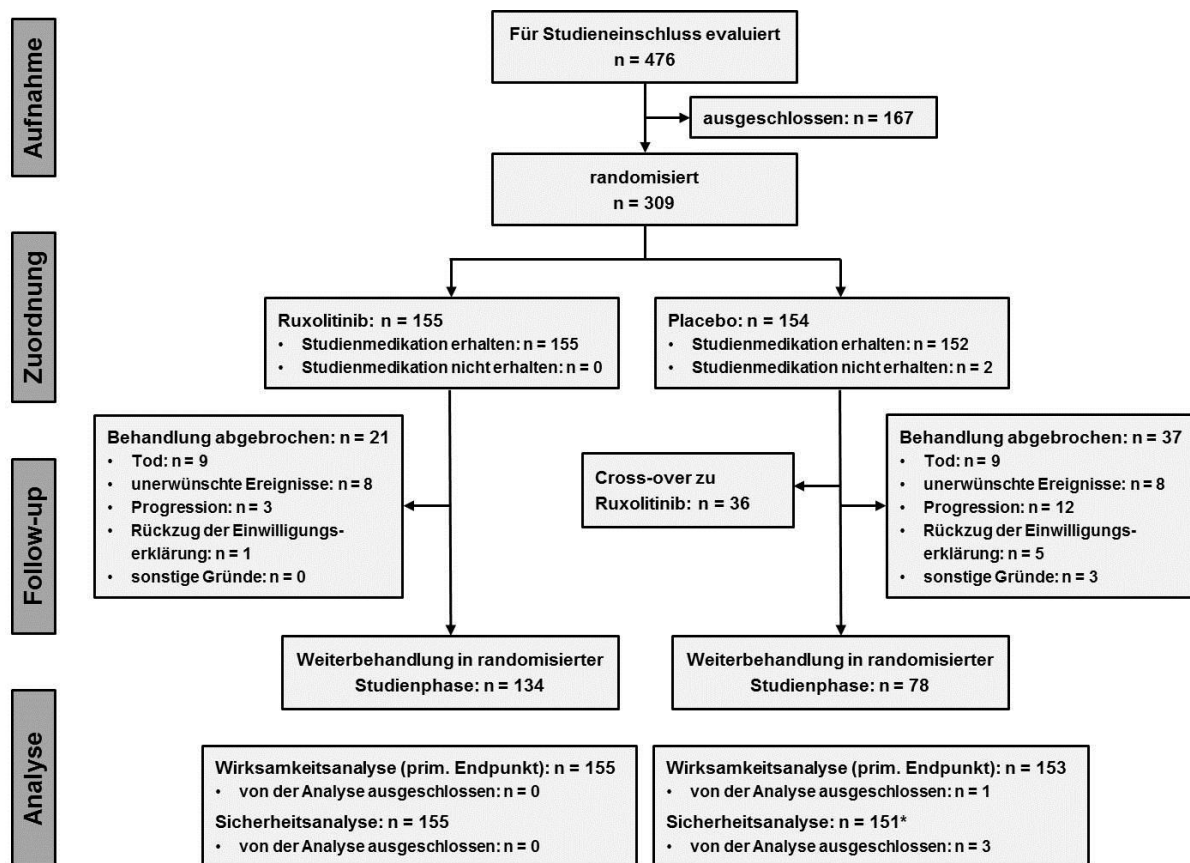
Item ^a	Charakteristikum	Studieninformation
		(nach einer Randomisierung im Verhältnis 1:1 erhielten 155 Patienten Ruxolitinib und 154 Patienten Placebo)
7a	Wie wurden die Fallzahlen bestimmt?	Primärer Endpunkt der Studie war der Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen. Unter Annahme einer Ansprechrate von mindestens 30% in der Interventionsgruppe bzw. 10% in der Kontrollgruppe ergab eine Fallzahl von 240 Patienten (je 120 Patienten/Gruppe) im Chi ² -Test bei einem Signifikanzniveau von 0,05 eine Power von 97%. Aufgrund eines plötzlichen Anstiegs von Patienten im Screening, denen die Studienteilnahme nicht verweigert werden sollte, überstieg die Zahl der tatsächlich randomisierten Patienten mit 309 die geplante Fallzahl.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte mit Hilfe eines IVRS (Interactive Voice Response System).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte im Verhältnis 1:1 (keine 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	Vor Therapiebeginn wurde jedem Patienten mit Hilfe des IVRS eine Nummer zugeteilt, durch die der Patient einem der beiden Studienarme zugewiesen wurde. Mit Hilfe dieser Nummer erfolgte auch die Zuteilung der ersten Studienmedikation, die ebenfalls durch die entsprechende Nummer gekennzeichnet 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?	Die Randomisierung erfolgte mit Hilfe eines IVRS-Systems. Nach Verifizierung aller verfügbaren Anforderungen kontaktierte das Studienpersonal des Zentrums das IVRS, durch das der Patient einem der beiden Studienarme zugewiesen wurde.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet. Lediglich Mitglieder des DSMB (Data Safety Monitoring Board) hatten für die regelmäßige Evaluation der Sicherheitsdaten Zugriff auf die Behandlungscodes. Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer.
11b	Falls relevant, Beschreibung der	Ruxolitinib- und Placebo-Tabletten waren von identischem

Item ^a	Charakteristikum	Studieninformation
	Ähnlichkeit von Interventionen	Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Es wurden folgende Analysepopulationen festgelegt: 1. Intention-to-treat (ITT)-Population - Auswertung aller randomisierten Patienten entsprechend dem Intention-to-treat (ITT)-Prinzip 2. Per-Protocol (PP)-Population - Auswertung aller Patienten der ITT-Population, bei denen es zu keiner relevanten Protokollverletzung gekommen war 3. Sicherheitspopulation - Auswertung aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten <u>Statistische Methoden zur Bewertung relevanter Zielkriterien:</u> 1. Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen - Vergleich der Ansprechraten mit Hilfe eines exakten Fisher-Tests mit einem zweiseitigen Signifikanzniveau von 0,05 - nur Patienten mit auswertbarer Milzvolumen-Messung zu Studienbeginn wurden in die Analyse einbezogen - Patienten ohne Milzvolumen-Bestimmung nach 24 Wochen wurden als Non-Responder klassifiziert 2. Dauer des Ansprechens (Milzvolumenreduktion) - Schätzung der Dauer des Ansprechens mit Hilfe der Kaplan-Meier-Methode (nur Responder) 3. Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) um $\geq 50\%$ nach 24 Wochen - Vergleich der Ansprechraten mit Hilfe eines χ^2-Tests mit einem zweiseitigen Signifikanzniveau von 0,05 - Patienten ohne TSS-Bestimmung nach 24 Wochen wurden als Non-Responder klassifiziert 4. Prozentuale/absolute Veränderung des Total Symptom Score (TSS) gegenüber Studienbeginn nach 24 Wochen - Vergleich der medianen Änderung gegenüber dem Ausgangswert mit Hilfe eines Wilcoxon-Rangsummen-Tests 5. Gesamtüberleben (overall survival, OS) - Schätzung des Gesamtüberlebens mit Hilfe der Kaplan-Meier-Methode - Vergleich der Kaplan-Meier-Kurven zwischen den Gruppen mittels eines Logrank-Tests und Berechnung der Hazard Ratios sowie zugehörigen 95%-Konfidenzintervalle mit einem Cox-Regressionsmodell
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<u>Subgruppenanalysen:</u> Zur Überprüfung der Konsistenz des Behandlungseffekts wurden für den primären Endpunkt Subgruppenanalysen durchgeführt.

Item ^a	Charakteristikum	Studieninformation
		Dabei wurden jeweils die Ansprechraten angegeben und die Odds Ratios sowie die zugehörigen 95%-Konfidenzintervalle mit Hilfe eines logistischen Regressionsmodells berechnet. Die Analyse des Zielkriteriums erfolgte dabei getrennt in folgenden a priori definierten Subgruppen: - Geschlecht (männlich, weiblich) - Alter (≤ 65 Jahre, > 65 Jahre) - Myelofibrose-Typ (PMF, PPV-MF, PET-MF) - Vorbehandlung mit Hydroxycarbamid in den letzten 3 Monaten vor Studienbeginn (ja, nein) - Vorliegen der V617F-Mutation zu Studienbeginn (ja, nein) - Milzvolumen zu Studienbeginn ($\leq$ Median, $>$ Median) - tastbare Milzlänge zu Studienbeginn (≤ 10 cm, > 10 cm) - Land (USA vs. Kanada und Australien) - Hautfarbe (Weiße, Nicht-Weiße) - Ethnie (Latino, Nicht-Latino) <u>Sensitivitätsanalysen:</u> Zur Einschätzung der Robustheit der Ergebnisse wurden verschiedene Sensitivitätsanalysen durchgeführt. So wurde z. B. der primäre Endpunkt „Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen“ sowie der sekundäre Endpunkt „Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) um $\geq 50\%$ nach 24 Wochen“ auch unter Berücksichtigung von Patienten mit fehlendem Baseline-Wert ausgewertet. Die Schätzung des Gesamtüberlebens erfolgte neben der ITT-Analyse zusätzlich auch unter Zensurierung von Patienten zum Zeitpunkt des Cross-overs in den Ruxolitinib-Arm. <u>Adjustierung für multiples Testen:</u> Die Analyse der sekundären Zielkriterien erfolgte nur bei Nachweis eines signifikanten Unterschieds in Bezug auf den primären Endpunkt. Die sekundären Zielkriterien wurden mit Hilfe einer hierarchischen Testprozedur in der folgenden Reihenfolge ausgewertet (jeweils mit einem Signifikanzniveau von 0,05): 1. Anteil der Patienten mit einer Verbesserung des Total Symptom Score (TSS) um $\geq 50\%$ nach 24 Wochen 2. Prozentuale/absolute Veränderung des Total Symptom Score (TSS) gegenüber Studienbeginn nach 24 Wochen 3. Gesamtüberleben
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,	a) Anzahl der randomisierten Patienten - Ruxolitinib: n = 155 - Placebo: n = 154

Item ^a	Charakteristikum	Studieninformation
	b) tatsächlich die geplante Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	b) Anzahl der tatsächlich behandelten Patienten: - Ruxolitinib: n = 155 - Placebo: n = 152 2 Patienten aus dem Placebo-Arm erhielten keine Dosis der Studienmedikation c) Anzahl der bei der Analyse des primären Zielkriteriums berücksichtigten Patienten: - Ruxolitinib: n = 155 - Placebo: n = 153 1 Patient aus der Placebo-Gruppe wurde aufgrund fehlender Milzvolumen-Messung zu Studienbeginn nicht in die Analyse des primären Zielkriteriums eingeschlossen
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Die primäre Datenanalyse fand am 2. November 2010 statt, nachdem alle Patienten 24 Wochen und 50% der Patienten 36 Wochen behandelt worden waren. Ein 2-Jahres-Update wurde am 1. März 2012 durchgeführt. Ein 3-Jahres-Update erfolgte am 25. Januar 2013, nachdem alle Patienten 144 Wochen behandelt worden waren. Anzahl der Patienten, die die Therapie in der randomisierten Phase beendet hatten (Sicherheitspopulation): <u>Primäranalyse vom 2. November 2010:</u> 1. Ruxolitinib: 21 von 155 (13,5%) - Tod: n = 9 - unerwünschte Ereignisse: n = 8 - Krankheitsprogression: n = 3 - Rückzug der Einwilligungserklärung: n = 1 - sonstige Gründe: n = 0 2. Placebo: 73 von 151 (48,3%) - Cross-over in den Ruxolitinib-Arm (laut Studienprotokoll nach Krankheitsprogression erlaubt): n = 36 - Tod: n = 9 - unerwünschte Ereignisse: n = 8 - Krankheitsprogression: n = 12 - Rückzug der Einwilligungserklärung: n = 5 - sonstige Gründe: n = 3 <u>2-Jahres-Update vom 1. März 2012:</u> 1. Ruxolitinib: 55 von 155 (35,5%) - Tod: n = 13 - unerwünschte Ereignisse: n = 11 - Krankheitsprogression: n = 12 - Rückzug der Einwilligungserklärung: n = 9 - sonstige Gründe: n = 10 2. Placebo: 151 von 151 (100%) - Cross-over in den Ruxolitinib-Arm: n = 111 - Tod: n = 10

Item ^a	Charakteristikum	Studieninformation
		- unerwünschte Ereignisse: n = 9 - Krankheitsprogression: n = 12 - Rückzug der Einwilligungserklärung: n = 6 - sonstige Gründe: n = 3 <u>3-Jahres-Update vom 25. Januar 2013:</u> 1. Ruxolitinib: 78 von 155 (50,3%) - Tod: n = 1 - unerwünschte Ereignisse: n = 29 - Krankheitsprogression: n = 18 - Rückzug der Einwilligungserklärung: n = 12 - sonstige Gründe: n = 18 2. Placebo: 151 von 151 (100%) - Cross-over in den Ruxolitinib-Arm: n = 111 - Tod: n = 0 - unerwünschte Ereignisse: n = 16 - Krankheitsprogression: n = 13 - Rückzug der Einwilligungserklärung: n = 7 - sonstige Gründe: n = 4 Anmerkung: 2 Patienten unter Placebo erhielten keine Dosis der Studienmedikation, die Daten eines weiteren Patienten gingen beim Umzug des Prüfzentrums verloren (insgesamt wurden deshalb im Placebo-Arm 3 Patienten nicht in die Sicherheits-Analyse eingeschlossen)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Der Einschreibungszeitraum war zwischen August 2009 und April 2010. Die Primäranalyse fand am 2. November 2010 statt. Zu diesem Zeitpunkt waren alle Patienten 24 Wochen und 50% der Patienten 36 Wochen behandelt worden, das mediane Follow-up betrug 32 Wochen. Beim 2-Jahres-Update am 1. März 2012 betrug das mediane Follow-up 102 Wochen. Ein 3-Jahres-Update erfolgte am 25. Januar 2013. Zu diesem Zeitpunkt waren alle Patienten 144 Wochen behandelt worden, das mediane Follow-up betrug 149 Wochen. Die geplante Studiendauer beträgt 5 Jahre.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie läuft noch und wird voraussichtlich im Januar 2015 beendet.
a: nach CONSORT 2010.		



* 2 Patienten erhielten keine Medikation; 1 Patient wurde von der Analyse ausgeschlossen, da seine Daten beim Umzug des Prüfzentrums verloren gingen

Abbildung 4-17: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (primäre Datenanalyse vom 2. November 2010)

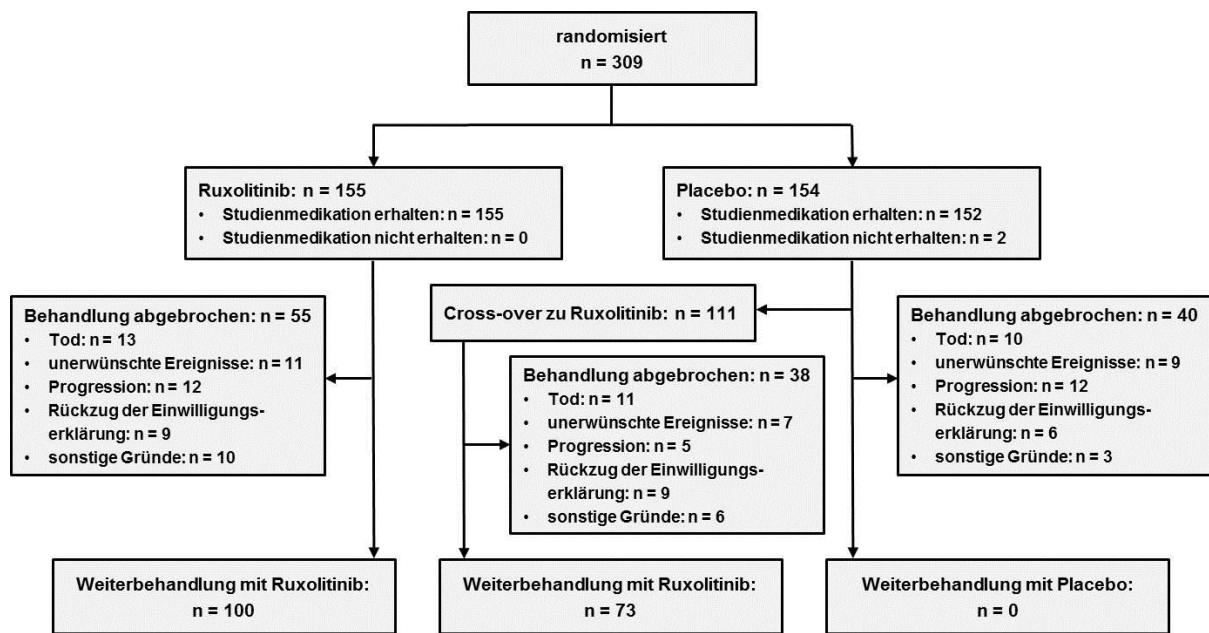


Abbildung 4-18: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (2-Jahres-Update vom 1. März 2012)

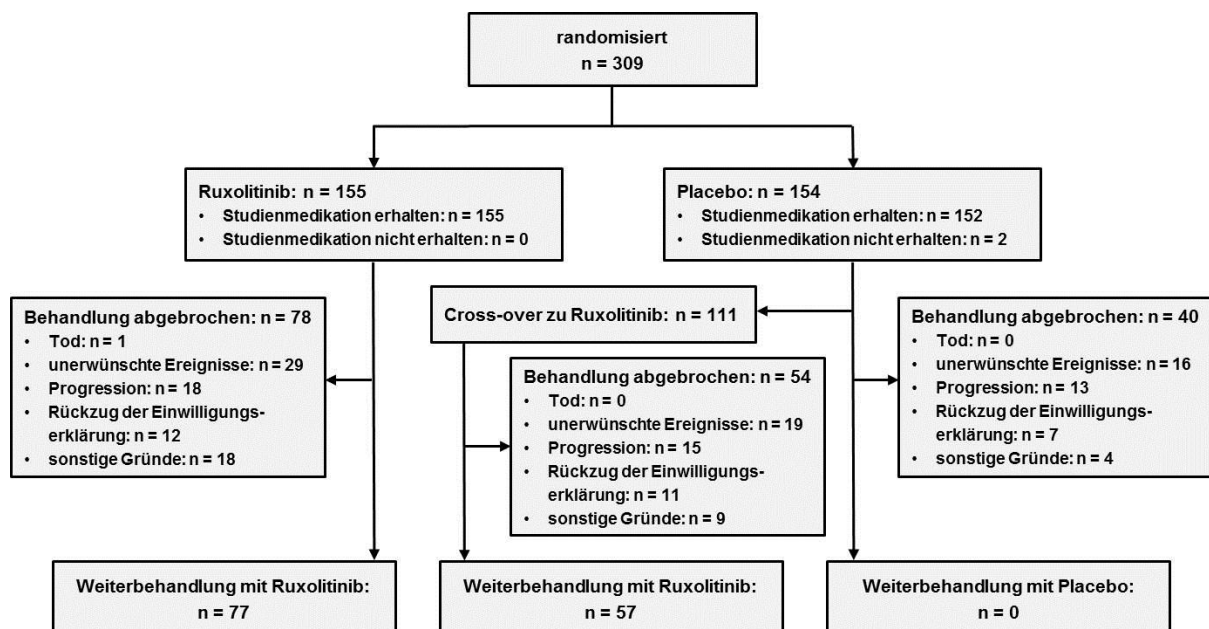


Abbildung 4-19: Flow-Chart zum Patientenfluss in der COMFORT-I-Studie (3-Jahres-Update vom 25. Januar 2013)

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. Quellenspezifische 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 4-70 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie COMFORT-II

Studie: COMFORT-II (CINC424A2352)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Oct 17.	(31)
Novartis Pharma GmbH. CINC424A2352: A randomized study of the JAK inhibitor INC424 tablets compared to best available therapy in subjects with primary myelofibrosis (PMF), post-polycythemia veramyelofibrosis (PPV-MF) or post-essential thrombocythemia-myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jun 19.	(62)
Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. N Engl J Med 2012 Mar 1;366(9):787-98.	(64)
Cervantes F, Vannucchi AM, Kiladjian JJ, Al-Ali HK, Sirulnik A, Stalbovskaya V, et al. Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3 study comparing ruxolitinib with best available therapy for myelofibrosis. Blood 2013 Dec 12;122(25):4047-53.	(63)
Harrison CN, Mesa RA, Kiladjian JJ, Al-Ali HK, Gisslinger H, Knoops L, et al. Health-related quality of life and symptoms in patients with myelofibrosis treated with ruxolitinib versus best available therapy. Br J Haematol 2013 Jul;162(2):229-39.	(65)
European Medicines Agency. European Public Assessment Report Jakavi. Procedure No.: EMEA/H/C/002464. 2012 Oct 4.	(20)

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:

Es liegt eine randomisierte kontrollierte Studie vor (20;31;62-64).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**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:**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:

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war (31).

behandelnde bzw. weiterbehandelnde Personen:

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war (31).

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Von einem relevanten Verzerrungspotenzial ist nicht auszugehen, da eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da alle wichtigen Endpunkte objektiv erfasst werden konnten und damit stabil gegenüber möglichen Verzerrungen sind. So wurde der Endpunkt „Reduktion der Milzgröße“ durch klar definierte Parameter erfasst, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann. Auch die Endpunkte „Gesamtüberleben“, „Leukämiefreies Überleben“ und „Progressionsfreies Überleben“ sind unabhängig von subjektiven Einflussfaktoren. Insgesamt ist das Verzerrungspotenzial für die COMFORT-II-Studie daher als „niedrig“ einzustufen.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**Endpunkt: Reduktion der Milzgröße – Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen****1. Verblindung der Endpunkterheber**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war. Darüber hinaus wurden die Ergebnisse der Milzvolumen-Messung den Prüfarzten nur bei einer Zunahme des Milzvolumens um 25% über den Nadir mitgeteilt (31).

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 bzw. 48 Wochen“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Milzvolumen durch klar definierte Parameter erfasst wurde, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden und somit stabil gegenüber möglichen Verzerrungen sind. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann.

Endpunkt: Reduktion der Milzgröße – Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 bzw. 48 Wochen**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war. Darüber hinaus wurden die Ergebnisse der Milzvolumen-Messung den Prüfarzten nur bei einer Zunahme des Milzvolumens um 25% über den Nadir mitgeteilt (31).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 bzw. 48 Wochen“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Milzvolumen durch klar definierte Parameter erfasst wurde, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden und somit stabil gegenüber möglichen Verzerrungen sind. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann.

Endpunkt: Reduktion der Milzgröße – Dauer des Ansprechens**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war. Darüber hinaus wurden die Ergebnisse der Milzvolumen-Messung den Prüfarzten nur bei einer Zunahme des Milzvolumens um 25% über den Nadir mitgeteilt (31).

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☐ nein

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

Bei der vorliegenden Untersuchung handelt es sich um einen Sonderfall, bei dem eine Auswertung nach dem ITT-Prinzip nicht sinnvoll wäre. Die Frage nach der adäquaten Umsetzung des ITT-Prinzips ist somit nicht zutreffend.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig**☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Dauer des Ansprechens“ als „niedrig“ eingestuft, da weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Milzvolumen durch klar definierte Parameter erfasst wurde, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden und somit stabil gegenüber möglichen Verzerrungen sind. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann.

Endpunkt: Reduktion der Milzgröße – Zeit bis zum ersten Ansprechen**1. Verblindung der Endpunkterheber**☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. Allerdings erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen für die Analyse des Milzvolumens durch einen unabhängigen, zentralen Reviewer, der in Bezug auf die Behandlungszuteilung verblindet war. Darüber hinaus wurden die Ergebnisse der Milzvolumen-Messung den Prüfern nur bei einer Zunahme des Milzvolumens um 25% über den Nadir mitgeteilt (31).

2. Adäquate Umsetzung des ITT-Prinzips☐ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Bei der vorliegenden Untersuchung handelt es sich um einen Sonderfall, bei dem eine Auswertung nach dem ITT-Prinzip nicht sinnvoll wäre. Die Frage nach der adäquaten Umsetzung des ITT-Prinzips ist somit nicht zutreffend.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Zeit bis zum ersten Ansprechen“ als „niedrig“ eingestuft, da weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Milzvolumen durch klar definierte Parameter erfasst wurde, die nicht von der subjektiven Beurteilung des Prüfers beeinflusst werden und somit stabil gegenüber möglichen Verzerrungen sind. Zudem erfolgte die Auswertung der MRT- bzw. CT-Aufnahmen durch einen zentralen Reviewer, der gegenüber der Behandlungszuteilung verblindet war, so dass bei der Analyse der entsprechenden Zielgrößen von einer unverzerrten Bewertung ausgegangen werden kann.

Endpunkt: Gesamtüberleben**1. Verblindung der Endpunkterheber**☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Der Endpunkt „Gesamtüberleben“ ist jedoch unabhängig von subjektiven Einflussfaktoren, so dass bei der Analyse des Endpunkts von einer unverzerrten Bewertung ausgegangen werden kann.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesamtüberleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da der Endpunkt „Gesamtüberleben“ unabhängig von subjektiven Einflussfaktoren ist.

Endpunkt: Leukämiefreies Überleben**1. Verblindung der Endpunkterheber**☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Das Auftreten einer leukämischen Transformation wurde jedoch durch klar definierte Laborparameter erfasst, so dass bei der Analyse des Endpunkts von einer unverzerrten Bewertung ausgegangen werden kann.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Leukämiefreies Überleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da das Auftreten einer leukämischen Transformation durch klar definierte Laborparameter erfasst wurde und somit wie die Erhebung des Gesamtüberlebens unabhängig von subjektiven Einflussfaktoren ist.

Endpunkt: Progressionsfreies Überleben**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Die Analyse der Krankheitsprogression wurde jedoch durch klar definierte Laborparameter sowie die Auswertung von MRT- bzw. CT-Aufnahmen durch einen zentralen, verblindeten Reviewer erfasst, so dass bei der Analyse des Endpunkts von einer unverzerrten Bewertung ausgegangen werden kann.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig**☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Progressionsfreies Überleben“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als äußerst gering einzuschätzen, da die Analyse der Krankheitsprogression durch klar definierte Laborparameter bzw. die Auswertung von MRT- bzw. CT-Aufnahmen erfasst wurde und somit wie die Erhebung des Gesamtüberlebens unabhängig von subjektiven Einflussfaktoren ist.

Endpunkt: Veränderung des ECOG-Performance-Status – Veränderung des ECOG-Performance-Status gegenüber Studienbeginn**1. Verblindung der Endpunkterheber**☐ **ja**☐ **unklar**☒ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Allerdings erfolgte die Bewertung des ECOG-Performance-Status anhand objektiver Kriterien (z.B. „gehfähig“, „Selbstversorgung möglich“) durch den Prüfarzt, so dass von einer unverzerrten Bewertung ausgegangen werden kann.

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Veränderung des ECOG-Performance-Status – Veränderung des ECOG-Performance-Status gegenüber Studienbeginn“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, bei der Analyse der entsprechenden Zielgrößen kann jedoch von einer unverzerrten Bewertung ausgegangen werden, da die Erhebung des ECOG-Performance-Status anhand objektiver Kriterien (z. B. „gehfähig“, „Selbstversorgung möglich“) durch den Prüfarzt erfolgte.

Endpunkt: Veränderung des ECOG-Performance-Status – Zeit bis zur Verschlechterung des ECOG-Performance-Status**1. Verblindung der Endpunkterheber**☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt. Allerdings erfolgte die Bewertung des ECOG-Performance-Status anhand objektiver Kriterien (z.B. „gehfähig“, „Selbstversorgung möglich“) durch den Prüfarzt, so dass von einer unverzerrten Bewertung ausgegangen werden kann.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Veränderung des ECOG-Performance-Status – Zeit bis zur Verschlechterung des ECOG-Performance-Status“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, bei der Analyse der entsprechenden Zielgrößen kann jedoch von einer unverzerrten Bewertung ausgegangen werden, da die Erhebung des ECOG-Performance-Status anhand objektiver Kriterien (z. B. „gehfähig“, „Selbstversorgung möglich“) durch den Prüfarzt erfolgte.

Endpunkt: Gesundheitsbezogene Lebensqualität – Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesundheitsbezogene Lebensqualität – Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn“ als „hoch“ eingestuft, da eine relevante Verzerrung der Ergebnisse aufgrund des offenen Studiendesigns nicht ausgeschlossen werden kann.

Endpunkt: Gesundheitsbezogene Lebensqualität – Veränderung des FACT-Lym gegenüber Studienbeginn**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☐ niedrig ☒ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesundheitsbezogene Lebensqualität – Veränderung des FACT-Lym gegenüber Studienbeginn“ als „hoch“ eingestuft, da eine relevante Verzerrung der Ergebnisse aufgrund des offenen Studiendesigns nicht ausgeschlossen werden kann.

Endpunkt: Verträglichkeit – Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Da alle Patienten mindestens eine Dosis der Studienmedikation erhalten hatten, entspricht die Sicherheitspopulation in diesem Fall der ITT-Population (31;62-64).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig**☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Unerwünschte Ereignisse“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74).

Endpunkt: Verträglichkeit – Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)**1. Verblindung der Endpunkterheber**☐ **ja**☐ **unklar**☒ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Da alle Patienten mindestens eine Dosis der Studienmedikation erhalten hatten, entspricht die Sicherheitspopulation in diesem Fall der ITT-Population (31;62-64).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74).

Endpunkt: Verträglichkeit – Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Da alle Patienten mindestens eine Dosis der Studienmedikation erhalten hatten, entspricht die Sicherheitspopulation in diesem Fall der ITT-Population (31;62-64).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Schwerwiegende unerwünschte Ereignisse“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74).

Endpunkt: Verträglichkeit – Unerwünschte Ereignisse, die zum Therapieabbruch führten**1. Verblindung der Endpunkterheber**☐ ja ☐ unklar ☒ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Da alle Patienten mindestens eine Dosis der Studienmedikation erhalten hatten, entspricht die Sicherheitspopulation in diesem Fall der ITT-Population (31;62-64).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Unerwünschte Ereignisse, die zum Therapieabbruch führten“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74).

Endpunkt: Verträglichkeit – Häufige unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

☐ ja ☐ unklar ☒ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie (31). Aufgrund des Einsatzes verschiedener Einzeltherapien im Kontrollarm (Best Available Therapy) mit unterschiedlichen Darreichungsformen, Dosierungen und Applikationsarten war eine Verblindung der Studie nicht möglich. Die Art der Studienmedikation war sowohl Patienten als auch Prüfern bekannt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Da alle Patienten mindestens eine Dosis der Studienmedikation erhalten hatten, entspricht die Sicherheitspopulation in diesem Fall der ITT-Population (31;62-64).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig**☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Häufige unerwünschte Ereignisse“ als „niedrig“ eingestuft, da die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen. Aufgrund des Einsatzes verschiedener Medikamente im Kontrollarm (Best Available Therapy) war eine Verblindung von Patienten und Behandlern zwar nicht durchführbar, das Verzerrungspotenzial durch die fehlende Verblindung ist jedoch als gering einzuschätzen, da zur Erhebung unerwünschter Ereignisse größten Teils objektive Kriterien herangezogen werden, wie zum Beispiel bestimmte Messparameter zur Erhebung von unerwünschten Ereignissen oder festgelegte Definitionen eines schwerwiegenden Ereignisses (74).

Tabelle 4-71 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie COMFORT-I

Studie: COMFORT-I (INCB 18424-351)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Primary Analysis). 2011 Apr 27	(23)
Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 1). 2011 May 4.	(55)
Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (Addendum 2). 2011 Oct 31.	(56)
Incyte Corporation. INCB 18424-351: A Randomized, Double-Blind, Placebo-Controlled Study of the JAK Inhibitor INCB018424 Tablets Administered Orally to Subjects With Primary Myelofibrosis (PMF), Post-Polycythemia Vera-Myelofibrosis (PPV-MF), Or Post-Essential Thrombocythemia-Myelofibrosis (PET-MF). Clinical Study Report (144-weeks follow-up). 2013 Jul 24.	(24)
Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. N Engl J Med 2012 Mar 1;366(9):799-807.	(22)
Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Efficacy, safety and survival with ruxolitinib treatment in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. Haematologica 2013 Dec;98(12):1865-71.	(59)
Mesa RA, Gotlib J, Gupta V, Catalano JV, Deininger MW, Shields AL, et al. Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: a randomized, double-blind, placebo-controlled trial. J Clin Oncol 2013 Apr 1;31(10):1285-92.	(16)
European Medicines Agency. European Public Assessment Report Jakavi. Procedure No.: EMEA/H/C/002464. 2012 Oct 4.	(20)

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:

Es liegt eine randomisierte kontrollierte Studie vor (20;22-24;59).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

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:

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:

Alle in die Studie involvierten Personen (Patienten, Prüfer, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

behandelnde bzw. weiterbehandelnde Personen:

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Alle in die Studie involvierten Personen (Patienten, Prüfer, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für die COMFORT-I-Studie als „niedrig“ eingestuft, da eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte, sowohl Patient als auch Behandler verblindet waren und es weder Hinweise auf ergebnis-gesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**Endpunkt: Reduktion der Milzgröße – Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen****1. Verblindung der Endpunkterheber**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüffärzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Anteil der Patienten mit einer Reduktion des Milzvolumens um $\geq 35\%$ nach 24 Wochen“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Reduktion der Milzgröße – Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 Wochen**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Prozentuale Veränderung des Milzvolumens gegenüber Studienbeginn nach 24 Wochen“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Reduktion der Milzgröße – Dauer des Ansprechens**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☐ nein

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

Bei der vorliegenden Untersuchung handelt es sich um einen Sonderfall, bei dem eine Auswertung nach dem ITT-Prinzip nicht sinnvoll wäre. Die Frage nach der adäquaten Umsetzung des ITT-Prinzips ist somit nicht zutreffend.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Dauer des Ansprechens“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Reduktion der Milzgröße – Zeit bis zum ersten Ansprechen**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☐ nein

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

Bei der vorliegenden Untersuchung handelt es sich um einen Sonderfall, bei dem eine Auswertung nach dem ITT-Prinzip nicht sinnvoll wäre. Die Frage nach der adäquaten Umsetzung des ITT-Prinzips ist somit nicht zutreffend.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Reduktion der Milzgröße – Zeit bis zum ersten Ansprechen“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Linderung der Krankheitssymptomatik – Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen**1. Verblindung der Endpunkterheber**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Linderung der Krankheitssymptomatik – Anteil der Patienten mit einer Verbesserung des TSS um $\geq 50\%$ nach 24 Wochen“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Linderung der Krankheitssymptomatik – Prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Linderung der Krankheitssymptomatik – Prozentuale/absolute Veränderung des TSS sowie der Einzelsymptome des MFSAF gegenüber Studienbeginn nach 24 Wochen“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Gesamtüberleben**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesamtüberleben“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Leukämiefreies Überleben**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüffärzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Leukämiefreies Überleben“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Veränderung des ECOG-Performance-Status**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfer, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Veränderung des ECOG-Performance-Status“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Gesundheitsbezogene Lebensqualität – Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüffärzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig**☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Gesundheitsbezogene Lebensqualität – Veränderung des EORTC-QLQ-C30 gegenüber Studienbeginn“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend dem ITT-Prinzip erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Verträglichkeit – Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja**☐ **unklar**☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Dies war nur für 307 der ursprünglich randomisierten 309 Patienten der Fall: 155 der 155 Patienten in der Ruxolitinib-Gruppe und 152 der 154 Patienten in der Placebo-Gruppe. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (22-24;59).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Unerwünschte Ereignisse“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Verträglichkeit – Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Dies war nur für 307 der ursprünglich randomisierten 309 Patienten der Fall: 155 der 155 Patienten in der Ruxolitinib-Gruppe und 152 der 154 Patienten in der Placebo-Gruppe. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (22-24;59).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Schwere unerwünschte Ereignisse (CTCAE Grad 3 bzw. 4)“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Verträglichkeit – Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfer, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Dies war nur für 307 der ursprünglich randomisierten 309 Patienten der Fall: 155 der 155 Patienten in der Ruxolitinib-Gruppe und 152 der 154 Patienten in der Placebo-Gruppe. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (22-24;59).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Schwerwiegende unerwünschte Ereignisse“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Verträglichkeit – Unerwünschte Ereignisse, die zum Therapieabbruch führten**1. Verblindung der Endpunkterheber**☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips☒ **ja** ☐ **unklar** ☐ **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Dies war nur für 307 der ursprünglich randomisierten 309 Patienten der Fall: 155 der 155 Patienten in der Ruxolitinib-Gruppe und 152 der 154 Patienten in der Placebo-Gruppe. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (22-24;59).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**☒ ja ☐ neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Unerwünschte Ereignisse, die zum Therapieabbruch führten“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

Endpunkt: Verträglichkeit – Häufige unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Alle in die Studie involvierten Personen (Patienten, Prüfarzte, zentraler Reviewer der MRT/CT-Daten, Sponsor) blieben vom Zeitpunkt der Randomisierung bis zum Datenbankschluss für die Primäranalyse verblindet (23). Die Identifikation der Patienten sowie die Zuweisung der Studienmedikation erfolgten ausschließlich über die entsprechende Randomisierungsnummer. Ruxolitinib- und Placebo-Tabletten waren von identischem Aussehen. Auch im Placebo-Arm wurden „Dosisanpassungen“ durchgeführt, um die Verblindung aufrecht zu erhalten (23).

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

Die Auswertung erfolgte entsprechend internationalen Standards anhand der Sicherheitspopulation, d.h. aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Dies war nur für 307 der ursprünglich randomisierten 309 Patienten der Fall: 155 der 155 Patienten in der Ruxolitinib-Gruppe und 152 der 154 Patienten in der Placebo-Gruppe. Außerdem gingen die Daten eines Patienten aus der Placebo-Gruppe beim Umzug des Prüfzentrums verloren, weshalb nur 151 Patienten aus der Placebo-Gruppe in die Sicherheitsanalyse eingeschlossen wurden („Sicherheitspopulation“) (22-24;59).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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

☒ ja ☐ nein

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

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

☒ niedrig ☐ hoch

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „Verträglichkeit – Häufige unerwünschte Ereignisse“ als „niedrig“ eingestuft, da Patienten und Behandler verblindet waren, die Auswertung der Daten entsprechend internationalen Standards anhand der Sicherheitspopulation erfolgte und weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorlagen.

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 erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

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

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

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

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

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

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

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

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

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

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

2.

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

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

☐ **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.☐ **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

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

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine 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 waren verblindet.☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

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 waren.

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 erfolgte.

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 Gruppenzuteilung 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“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung

bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

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

Begründung für die Einstufung:
