

Dokumentvorlage, Version vom 20.01.2011

**Dossier zur Nutzenbewertung
gemäß § 35a SGB V**

Aflibercept (Zaltrap[®])

Sanofi-Aventis Deutschland GmbH

Modul 4 A

Metastasiertes Kolorektalkarzinom

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 01.03.2013

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	7
Abkürzungsverzeichnis	8
4.1 Zusammenfassung der Inhalte von Modul 4.....	11
4.2 Methodik.....	17
4.2.1 Fragestellung.....	17
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	17
4.2.3 Informationsbeschaffung.....	20
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	20
4.2.3.2 Bibliografische Literaturrecherche.....	20
4.2.3.3 Suche in Studienregistern.....	21
4.2.3.4 Selektion relevanter Studien.....	22
4.2.4 Bewertung der Aussagekraft der Nachweise.....	22
4.2.5 Informationssynthese und -analyse.....	25
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	25
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	25
4.2.5.3 Meta-Analysen.....	28
4.2.5.4 Sensitivitätsanalysen.....	29
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	30
4.2.5.6 Indirekte Vergleiche.....	32
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	34
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	35
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	35
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	35
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	36
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	38
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	38
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	39
4.3.1.2.1 Studiendesign und Studienpopulationen.....	39
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	44
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	44
4.3.1.3.1 Gesamtüberleben (OS) – RCT.....	45
4.3.1.3.2 Progressionsfreies Überleben (PFS) – RCT.....	50
4.3.1.3.3 Objektive Ansprechrates (ORR) – RCT.....	56
4.3.1.3.4 Sicherheit (Unerwünschte Ereignisse) – RCT.....	59
4.3.1.3.5 Subgruppenanalysen – RCT.....	64
4.3.1.3.5.1 Präspezifizierte Subgruppenanalysen.....	65
4.3.1.3.5.2 Post-hoc durchgeführte Analysen.....	82

4.3.1.3.6	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	84
4.3.1.3.6.1	Zusammenfassung der Ergebnisse zum Gesamtüberleben (OS).....	84
4.3.1.3.6.2	Zusammenfassung der Ergebnisse zum Progressionsfreien Überleben (PFS)	86
4.3.1.3.6.3	Zusammenfassung der Ergebnisse zur Objektiven Ansprechrate (ORR)	87
4.3.1.3.6.4	Zusammenfassung der Ergebnisse zur Sicherheit (Unerwünschte Ereignisse)	87
4.3.2	Weitere Unterlagen.....	89
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	89
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	89
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	90
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	90
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	90
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	93
4.3.2.2	Nicht randomisierte vergleichende Studien	93
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	93
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	94
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	94
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	94
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	95
4.3.2.3	Weitere Untersuchungen.....	95
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	96
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	96
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	96
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	96
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	97
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	97
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	98
4.4.1	Beurteilung der Aussagekraft der Nachweise	98
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	99
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	104
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	105
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	106
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	106
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	106
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	106
4.5.4	Verwendung von Surrogatendpunkten	107
4.6	Liste der eingeschlossenen Studien.....	108

4.7 Referenzliste.....	108
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	113
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	114
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund	125
Anhang 4-D : Liste der abgebrochenen Studien	126
Anhang 4-E : Liste der laufenden Studien.....	127
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....	128
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	137

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Aflibercept.....	18
Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	35
Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	36
Tabelle 4-4: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	38
Tabelle 4-5: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	39
Tabelle 4-6: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.	40
Tabelle 4-7: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	41
Tabelle 4-8: Charakterisierung der Studienpopulationen (Alter, Geschlecht, geographische Region) – RCT mit dem zu bewertenden Arzneimittel.	41
Tabelle 4-9: Charakterisierung der Studienpopulationen (vorbestehender Hypertonus, ECOG-PS, Vorbehandlung mit Bevacizumab) – RCT mit dem zu bewertenden Arzneimittel.....	42
Tabelle 4-10: Charakterisierung der Studienpopulationen (Zeit zwischen Erstdiagnose und Randomisierung, Anzahl Organe, die von Metastasen betroffen sind, primäre Lage des Tumors) – RCT mit dem zu bewertenden Arzneimittel	42
Tabelle 4-11: Charakterisierung der Studienpopulationen (ausschließlich Lebermetastasen, beteiligte metastasierte Organe) – RCT mit dem zu bewertenden Arzneimittel.....	42
Tabelle 4-12: Charakterisierung der Studienpopulationen (Therapieabbrecher, Anzahl der Patienten, die weitere antineoplastische Therapie erhalten haben) – RCT mit dem zu bewertenden Arzneimittel	43
Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	44
Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	45
Tabelle 4-15: Operationalisierung von Gesamtüberleben (OS).....	45
Tabelle 4-16: Bewertung des Verzerrungspotenzials für Gesamtüberleben (OS) in RCT mit dem zu bewertenden Arzneimittel	46
Tabelle 4-17: Ergebnisse für Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel.....	47
Tabelle 4-18: Ergebnisse für Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalysen	49

Tabelle 4-19: Operationalisierung von Progressionsfreies Überleben (PFS)	51
Tabelle 4-20: Bewertung des Verzerrungspotenzials für Progressionsfreies Überleben (PFS) in RCT mit dem zu bewertenden Arzneimittel	51
Tabelle 4-21: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel	53
Tabelle 4-22: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse.....	55
Tabelle 4-23: Operationalisierung von Objektive Ansprechrates (ORR).....	56
Tabelle 4-24: Bewertung des Verzerrungspotenzials für Objektive Ansprechrates (ORR) in RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-25: Ergebnisse für Objektive Ansprechrates (ORR)aus RCT mit dem zu bewertenden Arzneimittel	58
Tabelle 4-26: Operationalisierung von Sicherheit (Unerwünschte Ereignisse).....	60
Tabelle 4-27: Bewertung des Verzerrungspotenzials für Sicherheit (Unerwünschte Ereignisse) in RCT mit dem zu bewertenden Arzneimittel	60
Tabelle 4-28: Ergebnisse für Sicherheit (Unerwünschte Ereignisse) aus RCT mit dem zu bewertenden Arzneimittel – Sicherheitspopulation	61
Tabelle 4-29: Unerwünschte Ereignisse mit mehr als 5 Prozentpunkten Unterschied zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe	62
Tabelle 4-30: Präspezifizierte unerwünschte Ereignisse von besonderem Interesse. Vergleich zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe.....	63
Tabelle 4-31: Interaktionstests der präspezifizierten Subgruppenanalyse für den primären Endpunkt Gesamtüberleben (OS).....	66
Tabelle 4-32: Effektmodifikationen (p-Wert<0,2) in den präspezifizierten Subgruppen für Gesamtüberleben (OS)	67
Tabelle 4-33: Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung für Gesamtüberleben (OS) innerhalb der ITT-Population	72
Tabelle 4-34: Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung für Gesamtüberleben (OS) innerhalb der Population unter Ausschluss der adjuvant vorbehandelten Patienten	73
Tabelle 4-35: Interaktionstests der präspezifizierten Subgruppenanalyse für den sekundären Endpunkt Progressionsfreies Überleben (PFS).....	74
Tabelle 4-36: Effektmodifikationen (p-Wert<0,2) in den präspezifizierten Subgruppen für Progressionsfreies Überleben (PFS).....	75
Tabelle 4-37: Subgruppenanalyse der Bevacizumab-Vorbehandlung für Progressionsfreies Überleben (PFS) innerhalb der ITT-Population	80
Tabelle 4-38: Subgruppenanalyse der Bevacizumab-Vorbehandlung für Progressionsfreies Überleben (PFS) innerhalb der Population unter Ausschluss der adjuvant behandelten Patienten	81

Tabelle 4-39: Interaktionstest der post-hoc Subgruppenanalyse für die Bewertung der Sicherheit (Unerwünschte Ereignisse) anhand der Sicherheitspopulation.....	82
Tabelle 4-40: Effektmodifikationen (p-Wert<0,2) in den post-hoc analysierten Subgruppen für den Endpunkt Sicherheit (Unerwünschte Ereignisse).....	83
Tabelle 4-41: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	90
Tabelle 4-42: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	91
Tabelle 4-43: Operationalisierung von <Endpunkt xxx>.....	91
Tabelle 4-44: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	92
Tabelle 4-45: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	92
Tabelle 4-46: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	94
Tabelle 4-47: Operationalisierung von <Endpunkt xxx>.....	94
Tabelle 4-48: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	95
Tabelle 4-49: Operationalisierung von <Endpunkt xxx>.....	96
Tabelle 4-50: Zusammenfassung der Ergebnisse zum Zusatznutzen auf Endpunktebene: FOLFIRI+Aflibercept vs. FOLFIRI+Placebo.....	104
Tabelle 4-51: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	105
Tabelle 4-52: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	106
Tabelle 4-53: Für die Nutzenbewertung berücksichtigte Studien und Untersuchungen.....	108
Tabelle 4-54 (Anhang): Im International Clinical Trials Registry Platform Search Portal (ICTRP) identifizierte Studien	115
Tabelle 4-55 (Anhang): In clinicaltrials.gov identifizierte Studien	116
Tabelle 4-56 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	126
Tabelle 4-57 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	127
Tabelle 4-58 (Anhang): Studiendesign und -methodik für Studie EFC10262 (VELOUR)...	129
Tabelle 4-59 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie VELOUR	138

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	37
Abbildung 2: Kaplan-Meier Kurven für Gesamtüberleben (OS) (in Monaten) in der ITT-Population.....	48
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	50
Abbildung 4: Kaplan-Meier-Kurven für Progressionsfreies Überleben (PFS) (in Monaten) in der ITT-Population.....	54
Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	55
Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	59
Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	64
Abbildung 8: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für Stratifizierungsfaktoren.....	69
Abbildung 9: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für demographische Faktoren	70
Abbildung 10: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für Baseline-Eigenschaften	71
Abbildung 11: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS). Subgruppenanalyse der ITT-Population für die Stratifizierungsfaktoren	77
Abbildung 12: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS). Subgruppenanalyse der ITT-Population für demografische Faktoren	78
Abbildung 13: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS). Subgruppenanalyse der ITT-Population für Baseline-Eigenschaften	79
Abbildung 14: Patientenfluss in der Studie VELOUR (Quelle: [33]).	136

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AIDS	Acquired Immune Deficiency Syndrome
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
AMNOG	Arzneimittelmarktneuordnungsgesetz
CONSORT	Consolidated Standards of Reporting Trials
Cox PH-Modell	Cox Proportional-Hazard-Modell
CPMP	Committee for Proprietary Medicinal Products
CR	Komplettes Ansprechen (complete response)
CSR	Studienbericht (clinical study report)
CYP3A4	Cytochrom P450 3A4
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
dl-LV	Leucovorin, exprimiert als d/l-Racemat
DMC	Komitee für Datenüberwachung (Data Monitoring Committee)
ECOG-PS	Eastern Cooperative Oncology Group-Performance Status
EG	Europäische Gemeinschaft
EMA	European Medicines Agency
EP	Evaluierbare Population
EU	Europäische Union
FDA	U.S. Food and Drug Administration
FOLFIRI	FOLinsäure + 5-Fluorouracil + IRInotecan
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
HCG	Humanes Choriongonadotropin
HIV	Humaner Immundefizienz-Virus
HR	Hazard Ratio
HRQoL	Gesundheitsbezogene Lebensqualität
ICTRP	International Clinical Trials Registry Platform
INR	International Normalized Ratio
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRC	Independent Review Committee
ITT	Intention to treat

IVRS	Interactive Voice Response System
KI	Konfidenzintervall
K-Ras	Kirsten Rat Sarcoma 2 viral oncogene homolog
mCRC	Metastasiertes Kolorektalkarzinom (metastatic colorectal cancer)
mm Hg	Millimeter Quecksilbersäule
MTC	Mixed Treatment Comparison
NCI	National Cancer Institute
NCI-CTC	National Cancer Institute Common Terminology Criteria
NYHA	New York Heart Association
OGR	Obere Grenze des Referenzbereiches
OR	Odds Ratio
ORR	Objektive Ansprechrate (objective response rate)
OS	Gesamtüberleben (overall survival)
PFS	Progressionsfreies Überleben (progression free survival)
PR	Teilweises Ansprechen (partial response)
PRES	Posteriores reversibles Enzephalopathie-Syndrom
PS	Performance Status
RCT	Randomized Controlled Trial
RECIST	Response Evaluation Criteria in Solid Tumors
RMP	Risk-Management -Plan
RR	Relatives Risiko
SD	Standardabweichung
SGB	Sozialgesetzbuch
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
VEGF	Vaskulo-endothelialer Wachstumsfaktor (vascular endothelial growth factor)
VerfO	Verfahrensordnung
WHO	Weltgesundheitsorganisation (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.

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Ziel des vorliegenden Dossiers ist die Darstellung und Quantifizierung des Nutzens, insbesondere des Zusatznutzens nach § 35a SGB V des Wirkstoffs Aflibercept (Zaltrap®) in Kombination mit einem Chemotherapie regime bestehend aus 5-Fluorouracil (5-FU), Folinsäure und Irinotecan (FOLFIRI) in der Behandlung von Erwachsenen mit metastasiertem kolorektalem Karzinom (mCRC), das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist, gegenüber der zweckmäßigen Vergleichstherapie FOLFIRI. Der Zusatznutzen von Aflibercept soll anhand patientenrelevanter Endpunkte der Kategorien „Mortalität“ (Gesamtüberlebensrate, OS), „Morbidität“ (progressionsfreies Überleben, PFS; objektive Ansprechrate, ORR) und „Sicherheit“ (unerwünschte Arzneimittelwirkungen) durch Auswertung publizierter Daten aus randomisierten, kontrollierten klinischen Studien (RCT) ermittelt werden.

Datenquellen

Zur Beantwortung der Fragestellung werden die Daten einer randomisierten, kontrollierten Vergleichsstudie von Aflibercept in Kombination mit FOLFIRI gegen FOLFIRI alleine bei Patienten mit mCRC nach Vorbehandlung mit einem Oxaliplatin-haltigen Therapieregime (VELOUR) herangezogen [1]. Die genannte Studie lieferte auch die wissenschaftliche Basis für die Zulassung von Aflibercept in der genannten Indikation.

Ein-/Ausschlusskriterien für Studien

Zur Identifizierung der für die Nutzenbewertung von Aflibercept relevanten klinischen Studien wurde eine Studienregistersuche durchgeführt (siehe Abschnitt 4.2.3.3 und Anhang 4-B). Tabelle 4-1 in Abschnitt 4.2.2 gibt eine Übersicht über die in der Studienselektion angewandten prädefinierten Ein- und Ausschlusskriterien:

- Das Kriterium zur **Zielpopulation** richtet sich nach den Einschränkungen gemäß Zulassung für Aflibercept.
- Hinsichtlich der **Intervention** werden nur Studien berücksichtigt, in denen Aflibercept gemäß der Empfehlungen der Fachinformation gegeben wurde.
- Die zweckmäßige **Vergleichstherapie** ist der alleinige Einsatz einer Kombinations-Chemotherapie aus 5-FU, Folinsäure und Irinotecan (FOLFIRI).

- Die Ein- und Ausschlusskriterien hinsichtlich der in den in Frage kommenden Studien zu betrachtenden **Endpunktkategorien** orientieren sich direkt an den Vorgaben der „Allgemeinen Methoden (Ver. 4.0)“ des Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) für die Nutzenbewertung (siehe Abschnitte 4.2.1 und 4.2.5.2). Endpunkte der folgenden Kategorien wurden als adäquat angesehen:
 - Mortalität
 - Morbidität
 - (gesundheitsbezogene) Lebensqualität
 - Sicherheit
- **Studiendesign:** Es wurden nur RCT für die Zusatznutzenbewertung eingeschlossen, da diese die bestmögliche Evidenzsicherheit für einzelne Studien liefern.
- **Publikation:** Es wurden nur Vollpublikationen eingeschlossen, die Erste-Hand-Resultate berichten und den Anforderungen des CONSORT Statements genügen.

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

Das Verzerrungspotenzial der Ergebnisse wurde für die in die Nutzenbewertung eingeschlossene Studie bewertet. Dazu wurden endpunktübergreifende und endpunktspezifische Aspekte, die das Verzerrungspotenzial beeinflussen, in Analogie zur Vorgehensweise des IQWiG systematisch extrahiert und bewertet.

Bewertung des Verzerrungspotenzials

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte wurden untersucht. Hinsichtlich Verzerrungen auf Endpunktebene wurde jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Entsprechend der Verfahrensordnung (VerfO) des Gemeinsamen Bundesausschusses (G-BA) zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V vom 20. Januar 2011 [2] wurden dabei folgende Aspekte bewertet:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnisgesteuerte Berichterstattung

- Sonstige das Verzerrungspotential beeinflussende endpunktübergreifende Aspekte

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des Intention-To-Treat (ITT)-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte: Behandlung fehlender Werte und Drop-outs

Die Vorgehensweise bei der Bewertung des Verzerrungspotentials für die Studie und jeden Endpunkt (Operationalisierung) richtete sich nach den Hinweisen der Bewertungsbögen in Anhang 4-G.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die Bewertung des Nutzens und des Zusatznutzen erfolgte auf Grundlage der Zulassungsstudie VELOUR. Eingeschlossen in dieser Studie waren mit einem Oxaliplatin-basiertem Therapieregime vorbehandelte Patienten mit metastasiertem Kolorektalkarzinom, welche unter oder nach diesem Regime progredient waren und für die in diesem fortgeschrittenem Krankheitsstadium nach derzeitigem Behandlungsstandard keine kurativen Therapieoptionen mehr bestehen (siehe auch Modul 3 A, Abschnitt 3.2.2 des vorliegenden Dossiers). Die Patienten wurden randomisiert auf die Kombinations-Chemotherapie FOLFIRI plus Placebo oder FOLFIRI in Kombination mit Aflibercept. Die Randomisierung erfolgte stratifiziert nach Bevacizumab-Vorbehandlung und Eastern Cooperative Oncology Group-Performance Status (ECOG-PS). Die Therapie wurde in einem 2-wöchigen Zyklus über eine mediane Behandlungsdauer von insgesamt 9 Zyklen verabreicht.

Bei dem primären Endpunkt **Gesamtüberleben (OS)** konnte eine signifikante Verlängerung des Überlebens für die Patienten im FOLFIRI+Aflibercept-Arm gegenüber den Patienten im FOLFIRI+Placebo-Arm gezeigt werden ($p=0,0032$). Das mediane Überleben betrug bei der FOLFIRI+Aflibercept-Gruppe 13,50 Monate, bei der FOLFIRI+Placebo Gruppe lediglich 12,06 Monate; dies entspricht einer signifikanten Verlängerung des Gesamtüberlebens um 1,44 Monate. Im 75%-Quantil der Überlebenszeit vergrößerte sich die Differenz zwischen den Behandlungsarmen sogar auf 4,56 Monate zugunsten von FOLFIRI+Aflibercept. Dieses Ergebnis wird auch durch die zugehörigen Kaplan-Meier-Kurven bestätigt, die sich kontinuierlich weiter distanzieren. Nach 18 Monaten lag die Überlebenswahrscheinlichkeit in der FOLFIRI+Aflibercept-Gruppe bei 38,5%, in der FOLFIRI+Placebo-Gruppe dagegen bei 30,9%. Auch nach 24 und nach 30 Monaten zeigte sich die Überlegenheit der FOLFIRI+Aflibercept-Therapie gegenüber FOLFIRI+Placebo (24 Monate: 28% vs. 18,7%; 30 Monate: 22,3% vs. 12%). Das Hazard Ratio beträgt 0,817 (95%-Konfidenzintervall (KI):

[0,714;0,934], 95,34%-KI: [0,713;0,937]) zugunsten von FOLFIRI+Aflibercept, was einer 18,3%-igen Reduzierung des Sterberisikos entspricht.

Auch die nachträglich durchgeführte Analyse des mittleren Überlebens unterstreicht den Überlebensvorteil durch die Behandlung mit FOLFIRI+Aflibercept. Modelle mit log-logistischer Verteilung ergeben eine mittlere Überlebenszeit von 22,8 Monaten im FOLFIRI+Aflibercept-Arm gegenüber 18,1 Monaten im FOLFIRI+Placebo-Arm. Dies entspricht einer Differenz von 4,7 Monaten zugunsten von FOLFIRI+Aflibercept. Als Sensitivitätsanalyse wurde ein Modell mit Weibull-Verteilungsannahme angepasst, welches eine verlängerte mittlere Überlebenszeit von 3 Monaten in der FOLFIRI+Aflibercept-Gruppe zeigt.

Darüber hinaus konnten auch bei dem sekundären Endpunkt **Progressionsfreies Überleben (PFS)** signifikante Unterschiede zwischen beiden Behandlungsarmen zugunsten der Behandlung mit FOLFIRI+Aflibercept festgestellt werden: Das mediane PFS betrug in der FOLFIRI+Aflibercept-Gruppe 6,90 Monate, in der FOLFIRI+Placebo-Gruppe hingegen nur 4,67 Monate. Der Log-Rank-Test zeigt einen signifikanten Unterschied zwischen beiden Behandlungen ($p < 0,0001$), insbesondere während der ersten neun Monate der Behandlung. Das Hazard Ratio des zugrundeliegenden Cox Proportional-Hazard-Modells beträgt 0,758 (99,99%-KI: [0,578;0,995], 95%-KI: [0,661;0,867]). Die FOLFIRI+Aflibercept-Gruppe erreichte somit eine 24,2%-ige Risikoreduktion hinsichtlich einer Progression des mCRC, verglichen mit der Behandlungsgruppe FOLFIRI+Placebo.

Hinsichtlich des sekundären Endpunktes **Objektive Ansprechrates (ORR)**, zusammengefasst aus kompletter und partieller Remission) wurde ebenfalls ein signifikanter Unterschied zwischen beiden Gruppen zugunsten des FOLFIRI+Aflibercept-Arms festgestellt. In der auswertbaren Population betrug die Ansprechrates in der FOLFIRI+Aflibercept-Gruppe 19,8% gegenüber 11,1% in der FOLFIRI+Placebo-Gruppe, was nahezu einer Verdopplung der Rate entspricht. Bei zusätzlicher Berücksichtigung der stabilen Krankheitsrate zeigt sich, dass unter FOLFIRI+Aflibercept lediglich 10,4% der Patienten einen fortschreitenden Progress hatten, während bei 89,6% eine direkt verlangsamende Wirkung auf die Krankheitsprogression beobachtet wurde. Das Odds Ratio für die Ansprechrates zwischen dem FOLFIRI+Aflibercept- und dem FOLFIRI+Placebo-Arm beträgt 1,968 (95%-KI: [1,394;2,778]), das relative Risiko 1,776 (95%-KI: [1,322;2,387]).

Um zu überprüfen, ob ein Therapieeffekt über Subgruppen hinweg konstant ist oder ob es Hinweise bzw. Belege für eine Effektmodifikation gibt, sind Subgruppenanalysen inklusive Interaktionstests durchgeführt worden. Auch die präspezifizierten Subgruppenanalysen für OS und PFS zeigten nahezu ausnahmslos über alle Subgruppen hinweg einen quantitativ übereinstimmenden Behandlungseffekt zugunsten von FOLFIRI+Aflibercept (Hazard Ratio < 1), auch in den wenigen Fällen, in denen Hinweise auf eine Interaktion gefunden wurden (siehe Abschnitt 4.3.1.3.5). Insgesamt weisen die patientenrelevanten Endpunkte OS, PFS und ORR in der randomisierten kontrollierten Phase III-Studie (VELOUR) somit konsistent auf einen signifikanten Vorteil zugunsten der Therapie FOLFIRI+Aflibercept hin.

Fast alle Teilnehmer der Studie gaben mindestens ein unerwünschtes Ereignis an (606 Patienten (99,2%) in der FOLFIRI+Aflibercept-Gruppe und 592 Patienten (97,9%) in der FOLFIRI+Placebo-Gruppe). Der Unterschied zwischen beiden Gruppen ist dabei nicht signifikant ($p=0,0645$). Die häufigsten unerwünschten Ereignisse mit mehr als 5 Prozentpunkten Unterschied zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe waren gastrointestinale Erkrankungen, die von 571 Patienten (93,5%) der FOLFIRI+Aflibercept-Gruppe und von 521 Patienten (86,1%) der FOLFIRI+Placebo-Gruppe berichtet wurden, gefolgt von allgemeinen Beschwerden, wie z. B. Müdigkeit oder Kraftlosigkeit, welche 437 Patienten (71,5%) in der FOLFIRI+Aflibercept-Gruppe und 378 Patienten (62,5%) in der FOLFIRI+Placebo-Gruppe berichteten. Auch bei den schwerwiegenden unerwünschten Ereignissen (FOLFIRI+Aflibercept 48,1% vs. FOLFIRI+Placebo 32,7%) wurden gastrointestinale Erkrankungen am häufigsten angegeben (FOLFIRI+Aflibercept 20,3% vs. FOLFIRI+Placebo 11,2%). Unerwünschte Ereignisse vom Grad 3 oder 4 wurden bei 510 Patienten (83,5%) in der FOLFIRI+Aflibercept-Gruppe und bei 378 Patienten (62,5%) in der FOLFIRI+Placebo-Gruppe beobachtet. Zwischen den Gruppen bestand kein statistisch signifikanter Unterschied bei unerwünschten Ereignissen, die zum Tod führen (FOLFIRI+Aflibercept 6,1% vs. FOLFIRI+Placebo 4,8%). 164 Patienten (26,8%) im FOLFIRI+Aflibercept-Arm und 73 Patienten (12,1%) im FOLFIRI+Placebo-Arm brachen die Therapie aufgrund eines unerwünschten Ereignisses ab. Die häufigsten Ursachen für einen Behandlungsabbruch waren in beiden Behandlungsarmen Müdigkeit und Asthenie, Infektionen, Durchfall, Myelosuppression, Hypertonie, Lungenembolie, Proteinurie und tiefe Venenthrombose.

Eine vergleichende Aufzählung aller unerwünschter Ereignisse unabhängig des Schweregrades, welche um mindestens 2% häufiger im FOLFIRI+Aflibercept-Arm als im FOLFIRI+Placebo-Arm auftraten, findet sich in der Fachinformation [3]. In diese Liste flossen jedoch auch Beobachtungen aus zwei weiteren klinischen Phase III Studien zu Aflibercept mit ein, welche in anderen Tumorentitäten erhoben wurden und in denen Aflibercept in Kombination mit anderen Chemotherapeutika angewendet wurde, weshalb die Resultate in einigen Punkten (z. B. dem Auftreten eines posterioren reversiblen Enzephalopathie-Syndroms, PRES) von denen der VELOUR-Studie alleine abweichen (siehe auch Abschnitt 4.3.1.3.4).

Insgesamt ergibt sich somit ein für eine wirksame antineoplastische Therapie zu erwartendes und von einem erfahrenen Onkologen beherrschbares Nebenwirkungsprofil.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Kolorektale Karzinome sind die europaweit am zweithäufigsten auftretende Krebserkrankung. Dass trotz weitreichender Früherkennungsmaßnahmen viele Kolorektalkarzinome erst in fortgeschrittenen Stadien erkannt werden, ist eine der Hauptursachen für die insgesamt im Vergleich mit anderen Krebserkrankungen besonders schlechte Prognose der Betroffenen. Speziell das mCRC weist eine mit Fortschreiten der Erkrankung zunehmend schlechte Prognose auf. Die Erkrankten besitzen mit einer durchschnittlichen Überlebenszeit von nur

noch ca. 12 Monaten in der Regel sehr schlechte Prognosen. Weitere Behandlungsalternativen, die das Patientenüberleben mit mCRC klinisch signifikant verlängern können, stellen daher einen echten und erheblichen medizinischen Bedarf dar. Keine zielgerichtete Therapie war bislang in der Lage, in Kombination mit einer Irinotecan-basierten Chemotherapie bei mCRC-Patienten, welche zuvor eine Oxaliplatin-haltige Therapie erhalten hatten, einen zusätzlichen Überlebensvorteil auf Basis einer breiten Patientenpopulation klinisch zu beweisen.

Der Einsatz von Aflibercept zusammen mit der Kombinations-Chemotherapie aus 5-FU, Folinsäure und Irinotecan (FOLFIRI) zur Behandlung von Patienten mit mCRC, die bereits eine Oxaliplatin-haltige Chemotherapie erhalten haben, ist die erste und einzige zielgerichtete Therapieoption für diese Patientengruppe, für die auf Basis klinischer Daten ein signifikanter Überlebensvorteil gezeigt werden konnte. Angesichts dessen, dass es sich um eine fortschreitende konsumierende Erkrankung im Endstadium handelt, eine Heilung nicht mehr möglich ist und den Erkrankten bei einem weiteren Therapieversagen in der Regel nur noch palliative, hinsichtlich einer Verlängerung des Überlebens weitestgehend wirkungslose Maßnahmen zur Verfügung stehen, sind die erbrachten Nachweise einer medianen Überlebensverlängerung um 1,44 Monate sowie einer 24,2%-igen Risikoreduktion hinsichtlich einer weiteren Progression des mCRC, ergänzt durch nahezu eine Verdopplung der Ansprechrate (19,8% gegenüber 11,1%) unter Behandlung mit FOLFIRI+Aflibercept als ein Beleg für einen erheblichen Zusatznutzen zu werten. Dabei wird die Bedeutung der Resultate zum PFS und der ORR nicht nur hinsichtlich der Unterstützung der Resultate zum OS und als Abbild verbesserter Morbidität sondern auch vor dem Hintergrund psychologischer Aspekte als bedeutsam für die gesundheitsbezogene Lebensqualität der Betroffenen gewertet. Patientenberichtete oder anderweitig erfasste Daten zur gesundheitsbezogenen Lebensqualität (HRQoL) wurden in der VELOUR-Studie nicht explizit erhoben. Um hierzu belastbare Erkenntnisse zu erlangen und gleichzeitig die Datengrundlage zur Sicherheit der Anwendung weiter auszubauen, wurde von Sanofi-Aventis eine einarmige Open-Label-Studie initiiert, in der diese Punkte derzeit speziell in Patienten mit mCRC nach Vorbehandlung mit einem Oxaliplatin-haltigen Regime erhoben werden [4].

Den Resultaten zur Wirksamkeit gegenüber steht ein für eine wirksame antineoplastische Therapie zu erwartendes und von einem erfahrenen Onkologen beherrschbares Nebenwirkungsprofil: Die unabhängig vom Schweregrad häufigsten beobachteten unerwünschten Ereignisse in der Studie VELOUR waren für die Behandlung von Krebserkrankungen im Endstadium typische wie gastrointestinale Beschwerden, gefolgt von allgemeinen Beschwerden, wie z. B. Müdigkeit oder Kraftlosigkeit. Die größten Unterschiede zwischen den Behandlungsgruppen FOLFIRI+Aflibercept und FOLFIRI+Placebo in der Analyse präspezifizierter unerwünschter Ereignisse (Klasseneffekte) lagen im Auftreten von Hämorrhagien, Hypertonien und Proteinurie.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

Ziel des vorliegenden Dossiers ist die Darstellung und Quantifizierung des Nutzens, insbesondere des Zusatznutzens nach § 35a SGB V des Wirkstoffs Aflibercept (Zaltrap®) in Kombination mit einem Chemotherapie regime bestehend aus 5-FU, Folinsäure und Irinotecan (FOLFIRI) in der Behandlung von Erwachsenen mit metastasiertem kolorektalem Karzinom (mCRC), das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist, gegenüber der zweckmäßigen Vergleichstherapie FOLFIRI. Der Zusatznutzen von Aflibercept soll anhand patientenrelevanter Endpunkte der Kategorien „Mortalität“ (Gesamtüberlebensrate, OS), „Morbidity“ (progressionsfreies Überleben, PFS; objektive Ansprechrate, ORR) und „Sicherheit“ (unerwünschte Arzneimittelwirkungen) durch Auswertung publizierter Daten aus randomisierten kontrollierten klinischen Studien (RCT) ermittelt werden.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

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

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

Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Afibercept

Kategorie	Prädefinierte Einschlusskriterien	Prädefinierte Ausschlusskriterien
Patienten (P)	Erwachsene Patienten beiderlei Geschlechts mit metastasiertem Kolorektalkarzinom (mCRC), das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist.	Patienten unter 18 Jahren; nicht metastasiertes CRC (Stadien <IV); nicht mit Oxaliplatin vorbehandelte Patienten
Intervention (I)	Behandlung mit Afibercept 4 mg/kg als intravenöse Infusion über 1 h alle zwei Wochen, in Kombination mit FOLFIRI, eingesetzt entsprechend der gültigen Fachinformationen der Einzelkomponenten	Behandlung mit anderen Medikamenten; Behandlung mit abweichenden Dosierungen von Afibercept oder FOLFIRI
Kontrollgruppe (C)	Behandlung mit FOLFIRI	Vom Einschlusskriterium abweichende/keine Kontrollgruppe
Endpunkte (O)	Mindestens ein patientenrelevanter Endpunkt der Kategorien <ul style="list-style-type: none"> - Mortalität - Morbidität - (gesundheitsbezogene) Lebensqualität - Sicherheit wird berichtet	Kein Endpunkt aus den genannten Kategorien
Studientyp	Randomisierte kontrollierte Studie (RCT)	Vom Einschlusskriterium abweichendes Studiendesign
Studiendauer	Keine Einschränkungen	Keine
Publikationstyp	Vollpublikation oder Bericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht. Doppelpublikationen ohne relevante Zusatzinformation werden ausgeschlossen.	Studien in der Rekrutierungsphase; ohne verfügbare Vollpublikation oder Bericht

Zielpopulation für Afibercept sind gemäß Zulassung erwachsene Patienten beiderlei Geschlechts mit mCRC, das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist. Studien, welche in signifikantem Ausmaß Patienten eingeschlossen hatten, die dieser Anforderung nicht entsprechen, wurden folglich nicht bei der Bewertung berücksichtigt.

Hinsichtlich der **Intervention** werden nur Studien berücksichtigt, in denen Aflibercept gemäß der Empfehlungen der Fachinformation [3] gegeben wurde. Das gleiche gilt für die zusätzlich verabreichte Kombinations-Chemotherapie FOLFIRI und die Dosierung der Einzelkomponenten 5-FU, Folinsäure/Leucovorin und Irinotecan.

Die zweckmäßige **Vergleichstherapie** ist, bestätigt durch die Beratung durch die G-BA-Geschäftsstelle [5], der alleinige Einsatz einer Kombinations-Chemotherapie aus 5-FU, Folinsäure und Irinotecan (FOLFIRI). Entsprechend werden nur Studien berücksichtigt, die a) eine Kontrollgruppe besitzen und in denen b) auch eine diesem Kriterium entsprechende Medikation verabreicht wurde.

Die Ein- und Ausschlusskriterien hinsichtlich der in den in Frage kommenden Studien zu betrachtenden **Endpunkt**kategorien orientieren sich direkt an den Vorgaben der „Allgemeinen Methoden (Ver. 4.0)“ des IQWiG für die Nutzenbewertung [6], die das Institut auch in der Operationalisierung des Zusatznutzens von Wirkstoffen gemäß Arzneimittelmarktneuordnungsgesetz (AMNOG) bereits herangezogen hat.

Untersuchungen zur Mortalität sind, gerade bei der Beurteilung von Wirkstoffen in der Krebstherapie, von außerordentlichem Interesse. Dies gilt somit auch für den klassischen Endpunkt dieser Kategorie, das Gesamtüberleben (OS). Aufgrund der eindeutigen Definition, der Belastbarkeit des gemessenen Ergebnisses sowie der Tatsache, dass bereits implizit eine Risiko/Nutzen-Abwägung enthalten ist, wurde und wird OS in vielen Bereichen der Onkologie als zuverlässigster Endpunkt eingeordnet [7-9].

Zur Beurteilung der Wirksamkeit hinsichtlich der Morbidität, also der Beeinflussung (insbesondere schwerer) Symptome der Erkrankung, kann eine Reihe unterschiedlicher Punkte betrachtet werden. In Studien zu onkologischen Indikationen finden hier besonders oft das progressionsfreie Überleben (PFS) sowie die objektive Ansprechrates (ORR) Verwendung [9-11]. Beide gelten laut IQWiG als nicht direkt durch Patienten erfahrbar und somit als Surrogatendpunkte [8]. Die Ansichten sind hier jedoch geteilt; gerade vor dem Hintergrund schwerer, nicht mehr heilbarer Krebserkrankungen mehren sich die Forderungen, unter anderem auch die psychologische Komponente, die die Wahrnehmung eines gesicherten Tumorsprechens bzw. einer ausbleibenden Progression durch den Patienten enthalten kann, zu berücksichtigen und diese Endpunkte in solchen Indikationen zu den patientenrelevanten hinzu zu nehmen [12-14]. Neuere retrospektive Untersuchungen in Studien zu fortgeschrittenen Krebserkrankungen unterstützen die Annahme, dass sich ein Progress der Erkrankung negativ auf die patientenberichtete Lebensqualität auswirkt [15]. Und selbst wenn man die direkte Patientenrelevanz von PFS nicht akzeptiert, ist es speziell beim Kolorektalkarzinom allgemein als valider Surrogatendpunkt für das Gesamtüberleben anerkannt, d. h. die Übertragbarkeit von Ergebnissen zum PFS auf das OS gilt als statistisch hinreichend gesichert [16]. Signifikante Resultate zum OS werden daher durch gleichgerichtete Effekte bezüglich des PFS bei denselben Patienten in besonderer Weise untermauert. FDA (U.S. Food and Drug Administration) und EMA (European Medicines Agency) akzeptieren PFS laut aktueller Guidance-Dokumente als Endpunkt für den Nachweis eines klinischen Nutzens, gerade in Bezug auf Patientenrelevanz [17-21], und haben bereits

mehrfach neue Medikamente unter anderem auch auf Basis einer PFS-Verbesserung zugelassen. Hintergrund ist hier vor allem die Erkenntnis, dass Daten zum PFS rascher erhoben werden können als solche zum OS und dass diese darüber hinaus auch nicht durch mögliche, nach Eintritt einer Progression verabreichte, Folgetherapien beeinflusst werden können [7].

Sofern verfügbar, zählen auch Daten zur Lebensqualität der Studienteilnehmer, typischerweise mittels Fragebögen direkt oder indirekt über die behandelnden Ärzte erhoben, zu den für die Nutzenbewertung relevanten Endpunkten [6]. Ein Erhalt oder sogar eine Verbesserung der Lebensqualität stellt – gerade in der palliativen Situation, die auf Dauer fast immer mit einer Verschlechterung der Lebensqualität einhergeht – neben den zuvor genannten Parametern ebenfalls einen patientenrelevanten Endpunkt dar, da die Therapie die Lebensqualität von Patienten möglichst verbessern, zumindest jedoch nicht verschlechtern sollte [22].

Ergebnisse zur Sicherheit des zu bewertenden Wirkstoffs, gemessen anhand der Rate an unerwünschten und/oder schwerwiegenden unerwünschten Ereignissen (UE/SUE), sind für eine umfassende Beurteilung zusätzlich zu den Ergebnissen zur Wirksamkeit grundsätzlich immer mit zu berücksichtigen [6].

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

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

– Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.

– Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

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

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

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

Da eine Zulassung von Aflibercept im Indikationsgebiet mCRC weltweit nicht länger als 12 Monate vor Einreichung des vorliegenden Dossiers erfolgt ist (Zulassung durch die FDA am 03.08.2012) und zudem für die Bewertung des Zusatznutzens eine direkte Vergleichsstudie mit der zweckmäßigen Vergleichstherapie FOLFIRI vorliegt, besteht laut Dossievorlage keine Verpflichtung zur Durchführung einer bibliografischen Literaturrecherche. Dementsprechend wurde auf diese verzichtet.

4.2.3.3 Suche in Studienregistern

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

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

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

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

Eine Registersuche nach für die Zusatznutzenbewertung relevanten klinischen Studien zu Aflibercept wurde am 10.01.2013 gemäß den Vorgaben der VerfO [2] im Internetauftritt des internationalen Studienregisters ClinicalTrials.gov und über das International Clinical Trials Registry Platform (ICTRP) Search Portal der World Health Organization (WHO) durchgeführt. Die verwendeten Suchstrategien und die erzielten Treffer sind in Anhang 4-B detailliert dargestellt.

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 durch die in Abschnitt 4.2.3.3 und Anhang 4-B beschriebene Recherche in Studienregistern identifizierten Treffer wurden von zwei unabhängigen Personen gesichtet und die aus dem jeweiligen Registereintrag bzw. eventuell vorliegenden Publikationen verfügbaren Informationen anhand der in Abschnitt 4.2.2 präzisierten Ein- und Ausschlusskriterien (Tabelle 4-1) evaluiert. Für den Fall von Diskrepanzen im Selektionsprozess war die Einbeziehung der Einschätzung einer unabhängigen dritten Person in die Selektion vorgesehen.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende

endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-G):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

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

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

Die zur Darstellung und Quantifizierung des Zusatznutzens eingeschlossene Studie wurde hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung der eingeschlossenen Studie im Rahmen des vorliegenden Dossiers erfolgte anhand der zur Verfügung stehenden Informationen aus dem Studienbericht der Studie [1].

Bewertung des Verzerrungspotenzials

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte wurden untersucht. Bei Verzerrungen auf Endpunktebene wurde jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Entsprechend der Verfahrensordnung des G-BA zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V vom 20. Januar 2011 [2] wurden dabei folgende Aspekte bewertet:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnigesteuerte Berichterstattung
- Sonstige das Verzerrungspotential beeinflussende endpunktübergreifende Punkte

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des Intention-To-Treat (ITT)-Prinzips
- Ergebnigesteuerte Berichterstattung
- Sonstige Aspekte: Behandlung fehlender Werte und Drop-outs

Die Bewertung des Verzerrungspotentials der Einzelstudie wurde tabellarisch zusammengefasst. Gemäß Verfahrensordnung wurde das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, 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.

Abschließend wurde unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auch auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem „niedrigen“ Verzerrungspotential wurde ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt sind.

Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes relevant verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotentials für die Studie und jeden Endpunkt (Operationalisierung) richtete sich nach den Hinweisen der Bewertungsbögen in Anhang 4-G.

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 für die Nutzenbewertung von Aflibercept in Kombination mit FOLFIRI herangezogene randomisierte kontrollierte Phase III-Studie VELOUR wird im Abschnitt 4.3.1.2 sowie in Anhang 4-F anhand von Design-Charakteristika beschrieben. Die Beschreibung erfolgte anhand der Punkte 2b bis 14 des CONSORT-Statements. Der Patientenfluss wurde gemäß CONSORT-Flow-Chart dargestellt (siehe Abbildung 14).

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

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.

werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

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

Die Studienpopulation der in die Nutzenbewertung eingehenden RCT VELOUR wurde anhand folgender Charakteristika beschrieben:

- Alter
- Geschlecht
- Geografische Region
- Vorbestehende Hypertonie ja/nein
- ECOG-PS
- Vorbehandlung mit Bevacizumab ja/nein
- Monate zwischen Erstdiagnose und Randomisierung
- Anzahl Organe, die von Metastasen betroffen sind
- Primäre Lage des Tumors
- Ausschließlich Lebermetastasen ja/nein
- Beteiligte metastasierte Organe
- Zahl der Therapieabbrecher
- Anzahl Patienten, die weitere antineoplastische Therapie erhalten haben

Die Ergebnisse der Studie wurden nach Zielgrößen geordnet gegenübergestellt und vergleichend beschrieben. Dabei wurden die folgenden aussagekräftigen und patientenrelevanten Endpunkte quantifiziert:

- Gesamtüberleben (OS, primärer Endpunkt)
- Progressionsfreies Überleben (PFS)
- Objektive Ansprechrate (ORR)
- Sicherheit

Alle aufgeführten Zielgrößen stellen nach Auffassung von Sanofi-Aventis patientenrelevante Endpunkte dar (siehe auch Diskussion unter 4.2.2). Resultate zum Endpunkt „Gesundheitsbezogene Lebensqualität“ wurden in VELOUR nicht erhoben.

Gesamtüberleben (OS): Bei der Behandlung fortgeschrittener onkologischer Erkrankungen, im Besonderen solcher, in denen nur eine kurze Zeit zwischen dem Eintritt einer Progression und dem Tod zu erwarten ist, stellt die Verlängerung des OS in der Regel ein herausragendes Therapieziel und somit einen der wichtigsten klinischen Endpunkte dar. Das OS wird in randomisierten klinischen Studien als Zeitspanne zwischen dem Zeitpunkt der Randomisierung und dem Tod aus jeglicher Ursache operationalisiert. Aufgrund der eindeutigen Definition, der Belastbarkeit des gemessenen Ergebnisses sowie der Tatsache, dass in diesem Parameter bereits implizit eine Risiko/Nutzen-Abwägung enthalten ist, wurde und wird das OS in vielen Bereichen der Onkologie als zuverlässigster und darüber hinaus zweifelsfrei patientenrelevanter Endpunkt eingeordnet [7-9].

Progressionsfreies Überleben (PFS): PFS wird meist als die Zeitspanne zwischen Studieneinschluss eines Patienten und dem Beginn der Progression der Erkrankung oder dem Todesdatum des Patienten (je nach Ersteintritt) definiert. Die Todesursache spielt dabei keine Rolle [7, 17]. Wie auch die Ansprechrate gilt PFS laut IQWiG als nicht direkt durch Patienten erfahrbar und somit als Surrogatendpunkt [8]. Die Ansichten sind hier jedoch geteilt; gerade vor dem Hintergrund schwerer, nicht mehr heilbarer Krebserkrankungen mehren sich die Forderungen, unter anderem auch die psychologische Komponente, die die Wahrnehmung eines gesicherten Tumorsprechens bzw. einer ausbleibenden Progression durch den Patienten enthalten kann, zu berücksichtigen und diese Endpunkte in solchen Indikationen zu den patientenrelevanten hinzu zu nehmen [12-14]. Neuere retrospektive Untersuchungen unterstützen die Annahme, dass sich bei fortgeschrittenen Krebserkrankungen ein Progress der Erkrankung negativ auf die patientenberichtete Lebensqualität auswirkt [15]. Und selbst wenn man eine direkte Patientenrelevanz von PFS nicht akzeptiert, ist es speziell beim Kolorektalkarzinom allgemein als valider Surrogatendpunkt für das OS anerkannt, d. h. die Übertragbarkeit von Ergebnissen zum PFS auf das OS gilt als statistisch hinreichend gesichert [16]. Signifikante Resultate zum OS werden daher durch gleichgerichtete Effekte bezüglich des PFS bei denselben Patienten in besonderer Weise untermauert. FDA und EMA akzeptieren PFS laut aktueller Guidance-Dokumente als Endpunkt für den Nachweis eines Nutzens, gerade in Bezug auf Patientenrelevanz [17-21], und haben bereits mehrfach neue Medikamente unter anderem auch auf Basis einer PFS-Verbesserung zugelassen. Hintergrund

ist hier vor allem die Erkenntnis, dass Daten zum PFS rascher erhoben werden können als solche zum OS und dass diese darüber hinaus auch nicht durch mögliche, nach Eintritt einer Progression verabreichte, Folgetherapien beeinflusst werden können [7].

Objektive Ansprechrate (ORR): Die Erhebung der ORR erfolgte innerhalb der Studie VELOUR durch eine objektive Bewertung nach RECIST (Response Evaluation Criteria in Solid Tumors). Falls möglich sollte bei Vorliegen einer messbaren Erkrankung diese zur Verlaufskontrolle hinzugezogen werden. Um eine Vergleichbarkeit über verschiedene Studien zu gewährleisten, wurden Ansprechen, stabile Erkrankung und Progression entsprechend der RECIST Kriterien (Version 1.0) des NCI (National Cancer Institute) definiert [23], die einen weit verbreiteten Standard darstellen. Die ORR misst die Verringerung der Tumorlast und umfasst sowohl komplette als auch partielle Remission, allerdings nicht den Stillstand der Erkrankung [9, 10, 17] und ist damit ein Parameter für reduzierte Morbidität. Das Wissen um ein Tumoransprechen und die Verringerung von Metastasen bedeuten eine psychische Entlastung und damit einen direkten Nutzen für den Patienten. Darüber hinaus dient auch dieser Endpunkt in vielen Studien als primärer Endpunkt zum Nachweis der Wirksamkeit und als Basis für die Erteilung der Zulassung durch die FDA [9, 17], und man geht davon aus, dass eine verbesserte ORR durch die Verkleinerung des Tumors die Symptome verringern kann und so gegebenenfalls auch ohne Verlängerung der Gesamtüberlebensrate einen patientenrelevanten klinischen Vorteil darstellt [24, 25]. In einigen onkologischen Indikationen gibt es bereits Hinweise darauf, dass sich ein gutes Ansprechen positiv auf die Überlebenszeit auswirkt [26].

Sicherheit (Unerwünschte Ereignisse): Ergänzend zum Nachweis des Nutzens dient die Erfassung unerwünschter Arzneimittelwirkungen zur Nutzen-Risiko-Bewertung. Neben den erwünschten Wirkungen eines Arzneimittels treten in vielen Fällen unvermeidlich auch unerwünschte Wirkungen auf, die den Patienten neben der eigentlichen Erkrankung physisch und psychisch belasten können. Daher wird der Nutzen einer Therapie hinsichtlich patientenrelevanter Zielkriterien wie OS oder PFS üblicherweise auch in Relation zur möglichen Toxizität bewertet.

(Gesundheitsbezogene) Lebensqualität: Patientenberichtete oder anderweitig erfasste Daten zur gesundheitsbezogenen Lebensqualität (HRQoL) wurden in der VELOUR-Studie nicht explizit erhoben. Um hierzu belastbare Erkenntnisse zu erlangen und gleichzeitig die Datengrundlage zur Sicherheit der Anwendung weiter auszubauen, wurde von Sanofi-Aventis eine einarmige Open-Label-Studie initiiert, in der diese Punkte derzeit speziell in Patienten mit mCRC nach Vorbehandlung mit einem Oxaliplatin-haltigen Regime erhoben werden [4].

4.2.5.3 Meta-Analysen

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

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen 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 anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten 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.

Im vorliegenden Dossier werden keine Meta-Analysen für die Bewertung des Zusatznutzens von Aflibercept eingesetzt.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden, wenn sich die Studien bezüglich dieser Faktoren unterscheiden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die

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.

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.

Sensitivitätsanalysen dienen der Überprüfung und Dokumentation der Robustheit der Ergebnisse und Schlussfolgerungen gegenüber der Variation methodischer Faktoren. Sie vergleichen unter anderem unterschiedliche Analysemethoden (zum Endpunkt OS wird neben dem medianen Überleben auch das mittlere Überleben betrachtet) oder unterschiedliche Zensierungsmethoden hinsichtlich des Einflusses auf das Resultat. Bei der Darstellung der Studienergebnisse im vorliegenden Dossier wurden die gemäß Studienprotokoll präspezifizierten Sensitivitätsanalysen berücksichtigt.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

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

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

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

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

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

Um zu überprüfen, ob ein Therapieeffekt über Subgruppen hinweg konstant ist oder ob es Hinweise auf eine Effektmodifikation gibt, sind Subgruppenanalysen inklusive Interaktionstests für die Stratifikationsfaktoren (ECOG-Performance Status und Vorbehandlung mit Bevacizumab) durchgeführt worden. Für den primären Endpunkt der VELOUR-Studie OS und den sekundären Endpunkt PFS wurden die Ergebnisse der präspezifizierten Subgruppenanalysen aus dem Studienbericht dargestellt. Darin wurde der Interaktionstest des Cox-Modells mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor verwendet. Für die Bewertung der Sicherheit wurden ebenfalls Subgruppenanalysen post-hoc durchgeführt. Hierfür wurde der Interaktionstest für Subgruppen basierend auf der Cochrane Q-Statistik angewendet.

Um die Stärke einer Interaktion einzuordnen gibt es hinsichtlich des p-Wertes unterschiedliche Argumentationen. In der schließenden Statistik ist für Testentscheidungen die Verwendung eines Signifikanzniveaus von 0,05 üblich [27]. Die Cochrane Collaboration empfiehlt jedoch in ihrem Handbuch eine Erhöhung auf 0,1 als Ausgleich für die geringe Power, die Subgruppenanalysen meistens mit sich bringen [28]. Andere Quellen erwägen sogar eine Erhöhung des Schwellenwertes auf bis zu 0,2 [6]. Dem gegenüber steht allerdings der Einwand, dass das grundsätzliche Problem der mangelnden Power dadurch auch nicht gelöst werden kann [29]. Im vorliegenden Dossier wurde, in Analogie zu den Empfehlungen des IQWiG, aufgrund oben genannter Argumentationen, ein p-Wert unterhalb von 0,05 als Beleg für eine Effektmodifikation gewertet, und ein p-Wert zwischen 0,05 und 0,2 als Hinweis einer Effektmodifikation definiert.

Generell sollte die Aussagekraft von Subgruppenanalysen, insbesondere auf Studienebene, kritisch beurteilt werden. Geringe Patientenzahlen innerhalb der Subgruppenkategorien, ein Ungleichgewicht zwischen den Subgruppenkategorien und die steigende Anzahl statistischer Tests erhöhen die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse (siehe hierzu auch das IQWiG Methodenpapier 4.0, S. 121ff) [6]. Subgruppenanalysen sind daher nicht dafür bestimmt, definitive Schlüsse über das Ausmaß des Behandlungseffektes in Subgruppen zu ziehen, sondern um die Konsistenz des Behandlungseffektes über Subgruppen

hinweg zu beurteilen. Diese Aspekte sollten bei der Bewertung und Interpretation von Subgruppenresultaten Berücksichtigung finden.

Folgende Subgruppen wurden, wie im Studienprotokoll präspezifiziert, für die Endpunkte OS und PFS hinsichtlich Effektmodifikation evaluiert:

- ECOG-PS (0/1/2)
- Alter (<65, ≥65)
- Geschlecht (männlich, weiblich)
- Ethnische Zugehörigkeit (Kaukasisch + weiß/andere)
- Geografische Region (West-/Osteuropa/Nord-/ Südamerika/andere Kontinente)
- Vorbehandlung mit Bevacizumab (ja/nein)
- Vorbestehende Hypertonie (ja/nein)
- Anzahl Organe, die von Metastasen betroffen sind (>1/≤1)
- Nur Lebermetastasen (ja/nein)
- Lokalisation des primären Tumors (Kolon/Rektosigmoid/andere vs. Rektum)

Für die Endpunkte OS und PFS wurde die Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung post-hoc zusätzlich auch anhand der Studienpopulation unter Ausschluss der adjuvant mit Oxaliplatin vorbehandelten Patienten (letztere stellten etwa 10,1% der gesamten Studienteilnehmer dar) durchgeführt [3]. Da ein Einsatz von Bevacizumab nach den Vorgaben der Fachinformation in der adjuvanten Therapiesituation nicht zulässig ist [30], wurde die Gefahr einer Verzerrung der Ergebnisse bei Betrachtung der ungefilterten Resultate zur Gesamtpopulation gesehen.

Für den Endpunkt Sicherheit wurden Subgruppenanalysen post-hoc für die durch die Verfahrensordnung vorgegebenen Merkmale Alter, Geschlecht und Schwere der Erkrankung (in dieser Studie repräsentiert durch den ECOG-PS) durchgeführt. Außerdem wurden auch hier die Subgruppen für das Merkmal Vorbehandlung mit Bevacizumab analysiert, da in der VELOUR-Studie zusätzlich (neben dem ECOG-PS) eine stratifizierte Randomisierung nach dieser Variable erfolgt war.

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

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

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

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

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

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

¹¹ B. Schöttker, D. Lümann, 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.

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

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

Aufgrund des Vorliegens einer direkten Vergleichsstudie mit der zweckmäßigen Vergleichstherapie war das Heranziehen indirekter Vergleiche nicht erforderlich.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

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

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

Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
VELOUR	ja	abgeschlossen	November 2007 – Februar 2011	Aflibercept/FOLFIRI, Placebo/FOLFIRI
AFLAME	ja*	laufend	Juli 2012 – Januar 2016	Aflibercept/FOLFIRI, Placebo/FOLFIRI

* Zulassungsstudie für den asiatisch-pazifischen Raum; umfasst nur Patienten asiatischer Herkunft

Geben Sie an, welchen Stand die Information in Tabelle 4-2 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information von Tabelle 4-2 ist der 10.01.2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-2 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
AFLAME	Studie befindet sich noch in der Rekrutierungsphase; es sind noch keine Daten verfügbar. Die als Zulassungsstudie für den asiatisch-pazifischen Raum konzipierte Studie umfasst nur Patienten asiatischer Herkunft und somit eine für den Dossierrahmen weitgehend nicht relevante Population.

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

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

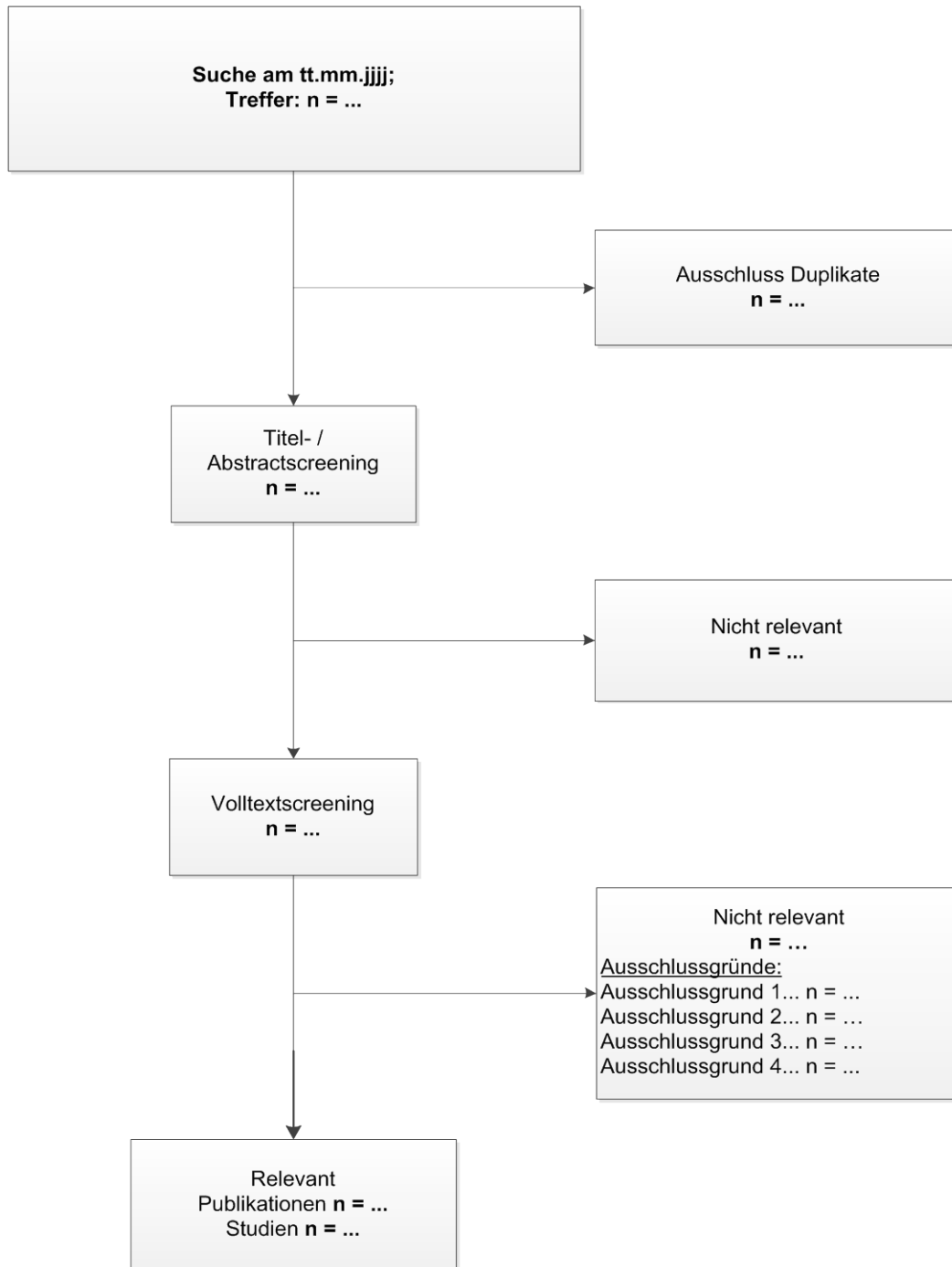


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Gemäß Verfahrensordnung wurde keine bibliografische Literaturrecherche durchgeführt, da zum Zeitpunkt der deutschen Markteinführung von Aflibercept weltweit keine Zulassung länger als 12 Monate vorlag.

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 4-2) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

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

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

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

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^b)
VELOUR	clinicaltrials.gov [http://www.clinicaltrials.gov/show/NCT00561470] [31], International Clinical Trials Registry Platform Search Portal [http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2007-000820-42-DE] [32]	ja	n. d.

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

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

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

Tabelle 4-5: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
VELOUR	ja	ja	nein	ja [1]	ja [31]	ja [33]

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

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

Tabelle 4-6: 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
VELOUR	RCT, doppelblind, parallel, prospektiv multinational (28 Länder), multizentrisch (176 Zentren), Randomisierung stratifiziert nach ECOG-PS (0/1/2) und nach Vorbehandlung mit Bevacizumab (ja/nein)	Erwachsene (>18 Jahre) mit histologisch oder zytologisch bestätigtem metastasiertem Adenokarzinom des Kolons oder Rektums, für die es nach Versagen eines Oxaliplatin-basiertem Therapieregimes keine kurative Behandlung gibt. Patienten mit messbarer oder nicht messbarer Krankheit (nach RECIST-Kriterien) waren akzeptabel. Eine Vorbehandlung mit Bevacizumab war erlaubt; eine Vorbehandlung mit Irinotecan war nicht erlaubt. ECOG-PS der Patienten 0-2.	FOLFIRI+Aflibercept (n= 612) FOLFIRI+Placebo (n= 614)	Behandlungsdauer: Behandlung bis Krankheitsprogression, inakzeptabler Toxizität, Abbruch der Studienmedikation durch den Patienten oder Arzt Beobachtungsdauer: Entsprechend der Behandlungsdauer; das Überleben wurde nachbeobachtet bis zum Tod oder zum Studienende (Cutoff-Datum), je nach Ersteintritt (maximale Nachbeobachtung 9 Monate nach Cutoff-Datum)	Westeuropa, Osteuropa, Nordamerika, Südamerika, Australien, Neuseeland, Südafrika, Korea 11/2007 – 02/2011	Primärer Endpunkt: Verlängerung des Gesamtüberlebens Patientenrelevante sekundäre Endpunkte: Progressionsfreies Überleben Therapieansprechen Sicherheit

RCT: Randomized Controlled Trial; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; RECIST: Response Evaluation Criteria in Solid Tumors

Tabelle 4-7: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertende Arzneimittel	Vergleichstherapie	Behandlungscharakteristika
VELOUR	FOLFIRI + Aflibercept 4 mg/kg i.v. über 1 Stunde	FOLFIRI + Placebo	Gabe des Therapieregime jeweils im 2-wöchigen Zyklus FOLFIRI Zusammensetzung: Folinsäure (dl-LV) 400 mg/m ² über 120 min Irinotecan 180 mg/m ² über 90 min 5-FU Bolus 400 mg/m ² über 2-4 min 5-FU Infusion 2.400 mg/m ² über 46 h Begleitmedikation: Alle unterstützenden Maßnahmen, die die optimale Pflege des Patienten gewährleisten, konnten während der Studie verabreicht werden.

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; dl-LV: Leucovorin (Folinsäure), als d/l-Racemat (im Falle der Verwendung der reinen l-Isoform sollte die Dosierung halbiert werden); 5-FU: Fluorouracil; i.v.: intravenös

Tabelle 4-8: Charakterisierung der Studienpopulationen (Alter, Geschlecht, geographische Region) – RCT mit dem zu bewertenden Arzneimittel.

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w /m (%)	Geographische Region n (%)				
				West-europa	Ost-europa	Nord-amerika	Süd-amerika	Andere Länder*
VELOUR								
FOLFIRI+ Aflibercept	612	59,5 (10,5)	40,4/59,6	208 (34,0)	161 (26,3)	63 (10,3)	62 (10,1)	118 (19,3)
FOLFIRI+ Placebo	614	60,2 (10,8)	42,5/57,5	217 (35,3)	136 (22,1)	75 (12,2)	56 (9,1)	130 (21,2)

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; SD: Standardabweichung

* andere Länder: Australien, Neuseeland, Südafrika, Korea

Tabelle 4-9: Charakterisierung der Studienpopulationen (vorbestehender Hypertonus, ECOG-PS, Vorbehandlung mit Bevacizumab) – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Vorbestehender Hypertonus, n (%)		ECOG-PS, n (%)			Vorbehandlung mit Bevacizumab, n (%)	
		ja	nein	0	1	2	ja	nein
VELOUR				349 (57)	250 (40,8)	13 (2,1)		
FOLFIRI+ Aflibercept	612	266 (43,5)	346 (56,5)				186 (30,4)	426 (69,6)
FOLFIRI+ Placebo	614	268 (43,6)	346 (56,4)	350 (57)	250 (40,7)	14 (2,3)	187 (30,5)	427 (69,5)

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status

Tabelle 4-10: Charakterisierung der Studienpopulationen (Zeit zwischen Erstdiagnose und Randomisierung, Anzahl Organe, die von Metastasen betroffen sind, primäre Lage des Tumors) – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Monate zw. Erstdiagnose & Randomisierung Mittelwert (SD)	Anzahl Organe, die von Metastasen betroffen sind n (%)			Primäre Lage n (%)			
			0	1	>1	Kolon	Rekto-sigmoid	Rektum	Andere
VELOUR									
FOLFIRI+ Aflibercept	612	20,98±24,08	2 (0,3)	256 (41,8)	354 (57,8)	289 (47,2)	123 (20,1)	197 (32,2)	3 (0,5)
FOLFIRI+ Placebo	614	20,88±21,10	6 (1,0)	271 (44,1)	337 (54,9)	302 (49,2)	136 (22,1)	174 (28,3)	2 (0,3)

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; SD: Standardabweichung

Tabelle 4-11: Charakterisierung der Studienpopulationen (ausschließlich Lebermetastasen, beteiligte metastasierte Organe) – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	nur Lebermetastasen n (%)		Beteiligte metastasierte Organe n (%)			
		ja	nein	Leber	Lunge	Lymphknoten	Peritoneum
VELOUR							
FOLFIRI+ Aflibercept	612	153 (25,0)	459 (75,0)	459 (75,0)	271 (44,3)	173 (28,3)	68 (11,1)
FOLFIRI+ Placebo	614	146 (23,8)	468 (76,2)	431 (70,2)	277 (45,1)	181 (29,5)	88 (14,3)

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan.

Tabelle 4-12: Charakterisierung der Studienpopulationen (Therapieabbrecher, Anzahl der Patienten, die weitere antineoplastische Therapie erhalten haben) – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Therapieabbrecher	Anzahl Patienten, die weitere antineoplastische Therapie erhalten haben
VELOUR			
FOLFIRI+Aflibercept	612	593 (96,9)	364 (59,5)
FOLFIRI+Placebo	614	598 (97,4)	366 (59,6)

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan.

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

VELOUR ist eine multizentrische, placebokontrollierte, randomisierte, doppelblinde, prospektive Phase-III-Zulassungsstudie, die bei 1.226 Patienten mit mCRC nach Versagen eines Oxaliplatin-basierten Therapieregimes in 176 Zentren in 28 Ländern durchgeführt wurde. Die Studienteilnehmer wurden in zwei Behandlungsgruppen randomisiert, stratifiziert nach ECOG-PS (0/1/2) und nach Vorbehandlung mit Bevacizumab (ja/nein).

Die Studienteilnehmer erhielten in Abständen von zwei Wochen zu Zyklusbeginn zusätzlich zu FOLFIRI entweder Aflibercept (4 mg/kg über eine Stunde intravenös) oder Placebo.

Jeder Patient wurde mit der Studienmedikation behandelt bis zum Krankheitsprogress, inakzeptabler Toxizität, Tod oder bis zum Abbruch der Studienmedikation durch den Patienten selbst oder den behandelnden Arzt. Patienten wurden bis zum Studienende (Cutoff-Datum) oder zum Tod, je nach Ersteintritt, beobachtet (maximale Nachbeobachtung: 9 Monate nach Daten-Cutoff). Primärer Endpunkt der Studie war eine Verlängerung des Gesamtüberlebens. Die sekundären Endpunkte umfassten progressionsfreies Überleben, Therapieansprechen (objektive Ansprechrates) und Sicherheit. Das Gesamtüberleben wurde gemäß Protokoll nach Eintritt von 863 Todesfällen ausgewertet.

Die demografischen und krankheitsspezifischen Charakteristika waren zu Beginn der Studie zwischen den Studienarmen vergleichbar. Es gab in jedem Behandlungsarm etwas mehr Männer als Frauen (FOLFIRI+Aflibercept: 57,5% und FOLFIRI+Placebo: 59,6%); die Studienteilnehmer waren im Median 61 Jahre alt und wurden durchschnittlich 21 Monate vor Randomisierung erstmals diagnostiziert. Fast alle Patienten hatten bereits Metastasen, viele schon mehr als eine (FOLFIRI+Aflibercept: 57,8% und FOLFIRI+Placebo: 54,9%) (Tabelle 4-8, Tabelle 4-10).

Die Zahl der Patienten, die die Studienmedikation abbrachen, war in beiden Behandlungsarmen hoch (FOLFIRI+Aflibercept: 96,9% und FOLFIRI+Placebo: 97,4%) (Tabelle 4-12). Der häufigste Grund hierfür war, dass die Patienten gemäß Protokoll nach Krankheitsprogress die Studie abbrechen mussten (FOLFIRI+Aflibercept: 49,8% und

FOLFIRI+Placebo: 71,2%). Im Studienverlauf erhielten 730 Patienten eine weitere antineoplastische Therapie (nach FOLFIRI+Aflibercept: 59,5% und nach FOLFIRI+Placebo: 59,6%) (Tabelle 4-12). Diese Zusatzbehandlungen (unter anderem Radiotherapien, Operationen oder systemische Behandlungen) waren in den beiden Behandlungsarmen vergleichbar.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

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

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei der VELOUR-Studie handelt es sich um eine randomisierte, doppelblinde Studie. Patienten wurden zentral über ein Spracherkennungssystem (IVRS; Interactive Voice Response System) mit einem Zuteilungsverhältnis 1:1 randomisiert (stratifiziert nach ECOG-PS und Vorbehandlung mit Bevacizumab). Patienten, die behandelnden Ärzte und das Studienpersonal waren verblindet bezüglich der Studienmedikation. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder weitere das Verzerrungspotenzial beeinflussende Punkte lagen nicht vor. Das Verzerrungspotenzial der Studie wird daher als **niedrig** bewertet.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

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

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

Studie	Gesamtüberleben (OS)	Progressions-freies Überleben (PFS)	Objektive Ansprechrate (ORR)	Sicherheit (Unerwünschte Ereignisse)
VELOUR	ja	ja	ja	ja

4.3.1.3.1 Gesamtüberleben (OS) – RCT

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

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

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

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

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

Tabelle 4-15: Operationalisierung von Gesamtüberleben (OS)

Studie	Operationalisierung
VELOUR	Das Gesamtüberleben ist definiert als der Zeitraum von der Randomisierung bis zum Tod (die Todesursache spielt keine Rolle).

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

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

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

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Der Endpunkt OS wurde definiert als Zeit von der Randomisierung in die Studie bis zum Tod und wurde innerhalb der ITT-Population ausgewertet. Die ITT-Population schloss alle Patienten mit ein, die durch das IVRS (interactive voice response system) randomisiert wurden, ungeachtet dessen, ob sie die Studienmedikation eingenommen hatten oder nicht. Verzerrende Faktoren bei der Umsetzung des ITT-Prinzips wurden nicht gefunden. Hinweise oder sonstige das Verzerrungspotential beeinflussende Punkte gab es ebenfalls nicht. Das Verzerrungspotential für diesen Endpunkt wurde daher mit **niedrig** bewertet.

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

Das OS wurde anhand eines Vergleichs der zwei Behandlungsgruppen mit dem stratifiziertem (nach ECOG-PS/vorangegangener Einnahme von Bevacizumab) Log-Rank-Test analysiert. Des Weiteren werden die Überlebenskurven mittels der Kaplan-Meier-Schätzer dargestellt (siehe Abbildung 2) sowie das Hazard Ratio mittels eines stratifizierten (nach denselben Faktoren wie im Log-Rank-Test) Cox-Proportional-Modells angegeben.

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

Studie			
VELOUR	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)	
Todesfälle, n (%)	403 (65,8)	460 (74,9)	
Zensierte Fälle, n (%)	209 (34,2)	154 (25,1)	
Diff. in Monaten			
25%-Quantil der Überlebenszeit in Monaten [95,34%*-KI] [95%-KI]	7,62 [6,57;8,48] [6,60;8,48]	6,83 [6,14;7,59] [6,24;7,59]	0,79
Mediane Überlebenszeit in Monaten [95,34%*-KI] [95%-KI]	13,50 [12,52;14,95] [12,52;14,95]	12,06 [11,07;13,11] [11,07;13,08]	1,44
75%-Quantil der Überlebenszeit in Monaten [95,34%*-KI] [95%-KI]	25,59 [22,01;31,70] [22,01;31,70]	21,03 [18,92;22,80] [18,92;22,80]	4,56
Kaplan-Meier Kenngrößen			
Patienten unter Risiko, n (%)			
t = 3 Monate	566 (92,5)	573 (93,3)	
t = 6 Monate	498 (81,4)	485 (79,0)	
t = 9 Monate	416 (68,0)	401 (65,3)	
t = 12 Monate	311 (50,8)	286 (46,6)	
t = 18 Monate	148 (24,2)	131 (21,3)	
t = 24 Monate	75 (12,3)	51 (8,3)	
t = 27 Monate	49 (8,0)	31 (5,0)	
t = 30 Monate	33 (5,4)	14 (2,3)	
Überlebenswahrscheinlichkeit [95%-KI]			
t = 3 Monate	0,931 [0,911;0,951]	0,935 [0,915;0,955]	
t = 6 Monate	0,819 [0,788;0,850]	0,791 [0,759;0,823]	
t = 9 Monate	0,687 [0,651;0,724]	0,654 [0,617;0,691]	
t = 12 Monate	0,561 [0,522;0,601]	0,503 [0,463;0,542]	
t = 18 Monate	0,385 [0,344;0,426]	0,309 [0,270;0,347]	
t = 24 Monate	0,280 [0,238;0,323]	0,187 [0,150;0,224]	
t = 27 Monate	0,238 [0,195;0,282]	0,158 [0,120;0,195]	
t = 30 Monate	0,223 [0,179;0,267]	0,120 [0,080;0,159]	
Hazard Ratio** [95,34%*-KI] [95%-KI]	0,817 [0,713;0,937] [0,714;0,934]		
p-Wert des Log-Rank***-Test	0,0032		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

Mediane Follow-up Zeit für die ITT-Population zum Cutoff-Datum (07. Februar 2011, bei 863 Todesfällen) war 22,28 Monate

* $\alpha=0,0466$ (O'Brien Fleming- α -Korrektur)

** stratifiziert nach ECOG-PS und Vorbehandlung mit Bevacizumab

Annahme der Proportionalität der Überlebenskurve für den Log-Rank-Test mittels Grafik (siehe Abbildung 2) bestätigt

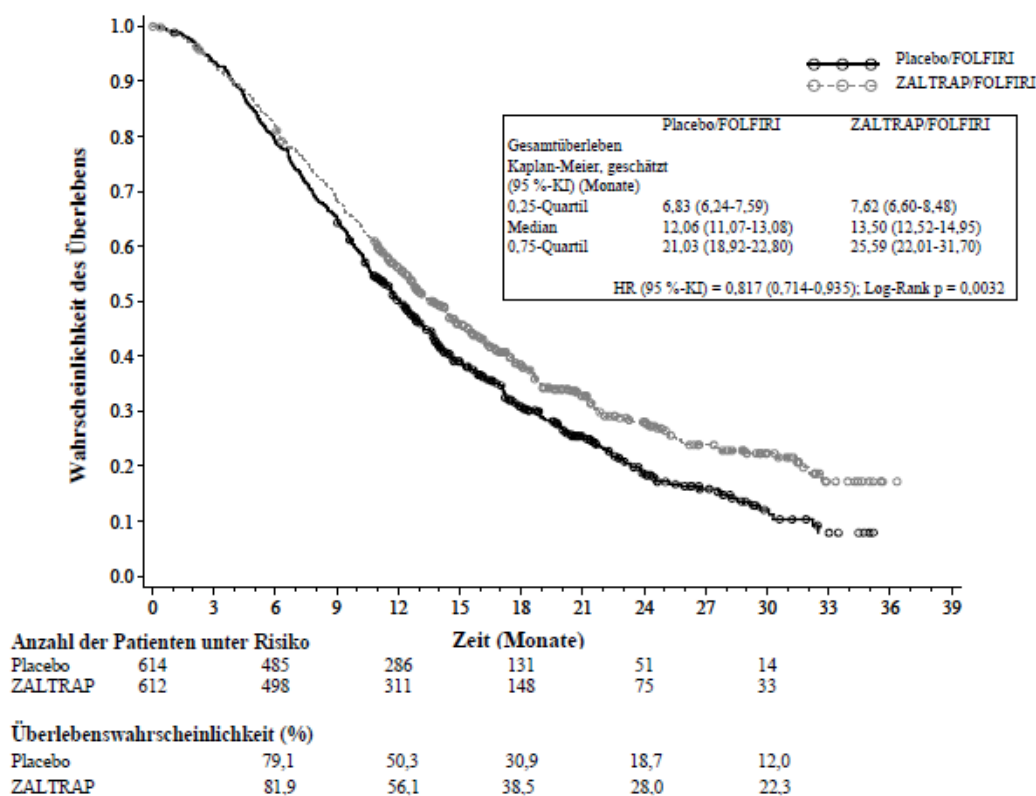


Abbildung 2: Kaplan-Meier Kurven für Gesamtüberleben (OS) (in Monaten) in der ITT-Population

Die Analyse des OS (siehe Tabelle 4-17) zeigte eine signifikante Verbesserung des Überlebens für Patienten mit einer FOLFIRI+Aflibercept-Behandlung gegenüber einer FOLFIRI+Placebo-Behandlung ($p=0,0032$ anhand des Log-Rank-Tests). Die mediane Überlebenszeit für die Patienten im FOLFIRI+Aflibercept-Arm lag bei 13,50 (95%-KI: [12,52;14,95]) Monaten, im Gegensatz dazu betrug die mediane Überlebenszeit im FOLFIRI+Placebo-Arm lediglich 12,06 (95%-KI: [11,07;13,08]) Monate. Dies entspricht einer positiven Differenz von 1,44 Monaten im Aflibercept-Arm. Das Hazard Ratio für die Mortalität betrug 0,817 (95%-KI: [0,714;0,934]) zugunsten der FOLFIRI+Aflibercept-Gruppe; das bedeutet eine 18,3%-ige Reduzierung des Sterberisikos für die Patienten, die mit FOLFIRI+Aflibercept behandelt wurden, gegenüber Patienten, die FOLFIRI+Placebo erhielten. Die frühe und anhaltende Separierung der Überlebenskurven im Kaplan-Meier-Diagramm (Abbildung 2) deutet auf einen robusten und anhaltenden Effekt hin. Die Überlebenswahrscheinlichkeit nach 18 Monaten beträgt in der FOLFIRI+Aflibercept-Gruppe 38,5% (95%-KI: [34,4;42,6]) und in der FOLFIRI+Placebo-Gruppe 30,9% (95%-KI: [27,0;34,7]). Nach 24 Monaten beträgt diese in der FOLFIRI+Aflibercept-Gruppe 28,0% (95%-KI: [23,8;32,3]) dagegen in der FOLFIRI+Placebo-Gruppe lediglich 18,7% (95%-KI: [15,0;22,4]). Auch nach 30 Monaten liegt die Überlebenswahrscheinlichkeit in der FOLFIRI+Aflibercept-Gruppe mit 22,3% (95%-KI: [17,9;26,7]) oberhalb derjenigen in der FOLFIRI+Placebo-Gruppe mit 12% (95%-KI: [8,0;15,9]).

Tabelle 4-18: Ergebnisse für Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalysen

Studie			
VELOUR			
Unstratifizierte Analyse:			
Hazard Ratio [95,34%*-KI] [95%-KI]	0,809 [0,706;0,927] [0,707;0,925]		
p-Wert des Log-Rank-Test	0,0019		
Multivariates** Cox-Proportional Hazard Modell:			
Hazard Ratio** [95,34%*-KI] [95%-KI]	0,817 [0,713;0,936] [0,714;0,934]		
p-Wert des Log-Rank-Test	0,0032		
Mittleres Gesamtüberleben (in Monaten)			
	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)	Differenz in Monaten [95%-KI]
Log-logistische Verteilung (beide Gruppen separat)	22,8	18,1	4,7 [2,1;6,1]
Log-logistische Verteilung (mit der Behandlung als Kovariable)	21,5	19,0	2,5
Weibull-Verteilung (beide Gruppen separat)	17,9	14,9	3,0 [1,2;4,2]
Weibull-Verteilung (mit der Behandlung als Kovariable)	17,7	15,0	2,7

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* $\alpha=0,0466$ (O'Brian Fleming- α -Korrektur)

** Auswahl prognostische Faktoren mittels *Edwards* Strategie: ECOG-PS, Vorbehandlung mit Bevacizumab, Alter, Hypertonie, Anzahl Organe mit Metastasen

Die Ergebnisse der gemäß Protokoll durchgeführten Sensitivitätsanalyse (siehe Tabelle 4-18) bestätigen die Robustheit der Analyse des primären Endpunktes.

Zusätzlich wurde auch die mittlere Überlebenszeit berechnet [34]. Dabei wurden die Kaplan-Meier-Kurven mit Survivalfunktionen extrapoliert, wobei nach gängigen Modellwahlkriterien (Akaikes Informationskriterium, Bayessesches Informationskriterium sowie graphische Methoden) ausgewählte Modelle mit log-logistischer Verteilung zur besten Anpassung an die Daten in der ITT-Population führten. In einem ersten Modell wurde für jede Behandlungsgruppe eine separate Survivalfunktion angepasst. Dadurch ergab sich in der FOLFIRI+Aflibercept-Gruppe eine mittlere Überlebenszeit von 22,8 Monaten gegenüber 18,1 Monaten in der FOLFIRI+Placebo-Gruppe, was einer Differenz von 4,7 Monaten

(95%-KI: [2,1;6,1]) zugunsten von Aflibercept entspricht. In einem zweiten Modell wurde eine Kovariable für die Behandlungsart hinzugefügt und eine Survivalfunktion an die Daten aus beiden Behandlungsarmen angepasst. In diesem Modell ergab sich in der FOLFIRI+Aflibercept-Gruppe eine mittlere Überlebenszeit von 21,5 Monaten gegenüber 19,0 Monaten in der FOLFIRI+Placebo-Gruppe, was einer Differenz von 2,5 Monaten zugunsten von Aflibercept entspricht. Als Sensitivitätsanalyse wurden die beiden genannten Modelle analog mit einer Weibull-Verteilung durchgeführt. Auch hier sind die mittleren Überlebenszeiten in der FOLFIRI+Aflibercept-Gruppe länger als in der FOLFIRI+Placebo-Gruppe (Differenz 3,0 Monate im Modell mit separaten Survivalfunktionen und 2,7 Monate im Modell mit Kovariable für Behandlung).

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

<Abbildung Meta-Analyse>

Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie zur Quantifizierung des Zusatznutzens von Aflibercept herangezogen wird.

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

4.3.1.3.2 Progressionsfreies Überleben (PFS) – RCT

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

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

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung

- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

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

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

Tabelle 4-19: Operationalisierung von Progressionsfreies Überleben (PFS)

Studie	Operationalisierung
VELOUR	Progressionsfreies Überleben (PFS) ist definiert als das Zeitintervall zwischen Randomisierung und dem ersten Tag, an dem eine Progression der Krankheit beobachtet wurde, bzw. Tod jedweder Ursache, je nachdem, welches Ereignis zuerst eintrat.

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

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

Tabelle 4-20: Bewertung des Verzerrungspotenzials für Progressionsfreies Überleben (PFS) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

PFS wurde als sekundärer Endpunkt der VELOUR-Studie erfasst und ist definiert als Zeitintervall zwischen der Randomisierung und dem ersten Tag, an dem eine Progression der Erkrankung beobachtet wurde, bzw. Tod jedweder Ursache, je nachdem, welches Ereignis

zuerst eintrat. Die PFS-Analyse wurde am Cutoff-Datum (06. Mai 2010) durchgeführt, welches dem Zeitpunkt der 2. Zwischenanalyse für das OS entspricht. Die Durchführung erfolgte innerhalb der ITT-Population und basiert auf einer radiologischen Tumorbewertung durch ein Independent Review Committee (IRC), welches den Zustand der Progression jedes einzelnen Patienten bestimmte. Die Verblindung des Endpunkterhebers war somit gewährleistet. Des Weiteren konnten keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Daher wird das Verzerrungspotential als **niedrig** beurteilt.

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

Tabelle 4-21: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel

Studie			
VELOUR	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)	
Ereignisse, n (%)	393 (64,2)	454 (73,9)	
Zensierte Fälle, n (%)	219 (35,8)	160 (26,1)	
Diff. in Monaten			
25%-Quantil für progressionsfreies Überleben in Monaten [99,99%*-KI] [95%-KI]	3,84 [2,79;4,99] [3,22;4,30]	2,60 [1,61;2,96] [1,97;2,76]	1,24
Medianes progressionsfreies Überleben in Monaten [99,99%*-KI] [95%-KI]	6,90 [5,88;7,85] [6,51;7,20]	4,67 [4,07;5,55] [4,21;5,36]	2,23
75%-Quantil für progressionsfreies Überleben in Monaten [99,99%*-KI] [95%-KI]	10,05 [8,77;11,79] [9,40;10,74]	9,10 [6,97;11,07] [7,89;9,89]	0,95
Kaplan-Meier Kenngrößen			
Patienten unter Risiko, n (%)			
t = 3 Monate	420 (68,6)	355 (78,2)	
t = 6 Monate	247 (40,4)	171 (37,7)	
t = 9 Monate	99 (16,2)	94 (20,7)	
t = 12 Monate	43 (7,0)	46 (10,1)	
t = 18 Monate	7 (1,1)	9 (2,0)	
Hazard Ratio** [99,99%*-KI] [95%-KI]	0,758 [0,578;0,995] [0,661;0,867]		
p-Wert des Log-Rank***-Test	<0,0001		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

Cutoff-Datum 06. Mai 2010

* $\alpha=0,0001$ (O'Brien Fleming- α -Korrektur)

** stratifiziert nach ECOG-PS und Vorbehandlung mit Bevacizumab

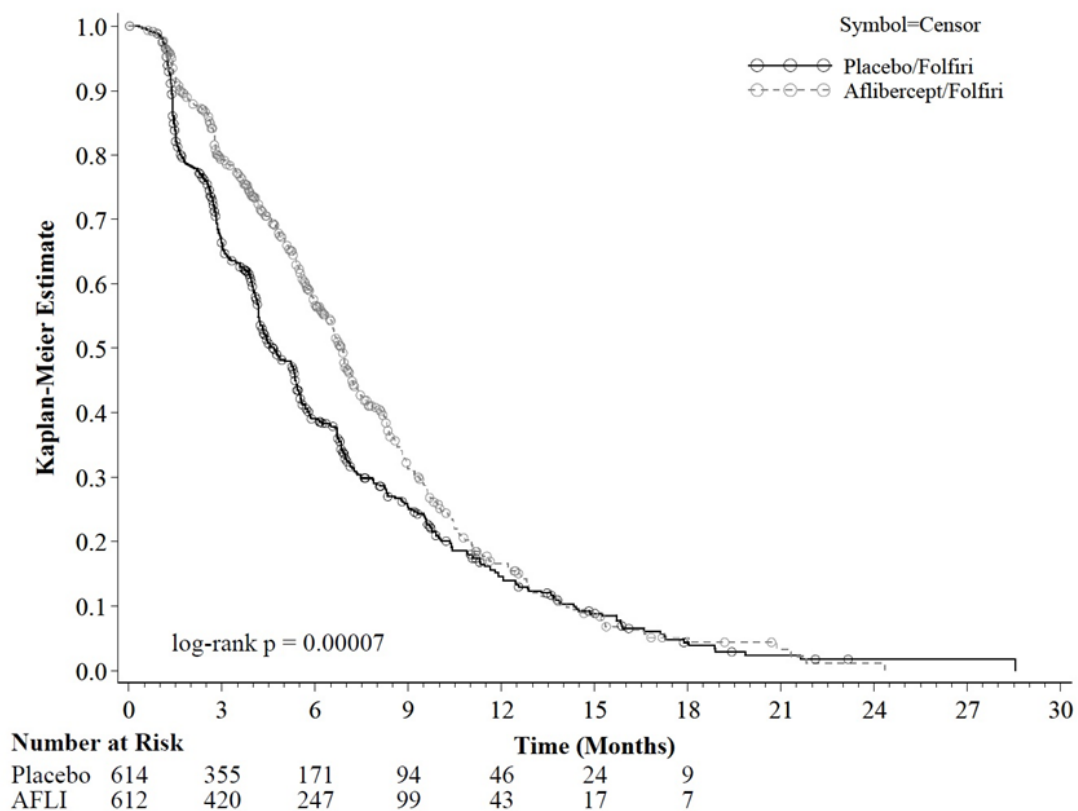


Abbildung 4: Kaplan-Meier-Kurven für Progressionsfreies Überleben (PFS) (in Monaten) in der ITT-Population

Die Analyse des PFS (Tabelle 4-21) basiert auf einer Gesamtanzahl von 847 Ereignissen bis zum Cutoff-Datum am 06. Mai 2010, davon 393 in der FOLFIRI+Aflibercept-Gruppe und 454 in der FOLFIRI+Placebo-Gruppe. Das mediane PFS betrug in der FOLFIRI+Aflibercept-Gruppe 6,90 Monate (95%-KI: [6,51;7,20]) und in der FOLFIRI+Placebo-Gruppe lediglich 4,67 Monate (95%-KI: [4,21;5,36]). Die PFS-Kurven der beiden Behandlungsgruppen wurden mithilfe eines Log-Rank-Tests verglichen, welcher signifikante Unterschiede ($p=0,00007$) zwischen den Gruppen, insbesondere während der ersten neun Beobachtungsmonate, zeigte (Abbildung 4). Das Hazard Ratio des zugrundeliegenden Cox Proportional-Hazard-Modells betrug 0,758 (95%-KI: [0,661;0,867]).

Tabelle 4-22: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse.

Studie		
VELOUR	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)
Sensitivitätsanalyse Nr. 1^a:		
Anzahl Patienten, n	281	353
Medianes progressionsfreies Überleben in Monaten [99,99%*-KI]	6,97 [6,045;8,509]	4,53 [4,074;5,684]
Hazard Ratio** [99,99%*-KI]	0,654 [0,477;0,895]	
p-Wert des Log-Rank-Test**	<0,0001	
Sensitivitätsanalyse Nr. 2^b:		
Anzahl Patienten, n	452	485
Medianes progressionsfreies Überleben in Monaten [99,99%*-KI]	6,24 [5,487;7,195]	4,50 [4,041;5,552]
Hazard Ratio** [99,99%*-KI]	0,814 [0,630;1,052]	
p-Wert des Log-Rank-Test**	0,0017	

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

Cutoff-Datum 6. Mai 2010

* $\alpha=0,0001$ (O'Brien Fleming- α -Korrektur)

** stratifiziert nach ECOG-PS und Vorbehandlung mit Bevacizumab

^a In Sensitivitätsanalyse Nr. 1 wurden diejenigen Patienten, bei denen eine Progression (oder Tod) mehr als neun Wochen nach der letzten Beurteilung des Tumors dokumentiert wurde, zensiert, und zwar zum Zeitpunkt dieser Tumorbeurteilung. Patienten, die weiterhin eine antineoplastische Therapie erhielten, ohne dass eine Progression dokumentiert wurde, wurden zum Zeitpunkt der letzten gültigen Tumorbeurteilung zensiert.

^b In Sensitivitätsanalyse Nr. 2 wurde das PFS unter Berücksichtigung der Läsionsbewertung des Prüfarztes bestimmt und analysiert, wobei zusätzlich die klinische Progression berücksichtigt wurde.

Die Ergebnisse der gemäß Protokoll durchgeführten Sensitivitätsanalysen (siehe Tabelle 4-22) bestätigen die Robustheit der Analyse des sekundären Endpunktes PFS.

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

<Abbildung Meta-Analyse>

Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie zur Quantifizierung des Zusatznutzens von Aflibercept herangezogen wird.

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.

4.3.1.3.3 Objektive Ansprechrate (ORR) – RCT

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

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

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

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

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

Tabelle 4-23: Operationalisierung von Objektive Ansprechrate (ORR)

Studie	Operationalisierung
VELOUR	Die objektive Ansprechrate (komplette Remission oder partielle Remission) wurde gemäß RECIST-Kriterien (Version 1.0) erhoben.

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

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

Tabelle 4-24: Bewertung des Verzerrungspotenzials für Objektive Ansprechrates (ORR) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Ermittlung der ORR obliegt einer objektiven Bewertung nach RECIST (Response Evaluation Criteria in Solid Tumors, Version 1.0 [23]) und wurde durch das IRC bewertet. Darin ist festgelegt, nach welchen Kriterien bei einem Patient eine komplette oder partielle Remission, bzw. eine stabile oder progressive Krankheit vorliegt. Es wurde der Anteil der Patienten („Responder“) berücksichtigt, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression einem kompletten Ansprechen (CR, complete response) oder partiellen Ansprechen (PR, partial response) entsprach. Die ORR wurde nur bei Patienten mit messbarer Erkrankung analysiert. Aufgrund verschiedener Ausschlusskriterien (z. B. das radiologische Nichtvorhandensein von Zielläsionen zu Studienbeginn) reduzierte sich die Patientenzahl in der FOLFIRI+Aflibercept-Gruppe von 612 auf 531 und in der FOLFIRI+Placebo-Gruppe von 614 auf 530. Das ITT-Prinzip wurde dennoch adäquat umgesetzt. Des Weiteren konnten keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Daher wird das Verzerrungspotential für diesen Endpunkt als **niedrig** beurteilt.

Stellen Sie die Ergebnisse für den Endpunkt Objektive Ansprechrate (ORR) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-25: Ergebnisse für Objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel

Studie		
VELOUR	FOLFIRI+Aflibercept (N = 531*)	FOLFIRI+Placebo (N = 530*)
Komplettes Ansprechen, n (%)	0	2 (0,4)
Partielles Ansprechen, n (%)	105 (19,8)	57 (10,8)
Stabile Krankheit, n (%)	350 (65,9)	344 (64,9)
Progressive Krankheit, n (%)	55 (10,4)	114 (21,5)
Nicht auswertbar, n (%)	21 (4,0)	13 (2,5)
Responder: Komplettes oder partielles Ansprechen, n (%)	105 (19,8)	59 (11,1)
Odds Ratio** [95%-KI]	1,968 [1,394;2,778]	
p-Wert des stratifizierten Cochran-Mantel-Haenszel Test***	0,0001	
Relatives Risiko** [95%-KI]	1,776 [1,322;2,387]	
p-Wert***	0,0001	

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Auswertbare Population

** Wert stammt aus eigener Berechnung

*** stratifiziert nach ECOG-PS und Vorbehandlung mit Bevacizumab

Die ORR (komplette + partielle Remission) in der auswertbaren Population beträgt 19,8% (95%-KI: [16,4;23,2]) in der FOLFIRI+Aflibercept-Gruppe gegenüber 11,1% (95%-KI: [8,5;13,8]) in der FOLFIRI+Placebo-Gruppe. Das Odds Ratio für die ORR zwischen dem FOLFIRI+Aflibercept- und dem FOLFIRI+Placebo-Arm beträgt 1,968 (95%-KI: [1,394;2,778]), das relative Risiko (RR) 1,776 (95%-KI: [1,322;2,387]). Der stratifizierte Cochran-Mantel-Haenszel Test zeigt einen signifikanten Unterschied (für OR und RR: $p=0,0001$) in der Ansprechrate zwischen beiden Gruppen, zugunsten der FOLFIRI+Aflibercept-Gruppe.

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

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

<Abbildung Meta-Analyse>

Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie zur Quantifizierung des Zusatznutzens von Aflibercept herangezogen wird.

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.

4.3.1.3.4 Sicherheit (Unerwünschte Ereignisse) – RCT

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

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

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

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

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

Tabelle 4-26: Operationalisierung von Sicherheit (Unerwünschte Ereignisse)

Studie	Operationalisierung
VELOUR	Anteil der Patienten, die während ihrer Behandlungszeit mindestens ein unerwünschtes Ereignis (UE) oder ein schwerwiegendes unerwünschtes Ereignis (SUE) erfahren haben Anteil der Patienten, die die Behandlung aufgrund des UE abgebrochen haben Anteil der Patienten, bei denen das UE zum Tod führte Anteil der Patienten mit einem unerwünschten Ereignis von besonderem Interesse

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

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

Tabelle 4-27: Bewertung des Verzerrungspotenzials für Sicherheit (Unerwünschte Ereignisse) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Bewertung des Verzerrungspotentials für unerwünschte Ereignisse erfolgte zusammenfassend für alle betrachteten Operationalisierungen. Die Meldung und Berichterstattung unerwünschter Ereignisse in der Studie erfolgte fortlaufend und unter Beachtung der GCP-Standards [35]. Von den 1.226 randomisierten Patienten erhielten 1.216 eine Behandlung (611 im FOLFIRI+Aflibercept-Arm und 605 im FOLFIRI+Placebo-Arm) und bildeten die Sicherheitspopulation. Es wurden keine Hinweise auf ereignisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte gefunden. Das Verzerrungspotential wird daher als **niedrig** bewertet.

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

Tabelle 4-28: Ergebnisse für Sicherheit (Unerwünschte Ereignisse) aus RCT mit dem zu bewertenden Arzneimittel – Sicherheitspopulation

VELOUR	Studie		Relatives Risiko* [95%-KI] p-Wert**	Odds Ratio* [95%-KI] p-Wert**
	FOLFIRI+ Aflibercept (N = 611)	FOLFIRI+ Placebo (N = 605)		
Patienten mit UE, n (%)	606 (99,2)	592 (97,9)	1,01 [1,00;1,03] 0,0557	2,66 [0,94;7,51] 0,0645
Patienten mit UE Grad 3 oder 4, n (%)	510 (83,5)	378 (62,5)	1,34 [1,24;1,43] <0,0001	3,03 [2,32;3,97] <0,0001
Patienten mit einem SUE***, n (%)	294 (48,1)	198 (32,7)	1,47 [1,28;1,69] <0,0001	1,91 [1,51;2,41] <0,0001
Patienten mit einem UE, welches zum Tod führte, n (%)	37 (6,1)	29 (4,8)	1,26 [0,79;2,03] 0,3327	1,28 [0,78;2,11] 0,3324
Patienten mit UE, welches zum Abbruch der Behandlung führte, n (%)	164 (26,8)	73 (12,1)	2,22 [1,73;2,86] <0,0001	2,67 [1,97;3,62] <0,0001

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall; UE: Unerwünschtes Ereignis; SUE: Schwerwiegendes unerwünschtes Ereignis

* Werte stammen aus eigener Berechnung

** p-Wert des Mantel-Haenszel Test für relatives Risiko/Odds-Ratio

*** Schweregrad des Ereignisses wurde durch NCI-CTC Version 3.0 bewertet.

Fast alle Studienteilnehmer aus dieser Population gaben mindestens ein UE an, und zwar 606 Patienten (99,2%) in der FOLFIRI+Aflibercept-Gruppe und 592 Patienten (97,9%) in der FOLFIRI+Placebo-Gruppe, wobei dieser Unterschied nicht signifikant ist ($p=0,0645$). Bei 37 Patienten (6,1%) der FOLFIRI+Aflibercept-Gruppe sowie bei 29 Patienten (4,8%) in der FOLFIRI+Placebo-Gruppe führte das UE zum Tod, wobei dieser Unterschied ebenfalls nicht signifikant ist ($p=0,3324$). UEs vom Grad 3 oder 4 traten bei Patienten in der FOLFIRI+Aflibercept-Gruppe (510 Patienten, 83,5%) signifikant häufiger als in der FOLFIRI+Placebo-Gruppe (378 Patienten, 62,5%) auf ($p<0,0001$). Ebenso traten SUEs bei Patienten in der FOLFIRI+Aflibercept-Gruppe (294 Patienten, 48,1%) signifikant häufiger als in der FOLFIRI+Placebo-Gruppe (198 Patienten, 32,7%) auf ($p<0,0001$). Im Behandlungsarm FOLFIRI+Aflibercept brachen 164 Patienten (26,8%) die Behandlung aufgrund des UE ab, im Behandlungsarm FOLFIRI+Placebo waren es 73 Patienten (12,1%). Dieser Unterschied ist ebenfalls signifikant ($p<0,0001$). Die häufigsten Ursachen für einen Behandlungsabbruch waren in beiden Behandlungsarmen Müdigkeit und Asthenie, Infektionen, Durchfall, Myelosuppression, Hypertonie, Lungenembolie, Proteinurie und tiefe Venenthrombose [1].

Tabelle 4-29: Unerwünschte Ereignisse mit mehr als 5 Prozentpunkten Unterschied zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe

Studie				
VELOUR	FOLFIRI+ Aflibercept (N = 611)	FOLFIRI+ Placebo (N = 605)	Relatives Risiko* [95%-KI] p-Wert**	Odds Ratio* [95%-KI] p-Wert**
Infektionen und parasitäre Erkrankungen, n (%)	282 (46,2)	198 (32,7)	1,41 [1,22;1,63] <0,0001	1,76 [1,40;2,22] <0,0001
Erkrankungen des Blutes und des Lymphgefäßsystems, n (%)	270 (44,2)	232 (38,3)	1,15 [1,01;1,32] 0,0391	1,27 [1,01;1,60] 0,0387
Stoffwechsel- und Ernährungsstörungen, n (%)	233 (38,1)	160 (26,4)	1,44 [1,22;1,70] <0,0001	1,71 [1,34;2,19] <0,0001
Erkrankungen des Nervensystems, n (%)	280 (45,8)	226 (37,4)	1,23 [1,07;1,40] 0,0029	1,42 [1,12;1,78] 0,0028
Gefäßerkrankungen, n (%)	299 (48,9)	137 (22,6)	2,16 [1,83;2,56] <0,0001	3,27 [2,55;4,19] <0,0001
Erkrankungen der Atemwege, des Brustraums und des Mediastinums, n (%)	359 (58,8)	215 (35,5)	1,65 [1,46;1,88] <0,0001	2,58 [2,05;3,26] <0,0001
Erkrankungen des Gastrointestinaltrakts, n (%)	571 (93,5)	521 (86,1)	1,09 [1,04;1,13] <0,0001	2,30 [1,55;3,42] <0,0001
Erkrankungen der Nieren oder Harnwege, n (%)	117 (19,1)	56 (9,3)	2,07 [1,54;2,79] <0,0001	2,32 [1,65;3,27] <0,0001
allgemeine Erkrankungen, wie z. B. Müdigkeit oder Kraftlosigkeit, n (%)	437 (71,5)	378 (62,5)	1,14 [1,06;1,24] 0,0009	1,51 [1,19;1,92] 0,0008

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Werte stammen aus eigener Berechnung

** p-Wert des Mantel-Haenszel Test für relatives Risiko/Odds Ratio

Die häufigsten UEs mit mehr als 5 Prozentpunkten Unterschied zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe waren gastrointestinale Erkrankungen, die 571 Patienten (93,5%) in der FOLFIRI+Aflibercept-Gruppe und 521 Patienten (86,1%) in der FOLFIRI+Placebo-Gruppe angaben, gefolgt von allgemeinen Beschwerden, wie z. B. Müdigkeit oder Kraftlosigkeit, welche 437 Patienten (71,5%) in der FOLFIRI+Aflibercept-Gruppe und 378 Patienten (62,5%) in der FOLFIRI+Placebo-Gruppe berichteten. Die deutlichsten Unterschiede in der Häufigkeit des Auftretens zwischen beiden Armen bestanden bei Gefäßerkrankungen (48,9% vs. 22,6%) sowie bei Erkrankungen der Atemwege, des Brustraums und des Mediastinums (58,8% vs. 35,5%), in beiden Fällen zuungunsten von Aflibercept. Für alle UEs, die in der FOLFIRI+Aflibercept-Gruppe mit mehr als 5 Prozentpunkten Unterschied zur FOLFIRI+Placebo-Gruppe auftraten, sind die

Odds Ratios signifikant (p -Wert $<0,05$) größer als Eins, d. h. das Risiko für ein UE in der FOLFIRI+Aflibercept-Gruppe ist signifikant höher als in der FOLFIRI+Placebo-Gruppe.

In der Erhebung der SUEs wurden gastrointestinale Erkrankungen von 124 Patienten (20,3%) der FOLFIRI+Aflibercept-Gruppe und von 68 Patienten (11,2%) der FOLFIRI+Placebo-Gruppe angegeben [1]. Das Odds Ratio beträgt 2,01 (95%-KI: [1,46;2,77]) und ist signifikant ($p<0,0001$) (eigene Berechnung des p -Wertes basierend auf dem Mantel-Haenszel Test für RR/Odds Ratio).

Tabelle 4-30: Präspezifizierte unerwünschte Ereignisse von besonderem Interesse. Vergleich zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe

Studie				
VELOUR	FOLFIRI+Aflibercept (N =611)	FOLFIRI+Placebo (N = 605)	Relatives Risiko* [95%-KI] p-Wert**	Odds Ratio* [95%-KI] p-Wert**
Hämorrhagie, n (%)	231 (37,8)	115 (19,0)	1,99 [1,64;2,41] <0,0001	2,59 [2,00;3,36] <0,0001
Hypertonie, n (%)	253 (41,4)	65 (10,7)	3,85 [3,01;4,94] <0,0001	5,87 [4,33;7,95] <0,0001
venöse thromboembolische Ereignisse, n (%)	57 (9,3)	44 (7,3)	1,28 [0,88;1,87] 0,1954	1,31 [0,87;1,98] 0,195
Proteinurie, n (%)	380 (62,2)	246 (40,8)	1,53 [1,36;1,71] <0,0001	2,40 [1,91;3,02] <0,0001

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Werte stammen aus eigener Berechnung

** p-Wert des Mantel-Haenszel Test für relatives Risiko/Odds Ratio

In der Analyse nach Studienprotokoll präspezifizierter UEs zeigt sich, dass die größten Unterschiede zwischen den Behandlungsgruppen FOLFIRI+Aflibercept und FOLFIRI+Placebo im Auftreten von Hämorrhagie, Hypertonie und Proteinurie lagen (p -Werte $<0,0001$) [1]. Diese Effekte sind auf einen sog. Klasseneffekt (häufig beobachtete UEs bei der Anwendung von anti-VEGF-Wirkstoffen) zurückzuführen, weshalb man bereits bei der Studienplanung ein besonderes Augenmerk darauf legte. Der erhöhte Anteil von Hämorrhagien in der FOLFIRI+Aflibercept-Gruppe resultiert auch aus dem vermehrten Auftreten von gastrointestinalen Beschwerden in dieser Gruppe.

Das Odds Ratio für Hämorrhagie beträgt 2,59 (95%-KI: [2,00;3,36]), für Hypertonie 5,87 (95%-KI: [4,33;7,95]) und für Proteinurie 2,40 (95%-KI: [1,91;3,02]), wobei die p -Werte jeweils kleiner als 0,0001 sind.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die

Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es existiert eine auf Basis einer Meta-Analyse zusammengefasste Auswertung der Ergebnisse zu UEs der Aflibercept-Studien VELOUR, VANILLA und VITAL [36]. Die Ergebnisse dieser Auswertung wurden auch in die Fachinformation von Aflibercept übernommen. Es ist jedoch zu betonen, dass sich die Studien VANILLA und VITAL auf andere Tumorentitäten beziehen, und Aflibercept dort in Kombination mit anderen Chemotherapeutika angewendet wurde. Daten der genannten Studien und damit auch die erwähnte Meta-Analyse konnten somit auf Basis der in Abschnitt 4.2.2 spezifizierten Ein- und Ausschlusskriterien auch nicht in die vorliegende Bewertung mit eingeschlossen werden. Vor diesem Hintergrund spielen die Ergebnisse der genannten Auswertung eine eher untergeordnete Rolle. Sie flossen zum Teil in die Beschreibung des Nebenwirkungsprofils in der Fachinformation von Aflibercept [3] mit ein, werden hier jedoch nicht im Detail aufgeführt oder diskutiert. In der Fachinformation wird z. B. das Auftreten eines posterioren reversiblen Enzephalopathie-Syndroms, PRES, zitiert, das in der VELOUR-Studie nicht beobachtet wurde. Die wichtigsten Aussagen der Metaanalyse sind kurz zusammengefasst [36]:

- Klasseneffekte, d. h. Risiken, welche bei Anwendung von VEGF-Inhibitoren häufig beobachtet werden, treten auch unter Aflibercept verstärkt auf.
- Die Häufigkeit des Auftretens von tiefen Venenthrombosen war im Gegensatz unter Aflibercept nicht erhöht.
- Einige generell bei Behandlung mit Chemotherapien oft auftretende Nebenwirkungen werden offenbar bei gleichzeitiger Gabe von Aflibercept verstärkt; speziell in der VELOUR-Studie betraf dies Neutropenie, Magenbeschwerden, Durchfall, Dehydratation sowie Hand-Fuß-Syndrom.

Es wurde keine Meta-Analyse zur Sicherheit von Aflibercept innerhalb der für das vorliegende Dossier relevanten Indikation durchgeführt.

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

4.3.1.3.5 Subgruppenanalysen – RCT

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

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

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

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

4.3.1.3.5.1 Präspezifizierte Subgruppenanalysen

Subgruppenanalysen dienen der Überprüfung des Effekts in einzelnen Untergruppen, und stellen somit ein Hilfsmittel dar, um die Konsistenz der Analysen aus der gesamten Studienpopulation zu festigen. Hierfür werden vor Beginn der Studie bestimmte Subgruppen spezifiziert, die sich nach demographischen und Baseline-Charakteristika sowie prognostischen Faktoren richten. Liegt in der Analyse der Wirksamkeit oder Sicherheit des Medikaments ein Unterschied in den Subgruppen vor, so spricht man von einer Interaktion zwischen Behandlungsarm und Subgruppenmerkmal. Die Aussagekraft von Subgruppenanalysen wird jedoch durch verschiedene Faktoren (z. B. Größe und Anzahl der Subgruppen) beeinflusst und ist deshalb mit Vorsicht zu betrachten. So kann es durch Bildung zahlreicher Subgruppen und dem damit verbundenem multiplen Testen zu signifikanten Unterschieden zwischen Subgruppen kommen, welche in Wahrheit auf Zufall basieren [6]. Zusätzliche Diskussionen, auch zum verwendeten Signifikanzniveau, finden sich in Abschnitt 4.2.5.5.

Für den primären Endpunkt OS und den sekundären Endpunkt PFS wurden in der VELOUR-Studie präspezifizierte Subgruppenanalysen für folgende Merkmale durchgeführt: ECOG-PS, Vorbehandlung mit Bevacizumab, Alter, Geschlecht, ethnische Zugehörigkeit, geografische Region, vorbestehender Hypertonus, Anzahl der von Metastasen betroffener Organe, ausschließlich Lebermetastasen zu Studienbeginn und Lokalisation des Primärtumors. Die präspezifizierten Subgruppenanalysen für OS und PFS zeigen fast ausnahmslos einen quantitativ übereinstimmenden Behandlungseffekt zugunsten FOLFIRI+Aflibercept (Hazard Ratio <1,0).

4.3.1.3.5.1.1 Subgruppenanalysen zum Gesamtüberleben (OS)

Tabelle 4-31: Interaktionstests der präspezifizierten Subgruppenanalyse für den primären Endpunkt Gesamtüberleben (OS)

Studie		
VELOUR		
Merkmal	Subgruppen	p-Wert der Interaktion*
ECOG-PS	0/1/2	0,7231
Vorbehandlung mit Bevacizumab	ja/nein	0,5668
Alter	<65/≥65	0,6830
Geschlecht	männlich/weiblich	0,5941
Ethnische Zugehörigkeit	Kaukasisch+weiß/andere	0,9687
Geographische Region	Westeuropa/Osteuropa/Nordamerika, Südamerika/andere Kontinente	0,6526
Vorbestehender Hypertonus	ja/nein	0,1309
Anzahl Organe, die von Metastasen betroffen sind	>1/≤1	0,6992
Ausschließlich Lebermetastasen	ja/nein	0,0899
Lokalisation des Primärtumors	Kolon/Rektosigmoid/andere vs. Rektum	0,1421

ECOG-PS: Eastern Cooperative Oncology Group-Performance Status

* Interaktionstest des Cox-Modell mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor

Für den Endpunkt OS wurde für alle Subgruppenmerkmale ein Interaktionstest des Cox-Modells mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor durchgeführt (siehe Tabelle 4-31). Dabei wurde jeweils geprüft, ob sich die Hazard Ratios der Subgruppen eines Merkmals signifikant unterscheiden.

Hinweis: Im CSR zur VELOUR-Studie wurden irrtümlicherweise die p-Werte der Interaktionstest zu den Subgruppenanalysen ECOG-PS 0/1/2 und Bevacizumab-Vorbehandlung ja/nein sowohl für OS als auch für PFS vertauscht (Tabelle 4-31 und Tabelle 4-35) [37].

Tabelle 4-32: Effektmodifikationen (p-Wert<0,2) in den präspezifizierten Subgruppen für Gesamtüberleben (OS)

Studie						
VELOUR						
Merkmal	FOLFIRI+Aflibercept (N = 612)		FOLFIRI+Placebo (N = 614)		Hazard Ratio [95,34%-KI*] [95%-KI]	
	Patienten, n	Todesfälle, n (%)	Patienten, n	Todesfälle, n (%)		
Vorbestehender Hypertonus						
<i>Nein</i>	346	245 (70,8)	346	262 (75,5)	0,883 [0,740;1,054] [0,742;1,051]	
<i>Ja</i>	266	158 (59,4)	268	198 (73,9)	0,714 [0,577;0,884] [0,574;0,879]	
Ausschließlich Lebermetastasen						
<i>Nein</i>	459	307 (66,9)	468	345 (73,7)	0,868 [0,742;1,015] [0,744;1,012]	
<i>Ja</i>	153	96 (62,7)	146	115 (78,8)	0,649 [0,492;0,855] [0,494;0,851]	
Lokalisation des Primärtumors						
<i>Kolon</i>	289	189 (65,4)	302	231 (76,5)	0,739 [0,607;0,899] [0,609;0,896]	
<i>Rektosigmoid + andere</i>	126	84 (66,7)	138	96 (69,6)	1,039 [0,772;1,400] [0,775;1,393]	
<i>Rektum</i>	197	130 (66,0)	174	133 (76,4)	0,806 [0,629;1,031] [0,632;1,027]	

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* $\alpha=0,0466$ (O'Brian Fleming- α -Korrektur)

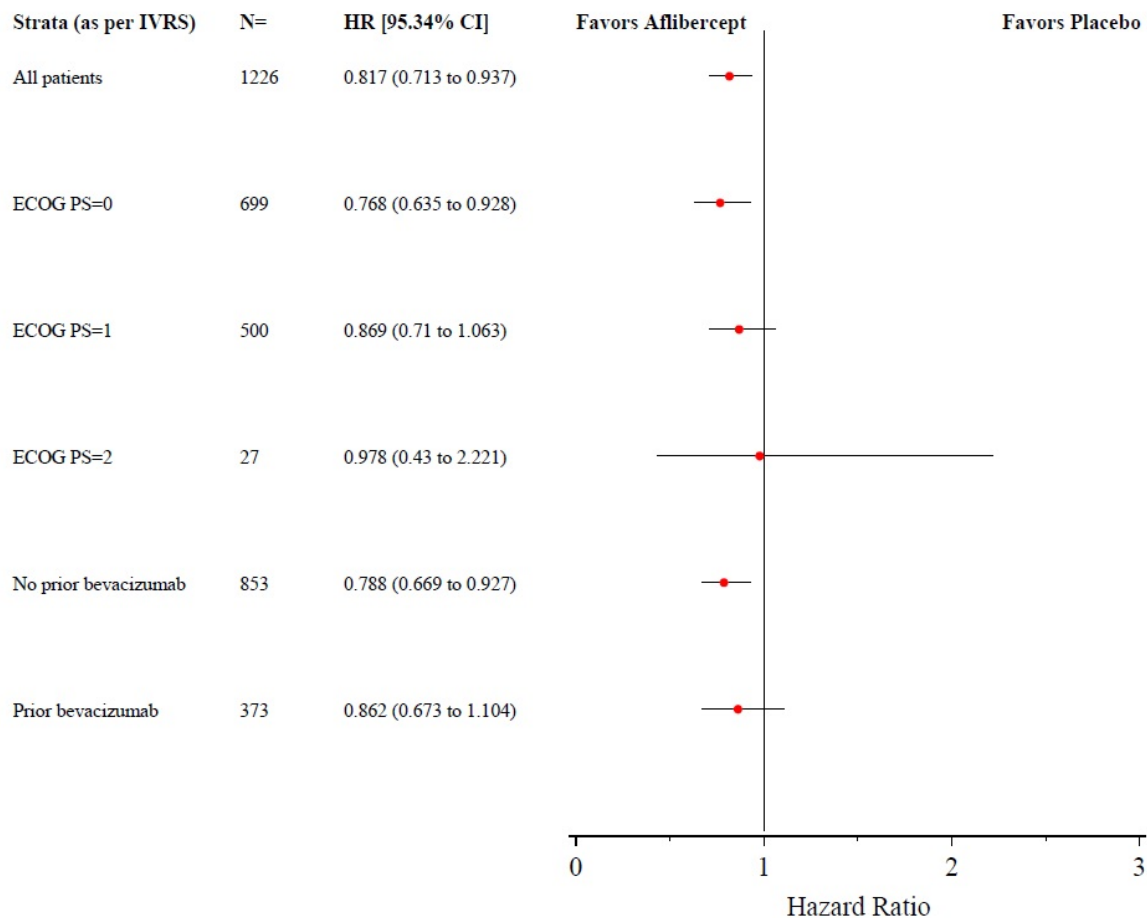
Bezüglich des primären Endpunkts OS bestehen Hinweise auf Effektmodifikation durch die Faktoren „vorbestehender Hypertonus“ (p=0,1309), „ausschließlich Lebermetastasen“ (p=0,0899) und „Lokalisation des Primärtumors“ (p=0,1421).

In der Subgruppe der Patienten ohne vorbestehenden Hypertonus beträgt das Hazard Ratio 0,883 (95%-KI: [0,742;1,051]), in der Subgruppe der Patienten mit vorbestehendem Hypertonus 0,714 (95%-KI: [0,574;0,879]). Zwischen den Patientengruppen ohne und mit vorbestehendem Hypertonus ist der Unterschied der Effektmaße also nur quantitativ (vgl. auch Abbildung 10). Im Rahmen von post-hoc durchgeführten Analysen ergaben sich keine Faktoren, die diese numerischen Unterschiede im Behandlungseffekt auf das OS hätten erklären können. Die Effektschätzer bleiben auch nach Adjustierung für mögliche *Confounder* gleich. Es ist daher davon auszugehen, dass es sich hierbei um eine zufällige Variation handelt.

In der Subgruppe der Patienten, bei denen sich Metastasen ausschließlich in der Leber befinden, beträgt das Hazard Ratio 0,649 (95%-KI: [0,494;0,851]), in der Subgruppe der Patienten, bei denen sich Metastasen auch in anderen Organen befinden, 0,868 (95%-KI: [0,744;1,012]). Die Ergebnisse dieser Subgruppenanalyse sind jedoch mit Vorsicht zu interpretieren, da zwischen den betrachteten Subgruppen ein deutliches Ungleichgewicht hinsichtlich der Patientenzahlen besteht (927 gegenüber 299). Zudem liegen beide zum Subgruppenmerkmal „ausschließlich Lebermetastasen“ zugehörigen Punktschätzer im Forest-Plot eindeutig auf der Aflibercept bevorzugenden Seite (Abbildung 10).

In der Subgruppe der Patienten, bei denen der Primärtumor im Kolon lokalisiert ist, ist das Hazard Ratio signifikant und beträgt 0,739 (95%-KI: [0,609;0,896]) während es in der Subgruppe der Patienten, bei denen der Primärtumor im Rektosigmoid oder an anderen Stellen lokalisiert ist, 1,039 (95%-KI: [0,775;1,393]) und in der Subgruppe mit Patienten, bei denen der Primärtumor im Rektum lokalisiert ist, 0,806 (95%-KI: [0,632;1,027]) beträgt. Eine mögliche Ursache für den beobachteten Gesamteffekt könnte hier in der dafür verantwortlichen Gruppe von Patienten liegen, deren Primärtumorlokalisierung als „Rektosigmoid oder an anderen Stellen“ charakterisiert wurde: abgesehen davon, dass es sich innerhalb der drei betrachteten Gruppen um diejenige mit den geringsten Patientenzahlen (und somit der höchsten statistischen Unsicherheit) handelt, stellt die Originalformulierung „rectosigmoid and other“ eine mangelhaft abgegrenzte, wissenschaftlich unsauber formulierte Definition dar, die Interpretationsspielraum lässt und zwischen den an der Studie beteiligten Ländern auch zu Inkonsistenzen bei der Umsetzung des Begriffs geführt hat. Es ist daher davon auszugehen, dass es sich bei den Patienten mit einer offiziellen Klassifizierung der primären Tumorlokalisierung als „Rektosigmoid oder an anderen Stellen“ um eine deutlich heterogenere Gruppe handelt als bei den beiden weiteren Gruppen „Kolon“ und „Rektum“. Zudem sollte als abschließende Einschränkung zu allen beobachteten Subgruppeneffekten an dieser Stelle nochmals darauf hingewiesen werden, dass Subgruppenanalysen generell das Risiko beinhalten, mit steigender Anzahl statistischer Tests auch die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse zu erhöhen [6].

Eine Interaktion der anderen Variablen (ECOG-PS, Vorbehandlung mit Bevacizumab, Alter, Geschlecht, ethnische Zugehörigkeit, geografische Region und Anzahl Organe, die von Metastasen betroffen sind) mit dem Behandlungstyp kann anhand des Signifikanzniveaus ausgeschlossen werden, d. h. die Subgruppen dieser Merkmale verhalten sich konsistent bezüglich des Behandlungseffekts.



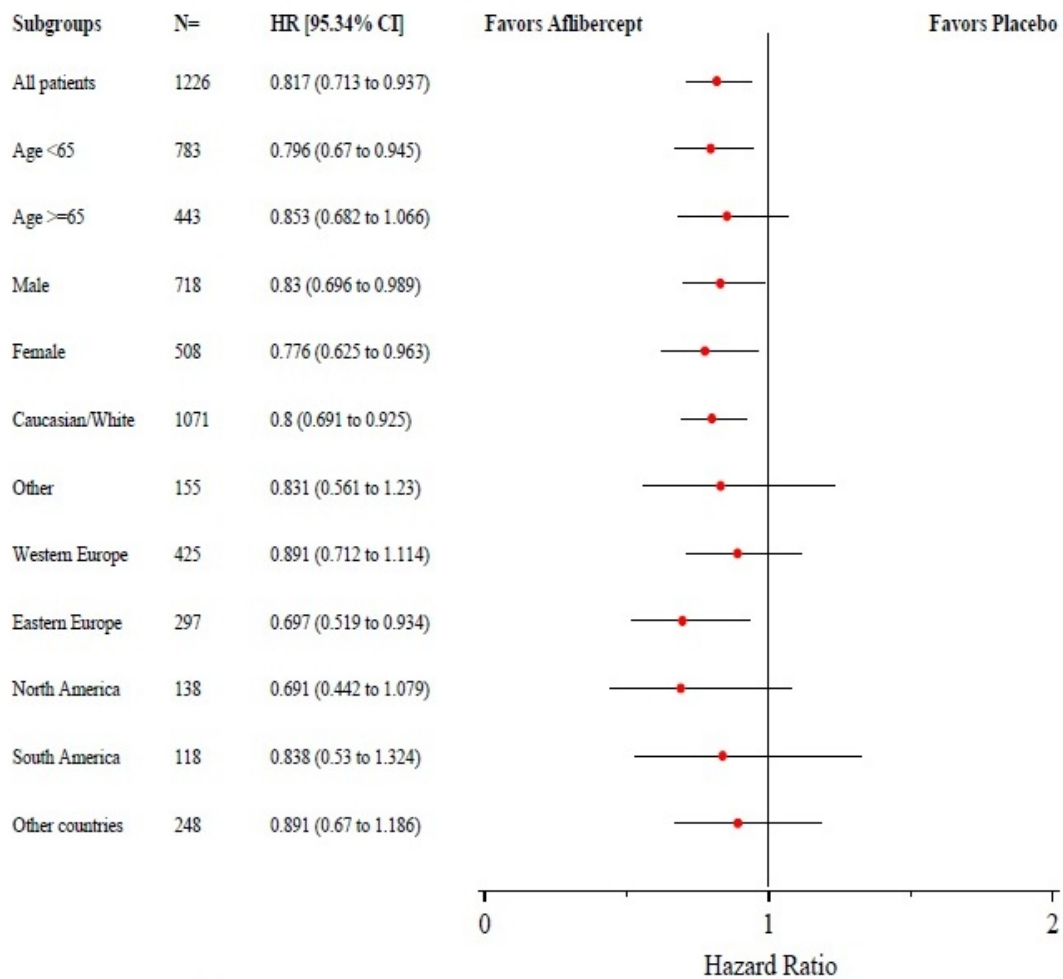
Cutoff date = 7 FEBRUARY 2011

Note: HR : Hazard Ratio, ECOG: Eastern Cooperative Oncology Group, PS: Performance status, IVRS: Interactive Response Voice System

PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_osforestplt_i_g.sas OUT=REPORT/OUTPUT/eff_osforestplt_i_g_i.rtf
(05MAY2011 - 16:03)

Abbildung 8: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für Stratifizierungsfaktoren

Alle Punktschätzer (Hazard Ratios) in Abbildung 8 sind kleiner als Eins und sprechen somit quantitativ auch in den Subgruppen „ECOG-PS“ und „Vorbehandlung mit Bevacizumab“ für ein verbessertes OS im Behandlungsarm FOLFIRI+Aflibercept. In der Subgruppe ECOG-PS=2 befindet sich nur eine geringe Patientenzahl, wodurch das KI verhältnismäßig breit und das zugehörige Hazard Ratio mit besonderer Vorsicht zu betrachten ist. Auch in der Subgruppe mit Bevacizumab-Vorbehandlung sind im Vergleich zur Gruppe ohne Bevacizumab-Vorbehandlung deutlich weniger Patienten; das Konfidenzintervall ist dementsprechend breiter.



Cutoff date = 7 FEBRUARY 2011

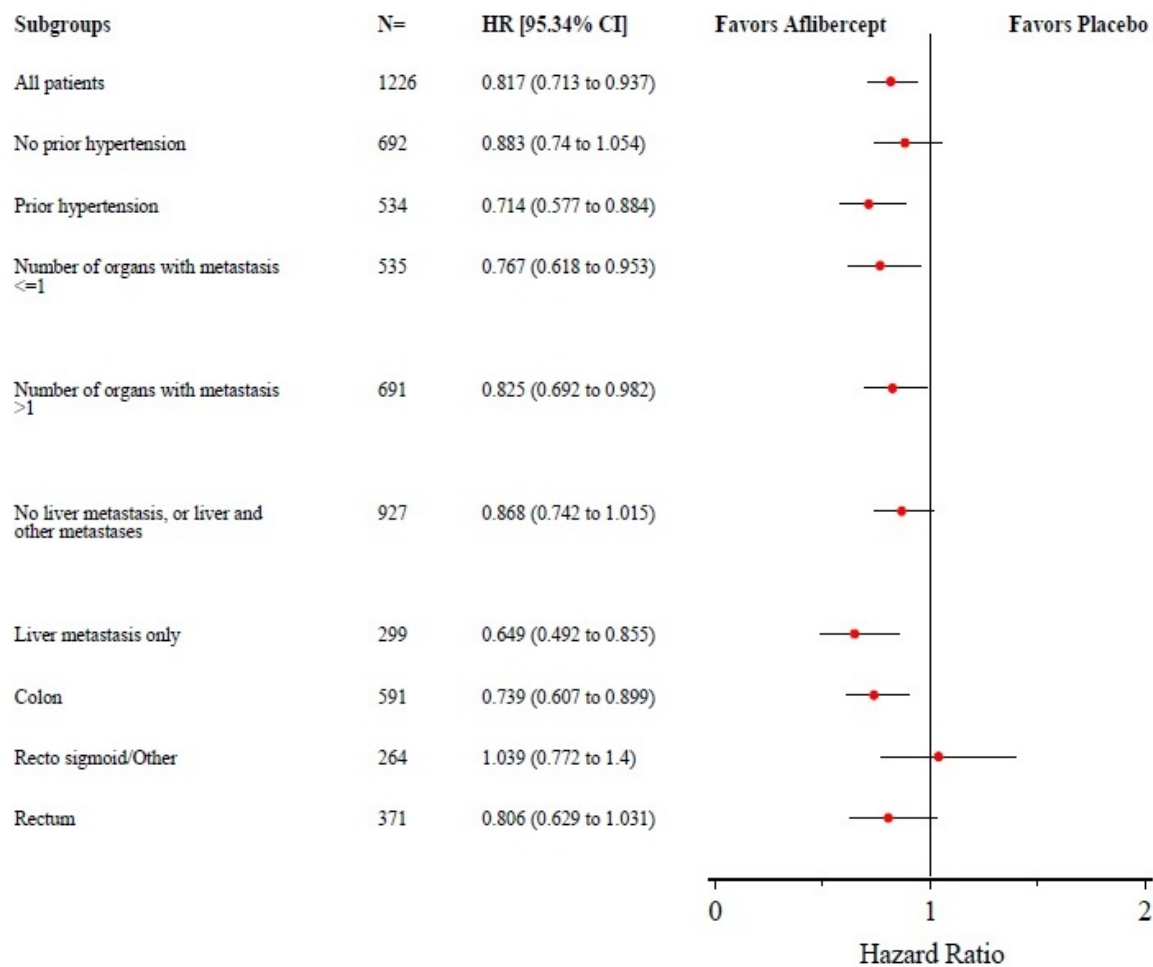
Note: HR : Hazard Ratio

Other countries = Australia, New Zeland, South Africa and Korea

PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_osforestplcov_i_g.sas OUT=REPORT/OUTPUT/eff_osforestplcov_i_g_i.rtf
(05MAY2011 - 16:05)

Abbildung 9: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für demographische Faktoren

Alle Punktschätzer (Hazard Ratios) in Abbildung 9 sind kleiner als Eins und sprechen somit quantitativ auch in den Subgruppen „Alter“, „Geschlecht“, „ethnische Zugehörigkeit“ und „geographische Region“ für ein verbessertes OS im Behandlungsarm FOLFIRI+Afibercept.



Cutoff date = 7 FEBRUARY 2011

Note: HR : Hazard Ratio

PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_osforestpltcov_i_g.sas OUT=REPORT/OUTPUT/eff_osforestpltcov1_i_g_i.rtf (05MAY2011 - 16:05)

Abbildung 10: Forest-Plot der Hazard Ratios für Gesamtüberleben (OS). Subgruppenanalyse der ITT-Population für Baseline-Eigenschaften

Die Punktschätzer (Hazard Ratios) in allen Subgruppen der Merkmale „vorbestehender Hypertonus“, „Anzahl der von Metastasen betroffenen Organe“ und „ausschließlich Lebermetastasen zu Studienbeginn“ sind kleiner als Eins und sprechen somit quantitativ auch für ein verbessertes OS im Behandlungsarm FOLFIRI+Aflibercept (siehe Abbildung 10). Für das Merkmal „Lokalisation des Primärtumors“ bildet lediglich die Subgruppe Rektosigmoid/Andere (deren mutmaßlich größere Heterogenität bereits angesprochen wurde; siehe oben) eine Ausnahme, wobei das Hazard Ratio mit 1,039 (95%-KI: [0,775;1,393]) nur geringfügig über dem Schwellenwert Eins liegt und nicht signifikant ist.

Tabelle 4-33: Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung für Gesamtüberleben (OS) innerhalb der ITT-Population

Studie			
VELOUR			
	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)	Diff. in Monaten
Patienten mit Bevacizumab-Vorbehandlung	186 (30,4)	187 (30,5)	
Mediane Überlebenszeit in Monaten [95%-KI]	12,5 [10,78;15,47]	11,7 [9,96;13,77]	0,8
Hazard Ratio [95%-KI]	0,862 [0,676;1,100]		
Patienten ohne Bevacizumab-Vorbehandlung	426 (69,6)	427 (69,5)	
Mediane Überlebenszeit in Monaten [95%-KI]	13,9 [12,72;15,64]	12,4 [11,17;13,54]	1,5
Hazard Ratio [95%-KI]	0,788 [0,671;0,925]		
p-Wert des Interaktionstests*	0,5668		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Interaktionstest des Cox-Modell mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor

Tabelle 4-34: Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung für Gesamtüberleben (OS) innerhalb der Population unter Ausschluss der adjuvant vorbehandelten Patienten

Studie			
VELOUR			
	FOLFIRI+Aflibercept (N = 552*)	FOLFIRI+Placebo (N = 550*)	Diff. in Monaten
Patienten mit Bevacizumab-Vorbehandlung, ohne adjuvante Therapie	177 (32,1)	179 (32,5)	
Mediane Überlebenszeit in Monaten [95%-KI]	13,8 [11,01;15,87]	11,7 [9,66;13,27]	2,1
Hazard Ratio [95%-KI]	0,812 [0,634;1,042]		
Patienten ohne Bevacizumab-Vorbehandlung, ohne adjuvante Therapie	375 (67,9)	371 (67,5)	
Mediane Überlebenszeit in Monaten [95%-KI]	13,7 [12,71;16,03]	12,4 [11,17;13,54]	1,3
Hazard Ratio [95%-KI]	0,766 [0,645;0,908]		
p-Wert des Interaktionstests**	0,7049		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Population unter Ausschluss der adjuvant behandelten Patienten

** Interaktionstest anhand Cochran Q-Statistik (Wert stammt aus eigener Berechnung im Rahmen der post-hoc Analyse)

Subgruppenanalysen zum Gesamtüberleben und zum PFS wurden auch nach Stratifikationsfaktoren durchgeführt und waren damit präspezifiziert (ECOG-PS und Bevacizumab-Vorbehandlung). Die Ergebnisse, aufgeschlüsselt nach Subgruppen der Patienten mit bzw. ohne Bevacizumab-Vorbehandlung innerhalb der ITT-Population der Studie sind in Tabelle 4-33 zusammengefasst. Es wurde eine zahlenmäßig geringere Behandlungswirkung des FOLFIRI+Aflibercept-Regime in Bezug auf das OS bei Patienten mit Bevacizumab-Vorbehandlung im Vergleich zu Patienten ohne Bevacizumab-Vorbehandlung gezeigt, ohne Beleg für eine Heterogenität in der Behandlungswirkung (nicht signifikanter Interaktionstest; p-Wert=0,5668). Zudem ist die Gruppe der Patienten mit Bevacizumab-Vorbehandlung deutlich kleiner, das Ergebnis demzufolge mit größerer Unsicherheit behaftet.

Zusätzliche post-hoc Analysen für Patienten mit bzw. ohne Bevacizumab-Vorbehandlung schlossen Patienten aus, deren Erkrankung unter oder innerhalb von 6 Monaten nach Abschluss der adjuvanten Therapie weiter fortschritt (etwa 10,1% der gesamten

Studienteilnehmer), und sind in Tabelle 4-34 zusammengefasst. Auch innerhalb der Teilpopulation besteht kein Hinweis auf eine Interaktion zwischen den Gruppen mit bzw. ohne Bevacizumab-Vorbehandlung ($p=0,7049$).

Diese zusätzlichen post-hoc Analysen wurden durchgeführt, da Bevacizumab laut Label und Fachinformation in der sogenannten adjuvanten Therapie keine Zulassung besitzt. Um ein erhöhtes Verzerrungsrisiko auszuschließen, wurden diese adjuvanten Patienten aus der entsprechenden Bewertung ausgeschlossen. Beide Analysen sind in ihrem Ergebnis konsistent.

4.3.1.3.5.1.2 Subgruppenanalysen zum Progressionsfreien Überleben (PFS)

Tabelle 4-35: Interaktionstests der präspezifizierten Subgruppenanalyse für den sekundären Endpunkt Progressionsfreies Überleben (PFS)

Studie		
VELOUR		
Merkmal	Subgruppen	p-Wert der Interaktion*
ECOG-PS	0/1/2	0,6954
Vorbehandlung mit Bevacizumab	ja/nein	0,1958
Alter	<65/≥65	0,9295
Geschlecht	männlich/weiblich	0,7120
Ethnische Zugehörigkeit	Kaukasisch+weiß/andere	0,7701
Geografische Region	Westeuropa/Osteuropa/Nordamerika, Südamerika/andere Kontinente	0,3250
Vorbestehender Hypertonus	ja/nein	0,7876
Anzahl Organe, die von Metastasen betroffen sind	>1/≤1	0,1455
Ausschließlich Lebermetastasen	ja/nein	0,0076
Lokalisation des Primärtumors	Kolon/Rektosigmoid/andere vs. Rektum	0,3539

ECOG-PS: Eastern Cooperative Oncology Group-Performance Status

* Interaktionstest des Cox-Modell mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor

Für den Endpunkt PFS wurden die Interaktionstests analog zum Endpunkt OS durchgeführt und bewertet (siehe Tabelle 4-35).

Hinweis: Wie in Abschnitt 4.3.1.3.5.1.1 dargestellt, wurden im CSR zur VELOUR-Studie irrtümlicherweise die p-Werte der Interaktionstests zu den Subgruppenanalysen ECOG-PS 0/1/2 und Bevacizumab-Vorbehandlung ja/nein sowohl für OS als auch für PFS vertauscht (Tabellen 29 und 38 des CSR) [37].

Tabelle 4-36: Effektmodifikationen (p-Wert<0,2) in den präspezifizierten Subgruppen für Progressionsfreies Überleben (PFS)

Studie						
VELOUR						
Merkmal	FOLFIRI+Aflibercept (N = 612)		FOLFIRI+Placebo (N = 614)		Hazard Ratio [99,99%-KI*] [95%-KI]	
	Patienten, n	Ereignisse, n (%)	Patienten, n	Ereignisse, n (%)		
Ausschließlich Lebermetastasen						
<i>Nein</i>	459	303 (66,0)	468	343 (73,3)	0,839 [0,617;1,143] [0,719;0,980]	
<i>Ja</i>	153	90 (58,8)	146	111 (76,0)	0,547 [0,313;0,956] [0,413;0,725]	
Vorbehandlung mit Bevacizumab						
<i>ja</i>	186	112 (60,2)	187	131 (70,1)	0,661 [0,399;1,095] [0,512;0,852]	
<i>nein</i>	426	281 (66,0)	427	323 (75,6)	0,797 [0,580;1,096] [0,679;0,936]	
Anzahl Organe, die von Metastasen betroffen sind						
<i>>1</i>	354	250 (70,6)	337	259 (76,9)	0,822 [0,581;1,162] [0,690;0,979]	
<i>≤1</i>	258	143 (55,4)	277	195 (70,4)	0,655 [0,426;1,009] [0,527;0,815]	

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Signifikanzlevel gesetzt auf $\alpha=0,0001$

Bezüglich des sekundären Endpunktes PFS zeigt sich eine belegte Interaktion zwischen dem Vorliegen von Metastasen ausschließlich in der Leber und der Art der Behandlung (p=0,0076). In der Subgruppe der Patienten, bei denen sich Metastasen ausschließlich in der Leber befinden, beträgt das Hazard Ratio 0,547 (95%-KI: [0,413;0,725]) während es in der Subgruppe der Patienten, bei denen sich Metastasen auch in anderen Organen befinden, 0,839 (95%-KI: [0,719;0,980]) beträgt. Die Behandlung mit FOLFIRI+Aflibercept scheint demnach bei Patienten, bei denen sich Metastasen ausschließlich in der Leber befinden eine größere Auswirkung auf das PFS im Vergleich zur Behandlung mit FOLFIRI+Placebo zu haben, jedoch zeigte sich für beide Subgruppen ein signifikantes Ergebnis zu Gunsten von FOLFIRI+Aflibercept. Grundsätzlich spielt auch bei dieser Beobachtung vermutlich das

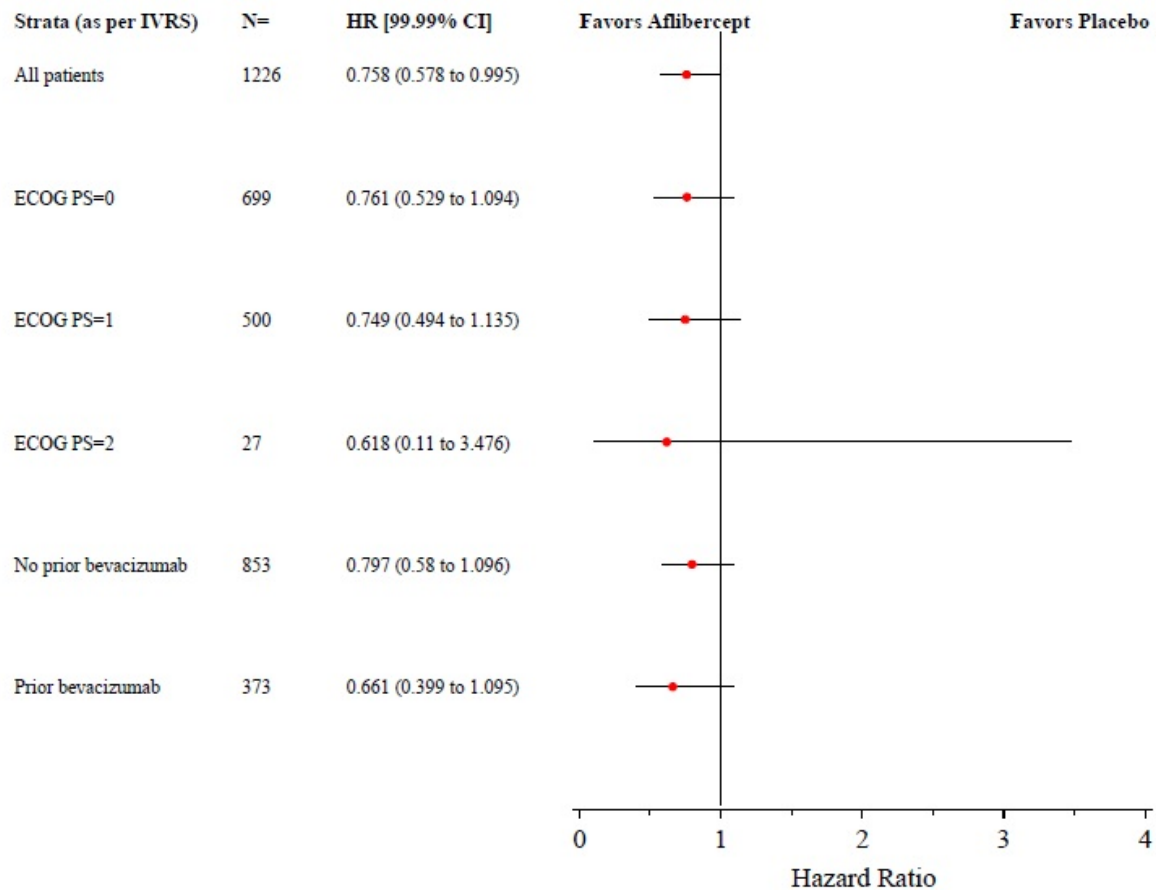
Ungleichgewicht der Patientenzahlen (1:3) innerhalb der beiden miteinander verglichenen Gruppen eine Rolle.

Für die Subgruppenanalyse nach Vorbehandlung mit Bevacizumab ergibt sich ein p-Wert des Interaktionstests von 0,1958, welcher somit nur knapp unterhalb der gewählten Grenze des Signifikanzniveaus für einen Anhaltspunkt auf Subgruppenheterogenität liegt. Diese wurde, wie bereits diskutiert, nicht auf dem international für solche Analysen üblichen, sondern dem durch das IQWiG geforderten strengeren Niveau von 0,2 angesetzt. Die Hazard Ratios in den Subgruppen „Vorbehandlung“ und „keine Vorbehandlung“ sprechen mit 0,661 (95%-KI: [0,512;0,852]) und 0,797 (95%-KI: [0,679;0,936]) dennoch in beiden Gruppen statistisch signifikant zugunsten von Aflibercept. Auch hier besteht ein erhebliches Ungleichgewicht zwischen den Patientenzahlen der beiden Subgruppen (ca. 1:2,5).

Des Weiteren existiert ein Hinweis auf eine Interaktion zwischen der Anzahl der Organe, die von Metastasen betroffen sind und der Behandlungsart ($p=0,1455$). In der Subgruppe der Patienten, bei denen höchstens ein Organ von Metastasen betroffen ist, beträgt das Hazard Ratio 0,655 (95%-KI: [0,527;0,815]), und in der Patientengruppe, bei denen mehr als ein Organ von Metastasen betroffen ist 0,822 (95%-KI: [0,690;0,979]). Da auch hier beide Punktschätzer auf ein signifikant besseres PFS im FOLFIRI+Aflibercept-Arm hindeuten, ist die mutmaßliche Interaktion lediglich quantitativ.

Insgesamt ist zu beachten, dass sich die Ergebnisse jeweils für alle Subgruppen statistisch signifikant positiv zu Gunsten von Aflibercept darstellen (vgl. auch die nachfolgenden Darstellungen der jeweiligen Punktschätzer in Forest Plots, Abbildung 11 - Abbildung 13). Als abschließende Einschränkung zu allen beobachteten Subgruppeneffekten wird an dieser Stelle nochmals darauf hingewiesen werden, dass Subgruppenanalysen generell das Risiko beinhalten, mit steigender Anzahl statistischer Tests auch die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse zu erhöhen [6].

Eine Interaktion der anderen Variablen (ECOG-PS, Alter, Geschlecht, ethnische Zugehörigkeit, geografische Region, vorbestehender Hypertonus und Lokalisation des Primärtumors) mit dem Behandlungstyp kann ausgeschlossen werden, d. h. die Subgruppen dieser Merkmale verhalten sich konsistent bezüglich des Effekts, der von der Art der Behandlung ausgeht.

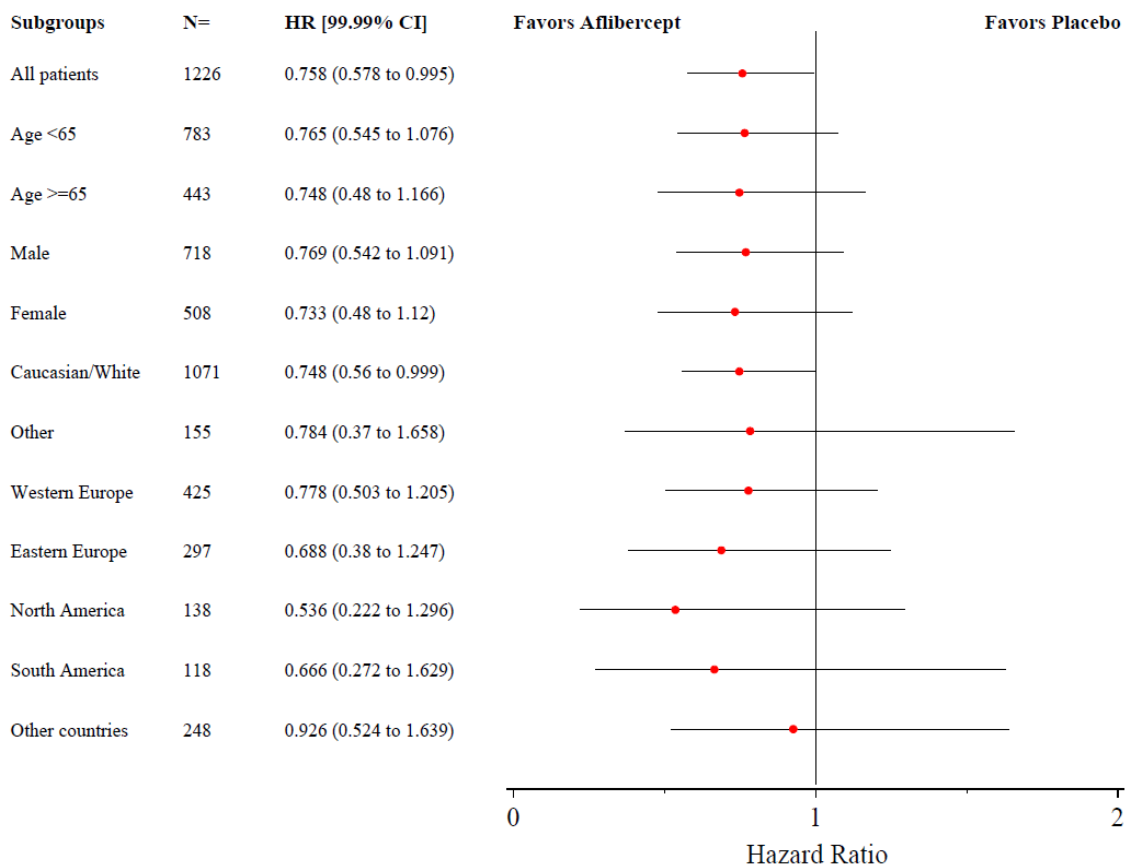


Cutoff date = 06 MAY 2010

Note: HR : Hazard Ratio, ECOG: Eastern Cooperative Oncology Group, IVRS: Interactive Response Voice System, PS: Performance status
 PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_pfsforestplt_i_g.sas OUT=REPORT/OUTPUT/eff_pfsforestplt_i_g_i.rtf
 (05MAY2011 - 16:10)

Abbildung 11: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS).
 Subgruppenanalyse der ITT-Population für die Stratifizierungsfaktoren

Alle Punktschätzer (Hazard Ratios) in Abbildung 11 sind kleiner als Eins und sprechen somit quantitativ auch in den Subgruppen für ein verbessertes PFS im Behandlungsarm FOLFIRI+Aflibercept. In der Subgruppe ECOG-PS=2 befindet sich nur eine geringe Patientenzahl (N=27), wodurch das Konfidenzintervall verhältnismäßig breit und das zugehörige Hazard Ratio mit Vorsicht zu betrachten ist. Die Ergebnisse für die Subgruppen mit ECOG-PS=0 und 1 sind bei Betrachtung des 95%-KI signifikant. Ebenso verhält es sich für die Subgruppen mit und ohne Bevacizumab-Vorbehandlung [33].



Cutoff date = 06 MAY 2010

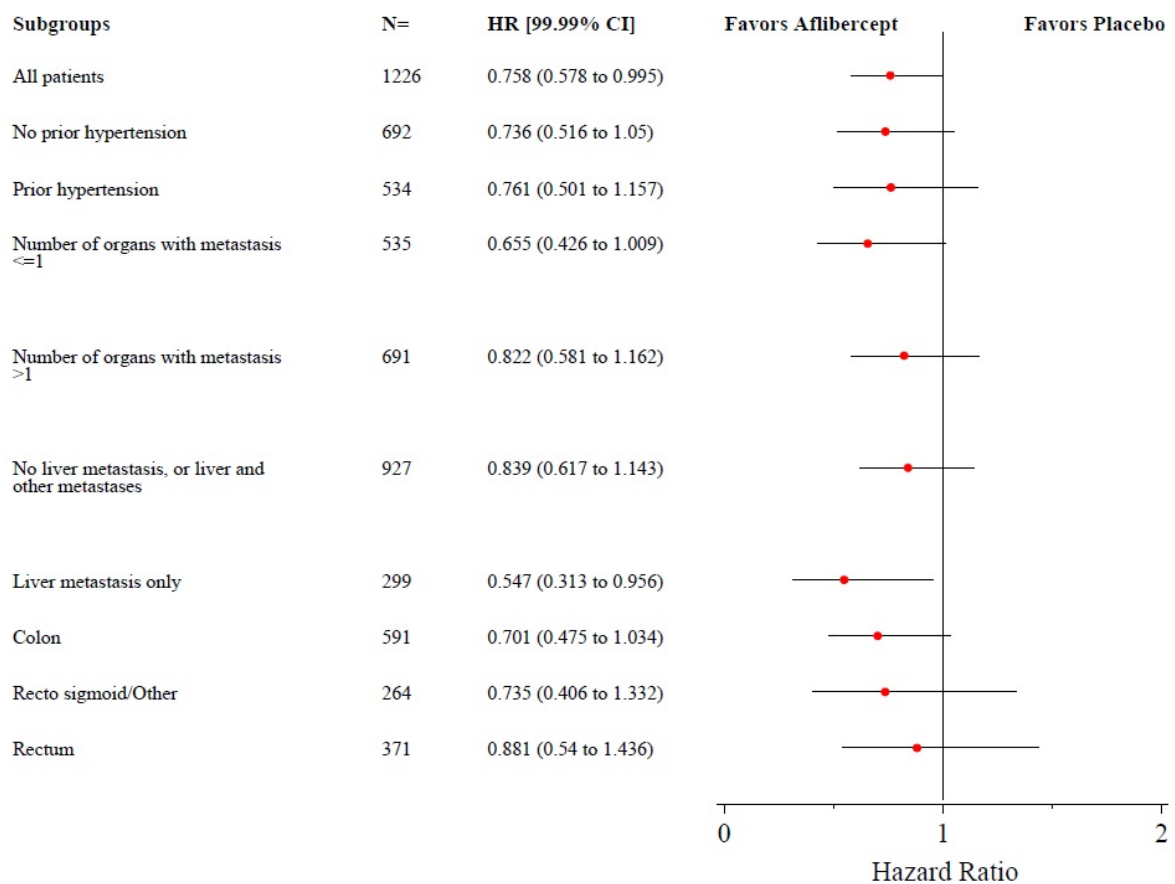
Note: HR : Hazard Ratio

Other countries = Australia, New Zeland, South Africa and Korea

PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_pfsforestplcov_i_g.sas OUT=REPORT/OUTPUT/eff_pfsforestplcov_i_g_i.rtf (05MAY2011 - 16:14)

Abbildung 12: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS). Subgruppenanalyse der ITT-Population für demografische Faktoren

Alle Punktschätzer (Hazard Ratios) in Abbildung 12 sind kleiner als Eins und sprechen somit quantitativ auch in den Subgruppen für ein verbessertes PFS im Behandlungsarm FOLFIRI+Aflibercept.



Cutoff date = 06 MAY 2010

Note: HR : Hazard Ratio

PGM=PRODOPS/AVE0005/EFC10262/CSR/REPORT/PGM/eff_pfsforestpltcov_i_g.sas OUT=REPORT/OUTPUT/eff_pfsforestpltcov1_i_g_i.rtf
(05MAY2011 - 16:14)

Abbildung 13: Forest-Plot der Hazard Ratios für Progressionsfreies Überleben (PFS).
Subgruppenanalyse der ITT-Population für Baseline-Eigenschaften

Alle Punktschätzer (Hazard Ratios) in Abbildung 13 sind kleiner als Eins und sprechen somit quantitativ auch in den Subgruppen für ein verbessertes PFS im Behandlungsarm FOLFIRI+Aflibercept.

Tabelle 4-37: Subgruppenanalyse der Bevacizumab-Vorbehandlung für Progressionsfreies Überleben (PFS) innerhalb der ITT-Population

Studie			
VELOUR			
	FOLFIRI+Aflibercept (N = 612)	FOLFIRI+Placebo (N = 614)	Diff. in Monaten
Patienten mit Bevacizumab-Vorbehandlung	186 (30,4)	187 (30,5)	
Medianes progressionsfreies Überleben in Monaten [95%-KI]	6,7 [5,75;8,21]	3,9 [3,02;4,30]	2,8
Hazard Ratio [95%-KI]	0,661 [0,512;0,852]		
Patienten ohne Bevacizumab-Vorbehandlung	426 (69,6)	427 (69,5)	
Medianes progressionsfreies Überleben in Monaten [95%-KI]	6,9 [6,37;7,20]	5,4 [4,53;5,68]	1,5
Hazard Ratio [95%-KI]	0,797 [0,679;0,936]		
p-Wert des Interaktionstests*	0,1958		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Interaktionstest des Cox-Modell mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor

Tabelle 4-38: Subgruppenanalyse der Bevacizumab-Vorbehandlung für Progressionsfreies Überleben (PFS) innerhalb der Population unter Ausschluss der adjuvant behandelten Patienten

Studie			
VELOUR			
	FOLFIRI+Aflibercept (N = 552*)	FOLFIRI+Placebo (N = 550*)	Diff. in Monaten
Patienten mit Bevacizumab-Vorbehandlung, ohne adjuvante Therapie	177 (32,1)	179 (32,5)	
Medianes progressionsfreies Überleben in Monaten [95%-KI]	6,7 [5,72;8,21]	3,9 [3,02;4,30]	2,8
Hazard Ratio [95%-KI]	0,645 [0,498;0,835]		
Patienten ohne Bevacizumab-Vorbehandlung, ohne adjuvante Therapie	375 (67,9)	371 (67,5)	
Medianes progressionsfreies Überleben in Monaten [95%-KI]	6,9 [6,24;7,20]	5,3 [4,50;5,55]	1,6
Hazard Ratio [95%-KI]	0,777 [0,655;0,921]		
p-Wert des Interaktionstests**	0,2377		

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; KI: Konfidenzintervall

* Population unter Ausschluss der adjuvant behandelten Patienten

** Interaktionstest anhand Cochran Q-Statistik (Wert stammt aus eigener Berechnung im Rahmen der post-hoc Analyse)

Analog zum Vorgehen beim primären Endpunkt OS wurde für den Endpunkt PFS eine Subgruppenanalyse zum Merkmal Bevacizumab-Vorbehandlung durchgeführt (im Studiendesign präspezifiziert/ Stratifikationsfaktor) und aus bereits erläuterten Gründen um eine post-hoc Analyse unter Ausschluss der adjuvant mit Oxaliplatin vorbehandelten Patienten ergänzt. Die Ergebnisse sind in den obigen beiden Tabellen dargestellt. Innerhalb der ITT-Population ergibt sich mit einem p-Wert von 0,1958, wie bereits diskutiert, äußerst knapp ein Anhaltspunkt auf Interaktion; der entsprechende Wert innerhalb der Teilpopulation unter Ausschluss der adjuvant mit Oxaliplatin vorbehandelten Patienten liegt mit 0,2377 hingegen jenseits des angewendeten Signifikanzniveaus von 0,2. Sowohl für die ITT-Population als auch für die Teilpopulation unter Ausschluss der adjuvant behandelten Patienten sprechen die Hazard Ratios für das PFS konsistent für einen Vorteil von FOLFIRI+Aflibercept gegenüber FOLFIRI+Placebo. Hierbei scheinen jeweils die mit Bevacizumab vorbehandelten Patienten etwas stärker von der Behandlung FOLFIRI+Aflibercept zu profitieren. Diese Beobachtung ist jedoch genau entgegengesetzt zu der entsprechenden Analyse zum Gesamtüberleben (vgl. Abschnitt 4.3.1.3.5.1.1, in der jeweils die Patientengruppe ohne Bevacizumab-Vorbehandlung ein stärkeres Profitieren von der Behandlung mit FOLFIRI+Aflibercept andeutet. Dies könnte insgesamt als zusätzliches

Indiz dafür aufgefasst werden, dass in der Realität kein echter Unterschied zwischen den beiden Subgruppen mit und ohne Bevacizumab-Vorbehandlung vorliegt.

4.3.1.3.5.2 Post-hoc durchgeführte Analysen

Für die Bewertung der Sicherheit wurden Subgruppenanalysen der VELOUR-Studie post-hoc für die durch die Verfahrensordnung vorgegebenen Merkmale Alter, Geschlecht und Schwere der Erkrankung (in dieser Studie repräsentiert durch den ECOG-Performance Status) durchgeführt (Tabelle 4-39). Außerdem wurden die Subgruppen für das Merkmal „Vorbehandlung mit Bevacizumab“ analysiert, da in der VELOUR-Studie neben dem ECOG-PS eine stratifizierte Randomisierung für diese Variable erfolgt ist. Für jedes Merkmal wurde ein Interaktionstest (Cochrane Q-Statistik) für Subgruppen durchgeführt. Dabei wurde jeweils geprüft, ob sich die Odds Ratios der Subgruppen eines Merkmals signifikant unterscheiden.

4.3.1.3.5.2.1 Subgruppenanalysen zur Sicherheit (Unerwünschte Ereignisse) innerhalb der Sicherheitspopulation

Tabelle 4-39: Interaktionstest der post-hoc Subgruppenanalyse für die Bewertung der Sicherheit (Unerwünschte Ereignisse) anhand der Sicherheitspopulation

Studie						
VELOUR						
Merkmal	Subgruppen	p-Wert der Interaktion*				
		UE	UE Grad 3-4	SUE	UE→ Tod	UE→ Abbruch Behandlg.
ECOG-PS	0/1/2	0,8122 ⁽¹⁾	0,9460	0,7650	0,7424	0,6514
Vorbehandlung mit Bevacizumab	ja/nein	0,1529	0,9869	0,5710	0,6607	0,6509
Alter	<65/≥65	0,4122	0,2574	0,0021	0,0222	0,2596
Geschlecht	männlich /weiblich	0,2501	0,5635	0,5269	0,5510	0,4398

ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; UE: Unerwünschtes Ereignis; SUE: Schwerwiegendes unerwünschtes Ereignis

* Chi-Quadrat-Test (Cochrane Q-Statistik) zur Untersuchung der Subgruppenunterschiede (Werte stammen aus eigener Berechnung)

⁽¹⁾ Test für Subgruppenunterschiede basiert nicht auf allen Subgruppen, bzw. ist nicht anwendbar (da Odds Ratio nicht in allen Subgruppen definiert)

Die Effektmodifikationen, bei denen der p-Wert des Interaktionstests unterhalb von 0,2 liegt, werden in Tabelle 4-40 dargestellt.

Tabelle 4-40: Effektmodifikationen (p-Wert<0,2) in den post-hoc analysierten Subgruppen für den Endpunkt Sicherheit (Unerwünschte Ereignisse)

Studie							
VELOUR							
Merkmal Subgruppe	FOLFIRI+ Aflibercept (N = 611)		FOLFIRI+Placebo (N = 605)		Relatives Risiko [95%-KI]	Odds Ratio [95%-KI]	
	Patienten, n	UE, n (%)	Patienten, n	UE, n (%)			
Patienten, bei denen ein unerwünschtes Ereignis auftrat							
Vorbehandlung	<i>nein</i>	424	419 (98,8)	421	414 (98,3)	1,00 [0,99;1,02]	1,42 [0,45;4,50]
mit Bevacizumab	<i>ja</i>	187	187 (100)	184	178 (96,7)	1,03 [1,00;1,06]	13,66 [0,76;244,17]
Patienten, bei denen ein schwerwiegendes unerwünschtes Ereignis auftrat							
Alter	<65	406	173 (42,6)	372	125 (33,6)	1,27 [1,06;1,52]	1,47 [1,10;1,96]
	≥65	205	121 (59,0)	233	73 (31,3)	1,88 [1,51;2,35]	3,16 [2,13;4,68]
Patienten, bei denen das unerwünschte Ereignis zum Tod führte							
Alter	<65	406	16 (3,9)	372	19 (5,1)	0,77 [0,40;1,48]	0,76 [0,39;1,51]
	≥65	205	21 (10,2)	233	10 (4,3)	2,39 [1,15;4,95]	2,55 [1,17;5,54]

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; UE: Unerwünschtes Ereignis; KI: Konfidenzintervall

Bezüglich der Bewertung der UEs gibt es einen Hinweis auf Interaktion zwischen der Vorbehandlung mit Bevacizumab und dem Behandlungsarm (p=0,1529). Unter allen Patienten, bei denen UEs auftraten, beträgt das Odds Ratio in der Subgruppe der mit Bevacizumab vorbehandelten Patienten 13,66 (95%-KI: [0,76;244,17]) und in der Subgruppe der Patienten, in der keine Behandlung mit Bevacizumab vorausging nur 1,42 (95%-KI: [0,45;4,50]) (Tabelle 4-40). Das große Ausmaß des Unterschieds zwischen den beiden Effektschätzern lässt sich dadurch erklären, dass in den Subgruppen jeweils nur eine sehr geringe Anzahl von Patienten (insgesamt nur 1,48%) keine UEs angab. Das Odds Ratio ist bezüglich dieser Datensituation empfindlich und dadurch wenig aussagekräftig. In diesem Fall stellt das RR ein für die Interpretation besser geeignetes Effektmaß dar. Es ergibt sich in der Patientengruppe mit Bevacizumab-Vorbehandlung ein RR von 1,00 (95%-KI: [0,99;1,02]) und in der Patientengruppe ohne Bevacizumab-Vorbehandlung ein RR von 1,03 (95%-KI: [1,00;1,06]). Auch wenn die Werte der RR aus beiden Subgruppen augenscheinlich näher zusammenliegen als die jeweiligen Werte der OR, ist zu beachten, dass die Interaktion zwischen Vorbehandlung mit Bevacizumab und dem Behandlungsarm auch unter Beobachtung des RR signifikant ist (p=0,09). Dieser Umstand ergibt sich aus der Eigenschaft der beiden Effektmaße RR und Odds Ratio, die nicht linear sind.

In der Bewertung der SUEs und der Bewertung der UEs, die zum Tod des Patienten führten, liegen signifikante Interaktionen zwischen dem Alter und der Behandlungsart vor ($p=0,0021$ und $p=0,0222$). Unter allen Patienten, bei denen SUEs auftraten, beträgt das Odds Ratio in der Altersgruppe unter 65 Jahren 1,47 (95%-KI: [1,10;1,96]) und in der Altersgruppe über 65 Jahren 3,16 (95%-KI: [2,13;4,68]) (Tabelle 4-40). Der Interaktionseffekt ist demnach nur quantitativ. In der Bewertung der UEs, die zum Tod des Patienten führten, beträgt das Odds Ratio in der Altersgruppe unter 65 Jahre 0,76 (95%-KI: [0,39;1,51]) und in der Altersgruppe über 65 Jahre 2,55 (95%-KI: [1,17;5,54]) (Tabelle 4-40), was auf einen qualitativen Interaktionseffekt hindeutet. Diese Resultate sprechen für einen Effekt des Alters und der Behandlungsart hinsichtlich des Auftretens von SUEs sowie von UEs, die zum Tod führten. Die Durchführung einer Beobachtungsstudie bei Patienten >65 Jahren wurde folglich auch als geplante Maßnahme in den EU-Risk-Management-Plan (EU-RMP) von Afibercept mit aufgenommen (siehe auch Modul 3 A, Abschnitt 3.4).

Aufgrund der Aufteilung des Endpunktes Sicherheit in verschiedene Subkategorien der UEs und die damit verbundene steigende Anzahl statistischer Tests erhöht sich, wie schon in 4.2.5.5 erwähnt, die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse.

In der Bewertung der UEs vom Grad 3-4 und der UEs, die zum Abbruch der Behandlung führten, gibt es keine signifikanten Interaktionen zwischen Subgruppen und Behandlungsart.

4.3.1.3.6 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.

Im Folgenden werden die Ergebnisse der VELOUR-Studie für die Gesamt- und Subgruppenanalysen des primären Endpunkts OS, des sekundären Endpunkts PFS sowie der ORR und der Bewertung der Sicherheit zusammengefasst.

4.3.1.3.6.1 Zusammenfassung der Ergebnisse zum Gesamtüberleben (OS)

Die Analyse des OS zeigte eine signifikante Verbesserung des Überlebens für Patienten mit einer FOLFIRI+Afibercept-Behandlung gegenüber einer FOLFIRI+Placebo-Behandlung ($p=0,0032$ anhand des Log-Rank-Tests). Die mediane Überlebenszeit für die Patienten im

FOLFIRI+Aflibercept-Arm war 13,50 (95%-KI: [12,52;14,95]) Monate, das 25%-Quantil betrug 7,62 (95%-KI: [6,60;8,48]) Monate und das 75%-Quantil des OS 25,59 (95%-KI: [22,01;31,70]) Monate. Im Gegensatz dazu betrug die mediane Überlebenszeit im FOLFIRI+Placebo-Arm 12,06 (95%-KI: [11,07;13,08]) Monate, das 25%-Quantil betrug 6,83 (95%-KI: [6,24;7,59]) Monate und das 75%-Quantil des OS 21,03 (95%-KI: [18,92;22,80]) Monate. Das mediane Überleben verlängerte sich mit FOLFIRI+Aflibercept signifikant um 1,44 Monate. Das Hazard Ratio für die Mortalität betrug 0,817 (95%-KI: [0,714;0,934]) zugunsten der FOLFIRI+Aflibercept-Gruppe; das entspricht einer 18,3%-igen Reduzierung des Sterberisikos für die Patienten, die mit FOLFIRI+Aflibercept behandelt wurden gegenüber den Patienten, die FOLFIRI+Placebo erhielten. Die zusätzliche Analyse des mittleren Überlebens bestätigt den Überlebensvorteil durch die FOLFIRI+Aflibercept-Therapie. Das mittlere Überleben ermöglicht eine aussagekräftigere Bewertung des individuellen Überlebens der Patienten und berücksichtigt insbesondere auch die Patienten, die eine deutliche Lebensverlängerung erfahren. Zwischen den Behandlungsarmen konnte so (je nach zugrunde liegendem Berechnungsmodell) eine Differenz von bis zu 4,7 Monate (95%-KI: [2,1;6,1]) zugunsten von FOLFIRI+Aflibercept gezeigt werden (FOLFIRI+Aflibercept 22,8 Monate vs. FOLFIRI+Placebo 18,1 Monate).

Die präspezifizierten Subgruppenanalysen für das OS zeigen bis auf eine Ausnahme einen quantitativ übereinstimmenden Behandlungseffekt zugunsten FOLFIRI+Aflibercept (Hazard Ratio <1).

Für alle Subgruppenmerkmale wurde ein Interaktionstest des Cox-Modells mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor durchgeführt. Bezüglich des primären Endpunkts OS bestehen Hinweise auf Effektmodifikationen durch die Faktoren „vorbestehender Hypertonus“ (p=0,1309), „ausschließlich Lebermetastasen“ (p=0,0899) und „Lokalisation des Primärtumors“ (p=0,1421). In der Subgruppe der Patienten ohne vorbestehenden Hypertonus beträgt das Hazard Ratio 0,883 (95%-KI: [0,742;1,051]), in der Subgruppe der Patienten mit vorbestehender Hypertonus 0,714 (95%-KI: [0,574;0,879]). Zwischen den Patientengruppen ohne und mit vorbestehendem Hypertonus ist der Unterschied der Effektmaße also nur quantitativ. In der Subgruppe der Patienten, bei denen sich Metastasen ausschließlich in der Leber befinden, beträgt das Hazard Ratio 0,649 (95%-KI: [0,494;0,851]), in der Subgruppe der Patienten, bei denen sich Metastasen auch in anderen Organen befinden, 0,868 (95%-KI: [0,744;1,012]). Das Vorliegen einer tatsächlichen Interaktion ist jedoch fraglich, da zwischen den betrachteten Subgruppen bereits ein deutliches Ungleichgewicht hinsichtlich der Patientenzahlen besteht (927 gegenüber 299). Zudem liegen auch hier beide zum Subgruppenmerkmal zugehörigen Punktschätzer im Forest-Plot eindeutig auf der Aflibercept bevorzugen Seite. In der Subgruppe der Patienten, bei denen der Primärtumor im Kolon lokalisiert ist, ist das Hazard Ratio signifikant und beträgt 0,739 (95%-KI: [0,609;0,896]) während es in der Subgruppe der Patienten, bei denen der Primärtumor im Rektosigmoid oder an anderen Stellen lokalisiert ist, 1,039 (95%-KI: [0,775;1,393]) und in der Subgruppe mit Patienten, bei denen der Primärtumor im Rektum lokalisiert ist, 0,806 (95%-KI: [0,632;1,027]) beträgt. Für die apparente qualitative Interaktion ist hier mutmaßlich die definitionsbedingt höhere Heterogenität der Subgruppe der Patienten

mit Lokalisation des Primärtumors im Rektosigmoid oder an anderen Stellen mitverantwortlich.

Zudem sollte als abschließende Einschränkung zu allen beobachteten Subgruppeneffekten an dieser Stelle nochmals darauf hingewiesen werden, dass Subgruppenanalysen generell das Risiko beinhalten, mit steigender Anzahl statistischer Tests auch die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse zu erhöhen [6].

Eine Interaktion der anderen subgruppenbildenden Merkmale mit dem Behandlungstyp kann zum zugrunde gelegten Signifikanzniveau ($p \leq 0,2$) ausgeschlossen werden. Bezüglich der Subgruppen mit und ohne Vorbehandlung mit Bevacizumab gilt dies sowohl für den Interaktionstest innerhalb der ITT-Population als auch für den unter Ausschluss der adjuvant mit Oxaliplatin vorbehandelten Patienten durchgeführten.

4.3.1.3.6.2 Zusammenfassung der Ergebnisse zum Progressionsfreien Überleben (PFS)

Das mediane PFS betrug in der FOLFIRI+Aflibercept-Gruppe 6,90 Monate (95%-KI: [6,51;7,20]), das 25%-Quantil betrug 3,84 (95%-KI: [3,22;4,30]) und das 75%-Quantil 10,05 (95%-KI: [9,40;10,74]) Monate. In der FOLFIRI+Placebo-Gruppe hingegen lag das mediane PFS nur bei 4,67 Monaten (95%-KI: [4,21;5,36]), das 25%-Quantil betrug 2,60 (95%-KI: [1,97;2,76]) und das 75%-Quantil 9,10 (95%-KI: [7,89;9,89]) Monate. Die PFS-Kurven der beiden Behandlungsgruppen wurden mithilfe eines Log-Rank-Tests verglichen, welcher einen signifikanten Unterschied ($p < 0,0001$) zwischen den Gruppen zeigte. Das Hazard Ratio des zugrundeliegenden Cox Proportional-Hazard-Modells betrug 0,758 (95%-KI: [0,661;0,867]) und deutet somit auf eine signifikante Verbesserung des PFS in der FOLFIRI+Aflibercept-Gruppe.

Die präspezifizierten Subgruppenanalysen für PFS zeigen alle ausnahmslos einen quantitativ übereinstimmenden Behandlungseffekt zugunsten FOLFIRI+Aflibercept (Hazard Ratio < 1). Für alle Subgruppenmerkmale wurde ein Interaktionstest des Cox-Modells mit Faktor, Behandlungseffekt und Interaktionseffekt von Behandlung und Faktor durchgeführt. Bezüglich des sekundären Endpunktes PFS liegt eine Interaktion zwischen dem Vorliegen von Metastasen ausschließlich in der Leber und der Art der Behandlung vor ($p = 0,0076$). In der Subgruppe der Patienten, bei denen sich Metastasen ausschließlich in der Leber befinden, ist das Hazard Ratio signifikant und beträgt 0,547 (95%-KI: [0,413;0,725]), während es in der Subgruppe der Patienten, bei denen sich Metastasen auch in anderen Organen befinden, 0,839 (95%-KI: [0,719;0,980]) beträgt und damit ebenfalls signifikant ist. Es handelt sich mit ziemlicher Sicherheit um den gleichen Effekt, der bei diesen Subgruppen auch hinsichtlich des OS beobachtet wurde, verbunden mit den gleichen genannten Zweifeln, ob es sich um eine tatsächliche Interaktion handelt (erhebliches Ungleichgewicht der Patientenzahlen in beiden Gruppen).

Auch beim Subgruppenmerkmal „Vorbehandlung mit Bevacizumab“ besteht ein erhebliches Ungleichgewicht zwischen den Patientenzahlen in den beiden Subgruppen, zudem liegt der

p-Wert mit 0,1958 auch nur gerade unterhalb der gewählten Signifikanzgrenze von 0,2. Die Hazard Ratios in den Subgruppen „Vorbehandlung“ und „keine Vorbehandlung“ unterscheiden sich mit 0,661 (95%-KI: [0,512;0,852]) und 0,797 (95%-KI: [0,679;0,936]) nur mäßig und sprechen beide zugunsten von Aflibercept. Die entsprechenden Werte, die sich bei Durchführung der gleichen Subgruppenanalyse unter Ausschluss der adjuvant mit Oxaliplatin vorbehandelten Patienten für die Hazard Ratios ergeben, sind nur minimal von denen der ITT-Population verschieden (Hazard Ratio bei Bevacizumab-Vorbehandlung 0,645 [0,498;0,835] gegenüber 0,777 [0,655;0,921] ohne Bevacizumab-Vorbehandlung), und der p-Wert des Interaktionstests liegt oberhalb der festgelegten Schwelle von 0,2 für einen Hinweis auf Interaktion.

Außerdem existiert noch ein Hinweis auf eine signifikante Interaktion zwischen der Anzahl der Organe, die von Metastasen betroffen sind, und der Behandlungsart ($p=0,1455$). In der Subgruppe der Patienten, bei denen höchstens ein Organ von Metastasen betroffen ist, beträgt das Hazard Ratio 0,655 (95%-KI: [0,527;0,815]), und in der Patientengruppe, bei denen mehr als ein Organ von Metastasen betroffen ist, 0,822 (95%-KI: [0,690;0,979]). Da beide Punktschätzer auf ein besseres PFS im FOLFIRI+Aflibercept-Arm hindeuten, ist die Interaktion lediglich quantitativ.

Eine Interaktion der anderen subgruppenbildenden Merkmale mit dem Behandlungstyp kann anhand des zugrunde liegenden Signifikanzniveaus ausgeschlossen werden.

4.3.1.3.6.3 Zusammenfassung der Ergebnisse zur Objektiven Ansprechrate (ORR)

Die ORR (komplette + partielle Remission) in der auswertbaren Population betrug 19,8% (95%-KI: [16,4;23,2]) in der FOLFIRI+Aflibercept-Gruppe gegenüber 11,1% (95%-KI: [8,5;13,8]) in der FOLFIRI+Placebo-Gruppe. Das Odds Ratio für die Ansprechrate zwischen dem FOLFIRI+Aflibercept-Arm und dem FOLFIRI+Placebo-Arm betrug 1,968 (95%-KI: [1,394;2,778]), das RR 1,776 (95%-KI: [1,322;2,387]). Der stratifizierte Cochran-Mantel-Haenszel-Test zeigte (bezüglich Odds Ratio und RR) einen signifikanten Unterschied zwischen beiden Behandlungsgruppen zugunsten der FOLFIRI+Aflibercept-Gruppe ($p=0,0001$).

Für die ORR wurden keine Subgruppenanalysen durchgeführt.

4.3.1.3.6.4 Zusammenfassung der Ergebnisse zur Sicherheit (Unerwünschte Ereignisse)

Insgesamt traten UEs in der FOLFIRI+Aflibercept-Gruppe häufiger als in der FOLFIRI+Placebo-Gruppe auf. Die häufigsten UEs mit mehr als 5 Prozentpunkten Unterschied zwischen der FOLFIRI+Aflibercept-Gruppe und der FOLFIRI+Placebo-Gruppe waren gastrointestinale Erkrankungen (93,5% in der FOLFIRI+Aflibercept-Gruppe vs. 86,1% in der FOLFIRI+Placebo-Gruppe) und allgemeine Beschwerden (71,5% in der

FOLFIRI+Aflibercept-Gruppe vs. 62,5% in der FOLFIRI+Placebo-Gruppe). Weitere UEs, bei denen der Unterschied zwischen beiden Behandlungsarmen bei mehr als 5 Prozentpunkten lag, waren Infektionen und parasitäre Erkrankungen (46,2% in der FOLFIRI+Aflibercept-Gruppe vs. 32,7% in der FOLFIRI+Placebo-Gruppe), Erkrankungen des Blutes und des Lymphgefäßsystems (44,2% in der FOLFIRI+Aflibercept-Gruppe vs. 38,3% in der FOLFIRI+Placebo-Gruppe), Stoffwechsel- und Ernährungsstörungen (38,1% in der FOLFIRI+Aflibercept-Gruppe vs. 26,4% in der FOLFIRI+Placebo-Gruppe), Störungen/Erkrankungen des Nervensystems (45,8% in der FOLFIRI+Aflibercept-Gruppe vs. 37,4% in der FOLFIRI+Placebo-Gruppe), Gefäßerkrankungen (48,9% in der FOLFIRI+Aflibercept-Gruppe vs. 22,6% in der FOLFIRI+Placebo-Gruppe), Erkrankungen der Atemwege, des Brustraums und Mediastinums (58,8% in der FOLFIRI+Aflibercept-Gruppe vs. 35,5% in der FOLFIRI+Placebo-Gruppe) sowie Erkrankungen der Nieren und Harnwege (19,1% in der FOLFIRI+Aflibercept-Gruppe vs. 9,3% in der FOLFIRI+Placebo-Gruppe). Bezüglich der Bewertung der SUEs lag der größte Unterschied zwischen beiden Behandlungsarmen im Auftreten von gastrointestinalen Erkrankungen (20,3% in der FOLFIRI+Aflibercept-Gruppe vs. 11,2% in der FOLFIRI+Placebo-Gruppe). UEs führten in der FOLFIRI+Aflibercept-Gruppe bei 164 Patienten (26,8%) und in der FOLFIRI+Placebo-Gruppe bei 73 Patienten (12,1%) zum Abbruch der Behandlung. Behandlungsübergreifend waren die häufigsten Ursachen für einen Behandlungsabbruch Müdigkeit und Asthenie, Infektionen, Durchfall, Myelosuppression, Hypertonie, Lungenembolie, Proteinurie und tiefe Venenthrombose.

Die größten Unterschiede zwischen den Behandlungsgruppen FOLFIRI+Aflibercept und FOLFIRI+Placebo in der Analyse nach Studienprotokoll präspezifizierter UEs (sog. Klasseneffekte, d. h. häufig beobachtete UEs bei der Anwendung von anti-VEGF-Wirkstoffen) lagen im Auftreten von Hämorrhagie, Hypertonie und Proteinurie.

Für die Bewertung der Sicherheit wurden Subgruppenanalysen der VELOUR-Studie post-hoc durchgeführt.

Bezüglich der Bewertung der UEs gibt es einen Hinweis auf Interaktion zwischen der Vorbehandlung mit Bevacizumab und dem Behandlungsarm ($p=0,1529$). Unter allen Patienten, bei denen UEs auftraten, beträgt das Odds Ratio in der Subgruppe der mit Bevacizumab vorbehandelten Patienten 13,66 (95%-KI: [0,76;244,17]) und in der Subgruppe der Patienten, in der keine Behandlung mit Bevacizumab vorausging, nur 1,42 (95%-KI: [0,45;4,50]). Das große Ausmaß des Unterschieds zwischen den beiden Effektschätzern lässt sich dadurch erklären, dass in den Subgruppen jeweils nur eine sehr geringe Anzahl von Patienten keine UEs angab. Insgesamt gaben nur 1,48% der Patienten keine UEs an. Das Odds Ratio ist bezüglich dieser Datensituation empfindlich und dadurch wenig aussagekräftig. In diesem Fall stellt das RR ein für die Interpretation besser geeignetes Effektmaß dar. Es ergibt sich in der Patientengruppe mit Bevacizumab-Vorbehandlung ein RR von 1,00 (95%-KI: [0,99;1,02]) und in der Patientengruppe ohne Bevacizumab-Vorbehandlung ein RR von 1,03 (95%-KI: [1,00;1,06]). Auch wenn die Werte des RR aus beiden Subgruppen augenscheinlich näher zusammenliegen als die jeweiligen Werte der Odds

Ratio, ist zu beachten, dass die Interaktion zwischen Vorbehandlung mit Bevacizumab und dem Behandlungsarm auch unter Beobachtung des RR signifikant ist ($p=0,09$). Dieser Umstand ergibt sich aus der Eigenschaft der beiden Effektmaße RR und Odds Ratio, die nicht linear sind.

In der Bewertung der SUEs und der Bewertung der UEs, die zum Tod des Patienten führten, liegen signifikante Interaktionen zwischen dem Alter und der Behandlungsart vor ($p=0,0021$ und $p=0,0222$). Unter allen Patienten, bei denen schwerwiegende unerwünschte Ereignisse auftraten, beträgt das Odds Ratio in der Altersgruppe unter 65 Jahren 1,47 (95%-KI: [1,10;1,96]) und in der Altersgruppe über 65 Jahren 3,16 (95%-KI: [2,13;4,68]). Der Interaktionseffekt ist demnach nur quantitativ. In der Bewertung der UEs, die zum Tod des Patienten führten, beträgt das Odds Ratio in der Altersgruppe unter 65 Jahre 0,76 (95%-KI: [0,39;1,51]) und in der Altersgruppe über 65 Jahre 2,55 (95%-KI: [1,17;5,54]), was auf einen qualitativen Interaktionseffekt hindeutet. Diese Resultate sprechen für einen Effekt des Alters und der Behandlungsart hinsichtlich des Auftretens von SUEs sowie von UEs, die zum Tod führten.

In der Bewertung der UEs vom Grad 3-4 und der UEs, die zum Abbruch der Behandlung führten, gibt es keine signifikanten Interaktionen zwischen subgruppenbildenden Merkmalen und Behandlungsart.

Insgesamt ergibt sich somit hinsichtlich der Sicherheit von Aflibercept ein für eine wirksame antineoplastische Therapie durchaus zu erwartendes und von einem erfahrenen Onkologen nichtsdestotrotz beherrschbares Nebenwirkungsprofil.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

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

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

Nicht 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 4-41: 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 4-42: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

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

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-43: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-44: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-45: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend.

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

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*

- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Nicht zutreffend.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

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. 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-G. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-46: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

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

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-47: 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-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-48: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

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

Nicht zutreffend.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

Nicht 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. Benennen Sie

1. Studien des pharmazeutischen Unternehmers
2. Studien aus der bibliografischen Literaturrecherche
3. Studien aus der Suche in Studienregistern
4. 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 Verzerrungspotenzial.

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

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

Nicht zutreffend.

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

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

Tabelle 4-49: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

Nicht zutreffend.

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

Nicht zutreffend.

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

4.3.2.3.2 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.

Bei der in die Nutzenbewertung eingeschlossenen Studie VELOUR handelt es sich um eine RCT der Phase III und Evidenzstufe Ib. Die Studie beinhaltet einen direkten Vergleich mit der zweckmäßigen Vergleichstherapie. Das Verzerrungspotential auf Studienebene ist als gering einzustufen; die verwendeten Endpunkte sind in den gewählten Operationalisierungen valide und patientenrelevant (siehe Abschnitte 4.2.2 und 4.2.5.2) und weisen ein geringes Verzerrungspotential auf.

Im Rahmen einer Nutzenbewertung gilt in der Regel für Aussagen eines Belegs die Anforderung, dass mindestens zwei unabhängig durchgeführte Studien mit endpunktbezogen hoher Ergebnissicherheit und entsprechendem statistisch signifikanten Effekt vorliegen müssen, deren Ergebnis nicht durch weitere vergleichbare, endpunktbezogene ergebnissichere Studien infrage gestellt wird. In Ausnahmefällen (insbesondere dann, wenn zuständige Ethikkommissionen und Zulassungsbehörden aufgrund der Schwere der Erkrankung der Zielpopulation die Genehmigung einer zweiten Phase-III-RCT im Anwendungsgebiet nicht befürworten) kann jedoch auch auf Basis einer einzelnen Studie ein Beleg abgeleitet werden. Hierfür ist die Erfüllung besonders strenger Kriterien erforderlich; die Erfordernisse beruhen auf den CPMP *Points to consider on application with: 1. meta-analyses; 2. one pivotal study* [38]. Die VELOUR-Studie als eine pivotale Studie war Grundlage der regulatorischen Zulassung von Aflibercept und erfüllt diese speziellen Kriterien in vollem Umfang:

1. Interne Validität: Die betrachtete Studie ist eine randomisierte doppelblinde Studie und weist für den patientenrelevanten primären Endpunkt OS eine hohe Ergebnissicherheit auf (Verzerrungspotential ist niedrig).
2. Externe Validität: Die Studie ist multizentrisch angelegt (176 aktive Zentren), international (28 Länder) und schließt eine große Anzahl an Patienten (1.226) ein, die zu deutlich über 80% aller Fälle die Zielpopulation laut Fachinformation von Aflibercept [3] abbilden.
3. Klinische Relevanz: Die Effekte sind für alle betrachteten Endpunkte klinisch relevant.
4. Signifikante Effekte: Die beobachteten Unterschiede für alle Wirksamkeitsendpunkte zwischen Behandlungs- und Vergleichstherapie sind statistisch signifikant.
5. Die Datenqualität der betrachteten Studie ist hoch.

6. Interne Konsistenz: Die beobachteten Effekte gelten über präspezifizierte Subgruppen hinweg.
7. Zentrumseffekte: Es wurden keine ausgeprägten Zentrumseffekte dokumentiert.
8. Die Plausibilität der getesteten Hypothesen ist in der VELOUR-Studie gegeben.

Da die oben beschriebenen hohen Qualitätsstandards erfüllt sind, können somit für alle in der vorliegenden Zusatznutzenbewertung diskutierten Endpunkte, im Falle statistisch signifikanter Ergebnisse, Aussagen getroffen werden, die als **Belege** gewertet werden können.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

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

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

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

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

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

Kolorektale Karzinome sind die europaweit am zweithäufigsten auftretende Krebserkrankung [39]. Dass trotz weitreichender Früherkennungsmaßnahmen viele Kolorektalkarzinome erst in fortgeschrittenen Stadien erkannt werden, ist eine der Hauptursachen für die insgesamt im

Vergleich mit anderen Krebserkrankungen besonders schlechte Prognose der Betroffenen. Speziell das mCRC weist mit 5-Jahres-Überlebensraten von nur noch ca. 5% im Stadium IV eine mit Fortschreiten der Erkrankung zunehmend schlechte Prognose auf (siehe auch Modul 3 A, Abschnitt 3.2.2).

Die vorhandenen Therapieoptionen beim mCRC umfassen grundsätzlich chirurgische Therapie, Radio-, bzw. Chemotherapie (letztere sowohl als Monotherapie als auch als Kombinationstherapie oder in Kombination mit zielgerichteten Wirkstoffen) sowie palliative Therapie (best supportive care). Die individuelle Auswahl und Reihenfolge der anzuwendenden Therapien erfolgt in Abhängigkeit vom Stadium der Erkrankung, der genauen Tumorlokalisation, dem Therapieziel sowie weiteren klinischen Faktoren des Patienten. Orientierung bei der Auswahl der Therapie geben Leitlinienempfehlungen. Grundsätzlich ist bei Diagnose einer metastasierten Erkrankung die Behandlung und bei Ausschluss einer operativen Resektabilität eine medikamentöse Therapie indiziert und der alleinigen Anwendung von Maßnahmen der sogenannten „best supportive care“ immer vorzuziehen [40, 41]. Die medikamentöse Therapie beinhaltet vor allem Chemotherapieregime, die mehrere Wirkstoffe kombinieren, z. B. Regime mit 5-FU (bzw. Capecitabin), Folinsäure + Oxaliplatin oder Irinotecan. Eine kombinierte Polychemotherapie wird mit dem Ziel einer Überlebensverlängerung angewendet. Mit derselben Zielsetzung können seit etwa 2004-2007 auch monoklonale Antikörper in Kombination mit diesen Kombinations-Chemotherapien eingesetzt werden. Diese sog. zielgerichteten Wirkstoffe bzw. Therapien wirken gezielt gegen VEGF (Bevacizumab) oder gegen EGFR (Cetuximab, Panitumumab), zwei der Schlüsselkomponenten in der Angiogenese bzw. der Signaltransduktion, und beeinflussen damit mittelbar das Wachstum solider Tumore (siehe Modul 3 A, Abschnitt 3.2.2).

Mehr als die Hälfte der Patienten mit mCRC, die in der Erstlinientherapie eine Standard-Chemotherapie mit oder ohne zielgerichteten Wirkstoff erhalten haben, wechseln neuesten Erkenntnissen zufolge danach auf Zweit- und später auch Drittlinienbehandlungen [42, 43]. Gerade der Mangel an wirksamen zielgerichteten Therapieoptionen in der Zweitlinie bei dieser Patientengruppe ist problematisch, da die Erkrankten mit einer durchschnittlichen Überlebenszeit von nur noch ca. 12 Monaten [44] in der Regel sehr schlechte Prognosen besitzen. Das Fehlen von effektiven zielgerichteten Behandlungsalternativen, die das Patientenüberleben mit mCRC nach bereits erfolgter Oxaliplatin-Vortherapie klinisch signifikant verlängern können, stellt daher einen echten und erheblichen medizinischen Bedarf dar [43]. Insbesondere war nach Vorbehandlung des mCRC mit einer Oxaliplatin-basierten Chemotherapie keine zielgerichtete Therapie bislang in der Lage, in Kombination mit einer Irinotecan-basierten Chemotherapie einen zusätzlichen Überlebensvorteil auf Basis einer breiten Patientenpopulation klinisch zu beweisen. Durch Aflibercept kann in dieser klinisch schwierigen Situation nun erstmalig die Überlebenszeit signifikant verlängert werden.

Zwei von drei der derzeit in Deutschland für die Behandlung des mCRC zugelassenen zielgerichteten Wirkstoffe stehen darüber hinaus nur für etwa 60-68% der Patientenpopulation

(diejenigen ohne Mutation des Kirsten Rat Sarcoma 2 viral oncogene homolog (K-Ras-Gens) [45]) zur Verfügung. Ein Bedarf an Alternativen, von denen alle Erkrankten unabhängig ihres K-Ras-Status profitieren, besteht also zusätzlich.

Obwohl die Verfügbarkeit neuartiger zielgerichteter Therapien die standardmäßige Behandlung von Patienten mit mCRC verändert hat, ließ sich bislang das Überleben von Erkrankten, die nach oder unter der Therapie mit Oxaliplatin einen Progress erlitten haben, nicht über den mit dem derzeitigen Behandlungsstandard FOLFIRI erzielbaren Rahmen hinaus verlängern. Dies unterstreicht den geschilderten therapeutischen Bedarf auch innerhalb der Zielpopulation von Aflibercept [43].

Eine Irinotecan-basierte Chemotherapie gilt als gute Option in mit Oxaliplatin vorbehandelten mCRC-Patienten. Bislang gab es jedoch keine Evidenz für eine signifikante Verlängerung des Überlebens der Patienten durch den kombinierten Einsatz einer zielgerichteten Therapie und einer Kombinations-Chemotherapie auf Irinotecan-Basis in der Zweitlinientherapie des mCRC. Mit Aflibercept und seiner spezifischen Wirkweise kann diesem medizinischen Bedarf an Therapieoptionen, die einen signifikanten Überlebensvorteil sichern, erstmals begegnet werden.

In der Zulassungsstudie von Aflibercept, VELOUR, konnte für die Behandlung von Patienten mit mCRC, nach Vorbehandlung mit einem Oxaliplatin-haltigem Therapieregime, mit Aflibercept in Kombination mit FOLFIRI im Vergleich zur Behandlung mit FOLFIRI+Placebo für alle erhobenen Wirksamkeitsendpunkte ein statistisch signifikanter Vorteil nachgewiesen werden. An herausragender Stelle ist hier vor allem das Resultat zum primären Wirksamkeitsendpunkt OS zu betrachten: Die mediane Überlebenszeit [95,34%-KI] für die Patienten im FOLFIRI+Aflibercept-Arm lag bei 13,50 [12,52;14,95] Monaten. Im Gegensatz dazu betrug die mediane Überlebenszeit im FOLFIRI+Placebo-Arm lediglich 12,06 [11,07;13,11] Monate. Dies entspricht einer signifikanten positiven Differenz von 1,44 Monaten im Aflibercept-Arm. Die frühe und anhaltende Separierung der Überlebenskurven für beide Arme im zugehörigen Kaplan-Meier-Diagramm (siehe Abbildung 2) deutet auf einen kontinuierlichen Effekt hin (25%/ 50%/ 75% Quantile: 0,79/ 1,44/ 4,56). Zur Unterstützung dieser Ergebnisse wurde zusätzlich die mittlere Überlebenszeit berechnet [34]. Dabei wurden die Kaplan-Meier-Kurven mit Survivalfunktionen extrapoliert, wobei nach gängigen Modellwahlkriterien (Akaikes Informationskriterium, Bayessches Informationskriterium sowie graphische Methoden) ausgewählte Modelle mit log-logistischer Verteilung zur besten Anpassung an die Daten in der ITT-Population führten. In einem ersten Modell wurde für jede Behandlungsgruppe eine separate Survivalfunktion angepasst. Dadurch ergab sich hier in der FOLFIRI+Aflibercept-Gruppe eine mittlere Überlebenszeit von 22,8 Monaten gegenüber 18,1 Monaten in der FOLFIRI+Placebo-Gruppe (4,7 Monate Differenz). In einem zweiten Modell wurde eine Kovariable für die Behandlungsart hinzugefügt und eine Survivalfunktion an die Daten aus beiden Behandlungsarmen angepasst. In diesem Modell ergab sich in der FOLFIRI+Aflibercept-Gruppe eine mittlere Überlebenszeit von 21,5 Monaten gegenüber 19,0 Monaten in der FOLFIRI+Placebo-Gruppe, was einer Differenz von 2,5 Monaten zugunsten von Aflibercept entspricht. Als Sensitivitätsanalyse wurden die beiden

genannten Modelle analog mit einer Weibull-Verteilung durchgeführt. Auch hier sind die mittleren Überlebenszeiten in der FOLFIRI+Aflibercept-Gruppe länger als in der FOLFIRI+Placebo-Gruppe (3,0 Monate im Modell mit separaten Survivalfunktionen und 2,7 Monate im Modell mit Kovariable für Behandlung). Angesichts der Schwere der Erkrankung, in der eine Heilung nicht mehr möglich ist und bei der den Erkrankten bei einem weiteren Therapieversagen in der Regel nur noch palliative, hinsichtlich einer Verlängerung des Überlebens weitestgehend wirkungslose Maßnahmen zur Verfügung stehen, die zudem auch noch eine unter Umständen belastende Umstellungsphase erforderlich machen, sind die erbrachten Belege einer Verlängerung des Überlebens der Patienten durch Aflibercept in jedem Fall als Zusatznutzen von **erheblichem** Ausmaß einzustufen.

Die Studienergebnisse zum OS werden substantiell durch die Resultate der sekundären Endpunkte PFS und ORR unterstützt. So betrug das mediane PFS in der FOLFIRI+Aflibercept-Gruppe 6,90 Monate (99,99%-KI: [5,88;7,85], 95%-KI: [6,51;7,20]) und in der FOLFIRI+Placebo-Gruppe lediglich 4,67 Monate (99,99%-KI: [4,07;5,55], 95%-KI: [4,21;5,36]). Dies entspricht einer Verlängerung des medianen PFS von mehr als zwei Monaten. Die ORR (komplette + partielle Remission) in der auswertbaren Population war mit 19,8% (95%-KI: [16,4;23,2]) in der FOLFIRI+Aflibercept-Gruppe gegenüber 11,1% (95%-KI: [8,5;13,8]) in der FOLFIRI+Placebo-Gruppe fast doppelt so hoch. Unter zusätzlicher Berücksichtigung der stabilen Krankheitsrate (stable disease rate) [33] lässt sich zusammenfassend berichten, dass unter FOLFIRI+Aflibercept lediglich 10,4% der Patienten einen fortschreitenden Progress hatten, während bei 89,6% eine direkte verlangsamende Wirkung auf die Krankheitsprogression beobachtet wurde. Ungeachtet einer Diskussion bezüglich einer Verwendung als patientenrelevante Endpunkte oder reine Surrogatendpunkte (siehe auch Abschnitt 4.2.2 und 4.2.5.2) unterstreichen die Resultate der sekundären Endpunkte zusätzlich auch die Robustheit der Resultate zum OS.

Patientenberichtete oder anderweitig erfasste Daten zur gesundheitsbezogenen Lebensqualität (HRQoL) der Teilnehmer wurden innerhalb der VELOUR-Studie nicht erhoben. Neueste Untersuchungen auf diesem Gebiet lassen jedoch Zweifel zu, dass die typischerweise in den etablierten Erhebungsinstrumenten betrachteten Kriterien noch zeitgemäß die aus Sicht der Patienten wirklich wichtigsten Aspekte ansprechen: Therapeutische Maßnahmen gerade auf dem Gebiet der am häufigsten innerhalb von Krebsfragebögen adressierten Symptome wie Schmerzen, Übelkeit und Erbrechen haben innerhalb der letzten zwanzig Jahre deutliche Fortschritte gemacht. Es ist daher vielleicht nicht sonderlich überraschend, dass sich neueren Umfragen zufolge der Fokus der von Krebspatienten als besonders belastend beschriebenen Beschwerden auf Symptome wie Erschöpfung und Müdigkeit, Zukunftsängste und die Angst vor Haarausfall verschoben zu haben scheint [46]. Dieser Verlagerung auf psychologische Aspekte der Krebserkrankung muss erst noch durch neu konzipierte und validierte Erhebungsmethoden Rechnung getragen werden. Gerade vor dem Hintergrund der psychologischen Komponente erscheint das Wissen eines schwer an Krebs erkrankten Patienten um ein Ansprechen seiner Erkrankung auf die angewandte Therapie oder um eine Verlangsamung/ein Anhalten der Progression als durchaus wichtiger Aspekt. Damit ergibt sich ein weiteres Argument, die sekundären Endpunkte der VELOUR-Studie, PFS und ORR

eben nicht nur als bloße Surrogatendpunkte für das OS zu sehen, sondern dafür, sie abhängig von der betrachteten Patientengruppe als harte patientenrelevante Endpunkte zu betrachten – und zwar nicht nur bezüglich der Morbidität sondern durchaus auch mit Auswirkungen auf die Lebensqualität der Betroffenen. Um die Datenlage zur Lebensqualität aus der VELOUR-Studie zu verbessern und auch auf diesem Gebiet belastbare Erkenntnisse zu gewinnen, wurde von Sanofi-Aventis eine einarmige Open-Label-Studie initiiert, in der dieser Endpunkt derzeit speziell in Patienten mit mCRC nach Vorbehandlung mit einem Oxaliplatin-haltigen Regime erhoben wird [4]. Zusätzlich werden auch noch weitere Daten zur Sicherheit von Aflibercept in dieser Studie erfasst.

Den Resultaten zur Wirksamkeit gegenüber steht ein für eine wirksame antineoplastische Therapie zu erwartendes und in Händen eines erfahrenen Onkologen steuerbares Nebenwirkungsprofil (siehe Abschnitt 4.3.1.3.4). Lebensbedrohliche UEs schlagen sich zudem direkt in der beobachteten Überlebensrate nieder und sind somit auch in den Ergebnissen zum primären Wirksamkeitsendpunkt OS bereits mit abgebildet. Die häufigsten UEs, bei denen Unterschiede von mehr als 5% zwischen beiden Behandlungsarmen der Studie zuungunsten von FOLFIRI+Aflibercept festgestellt wurden, waren gastrointestinale und allgemeine Beschwerden wie z. B. Müdigkeit oder Kraftlosigkeit. Die deutlichsten Unterschiede in der Häufigkeit des Auftretens zwischen beiden Armen bestanden bei Gefäßerkrankungen sowie bei Erkrankungen der Atemwege, des Brustraums und des Mediastinums. Eine indikationsübergreifende Meta-Analyse von Ergebnissen zur Sicherheit von Aflibercept aus drei großen Studien [36], deren Erkenntnisse auch in die Fachinformation mit aufgenommen wurden, kam zu dem Resultat, dass Klasseneffekte, d. h. Risiken, welche bei Anwendung von VEGF-Inhibitoren häufig beobachtet werden, auch unter Aflibercept verstärkt auftreten und dass einige generell bei Behandlung mit Chemotherapien oft auftretende Nebenwirkungen offenbar bei gleichzeitiger Gabe von Aflibercept verstärkt werden. Dadurch, dass andere Tumorentitäten sowie die Anwendung von Aflibercept zusammen mit anderen Chemotherapieregimen in die Zusammenschau aufgenommen wurden, weichen die Resultate der Meta-Analyse in einigen Punkten von denen der VELOUR-Studie alleine ab. Auch in der Fachinformation werden aufgrund dessen einige Nebenwirkungen thematisiert, die in der Indikation, für die das vorliegende Dossier erstellt wurde, nicht beobachtet wurden (wie z. B. das Auftreten eines posterioren reversiblen Enzephalopathie-Syndroms, PRES) und die folglich auch in der vorliegenden Auswertung nicht im Detail aufgeführt oder diskutiert werden.

Den Ausführungen in Abschnitt 4.4.1 folgend wird die Beweissicherheit der Resultate trotz des Vorliegens von nur einer RCT als **Beleg** eingestuft.

Zur besseren Veranschaulichung der erzielten Ausmaße der beobachteten Effekte werden diese nachfolgend in Tabelle 4-50 zusammenfassend dargestellt.

Abweichend von der Methodik der durch das IQWiG bereits angewandten „Methodik der Operationalisierung“ [47] werden jedoch nicht, wie dort vorgeschlagen, die Obergrenzen der 95%-Konfidenzintervalle als Schwellenwerte für die Kategorisierung des Zusatznutzens herangezogen, da diese Methode bislang weder in der Verfahrensordnung legitimiert noch zur

öffentlichen Anhörung gestellt wurde. Zudem werden über dieses Konstrukt die Vorgaben von § 5 Satz 7 der AM-NutzenV [48], nach der der jeweilige Zusatznutzen eines Wirkstoffs gegenüber seiner zweckmäßigen Vergleichstherapie **unter Berücksichtigung des Schweregrades der Erkrankung** zu quantifizieren ist, nicht in ausreichendem Maße berücksichtigt.

Tabelle 4-50: Zusammenfassung der Ergebnisse zum Zusatznutzen auf Endpunktebene: FOLFIRI+Aflibercept vs. FOLFIRI+Placebo

Endpunktkategorie Endpunkt	Effektschätzer FOLFIRI+Aflibercept vs. FOLFIRI+Placebo
Mortalität	
Gesamtüberleben (OS)	HR: 0,817; 95,34% *-KI: [0,713;0,937] 95%-KI: [0,714;0,934] Verlängerung des medianen OS um 1,44 (13,5 vs. 12,06) Monate Verlängerung des mittleren OS um 2,7 – 6,6 Monate (je nach Berechnungsmodell)
Morbidität	
Progressionsfreies Überleben (PFS)	HR: 0,758; 99,99% *-KI: [0,578;0,995] 95%-KI: [0,661;0,867] Verlängerung der medianen PFS-Dauer um 2,23 (6,9 vs. 4,67) Monate
Objektive Ansprechrate (ORR)	OR: 1,968; 95%-KI: [1,394;2,778] RR: 1,776; 95%-KI: [1,322;2,387] Signifikant höhere Zahl an Therapieansprechern (19,8% vs. 11,1%) Für Nicht-Ansprechen: OR: 0,508; 95%-KI:[0,360;0,717] RR: 0,903; 95%-KI:[0,857;0,951]

FOLFIRI: 5-Fluorouracil + Folinsäure + Irinotecan; HR: Hazard Ratio; OR: Odds Ratio; KI: Konfidenzintervall; RR: Relatives Risiko

* O'Brian Fleminging- α -Korrektur

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 4-51: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene mit metastasiertem Kolorektalkarzinom (mCRC), das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist.	erheblich

In der Gesamtschau der in Abschnitt 4.3.1.3 dargestellten Ergebnisse und Subgruppenanalysen ergeben sich keine hinreichend konsistenten Hinweise auf Patientengruppen innerhalb der Zielpopulation, für die ein von der Gesamtheit signifikant abweichender Zusatznutzen besteht.

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

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

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

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

Nicht relevant.

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

Tabelle 4-52: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

Es werden keine indirekten Vergleiche vorgelegt.

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.

Es werden weder nichtrandomisierte vergleichende Studien noch weitere Untersuchungen vorgelegt.

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.

Zu allen berücksichtigten patientenrelevanten Endpunkten lagen zum Zeitpunkt der Dossierabfassung valide Daten vor.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹³, Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁵) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁶) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

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

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

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

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

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

Die im vorliegenden Dossier getroffene Kernaussage zum Zusatznutzen von Aflibercept in Kombination mit FOLFIRI in der Anwendung bei Patienten mit mCRC, das unter oder nach einem Oxaliplatin-haltigen Regime fortgeschritten ist, im Vergleich zur Anwendung der zweckmäßigen Vergleichstherapie FOLFIRI basiert auf den Ergebnissen der Zulassungsstudie VELOUR im patientenrelevanten Endpunkt OS. Die Wirksamkeit wurde in VELOUR zusätzlich mittels der sekundären Endpunkte PFS und ORR erhoben, die ebenfalls für die Bewertung von Aflibercept im vorliegenden Dossier mit herangezogen werden. Beides sind nach Auffassung des IQWiG Surrogatendpunkte [6]. Die Gründe für die Patientenrelevanz dieser Endpunkte nach Auffassung von Sanofi-Aventis und den Miteinschluss in die Bewertung sind in den Abschnitten 4.2.2 und 4.2.5.2 dargelegt.

4.6 Liste der eingeschlossenen Studien

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

Zur Beantwortung der in Abschnitt 4.2.1 präzisierten Fragestellung wurde genau eine RCT als relevant identifiziert und im vorliegenden Dossier ausgewertet. Hierbei handelt es sich um die Zulassungsstudie zum zu bewertenden Wirkstoff Aflibercept, VELOUR.

Tabelle 4-53: Für die Nutzenbewertung berücksichtigte Studien und Untersuchungen

Studie	Titel	Datenquelle
VELOUR	A Multinational, Randomized, Double-blind Study, Comparing the Efficacy of Aflibercept Once Every 2 Weeks versus Placebo in Patients with Metastatic Colorectal Cancer (MCRC) Treated with Irinotecan / 5-FU Combination (FOLFIRI) after Failure of an oxaliplatin based regimen	[1], [37], [33]

4.7 Referenzliste

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

- [1] sanofi-aventis groupe. A Multinational, Randomized, Double-blind Study, Comparing the Efficacy of Aflibercept Once Every 2 Weeks versus Placebo in Patients with

- Metastatic Colorectal Cancer (MCRC) Treated with Irinotecan / 5-FU Combination (FOLFIRI) after Failure of an oxaliplatin based regimen. Clinical Study Report. 2011.
- [2] Gemeinsamer Bundesausschuss. 5. Kapitel der Verfahrensordnung. Bewertung des Nutzens von Arzneimitteln nach § 35a SGB V. Berlin: Gemeinsamer Bundesausschuss; 2011.
- [3] sanofi-aventis groupe. Fachinformation Zaltrap® 25mg/ml Konzentrat zur Herstellung einer Infusionslösung. 2013.
- [4] sanofi-aventis groupe. A Multicenter, Single arm, Open Label Clinical Trial to Evaluate the Safety and Health-Related Quality of Life of Aflibercept in Patients with Metastatic Colorectal Cancer (mCRC) Previously Treated with an Oxaliplatin-Containing Regimen. Amended Clinical Trial Protocol. 2012.
- [5] Gemeinsamer Bundesausschuss (G-BA). Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV, Beratungsanforderung 2012-B-019, Aflibercept zur Behandlung des metastasierten kolorektalen Karzinoms. 2012.
- [6] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden Version 4.0. 23.09.2011: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; 2011.
- [7] Committee for Medicinal Products for Human Use (CHMP). Draft Guideline on the Evaluation of Anticancer Medicinal Products in Man. EMA/CHMP/205/95/rev4. London: European Medicines Agency; 2011.
- [8] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Aussagekraft von Surrogatendpunkten in der Onkologie. Köln: IQWiG, IQWiG-Berichte – Jahr: 2011. 2011.
- [9] Pazdur R. Endpoints for assessing drug activity in clinical trials. *The oncologist*. 2008;13 Suppl 2:19-21.
- [10] McCain jr. JA. The Ongoing Evolution of Endpoints in Oncology. *Managed Care* [Internet]. 2010 [Zugriffsdatum: 11.10.2011]; May, 2010(Supplement). Verfügbar unter:
<http://www.nxtbook.com/nxtbooks/medimedia/oncologyendpoints/index.php?startid=Cover1>.
- [11] Saad ED, Katz A, Hoff PM, Buyse M. Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2010;21(1):7-12.
- [12] Haerting J, Victor A, Hauschke D. Kommentar zum IQWiG-Bericht A10-05 "Aussagekraft von Surrogatendpunkten in der Onkologie" (Version 1.0, Stand: 31.01.2011). In: Deutsche Gesellschaft für Medizinische Informatik Biometrie und Epidemiologie (gmds) e. V., Hrsg. Köln: gmds; 2011.
- [13] Verband Forschender Arzneimittelhersteller e.V. (vfa). vfa-Beitrag zur G-BA Diskussionsrunde "Frühe Nutzenbewertung onkologischer Arzneimittel". Berlin: Verband Forschender Arzneimittelhersteller e.V.; 2010.
- [14] Verband Forschender Arzneimittelhersteller e.V. (vfa). vfa-Position zu Endpunkten bei onkologischen Therapien - Positionspapier. Berlin: Verband Forschender Arzneimittelhersteller e.V.; 2012.
- [15] Walker MS, Hasan M, Yim YM, Yu E, Stepanski EJ, Schwartzberg LS. Retrospective study of the effect of disease progression on patient reported outcomes in HER-2 negative metastatic breast cancer patients. *Health and quality of life outcomes*. 2011;9:46.

- [16] Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2007;25(33):5218-24.
- [17] Center for Drug Evaluation and Research (CDER). Guidance for Industry - Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. Food and Drug Administration (FDA). 2007:1-19.
- [18] Committee for Medicinal Products for Human Use (CHMP). Methodological Considerations for using Progression-free Survival (PFS) as Primary Endpoint in Confirmatory Trials for Registration - Appendix 1 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man. London: European Medicines Agency, 2008.
- [19] Committee for Medicinal Products for Human Use (CHMP). Appendix 2 to the Guideline on the evaluation of Anticancer Medicinal Products in Man (CPMP/EWP/205/95 Rev. 3) on Confirmatory studies in Haematological Malignancies. London: European Medicines Agency; 2010.
- [20] Gastineau T. FDA Guidance Document Focuses on Cancer Trial Endpoints. *Applied Clinical Trials Online* [Internet]. 2006 [Zugriffdatum: 01.06.2012]. Verfügbar unter: <http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/article/articleDetail.jsp?id=324334>.
- [21] Mayfield E. Progression-Free Survival: Patient Benefit or Lower Standard? *NCI Cancer Bulletin*. 2008.
- [22] Schlesinger-Raab A, Schubert-Fritschle G, Hein R, Stolz W, Volkenandt M, Holzel D, et al. Quality of life in localised malignant melanoma. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2010;21(12):2428-35.
- [23] Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *Journal of the National Cancer Institute*. 2000;92(3):205-16.
- [24] Bureau of Metabolism Oncology and Reproductive Science (BMORS), Oncology Division. The use of progression-free survival as the efficacy endpoint for approval of targeted and chemotherapeutic agents for advanced cancer. Health Canada, Issues Analysis Summary. 2007.
- [25] Saad ED, Katz A. Progression-free survival and time to progression as primary end points in advanced breast cancer: often used, sometimes loosely defined. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2009;20(3):460-4.
- [26] Dy GK, Krook JE, Green EM, Sargent DJ, Delaunoy T, Morton RF, et al. Impact of complete response to chemotherapy on overall survival in advanced colorectal cancer: results from Intergroup N9741. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2007;25(23):3469-74.
- [27] Committee for Proprietary Medicinal Products (CPMP). ICH Topic E 9: Statistical Principles for Clinical Trials - Step 5: Note for Guidance on Statistical Principles for Clinical Trials (CPMP/ICH/363/96). London: European Medicines Agency; 1998.
- [28] Higgins JPT, Green Se. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration; 2008; Verfügbar unter: <http://www.cochrane-handbook.org/>.

- [29] Jackson D. The power of the standard test for the presence of heterogeneity in meta-analysis. *Statistics in medicine*. 2006;25(15):2688-99.
- [30] Roche Pharma AG. Fachinformation Avastin®. 2012.
- [31] [clinicaltrials.gov](http://www.clinicaltrials.gov). Aflibercept Versus Placebo in Combination With Irinotecan and 5-FU in the Treatment of Patients With Metastatic Colorectal Cancer After Failure of an Oxaliplatin Based Regimen (VELOUR). 2013 [updated 25.01.2013]; Verfügbar unter: <http://www.clinicaltrials.gov/show/NCT00561470>.
- [32] International Clinical Trials Registry Platform Search Portal. A Multinational, Randomized, Double-blind Study, Comparing the Efficacy of Aflibercept Once Every 2 Weeks versus Placebo in Patients with Metastatic Colorectal Cancer (MCRC) Treated with Irinotecan / 5-FU Combination (FOLFIRI) after failure of an oxaliplatin based regimen. - VELOUR. 2013 [updated 25.01.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2007-000820-42-DE>.
- [33] Van Cutsem E, Tabernero J, Lakomy R, Prenen H, Prausova J, Macarulla T, et al. Addition of Aflibercept to Fluorouracil, Leucovorin, and Irinotecan Improves Survival in a Phase III Randomized Trial in Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin-Based Regimen. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2012;30(28):3499-506.
- [34] Joulain F, Van Cutsem E, Iqbal SU, Hoyle M, Allegra CJ. Aflibercept versus placebo in combination with FOLFIRI in previously treated metastatic colorectal cancer (mCRC): Mean overall survival (OS) estimation from a phase III trial (VELOUR). ASCO Scientific Meeting Abstract; Chicago2012.
- [35] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). ICH Tripartite Guideline. Guideline for Good Clinical Practice 1996.
- [36] Allegra C, Tabernero J, Rougier P, Scagliotti G, Philip PA, Lakomy R, et al. Meta-Analysis of Anti VEGF Class Adverse Events from 3 Double Blind (Db) Placebo (Pbo) Controlled Phase 3 Trials with IV Aflibercept (Afl). *Gastrointestinal Cancers Symposium (ASCO GI)*; San Francisco, CA2012.
- [37] sanofi-aventis groupe. A Multinational, Randomized, Double-blind Study, Comparing the Efficacy of Aflibercept Once Every 2 Weeks versus Placebo in Patients with Metastatic Colorectal Cancer (MCRC) Treated with Irinotecan / 5-FU Combination (FOLFIRI) after Failure of an oxaliplatin based regimen. Erratum for Clinical Study Report. 2012.
- [38] Committee for Proprietary Medicinal Products (CPMP). Points to Consider on Application with 1. Meta-Analyses; 2. One Pivotal Study. London: European Medicines Agency, 2001.
- [39] Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2007;18(3):581-92.
- [40] Heinemann V, Engel J, Gross M, Kleespies A, Kolligs F, Kopp R, et al. Kolonkarzinom. In: Bruns CJ, Hrsg. *Gastrointestinale Tumoren Empfehlungen zur Diagnostik, Therapie und Nachsorge*. 8 Auflage. München: Tumorzentrum München; 2010. S. 139-85.
- [41] Kreis ME, Tymptner C, Geinitz H, Adam M, Gross M, Michl GM, et al. Rektumkarzinom. In: Bruns CJ, Hrsg. *Gastrointestinale Tumoren Empfehlungen zur Diagnostik, Therapie und Nachsorge*. 8 Auflage. München: Tumorzentrum München; 2010. S. 186-210.

- [42] Kantar Health, sanofi-aventis groupe. mCRC chart pull OMT Germany Report. 2012.
- [43] Thompson SF, Iqbal SU, Naoshy S, Ng DB, Andria ML, Sherman SA, et al. Current practices in 1st- (1L), 2nd- (2L) and 3rd-line (3L) treatment for metastatic colorectal cancer (mCRC). ASCO Scientific Meeting Abstract; Chicago2012.
- [44] Giantonio BJ, Catalano PJ, Meropol NJ, O'Dwyer PJ, Mitchell EP, Alberts SR, et al. Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: results from the Eastern Cooperative Oncology Group Study E3200. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2007;25(12):1539-44.
- [45] De Roock W, De Vriendt V, Normanno N, Ciardiello F, Tejpar S. KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. *The lancet oncology*. 2011;12(6):594-603.
- [46] Tauchert FK, Hofheinz R-D, Quidde J, Marschner N, Hipp M, Weber M, et al. Quality of life incriminating symptoms in cancer patients and their weight in the doctor-patient-talk: A survey of the "Quality of Life" Working Group of the "Arbeitsgemeinschaft Internistische Onkologie" (AIO). 37th ESMO Congress; Wien 2012.
- [47] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Ticagrelor: Nutzenbewertung gemäß § 35a SGB V - Anhang A: Erläuterungen zur Operationalisierung des Ausmaßes des Zusatznutzens gemäß AM-NutzenV auf Endpunktebene. Köln: 2011.
- [48] Bundesanzeiger. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV). BGBl I S 2324. Berlin: Bundesgesetzblatt; 2010. S. 2324-8.
- [49] [clinicaltrials.gov](http://www.clinicaltrials.gov). A Study of Aflibercept Versus Placebo With FOLFIRI in Asian Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin Chemotherapy (AFLAME). 2013 [28.01.2013]; Verfügbar unter: <http://www.clinicaltrials.gov/ct2/show/NCT01661270>.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

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

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

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

Eine bibliografische Literaturrecherche wurde nicht durchgeführt (siehe Abschnitt 4.3.1.1.2).

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.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

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

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

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

Eine Registersuche nach für die Zusatznutzenbewertung relevanten klinischen Studien zu Aflibercept wurde am 10.01.2013 gemäß den Vorgaben der VerFO im Internetauftritt des internationalen Studienregisters ClinicalTrials.gov und über das International Clinical Trials Registry Platform (ICTRP) Search Portal der World Health Organization (WHO) durchgeführt.

Alle identifizierten Studien wurden anhand der in Abschnitt 4.2.2 prädefinierten Ein- und Ausschlusskriterien selektiert. Als für die Zusatznutzenbewertung von Aflibercept plus FOLFIRI gegenüber FOLFIRI relevant eingestufte Studien wurden in den in Abschnitt 4.3.1.1.3 ausführlich dargestellten Studienpool aufgenommen.

Studienregister	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	10. Januar 2013
Suchstrategie	colorectal carcinoma OR colorectal cancer OR colorectal neoplasm OR colon carcinoma OR colon cancer OR colon neoplasm OR rectal carcinoma OR rectal cancer OR rectal neoplasm [condition] AND aflibercept OR AVE0005 OR zaltrap [intervention] AND [Recruitment status] is ALL
Treffer	11

Tabelle 4-54 (Anhang): Im International Clinical Trials Registry Platform Search Portal (ICTRP) identifizierte Studien

Recruitment status	Main ID	Public Title	Date of Registration
Not recruiting	NCT01669720	<u>Adjuvant Aflibercept for Metastatic Colorectal Cancer</u>	07/08/2012
Recruiting	NCT01661270	<u>A Study of Aflibercept Versus Placebo With FOLFIRI in Asian Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin Chemotherapy</u>	07/08/2012
Recruiting	NCT01661972	<u>Phase I/II Study of Capecitabine Plus Aflibercept to Treat Metastatic Colorectal Cancer</u>	31/07/2012
Not recruiting	NCT01652196	<u>Aflibercept and FOLFOX6 Treatment for Previously Untreated Stage IV Colorectal Cancer</u>	25/07/2012
Not recruiting	NCT01646554	<u>Efficacy of FOLFOX Versus FOLFOX Plus Aflibercept in K-ras Mutant Patients With Resectable Liver Metastases</u>	16/07/2012
Recruiting	NCT01670721	<u>Colorectal Cancer Metastatic</u>	13/07/2012
Recruiting	NCT01571284	<u>Safety and Quality of Life Study of Aflibercept in Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin-Based Regimen</u>	03/04/2012
Completed	NCT00921661	<u>Study of Intravenous Aflibercept in Combination With FOLFIRI in Japanese Patients With Metastatic Colorectal Cancer</u>	15/06/2009
Not recruiting	EUCTR2008-004178-41-DE	<u>Randomized, Multinational, Study Of Aflibercept And Modified FOLFOX6 As First-Line Treatment In Patients With Metastatic Colorectal Cancer - AFFIRM</u>	30/09/2008

Recruitment status	Main ID	Public Title	Date of Registration
Not recruiting	EUCTR2007-000820-42-DE	<u>A Multinational, Randomized, Double-blind Study, Comparing the Efficacy of Aflibercept Once Every 2 Weeks versus Placebo in Patients with Metastatic Colorectal Cancer (MCRC) Treated with Irinotecan / 5-FU Combination (FOLFIRI) after failure of an oxaliplatin based regimen. - VELOUR</u>	28/06/2007
Not recruiting	NCT00407654	<u>VEGF Trap in Treating Patients With Previously Treated Metastatic Colorectal Cancer</u>	04/12/2006

Studienregister clinicaltrials.gov

Internetadresse <http://www.clinicaltrials.gov>

Datum der Suche 10. Januar 2013

Suchstrategie colorectal carcinoma OR colorectal cancer OR colorectal neoplasm OR colon carcinoma OR colon cancer OR colon neoplasm OR rectal carcinoma OR rectal cancer OR rectal neoplasm [condition] AND aflibercept OR AVE0005 OR zaltrap [intervention] AND ALL studies [recruitment]

Treffer 13

Tabelle 4-55 (Anhang): In clinicaltrials.gov identifizierte Studien

Registereintrag	Studie
NCT01652196	<p>Titel: Aflibercept and FOLFOX6 Treatment for Previously Untreated Stage IV Colorectal Cancer</p> <p>Rekrutierungsstatus: Recruiting</p> <p>Erkrankung: Mucinous Adenocarcinoma of the Colon; Mucinous Adenocarcinoma of the Rectum; Signet Ring Adenocarcinoma of the Colon; Signet Ring Adenocarcinoma of the Rectum; Stage IV Colon Cancer; Stage IV Rectal Cancer</p>

Registereintrag	Studie
	<p>Interventionen: Biological: aflibercept; Drug: oxaliplatin; Drug: leucovorin calcium; Drug: fluorouracil; Other: laboratory biomarker analysis; Procedure: dynamic contrast-enhanced magnetic resonance imaging; Radiation: fludeoxyglucose F 18; Procedure: positron emission tomography</p> <p>Sponsor/Collaborators: Ohio State University Comprehensive Cancer Center; Richard Goldberg; Sanofi</p> <p>Phase: Phase 2</p> <p>Enrollment: 70</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Endpoint Classification: Efficacy Study; Intervention Model: Single Group Assignment; Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: OSU 11182, NCI-2012-01167</p> <p>Startdatum: August 2012</p> <p>Abschlussdatum:</p> <p>Acronym:</p> <p>URL: http://www.clinicaltrials.gov/ct2/show/NCT01652196</p>
NCT01646554	<p>Titel: Efficacy of FOLFOX Versus FOLFOX Plus Aflibercept in K-ras Mutant Patients With Resectable Liver Metastases</p> <p>Rekrutierungsstatus: Not yet recruiting</p> <p>Erkrankung: Colorectal Cancer Metastatic; Liver Metastases; KRAS Mutated Colorectal Cancer</p> <p>Interventionen: Drug: Modified FOLFOX6; Biological: Aflibercept; Procedure: Surgery</p> <p>Sponsor/Collaborators: European Organisation for Research and Treatment of Cancer - EORTC; Sanofi</p> <p>Phase: Phase 2 / Phase 3</p> <p>Enrollment: 240</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Allocation: Randomized; Endpoint Classification: Safety/Efficacy Study; Intervention Model: Parallel Assignment; Masking: Open Label; Primary Purpose: Treatment</p>

Registereintrag	Studie	
	Andere IDs:	EORTC-1207, 2012-002317-18
	Startdatum:	December 2012
	Abschlussdatum:	December 2016
	Acronym:	BOS3
	URL:	http://www.clinicaltrials.gov/ct2/show/NCT01646554
NCT01669720	Titel:	Adjuvant Aflibercept for Metastatic Colorectal Cancer
	Rekrutierungsstatus:	Not yet recruiting
	Erkrankung:	Metastatic Colorectal Cancer
	Interventionen:	Drug: Aflibercept
	Sponsor/Collaborators:	Brown University; Rhode Island Hospital; The Miriam Hospital; University of Vermont; Virginia Mason Hospital/Medical Center; Stanford University; Comprehensive Cancer Center of Wake Forest University; Thomas Jefferson University
	Phase:	Phase 2
	Enrollment:	69
	Studientypus:	Interventional
	Studiendesign:	Allocation: Randomized; Endpoint Classification: Efficacy Study; Intervention Model: Parallel Assignment; Masking: Open Label; Primary Purpose: Treatment
	Andere IDs:	BrUOG C261
	Startdatum:	December 2012
	Abschlussdatum:	January 2016
	Acronym:	C261
	URL:	http://www.clinicaltrials.gov/ct2/show/NCT01669720
NCT01661972	Titel:	Phase I/II Study of Capecitabine Plus Aflibercept to Treat Metastatic Colorectal Cancer
	Rekrutierungsstatus:	Recruiting
	Erkrankung:	Metastatic Colorectal Cancer
	Interventionen:	Drug: Capecitabine and aflibercept
	Sponsor/Collaborators:	Duke University; Fatima Rangwala; Sanofi
	Phase:	Phase 1 / Phase 2

Registereintrag	Studie
	<p>Enrollment: 60</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Endpoint Classification: Safety/Efficacy Study; Intervention Model: Single Group Assignment; Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: Pro00037688</p> <p>Startdatum: August 2012</p> <p>Abschlussdatum: August 2017</p> <p>Acronym: X-TRAP</p> <p>URL: http://www.clinicaltrials.gov/ct2/show/NCT01661972</p>
NCT00851084	<p>Titel: Study of Aflibercept And Modified FOLFOX6 As First-Line Treatment In Patients With Metastatic Colorectal Cancer</p> <p>Rekrutierungsstatus: Completed</p> <p>Erkrankung: Colorectal Neoplasms; Neoplasm Metastasis</p> <p>Interventionen: Drug: aflibercept (AVE0005); Drug: oxaliplatin; Drug: 5-FU; Drug: Folinic Acid</p> <p>Sponsor/Collaborators: Sanofi</p> <p>Phase: Phase 2</p> <p>Enrollment: 268</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Allocation: Randomized; Endpoint Classification: Efficacy Study; Intervention Model: Parallel Assignment; Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: EFC10668, EudraCT 2008-004178-41</p> <p>Startdatum: February 2009</p> <p>Abschlussdatum: January 2012</p> <p>Acronym: AFFIRM</p> <p>URL: http://clinicaltrials.gov/show/NCT00851084</p>
NCT01670721	<p>Titel: Colorectal Cancer Metastatic</p> <p>Rekrutierungsstatus: Recruiting</p> <p>Erkrankung: Colorectal Cancer Metastatic</p>

Registereintrag	Studie
	<p>Interventionen: Drug: AFLIBERCEPT AVE0005; Drug: irinotecan; Drug: fluorouracil; Drug: Leucovorin</p> <p>Sponsor/Collaborators: Sanofi; Regeneron Pharmaceuticals</p> <p>Phase: Phase 3</p> <p>Enrollment: 200</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Endpoint Classification: Safety Study; Intervention Model: Single Group Assignment; Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: AFLIBL06266, 2012-000048-89, U1111-1128-9325</p> <p>Startdatum: August 2012</p> <p>Abschlussdatum: June 2014</p> <p>Acronym: AFEQT</p> <p>URL: http://www.clinicaltrials.gov/ct2/show/NCT01670721</p>
NCT01571284	<p>Titel: Safety and Quality of Life Study of Aflibercept in Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin-Based Regimen</p> <p>Rekrutierungsstatus: Recruiting</p> <p>Erkrankung: Colorectal Cancer Metastatic</p> <p>Interventionen: Drug: AFLIBERCEPT AVE0005; Drug: FOLFIRI</p> <p>Sponsor/Collaborators: Sanofi; Regeneron Pharmaceuticals</p> <p>Phase: Phase 3</p> <p>Enrollment: 900</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Endpoint Classification: Safety Study; Intervention Model: Single Group Assignment; Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: AFLIBC06097, 2011-005724-17, U1111-1125-8949</p> <p>Startdatum: May 2012</p> <p>Abschlussdatum: June 2014</p> <p>Acronym:</p> <p>URL: http://clinicaltrials.gov/show/NCT01571284</p>

Registereintrag	Studie
NCT01661270	<p>Titel: A Study of Aflibercept Versus Placebo With FOLFIRI in Asian Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin Chemotherapy</p> <p>Rekrutierungsstatus: Recruiting</p> <p>Erkrankung: Colorectal Cancer Metastatic</p> <p>Interventionen: Drug: Aflibercept AVE0005; Drug: Placebo</p> <p>Sponsor/Collaborators: Sanofi; Regeneron Pharmaceuticals</p> <p>Phase: Phase 3</p> <p>Enrollment: 300</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Allocation: Randomized; Endpoint Classification: Efficacy Study; Intervention Model: Parallel Assignment; Masking: Double Blind (Subject, Investigator); Primary Purpose: Treatment</p> <p>Andere IDs: EFC11338, U1111-1115-7227</p> <p>Startdatum: July 2012</p> <p>Abschlussdatum: January 2016</p> <p>Acronym: AFLAME</p> <p>URL: http://www.clinicaltrials.gov/ct2/show/NCT01661270</p>
NCT00407654	<p>Titel: VEGF Trap in Treating Patients With Previously Treated Metastatic Colorectal Cancer</p> <p>Rekrutierungsstatus: Unknown</p> <p>Erkrankung: Colorectal Cancer</p> <p>Interventionen: Biological: aflibercept; Other: immunoenzyme technique; Other: pharmacological study</p> <p>Sponsor/Collaborators: National Cancer Institute (NCI); Princess Margaret Hospital, Canada</p> <p>Phase: Phase 2</p> <p>Enrollment: 80</p> <p>Studientypus: Interventional</p> <p>Studiendesign: Masking: Open Label; Primary Purpose: Treatment</p> <p>Andere IDs: CDR0000518293, PMH-PHL-050, NCI-7498</p> <p>Startdatum: October 2006</p>

Registereintrag	Studie
	Abschlussdatum: Acronym: URL: http://clinicaltrials.gov/show/NCT00407654
NCT00561470	Titel: Aflibercept Versus Placebo in Combination With Irinotecan and 5-FU in the Treatment of Patients With Metastatic Colorectal Cancer After Failure of an Oxaliplatin Based Regimen Rekrutierungsstatus: Completed, Has Results Erkrankung: Colorectal Neoplasms; Neoplasm Metastasis Interventionen: Drug: Placebo; Drug: Aflibercept (ziv-aflibercept, AVE0005, VEGF trap, ZALTRAP®); Drug: FOLFIRI (Irinotecan, 5-Fluorouracil, and Leucovorin) Sponsor/Collaborators: Sanofi; Regeneron Pharmaceuticals; National Surgical Adjuvant Breast and Bowel Project (NSABP) Phase: Phase 3 Enrollment: 1226 Studientypus: Interventional Studiendesign: Allocation: Randomized; Endpoint Classification: Efficacy Study; Intervention Model: Parallel Assignment; Masking: Double Blind (Subject, Caregiver, Investigator); Primary Purpose: Treatment Andere IDs: EFC10262, EudraCT 2007-000820-42 Startdatum: November 2007 Abschlussdatum: June 2012 Acronym: VELOUR URL: http://clinicaltrials.gov/show/NCT00561470
NCT01749956	Titel: 5-FU, Aflibercept, and RT for Preoperative and Postoperative Patients With Stage II/III Rectal Cancer Rekrutierungsstatus: Not yet recruiting Erkrankung: Rectal Cancer Interventionen: Radiation: Radiation Therapy; Drug: Aflibercept; Drug: 5-Fluorouracil; Drug: Leucovorin; Drug: Oxaliplatin; Procedure: Surgery Sponsor/Collaborators: Sarah Cannon Research Institute; Sanofi

Registereintrag	Studie	
	Phase:	Phase 2
	Enrollment:	39
	Studientypus:	Interventional
	Studiendesign:	Intervention Model: Single Group Assignment; Masking: Open Label; Primary Purpose: Treatment
	Andere IDs:	SCRI GI 168
	Startdatum:	December 2012
	Abschlussdatum:	March 2015
	Acronym:	
	URL:	http://www.clinicaltrials.gov/ct2/show/NCT01749956
NCT00921661	Titel:	Study of Intravenous Aflibercept in Combination With FOLFIRI in Japanese Patients With Metastatic Colorectal Cancer
	Rekrutierungsstatus:	Completed
	Erkrankung:	Colorectal Neoplasms; Neoplasm Metastasis
	Interventionen:	Drug: AVE0005 (aflibercept)
	Sponsor/Collaborators:	Sanofi; Regeneron Pharmaceuticals
	Phase:	Phase 1
	Enrollment:	16
	Studientypus:	Interventional
	Studiendesign:	Allocation: Non-Randomized Endpoint Classification: Safety Study Intervention Model: Single Group Assignment Masking: Open Label Primary Purpose: Treatment
	Andere IDs:	TCD10794
	Startdatum:	June 2009
	Abschlussdatum:	June 2012
	Acronym:	
	URL:	http://clinicaltrials.gov/show/NCT00921661
NCT01754272	Titel:	A Non-Interventional Follow-Up to the VELOUR Study - Translational Research
	Rekrutierungsstatus:	Active, not recruiting
	Erkrankung:	Metastatic Colorectal Cancer
	Interventionen:	

Registereintrag	Studie
	Sponsor/Collaborators: Katholieke Universiteit Leuven
	Phase: 1226
	Enrollment:
	Studientypus: Observational
	Studiendesign: Time Perspective: Retrospective
	Andere IDs: ADX11080
	Startdatum: January 2012
	Abschlussdatum: December 2013
	Acronym:
	URL: http://www.clinicaltrials.gov/ct2/show/NCT01754272

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

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

Es wurden keine Studien im Volltext gesichtet und ausgeschlossen.

Anhang 4-D: Liste der abgebrochenen Studien

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

Tabelle 4-56 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						

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

Nicht zutreffend: der in Abschnitt 4.3.1.1 identifizierte Studienpool umfasst die beschriebene Studie (VELOUR) und die Studie AFLAME, die sich noch in der Rekrutierungsphase befindet.

Anhang 4-E: Liste der laufenden Studien

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

Tabelle 4-57 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
AFLAME	nein	ja	nein	nein	ja [49]	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

Die im Abschnitt 4.3.1.1 identifizierte Studie AFLAME befindet sich noch in der Rekrutierungsphase.

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-58 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 4-58 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-58 (Anhang): Studiendesign und -methodik für Studie EFC10262 (VELOUR)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primäres Ziel: Untersuchung des verbessernden Effektes einer Aflibercept-Behandlung im Vergleich zu Placebo auf die Überlebenszeit von Patienten mit kolorektaler Krebserkrankung in Zweitbehandlung mit FOLFIRI</p> <p>Sekundäre Ziele: Vergleich der progressionsfreien Überlebenszeit bei Patienten, die Aflibercept eingenommen haben, versus Placebo; Untersuchung des Effekts einer Aflibercept-Behandlung im Vergleich zu Placebo auf die RECIST-Responderate;</p> <p>Sicherheitsziele: Unerwünschte Ereignisse während der Studie Laborwerte, klinische Untersuchungen, Begleitmedikation</p> <p>Immunogenität Anti-Aflibercept-Antikörper-Tests</p> <p>Pharmakokinetik: Analyse des endogenen vaskulo-endothelialen Wachstumsfaktors (VEGF)</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, randomisierte, parallele, placebokontrollierte und doppelblinde Phase III-Studie Zuteilungsverhältnis 1:1 für Aflibercept 25 mg/ml : Placebo
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Änderungen vom 03.08.2007: ECOG-PS der Probanden wurde geändert (von 0-1 versus 2 auf 0 versus 1 versus 2) auf Anfrage des FDA Die objektive Ansprechrage (Response rate) wurde als sekundärer Endpunkt hinzugefügt Untersuchung auf Anti-Aflibercept-Antikörper <i>sämtlicher</i> Studienteilnehmer (anstelle der anfangs geplanten 300 Patienten) und anschließende Messung zirkulierender Werte von freien und gebundenen Aflibercept Messungen des zirkulierenden endogenen VEGF <i>sämtlicher</i> Studienteilnehmer (anstelle der anfangs geplanten 300 Patienten)</p> <p>Änderungen vom 23.04.2008: Berücksichtigung für die Stratifizierung der Follow-up Phase, derjeniger Patienten mit Bevacizumab-Vorbehandlung Bewertung der Immunogenität von Aflibercept zu</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Studienbeginn sowie 30 und 90 Tage nach Einstellung der Einnahme von Aflibercept/Placebo</p> <p>Weitere geringfügige Änderungen (wie z. B. Hinzufügen einer alternativen Methode zur erleichterten Vorbereitung von Aflibercept/Placebo)</p> <p>Änderungen vom 16.11.2009:</p> <p>Auf Anfrage des DMC wurde OS sowie PFS deskriptiv mittels relativen Risiko berechnet (die Analyse erfolgte nach dem Eintritt von 315 Sterbefällen)</p> <p>Weitere geringfügige Änderungen</p> <p>Änderung vom 11.02.2010:</p> <p>Nach Review des DMC: Insgesamt 880 Patienten, die mindestens den ersten Studienzyklus beendet hatten, wurden in die Sicherheitsanalyse aufgenommen</p> <p>Änderung vom 27.04.2011:</p> <p>Verlängerung der Studienteilnahme über das ursprüngliche Studienende (Cutoff) um ca. 9 Monate - für primäre Analyse des OS und für die Registrierungsbehörden</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>Histologisch oder zytologisch bestätigte Diagnose von Adenokarzinom des Dick- oder Enddarms</p> <p>Inoperable Metastasen</p> <p>Messbarer oder nicht-messbarer Tumorbefall (nach RECIST-Kriterien)</p> <p>Rezidiv nach 6 Monaten nach Beendigung genau einer vorangegangenen Oxaliplatin-haltigen Chemotherapie</p> <p>Ausschlusskriterien</p> <p>Vorangegangene Therapie mit Irinotecan</p> <p>Beendigung einer Strahlentherapie, Operation oder Chemotherapie weniger als 28 Tage vor Randomisierung</p> <p>Beendigung einer größeren Operation weniger als 42 Tage vor Randomisierung</p> <p>Unerwünschte Ereignisse durch eine beliebige vorangegangene Krebsbehandlung (Grad >1); Ausnahmen: Alopecia, periphere Neuropathie</p> <p>Alter <18</p> <p>ECOG-PS >2</p> <p>Hirnmetastasen, unkontrollierte Kompression des Rückenmarks, karzinomatöse Meningitis, neue Hinweise auf Hirn- oder leptomeningeale Erkrankungen</p> <p>Andere vorangegangene Malignität, falls nicht >5 Jahre beschwerdefrei</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Teilnahme an anderen klinischen Studien mit Medikation innerhalb von 30 Tagen vor Randomisierung</p> <p>Innerhalb von 6 Monaten vor Randomisierung: Myokardinfarkt, schwere Angina pectoris, Transplantation eines Koronar-/peripheren Arterien-Bypasses, Kongestive Herzinsuffizienz der NYHA Klasse III oder IV, Schlaganfall, vorübergehende Durchblutungsstörung</p> <p>Innerhalb von 3 Monaten vor Randomisierung: gastrointestinale Blutung (Grad 3-4), behandlungsresistentes Magengeschwür, erosive Ösophagitis oder Gastritis, infektiöse oder entzündete Darmerkrankung, Divertikulitis, Lungenembolie, andere unkontrollierte thromboembolische Ereignisse</p> <p>Tiefe Venenthrombose innerhalb 4 Wochen vor Randomisierung</p> <p>Antiretrovirale Behandlung erfordernde AIDS- oder HIV-Erkrankung</p> <p>Schwere akute oder chronische Erkrankung, die die Teilnahme an der Studie oder die Interpretation der Studienergebnisse beeinträchtigen könnten</p> <p>Schwangere oder stillende Frauen; positiver Schwangerschaftstest (Serum oder β-HCG); gebärfähige, in Österreich angeworbene Frauen, die einem monatlichen Schwangerschaftstest nicht zustimmten</p> <p>Fortpflanzungsfähige Patienten (m/w), die einer effektiven Verhütungsmethode während und 6 Monate nach Beendigung der Studie nicht zustimmten</p> <p>Keine schriftliche, informierte Einverständniserklärung</p> <p>Ausschlusskriterien bzgl. Aflibercept</p> <p>Protein-Kreatinin-Ratio im Morgenurin >1 oder Proteinurie >500 mg/24 h</p> <p>Serum-Kreatinin $>1,5 \times$ die obere Grenze des Referenzbereiches (OGR). Bei Kreatininlevel von $1,0$-$1,5 \times$ OGR, führte eine Kreatinin-Clearance <60 ml/min (Cockcroft-Gault Formel) zum Ausschluss des Patienten</p> <p>Anamnestisch bekannter unkontrollierter Bluthochdruck, definiert als Blutdruck $>150/100$ mmHg oder systolischer Blutdruck >180 mmHg bei diastolischem Blutdruck <90 mmHg an mindestens zwei aufeinanderfolgenden Untersuchungen innerhalb 3 Monate vor Patientenaufnahme</p> <p>Patienten in Antikoagulantientherapie mit instabiler Dosis von Warfarin und/oder International Normalized Ratio (INR) >3 innerhalb 4 Wochen vor Studieneintritt</p> <p>Klinisch signifikante Blutungsneigung oder Gerinnungsstörung (INR $>1,5$ ohne Vitamin K-Antagonisten Therapie), nicht-heilende</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Wunde</p> <p>Ausschlusskriterien bzgl. FOLFIRI</p> <p>Bekannter Dihydropyrimidin-Dehydrogenase-Mangel</p> <p>Inadäquate Knochenmarksfunktion:</p> <p>Absolutzahl Neutrophile $<1,5 \times 10^9/l$</p> <p>Anzahl Blutplättchen $<100 \times 10^9/l$</p> <p>Hämoglobin $<9,0$ g/dl</p> <p>Inadäquater Leberfunktionstest:</p> <p>Totales Bilirubin $>1,5 \times OGR$</p> <p>Transaminasen $>3,0 \times OGR$ (bei Lebermetastasen $>5,0 \times OGR$)</p> <p>Alkalische Phosphatase $>3,0 \times OGR$ (bei Lebermetastasen $>5,0 \times OGR$)</p> <p>Prädisponierende Dünn-/Dickdarmerkrankung mit unkontrollierten Symptomen (>3 weicher Stuhl täglich)</p> <p>Vorgeschichte von: chronischer Enteropathie, entzündlicher Enteropathie, chronischer Diarrhoe, unbewältigter Darm-(Teil-)Verschluss, operative Entfernung von mehr als der Hälfte des Dickdarms, umfangreiche Dünndarmresektion mit chronischer Diarrhoe</p> <p>Vorgeschichte von Anaphylaxie oder bekannter Atropin-Sulfat-/Loperamid-/Antiemetika (verabreicht in Verbindung mit FOLFIRI) -Intoleranz</p> <p>Gleichzeitige Behandlung mit Antikonvulsiva (CYP3A4 Induktoren), die nicht >7 Tage eingestellt wurde</p> <p>Patienten mit Gilbert-Syndrom</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>Studie wurde in 178 Zentren in 28 Ländern durchgeführt.</p> <p>Ein Lenkungsausschuss war für die Betreuung der Studie verantwortlich, der nach den Empfehlungen des unabhängigen Komitee für Datenüberwachung (DMC) gehandelt hat.</p> <p>Dieser Ausschuss setzte sich zusammen aus Dr. Carmen Allegra, Vorstandsvorsitzende, University of Florida, Department of Medicine - Hämatologie / Onkologie, USA, Prof. Josep Tabernero (Vall d'Hebron University Hospital, Barcelona, Spanien), Prof. Eric Van Cutsem (Universitätsklinikum Gasthuisberg, Leuven, Belgien), und die Sponsorenvertreter Rémi Castan (Direktor für klinische Studien) sowie Emmanuelle Boelle (führende Biostatistikerin).</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur	Intravenöse Gabe über 1 Stunde von 4 mg/kg Aflibercept in der Konzentration 25 mg/ml alle 2 Wochen oder Placebo;

Item ^a	Charakteristikum	Studieninformation
	Administration etc.	direkt im Anschluss an die Aflibercept- bzw. Placebo-Behandlung: FOLFIRI (di-LV 400 mg/m ² Infusion über 2 Stunden, Irinotecan 180 mg/m ² Infusion über 90 Minuten, 5-FU 400 mg/m ² Bolus, 5-FU 2.400 mg/m ² Infusion über 46 Stunden)
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Ziel: Gesamtüberleben</p> <p>Sekundäre Ziele: Progressionsfreies Überleben; Objektive Ansprechrate (nach RECIST-Kriterien)</p> <p>Sicherheitsziele: Unerwünschte Ereignisse während der Studie Laborwerte, klinische Untersuchungen, Begleitmedikation</p> <p>Immunogenität Anti-Aflibercept-Antikörper-Tests bei allen Patienten während der gesamten Studie</p> <p>Pharmakokinetik: Analyse des endogenen vaskulo-endothelialen Wachstumsfaktors (VEGF) bei allen Patienten, die an einem mit 4°C-Zentrifuge ausgestatteten Zentrum randomisiert wurden</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Annahmen, um eine Differenz zwischen FOLFIRI+Aflibercept und FOLFIRI+Placebo hinsichtlich der Überlebenszeit zu entdecken</p> <p>Annahme einer 20%-igen Risikoreduktion für FOLFIRI+Aflibercept im Vergleich zu Placebo, Power von 90%, α-Fehler (2-seitig)=0,05 → Fallzahl: 600 Patienten pro Gruppe</p> <p>Berechnung basierend auf der O'Brien-Fleming Alpha spending Funktion mit Möglichkeit des Studienabbruchs aufgrund von Wirksamkeit nach 36,5% und 65% der Informationsbeschaffung</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Zwei Interimsanalysen (nach 36,5% und 65% der Informationsbeschaffung)</p> <p>Stratifizierter Log-Rang-Test der Überlebenszeit beider Studienarme; Subgruppenanalysen</p> <p>Nutzlosigkeitsgrenze (erste Zwischenanalyse) basierend auf der Gamma(-5) β-spending function: HR $\geq 1,084$</p> <p>Wirksamkeitsgrenze basierend auf O'Brien-Fleming: $\alpha=0,00042$ (erste Zwischenanalyse) bzw. $\alpha=0,0107$ (zweite Zwischenanalyse)</p>
8	Randomisierung, Erzeugung der	

Item ^a	Charakteristikum	Studieninformation
	Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zufällige Zuteilung mittels Interactive Voice Response System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Permutierte Block-Randomisierung; Stratifizierung nach vorangehender Bevacizumab-Therapie (ja/nein) und ECOG-PS (0/1/2)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nummerierte, identisch aussehende Behandlungssets wurden geliefert.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Zentrale Zuteilung durch IVRS. Patienten, die zum Zeitpunkt der Randomisierung im Follow-up einer doppelblinden, kontrollierten Studie mit Bevacizumab teilnahmen und noch verblindet waren, konnten für diese Studie randomisiert werden.
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?	a) ja b) ja c) ja Randomisierung und identisches Aussehen und Verpackung der Studienmedikation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Verwendung von identisch aussehenden Studienmedikationen und -verpackung für Patienten und medizinisches Personal
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Analysepopulationen: ITT-Population (Intention To Treat): alle Patienten, die die Einverständniserklärung gaben und erfolgreich eine Nummer via IVRS zugewiesen bekommen haben. Primärpopulation für Wirksamkeitsanalysen (außer Ansprechrates). Evaluierbare Population für Ansprechen des Tumors (EP): alle randomisierten Patienten mit messbarem Tumorbefall (IRC) bei Studienbeginn und mindestens einer späteren validen Tumorevaluation. Patienten, die vor der ersten geplanten, bildgebenden Diagnostik durch Krankheitsprogression verstarben oder radiologisch dokumentierte Krankheitsprogression aufwiesen, wurden nicht ausgeschlossen. Patienten mussten <i>third party review</i> zugestimmt haben. Sicherheitspopulation: alle Patienten der ITT-Population, die mindestens eine Dosis der doppelblinden, randomisierten Behandlung (Aflibercept oder Placebo) erhalten haben. Analysen dieser Population basieren auf der

Item ^a	Charakteristikum	Studieninformation
		<p>tatsächlich erhaltenen Therapie. Patienten, die mindestens eine Dosis Afibercept erhalten haben, zählen zum Afibercept-Studienarm.</p> <p>Primäres Zielkriterium: Zweiseitiger ($\alpha=0,0499$) Log-Rang-Test, stratifiziert nach Bevacizumab (ja/nein) und ECOG-PS (0/1/2) zum Vergleich der Überlebenszeit. Hazard Ratios wurden mittels stratifiziertem Cox PH-Modell geschätzt.</p> <p>Sekundäre Zielkriterien: Zweiseitiger ($\alpha=0,0001$), stratifizierter Log-Rang-Test zum Vergleich der progressionsfreien Überlebenszeit zum Zeitpunkt der zweiten Interimsanalyse. Hazard Ratios wurden mittels stratifiziertem Cox PH-Modell geschätzt. Sensitivitätsanalysen wurden durchgeführt.</p> <p>Die Ansprechrate wurde nach positivem Testergebnis der (progressionsfreien) Überlebenszeit getestet. Es wurden deskriptive Analysen durchgeführt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es wurden Subgruppenanalysen für ECOG-PS (0/1/2) und Bevacizumab (ja/nein), Alter (<65, ≥65), Geschlecht (männlich/weiblich), ethnische Zugehörigkeit (kaukasisch & weiß/andere), Lage des primären Tumors (Kolon/Rektum/ Rektosigmoid & andere), Anzahl Organe, die von Metastasen betroffen sind (>1, ≤1), nur Lebermetastasen (ja/nein), vorangegangener Bluthochdruck (ja/nein) und geographische Region (Westeuropa/Osteuropa/Nordamerika/ Südamerika) durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) FOLFIRI+Placebo: 614, FOLFIRI+Afibercept: 612 b) FOLFIRI+Placebo: 609, FOLFIRI+Afibercept: 607 c) ITT: FOLFIRI+Placebo: 614, FOLFIRI+Afibercept: 612
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flowchart (Abbildung 14)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Erster Patient eingeschrieben: 19. November 2007 Letzter Patient eingeschrieben: 16. März 2010 Studienende: 07. Februar 2011
14b	Informationen, warum die Studie endete	Die Studie endete regulär nach 863 Todesfällen.

Item ^a	Charakteristikum	Studieninformation
	oder beendet wurde	
a: nach CONSORT 2010.		

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

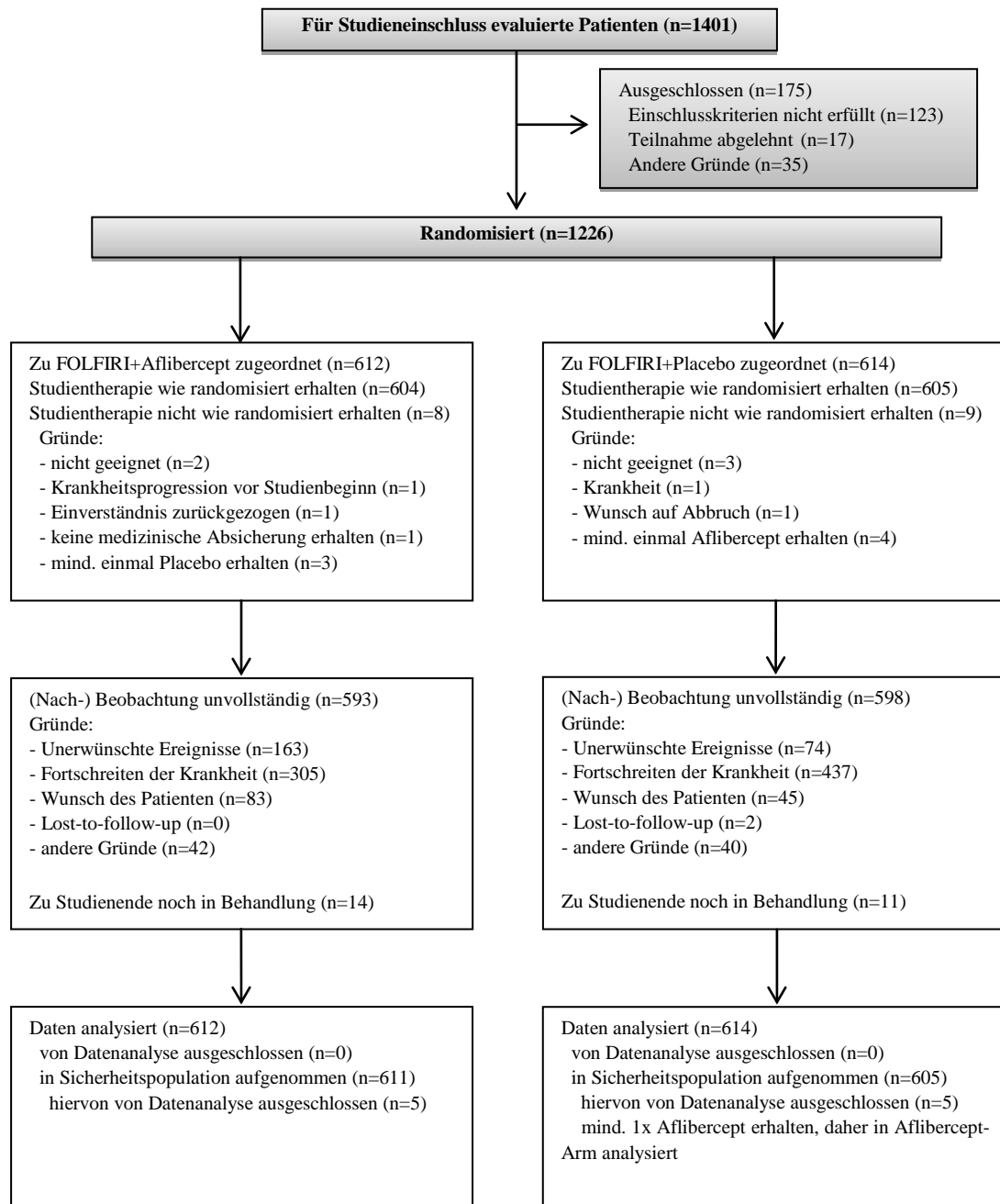


Abbildung 14: Patientenfluss in der Studie VELOUR (Quelle: [33]).

Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

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

Tabelle 4-59 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie VELOUR

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

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

Der Bogen gliedert sich in zwei Teile:

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

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

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

Studie: VELOUR

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Studienbericht VELOUR	CSR
Volltextpublikation	Van Cutsem_2012

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

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

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

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

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie (CSR)

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 kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

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

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

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

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

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

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

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

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es handelt sich um ein doppelblindes Studiendesign (CSR)

Behandler bzw. Weiterbehandler

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

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es handelt sich um ein doppelblindes Studiendesign (CSR)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

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

Beispiele zu a und b:

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

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

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

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

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

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

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

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

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

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

Endpunkt: Gesamtüberleben (OS)

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
Der Endpunkterheber war verblindet (CSR)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

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

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

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Der Endpunkterheber war verblindet (CSR)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

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

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

- ja**
- nein**

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Ansprechrate

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
Der Endpunkterheber war verblindet (CSR)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

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

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Sicherheit

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
Der Endpunkterheber war verblindet (CSR)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
