

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Regorafenib (Stivarga®)

Bayer Vital GmbH

Modul 4 A

Behandlung von erwachsenen Patienten mit metastasiertem Kolorektalkarzinom, die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind; diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	5
Abbildungsverzeichnis	9
Abkürzungsverzeichnis	10
4 Modul 4 – allgemeine Informationen	15
4.1 Zusammenfassung der Inhalte von Modul 4.....	16
4.2 Methodik.....	37
4.2.1 Fragestellung.....	37
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	39
4.2.3 Informationsbeschaffung.....	40
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	41
4.2.3.2 Bibliografische Literaturrecherche.....	41
4.2.3.3 Suche in Studienregistern.....	42
4.2.3.4 Selektion relevanter Studien.....	42
4.2.4 Bewertung der Aussagekraft der Nachweise.....	43
4.2.5 Informationssynthese und -analyse.....	45
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	45
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	46
4.2.5.3 Meta-Analysen.....	53
4.2.5.4 Sensitivitätsanalysen.....	54
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	55
4.2.5.6 Indirekte Vergleiche.....	57
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	58
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	59
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	59
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	59
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	60
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	62
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	62
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	63
4.3.1.2.1 Studiendesign und Studienpopulationen.....	63
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	70
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	72
4.3.1.3.1 Endpunkt Mortalität: Gesamtüberleben (OS) – RCT.....	73
4.3.1.3.2 Endpunkt Mortalität/Morbidität: Progressionsfreies Überleben (PFS) – RCT.....	78
4.3.1.3.3 Endpunkt Morbidität: Objektive Tumoransprechrte (ORR) – RCT.....	84
4.3.1.3.4 Endpunkt Morbidität: Krankheitskontrollrate (DCR) – RCT.....	86
4.3.1.3.5 Endpunkt Morbidität: Dauer der Tumorstabilisierung – RCT.....	90

4.3.1.3.6	Endpunkt Gesundheitsbezogene Lebensqualität anhand EORTC QLQ C30 – RCT	93
4.3.1.3.7	Endpunkt Verträglichkeit: Unerwünschte Ereignisse – RCT	99
4.3.1.3.8	Subgruppenanalysen – RCT.....	113
4.3.1.3.9	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	134
4.3.2	Weitere Unterlagen.....	138
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	138
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	138
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	139
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	139
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	139
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	141
4.3.2.2	Nicht randomisierte vergleichende Studien	141
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	141
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	142
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	142
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	142
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	143
4.3.2.3	Weitere Untersuchungen.....	143
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	143
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	144
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	144
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	144
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	145
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	145
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	145
4.4.1	Beurteilung der Aussagekraft der Nachweise	145
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	148
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	151
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	152
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	153
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	153
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	153
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	153
4.5.4	Verwendung von Surrogatendpunkten	153
4.6	Liste der eingeschlossenen Studien.....	155
4.7	Referenzliste.....	155
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		160
Anhang 4-B : : Suchstrategien – Suche in Studienregistern		162

Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund	173
Anhang 4-D : Liste der abgebrochenen Studien	174
Anhang 4-E : Liste der laufenden Studien.....	175
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....	176
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	191

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Einschlusskriterien für Studien zur Bewertung des medizinischen Nutzens und Zusatznutzens von Regorafenib	18
Tabelle 4-2: Berücksichtigte Endpunkte bei der Bewertung des Zusatznutzens von Regorafenib im Anwendungsgebiet – Patientenrelevanz und Validität.....	20
Tabelle 4-3: Prospektiv geplante Subgruppenanalysen der Studienpopulation (CORRECT).....	21
Tabelle 4-4: Wesentliche Ein- und Ausschlusskriterien der eingeschlossenen Studie 14387 CORRECT	25
Tabelle 4-5: Zusammenfassung der Ergebnisse der patientenrelevanten Endpunkte	27
Tabelle 4-6: Zusammenfassende Bewertung des medizinischen Zusatznutzens von Regorafenib gegenüber BSC.....	35
Tabelle 4-7: Prädefinierte Einschlusskriterien für Studien zur Bewertung des medizinischen Nutzens und Zusatznutzens von Regorafenib	40
Tabelle 4-8: Für das Gesamtüberleben prädiktive Schwellenwerte für Effektgrößen von Veränderungen in Wirksamkeitsparametern, nach Stichprobengröße (nach Johnson et al.)... 51	51
Tabelle 4-9: Korrelation zwischen Surrogatendpunkten und Gesamtüberleben (nach Tange et al. (2009))	52
Tabelle 4-10: In der CORRECT-Studie geplante Subgruppenanalysen – RCT mit dem zu bewertenden Arzneimittel	56
Tabelle 4-11: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	59
Tabelle 4-12: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	60
Tabelle 4-13: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-14: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	63
Tabelle 4-15: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-16: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	66
Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	73
Tabelle 4-20: Operationalisierung von Gesamtüberleben.....	74

Tabelle 4-21: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel.....	74
Tabelle 4-22: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel (21. Juli 2012)	75
Tabelle 4-23: Sensitivitätsanalyse des Gesamtüberlebens (ITT)	77
Tabelle 4-24: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel (deskriptive Zusatzanalyse).....	78
Tabelle 4-25: Operationalisierung von Progressionsfreiem Überleben	80
Tabelle 4-26: Bewertung des Verzerrungspotenzials für progressionsfreies Überleben in RCT mit dem zu bewertenden Arzneimittel	80
Tabelle 4-27: Ergebnisse für progressionsfreies Überleben aus RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-28: Sensitivitätsanalyse des progressionsfreien Überlebens (ITT) aus RCT mit dem zu bewertenden Arzneimittel.....	83
Tabelle 4-29: Operationalisierung von Objektive Tumoransprechrates.....	85
Tabelle 4-30: Bewertung des Verzerrungspotenzials für Objektive Tumoransprechrates in RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4-31: Ergebnisse für Objektive Tumoransprechrates aus RCT mit dem zu bewertenden Arzneimittel	86
Tabelle 4-32: Operationalisierung von Krankheitskontrollrate	88
Tabelle 4-33: Bewertung des Verzerrungspotenzials für die Krankheitskontrollrate in RCT mit dem zu bewertenden Arzneimittel	88
Tabelle 4-34: Ergebnisse für Krankheitskontrollrate aus RCT mit dem zu bewertenden Arzneimittel.....	89
Tabelle 4-35: Operationalisierung von Dauer der Tumorstabilisierung	91
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Dauer der Tumorstabilisierung in RCT mit dem zu bewertenden Arzneimittel	91
Tabelle 4-37: Ergebnisse für Dauer der Tumorstabilisierung aus RCT mit dem zu bewertenden Arzneimittel	92
Tabelle 4-38: Operationalisierung von Gesundheitsbezogener Lebensqualität.....	93
Tabelle 4-39: Bewertung des Verzerrungspotenzials für gesundheitsbezogene Lebensqualität anhand EORTC QLQ-C30 in RCT mit dem zu bewertenden Arzneimittel....	95
Tabelle 4-40: Ergebnisse für Gesundheitsbezogene Lebensqualität anhand EORTC QLQ-C30 - RCT mit dem zu bewertenden Arzneimittel – Veränderungen im EORTC QLQ-C30 von Baseline an Zyklen 2, 3, 4 und Ende der Therapie (ITT).....	96
Tabelle 4-41: Ergebnisse für Behandlungseffekte in der Gesundheitsbezogenen Lebensqualität anhand EORTC QLQ-Gesamtskalen der Zeit-adjustierten Fläche unter der Kurve (ITT).....	99
Tabelle 4-42: Operationalisierung von unerwünschten Ereignissen.....	100

Tabelle 4-43: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse mit dem zu bewertenden Arzneimittel	101
Tabelle 4-44: Ergebnisse für unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel mit abgeleitetem relativem Risiko	102
Tabelle 4-45: Überblick über UE nach CTCAE-Grad und Zusammenfassung von UE-Kategorien in der CORRECT-Studie in RCT mit dem zu bewertenden Arzneimittel und abgeleiteter Odds Ratio	103
Tabelle 4-46: Studie 14387 (CORRECT) - Inzidenzen von unerwünschten Ereignissen (alle Grade), die in mdst. 10 % aller Patienten in einer der Behandlungsgruppen auftraten (SAF).....	105
Tabelle 4-47: Inzidenzen von UE von CTCAE-Grad 3 und 4, die in mdst. 2 % aller Patienten in einer der Behandlungsgruppen in RCT mit dem zu bewertenden Arzneimittel auftraten.....	107
Tabelle 4-48: Inzidenzen von UE, die zum Tode führten, (Grad 5), die bei mdst. einem Patienten in RCT mit dem zu bewertenden Arzneimittel auftraten	109
Tabelle 4-49: Subgruppenanalyse für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel.....	113
Tabelle 4-50: Subgruppenanalyse für PFS in RCT mit dem zu bewertenden Arzneimittel..	116
Tabelle 4-51: Subgruppenanalyse für Gesamtüberleben zum Zeitpunkt des Cross-overs (<i>data cut-off</i> 13.11.2011).....	119
Tabelle 4-52: Subgruppenanalyse für unerwünschte Ereignisse (UE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 - SAF-Population	122
Tabelle 4-53: Subgruppenanalyse für schwerwiegende unerwünschte Ereignisse (SUE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 - SAF-Population.....	125
Tabelle 4-54: Subgruppenanalyse für unerwünschte Ereignisse (UE), die zur dauerhaften beendigung der Studienmedikation führten – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF Population.....	128
Tabelle 4-55: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	139
Tabelle 4-56: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	139
Tabelle 4-57: Operationalisierung von <Endpunkt xxx>.....	140
Tabelle 4-58: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	140
Tabelle 4-59: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	140
Tabelle 4-60: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	142
Tabelle 4-61: Operationalisierung von <Endpunkt xxx>.....	142
Tabelle 4-62: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	143

Tabelle 4-63: Operationalisierung von <Endpunkt xxx>.....	144
Tabelle 4-64: Regorafenib + BSC vs. Placebo + BSC - Ausmaß des Zusatznutzens auf Endpunktebene	149
Tabelle 4-65: Begleitende Resultate zur Gesamtaussage zum Zusatznutzen	151
Tabelle 4-66: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	152
Tabelle 4-67: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	153
Tabelle 4-68 (Anhang): Abgeschlossene, laufende und abgebrochene Studien aus der Studienregisterrecherche und Kriterien für Studienausschluss	164
Tabelle 4-69 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	174
Tabelle 4-70 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	175
Tabelle 4-71 (Anhang): Studiendesign und -methodik für Studie CORRECT.....	177
Tabelle 4-72 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 14387 CORRECT	192

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Kaplan-Meier-Kurve für das Gesamtüberleben in der Studie 14387 CORRECT (ITT-Population).....	29
Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	61
Abbildung 4-3: Kaplan-Meier-Kurve für das mediane Gesamtüberleben (CORRECT; ITT-Population).....	76
Abbildung 4-4: Kaplan-Meier-Kurve für das mediane progressionsfreie Überleben (CORRECT; ITT-Population).....	82
Abbildung 4-5: Forest-Plot der Subgruppenanalyse für Gesamtüberleben (ITT) in RCT mit dem zu bewertenden Arzneimittel	115
Abbildung 4-6: Forest-Plot der Subgruppenanalyse für PFS (ITT) in RCT mit dem zu bewertenden Arzneimittel	118
Abbildung 4-7: Subgruppenanalyse für Gesamtüberleben zum Zeitpunkt des Cross-overs (data cut-off 13.11.2011) – Darstellung der Ergebnisse im Forest-Plot	121
Abbildung 4-8: Subgruppenanalyse für unerwünschte Ereignisse (UE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF-Population, Forest- Plot-Darstellung mit logarithmischer Darstellung der x-Achse	124
Abbildung 4-9: Subgruppenanalyse für schwerwiegende unerwünschte Ereignisse (SUE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF- Population, Forest-Plot-Darstellung mit logarithmischer Darstellung der x-Achse	127
Abbildung 4-10: Subgruppenanalyse für unerwünschte Ereignisse, die zur dauerhaften Beendigung der Studienmedikation führten – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF-Population, Forest-Plot- Darstellung mit logarithmischer Darstellung der x-Achse.....	130
Abbildung 4-11: Patientenfluss CORRECT Studie Datenauswertungspunkt 1.....	190

Abkürzungsverzeichnis

Abkürzung	Bedeutung
5-FU	5-Fluorouracil
Abs.	Absatz
AKT	= PKB (Proteinkinase B)
ALP	Alkalische Phosphatase
ALT	Alaninaminotransferase
ARDS	Acute respiratory distress syndrome
AST	Aspartataminotransferase
AUC	Area under the curve (Fläche unter der Kurve)
BRAF	B-Rapidly Accelerated Fibrosarcoma-Protein
BSC	Best Supportive Care
bspw.	beispielsweise
bzw.	beziehungsweise
CONSORT	Consolidated standards of reporting trials
CORRECT	Colorectal cancer treated with Regorafenib or placebo after failure of standard therapy (Phase-III-Studie)
CR	Complete response (Komplette Remission)
CRF	Case report form
CRO	Contract research organisation (klinisches Forschungszentrum)
CTC	Common toxicity criteria
CTCAE	Common terminology criteria for adverse event (geläufige Terminologiekriterien für unerwünschte Ereignisse)
d.h.	das heißt
DCR	Disease control rate (Krankheitskontrollrate)
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
dl	Deziliter
DNA	Deoxyribonucleic acid (Desoxyribonukleinsäure)
EbM	Evidenzbasierte Medizin
ECOG	Eastern Cooperative Oncology Group
EdT	Ende der Therapie
EG	Europäische Gemeinschaft
EGF	Epidermal growth factor

EGRF	Epidermal growth factor receptor
EKG	Elektrokardiogramm
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D	European Quality of Life Group Questionnaire 5D
ERK	Extracellular signal-regulated kinase
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FOIB	Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab
FOIBE	Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper
PPFV	First patient first visit
g	Gramm
G-BA	Gemeinsamer Bundesausschuss
GFR	Glomeruläre Filtrationsrate
ggf.	gegebenenfalls
GI	Gastro-intestinal
HIV	Humanes Immundefizienz-Virus
HR	Hazard Ratio
HRQoL	Health-related quality of life (gesundheitsbezogene Lebensqualität)
ICH	International Conference of Harmonisation
IDMC	Independent Data Monitoring Committee
i. d. R.	in der Regel
inkl.	inklusive
INR	International normalized ratio
i. S.	im Sinne
ITT	Intention to treat
IVRS	Interactive voice response system (Sprachdialogsystem)
Kap.	Kapitel
KI	Konfidenzintervall
KM	Kaplan-Meier
KRAS	Kirsten rat sarcoma viral oncogene homolog
LPLV	Last patient last visit
mCRC	Metastatic colorectal cancer (metastasierter Kolorektalkarzinom)

MDRD	Modification of diet in renal disease
MEK	Mitogen-activated protein kinase / extracellular signal-regulated kinase
mdst.	mindestens
min	Minute(n)
mg	Milligramm
ml	Milliliter
mm	Millimeter
MTC	Mixed treatment comparison
mTOR	Mammalian target of Rapamycin (Ziel des Rapamycins im Säugetier)
MUGA	Multi gated acquisition
n	Anzahl
NCI	National Cancer Institute
Nr.	Nummer
NYHA	New York Heart Association
o.g.	oben genannt
ONW	Oberer Normalwert
OR	Odds ratio
ORR	Objective response rate (Objektive Tumoransprechrage)
OS	Overall survival (Gesamtüberleben)
p	Überschreitungswahrscheinlichkeit
PD	Progressive disease (progressive Erkrankung)
PDGF	Platelet-derived growth factor
PDGFR	Platelet-derived growth factor receptor
PFS	Progression-free survival (progressionsfreies Überleben)
PI3K	Phosphoinositid-3-Kinasen
PKB	Protein-Kinase B
Pla	Placebo
PR	Partial response (partielle Remission)
PTEN	Phosphatase and Tensin homolog
PTT	Partielle Thromboplastinzeit
QLQ-C30	Core quality of life questionnaire C30
QOL-CS	Cancer-specific health-related quality of life questionnaires

RAF	Rapidly Accelerated Fibrosarcoma-Protein
RAS	Rat Sarcoma-Protoonkogen
RCT	Randomized controlled trial
RECIST	Response evaluation criteria in solid tumors
Reg	Regorafenib
RR	Relatives Risiko
SAF	Safety population (Sicherheitspopulation)
SAP	Statistical analysis plan (statistischer Analyseplan)
SD	Stable disease (Krankheitsstabilisierung)
SF-36	Short-Form health survey mit 36 Fragen
SGB	Sozialgesetzbuch
STD	Standard deviation (Standardabweichung)
STE	Surrogate threshold effects
STROBE	Strengthening the reporting of observational studies in epidemiology
SUAW	Schwerwiegende unerwünschte Arzneimittelwirkung
SUE	Schwerwiegende(s) unerwünschte(s) Ereignis(se)
SUSAR	Suspected unexpected serious adverse reaction
Tabl.	Tablette(n)
TREND	Transparent reporting of evaluations with non-randomized design
TTP	Time to progression (Zeit bis zur Progression)
u. a.	unter anderem, unter anderen
UAW	Unerwünschte Arzneimittelwirkung
UE	Unerwünschte(s) Ereignis(se)
usw.	und so weiter
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
VerfO-GBA	Verfahrensordnung des Gemeinsamen Bundesausschusses
vgl.	vergleiche
vs.	versus
VWF	Von-Willebrand-Faktor
WHO	World Health Organization
WT	Wildtyp
z. B.	zum Beispiel

Z. n.	Zustand nach
-------	--------------

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

Es wurde untersucht, ob die Behandlung mit dem Multikinase-Inhibitor Regorafenib im Anwendungsgebiet zu einem Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie im Sinne des § 35a SGB V führt. Das Anwendungsgebiet bzw. die Zielpopulation beinhaltet dabei „Patienten mit metastasiertem Kolorektalkarzinom (KRK), die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie.“ Der Nachweis des medizinischen Zusatznutzens erfolgt auf Basis randomisierter, kontrollierter klinischer Studien (RCT), die die Endpunkte Mortalität (Gesamtüberleben), Mortalität/Morbidität (progressionsfreies Überleben), Morbidität (Objektive Tumoransprechrage, Krankheitskontrollrate, Dauer der Tumorstabilisierung), gesundheitsbezogene Lebensqualität sowie Verträglichkeit untersuchten.

Die Bewertung des Zusatznutzens von Regorafenib erfolgt im Vergleich zu Best-Supportive-Care (BSC). BSC wurde am 08.08.2012 im Beratungsgespräch vom G-BA als zweckmäßige Vergleichstherapie bestimmt (1). Unter „Best-Supportive-Care“ wird die Therapie verstanden, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirkt.

Um nachzuweisen, dass die Behandlung mit Regorafenib einen medizinischen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie hat, werden folgende Endpunkte berücksichtigt:

- Wichtigster patientenrelevanter Endpunkt in Bezug auf die **Mortalität** ist die **Verlängerung des Gesamtüberlebens**, gemessen als Zeitraum von der Randomisierung bis zum Tod (nicht krankheitsspezifisch).
- Eng verknüpft mit diesem Endpunkt sind der Endpunkt zur **Mortalität/Morbidität progressionsfreies Überleben** bzw. die Endpunkte zur **Morbidität Objektive Tumoransprechrage, Krankheitskontrollrate** und **Dauer der Tumorstabilisierung**.
- Daten zur **gesundheitsbezogenen Lebensqualität** während der Behandlung werden ebenfalls berücksichtigt.
- Zusätzlich wird der Parameter **Verträglichkeit**, beurteilt anhand der Rate an unerwünschten Ereignissen während der Behandlung, in die Beurteilung eines medizinischen Zusatznutzens mit einbezogen.

Datenquellen

Die Bayer Vital GmbH listet gemäß den Kriterien der Verfahrensordnung des Gemeinsamen Bundesausschusses (VerfO-GBA) alle Studien mit Regorafenib im Anwendungsgebiet auf, die an die Zulassungsbehörden übermittelt wurden. Es wurden dabei alle randomisierten Studien mit Regorafenib in der Indikation metastasiertes kolorektales Karzinom berücksichtigt (siehe Tabelle 4-11).

Des Weiteren wurde eine systematische Recherche in Studienregistern (*clinicaltrials.gov*, *EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu)*, *Klinische Prüfungen Pharmnet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm)* sowie *International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http://apps.who.int/trialsearch/)* am 13. August 2013 durchgeführt (siehe Tabelle 4-14 und Anhang 4-B).

Da eine Zulassung von Regorafenib weltweit noch nicht länger als zwölf Monate besteht und keine indirekten Vergleiche herangezogen werden, ist nach den Kriterien der VerfO-GBA eine bibliographische Literaturrecherche in Literaturdatenbanken nicht notwendig und wurde aus diesem Grund auch nicht durchgeführt.

Ein-/Ausschlusskriterien für Studien

Kriterien für den Ein- und Ausschluss wurden entsprechend der Angabe zum Modul 4.2.2 der VerfO-GBA gewählt und in Bezug auf das zugelassene Anwendungsgebiet präzisiert. Die Studienpopulation beinhaltet gemäß dem zugelassenen Anwendungsgebiet erwachsene Patienten mit metastasiertem Kolorektalkarzinom (KRK), die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie. Als Intervention wird Regorafenib 160 mg einmal täglich für drei Wochen mit anschließend einer Woche Pause definiert.

Mindestens einer der patientenrelevanten Endpunkte Mortalität (Gesamtüberleben), Morbidität/Mortalität (progressionsfreies Überleben), Morbidität (Objektive Tumorsprechrage, Krankheitskontrollrate, Dauer der Tumorstabilisierung), gesundheitsbezogene Lebensqualität (HRQoL) oder therapiebedingte Morbidität (Verträglichkeit) sollte untersucht worden sein.

Eine weitere Begrenzung erfolgte auf abgeschlossene randomisierte, kontrollierte Studien (RCT). Vergleichstherapie ist BSC, d. h. Therapieoptionen, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirken. Tabelle 4-1 fasst diese prädefinierten Einschlusskriterien zusammen.

Tabelle 4-1: Prädefinierte Einschlusskriterien für Studien zur Bewertung des medizinischen Nutzens und Zusatznutzens von Regorafenib

Kürzel	Kategorie	Prädefinierte Kriterien (Einschlusskriterien, sofern nicht anders vermerkt)
E1	Patientenpopulation	<ul style="list-style-type: none"> • Patienten mit metastasiertem kolorektalen Karzinom • O. g. Patienten, die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie.
E2	Intervention	Regorafenib 160 mg einmal täglich für drei Wochen mit anschließend einer Woche Pause
E3	Endpunkte	<p>Mindestens einer der patientenrelevanten Endpunkte wird berichtet:</p> <ul style="list-style-type: none"> • Mortalität: Gesamtüberleben (OS) • Mortalität/Morbidität: progressionsfreies Überleben (PFS) • Morbidität: Objektive Tumoransprechrage (ORR), Krankheitskontrollrate (DCR), Dauer der Tumorstabilisierung • Gesundheitsbezogene Lebensqualität (HRQoL) • Therapiebedingte Morbidität/Verträglichkeit: Rate an unerwünschten Ereignissen (UE)
E4	Studientyp	Abgeschlossene randomisierte, kontrollierte Studien (RCT) Phase II, III oder IV
E5	Vergleichsintervention	BSC, d. h. Therapieoptionen, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirken
E6	Publikationstyp	keine Einschränkung
E7	Studiendauer	keine Einschränkung
E8	Sprache	keine Einschränkung

Die Selektion der relevanten Studien aus der Auflistung der von BAYER gesponsorten und in Auftrag gegebenen Studien und der Suche in Studienregistern wurde von zwei Reviewern unabhängig voneinander durchgeführt (für den resultierenden Studienpool siehe 4.3.1.1.4).

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

Gemäß §5 Abs. 5 Satz 1 5. Kap. VerfO-GBA ist ein Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie dann festgestellt, wenn die Verbesserung patientenrelevanter Endpunkte gegenüber der zweckmäßigen Vergleichstherapie nachgewiesen werden kann (2). Vorrangig sind für den Nachweis randomisierte, verblindete und kontrollierte direkte Vergleichsstudien nach internationalem Standard der evidenzbasierten Medizin (EbM) zu berücksichtigen, die für die Behandlungssituation in der klinischen Routine repräsentativ und relevant sind (§ 5 Abs. 5 Satz 2 5. Kap. VerfO-GBA) (2).

Darstellung des Designs und der Methodik: Die als relevant identifizierten Studien werden gemäß der *items* 1 bis 14 des CONSORT-Statements beurteilt (3). Dabei werden sowohl die Checkliste des CONSORT-Statements als auch die weiteren Ausführungen zu einzelnen *items* zur besseren Beurteilung der internen und externen Validität der Studie herangezogen.

Gegenüberstellung der Ergebnisse: Die Charakteristika der Studienpopulation sowie die Ergebnisse aller berichteten patientenrelevanten Endpunkte werden dargestellt. Für die patientenrelevanten Endpunkte werden Angaben zur Validität gemacht. In diesem Dossier müssen die Morbiditätsendpunkte nicht als Surrogat für das Gesamtüberleben validiert werden, da das Gesamtüberleben der primäre Endpunkt der Studie war und somit bereits enthalten ist. Die Morbiditätsendpunkte werden in dieser Indikation als separate valide patientenrelevante Endpunkte angesehen.

Als patientenrelevant im Anwendungsgebiet wurden folgende Endpunkte identifiziert und in der Darstellung der Ergebnisse berücksichtigt:

Tabelle 4-2: Berücksichtigte Endpunkte bei der Bewertung des Zusatznutzens von Regorafenib im Anwendungsgebiet – Patientenrelevanz und Validität

Endpunkt	Patientenrelevanz, Validität
Mortalität: medianes Gesamtüberleben (OS)	Das mediane Gesamtüberleben ist ein akzeptierter patientenrelevanter Endpunkt.
Morbidität/Mortalität: progressionsfreies Überleben (PFS)	Das Überleben ohne Fortschreiten der Erkrankung, progressionsfreies Überleben, ist für die Indikation kolorektales Karzinom ein anerkannter eigenständiger patientenrelevanter Endpunkt.
Morbidität: Objektive Tumoransprechrates (ORR)	Die Objektive Tumoransprechrates als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum wird bestimmt anhand der Summe von Patienten, die eine komplette (CR) oder partielle Remission (PR) unter der therapeutischen Intervention erfahren (ORR = CR+PR).
Morbidität: Krankheitskontroll- rate (DCR)	Die Krankheitskontrollrate als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum wird bestimmt anhand der Summe von Patienten, die eine komplette (CR) oder partielle (PR) Remission oder eine Stabilisierung der Erkrankung (SD) unter der therapeutischen Intervention erfahren (DCR = CR + PR + SD).
Morbidität: Dauer der Tumorstabilisierung	Die Dauer der Tumorstabilisierung als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum über die Zeit ist definiert als Zeit (in Tagen) von der Randomisierung bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression auftrat). Einbezogen wurden alle Patienten mit einer stabilen Erkrankung (SD) über mindestens 6 Wochen, sowie der Wertung als non-CR/non-PR, aber nicht solche mit einer Progression der Erkrankung (PD).
Gesundheitsbezogene Lebensqualität (HRQoL)	Die Lebensqualität ist ein akzeptierter patientenrelevanter Endpunkt; die eingesetzten Fragebögen werden hinsichtlich ihrer Validität beurteilt. Nur validierte Fragebögen (EORTC QLQ-C30) werden zur Bewertung des Zusatznutzens herangezogen.
Verträglichkeit: unerwünschte Ereignisse (UE)	Unerwünschte Ereignisse reflektieren vor allem in Form der therapieassoziierten unerwünschten Ereignisse die Verträglichkeit der Therapie. Die Schwere und Häufigkeit sind patientenrelevant. Die Erfassung der unerwünschten Ereignisse in der vorliegenden Studie folgt den internationalen Standards und ist somit validiert.

Bewertung der Aussagekraft: Ausschlaggebend für eine Bewertung der Studien bezüglich der unverzerrten Aussagekraft sind die angewandten Methoden zur Fallzahlbestimmung, die

Art und Weise der Randomisierung, insbesondere die Generierung der Randomisierungssequenz und die verdeckte Zuteilung der Patienten zu den Gruppen, die Verblindung und die adäquate Anwendung von statistischen Analysemethoden. Weiterhin werden die Darstellung der Ergebnisse, die Vollständigkeit berichteter Ergebnisse und der Patientenfluss zur Beurteilung eines potentiellen Verzerrungspotentials herangezogen.

Subgruppenmerkmale und -effekte: Für die zur Bewertung identifizierten Studien wird geprüft, inwiefern die dargestellten Patientencharakteristika für die Studienpopulation als relevant zu betrachten sind und inwiefern diese mit den in den Studien a priori gewählten Subgruppenkategorien übereinstimmen. Für die zu untersuchende Patientenpopulation werden folgende prospektiv geplanten Subgruppenmerkmale betrachtet:

Tabelle 4-3: Prospektiv geplante Subgruppenanalysen der Studienpopulation (CORRECT)

Subgruppe	Subgruppenlevel
Geschlecht	Männlich, weiblich
Alter (Jahre)	< 65, ≥ 65
Ethnische Zugehörigkeit	Kaukasier, Asiatisch
ECOG bei Einschluss	0, 1
Zeit seit der Diagnose der metastasierten Erkrankung	< 18 Monate, ≥ 18 Monate
Region ^a	1, 2, 3
Vorherige Anti-VEGF-Therapie	Ja, nein
Vorherige antitumoröse Therapie	FOIB, FOIBE
Anzahl vorheriger Therapielinien	≤ 3, > 3
Anzahl vorheriger Therapielinien nach Diagnose der metastasierten Erkrankung	≤ 3, > 3
KRAS Mutation	Nein, ja
Primärer Ort der Erkrankung	Kolon, Rektum, Kolon und Rektum
^a Region 1: Nordamerika, Westeuropa, Israel, Australien; Region 2: Asien; Region 3: Südamerika, Türkei, Osteuropa FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper	

Sensitivitätsanalysen: Für die identifizierten Studien wird geprüft, welche Art von Sensitivitätsanalysen geplant und welche Sensitivitätsanalysen post hoc durchgeführt wurden. Die für die patientenrelevanten Endpunkte durchgeführten Sensitivitätsanalysen werden im Dossier unter Beschreibung der Methoden angegeben.

Biomarkeranalysen: Für die identifizierten Studien wird geprüft, welche Art von Biomarkeranalysen durchgeführt wurden, die mit in die Bewertung des Zusatznutzens des Arzneimittels einbezogen werden können.

Patientenrelevanz von PFS und Parametern der Tumorkontrolle (Objektive Tumoransprechrates, Krankheitskontrollrate, Dauer der Tumorstabilisierung)

Vorbemerkung: In diesem Dossier müssen die Morbiditätsendpunkte nicht als Surrogat für das Gesamtüberleben validiert werden, da das Gesamtüberleben der primäre Endpunkt der Studie war und somit bereits enthalten ist. Vielmehr werden die Morbiditätsendpunkte in dieser Indikation als separate valide patientenrelevante Endpunkte gesehen. Sie unterstützen und bestätigen den signifikanten und klinisch relevanten Vorteil beim Gesamtüberleben. Eine ausführliche Darstellung der Patientenrelevanz dieser Endpunkte erfolgt in Kapitel 4.2.5.2.

Eine Therapie, die effektiv und über eine relevante Zeit eine Tumorstabilisierung verhindert, ohne einen negativen Effekt auf das Überleben durch eine intrinsische Toxizität zu haben, mündet in der Regel auch in einen Überlebensvorteil. Genauso sollte eine Verlängerung der Tumorkontrolle ohne zusätzliche Toxizität zu einer Verbesserung der krebisbedingten Symptome und der Lebensqualität führen (4, 5). Es ist daher nachzuvollziehen, dass eine länger andauernde Kontrolle des Tumorstabilisierungs zu positiven Effekten auf das physische und psychische Wohlbefinden des Patienten führt und auch im Interesse des Gesundheitswesens liegt. Dies wird bestätigt durch Studien, die belegen, dass der Status der Progressionsfreiheit ein Wert für die Patienten an sich darstellt, der mit einer höheren Lebensqualität und einem besseren Gesundheitszustand einhergeht (6-8).

Dass PFS einen wertvollen Endpunkt für klinische Studien mit antitumorös wirkenden Substanzen darstellt, haben auch die meisten Arzneimittel-Zulassungsbehörden anerkannt und entsprechende Richtlinien zu dessen Verwendung herausgegeben (9-11). Mehrere Arzneimittel sind auf der Basis von signifikanten Verbesserungen im PFS zugelassen worden. Beispiele sind die Zulassung von Sorafenib und Pazopanib beim metastasierten Nierenzellkarzinom, und Gemcitabin in Kombination mit Carboplatin zur Therapie des fortgeschrittenen Ovarialkarzinoms. Auch wenn die Erhebung von PFS mit einigen Nachteilen verbunden ist, haben insbesondere in Indikationen, in denen es sequentielle Therapien gibt, die dadurch erreichten stufenweisen Verbesserungen in der Tumorkontrolle dazu geführt, dass die Krebserkrankung über längere Zeit kontrollierbar wird.

Der Parameter Progressionsfreies Überleben misst direkt die Effekte der jeweiligen Behandlung auf das Tumorstabilisierungs unter Benutzung einer objektiven und quantitativen Methodik. Dabei werden sowohl zytostatische (Tumorstabilisierung) als auch zytotoxische (Tumorrückgang) Effekte gemessen. Eng zusammen mit dem progressionsfreien Überlebensstatus steht das Tumorstabilisieren und die Tumorstabilisierung bzw. das Nicht-Fortschreiten der Erkrankung, auf dem das Kriterium „PFS“ zum Teil basiert. Daher sollte im Kontext einer positiven Risiko-Nutzen-Abschätzung eine klinisch relevante Verbesserung des PFS sowie von Parametern der Tumorkontrolle (Objektive Tumoransprechrates,

Krankheitskontrollrate, Dauer der Tumorstabilisierung) patientenrelevante Nutzenparameter mit einem eigenen Stellenwert darstellen.

Viele Patienten und ihre Behandler beurteilen es als Vorteil, wenn eine schnelle und effektive Tumorkontrolle erfolgt und eine Progression des Tumors verzögert werden kann, und zwar unabhängig von einer etwaigen nachgewiesenen Verlängerung des Gesamtüberlebens. (4, 12, 13) Dies trifft insbesondere auf schnell fortschreitende Tumoren oder solche zu, die ein lebenswichtiges Organ bedrohen.

In einer Metaanalyse von 21 Studien in verschiedenen Tumorentitäten wurde die Korrelation von radiologischem Tumoransprechen und der gesundheitsbezogenen Lebensqualität untersucht (14). Ein objektives Tumoransprechen ist demnach signifikant mit einem Symptomansprechen assoziiert. Auch Patienten, die kein objektives Tumoransprechen im Sinne einer kompletten oder partiellen Remission, sondern „nur“ eine Stabilisierung der Erkrankung erfuhren, hatten ein Symptomansprechen. Die medikamentöse Tumorthherapie hatte somit den größten Effekt für die Steigerung der gesundheitsbezogenen Lebensqualität bei Patienten, die eine komplette oder partielle Remission erfuhren, gefolgt von denen, die eine Stabilisierung ihrer Erkrankung als bestes Ansprechen aufwiesen.

Daher ist es sinnvoll, mit zusätzlichen Endpunkten die Verlängerung der Lebenszeit in einen biologischen Kontext zu bringen und patientenrelevante Therapieziele, wie die Reduktion der Morbidität bzw. der physischen und psychischen Belastung durch die Tumorerkrankung, sowie eine Stabilisierung des Patienten über einen relevanten Zeitraum durch Verhinderung des weiteren Tumorwachstums mit zu betrachten. Viele Patienten sind auch oder gerade bei einer infausten Erkrankung bereit, für einen kleinen Gewinn an Lebenszeit oder Lebensqualität eine medikamentöse, antitumorös wirkende Therapie mit den damit verbundenen Toxizitäten auf sich zu nehmen, und zwar für einen kleineren Unterschied in der Wirksamkeit, als Gesunde oder die sie betreuenden Personen annehmen würden. (13)

Fazit:

1. Progressionsfreies Überleben ist beim fortgeschrittenen kolorektalen Karzinom, insbesondere in einer Dritt- und Viertliniensituation in einer Phase-III-Studie nicht nur ein valider Surrogatparameter für das Gesamtüberleben, sondern stellt einen patientenrelevanten Endpunkt aus eigenem Recht dar.
2. In dieser Indikation ist das PFS als kombinierter Mortalitäts- und Morbiditätsendpunkt ein klinisch-medizinischer, sozialmedizinischer und auch für den Patienten sehr relevanter Parameter, der einen beträchtlichen Nutzen hinsichtlich Morbidität und Lebensqualität an sich darstellt.
3. Das Tumoransprechen im Sinne von Remissionen und Tumorstabilisierungen ist ebenfalls ein patientenrelevanter Parameter für Patienten, indem es die physische und psychische Morbiditätslast senkt.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Zur Bewertung des Zusatznutzens von Regorafenib gegenüber BSC in der Behandlung von erwachsenen Patienten mit metastasiertem kolorektalen Karzinom, die eine vorangegangene Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie, und, sofern ein KRAS-Wildtyp vorliegt, eine Anti-EGFR-Therapie gehabt haben, oder für eine solche Therapie nicht infrage kommen, kann die prospektive, zwei-armige, randomisierte, kontrollierte, doppelt verblindete, multizentrische Phase-III-Zulassungsstudie 14387 (CORRECT) gemäß den oben beschriebenen Ein- und Ausschlusskriterien herangezogen werden. Hierbei handelt es sich um einen Nachweis der Evidenzstufe Ib i.S.d. §5 Abs. 6 Satz 4 Nr. 2 VerfO-GBA.

Intervention und Vergleichstherapie

Die Patienten in dieser Studie wurden im Verhältnis 2:1 randomisiert und mit Regorafenib + BSC oder Placebo + BSC behandelt.

Für die Bewertung des Zusatznutzens wird aus der Studie gemäß dem zugelassenen Anwendungsgebiet der Vergleich des Regorafenib + BSC-Armes mit dem Placebo + BSC-Arm herangezogen. Dieser Vergleich war prospektiv in der Studie geplant. Die Studie wurde nach einer geplanten zweiten Zwischenanalyse anhand der Daten von 432 Todesereignissen vorzeitig auf Empfehlung des Data Monitoring Committee (DMC) beendet, da eine signifikante Überlegenheit des Regorafenib + BSC-Armes gegenüber dem Placebo + BSC-Arm gezeigt werden konnte.

Studienpopulation

Die Einschlusskriterien der Studie 14387 CORRECT umfassten erwachsene Patienten mit einem metastasierten kolorektalen Karzinom, die mindestens eine vorangegangene Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie, und, sofern ein KRAS-Wildtyp vorliegt, eine Anti-EGFR-Therapie gehabt hatten, oder für eine solche Therapie nicht infrage kamen. Unter den vorangegangenen Therapien kam es entweder zum Nichtansprechen oder zum Fortschreiten der Tumorerkrankung, oder es bestand eine Unverträglichkeit, so dass somit alle evidenzbasierten Therapieoptionen ausgeschöpft waren, jedoch weiterhin eine Behandlungsindikation bestand. Die folgende Tabelle 4-4 fasst die wesentlichen Ein- und Ausschlusskriterien der Studie 14387 CORRECT zusammen (alle Details sind in Tabelle 4-71 wiedergegeben).

Tabelle 4-4: Wesentliche Ein- und Ausschlusskriterien der eingeschlossenen Studie 14387 CORRECT

Einschlusskriterien	
1	Unterschiedene Einverständniserklärung vor der ersten studien-spezifischen Maßnahme. Patienten müssen die Einverständniserklärung verstehen können und diese unterzeichnen wollen.
2	Histologische oder zytologische Dokumentation eines Adenokarzinoms des Kolons oder Rektums. Alle anderen histologischen Typen sind ausgeschlossen.
3	Patienten mit einem metastasierten Kolorektalkarzinom (Stadium IV).
4	Progression während oder binnen drei Monate nach der letzten Gabe einer zugelassenen Standardtherapie, die Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab und Cetuximab oder Panitumumab (wenn KRAS-WT vorliegt) umfassen muss. Patienten, die mit Oxaliplatin in einem adjuvanten Setting behandelt wurden, sollten während der Behandlung oder binnen sechs Monaten nach Beendigung der Behandlung eine Progression erlitten haben. Patienten, die nach sechs Monaten nach Beendigung einer adjuvanten Oxaliplatin-haltigen Therapie progredient waren, müssen erneut mit einer Oxaliplatin-basierten Therapie behandelt worden sein, um eingeschlossen werden zu können. Patienten, die eine Standardtherapie wegen intolerabler Toxizität vor einer Progression beendeten und bei denen eine erneute Therapie mit der gleichen Substanz nicht indiziert war, konnten ebenfalls in die Studie aufgenommen werden. Patienten mit einem unbekanntem KRAS-Status zum Screening mussten eine vorangegangene anti-EGFR-Therapie gehabt haben.
5	Patienten mit einem mCRC mit messbaren oder nicht messbaren Tumorknoten, entsprechend RECIST-Kriterien, Version 1.1
6	Männer oder Frauen ≥ 18 Jahre
7	Eastern Cooperative Oncology Group (ECOG) Performance Status ≤ 1
8	Lebenserwartung von mindestens drei Monaten
9	Frauen in gebärfähigem Zustand und Männer müssen zustimmen, eine adäquate Verhütungsmaßnahme zu verwenden von der Unterzeichnung der Einwilligungserklärung bis mindestens drei Monate nach dem Ende der Einnahme der Studienmedikation.
10	Adäquate Knochenmarks-, Leber- und Nierenfunktion, bestimmt anhand der folgenden Laborparameter innerhalb von sieben Tagen vor Beginn der Studienbehandlung:
11	Gesamtbilirubin $\leq 1,5$ fache des oberen Normwertes (ONW)
12	Alaninaminotransferase (ALT) und Aspartataminotransferase (AST) $\leq 2,5$ fache des ONW
13	Amylase und Lipase $\leq 1,5$ fache des ONW
14	Serumkreatinin $\leq 1,5$ fache des ONW
15	Glomeruläre Filtrationsrate (GFR) ≥ 30 ml/min/1,73m ² entsprechend der verkürzten MDRD-Formel
16	International Normalized Ratio (INR) / Partielle Thromboplastinzeit (PTT) $\leq 1,5$ fache des ONW
17	Thrombozytenzahl ≥ 100.000 /mm ³ , Hämoglobin ≥ 9 g/dl, absolute Neutrophilenzahl ≥ 1.500 /mm ³ . Transfusionen, um die genannten Werte zu erreichen, waren nicht erlaubt.
18	Alkalische Phosphatase (ALP) $\leq 2,5$ fache des ONW (≤ 5 fache des ONW bei Patienten mit Lebermetastasen)
Ausschlusskriterien	
1	Vorangegangene Therapie mit Regorafenib
2	Vorherige Behandlung innerhalb dieser Studie
3	Vorangegangene oder gleichzeitig bestehende andere Tumorerkrankung innerhalb von 5 Jahren vor der Randomisierung AUSSER kurativ behandeltes carcinoma-in-situ der Zervix, nicht-melanomartige Hauttumoren und oberflächliches Harnblasenkarzinom
4	Extended-Field-Radiotherapie innerhalb von vier Wochen oder Limited-Field-Radiotherapie innerhalb von zwei Wochen vor der Randomisierung. Patienten müssen sich von allen therapiebedingten Nebenwirkungen erholt haben. Die bestrahlte Läsion muss eine Progression aufweisen, wenn sie die einzige Tumorknoten war.

- | | |
|----|--|
| 5 | Große chirurgische Eingriffe, offene Biopsien, oder signifikante Traumata innerhalb von 28 Tagen vor Beginn der Studienbehandlung |
| 6 | Schwangere oder stillende Frauen. Frauen mit gebärfähigem Potential müssen einen negativen Schwangerschaftstest innerhalb von sieben Tagen aufweisen. |
| 7 | Chronische Herzinsuffizienz \geq New York Heart Association (NYHA)-Klasse II |
| 8 | Instabile Angina pectoris, neu aufgetretene Angina, oder Myokardinfarkt innerhalb von sechs Monaten vor Beginn der Studienbehandlung |
| 9 | Herzrhythmusstörungen, die einer antiarrhythmischen Therapie bedürfen (Betablocker und Digoxin sind erlaubt) |
| 10 | Unkontrollierte Hypertonie (systolischer Blutdruck >150 mmHg oder diastolischer Blutdruck >90 mmHg trotz optimaler Behandlung) |
| 11 | Patienten mit Phäochromozytom |
| 12 | Pleuraerguß oder Aszites die zur Atembehinderung führen (\geq CTCAE-Grad 2 in Dyspnoe) (CTCAE: Common Terminology Criteria for Adverse Events) |
| 13 | Arterielle oder venöse Thrombosen oder Embolien, wie z. B. ischämischer Hirninfarkt (inkl. einer transitorisch-ischämischen Attacke), tiefe Beinvenenthrombosen oder Lungenembolien innerhalb von sechs Monaten vor Beginn der Studienbehandlung |
| 14 | Bestehende Infektion $>$ Grad 2 CTCAE Version 3.0 |
| 15 | Bekannte HIV-Infektion |
| 16 | Aktive Hepatitis B oder C, oder chronische Hepatitis B oder C, die einer Behandlung mit antiviralen Substanzen bedarf |
| 17 | Patienten mit Epilepsie, die medikamentös behandlungsbedürftig sind |
| 18 | Anamnese von oder bestehende Hirnmetastasen |
| 19 | Z. n. Organtransplantation |
| 20 | Patienten mit Anamnese oder bestehender Blutungsdiathese |
| 21 | Blutungen \geq CTCAE-Grad 3 innerhalb von vier Wochen vor Beginn der Studienbehandlung |
| 22 | Nicht heilende Wunden, Ulcera oder Knochenfrakturen |
| 23 | Nierenversagen mit Erfordernis einer Hämo- oder Peritonealdialyse |
| 24 | Dehydratation \geq Grad 1 CTCAE Version 3.0 |

Beleg des medizinischen Zusatznutzens

Die folgende Tabelle 4-5 fasst die Ergebnisse der patientenrelevanten Endpunkte aus der Studie 14387 CORRECT zum Nachweis des Zusatznutzens von Regorafenib + BSC gegenüber Placebo + BSC zusammen.

Tabelle 4-5: Zusammenfassung der Ergebnisse der patientenrelevanten Endpunkte

Studie		Behandlungsarm	
CORRECT 14387		Regorafenib + BSC	Placebo + BSC
n		505	255
Gesamtüberleben (OS)	Median in Tagen (95 %-KI)	196 (178; 222)	151 (134; 177)
	Hazard Ratio (95 %-KI) p-Wert	0,774 (0,636; 0,942) 0,0052	
Progressionsfreies Überleben (PFS)	Median in Tagen (95 %-KI)	59 (57; 65)	52 (51; 63)
	Hazard Ratio (95 %-KI) p-Wert	0,494 (0,419; 0,582) <0,000001	
Objektive Tumoransprechrates (ORR = CR + PR)	n (%) (95% KI)	5 (1,0) (0,3; 2,3)	1 (0,4) (0,0; 2,2)
	p-Wert	0,1884	
Krankheitskontrollrate (DCR = CR + PR + SD)	n (%) (95 %-KI)	207 (41,0) (36,7; 45,4)	38 (14,9) (10,8; 19,9)
	p-Wert	<0,000001	
Dauer der Tumorstabilisierung	Median in Tagen (95 %-KI)	60 (57; 70)	52 (51; 53)
	HR (95 %-KI)	0,471 (0,398; 0,558)	
Gesundheitsbezogene Lebensqualität	EORTC QLQ-C30 Zeit- adjustierte Fläche unter der Kurve, Mittelwerte der Fläche (95 % KI)	56,93 (54,79; 59,08)	58,13 (55,72; 60,53)
	Differenz (95 % KI)	-1,19 (-3,13; 0,75)	
Verträglichkeit Gesamtanzahl Patienten mit UE	n (%)	498 (99,60)	245 (96,84)
	Relatives Risiko (RR) (95 %-KI)	0,97 (Placebo+BSC/Regorafenib+BSC) (0,95; 0,99)	
Verträglichkeit Gesamtanzahl Patienten mit UE vom CTAE-Grad 1	n (%)	12 (2,40)	45 (17,79)
	RR (95 %-KI)	0,84 (Regorafenib+BSC/ Placebo+BSC) (0,79; 0,89)	
Verträglichkeit Gesamtanzahl Patienten mit UE vom CTAE-Grad 2	n (%)	96 (19,20)	76 (30,04)
	RR (95 %-KI)	0,87 (Regorafenib+BSC/ Placebo+BSC) (0,79; 0,95)	
Verträglichkeit Gesamtanzahl Patienten mit UE vom CTAE-Grad 3	n (%)	280 (56,00)	67 (26,48)
	RR (95 %-KI)	0,47 (Placebo+BSC/Regorafenib+BSC) (0,38; 0,59)	
Verträglichkeit Gesamtanzahl	n (%)	43 (8,60)	20 (7,91)
	RR	0,92 (Regorafenib+BSC/ Placebo+BSC)	

Patienten mit UE vom CTAE-Grad 4	(95 %-KI)	(0,55; 1,53)	
Verträglichkeit Gesamtanzahl Patienten mit UE vom CTAE-Grad 5	n (%)	67 (13,40)	37 (14,62)
	RR (95 %-KI)	0,99 (Regorafenib+BSC/ Placebo+BSC) (0,93; 1,05)	
Verträglichkeit Gesamtanzahl Patienten mit SUE	n (%)	219 (43,80)	100 (39,53)
	RR (95 %-KI)	0,90 (Placebo+BSC/Regorafenib+BSC) (0,75; 1,08)	
Verträglichkeit Gesamtanzahl Patienten mit UE, die zum Studienabbruch geführt haben	n (%)	88 (17,60)	32 (12,65)
	RR (95 %-KI)	0,72 (Placebo+BSC/Regorafenib+BSC) (0,49; 1,05)	
Quelle: CORRECT 14387 Studienbericht und ZA1			

Aus diesen Ergebnissen wird ein Beleg für einen beträchtlichen Zusatznutzen abgeleitet. Die Ergebnissicherheit „Beleg“ ergibt sich dabei aus der hohen internen und externen Validität der Studie, der hohen klinischen Relevanz des Therapieeffektes, dem Grad der statistischen Signifikanz sowie der hohen Datenqualität. Eine ausführliche Begründung zur Aussagekraft der Nachweise mit einer weitergehenden Begründung für die Ergebnissicherheit „Beleg“ erfolgt im Kapitel 4.4.1.

Patientenrelevanter Endpunkt Mortalität (Gesamtüberleben)

Regorafenib zeigt in der Behandlung von erwachsenen Patienten mit einem metastasierten kolorektalen Karzinom, die eine vorangegangene Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie, und, sofern ein KRAS-Wildtyp vorliegt, eine Anti-EGFR-Therapie gehabt haben, oder für eine solche Therapie nicht infrage kommen, hinsichtlich des Gesamtüberlebens einen Beleg für einen beträchtlichen Zusatznutzen. Die Kategorie „Beleg“ bezüglich der Ergebnissicherheit wird reklamiert, da alle im Papier „Points to Consider on Application with 1. Meta-analyses; 2. One Pivotal Study“ der European Medicines Agency (EMA) genannten Kriterien, wann nur eine Phase III-Studie als ausreichend angesehen wird, erfüllt sind (15). Eine ausführliche Begründung erfolgt im Kapitel 4.4.2.

Das mediane Gesamtüberleben betrug im Regorafenib + BSC-Arm 196 Tage (95 %-KI: 178-222) gegenüber 151 Tagen (95 %-KI: 134-177) im Placebo + BSC-Arm. Zwischen dem Regorafenib + BSC- und dem Placebo + BSC-Arm ergibt sich somit ein Unterschied von 45 Tagen. Der Vergleich zwischen der Regorafenib + BSC- und der Placebo + BSC-Gruppe zeigte hinsichtlich des Gesamtüberlebens einen statistisch signifikanten Vorteil mit einer Hazard Ratio von 0,774 (95 %-KI: 0,636-0,942; p=0,005178) und geht somit mit einer Risikoreduktion von 22,6 % einher.

Bezüglich des Gesamtüberlebens ergaben Sensitivitätsanalysen konsistente Ergebnisse (siehe Abschnitt 4.3.1.3.1).

Die Kaplan-Meier-Überlebenskurve (siehe Abbildung 4-1) zeigt den Kurvenverlauf der Mortalitätsrate über den Studienverlauf. Nach anfänglichem Gleichlauf trennen sich die Kurven ab ca. Tag 50 und bleiben bis zum Ende der Beobachtungsdauer getrennt.

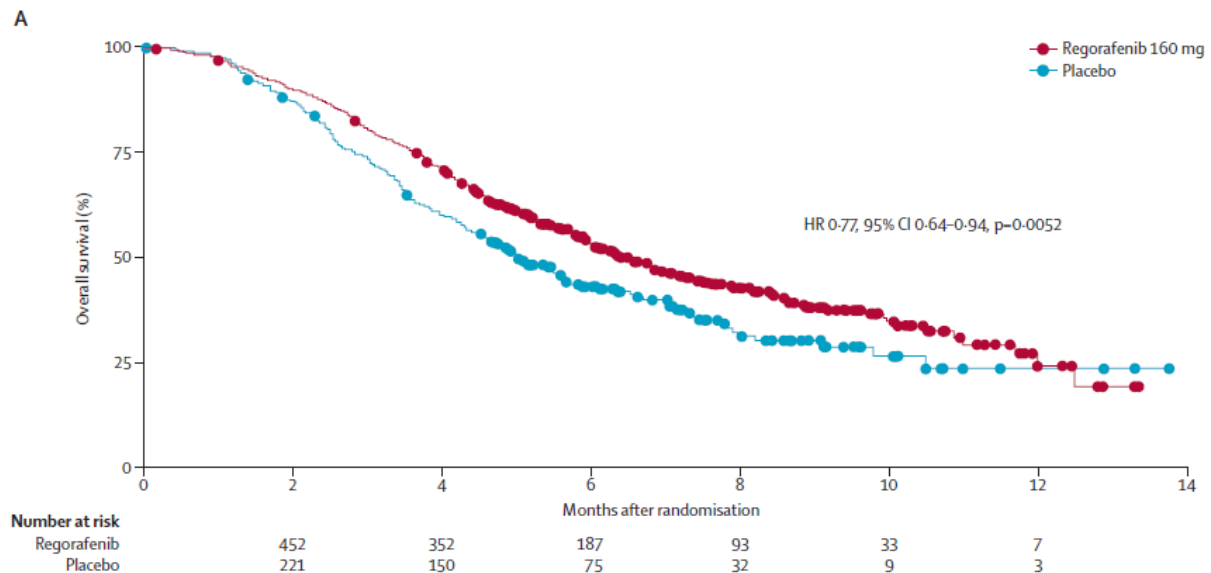


Abbildung 4-1: Kaplan-Meier-Kurve für das Gesamtüberleben in der Studie 14387 CORRECT (ITT-Population)

Quelle: (16)

Die 6-Monats-Überlebensrate betrug 52,5 % in der Regorafenib + BSC-Gruppe und 43,5 % in der Placebo + BSC-Gruppe und entspricht somit einer relativen Verbesserung von 20,7 %. Die relative Verbesserung betrug 24 % in der 9-Monats-Überlebensrate (38,2 % in der Regorafenib-Gruppe und 30,8 % in der Placebo-Gruppe). Die Wirksamkeit von Regorafenib + BSC auf das Gesamtüberleben war weitgehend konsistent über die Subgruppen. Die Signifikanz des Behandlungseffektes auf das Gesamtüberleben war nach der Adjustierung für die Stratifizierungsfaktoren in einer multivariaten Analyse robust: die adjustierte Hazard Ratio betrug 0,773 (95 %-KI 0,634-0,942; $p=0,005255$).

Zusammenfassend belegt die Studie 14387 CORRECT gemäß §5 Abs. 7 Satz 1 Nr. 1 5. Kap. Verfo-GBA für Regorafenib im bedeutendsten patientenrelevanten Endpunkt Mortalität eine nachhaltige und gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie BSC bisher nicht erreichte Verbesserung des therapie relevanten Nutzens durch die deutliche Verlängerung des medianen Gesamtüberlebens. Der Zusatznutzen wird als beträchtlich eingestuft.

Patientenrelevanter Endpunkt Mortalität/Morbidität Progressionsfreies Überleben (PFS)

Regorafenib zeigt bei den behandelten Patienten einen erheblichen Zusatznutzen bei der Verbesserung des progressionsfreien Überlebens und bestätigt somit die Ergebnisse des primären Endpunkts Gesamtüberleben.

Die Behandlung mit Regorafenib + BSC resultierte bezüglich des progressionsfreien Überlebens in einer Risikominderung von 50,6 % im Vergleich zu Placebo + BSC. Die Hazard Ratio betrug 0,494 (95 %-KI: 0,419-0,582; $p < 0,000001$). Das mediane progressionsfreie Überleben betrug 59 Tage in der Regorafenib + BSC-Gruppe und 52 Tage in der Placebo + BSC-Gruppe. Das mediane progressionsfreie Überleben verbesserte sich um 13,5 %. Die relative Verbesserung betrug 211 % in der 3-Monats-progressionsfreien Überlebensrate (42 % in der Regorafenib + BSC-Gruppe und 13,5 % in der Placebo + BSC - Gruppe). Die relative Verbesserung betrug 519 % in der 6-Monats-Überlebensrate (13 % in der Regorafenib + BSC-Gruppe und 2,1 % in der Placebo + BSC-Gruppe). Die Wirksamkeit von Regorafenib + BSC auf das progressionsfreie Überleben war konsistent über die Subgruppen hinweg. Die Signifikanz des Behandlungseffektes auf das progressionsfreie Überleben war nach der Adjustierung für die Stratifizierungsfaktoren in einer multivariaten Analyse robust: die adjustierte Hazard Ratio betrug 0,479 (95 %-KI 0,405; 0,565; $p < 0,000001$). Die punktuelle Betrachtung der PFS-Differenz an den Medianen ist in diesem Fall für den erheblichen Zusatznutzen nicht repräsentativ, der sich aus der Hazard Ratio für PFS ergibt, die den gesamten Verlauf der Kaplan-Meier-Kurven reflektiert (siehe Abbildung Abbildung 4-4). Die protokollgemäße erste radiologische Tumorauswertung erst nach 8 Wochen Therapie bedingt eine geringe Diskriminierbarkeit der beiden Studiengruppen vor diesem Zeitpunkt mittels des Kriteriums Progression nach RECIST. Insofern ist die Hazard Ratio für PFS zur Bestimmung des Zusatznutzens maßgeblich.

Patientenrelevanter Endpunkt Objektive Tumorausprechrates (ORR)

Regorafenib zeigt bei den behandelten Patienten keinen Zusatznutzen bei der Verbesserung der Objektive Tumorausprechrates.

Für die Objektive Tumorausprechrates (PR + CR) wurde eine Korrelation mit dem Gesamtüberleben bei kolorektalem Karzinom nachgewiesen (17, 18). Die von Johnson et al. (18) gefundene Korrelation von Tumorausprechen und Gesamtüberleben (siehe Kapitel 4.2.5.2) bei einer Stichprobengröße von $n=750$ von 20 % wird in der CORRECT-Studie ($n=760$) nicht erfüllt (Ansprechrates im Regorafenib + BSC-Arm: 1 %, im Placebo + BSC-Arm: 0,4 %; Unterschied 0,6 %). Tumorausprechen nach RECIST- oder WHO-Kriterien steht für viele zielgerichtete Arzneimittel im Allgemeinen und für Multikinaseinhibitoren im Besonderen nicht im Vordergrund ihrer biologischen Wirkung. Vielmehr geht es um eine Tumorstabilisierung und Veränderungen im Tumor, in der Tumorumgebung bzw. in den einen Tumor versorgenden Blutgefäßen, die günstige Wirkungen auf die Tumorprogression bewirken. Daher werden für diese Arzneimittel andere Methoden des Wirksamkeitsnachweises diskutiert und angewandt (19-22).

Patientenrelevanter Endpunkt Krankheitskontrollrate (DCR)

Regorafenib zeigt bei den behandelten Patienten einen beträchtlichen Zusatznutzen bei der Verbesserung der Krankheitskontrollrate und belegt somit die Konsistenz des Vorteils einer Behandlung mit Regorafenib für das Gesamtüberleben gegenüber der Kontrollgruppe.

Die Krankheitskontrollrate (CR + PR + SD >6 Wochen) in der ITT-Population war höher in der Regorafenib + BSC-Gruppe als in der Placebo + BSC-Gruppe. Die Krankheitskontrollrate betrug 41,0 % (95 %-KI: 36,7;45,4 %) in der Regorafenib + BSC-Gruppe und 14,9 % (95 %-KI: 10,8;19,9 %) in der Placebo + BSC-Gruppe. Der Unterschied von -25,94 % (95 %-KI: -32,06;19,82 %) zwischen den beiden Gruppen war statistisch signifikant (einseitiges $p < 0,000001$; präspezifiziertes alpha 0,025). Der Unterschied war hauptsächlich durch die Rate an Krankheitsstabilisierungen (SD) bedingt. Die Unterschiede in den Ansprechraten und in der Krankheitskontrollrate sowie die zugehörigen 95 %-Konfidenzintervalle basieren auf dem Cochran-Mantel-Haenszel-Test, adjustiert auf dieselben Stratifizierungsfaktoren wie für den primären Endpunkt.

Patientenrelevanter Endpunkt Dauer der Tumorstabilisierung

Regorafenib zeigt bei den behandelten Patienten einen beträchtlichen Zusatznutzen bei der Verbesserung der Dauer der Tumorstabilisierung und belegt somit die Konsistenz des Vorteils einer Behandlung mit Regorafenib für das Gesamtüberleben gegenüber der Kontrollgruppe.

Die mediane Dauer der Tumorstabilisierung betrug im Median 60 Tage (95 %-KI: 57; 70) in der Regorafenib + BSC-Gruppe und 52 Tage (95 % KI: 51; 53) in der Placebo + BSC-Gruppe. Die Hazard Ratio für die Dauer der Tumorstabilisierung betrug 0,471 (95 %-KI: 0,398; 0,558). Der geringe Unterschied im Median ist hauptsächlich durch die primär progredienten Patienten (250 (49,5 %) in der Regorafenib + BSC-Gruppe bzw. 204 (80,0 %) in der Placebo + BSC-Gruppe) bedingt.

Patientenrelevanter Endpunkt Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität nahm in den beiden Studiengruppen während der Behandlungsphase jeweils nur wenig ab: Für die verschiedenen Dimensionen der Lebensqualität, die mit dem Fragebogen der EORTC QLQ-C30 erfasst wurden, bewegen sich die Veränderungen während der Behandlungsphase im Bereich keine, wenig oder moderat; die Kategorie mit der stärksten Ausprägung der Veränderung (viel) war nicht besetzt. Die Globalbewertung anhand der Zeit-adjustierten Berechnung der Fläche unter der Kurve ergab keinen signifikanten Unterschied (Differenz Regorafenib – Placebo: -1,19 (95 % KI: -3,13; 0,75).

In der Placebo + BSC-Gruppe kam es an sieben von 18 Zeitpunkten zu Verschlechterungen der Lebensqualität in der Kategorie „wenig“, während es in der Regorafenib + BSC-Gruppe an zehn Zeitpunkten zu Verschlechterungen in der Kategorie „wenig“ und an zwei Zeitpunkten zu Verschlechterungen in der Kategorie „moderat“ kam. Betroffen waren in beiden Behandlungsgruppen die Module physische Funktionalität, Rollenfunktionalität, soziale Funktionalität und globaler Gesundheitsstatus. In den Modulen emotionale Funktionalität und kognitive Funktionalität kam es in beiden Gruppen zu keinen Änderungen.

Bei Betrachtung der absoluten Unterschiede mit ihren errechneten 95 %-Konfidenzintervallen und der daraus sich ergebenden Ableitung eines signifikanten Unterschiedes gibt es einen solchen zu einem Zeitpunkt (Zyklus 2) im Modul „Körperliche Funktion“ und zu allen Zeitpunkten im Modul „Soziale Funktion“. Dieses Modul besteht im Fragebogen aus 2

Fragen, eine Beeinträchtigung des Familienlebens bzw. das Zusammensein mit anderen Menschen betreffend. Hier muss konstatiert werden, dass eine Therapie mit Regorafenib offenbar zu einer Beeinträchtigung dieser beiden sozialen Bereiche führt.

Aufgrund der geringen Rücklauftrate wird das Verzerrungspotenzial als hoch eingestuft, sodass folglich nur ein Hinweis auf einen möglichen Zusatznutzen abgeleitet werden könnte. Es wird jedoch konstatiert, dass Regorafenib in Bezug auf die gesundheitsbezogene Lebensqualität keinen Zusatznutzen hat.

Patientenrelevanter Endpunkt therapiebedingte Morbidität (Verträglichkeit)

Die Verträglichkeit wurde anhand von allen unerwünschten Ereignissen (UE), UE nach Schweregradeinteilung gemäß CTCAE-Klassifikation, schwerwiegenden unerwünschten Ereignissen (SUE), sowie UE, die zum Studienabbruch oder zu einer Dosismodifikation führten, operationalisiert.

Es gab keinen relevanten Unterschied in der Gesamtinzidenz von UE zwischen den beiden Behandlungsgruppen: mindestens ein UE wurde während der gesamten Studie von den meisten Patienten der Sicherheitspopulation (SAF) berichtet: 498 (99,6 %) in der Regorafenib + BSC-Gruppe und 245 (96,8 %) in der Placebo + BSC-Gruppe. UE wurden als in ursächlichem Zusammenhang mit der Behandlung stehend bei 465 (93,0 %) in der Regorafenib + BSC-Gruppe und 154 (60,9 %) in der Placebo + BSC-Gruppe angesehen.

Während UE vom Grad 1 und 2 häufiger in der Placebo + BSC-Gruppe registriert wurden (Grad 1: 17,8 % vs. 2,4 %; Grad 2: 30,0 % vs. 19,2 %), waren UE vom Grad 3 häufiger in der Regorafenib + BSC-Gruppe (56,0 % vs. 26,5 %). Die Unterschiede im Grad 1 bis Grad 3 waren jeweils signifikant. Die höhere Inzidenz an UE vom Grad 3 ist im Wesentlichen auf das substanzspezifische Nebenwirkungsprofil von Regorafenib zurückzuführen, bei dem spezifische Nebenwirkungen wie bspw. die Hand-Fuß-Hautreaktion oft den Grad 3 erreichen. Die Inzidenz von UE der Grade 4 und 5 war nicht signifikant unterschiedlich in der Regorafenib + BSC-Gruppe im Vergleich zur Placebo + BSC-Gruppe (8,6 % vs. 7,9 % bzw. 13,4 % vs. 14,6 %).

Schwerwiegende unerwünschte Ereignisse (SUE) wurden bei 219 Patienten (43,8 %) in der Regorafenib + BSC- und bei 100 Patienten (39,5 %) in der Placebo + BSC-Gruppe berichtet, was nicht signifikant unterschiedlich war. Die Gesamtinzidenz der SUE, die als in ursächlichem Zusammenhang mit der Studienmedikation stehend angesehen wurden, war relativ gering in beiden Gruppen, jedoch höher in der Regorafenib + BSC verglichen mit der Placebo + BSC-Gruppe (11,8 % vs. 3,6 %).

Obwohl signifikant mehr Patienten im Regorafenib + BSC-Arm Dosismodifikationen aufgrund eines UE hatten (66,6 % vs. 22,5 %), war der Unterschied in den UE, die zu einer dauernden Beendigung der Therapie führten, relativ klein (17,6 % vs. 12,6 %) und nicht signifikant unterschiedlich. Von diesen Patienten beendeten 41 (8,2%) im Regorafenib + BSC-Arm und 3 (1,2%) im Placebo + BSC-Arm dauerhaft die Studientherapie, da das jeweilige UE als in ursächlichem Zusammenhang mit der Studienmedikation stehend

angesehen wurde. Dies weist darauf hin, dass die meisten UE mit Dosismodifikationen beherrscht werden können und nicht zur dauernden Therapiebeendigung führten.

Zusammenfassend lässt sich somit folgendes in Bezug auf die Verträglichkeit feststellen:

- Es besteht ein geringerer Schaden mit Ausmaß beträchtlich von Regorafenib gegenüber Placebo bei den unerwünschten Ereignissen vom Schweregrad 1.
- Es besteht ein geringerer Schaden mit Ausmaß gering von Regorafenib gegenüber Placebo bei den unerwünschten Ereignissen von Schweregrad 2.
- Es besteht ein größerer Schaden mit Ausmaß erheblich von Regorafenib gegenüber Placebo bei den unerwünschten Ereignissen vom Schweregrad 3.
- Keine signifikanten Unterschiede zwischen Regorafenib und Placebo bestehen bei der Gesamtzahl unerwünschter Ereignisse, bei schwerwiegenden unerwünschten Ereignissen, bei unerwünschten Ereignissen der Schweregrade 4 und 5 sowie bei unerwünschten Ereignissen, die zu einem Studienabbruch geführt haben.
- Das Verträglichkeitsprofil zeigt die für die Substanz(klasse) erwartbaren unerwünschten Ereignisse, die gegenüber dem nicht antineoplastisch wirkenden Placebo in den substanzspezifischen Nebenwirkungskategorien signifikant häufiger vorkommen.
- Da Placebo und BSC *per se* nicht auf den Tumor wirkt und somit keine Nebenwirkungen einer antitumorösen Therapie auftreten, kann dies den Nutzen von Regorafenib jedoch nicht infrage stellen, da für einen patientenrelevanten und zweckmäßigen Vergleich der Verträglichkeit Substanzen herangezogen werden sollten, die im Anwendungsgebiet das gleiche Therapieziel, nämlich die Verlängerung des Gesamtüberlebens, haben.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen

Zusatznutzen

Für die Bewertung des Ausmaßes und der Wahrscheinlichkeit des Zusatznutzens von Regorafenib wurde die Studie 14387 CORRECT herangezogen. Diese Studie ist ein RCT von hoher Evidenz und Ergebnissicherheit, da die methodischen Anforderungen an eine qualitativ gut durchgeführte, randomisierte kontrollierte Studie erfüllt wurden und sich somit nach Überprüfung ein niedriges Verzerrungspotential auf Studienebene ergibt.

Die Ergebnisse zum patientenrelevanten Endpunkt Mortalität (Gesamtüberleben) belegen für die Patienten unter Behandlung mit Regorafenib gegenüber der zweckmäßigen Vergleichstherapie BSC in diesem Therapiesetting erstmalig einen signifikanten Überlebensvorteil. Die bislang nicht erreichte Verbesserung des therapielevanten Nutzens ist gegeben durch die beträchtliche, sogar erstmalig überhaupt nachgewiesene signifikante Verlängerung des medianen Überlebens gegenüber BSC.

Die Ergebnisse zu den patientenrelevanten Endpunkten progressionsfreies Überleben, Krankheitskontrollrate und Dauer der Tumorstabilisierung zeigen jeweils eine signifikante Überlegenheit der Patienten in der Regorafenib + BSC-Gruppe gegenüber der Placebo + BSC-Gruppe und zeigen somit die Konsistenz des Vorteils einer Behandlung mit Regorafenib für das Gesamtüberleben gegenüber der Kontrollgruppe. Der therapierelevante Nutzen wird als erheblich (PFS) bzw. beträchtlich (DCR, Dauer der Tumorstabilisierung) eingestuft.

Für den patientenrelevanten Endpunkt gesundheitsbezogene Lebensqualität ergeben sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen in den einzelnen Modulen der angewandten Fragebogen sowie in der Zeit-adjustierten Analyse der Fläche unter der Kurve des Instruments EORTC QLQ-C30. Generell nimmt unter beiden Interventionen die gesundheitsbezogene Lebensqualität jeweils geringfügig ab. Ausnahme bildet das Submodul „Soziale Funktion“, das das Familienleben sowie das Zusammensein mit anderen Menschen erfasst, und bei dem es zu einer signifikant schlechteren Beurteilung in der Regorafenib + BSC-Gruppe kam. Hierbei ist jedoch das hohe Verzerrungspotential zu berücksichtigen.

Die Ergebnisse zu dem patientenrelevanten Endpunkt Verträglichkeit zeigen eine bezüglich der Inzidenzraten vergleichbare Häufigkeit (96,8 % in der Placebo + BSC-Gruppe und 99,6 % in der Regorafenib + BSC-Gruppe). Es liegen erwartbare höhere Häufigkeiten von bestimmten UE vor, die auf den Wirkmechanismus von Regorafenib zurückzuführen sind. Allerdings gibt es keine signifikant höhere Häufigkeit von SUE und UE, die zum Studienabbruch führten. Bei der Bewertung ist zu berücksichtigen, dass sich trotz des häufigeren Auftretens von UE die Lebensqualität in der Regorafenib + BSC-Gruppe nicht signifikant von der in der Placebo + BSC-Gruppe unterschied. Zweckmäßig und patientenrelevant ist ferner, das Verträglichkeitsprofil von Substanzen zu berücksichtigen, die in der klinischen Routine alternativ zu Regorafenib eingesetzt werden.

Die folgende Tabelle 4-6 fasst die Bewertung des Ausmaßes und der Wahrscheinlichkeit des medizinischen Zusatznutzens von Regorafenib gegenüber der zweckmäßigen Vergleichstherapie zusammen.

Tabelle 4-6: Zusammenfassende Bewertung des medizinischen Zusatznutzens von Regorafenib gegenüber BSC

Patientenrelevanter Endpunkt	Zusammenfassende Bewertung		
	Ausmaß	Wahrscheinlichkeit	Begründung
Medianes Gesamtüberleben	beträchtlich	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential. Nachhaltige und gegenüber der zweckmäßigen Vergleichstherapie BSC bisher nicht erreichte große Verbesserung des therapielevanten Nutzens Gesamtüberleben durch die erhebliche Risikominimierung, an der Erkrankung zu versterben.
Progressionsfreies Überleben	erheblich	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential. Nachhaltige und gegenüber der zweckmäßigen Vergleichstherapie BSC bisher nicht erreichte große Verbesserung des therapielevanten Nutzens progressionsfreies Überleben durch die erhebliche Risikominimierung, an der Erkrankung zu versterben oder eine Progression zu erleiden.
Objektive Tumoransprechrate	Kein Zusatznutzen	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential. Kein Unterschied in der objektiven Tumoransprechrate (PR/CR).
Krankheitskontrollrate	beträchtlich	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential. Nachhaltige und gegenüber der zweckmäßigen Vergleichstherapie BSC bisher nicht erreichte große Verbesserung des therapielevanten Nutzens Krankheitskontrollrate durch die erhebliche Erhöhung des Prozentsatzes an Patienten, die unter

			der Therapie mit Regorafenib ein objektives Tumoransprechen oder eine Stabilisierung der Erkrankung erfahren.
Dauer der Tumorstabilisierung	beträchtlich	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential. Nachhaltige und gegenüber der zweckmäßigen Vergleichstherapie BSC bisher nicht erreichte große Verbesserung des therapielevanten Nutzens Dauer der Tumorstabilisierung durch die erhebliche Erhöhung des Prozentsatzes an Patienten, die unter der Therapie mit Regorafenib eine Verlängerung der Dauer der Stabilisierung ihrer Tumorerkrankung erfahren.
Gesundheitsbezogene Lebensqualität	Kein Zusatznutzen	Hinweis	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Hohes Verzerrungspotential, da nur Daten von 92 % der Patienten zum Ausgangszeitpunkt und von 54 % zum Zeitpunkt „Ende der Therapie“ vorlagen. Die Verschlechterung in der gesundheitsbezogenen Lebensqualität bewegt sich für beide Gruppen in den Kategorien „kein Unterschied“, „wenig“ und „moderat“. In der Zeit-adjustierten Analyse der Fläche unter der Kurve fand sich kein signifikanter Unterschied.
Therapiebezogene Morbidität (Verträglichkeit)			
UE	Kein grösserer Schaden belegt	Beleg	Die Studie entspricht den Kriterien für eine adäquate Zulassungsstudie. Niedriges Verzerrungspotential.
SUE	Kein grösserer Schaden belegt	Beleg	Bei der Gesamtinzidenz von UE ist bei der Berechnung des Relativen Risikos der obere Wert des KI >0,90 und wird nach der Methodik des IQWiG trotz formaler statistischer Signifikanz als klinisch nicht relevant gewertet.
UE, die zum Studienabbruch geführt haben	Kein grösserer Schaden	Beleg	Es wurden keine signifikanten Unterschiede im Vergleich zu Placebo

	belegt		in den Raten an Patienten mit SUE, UE der Schweregrade 4 und 5 sowie von UE, die zum Studienabbruch führten, gefunden.
UE CTCAE-Grade 4 und 5	Kein grösserer Schaden belegt	Beleg	
UE CTCAE-Grad 1	Geringerer Schaden, Ausmaß: beträchtlich	Beleg	Für Regorafenib wurde eine signifikant geringere Rate an Patienten mit UE vom CTCAE-Grad 1 gefunden.
UE CTCAE-Grad 2	Geringerer Schaden, Ausmaß: gering	Beleg	Für Regorafenib wurde eine signifikant geringere Rate an Patienten mit UE vom CTCAE-Grad 2 gefunden.
UE CTCAE-Grad 3	Grösserer Schaden, Ausmaß: erheblich	Beleg	Für Regorafenib wurde eine signifikant höhere Rate an Patienten mit UE vom CTCAE-Grad 3 gefunden.

Zusammenfassung

Zusammenfassend zeigen die Ergebnisse der Bewertung einen **Beleg für einen beträchtlichen Zusatznutzen** von Regorafenib gegenüber der zweckmäßigen Vergleichstherapie BSC für erwachsene Patienten mit metastasiertem kolorektalen Karzinom, die eine vorangegangene Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie, und, sofern ein KRAS-Wildtyp vorliegt, eine Anti-EGFR-Therapie gehabt haben, oder für eine solche Therapie nicht infrage kommen.

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.

Es wurde untersucht, ob die Behandlung mit dem Multikinase-Inhibitor Regorafenib im Anwendungsgebiet zu einem Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie im Sinne des §35a SGB V führt. Das Anwendungsgebiet bzw. die Zielpopulation beinhaltet dabei „Patienten mit metastasiertem Kolorektalkarzinom (KRK), die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie“ (23). Der Nachweis des Zusatznutzens erfolgt auf Basis randomisierter, kontrollierter klinischer Studien (RCT), die die Endpunkte Mortalität (Gesamtüberleben), Mortalität/Morbidität (progressionsfreies Überleben), Morbidität (Objektive Tumoransprechrates, Krankheitskontrollrate, Dauer der Tumorstabilisierung), gesundheitsbezogene Lebensqualität sowie Verträglichkeit untersuchten.

Die Bewertung des Zusatznutzens von Regorafenib erfolgt im Vergleich zu Best-Supportive-Care (BSC). Diese Therapie wurde vom G-BA als zweckmäßige Vergleichstherapie bestimmt. Als „Best-Supportive-Care“ wird die Therapie verstanden, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirkt (1).

Um nachzuweisen, dass die Behandlung mit Regorafenib einen medizinischen Zusatznutzen im Vergleich zur zweckmäßigen Vergleichstherapie hat, werden folgende Endpunkte berücksichtigt:

- Wichtigster patientenrelevanter Endpunkt in Bezug auf die **Mortalität** ist die **Verlängerung der Gesamtüberlebenszeit**, gemessen als Zeitraum von der Randomisierung bis zum Tod (nicht krankheitsspezifisch).
- Eng verknüpft mit diesem Endpunkt sind der Endpunkt zur **Mortalität/Morbidität progressionsfreies Überleben**, bzw. die Endpunkte zur **Morbidität Objektive Tumoransprechrates, Krankheitskontrollrate und Dauer der Tumorstabilisierung**.
- Daten zur **gesundheitsbezogenen Lebensqualität (HRQoL)** während der Behandlung werden ebenfalls berücksichtigt.
- Zusätzlich wird der Parameter **Verträglichkeit**, beurteilt anhand der Rate an unerwünschten Ereignissen während der Behandlung, in die Beurteilung eines medizinischen Zusatznutzens mit einbezogen.

Die Fragestellung lautet daher: Wie sind Ausmaß und Wahrscheinlichkeit des medizinischen Zusatznutzens von Regorafenib hinsichtlich der Endpunkte Mortalität, progressionfreies Überleben, objektive Tumoransprechrates, Krankheitskontrollrate, Dauer der Tumorstabilisierung, gesundheitsbezogene Lebensqualität und Verträglichkeit im Vergleich zu BSC zu bewerten.

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.

Folgende Selektionskriterien zum Einschluss wurden definiert: Patientenpopulation, Intervention, Endpunkte (Zielgrößen), Studientyp, Vergleichstherapie, Studiendauer und Sprache. Sprache und Studiendauer sind dabei keine relevanten Ein- und Ausschlusskriterien.

Tabelle 4-7: Prädefinierte Einschlusskriterien für Studien zur Bewertung des medizinischen Nutzens und Zusatznutzens von Regorafenib

Kürzel	Kategorie	Prädefinierte Kriterien (Einschlusskriterien, sofern nicht anders vermerkt)
E1	Patientenpopulation	<ul style="list-style-type: none"> • Patienten mit metastasiertem kolorektalen Karzinom • O. g. Patienten, die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie.
E2	Intervention	Regorafenib 160 mg einmal täglich für drei Wochen mit anschließend einer Woche Pause
E3	Endpunkte	<p>Mindestens einer der patientenrelevanten Endpunkte wird berichtet:</p> <ul style="list-style-type: none"> • Mortalität: Gesamtüberleben (OS) • Mortalität/Morbidität: progressionsfreies Überleben (PFS) • Morbidität: Objektive Tumoransprechrate (ORR), Krankheitskontrollrate (DCR), Dauer der Tumorstabilisierung • Gesundheitsbezogene Lebensqualität (HRQoL) • Therapiebedingte Morbidität/Verträglichkeit: Rate an unerwünschten Ereignissen (UE)
E4	Studientyp	Abgeschlossene randomisierte, kontrollierte Studien (RCT) Phase II, III oder IV
E5	Vergleichsintervention	BSC, d. h. Therapieoptionen, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirken
E6	Publikationstyp	keine Einschränkung
E7	Studiendauer	keine Einschränkung
E8	Sprache	keine Einschränkung

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 Regorafenib weltweit noch nicht länger als zwölf Monate besteht und keine indirekten Vergleiche herangezogen werden, ist nach den Kriterien der Verfo-GBA eine bibliographische Literaturrecherche in Literaturdatenbanken nicht notwendig und wurde aus diesem Grund nicht durchgeführt.

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.

Die Recherche erfolgte am 13. August 2013 in den Studienregistern clinicaltrials.gov, EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen Pharmnet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>).

Die Suchstrategie und die vollständigen Rechercheergebnisse sind im Anhang 4-B ersichtlich (24).

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 Rechteschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Selektion relevanter Studien erfolgt gemäß der unter 4.2.2. beschriebenen Ein- und Ausschlusskriterien und wird anhand ihrer Einträge in den Registern von zwei Reviewern

unabhängig voneinander durchgeführt. Abweichungen in der Selektion durch beide Reviewer werden im Konsensverfahren aufgelöst. Die jeweilige Webseite der ausgeschlossenen Studien wurde als PDF Format abgespeichert (25) und in eine ris-datei (M4A_Informationsbeschaffung_Studienregister-ausgeschlossene-Studien_Abschnitt4-B.ris) aufgenommen .

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.

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise wurden Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie und separat für jeden patientenrelevanten Endpunkt beschrieben. Zusätzlich wurde die Präzision der Ergebnisse überprüft.

A - Aspekte des Verzerrungspotenzials auf Studienebene

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des Behandlers
- Ergebnigesteuerte Berichterstattung
- Sonstige Aspekte

Auf die Aspekte „zeitliche Parallelität der Gruppen“ und „Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren“ wurde nicht eingegangen, da keine nicht randomisierten vergleichenden Studien eingeschlossen wurden.

B - Aspekte des Verzerrungspotenzials auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-treat (ITT) Prinzips
- Ergebnigesteuerte Berichterstattung
- Sonstige Aspekte

Die Studien sind im Einzelnen mit Hilfe des Bewertungsbogens in Anhang 4-G evaluiert worden.

Das Verzerrungspotenzial wurde als entweder niedrig oder hoch eingestuft. Ein niedriges Verzerrungspotenzial lag vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt waren. Unter einer hohen Verzerrung wurde verstanden, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts wurde zunächst das Verzerrungspotenzial endpunktübergreifend (= auf Studienebene) anhand der oben aufgeführten Aspekte als niedrig

oder hoch eingestuft. Falls bereits dieses als hoch eingestuft wurde, wurde das Verzerrungspotenzial für den Endpunkt i. d. R. auch als hoch bewertet. Ansonsten fanden die unter B genannten endpunktspezifischen Aspekte Berücksichtigung.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als hoch führte nicht zum Ausschluss aus der Nutzenbewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Schlussfolgerungen.

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 selektierten randomisierten kontrollierten Studien (RCT) werden gemäß dem CONSORT-Statement beurteilt. Beim CONSORT-Statement handelt es sich um ein Instrument zur Analyse von RCTs anhand von 25 *items*.

In den ersten sechs *items* soll das Ziel der Studie (*item* 2b), die Studienpopulation (*item* 4), die Interventionen (*item* 5) und die Zielgrößen (*item* 6) mittels entsprechender Kriterien klar definiert sein. Ausschlaggebend für eine Bewertung der Studie hinsichtlich der unverzerrten Aussagekraft sind die angewandten Methoden zur Bestimmung der Fallzahl, die Randomisierung, insbesondere die Erzeugung der Behandlungsfolge und die verdeckte Zuteilung der Patienten zu den Gruppen, die Verblindung und die adäquate Anwendung von statistischen Analysemethoden (*items* 7, 8-12). Da letztere Punkte eine zentrale Größe in der

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.

Beurteilung des Verzerrungspotenzials der Aussagekraft von Studien haben, werden hierzu zusätzlich die vom IQWiG herausgegebenen Methoden herangezogen (26).

In den *items* 13-16 müssen die Ergebnisse nach den Angaben zu der Zahl an geplanten und tatsächlich eingeschlossenen Patienten, zu relevanten Basis-Patientencharakteristika, der grafischen Darstellung des Patientenflusses sowie hinsichtlich Studienabbrüche und evtl. Therapiewechsel dargestellt und transparent beschrieben sein. Die Ergebnisse selbst müssen anschaulich dargestellt und beschrieben sein. Zur Beurteilung der Wirksamkeit ist es erforderlich, dass Angaben der statistischen Analysen vorliegen, die über die Angaben eines p-Wertes hinausgehen, wie z. B. Konfidenzintervalle der Effektschätzer. Weiterhin ist eine Überprüfung auf Vollständigkeit der berichteten Ergebnisse in Bezug auf die geplanten Zielgrößen empfehlenswert. Sofern Subgruppenanalysen durchgeführt wurden, sollten sie dahingehend überprüft werden, inwieweit die Subgruppen bereits bei der Randomisierung bspw. durch Stratifizierung definiert wurden. Abschließend wird beurteilt, inwiefern die dargestellten Kriterien eingehalten wurden. Daraus wird das Verzerrungspotential für die berichteten Endpunkte und für die Studie abgeleitet.

Alle relevanten Informationen sind in Anhang 4-F hinterlegt.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

Patientencharakteristika

Die Studienpopulation umfasst laut den definierten Einschlusskriterien zur Studienselektion Patienten mit metastasiertem kolorektalen Karzinom, die bereits zuvor mit einer Fluoropyrimidin-basierten Chemotherapie, einer Anti-VEGF-Therapie, und, sofern ein KRAS-Wildtyp vorliegt, einer Anti-EGFR-Therapie behandelt wurden, oder für eine solche Therapie nicht infrage kamen.

Bestimmte Patientencharakteristika können prognostische Faktoren für den Krankheitsverlauf darstellen und somit potentiell Ergebnisse von Interventionsstudien beeinflussen. In

Subgruppenanalysen kann der mögliche Einfluss von Patientencharakteristika untersucht werden.

Therapeutisch relevante Charakteristika sind Geschlecht, Alter, ECOG zum Therapiebeginn, Zeit seit der Diagnose der metastatischen Erkrankung, vorherige anti-VEGF-Therapie, vorherige antitumoröse Therapie, Anzahl der Therapielinien, Anzahl der Therapielinien nach Diagnose der metastatischen Erkrankung, KRAS-Mutationsstatus und Primärlokation der Erkrankung. Die Details der potentiellen Effektmodifikatoren finden sich in Kapitel 4.2.5.5.

Patientenrelevante Endpunkte

Zur Bewertung des Ausmaßes des medizinischen Zusatznutzens werden gemäß §3 Abs. 2 und 1 5. Kap. Verfo-GBA Ergebnisse bezüglich der patientenrelevanten Endpunkte Mortalität, Morbidität und Lebensqualität herangezogen.

In diesem Dossier müssen die Morbiditätsendpunkte nicht als Surrogat für das Gesamtüberleben validiert werden, da das Gesamtüberleben der primäre Endpunkt der Studie war und somit bereits enthalten ist. Die Morbiditätsendpunkte werden in dieser Indikation als separate valide patientenrelevante Endpunkte angesehen.

Mortalität: Das mediane Gesamtüberleben (OS) gilt in der Onkologie unbestritten als gesicherter patientenrelevanter Endpunkt. Auch in einem kürzlich veröffentlichten Rapid Report des IQWiG zur Bewertung von Surrogatendpunkten in der Onkologie wird das Gesamtüberleben als herausragendes Therapieziel beschrieben (27). Die Darstellung der Ergebnisse zum Endpunkt Mortalität erfolgt, wie in Studien dieser Art üblich, über eine Überlebenszeitanalyse nach der Kaplan-Meier-Methode ebenso wie durch eine Darstellung von Überlebensraten für bestimmte Zeitintervalle (bspw. 1-Jahres-Überlebensrate).

Morbidität/Mortalität: Das progressionsfreie Überleben, d.h. das Überleben ohne Fortschreiten der Erkrankung, ist für die Indikation kolorektales Karzinom als Surrogatparameter anerkannt (28-30). Es wird ebenfalls über die Kaplan-Meier-Methode bestimmt und beschrieben. Darüber hinaus stellt das progressionsfreie Überleben jedoch einen patientenrelevanten Parameter aus eigenem Recht dar.

Morbidität: Patientenrelevante Morbiditätsparameter sind die Objektive Tumorsprechrage, die Krankheitskontrollrate sowie die Dauer der Tumorstabilisierung. Die Objektive Tumorsprechrage als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum ist die Summe der Patienten, die eine komplette oder partielle Remission erfahren. Die Krankheitskontrollrate als erweitertes Maß für die direkte Wirkung der Therapie auf das Tumorwachstum wird bestimmt anhand der Summe von Patienten, die eine komplette oder partielle Remission oder eine Stabilisierung der Erkrankung unter der therapeutischen Intervention erfahren. Die Dauer der Tumorstabilisierung als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum über die Zeit war definiert als Zeit (in Tagen) von der Randomisierung bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression auftrat). Einbezogen wurden alle Patienten, die keine Partielle Remission (PR)

oder Komplette Remission (CR) erreichten. Ausgeschlossen wurden Patienten, die keine Tumorevaluierung nach Studienbeginn hatten.

Gesundheitsbezogene Lebensqualität: Die gesundheitsbezogene Lebensqualität, gemessen mit dem Instrument EORCT QLQ-C30, ist ein akzeptierter patientenrelevanter Endpunkt; die eingesetzten Fragebögen des EORCT QLQ-C30 sind validiert und werden routinemäßig in der Onkologie eingesetzt.

Unerwünschte Ereignisse/Verträglichkeit: Die Verträglichkeit (i. S. UE, UE, die zum Studienabbruch geführt haben, und SUE) ist ein patientenrelevanter Endpunkt.

Die oben genannten Endpunkte werden für die Bewertung eines größeren Nutzens bzw. Schadens herangezogen.

Patientenrelevanz von PFS und Parametern der Tumorkontrolle (Objektive Tumorsprechrate, Krankheitskontrollrate, Dauer der Tumorstabilisierung)

Sowohl Patienten als auch Ärzte sehen eine Tumorerkrankung als potentiell tödliche Erkrankung an (31-33). Daher beginnen sie, wenn eine Heilung nicht möglich erscheint, eine Therapie, die durch eine Reduktion der Tumorgöße eine Progression verhindern soll in der Hoffnung, dass dadurch das Leben des Patienten verlängert und das Leiden, das mit einer kontinuierlichen Tumorprogression einhergeht, vermindert wird.

Wenn ein Tumor fortschreitet, ist dies i. d. R., wenn auch nicht immer sofort, mit einer Verkürzung der Überlebenszeit, einer Verschlechterung der physischen Leistungsfähigkeit, und/oder einer Verschlechterung des körperlichen und seelischen Befindens sowie der Lebensqualität verbunden. Die Verlängerung des progressionsfreien Überlebens, aber auch ein Ansprechen des Tumors bzw. eine Stabilisierung der Erkrankung haben dabei weniger fassbare Auswirkungen wie bei den oben genannten, da sie den emotionalen Stress, der mit der dauernden Angst vor einem Fortschreiten der Erkrankung assoziiert ist, vermindern. Der Wert dieser Parameter hängt natürlich sehr stark vom Ausmaß der Verzögerung der Progression bzw. des Ausmaßes und der Dauer des Tumorsprechens, der Verträglichkeit der angewandten Medikation sowie der Erwartung des Patienten ab.

Eine Therapie, die effektiv und über eine relevante Zeit eine Tumorprogression verhindert, ohne einen negativen Effekt auf das Überleben durch eine intrinsische Toxizität zu haben, mündet i. d. R. auch in einen Überlebensvorteil. Genauso sollte eine Verlängerung der Tumorkontrolle ohne zusätzliche Toxizität zu einer Verbesserung der krebsbedingten Symptome und der Lebensqualität führen (4, 5). Es ist daher nachzuvollziehen, dass eine länger andauernde Kontrolle des Tumorwachstums zu positiven Effekten auf das physische und psychische Wohlbefinden des Patienten führt und auch im Interesse des Gesundheitswesens liegt. Dies wird bestätigt durch Studien, die belegen, dass der Status der Progressionsfreiheit ein Wert für die Patienten an sich darstellt, der mit einer höheren Lebensqualität und einem besseren Gesundheitszustand einhergeht (6-8). In der Studie von Siena et al. (7) wurde die Assoziation von PFS mit Symptomen, die durch Darmkrebs bedingt

sind, mit der Lebensqualität und dem Gesamtüberleben untersucht. Das Fehlen einer Krankheitsprogression war mit einer signifikant und klinisch relevant niedrigeren Symptomschwere und einer höheren gesundheitsbezogenen Lebensqualität assoziiert als bei Patienten mit einer Progression der Erkrankung. Thong et al. (8) verglichen die Lebensqualität von Patienten mit und ohne Krankheitsprogression anhand der validierten Fragebogen SF-36 und QOL-CS. Im Ergebnis wurde festgestellt, dass Patienten mit einer Progression sowohl einen schlechteren Gesundheitsstatus als auch eine schlechtere Lebensqualität aufweisen. In der Studie von Hurvitz et al. (6) wurde insbesondere die psychische Befindlichkeit untersucht. Frauen mit Brustkrebs wurden zu zwei hypothetischen medikamentösen Behandlungen befragt; die eine sei assoziiert mit einem durchschnittlichen PFS von zwölf Monaten, die andere mit einem von 16 Monaten. Das Nebenwirkungsprofil und das Gesamtüberleben wären bei beiden Therapien gleich. 88 % der befragten Patientinnen bevorzugten die Therapie mit dem längeren PFS. Eine weitere Erhebung des psychischen Befindens in dieser Studie bei Patientinnen, deren Erkrankung progredient war versus denen ohne Progression ergab, dass erstere mehr in Apathie und Interessenlosigkeit fallen, fürchten, dass die Therapie nicht wirkt, und sich sorgen, dass sie ihren alltäglichen Aufgaben nicht nachkommen können.

In engem Zusammenhang mit dem progressionsfreien Überlebensstatus steht das **Tumoransprechen und die Tumorstabilisierung** bzw. das Nicht-Fortschreiten der Erkrankung, auf dem das Kriterium PFS zum Teil basiert.

Es ist mehrfach publiziert worden, dass Patienten, die schlechte Prognosefaktoren für ihre Tumorerkrankung aufweisen, eine schnellere Progression erleiden als Patienten mit guten Prognosefaktoren. Gemeinhin bestimmt die Geschwindigkeit der Tumorprogression, in welchem Ausmaß der Patient unter Krebs-bedingten Symptomen leidet und wie lange er noch zu leben hat. Das heißt, dass i. d. R. die Verkürzung der Lebenserwartung bei Patienten mit Krebs auf die Progression des Tumors zurückgeführt werden kann (34-36).

Viele Patienten und ihre Behandler beurteilen es als Vorteil, wenn eine schnelle und effektive Tumorkontrolle erfolgt und eine Progression des Tumors verzögert werden kann, und zwar unabhängig von einer etwaigen nachgewiesenen Verlängerung der Gesamtüberlebenszeit (4, 12, 13). Dies trifft insbesondere auf schnell fortschreitende Tumoren oder solche zu, die ein lebenswichtiges Organ bedrohen. In einer Metaanalyse von 21 Studien in verschiedenen Tumorentitäten wurde die Korrelation von radiologischem Tumoransprechen und der gesundheitsbezogenen Lebensqualität untersucht (14). Ein objektives Tumoransprechen ist demnach signifikant mit einem Symptomansprechen assoziiert. Auch Patienten, die kein objektives Tumoransprechen im Sinne einer kompletten oder partiellen Remission, sondern eine Stabilisierung der Erkrankung erfuhren, hatten ein Symptomansprechen. Die medikamentöse Tumorthherapie hatte somit den größten Effekt für die Steigerung der gesundheitsbezogenen Lebensqualität bei Patienten, die eine komplette oder partielle Remission erfuhren, gefolgt von denen, die eine Stabilisierung ihrer Erkrankung als bestes Ansprechen aufwiesen.

Der Parameter **progressionsfreies Überleben** misst direkt die Effekte der jeweiligen Behandlung auf das Tumorwachstum unter Benutzung einer objektiven und quantitativen Methodik. Dabei werden sowohl zytostatische (Tumorstabilisierung) als auch zytotoxische (Tumorrückgang) Effekte gemessen. Daher sollten im Kontext einer positiven Risiko-Nutzen-Abschätzung eine klinisch relevante Verbesserung des PFS sowie von Parametern der Tumorkontrolle (Objektive Tumoransprechrates, Krankheitskontrollrate, Dauer der Tumorstabilisierung) patientenrelevante Nutzenparameter mit einem eigenen Stellenwert darstellen.

Dass zumindest PFS einen wertvollen Endpunkt für klinische Studien mit antitumorös wirkenden Substanzen darstellt, haben auch die meisten Arzneimittel-Zulassungsbehörden anerkannt und entsprechende Richtlinien zu dessen Verwendung herausgegeben (9-11). Mehrere Arzneimittel sind auf der Basis von signifikanten Verbesserungen im PFS zugelassen worden. Beispiele sind die Zulassung von Sorafenib und Pazopanib beim metastasierten Nierenzellkarzinom, und Gemcitabin in Kombination mit Carboplatin zur Therapie des fortgeschrittenen Ovarialkarzinoms. Auch wenn die Erhebung von PFS mit einigen Nachteilen verbunden ist, haben insbesondere in Indikationen, in denen es sequentielle Therapien gibt, die dadurch erreichten stufenweisen Verbesserungen in der Tumorkontrolle dazu geführt, dass die Krebserkrankung über längere Zeit kontrollierbar wird.

Primärer Endpunkt der eingereichten Studie (CORRECT) war das Gesamtüberleben. Das Gesamtüberleben ist aufgrund seiner eindeutigen Definition der Art und des Zeitpunktes des Ereignisses sowie des Fehlens von methodischen Unschärfen ein direktes und zuverlässiges Maß zur Darstellung eines Therapieeffektes. Sowohl die Sicherheit als auch die Wirksamkeit der zu untersuchenden Therapie fließt in diesen Endpunkt ein. Die Wichtigkeit des Gesamtüberlebens steht außer Frage, und es ist daher als valider Endpunkt anerkannt. Jedoch hat dieser Parameter in Bezug auf das zu untersuchende Arzneimittel auch Nachteile, indem er durch nachfolgende Therapien beeinflusst wird und auch Todesursachen mit berücksichtigt werden, die mit der eigentlich zu untersuchenden Erkrankung nicht in Verbindung stehen (sogenannte *confounding factors*).

Daher ist es sinnvoll, mit zusätzlichen Endpunkten die Verlängerung der Lebenszeit in einen biologischen Kontext zu bringen und patientenrelevante Therapieziele, wie die Reduktion der Morbidität bzw. der physischen und psychischen Belastung durch die Tumorerkrankung, sowie eine Stabilisierung des Patienten über einen relevanten Zeitraum durch Verhinderung des weiteren Tumorwachstums mit zu betrachten. Viele Patienten sind auch oder gerade bei einer infausten Erkrankung bereit, für einen kleinen Gewinn an Lebenszeit oder Lebensqualität eine antitumoröse Therapie mit den damit verbundenen Toxizitäten auf sich zu nehmen, und zwar für einen kleineren Unterschied, als Gesunde oder die sie betreuenden Personen annehmen würden (13).

Dem progressionsfreien Überleben kommt hierbei eine duale Rolle zu: Es ist ein direkter Parameter der Krankheitslast und ein Endpunkt, für den eine Korrelation mit dem Gesamtüberleben bei bestimmten Tumorerkrankungen validiert wurde. So wurde beim fortgeschrittenen kolorektalen Karzinom die Validität von PFS als Surrogatendpunkt für das

Gesamtüberleben nachgewiesen (18, 28-30, 37, 38). Auch für die Tumoransprechrates wurde eine Korrelation mit dem Gesamtüberleben bei kolorektalem Karzinom nachgewiesen (17, 18).

Johnson et al. (18) benutzten die Daten aus 146 randomisierten Studien in der Erstlinientherapie des kolorektalen Karzinoms, um sie mittels linearer Regression zu analysieren. Sowohl das Tumoransprechen als auch die Zeit bis zur Progression korrelierten mit einer Verbesserung des Gesamtüberlebens beim kolorektalem Karzinom ($p < 0,0001$ für beide Parameter). Für die Endpunkte mediane Zeit bis zur Progression und Tumoransprechrates wurden Schwellenwerte für bestimmte Stichprobengrößen ermittelt, die für einen signifikanten Unterschied im Gesamtüberleben prädiktiv sein würden (vgl. Tabelle 4-8).

Tabelle 4-8: Für das Gesamtüberleben prädiktive Schwellenwerte für Effektgrößen von Veränderungen in Wirksamkeitsparametern, nach Stichprobengröße (nach Johnson et al.)

	Stichprobengröße		
	250 Patienten	500 Patienten	750 Patienten
Unterschied im Tumoransprechen (%)	38 %	26 %	20 %
Unterschied in der Zeit bis zur Progression (Monate)	3,2	2,3	1,9

Quelle: (18)

Buyse et al. (37) nahmen die Daten von zehn historischen klinischen Studien, die entweder 5-Fluorouracil (5-FU) in Kombination mit Leucovorin (LV) mit entweder 5-FU allein (N=1.744) oder mit 5-FU in Kombination mit Raltitrexed (N=1.345) verglichen. Als Validierungsstudien dienten drei Studien zum Vergleich von 5-FU + LV mit oder ohne Irinotecan oder Oxaliplatin (N=1.263). Es wurden die Korrelationskoeffizienten sowohl zwischen den Endpunkten PFS und OS als auch zwischen den Behandlungseffekten auf diese Endpunkte in den historischen Kontrollen bestimmt. Die Behandlungseffekte auf das OS wurden für die Validierungsstudien vorausberechnet, und schließlich mit den beobachteten Effekten verglichen. Als Ergebnis der Korrelationsanalysen ergab sich, dass eine Hazard Ratio von 0,77 oder weniger für den Parameter PFS prädiktiv für einen Vorteil im Gesamtüberleben wäre (37).

Tang et al. (38) identifizierten 39 randomisierte Studien mit mindestens 100 Patienten pro Arm, die bei Patienten zur Erstlinientherapie des metastasierten kolorektalen Karzinoms durchgeführt wurden. Korrelationsanalysen für die Endpunkte PFS, TTP und Ansprechrates mit dem Gesamtüberleben wurden durchgeführt. Es fanden sich folgende, zum Teil sehr ausgeprägte Korrelationen zwischen den Wirksamkeitsendpunkten und dem Gesamtüberleben (Tabelle 4-9) (38).

Tabelle 4-9: Korrelation zwischen Surrogatendpunkten und Gesamtüberleben (nach Tange et al. (2009))

Surrogatendpunkt	Nicht-parametrischer Spearman Rank Korrelationskoeffizient	95 %-KI	p-Wert
PFS vs. OS	0,79	0,65 – 0,87	<0,000001
TTP vs. OS	0,24	-0,13 – 0,55	0,21
ORR vs. OS	0,59	0,42 – 0,72	<0,000001
Δ PFS vs. Δ OS ^a	0,74	0,47 – 0,88	0,00004
Δ TTP vs. Δ OS ^a	0,52	0,004 – 0,81	0,05
Δ ORR vs. Δ OS ^a	0,39	0,08 – 0,63	0,015

^aUnterschied zwischen Behandlungsarmen innerhalb einer Studie
ORR: objektive Tumoransprechrates, OS: Gesamtüberleben, PFS: progressionsfreies Überleben, TTP: Zeit bis zur Progression

Quelle: (38)

Eine Korrelationsanalyse der Hazard Ratios von PFS und OS ergab, dass der Abfall der Regressionslinie 0,54 (\pm 0,10) betrug. Dies bedeutet, dass eine neue Therapie, die eine 10%ige Risikoreduktion im PFS erreicht, eine Risikoreduktion um 5,4 % (\pm 1 %) im Gesamtüberleben bewirkt.

Wenn man die Ergebnisse der CORRECT-Studie (siehe Abschnitt 4.3.1.3) mit den Ergebnissen dieser drei Arbeiten in Bezug setzt, ergibt sich Folgendes:

Die von Johnson et al. gefundene Korrelation von Tumoransprechen und Gesamtüberleben bei einer Stichprobengröße von 750 von 20 % wird in der CORRECT-Studie (n=760) nicht erfüllt (Unterschied 0,6 %). Tumoransprechen nach RECIST- oder WHO-Kriterien steht für viele zielgerichtete Arzneimittel im Allgemeinen und für Multikinaseinhibitoren im Besonderen nicht im Vordergrund ihrer biologischen Wirkung. Vielmehr geht es um eine Tumorstabilisierung und Veränderungen im Tumor, in der Tumorumgebung bzw. in den einen Tumor versorgenden Blutgefäßen, die günstige Wirkungen auf die Tumorumgebung bewirken. Daher werden für diese Arzneimittel andere Methoden des Wirksamkeitsnachweises diskutiert und angewandt (19-22).

Der Parameter TTP ist in der CORRECT-Studie nicht mitgeführt worden; daher ist eine Einordnung nicht möglich. Der Parameter TTP hat in der Analyse von Tang et al. die schwächste Korrelation mit OS (Korrelationskoeffizient 0,24, p-Wert 0,21), nur die Differenz Δ TTP vs. Δ OS erreicht eine stärkere und signifikante Korrelation (Korrelationskoeffizient 0,52, p-Wert 0,05). Die Überlegenheit von PFS gegenüber TTP ist höchstwahrscheinlich durch die Überlappung von PFS und OS bedingt, da bei PFS im Gegensatz zu TTP zusätzlich zur Progression auch die Todesfälle jeglicher Ursache mit einbezogen werden.

Der von Buyse et al. (37) gefundene Schwellenwert von 0,77 in der HR von PFS wird in der CORRECT-Studie deutlich unterschritten (HR=0,494).

Nach Tang et al. sollte eine 10%ige Risikoreduktion im PFS eine Risikoreduktion um 5,4 % (± 1 %) im Gesamtüberleben zur Folge haben. In der CORRECT-Studie gibt es eine ca. 50%ige Reduktion des Risikos für das PFS, was Risikoreduktion um eine ca. 27 % (± 5 %) im OS bewirken sollte. Tatsächlich beträgt die Risikoreduktion für das Gesamtüberleben 22,6 % (HR=0,774) und liegt damit im kalkulierten Bereich.

Somit bestätigen die Ergebnisse der CORRECT-Studie zumindest die Korrelation zwischen PFS und OS, so wie sie von Buyse et al. (37) und Tang et al. (38) gefunden wurden.

Fazit:

1. Progressionsfreies Überleben ist beim fortgeschrittenen kolorektalen Karzinom, insbesondere in einer Dritt- und Viertliniensituation in einer Phase III-Studie nicht nur ein valider Surrogatparameter für das Gesamtüberleben, sondern stellt einen patientenrelevanten Endpunkt aus eigenem Recht dar.
2. In dieser Indikation ist das PFS als kombinierter Mortalitäts- und Morbiditätspunkt ein klinisch-medizinischer, sozialmedizinischer und auch für den Patienten sehr relevanter Parameter, der einen beträchtlichen Nutzen hinsichtlich Morbidität und Lebensqualität an sich darstellt.
3. Das Tumoransprechen im Sinne von Remissionen und Tumorstabilisierungen ist ebenfalls ein patientenrelevanter Parameter für Patienten, indem es die physische und psychische Morbiditätslast senkt.

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.

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

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.

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

4.2.5.4 Sensitivitätsanalysen

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

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

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

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.

Für den primären Wirksamkeitsendpunkt Gesamtüberleben und den sekundären Wirksamkeitsendpunkt PFS wurden zwei Sensitivitätsanalysen für den Schnittpunkt 21. Juli 2011 prospektiv geplant und durchgeführt. Die erste Sensitivitätsanalyse wurde auf Basis unstratifizierter Daten durchgeführt, die zweite mit der Stratifizierung gemäß Sprachdialogsystem (interactive voice response system (IVRS)). Grund war die Misstratifizierung einiger Patienten bei der Randomisierung. Die Analysen wurden identisch zur primären Wirksamkeitsanalyse auf Basis eines Cox-Regressionsmodells berechnet, wobei Hazard Ratios mit 95 %-Konfidenzintervallen sowie einseitige log-rank-Tests berechnet wurden.

Es wurden keine weiteren Sensitivitätsanalysen durchgeführt.

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.

In der CORRECT-Studie wurden *a priori* im statistischen Analyseplan definierte Subgruppenanalysen durchgeführt, um den Einfluss potentieller Effektmodifikatoren auf den Behandlungseffekt bezüglich der Wirksamkeitsendpunkte des medianen Gesamt- und des medianen progressionsfreien Überlebens zu untersuchen. Für alle anderen beschriebenen Endpunkte wurden die Subgruppenanalysen *post hoc* durchgeführt.

Tabelle 4-10: In der CORRECT-Studie geplante Subgruppenanalysen – RCT mit dem zu bewertenden Arzneimittel

Subgruppe	Subgruppenlevel
Geschlecht	Männlich, weiblich
Alter (Jahre)	< 65, ≥ 65
Ethnische Zugehörigkeit	Kaukasier, Asiatisch
ECOG bei Einschluss	0, 1
Zeit seit der Diagnose der metastasierten Erkrankung	< 18 Monate, ≥ 18 Monate
Region ^a	1, 2, 3
Vorherige Anti-VEGF-Therapie	Ja, nein
Vorherige antitumoröse Therapie	FOIB, FOIBE
Anzahl vorheriger Therapielinien	≤3, >3
Anzahl vorheriger Therapielinien nach Diagnose der metastasierten Erkrankung	≤3, >3
KRAS-Mutation	Nein, ja
Primärer Ort der Erkrankung	Kolon, Rektum, Kolon und Rektum
^a Region 1: Nordamerika, Westeuropa, Israel, Australien; Region 2: Asien; Region 3: Südamerika, Türkei, Osteuropa FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper	

Die Ergebnisse aus den Subgruppenanalysen sollten wegen der zum Teil kleinen Stichproben mit Vorsicht interpretiert werden, da die Punktschätzer bei sehr kleinen Gruppengrößen stark schwanken können. Die Analysen pro Endpunkt waren vom gleichen Typ wie die Analysen der Gesamtpopulation.

Biomarkeranalysen: In der CORRECT-Studie wurden drei genetische und 15 nicht-genetische Biomarker analysiert. Da sowohl Ergebnisse präklinischer Studien als auch die

von Phase I- und II-Studien keine starken Hinweise auf einen prädiktiven Biomarker in Bezug auf die populationsspezifische Wirkung von Regorafenib lieferten, war die Zulassungsstudie CORRECT nicht spezifisch dafür gepowert, Biomarkerfragen zu adressieren. Die Biomarkeranalysen in dieser Studie wurden retrospektiv durchgeführt, und werden als explorativ und Hypothesen-generierend angesehen. Somit wurden in den Analysen die p-Werte nicht in Bezug auf multiples Testen korrigiert, um nicht eine potentiell wichtige Korrelation zu übersehen. Daher werden p-Werte $<0,05$ formal nicht als statistisch signifikant angesehen, sondern sollten als potentiell interessantes Signal betrachtet werden.

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. Lüthmann, 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.*

Indirekte Vergleiche wurden nicht durchgeführt, da die zweckmäßige Vergleichstherapie, (BSC) in der randomisierten, kontrollierten Studie CORRECT 14387 als Vergleichsmedikation in der Kontrollgruppe verwendet wurde und somit direkte Vergleiche vorliegen.

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 4-11: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
CORRECT 14387	Ja	Abgeschlossen	April 2010 (FPFV) – Juli 2011 (LPLV)	Regorafenib + BSC, Placebo + BSC,
CONCUR 15808	Ja*	Laufend	April 2012 (FPFV) Mai 2014 (LPLV-geplant)	Regorafenib, Placebo

*zulassungsrelevante Studie für Asien, insbesondere für China

Geben Sie an, welchen Stand die Information in Tabelle 4-11 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.

Die Information der Tabelle 4-11 bildet den Studienstatus mit Datum 26.08.2013 ab.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-12 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-12: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
CONCUR 15808	Studie begann Rekrutierung im April 2012. Es liegen noch keine Ergebnisse für diese Studie vor. Zulassungsstudie insbesondere für China.

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

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

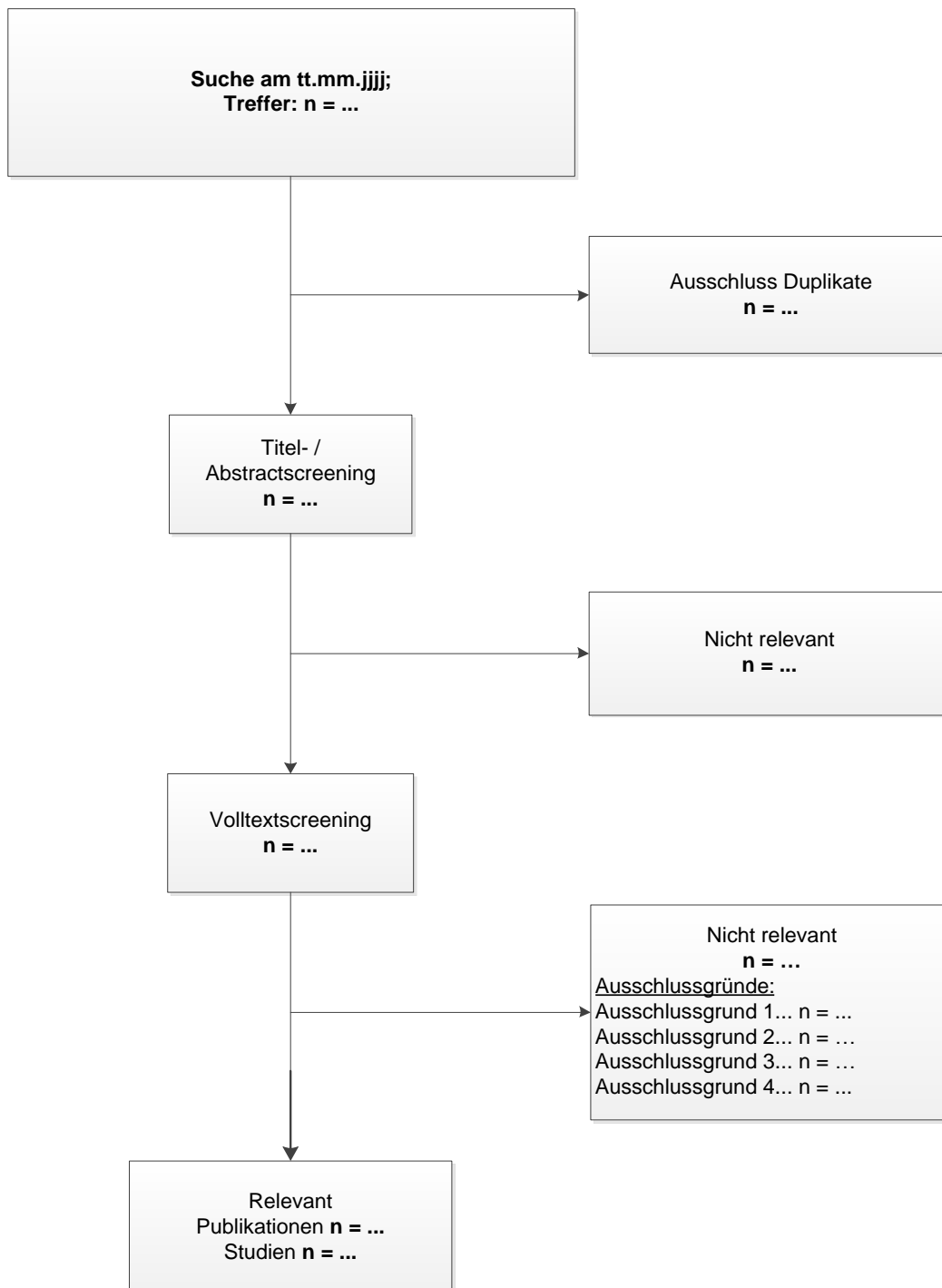


Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Nicht zutreffend: Da eine Zulassung von Regorafenib weltweit noch nicht länger als zwölf Monate besteht und keine indirekten Vergleiche herangezogen werden, ist nach den Kriterien der VerfO des G-BA eine bibliographische Literaturrecherche in Literaturdatenbanken nicht notwendig und wurde aus diesem Grund nicht durchgeführt.

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 4-11) 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 4-13: 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)
CORRECT 14387	Clinicaltrials.gov (NCT01103323) – Studienregistereintrag (39)	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 4-14: 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])
Placebokontrolliert						
CORRECT 14387	Ja	Ja	Nein	Ja (40)	Ja (39)	Ja (16)
Aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Nicht zutreffend.						
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 4-15: 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
CORRECT 14387	RCT, doppelblind, parallel	Erwachsene mit metastasiertem kolorektalen Karzinom (Stadium IV)	Regorafenib + BSC: (n=505) Placebo + BSC (n=255)	Behandlung bis zur Progression oder dem Tod, Nachbeobachtung bis zum Tod durch monatliche Nachverfolgung via Telefon	Multinational (114 Zentren in 16 Ländern auf 4 Kontinenten) 04/2010 – 07/2011 (FPFV-LPLV)	Gesamtüberleben; progressionsfreies Überleben, Objektive Tumoransprechrates, Krankheitskontroll- rate, Dauer der Krankheitsstabilis- ierung, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse/Verträglich- keit

Tabelle 4-16: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Regorafenib + BSC	Placebo + BSC
CORRECT 14387	Regorafenib 160 mg (4 Tbl. à 40 mg) einmal täglich für drei Wochen, gefolgt von einer Woche Pause + BSC	Placebo 4 Tbl. einmal täglich für drei Wochen, gefolgt von einer Woche Pause + BSC

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Patientencharakteristikum	Regorafenib + BSC N=505	Placebo + BSC N=255
Geschlecht, n (%)		
Männlich	311 (61,6)	153 (60,0)
Weiblich	194 (38,4)	102 (40,0)
Alter (Jahre) bei Einschluss		
Mean (Bereich)	60,7 (22-82)	60,1 (25-85)
Median	61,0	61,0
Altersgruppen, n (%)		
<65	309 (61,2)	166 (65,1)
≥65	196 (38,8)	89 (34,9)
Mody Mass Index (kg/m ²)		
Mean (Bereich)	25,405 (14,36-47,33)	26,339 (13,92-50,36)
Median	24,82	25,65
Region, n (%)		
Nordamerika, Westeuropa, Israel, Australien	420 (83,2)	212 (83,1)
Asien	69 (13,7)	35 (13,7)
Südamerika, Türkei, Osteuropa	16 (3,2)	8 (3,1)
Ethnische Gruppe, n (%)		
Kaukasier	392 (77,6)	201 (78,8)
Schwarze oder Afroamerikaner	6 (1,2)	8 (3,1)
Asiaten	76 (15,0)	35 (13,7)
Amerikanische Indianer oder Ureinwohner Alaskas	1 (0,2)	1 (0,4)
Nicht berichtet	29 (5,7)	10 (3,9)
Multiple	1 (0,2)	0
ECOG Performance Status vor Therapiebeginn, n (%)		
0	265 (52,5)	146 (57,3)
1	240 (47,5)	109 (42,7)
Histologie, n (%)		
Adenokarzinom in situ	2 (0,4)	3 (1,2)
Adenokarzinom	493 (97,6)	245 (96,1)
Mucinöses Karzinom	5 (1,0)	4 (1,6)
Adenosquamöses Karzinom	1 (0,2)	1 (0,4)
Undifferentiertes Karzinom	0	1 (0,4)
Karzinom, nicht näher bezeichnet	4 (0,8)	1 (0,4)
Lokalisation des Primärtumors, n (%)		
Kolon	323 (64,0)	172 (67,5)
Rektum	151 (29,9)	69 (27,1)
Kolon und Rektum	30 (5,9)	14 (5,5)

Patientencharakteristik	Regorafenib + BSC N=505	Placebo + BSC N=255
Nicht angegeben	1 (0,2)	0
Stadium bei Studieneintritt, n (%)		
Stadium IV	505 (100,0)	255 (100,0)
Vorherige Therapie mit Anti-VEGF Medikamenten, n (%)		
Ja	505 (100,0)	255 (100,0)
Nein	0	0
Zeit von der Erstdiagnose der metastasierten Erkrankung bis Randomisierung, Wochen		
Mean (Bereich)	151,7 (18,1 – 837)	150,3 (10,4 – 553)
Median	133,1	128,5
Zeit von der Erstdiagnose der metastasierten Erkrankung bis Randomisierung, kategorisiert, n (%)		
<18 Monate	91 (18,0)	49 (19,2)
≥18 Monate	414 (82,0)	206 (80,8)
Zeit von der letzten Progression der Erkrankung bis zur Randomisierung (Wochen)		
Mean (Bereich)	6,46 (0,1 – 50,0)	6,16 (0,3 – 52,1)
Median	4,99	4,56
Unbekannt	30 (5,9)	9 (3,5)
KRAS-Mutation, n (%)		
Nein	205 (40,6)	94 (36,9)
Ja	273 (54,1)	157 (61,6)
Unbekannt	27 (5,3)	4 (1,6)
BRAF-Mutation, n (%)		
Nein	41 (8,1)	25 (9,8)
Ja	4 (0,8)	2 (0,8)
Unbekannt	460 (91,1)	228 (89,4)
Quelle: CORRECT 14387 Studienbericht: Tabelle 8-5, 8-6, 14.1/10, 14.1/11		

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

Die CORRECT-Studie ist eine multinationale, multizentrische, randomisierte, doppelblinde, zweiarmige, Placebo-kontrollierte Phase-III-Studie, welche in 114 Zentren in 16 Ländern in Europa, Asien, Nord- und Südamerika und Australien durchgeführt wurde. In 105 Zentren wurde mindestens ein Patient eingeschlossen, in Südamerika und in der Türkei wurde kein Patient eingeschlossen. In Deutschland wurden in 15 Zentren 64 Patienten in die Studie eingeschlossen. Untersucht wurde die Wirksamkeit und Verträglichkeit von Regorafenib + BSC im Vergleich zu Placebo + BSC bei erwachsenen Patienten mit einem

mCRC, deren Erkrankung nach Erhalt aller zugelassenen Standardtherapeutika progredient war. Die Patienten mussten bei Einschluss eine Tumorprogression während oder binnen drei Monaten nach der letzten Gabe einer zugelassenen Standardtherapie aufweisen. Die Standardtherapie musste Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab und – im Falle eines KRAS-Wildtyps – Cetuximab oder Panitumumab umfassen. Die geplante Zahl zu randomisierender Patienten lag bei 690. Die Patienten wurden nach vorangegangener VEGF-basierter Therapie (ja/nein), der Zeit seit der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate) sowie nach der geographischen Region (Region 1: Nordamerika, Westeuropa, Israel, Australien, vs. Region 2: Asien, vs. Region 3: Südamerika, Türkei, Osteuropa) stratifiziert. Für die Studie wurden 1.052 Patienten gescreent, von denen 760 Patienten randomisiert wurden. Alle randomisierten Patienten bilden die Intention-to-Treat-Population (ITT). Die Safety-Population umfasst 753 Patienten und ist definiert als alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben.

Die Zuteilung zu den Behandlungsarmen erfolgte randomisiert im Verhältnis 2:1 (Regorafenib + BSC zu Placebo + BSC).

Die jeweils den Behandlungsarmen zugeordneten Patienten erhielten entweder 160 mg (vier Tabletten zu 40 mg) Regorafenib einmal täglich für drei Wochen, gefolgt von einer Woche Pause, oder vier Placebotabletten einmal täglich für drei Wochen, gefolgt von einer Woche Pause. Die Studienmedikation wurde morgens mit ca. 240 ml Flüssigkeit nach einem fettarmen Frühstück ($< 30\%$ Fettanteil) eingenommen. In beiden Armen wurde zusätzlich BSC angewandt. BSC umfasste jedwede Medikation oder Behandlung, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und eine Verbesserung der Lebensqualität bewirkt, wie z. B. Antibiotika, Analgetika, Strahlentherapie zur Schmerzlinderung (auf Knochenmetastasen beschränkt), Kortikosteroide, Transfusionen, Psychotherapie, Wachstumsfaktoren, palliative Chirurgie. Ausgenommen waren andere medikamentöse Tumorthérapien (experimentell oder zugelassen).

Die Patienten wurden bis zur Krankheitsprogression behandelt, oder bis eines der folgenden Kriterien eintrat:

- Abbruch auf eigenen Wunsch des Patienten,
- wenn nach Meinung des Prüfarztes ein weiterer Verbleib in der Studie für den Patienten nachteilig wäre,
- bei Auftreten schwerer allergischer Reaktionen,
- bei Auftreten anderer unerwünschter Ereignisse, die als so schwer angesehen wurden, dass sie die Beendigung der Studienbehandlung erforderlich machten,
- bei substantieller Non-Compliance des Patienten,
- bei Patientinnen, die einen positiven Schwangerschaftstest im Verlauf der Studie aufwiesen,
- bei Gebrauch illegaler Drogen oder Substanzen, die nach Einschätzung des Prüfarztes entweder eine Unverträglichkeit auslösen oder anderweitig zur Beeinflussung der Studienergebnisse beitragen würden,

- bei Auftreten einer Zweitneoplasie,
- Patienten, die aus der Beobachtung verloren gingen (Loss to follow-up),
- Patienten, die mehr als 28 aufeinanderfolgende Tage die Einnahme der Studienmedikation unterbrechen mussten,
- wenn mehr als zwei Dosisreduktionen der Studienmedikation notwendig waren (Dosismodifikationen folgten prädefinierten Dosisreduktionsstufen und waren im Prüfplan ausführlich dargelegt)

Primärer Endpunkt der Studie war das Gesamtüberleben (OS). Sekundäre Endpunkte waren das progressionsfreie Überleben (PFS), die Objektive Tumoransprechrate, und die Krankheitskontrollrate. Tertiäre Endpunkte umfassten die Dauer des Tumoransprechens und der Tumorstabilisierung, gesundheitsbezogene Lebensqualität gemessen anhand der Instrumente EORTC QLQ-C30 und EQ-5D, die Pharmakokinetik von Regorafenib sowie die explorative Evaluierung von Biomarkern, anhand von drei genetischen und 15 nicht-genetischen Parametern. Die Sicherheitsbewertung basierte auf der Anamnese, Vitalparametern, körperlicher Untersuchung, Überprüfung der Begleitmedikation und unerwünschter bzw. schwerwiegender unerwünschter Ereignisse, einschließlich der Laborparameter sowie kardialer Parameter aus EKG, Echokardiographie oder MUGA-Scan.

Die Studie bestand aus einer Screeningperiode (28 Tage bis einen Tag vor Beginn des ersten Zyklus⁴), der Behandlungsperiode (Zyklen von vier Wochen Dauer), einer Nachbeobachtungsperiode zur Sicherheit (bis 30 Tage nach der Visite zum Ende der Therapie), sowie einer Nachbeobachtungsperiode (einmal monatlich per Telefonkontakt bis Studienende). Innerhalb eines jeden Zyklus waren zwei Visiten im Studienzentrum vorgesehen (Tag 1 und Tag 15). Hier wurden Sicherheitsbewertungen anhand der Anamnese, körperlicher und laborchemischer Untersuchungen sowie Untersuchungen mittels EKG und Echokardiographie bzw. MUGA-Scan vorgenommen. Eine genaue Vorgabe, wann welche Untersuchung bzw. Maßnahme vorzunehmen war, enthielt der Prüfplan. Zusätzlich wurden bei ungefähr 150 Patienten Blutproben zur Bestimmung pharmakokinetischer Parameter gewonnen. Darüber hinaus waren die Patienten aufgefordert, jeweils zu Beginn der Visite am Tag 1 eines jeden Zyklus die Fragebogen zur gesundheitsbezogenen Lebensqualität EORTC QLQ-C30 sowie EQ-5D auszufüllen.

Das Data Monitoring Committee (DMC), besetzt mit drei Experten aus der medizinischen Onkologie und einem Biostatistiker, evaluierte die Sicherheitsdaten während der Studie in regelmäßigen Intervallen. Primärer Zweck des DMC war die Gewährleistung der Sicherheit der Studienteilnehmer in der Studie. Das DMC überprüfte ebenfalls die Ergebnisse der beiden geplanten Zwischenanalysen. Die erste Zwischenanalyse, durchgeführt nach 175 Todesfällen (ca. 30 % der benötigten Gesamtanzahl von 582 Todesfällen) diente ausschließlich der Überprüfung auf Vergeblichkeit (Futility). Die zweite geplante Zwischenanalyse nach 408 Todesfällen (ca. 70 % der benötigten Todesfälle) diente der Prüfung auf Vergeblichkeit sowie Wirksamkeit. Das Cut-off-Datum für die zweite geplante Zwischenanalyse war der 21. Juli 2011. Wären die vorgegebenen Monitoring-Grenzen für das Gesamtüberleben nicht überschritten worden, wären die Gesamtüberlebensdaten als unreif angesehen worden, und die Studie wäre

bis zur geplanten Todesfallzahl von ca. 582 fortgeführt worden. Da die Studie jedoch bei der zweiten geplanten Zwischenanalyse positiv in den Daten zur Gesamtüberlebenszeit ausfiel, wurde sie zu diesem Zeitpunkt gestoppt. Nachdem der primäre Endpunkt der Studie erreicht wurde und die Ergebnisse eine positive Nutzen-Risiko-Relation zeigte, was sowohl durch das DMC als auch durch den Sponsor unterstützt wurde, wurde den Patienten im Placeboarm, die noch keine Progression aufwiesen, angeboten, Regorafenib in einer offenen Behandlung zu erhalten. Dies wurde mittels eines Prüfplan-Amendments ermöglicht (41).

Da zwischen *data cut-off* (21. Juli 2011) und der Auswertung bzw. Kommunikation der Ergebnisse an die Prüfzentren mehrere Monate vergingen, wurde mit den Zulassungsbehörden vereinbart, eine weitere Analyse nur zum Gesamtüberleben zum Zeitpunkt unmittelbar vor Beginn des Cross-overs durchzuführen. Das Cut-off-Datum dafür war der 13. November 2011. Diese Analyse basierte auf 566 der ursprünglich 582 geplanten Todesfälle (97 %).

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 4-18: 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			
CORRECT 14387	Ja	Ja	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Randomisierung und Verdeckung der Gruppenzugehörigkeit

Am Ende der Screeningperiode wurden geeignete Patienten zufällig im Verhältnis 2:1 in einer verblindeten Art und Weise dem Arm Regorafenib + BSC oder dem Arm Placebo+ BSC zugeteilt. Um dies zu gewährleisten, wurde eine vom BAYER Randomisierungsmanager durch ein Computerprogramm erstellte Randomisierungsliste erzeugt, und zwar mit einem unizentrischen Randomisierungsschema. Die Randomisierungsnummer für jeden einzelnen Patienten erfuhr der Prüfarzt durch ein Interaktives Voice Response System (IVRS). Der Prüfarzt oder sein Beauftragter rief das IVRS an, nachdem die Erfüllung aller Ein- und

Ausschlusskriterien bestätigt wurde. Das IVRS vergab eine Randomisierungsnummer für den Patienten, die dazu benutzt wurde, den Patienten einem Behandlungsarm zuzuordnen. Patienten mussten innerhalb von sieben Tagen nach der Randomisierung mit der Studienbehandlung beginnen. Wenn dies nicht der Fall war, musste der Sponsor unter Angabe des Grundes, warum die Behandlung nicht begonnen wurde, benachrichtigt werden.

Die Randomisierung wurde stratifiziert nach:

- Vorangegangener Anti-VEGF-Therapie (ja/nein),
- Zeit von der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate)
- Geographische Region 1 (Nordamerika, Westeuropa, Israel, und Australien) vs. Region 2 (Asien) vs. Region 3 (Südamerika, Türkei und Osteuropa). Um eine ausgewogene Repräsentation der drei Regionen zu gewährleisten, sollten in der Region 2 (Asien) nicht mehr als 250 Patienten randomisiert werden.

Die Stratifizierung erfolgte ebenfalls über das IVRS und dem darin hinterlegten Randomisierungs- und Stratifizierungsschema.

Die Geheimhaltung der Gruppenzugehörigkeit war durch das zentralisierte Randomisierungsverfahren und durch die zentrale Vergabe der Patientennummern über Computer (IVRS) gewährleistet.

Somit ist sowohl die Generierung der Randomisierungssequenz als auch die Verdeckung der Zuteilung zu den Behandlungsarmen als angemessen zu bewerten.

Verblindung

Patienten wurden randomisiert, um Regorafenib oder ein identisch aussehendes Placebo in einer doppel-blinden Art und Weise zu erhalten, d. h. weder der Prüfarzt, Sponsor, noch der Patient wussten, welcher Behandlungsmethode der Patient zugeteilt wurde. Regorafenib und Placebo waren identisch im Aussehen, um die Verblindung zu gewährleisten. Patienten nahmen vier 40 mg-Tabletten einmal pro Tag oral ein. Um hier die Verblindung zu sichern, wurde die Studienmedikation (Regorafenib oder Placebo) mit einer nur einmal vorkommenden Packungsnummer versehen, die auf jede Flasche aufgedruckt wurde und durch das IVRS dem jeweiligen Patienten zugeordnet war. Es war beabsichtigt, dass alle Patienten, ihre Familien, die Studienteammitglieder (im Studienzentrum, in den teilnehmenden Klinischen Forschungszentren (CRO)) und die Mitglieder des Data Monitoring Committee (DMC) gegenüber der jeweiligen Gruppenzuteilung bis zum Studienende verblindet waren, mit folgenden Ausnahmen:

- der unabhängige Biostatistiker und der unabhängige Statistikprogrammierer, die für die vorläufigen Tabellen, Listen und Grafiken für den DMC-Review verantwortlich waren. Die Personenzahl mit diesen Aufgaben wurde so gering wie möglich gehalten.

Diesen Personen wurden keine anderen Aufgaben im Zusammenhang mit der Studie zugeteilt und sie wurden aus der Aufdeckung der Zuteilung zu den Behandlungsgruppen ausgeschlossen.

- DMC – im Falle, dass eine Entblindung zur umfassenden Evaluierung eines frühzeitigen Studienabbruchs bezüglich der Wirksamkeit/Unwirksamkeit oder Sicherheit nötig würde. Die Prozeduren zur Entblindung und die Kontrolle der unverblindeten Daten wurden in der DMC-Charta beschrieben.
- Laborpersonal, welches die Blut-Serumanalysen für die pharmakokinetischen Analysen und weitere Labortests durchführte
- unverblindete Sicherheitsbeauftragte und das Ethikkomitee, welche die Schwerwiegenden Unerwünschten Ereignisse berichteten, mit Ausnahme von Tod, da das Gesamtüberleben der primäre Endpunkt war

Die Übermittlung von Ergebnissen an die jeweiligen Studienzentren war nicht zugelassen. Der Datentransfer zu den klinischen Datenbanken erfolgte bis zum Datenbankschluss (Lock). Bei unbeabsichtigter Entblindung wurde die vorgesehene Behandlung weitergeführt und die unverblindete Information im CRF erfasst.

Vor Datenbankschluss und Entblindung der Studie wurde ein statistischer Analyseplan erstellt, auf Basis dessen die Analyse durchgeführt wurde. Somit ist eine ergebnisgesteuerte Analyse auszuschließen.

Es wurden keine weiteren das Verzerrungspotential beeinflussenden Punkte gefunden.

Insgesamt ist das Verzerrungspotential auf Studienebene als niedrig einzuschätzen

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 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	1	2	3a	3b	3c	4	5
CORRECT 14387	ja	ja	ja	ja	ja	ja	ja
1 – Mortalität: Gesamtüberleben (OS)							
2 – Morbidität/Mortalität: Progressionsfreies Überleben (PFS)							
3 – Morbidität							
3a Objektive Tumoransprechrage (ORR)							
3b Krankheitskontrollrate (DCR)							
3c Dauer der Tumorstabilisierung							
4 – Gesundheitsbezogene Lebensqualität anhand EORTC QLQ-C30							
5 – Unerwünschte Ereignisse/Verträglichkeit							

4.3.1.3.1 Endpunkt Mortalität: 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 4-20: Operationalisierung von Gesamtüberleben

Studie	Operationalisierung
CORRECT 14387	<p>Primärer Endpunkt war das Gesamtüberleben (OS) als Maß für die Mortalität gemessen vom Zeitpunkt der Randomisierung bis zum Tod (aufgrund jeglicher Ursache).</p> <p>Nach Beendigung der aktiven Studienbehandlung wurde das Gesamtüberleben in der Nachbeobachtungszeit monatlich per Telefoninterview erhoben.</p> <p>Die Überlebenszeit von Patienten, die zum Zeitpunkt der Analyse nicht verstorben waren, wurde zum letzten bekannten Lebensdatum oder Datum der verloren gegangenen Nachbeobachtung (Lost-to-Follow up) zensiert.</p> <p>Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren. Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden unabhängig von der wirklichen Behandlung.</p>

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

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

Tabelle 4-21: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
CORRECT 14387	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der primäre Endpunkt Mortalität (Gesamtüberleben (OS)) wurde verblindet erhoben. Die Berechnungen erfolgten nach dem ITT-Prinzip.

Auf der Grundlage des niedrigen Verzerrungspotenzials auf Studienebene ist auch das Verzerrungspotential des Endpunkts OS als niedrig einzustufen. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Fehlende Daten wurden zensiert. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Faktoren vor.

Das Verzerrungspotenzial für den Endpunkt Mortalität ist somit als niedrig einzustufen.

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

Tabelle 4-22: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel (21. Juli 2012)

Studie	Behandlungsarm	N	Gesamtüberleben (OS) in Tagen (Median 95 %-KI)	Mortalität (Todesfälle in der ITT-Population)	Reduziertes Risiko (%)	Hazard Ratio ^a (95 %-KI) p-Wert*
CORRECT 14387	Regorafenib + BSC	505	196 (178; 222)	275 (54,5 %)	22,6	0,774 (0,636,0,942) 0.0052
	Placebo + BSC	255	151 (134; 177)	157 (61,6 %)		

a – stratifiziert nach Vorbehandlung mit VEGF Medikamenten (ja/nein), Zeit seit Diagnose der metastasierten Erkrankung (≥ 18 Monate, < 18 Monate) und geographischer Region (1= Nordamerika, Westeuropa, Israel, Australien; 2 = Asien; 3 = Südamerika, Türkei, Osteuropa); (Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/1 und 14.2/2)

Für das Gesamtüberleben wurden die Anzahl an Ereignissen sowie die medianen Überlebenszeiten extrahiert. Für die Gruppenvergleiche (Cox-Proportional-Hazard-Model) wurden Hazard Ratios mit den zugehörigen 95 %-KI und p-Werten extrahiert. Dabei wurden sowohl unadjustierte als auch adjustierte Hazard Ratios berücksichtigt. Adjustierte Hazard Ratios berücksichtigten Faktoren, nach denen die Randomisierung stratifiziert war.

Der 21. Juli 2011 war das Cut-off-Datum für die zweite geplante Zwischenanalyse. Insgesamt verstarben bis zu diesem Zeitpunkt 432 Patienten der ITT-Population: 275 Patienten (54,5 %) in der Regorafenib + BSC-Gruppe und 157 Patienten (61,6 %) in der Placebo + BSC-Gruppe. Zu diesem Zeitpunkt resultierte die Behandlung mit Regorafenib bezüglich des Gesamtüberlebens in einer Risikominderung von 22,6 % im Vergleich zu Placebo. Die Hazard Ratio betrug 0,774, das 95 %-Konfidenzintervall lag zwischen 0,636 und 0,942 ($p=0,0052$). Das mediane Gesamtüberleben betrug 196 Tage in der Regorafenib-Gruppe und 151 Tage in der Placebo-Gruppe. Das mediane Gesamtüberleben verbesserte sich somit um 29,8 %. Die relative Verbesserung betrug in der 6-Monats-Überlebensrate 20,7 % (52,5 % in der Regorafenib + BSC-Gruppe und 43,5 % in der Placebo + BSC-Gruppe), in der 9-Monats-Überlebensrate 24 % (38,2 % in der Regorafenib + BSC-Gruppe und 30,8 % in der Placebo + BSC-Gruppe). Die Wirksamkeit von Regorafenib + BSC auf das Gesamtüberleben war weitgehend konsistent über die Subgruppen (siehe Kapitel 4.3.1.3.8). Die Signifikanz des Behandlungseffektes auf das Gesamtüberleben war nach der Adjustierung für die Stratifizierungsfaktoren in einer multivariaten Analyse robust: das adjustierte Hazard Ratio betrug 0,773 (95 %-KI 0,634-0,942; $p=0,0053$).

Die Kaplan-Meier-Kurven für beide Therapiearme verlaufen über den ersten Zyklus (28 Tage) zusammen, bevor sie sich deutlich separieren (vgl.).

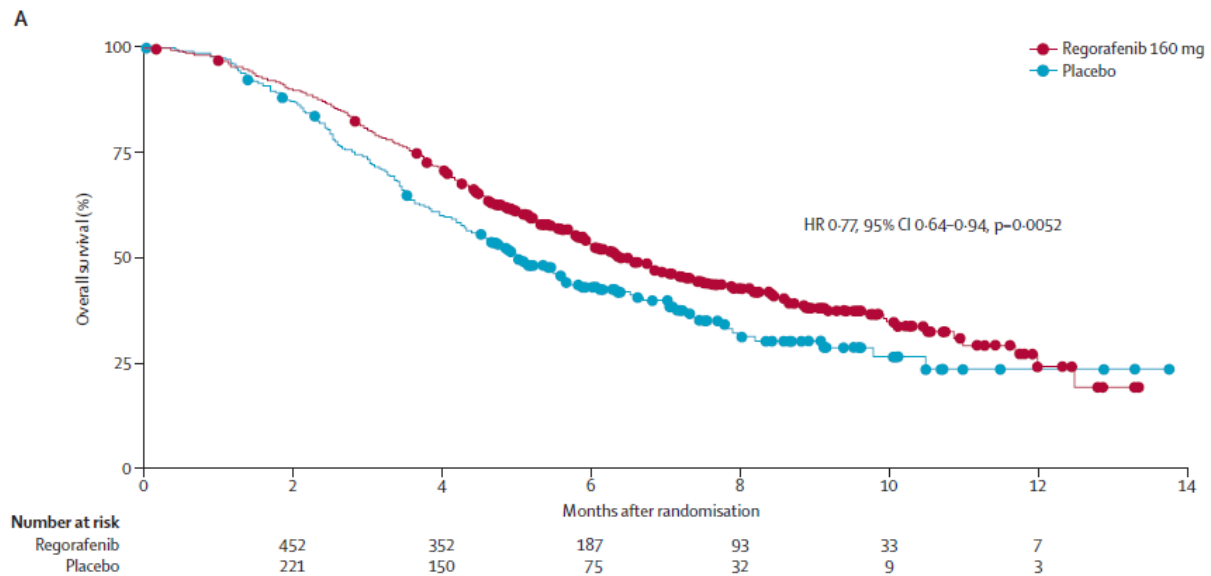


Abbildung 4-3: Kaplan-Meier-Kurve für das mediane Gesamtüberleben (CORRECT; ITT-Population)

Quelle: (16)

Sensitivitätsanalyse bezüglich des Endpunktes medianes Gesamtüberleben (OS)

Zwei Sensitivitätsanalysen des Gesamtüberlebens wurden durchgeführt: eine an unstratifizierten Daten und die andere anhand der Stratifizierungsinformationen aus dem IVRS. Dieser Ansatz wurde aufgrund der Misstratifikation einiger Patienten gewählt. Insgesamt wurden für 33 Patienten (15 in der Placebo + BSC-Gruppe und 18 in der Regorafenib + BSC-Gruppe) nicht zutreffende Stratifizierungsinformationen im IVRS hinterlegt. Für fünf Patienten (drei in der Placebo + BSC-Gruppe und zwei in der Regorafenib + BSC-Gruppe) war die Stratifizierungsvariable „geographische Region“ aufgrund einer Fehlprogrammierung im IVRS zu Studienbeginn betroffen. Nachdem dieser Fehler behoben wurde, kam es zu keinen weiteren Misstratifizierungen in Bezug auf diese Variable. Für die Stratifizierungsvariable „vorherige Anti-VEGF-Therapie“ kam es bei sieben Patienten (drei in der Placebo + BSC-Gruppe und vier in der Regorafenib + BSC-Gruppe) zu einem Mismatch zwischen IVRS und den Angaben im CRF. Für die Stratifizierungsvariable „Zeit seit der Diagnose des Stadiums IV“ kam es bei 23 Patienten (elf in der Placebo + BSC-Gruppe und zwölf in der Regorafenib + BSC-Gruppe) zu einem Mismatch zwischen IVRS und den Angaben im CRF. Zwei der insgesamt 33 Patienten wurden für mehr als einen Faktor misstratifiziert.

Die Ergebnisse der Sensitivitätsanalysen sind in Tabelle 4-23 dargestellt. Diese unterstützten die Ergebnisse der primären Analyse und waren mit ihnen konsistent. Daher sind die Ergebnisse der primären Analyse des Gesamtüberlebens schlüssig für einen definitiven Überlebensvorteil bei Patienten, die mit Regorafenib behandelt wurden.

Tabelle 4-23: Sensitivitätsanalyse des Gesamtüberlebens (ITT)

CORRECT 14387		Differenz der Überlebensraten
OS – stratifiziert ^a	Regorafenib - Placebo nach Monat 3 (95 %-KI)	7,6 % (1,1 %; 14,1 %)
	Regorafenib - Placebo nach Monat 6 (95 %-KI)	9,1 % (1,2 %; 16,9 %)
	Regorafenib - Placebo nach Monat 9 (95 %-KI)	7,4 % (-1,3 %; 16,2 %)
	Regorafenib - Placebo nach Monat 12 (95 %-KI)	0,2 % (-12,3 %; 12,8 %)
	Hazard Ratio (Regorafenib/Placebo)	0,773
	95 %-KI für Hazard Ratio	0,634; 0,942
	Einseitiger p-Wert aus Log rank Test (stratifiziert) ^b	0,0053
OS – unstratifiziert	Regorafenib - Placebo nach Monat 3 (95 %-KI)	7,6 % (1,1 %; 14,1 %)
	Regorafenib - Placebo nach Monat 6 (95 %-KI)	9,1 % (1,2 %; 16,9 %)
	Regorafenib - Placebo nach Monat 9 (95 %-KI)	7,4 % (-1,3 %; 16,2 %)
	Regorafenib - Placebo nach Monat 12 (95 %-KI)	0,2 % (-12,3 %; 12,8 %)
	Hazard Ratio (Regorafenib/Placebo)	0,767
	95 %-KI für Hazard Ratio	0,630; 0,933
	Einseitiger p-Wert aus Log rank-Test (unstratifiziert) ^c	0,0039

a – stratifiziert nach Randomisierung basierend auf Daten des IVRS

b – Eine Hazard Ratio <1 gibt eine Überlegenheit von Regorafenib 160 mg (experimenteller Arm) über Placebo (Kontrollarm) an. Die Hazard Ratio und ihr 95 %-KI basierte auf einem Cox-Regressionsmodell, stratifiziert nach vorheriger Therapie mit Anti-VEGF-Arzneimitteln (ja/nein), der Zeit seit der Diagnosestellung der metastasierten Erkrankung (≥18 Monate vs. <18 Monate) und der geographischen Region 1 (Nordamerika, Westeuropa, Israel, Australien) vs. Region 2 (Asien) vs. Region 3 (Südamerika, Türkei, Osteuropa).

c – Die Hazard Ratio und ihr 95 %-KI basierte auf einem Cox-Regressionsmodell, unstratifiziert.

(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/3 und 14.2/4)

Zusatzanalyse des Gesamtüberlebens zum Zeitpunkt des Cross-overs (*data cut-off* 13. November 2011)

Da zwischen *data cut-off* für die zweite geplante Zwischenanalyse (21. Juli 2011) und der Auswertung bzw. Kommunikation der Ergebnisse an die Prüfzentren mehrere Monate vergingen, wurde mit den Zulassungsbehörden vereinbart, eine weitere deskriptive Analyse

nur zum Gesamtüberleben zum Zeitpunkt unmittelbar vor Beginn des Cross-overs durchzuführen. Das Cut-off-Datum dafür war der 13. November 2011. Diese Analyse basierte auf 566 der ursprünglich 582 geplanten Todesfälle (97 %). Die Ergebnisse sind in Tabelle 4-24 dargestellt.

Tabelle 4-24: Ergebnisse für Gesamtüberleben aus RCT mit dem zu bewertenden Arzneimittel (deskriptive Zusatzanalyse)

Studie	Behandlungsarm	N	Gesamtüberleben (OS) in Tagen (Median 95 %-KI)	Mortalität (Todesfälle in der ITT-Population)	Reduziertes Risiko (%)	Hazard Ratio ^a (95 %-KI) p-Wert
CORRECT 14387	Regorafenib + BSC	505	194 (177; 214)	369 (73,1 %)	21,0	0,790 (0,664; 0,939) 0,0038
	Placebo + BSC	255	152 (134; 178)	197 (77,3 %)		

a – stratifiziert nach Vorbehandlung mit VEGF Medikamenten (ja/nein), Zeit seit Diagnose der metastasierten Erkrankung (≥ 18 Monate, < 18 Monate) und geographischer Region (1= Nordamerika, Westeuropa, Israel, Australien; 2 = Asien; 3 = Südamerika, Türkei, Osteuropa); (Quelle: CORRECT 14387 ZA6: Tabelle 14.2/1 und Tabelle 14.2/2 und (42))

Die ermittelte HR dieser zusätzlichen deskriptiven Analyse des Gesamtüberlebens weicht nur geringfügig von der HR der primären Analyse ab (0,790 vs. 0,774). Die Konfidenzintervalle sind ebenfalls sehr ähnlich (0,664 bis 0,939 vs. 0,636 bis 0,942). Damit bestätigt diese zusätzliche deskriptive Analyse anhand von 97 % der ursprünglich geplanten Todesfälle den Überlebensvorteil von Regorafenib + BSC gegenüber Placebo + BSC.

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.2 Endpunkt Mortalität/Morbidität: 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.

Der Parameter progressionsfreies Überleben ist ein direkter Parameter der Krankheitslast und ein Endpunkt, für den eine Korrelation mit dem Gesamtüberleben bei bestimmten Tumorerkrankungen validiert wurde. So wurde beim fortgeschrittenen kolorektalen Karzinom die Validität von PFS als Surrogatendpunkt für das Gesamtüberleben nachgewiesen (18, 37) (siehe auch Kapitel 4.2.5.2). Er misst direkt die Effekte der jeweiligen Behandlung auf das Tumorstadium unter Benutzung einer objektiven und quantitativen Methodik. Dabei werden sowohl zytostatische (Tumorstabilisierung) als auch zytotoxische (Tumorrückgang) Effekte gemessen. Daher sollte im Kontext einer positiven Risiko-Nutzen-Abschätzung eine klinisch relevante Verbesserung des PFS einen patientenrelevanten Endpunkt aus eigenem Recht darstellen.

Studien belegen, dass der Status der Progressionsfreiheit tatsächlich ein Wert für die Patienten an sich darstellt, der mit einer höheren Lebensqualität und einem besseren Gesundheitszustand einhergeht (6-8). In einer Studie von Siena et al. (7) wurde die Assoziation von PFS mit Symptomen, die durch Darmkrebs bedingt sind, mit der Lebensqualität und mit dem Gesamtüberleben untersucht. Das Fehlen einer Krankheitsprogression war mit einer signifikant und klinisch relevant niedrigeren Symptomschwere und einer höheren gesundheitsbezogenen Lebensqualität aufgrund der Krebserkrankung assoziiert als bei Patienten mit einer Progression der Erkrankung. Thong et al. (8) verglichen die Lebensqualität von Patienten mit und ohne Krankheitsprogression anhand der validierten Fragebogen SF-36 und QOL-CS. Im Ergebnis wurde festgestellt, dass Patienten mit einer Progression sowohl einen schlechteren Gesundheitsstatus als auch eine schlechtere Lebensqualität aufweisen. Insbesondere die psychische Befindlichkeit wurde in einer Studie von Hurvitz et al. (6) untersucht. Frauen mit Brustkrebs wurden zu zwei hypothetischen medikamentösen Behandlungen befragt; die eine sei assoziiert mit einem

durchschnittlichen PFS von 12 Monaten, die andere mit einem von 16 Monaten. Das Nebenwirkungsprofil und das Gesamtüberleben wären bei beiden Therapien gleich. 88 % der befragten Patientinnen bevorzugten die Therapie mit dem längeren PFS. Eine weitere Erhebung des psychischen Befindens in dieser Studie bei Patientinnen, deren Erkrankung progredient war versus denen ohne Progression ergab, dass erstere mehr in Apathie und Interessenlosigkeit fallen, fürchten, dass die Therapie nicht wirkt, und sich sorgen, dass sie ihren alltäglichen Aufgaben nicht nachkommen können.

Tabelle 4-25: Operationalisierung von Progressionsfreiem Überleben

Studie	Operationalisierung
CORRECT 14387	<p>Sekundärer Endpunkt war das progressionsfreie Überleben (PFS) als Maß für die Zeit ohne Krankheitsprogression oder Tod gemessen vom Zeitpunkt der Randomisierung bis zur ersten beobachteten Krankheitsprogression (radiologisch oder klinisch) oder dem Tod.</p> <p>Die Analyse des PFS basierte auf der Beurteilung durch den Prüfarzt anhand der RECIST-Kriterien, Version 1.1. Die erste Evaluierung des PFS nach der Randomisierung fand nach ca. 56 Tagen statt. Tumorevaluierungen alle acht Wochen sind ein von den Zulassungsbehörden anerkanntes Zeitintervall.</p> <p>Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren. Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden, unabhängig von der wirklichen Behandlung.</p>

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

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

Tabelle 4-26: Bewertung des Verzerrungspotenzials für progressionsfreies Überleben 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
CORRECT 14387	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der sekundäre Endpunkt PFS wurde verblindet erhoben. Die Berechnungen erfolgten nach dem ITT-Prinzip.

Auf der Grundlage des niedrigen Verzerrungspotenzials auf Studienebene ist auch das Verzerrungspotential des Endpunkts Progressionsfreies Überleben als niedrig einzustufen. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Fehlende Daten wurden zensiert. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Faktoren vor.

Das Verzerrungspotenzial für den Endpunkt PFS ist somit als niedrig einzustufen.

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

Tabelle 4-27: Ergebnisse für progressionsfreies Überleben aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlungssarm	N	PFS in Tagen (Median, 95 %-KI)	PFS-Ereignis (% der ITT-Population)	Reduziertes Risiko (%)	Hazard Ratio (95 %-KI) p-Wert
CORRECT 14387	Regorafenib + BSC	505	59 (57, 65)	85,1	50,6	0,494 (0,419, 0,582) <0,0001
	Placebo + BSC	255	52 (51, 53)	94,5		

(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/7 und 14.2/8)

Für das PFS wurden die Anzahl an Ereignissen sowie die medianen progressionsfreien Überlebenszeiten extrahiert. Für die Gruppenvergleiche (Cox-Proportional-Hazard-Model) wurden Hazard Ratios mit den zugehörigen 95 %-Konfidenzintervallen und p-Werten extrahiert. Dabei wurden sowohl unadjustierte als auch adjustierte Hazard Ratios berücksichtigt. Adjustierte Hazard Ratios berücksichtigten Faktoren, nach denen die Randomisierung stratifiziert war.

Der 21. Juli 2011 war Cut-off-Datum für die zweite geplante Zwischenanalyse. Insgesamt hatten bis zu diesem Zeitpunkt 671 Patienten der ITT-Population eine Krankheitsprogression oder verstarben (vor einer festgestellten Progression): 430 Patienten (85,1 %) in der Regorafenib + BSC-Gruppe und 241 Patienten (94,5 %) in der Placebo + BSC-Gruppe. Zu diesem Zeitpunkt resultierte die Behandlung mit Regorafenib bezüglich des progressionsfreien Überlebens in einer Risikominderung von 50,6 % im Vergleich zu Placebo. Die Hazard Ratio betrug 0,494, das 95 %-Konfidenzintervall lag zwischen 0,419 und 0,582 ($p < 0,0001$). Das mediane progressionsfreie Überleben betrug 59 Tage in der Regorafenib + BSC-Gruppe und 52 Tage in der Placebo + BSC-Gruppe. Das mediane progressionsfreie Überleben verbesserte sich um 13,5 %. Die relative Verbesserung betrug 211 % in der progressionsfreien Überlebensrate nach drei Monaten (42 % in der Regorafenib + BSC-Gruppe und 13,5 % in der Placebo + BSC-Gruppe). Die relative Verbesserung betrug 519 % in der progressionsfreien Überlebensrate nach sechs Monaten (13 % in der Regorafenib-Gruppe und 2,1 % in der Placebo-Gruppe). Die Wirksamkeit von

Regorafenib + BSC auf das progressionsfreie Überleben war konsistent über die Subgruppen. Die Signifikanz des Behandlungseffektes auf das progressionsfreie Überleben war nach der Adjustierung für die Stratifizierungsfaktoren in einer multivariaten Analyse robust: das adjustierte Hazard Ratio betrug 0,479 (95 %-KI 0,405 bis 0,565; p-Wert <0,0001).

Die Kaplan-Meier-Kurven für das PFS für beide Therapiearme verlaufen bis zur ersten Tumorevaluierung zusammen, bevor sie sich deutlich separieren (vgl. Abbildung 4-4). Zu diesem Zeitpunkt wurde der Median bereits erreicht. Die punktuelle Betrachtung der PFS-Differenz an den Medianen ist in diesem Fall für den erheblichen Zusatznutzen nicht repräsentativ, der sich aus der Hazard Ratio für PFS ergibt, die den gesamten Verlauf der Kaplan-Meier-Kurven reflektiert. Die protokollgemäße erste radiologische Tumorauswertung erst nach 8 Wochen Therapie bedingt eine geringe Diskriminierbarkeit der beiden Studiengruppen vor diesem Zeitpunkt mittels des Kriteriums Progression nach RECIST. Insofern ist die Hazard Ratio für PFS zur Bestimmung des Zusatznutzens maßgeblich. Die PFS-Raten waren nach drei, sechs und neun Monaten nach der Randomisierung konsistent höher in der Regorafenib + BSC-Gruppe im Vergleich zur Placebo + BSC-Gruppe.

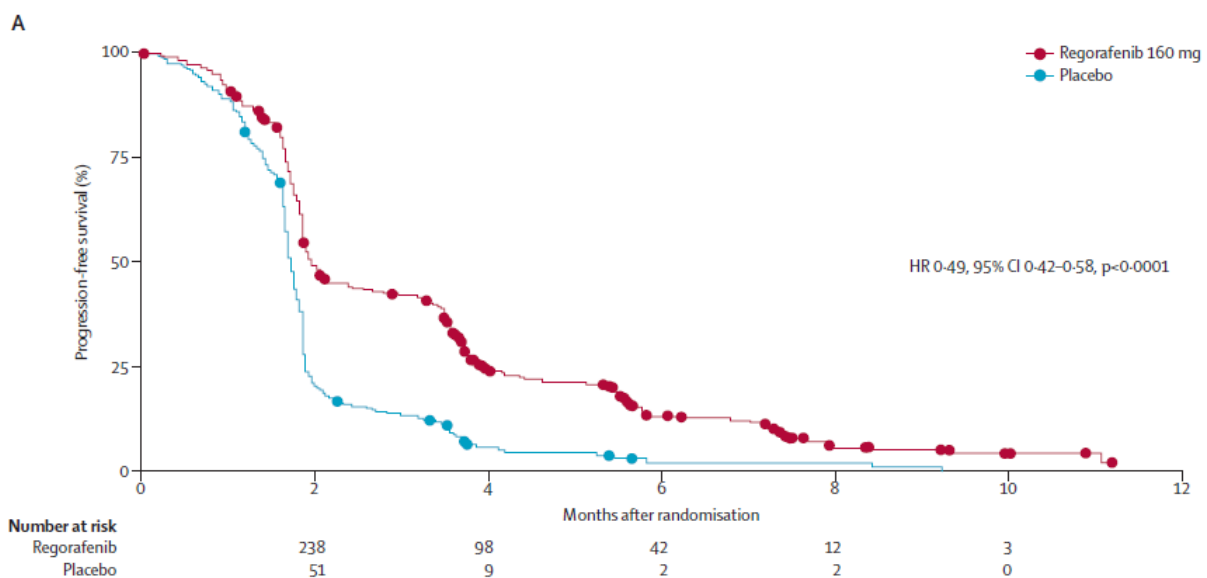


Abbildung 4-4: Kaplan-Meier-Kurve für das mediane progressionsfreie Überleben (CORRECT; ITT-Population)

Quelle: (16)

Sensitivitätsanalyse bezüglich des Endpunktes progressionsfreies Überleben

Aufgrund der Misstratifikation einiger Patienten wurden zwei Sensitivitätsanalysen des Progressionsfreien Überlebens durchgeführt: eine an unstratifizierten Daten und die andere anhand der Stratifizierungsinformationen aus dem IVRS. Die Ergebnisse der Sensitivitätsanalysen sind in Tabelle 4-28 dargestellt. Diese unterstützten die Ergebnisse der primären Analyse des progressionsfreien Überlebens und sind mit ihnen konsistent, indem sie eine statistisch signifikante Verbesserung des progressionsfreien Überlebens im Regorafenib + BSC-Arm gegenüber dem Placebo + BSC-Arm nachweisen.

Tabelle 4-28: Sensitivitätsanalyse des progressionsfreien Überlebens (ITT) aus RCT mit dem zu bewertenden Arzneimittel

	PFS – stratifiziert nach Randomisierung basierend auf Daten des IVRS	Differenz der progressionsfreien Überlebensraten
PFS – stratifiziert^a	Regorafenib - Placebo nach Monat 3 (95 %-KI)	28,5 % (22,4 %; 34,6 %)
	Regorafenib - Placebo nach Monat 6 (95 %-KI)	11,0 % (6,9 %; 15,0 %)
	Regorafenib - Placebo nach Monat 9 (95 %-KI)	4,0 % (0,9 %; 7,2 %)
	Hazard Ratio (Regorafenib/Placebo)	0,479
	95 %-KI für Hazard Ratio	(0,405; 0,565)
	Einseitiger p-Wert aus Log rank Test (stratifiziert) ^b	<0,0001
PFS – unstratifiziert	Regorafenib - Placebo nach Monat 3 (95 %-KI)	28,5 % (22,4 %; 34,6 %)
	Regorafenib - Placebo nach Monat 6 (95 %-KI)	11,0 % (6,9 %; 15,0 %)
	Regorafenib - Placebo nach Monat 9 (95 %-KI)	4,0 % (0,9 %; 7,2 %)
	Hazard Ratio (Regorafenib/Placebo)	0,501
	95 %-KI für Hazard Ratio	(0,425; 0,590)
	Einseitiger p-Wert aus Log rank Test (unstratifiziert) ^c	<0,0001
<p>a – stratifiziert nach Randomisierung basierend auf Daten des IVRS</p> <p>b – Eine Hazard Ratio < 1 gibt eine Überlegenheit von Regorafenib 160 mg (experimenteller Arm) über Placebo (Kontrollarm) an. Die Hazard Ratio und ihr 95 %-KI basierte auf einem Cox-Regressionsmodell, stratifiziert nach vorheriger Therapie mit Anti-VEGF-Arzneimitteln (ja/nein), der Zeit seit der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate) und der geographischen Region 1 (Nordamerika, Westeuropa, Israel, Australien) vs. Region 2 (Asien) vs. Region 3 (Südamerika, Türkei, Osteuropa).</p> <p>c – Die Hazard Ratio und ihr 95 %-KI basierte auf einem Cox-Regressionsmodell, unstratifiziert.</p> <p>(Quelle: CORRECT 14387 Studienbericht: Tabelle 9-5, Tabelle 14.2/9 und 14.2/10)</p>		

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.3 Endpunkt Morbidität: Objektive Tumoransprechrates (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.

Für die Tumoransprechrates wurde eine Korrelation mit dem Gesamtüberleben beim kolorektalen Karzinom nachgewiesen. (17, 18) Die von Johnson et al. (18) gefundene Korrelation von Tumoransprechen und Gesamtüberleben (siehe Kapitel 4.2.5.2) bei einer Stichprobengröße von n=750 von 20 % wird in der CORRECT-Studie (n=760) nicht erfüllt (Unterschied 0,6 %). Tumoransprechen nach RECIST- oder WHO-Kriterien steht für viele zielgerichtete Arzneimittel im Allgemeinen und für Multikinaseinhibitoren im Besonderen im Gegensatz zu den klassischen zytotoxischen Arzneimitteln nicht im Vordergrund ihrer biologischen Wirkung. Vielmehr geht es um eine Tumorstabilisierung und Veränderungen im

Tumor, in der Tumorumgebung bzw. in den einen Tumor versorgenden Blutgefäßen, die hemmende Wirkungen auf die Tumorprogression auslösen. Daher werden für diese Arzneimittel andere Methoden des Wirksamkeitsnachweises diskutiert und angewandt. (19-22)

Tabelle 4-29: Operationalisierung von Objektive Tumoransprechrates

Studie	Operationalisierung
CORRECT 14387	<p>Sekundärer Endpunkt war die Objektive Tumoransprechrates (ORR) als Maß für die direkte Wirkung der Therapie auf das Tumorstadium. Sie wird errechnet aus der Summe der Patienten mit einer kompletten (CR) und der Patientent mit einer partiellen Remission (PR) – ORR = CR+PR.</p> <p>Die Analyse der ORR basierte auf der Beurteilung durch den Prüfarzt anhand der RECIST-Kriterien, Version 1.1. Die erste Evaluierung der ORR nach der Randomisierung fand nach ca. 56 Tagen statt. Die Einstufung in CR und PR basierte auf der radiologischen Bewertung des Prüfarztes.</p> <p>Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren, die eine PR oder CR als bestes Ansprechen aufwiesen. Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden, unabhängig von der wirklichen Behandlung.</p>

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

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

Tabelle 4-30: Bewertung des Verzerrungspotenzials für Objektive Tumoransprechrates 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
CORRECT 14387	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der sekundäre Endpunkt ORR wurde verblindet erhoben. Die Berechnungen erfolgten nach dem ITT-Prinzip.

Auf der Grundlage des niedrigen Verzerrungspotenzials auf Studienebene ist auch das Verzerrungspotential des Endpunkts ORR als niedrig einzustufen. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Es wurden nur Patienten in die Analyse einbezogen, in denen eine Tumorevaluierung stattfand und bei denen eine komplette Remission (CR) oder eine partielle Remission (PR) als bestes Ansprechen dokumentiert wurde. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte vor.

Das Verzerrungspotenzial für den Endpunkt ORR ist somit als niedrig einzustufen.

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

Tabelle 4-31: Ergebnisse für Objektive Tumoransprechrates aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlungssarm	N	CR	PR	Differenz Placebo + BSC – Regorafenib + BSC (%)
			n (%) 95 %-KI	n (%) 95 %-KI	
CORRECT 14387	Regorafenib + BSC	505	0 (0,0) (0,0; 0,6)	5 (1,0) (0,3; 2,3)	-0,60 (-1,74; 0,53)
	Placebo + BSC	255	0 (0,0) (0,0; 1,2)	1 (0,4) (0,0; 2,2)	

(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/22 und 14.2/23)

Die Objektive Tumoransprechrates (CR + PR) in der ITT-Population betrug 1,0 % (95 %-KI: 0,3 bis 2,3 %) in der Regorafenib + BSC-Gruppe und 0,4 % (95 %-KI: 0,0 bis 2,2 %) in der Placebo + BSC-Gruppe. Der Unterschied von -0,60 % (95 %-KI: -1,74 bis 0,53 %) zwischen den beiden Gruppen war statistisch nicht signifikant (einseitiges p=0,1884; präspezifiziertes alpha 0,025).

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.4 Endpunkt Morbidität: Krankheitskontrollrate (DCR) – 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.

Eng zusammen mit dem progressionsfreien Überlebensstatus steht das Tumoransprechen und die Tumorstabilisierung bzw. das Nicht-Fortschreiten der Erkrankung, hier als Krankheitskontrollrate operationalisiert.

Es ist mehrfach publiziert worden, dass Patienten, die schlechte Prognosefaktoren für ihre Tumorerkrankung aufweisen, eine schnellere Progression erleiden als Patienten mit guten Prognosefaktoren. Gemeinhin bestimmt die Geschwindigkeit der Tumorprogression, in welchem Ausmaß der Patient unter Krebs-bedingten Symptomen leidet und wie lange er noch zu leben hat. Das heißt, dass i. d. R. die Verkürzung der Lebenserwartung bei Patienten mit Krebs auf die Geschwindigkeit der Progression des Tumors zurückgeführt werden kann (34-36).

Viele Patienten und ihre Behandler beurteilen es als Vorteil, wenn eine schnelle und effektive Tumorkontrolle erfolgt und eine Progression des Tumors verzögert werden kann, und zwar unabhängig von einer etwaigen nachgewiesenen Verlängerung der Gesamtüberlebenszeit (4). Dies trifft insbesondere auf schnell fortschreitende Tumoren oder solche zu, die ein lebenswichtiges Organ bedrohen. In einer Meta-Analyse von 21 Studien in verschiedenen Tumorentitäten wurde die Korrelation von radiologischem Tumoransprechen und der

gesundheitsbezogenen Lebensqualität untersucht (14). Ein objektives Tumoransprechen ist demnach signifikant mit einem Symptomansprechen assoziiert. Auch Patienten, die kein objektives Tumoransprechen im Sinne einer kompletten oder partiellen Remission, sondern eine Stabilisierung der Erkrankung erfuhren, hatten ein Symptomansprechen. Die medikamentöse Tumorthherapie hatte somit den größten Effekt für die Steigerung der gesundheitsbezogenen Lebensqualität bei Patienten, die eine komplette oder partielle Remission erfuhren, gefolgt von denen, die eine Stabilisierung ihrer Erkrankung als bestes Ansprechen aufwiesen. Der zusammenfassende Parameter ist die Krankheitskontrollrate.

Tabelle 4-32: Operationalisierung von Krankheitskontrollrate

Studie	Operationalisierung
CORRECT 14387	<p>Sekundärer Endpunkt war die Krankheitskontrollrate (Disease Control Rate - DCR) als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum, inklusive Tumorstabilisierungen. Die DCR setzt sich zusammen aus der Summe der Patienten mit CR, PR und SD – $DCR = CR + PR + SD$.</p> <p>Die Analyse der DCR basierte auf der Beurteilung durch den Prüfarzt anhand der RECIST-Kriterien, Version 1.1. Die erste Evaluierung der DCR nach der Randomisierung fand nach ca. 56 Tagen statt. Eine Stabilisierung der Erkrankung (SD) wurde in die Berechnung der DCR nur dann einbezogen, wenn diese mindestens sechs Wochen nach der Randomisierung gemessen wurde.</p> <p>Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren, und die eine CR, PR oder SD als bestes Ansprechen aufwiesen. Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden, unabhängig von der wirklichen Behandlung.</p>

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

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

Tabelle 4-33: Bewertung des Verzerrungspotenzials für die Krankheitskontrollrate 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
CORRECT 14387	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der sekundäre Endpunkt Krankheitskontrollrate (DCR) wurde verblindet erhoben. Die Berechnungen erfolgten nach dem ITT-Prinzip.

Auf der Grundlage des niedrigen Verzerrungspotenzials auf Studienebene ist auch das Verzerrungspotential des Endpunkts Krankheitskontrollrate als niedrig einzustufen. Verblindung und ITT-Prinzip wurden adäquat umgesetzt. Es wurden nur Patienten in die Analyse einbezogen, bei denen eine Tumorevaluierung stattfand und bei denen eine CR, PR oder eine Stabilisierung der Erkrankung (SD) als bestes Ansprechen dokumentiert wurde. SD wurde in die Berechnung der DCR nur dann einbezogen, wenn diese mindestens sechs Wochen nach der Randomisierung gemessen wurde. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte vor.

Das Verzerrungspotenzial für den Endpunkt Krankheitskontrollrate ist somit als niedrig einzustufen.

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

Tabelle 4-34: Ergebnisse für Krankheitskontrollrate aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlungsarm	N	CR	PR	SD	DCR ^a	Differenz DCR Placebo + BSC – Regorafenib + BSC 95 %-KI p-Wert
			n (%) 95 %-KI	n (%) 95 %-KI	n (%) 95 %-KI	n (%) 95 %-KI	
CORRECT 14387	Regorafenib + BSC	505	0 (0,0) 0,0; 0,6	5 (1,0) 0,3; 2,3	216 (42,8) 38,4; 47,2	207 (41,0) 36,7; 45,4	-25,94 -32,06; -19,82 <0,0001
	Placebo + BSC		255	0 (0,0) 0,0; 1,2	1 (0,4) 0,0; 2,2	37 (14,5) 10,4; 19,4	38 (14,9) 10,8; 19,9
<p>a – Prozentzahl an Patienten mit CR, PR oder SD entsprechend RECIST-Kriterien, Version 1.1. Patienten mit einer SD, deren Evaluation vor der ersten, im Prüfplan vorgesehenen Tumorevaluation (sechs Wochen nach Randomisierung) durchgeführt wurde, sind in die Berechnung der DCR nicht mit einbezogen worden. Non-CR/non-PD wurden in die DCR-Kalkulation mit einbezogen und folgten denselben Kriterien wie für SD. Der Denominator für die Raten (%) und die 95 %-KI basieren auf der ITT-Population (alle randomisierten Patienten).</p> <p>(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/22 und 14.2/23 und ZA5: Tabelle 16.1.9.2/1)</p>							

Die Krankheitskontrollrate (CR + PR + SD >6 Wochen) in der ITT-Population war höher in der Regorafenib + BSC-Gruppe als in der Placebo + BSC-Gruppe. Die DCR betrug 41,0 % (95 %-KI: 36,7 bis 45,4 %) in der Regorafenib + BSC-Gruppe und 14,9 % (95 %-KI: 10,8 bis 19,9 %) in der Placebo + BSC-Gruppe. Der Unterschied von -25,94 % (95 %-KI: -32,06

bis -19,82 %) zwischen den beiden Gruppen war statistisch signifikant (einseitiges $p < 0,0001$; präspezifiziertes alpha 0,025). Der Unterschied war hauptsächlich durch die Rate an Krankheitsstabilisierungen (SD) bedingt. Die Unterschiede in den Ansprechraten und in der Krankheitskontrollrate sowie den zugehörigen 95 %-Konfidenzintervallen basieren auf dem Cochran-Mantel-Haenszel-Test adjustiert auf dieselben Stratifizierungsfaktoren wie für den primären Endpunkt.

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.5 Endpunkt Morbidität: Dauer der Tumorstabilisierung – 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.

In einer Metaanalyse von 21 Studien in verschiedenen Tumorentitäten wurde die Korrelation von radiologischem Tumoransprechen und der gesundheitsbezogenen Lebensqualität untersucht. (14) Ein objektives Tumoransprechen ist demnach signifikant mit einem Symptomansprechen assoziiert. Auch Patienten, die kein objektives Tumoransprechen im Sinne einer kompletten oder partiellen Remission, sondern eine Stabilisierung der Erkrankung erfuhren, hatten ein Symptomansprechen. Die Dauer dieser Stabilisierung der Erkrankung mit den dadurch bedingten positiven Effekten auf die psychische und physische Befindlichkeit des einzelnen Patienten ist daher patientenrelevant (siehe auch Abschnitt 4.2.5.2).

Tabelle 4-35: Operationalisierung von Dauer der Tumorstabilisierung

Studie	Operationalisierung
CORRECT 14387	<p>Tertiärer Endpunkt war die die Dauer der Tumorstabilisierung als Maß für die direkte Wirkung der Therapie auf das Tumorwachstum über die Zeit.</p> <p>Die Dauer der Tumorstabilisierung war definiert als Zeit (in Tagen) von der Randomisierung bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression auftrat). Einbezogen wurden alle Patienten mit einer SD von mindestens 6 Wochen, non-CR/non-PR und ein Patient, der als „nicht zutreffend“ im Besten Ansprechen bewertet wurde.</p> <p>Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren. Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden, unabhängig von der wirklichen Behandlung.</p>

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

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

Tabelle 4-36: Bewertung des Verzerrungspotenzials für Dauer der Tumorstabilisierung 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
CORRECT 14387	Niedrig	Ja	Nein	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der tertiäre Endpunkt Dauer der Tumorstabilisierung wurde auf der Basis der verblindeten Erhebung des Tumoransprechens erhoben. Die Berechnungen erfolgten nach dem ITT-Prinzip.

Auf der Grundlage des niedrigen Verzerrungspotenzials auf Studienebene ist auch das Verzerrungspotential des Endpunkts Dauer der Tumorstabilisierung als niedrig einzustufen. Die Verblindung wurde adäquat umgesetzt. Es wurden nur Patienten in die Analyse einbezogen, bei denen eine Tumorevaluierung stattfand und bei denen eine CR, PR oder SD als bestes Ansprechen dokumentiert wurde. Da nur für 712 von 760 (93,7 %) der Patienten der ITT-Population Daten zu diesem Endpunkt vorlagen, liegt eine Abweichung um mehr als fünf Prozent von der ITT-Population vor. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte vor.

Das Verzerrungspotenzial für den Endpunkt Dauer der Tumorstabilisierung ist somit als niedrig einzustufen.

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

Tabelle 4-37: Ergebnisse für Dauer der Tumorstabilisierung aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlungsarm	Auswertbar für Endpunkt n (% der Gesamt-population)	Zeit der Tumorstabilisierung in Tagen (Median) 95 %-KI	Gesamtzahl der Patienten mit Ereignis n (%)	Gesamtzahl zensierte Patienten n (%)	Hazard Ratio (95 %-KI)
CORRECT 14387	Regorafenib + BSC	470 (93,1)	60 (57; 70)	409 (87,0)	61 (13,0)	0,471 (0,398; 0,558)
	Placebo + BSC	242 (94,9)	52 (51; 53)	230 (95,0)	12 (5,0)	

(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/20 und 14.2/21; ZA5: Tabelle 16.1.9.2/1)

Die Hazard Ratio für die **Dauer der Tumorstabilisierung** betrug 0,471 (95% KI 0,398, 0,558) zugunsten der Regorafenib + BSC-Gruppe. Die mediane Dauer der Tumorstabilisierung betrug 60 Tage (95 %-KI: 57, 70) in der Regorafenib + BSC-Gruppe und 52 Tage (95 %-KI: 51 bis 53) in der Placebo + BSC-Gruppe. Die mediane Dauer der Tumorstabilisierung verbesserte sich somit um 15,3 %.

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.6 Endpunkt Gesundheitsbezogene Lebensqualität anhand EORTC QLQ C30 – 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 4-38: Operationalisierung von Gesundheitsbezogener Lebensqualität

Studie	Operationalisierung
CORRECT 14387	Der Fragebogen QLQ C30 der <i>European Organisation for Research and Treatment of Cancer</i> (EORTC) wurde zur Messung der gesundheitsbezogenen Lebensqualität benutzt. Insgesamt handelt es sich bei dem QLQ-C30 um einen Fragebogen für Tumorpatienten mit 30 Fragen (die aktuelle Standard-Version ist die Version 3.0). Insgesamt können 100 Punkte erreicht werden; das stünde dann für perfekte Gesundheit und Wohlbefinden. Die

Ergebnisse werden auf fünf Funktionsskalen (körperliche, kognitive, emotionale, soziale und Rollenfunktionalität), einer Skala zur Bewertung der allgemeinen Lebensqualität bzw. des Gesundheitsstatus sowie einer Skala hinsichtlich Symptomen dargestellt (43).

Der Fragebogen wurde an die Patienten bei Beginn der Visite ausgegeben, bevor sie den Prüfartz traf oder andere studienspezifische Maßnahmen begonnen wurden. Der Fragebogen sollte am Tag 1 des Zyklus 1, am Tag 1 der Zyklen 2 bis 4, und danach am Tag 1 jedes zweiten Zyklus (Zyklus 6, 8 usw.), sowie bei der Visite zum Therapieende ausgefüllt werden.

Mit dem Fragebogen werden die wichtigsten Dimensionen der gesundheitsbezogenen Lebensqualität (allgemeiner Gesundheitszustand/Lebensqualität, physische Funktionalität, Rollen-Funktionalität, kognitive Funktionalität, emotionale Funktionalität, soziale Funktionalität), die häufigsten Symptome der Tumorpatienten (Erschöpfung, Übelkeit und Erbrechen, Schmerzen, Dyspnoe/Atembeschwerden, Schlaf-Störungen, Appetit-Mangel, Obstipation, Diarrhö) und Effekte auf die finanzielle Situation erfasst.

Für die Mehrheit der EORTC QLQ-C30 Items wird eine 4-Punkt Likert-Skala verwendet. Nur für den globalen Gesundheitszustand wird eine 7-Punkt Likert-Skala verwendet. Zur einfacheren Verwendung wurden die Werte linear auf eine Skala von 0 bis 100 konvertiert. Für die funktionalen und globalen Skalen entspricht ein hoher Wert einer besseren Funktionalität. Bei den Symptomskalen entspricht ein hoher Wert einem hohen Grad an Symptomatik.

Für einzelne Bereiche wurden folgende Berechnungen durchgeführt:

- Der Wert für den allgemeinen Gesundheitszustand/Lebensqualität ist eine Transformation aus den durchschnittlichen Rohdaten der Fragen 29 und 30 mit 0 als schlechtestem Wert und 100 als bestem Wert.
- Der Wert für die funktionalen Skalen ist eine Transformation aus den durchschnittlichen Rohdaten der Fragen mit 0 als schlechtestem Wert und 100 als bestem Wert: Frage 1-5: physische Funktionalität, Frage 6 und 7: Rollen-Funktionalität, Frage 21-24: Emotionale Funktionalität, Frage 20 und 25: Kognitive Funktionalität, Frage 26 und 27: Soziale Funktionalität.
- Der Wert für die Symptomskala ist eine Transformation aus den durchschnittlichen Rohdaten der Fragen mit 0 als keine Symptome und 100 als schlechtestem Wert. Frage 10, 12 und 18: Erschöpfung, Frage 14 und 15: Übelkeit und Erbrechen, Frage 9 und 19: Schmerzen, Frage 8: Dyspnoe/Atembeschwerden, Frage 11: Schlaf-Störungen, Frage 13: Appetit-Mangel, Frage 16: Obstipation, Frage 17: Diarrhö, Frage 28: Finanzielle Schwierigkeiten.

Änderungen der Lebensqualität werden, da dies eine akzeptierte Abstufung ist, in die Kategorien „wenig“ (Scorewertunterschied >5 bis 10), „moderat“ (Scorewertunterschied >10 bis 20), und „viel“ (Scorewertunterschied >20) eingeteilt (44). Ein Scorewertunterschied bis 5 wird als klinisch bedeutungslos eingestuft.

Andere Publikationen gehen davon aus, dass erst ein Unterschied von mindestens 10 Punkten klinisch bedeutsam ist (44-46). In einem deutschen Beitrag aus dem Jahr 2005 (47) heißt es aber: „Zugleich bleibt bislang eine eindeutige Antwort auf die Frage, was eine bedeutende Änderung der Lebensqualität ist und wie diese quantifiziert werden kann, unbeantwortet. (...) Die Unterscheidung, wann eine Änderung der Lebensqualität, die durch eine Behandlung hervorgerufen wird, bedeutsam und wichtig oder statistisch signifikant, aber klinisch irrelevant ist, ist gegenwärtig noch weitgehend ungelöst.“

Die Analyse basierte auf der ITT-Population, in der alle randomisierten Patienten eingeschlossen waren. Die Analyse wurde nur für die Zyklen 2, 3 und 4 sowie zum Zeitpunkt „Ende der Therapie“ (EdT) durchgeführt, da bei späteren Zyklen die Patientenzahlen keine aussagekräftigen Analysen erlaubten.

Die Patienten wurden in der Gruppe analysiert, in die sie randomisiert wurden, unabhängig von der wirklichen Behandlung.

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

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

Tabelle 4-39: Bewertung des Verzerrungspotenzials für gesundheitsbezogene Lebensqualität anhand EORTC QLQ-C30 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
CORRECT 14387	Niedrig	Ja	Nein	Nein	Nein	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial des Endpunktes gesundheitsbezogene Lebensqualität (HRQoL) wird als hoch eingestuft. Die Untersucher und Erheber der Endpunkte waren verblindet gegenüber der Behandlung der Patienten, die Auswertung fand innerhalb der Sicherheitspopulation statt und alle Ergebnisse wurden adäquat berichtet. Von den 760 randomisierten Patienten hatten 697 (92 %) analysierbare Daten anhand des EORTC QLQ-C30-Fragebogens zur Baseline-Visite; 565 (74 %) hatten analysierbare Daten zum Zyklus 2, 256 (34 %) zum Zyklus 3, 206 (27 %) zum Zyklus 4, und 412 (54 %) zum jeweiligen Ende der Therapie. Ab Zyklus 5 wurden die Patientenzahlen pro Gruppe zu klein, um verwertbare Analysen der Lebensqualität durchzuführen. Die abnehmende Zahl der Fragebogen ausfüllender Patienten ist nicht ungewöhnlich, da im Verlauf der Studie die Ausfüllraten bedingt durch verschiedene Gründe, hauptsächlich durch Beendigung der Studie oder Eintritt des Todes, im allgemeinen rückläufig sind. Damit gab es eine Abweichung von der adäquaten Umsetzung des ITT-Prinzips, da für mehr als fünf Prozent der ITT-Population keine Daten vorlagen. Der häufigste Grund, warum die Fragebogen den Patienten sowohl in der Regorafenib- als auch in der Placebogruppe nicht ausgehändigt wurden, waren „Institution Errors“ sowie „Other“. Generell gab es jedoch zwischen den beiden Behandlungsgruppen keine deutlichen Unterschiede in der Anzahl der Patienten, die die Fragebogen zur Lebensqualität nicht ausgefüllt haben (siehe Clinical Study Report, Tabelle 14.2/29) (40).

Die Tabelle 4-40 stellt die Ergebnisse für HRQoL ermittelt mit dem Fragebogen der EORTC QLQ-C30 anhand des Vergleiches der mittleren Differenzen zum Ausgangswert über die Behandlungszyklen bis zum Zeitpunkt „Ende der Therapie“ dar.

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

Tabelle 4-40: Ergebnisse für Gesundheitsbezogene Lebensqualität anhand EORTC QLQ-C30 - RCT mit dem zu bewertenden Arzneimittel – Veränderungen im EORTC QLQ-C30 von Baseline an Zyklen 2, 3, 4 und Ende der Therapie (ITT)

Studie	Parameter	n	Regorafenib + BSC (N=505) Mean ± STD	n	Placebo + BSC (N=255) Mean ± STD	Differenz (95 % KI)
COR- RECT 14387	Körperliche Funktionalität					
	Zyklus 2	371	-8,57±17,38	194	-5,60±15,25	-2,96 (-5,86; -0,06)
	Zyklus 3	197	-6,64±17,84	59	-5,36±14,98	-1,29 (-6,32; 3,75)
	Zyklus 4	175	-5,88±18,94	31	-4,03±16,89	-1,84 (-9,01; 5,32)
	Ende der Therapie (EdT)	270	-16,47±21,63	142	-14,51±20,21	-1,96 (-6,27; 2,35)
	Rollenfunktionalität					
	Zyklus 2	372	-13,44±28,57	193	-9,84±25,08	-3,60 (-8,38; 1,18)
	Zyklus 3	198	-9,68±27,34	59	-4,52±20,49	-5,16 (-12,74; 2,42)
	Zyklus 4	175	-11,52±27,31	31	-4,84±25,16	-6,69 (-17,06; 3,69)
	EdT	270	-24,07±31,07	142	-19,48±27,65	-4,59 (-10,69; 1,51)
	Emotionale Funktionalität	369	-2,48±18,42	189	-0,50±19,44	-1,98 (-5,28; 1,32)
	Zyklus 2	197	-2,28±18,36	57	0,58±16,36	-1,70 (-7,01; 3,61)
	Zyklus 3	173	-1,85±21,32	31	2,96±20,70	-4,80 (-12,97; 3,36)
	Zyklus 4	268	-10,63±22,43	142	-8,06±19,87	-2,57 (-6,98; 1,83)
	EdT					
	Soziale Funktionalität	369	-7,23±25,46	189	-2,20±23,12	-5,02 (-9,36; -0,68)
	Zyklus 2	197	-5,58±23,51	57	1,17±16,92	-6,75 (-13,33; -0,17)
	Zyklus 3	173	-7,90±27,23	31	5,38±14,52	-13,28 (-23,18; -3,38)
	Zyklus 4	268	-20,02±31,29	142	-11,50±25,74	-8,52 (-14,54; -2,51)
	EdT					
Kognitive Funktionalität	369	-3,48±16,69	189	-1,41±18,06	-2,07 (-5,08; 0,95)	
Zyklus 2	197	-2,62±15,06	57	-1,17±14,04	-1,45 (-5,85; 2,94)	
Zyklus 3	173	-2,99±14,12	31	-2,15±14,10	-0,84 (-6,26; 4,59)	
Zyklus 4	269	-9,91±21,47	142	-6,34±18,70	-3,58 (-7,77; 0,62)	
EdT						
Globaler Gesund- heitsstatus						
Zyklus 2	367	-8,20±20,59	189	-6,61±19,40	-1,58 (-5,13; 1,97)	
Zyklus 3	196	-6,55±20,94	55	-8,64±17,93	2,09 (-4,02; 8,20)	
Zyklus 4	172	-7,41±19,65	31	-6,45±18,72	-0,96 (-8,47; 6,55)	

EdT	266	-14,94±21,05	140	-16,07±21,71	1,13 (-3,24; 5,50)
-----	-----	--------------	-----	--------------	--------------------

(Quelle: CORRECT 14387 Studienbericht: Tabelle 9-10, Tabelle 14.2/26 und ZA 2: Tabelle 16.1.9.2/1)

Der durchschnittliche EORTC QLQ-C30 Score in der ITT-Population als Ausgangswert (Zyklus 1, Tag 1) betrug 62,64 in der Regorafenib + BSC-Gruppe und 64,65 in der Placebo + BSC-Gruppe. Der durchschnittliche Score am Ende der Therapie (EdT) betrug 48,94 bzw. 51,85.

Legt man der Auswertung die Einteilung in „wenig“, „moderat“ und „viel“ zugrunde (44), lässt sich zusammenfassend sagen, dass es bei den Erhebungen *während* der Behandlung (Zyklen 2-4) in der Placebo + BSC-Gruppe an sieben von 18 Zeitpunkten zu Verschlechterungen der Lebensqualität in der Kategorie „wenig“ kam, während es in der Regorafenib + BSC-Gruppe an zehn Zeitpunkten zu Verschlechterungen in der Kategorie „wenig“ und an zwei Zeitpunkten zu Verschlechterungen in der Kategorie „moderat“ kam. Betroffen waren in beiden Behandlungsgruppen die Module physische Funktionalität, Rollenfunktionalität, soziale Funktionalität und globaler Gesundheitsstatus. Zu keinen Änderungen in beiden Gruppen kam es in den Modulen emotionale Funktionalität und kognitive Funktionalität.

Zwischen den Gruppen gab es in zwei Modulen (Rollenfunktionalität und soziale Funktionalität) zu vier Zeitpunkten Unterschiede, die in die Kategorie „wenig“ fallen, und zu einem Zeitpunkt (Zyklus 4) im Modul soziale Funktionalität einen Unterschied in der Kategorie „moderat“. Zu allen anderen Zeitpunkten war in den Modulen kein Unterschied zwischen den Gruppen feststellbar.

Zum Untersuchungszeitpunkt „Ende der Therapie“ wurde nochmals die Lebensqualität erhoben. Die medikamentöse Behandlung wurde bei der Mehrheit der Patienten aufgrund einer Progression der Erkrankung, aber auch aufgrund unerwünschter Ereignisse oder anderer Gründe beendet. Die Erhebung fand 14 Tage nach der letzten Medikamenteneinnahme statt. Hierbei kam es in allen Kategorien zu einer Verschlechterung der Lebensqualität: In der Placebo + BSC-Gruppe in zwei Modulen in der Kategorie „wenig“ und in vier Modulen in der Kategorie „moderat“; in der Regorafenib + BSC-Gruppe in einem Modul in der Kategorie „wenig“, in drei Modulen in der Kategorie „moderat“ und in zwei Kategorien (Rollenfunktionalität, Soziale Funktion) in der Kategorie „viel“.

Der Unterschied zwischen den beiden Gruppen zu diesem Erhebungszeitpunkt betrug jeweils weniger als 5 Punkte, außer im Modul „soziale Funktionalität“, wo der Unterschied zwischen den beiden Gruppen in die Kategorie „wenig“ fällt (Δ 8,52 Punkte).

Bei Betrachtung der absoluten Unterschiede mit ihren errechneten 95 % Konfidenzintervallen und der sich daraus ergebenden Ableitung eines signifikanten Unterschiedes gibt es einen solchen zu einem Zeitpunkt (Zyklus 2) im Modul „Körperliche Funktion“ und zu allen Zeitpunkten im Modul „Soziale Funktion“. Dieses Modul besteht im Fragebogen aus 2 Fragen, eine Beeinträchtigung des Familienlebens bzw. das Zusammensein mit anderen

Menschen betreffend. Hier muss konstatiert werden, dass eine Therapie mit Regorafenib offenbar zu einer Beeinträchtigung dieser beiden sozialen Bereiche führen kann.

Der Abfall der Scorewerte in beiden Gruppen in allen Modulen zum Zeitpunkt „Ende der Therapie“ ist vor dem Hintergrund zu sehen, dass ein Fortschreiten der Erkrankung bzw. die Notwendigkeit einer Therapiebeendigung aufgrund von unerwünschten Ereignissen Umstände sind, die *per se* mit physischen und insbesondere psychischen Beeinträchtigungen einhergehen.

Zusammenfassend lässt sich konstatieren, dass die durchschnittlichen Veränderungen vom Ausgangswert im Globalen Gesundheitsstatus und in vier von fünf funktionalen Dimensionen eine Verschlechterung der Lebensqualität der Patienten von ähnlichem Ausmaß in der Regorafenib + BSC- und der Placebo + BSC-Gruppe zeigten, wobei die Veränderungen zum Ausgangswert in der Regorafenib + BSC-Gruppe etwas ausgeprägter waren als in der Placebo + BSC-Gruppe. Es gab jedoch keine gegenläufigen Tendenzen in den beiden Behandlungsgruppen. Dem gegenüber steht eine stärkere Beeinträchtigung der sozialen Funktion, das Familienleben bzw. das Zusammensein mit anderen Menschen betreffend, unter einer Therapie mit Regorafenib.

Die Erkrankung des metastasierten kolorektalen Karzinoms als solche mit all seinen physischen und psychischen Auswirkungen beeinträchtigt über den Verlauf hinweg die Lebensqualität negativ. Die Gabe des Medikamentes Regorafenib hatte auf diesen negativen Verlauf keinen gegensteuernden Effekt, trug jedoch, trotz auftretender Toxizitäten, nicht zu einer darüber hinausgehenden relevanten Verschlechterung der Lebensqualität bei. Ausgenommen hiervon ist das Familienleben und das Zusammensein mit anderen Menschen („Soziale Funktion“), wo es zu einer stärkeren Beeinträchtigung unter einer Therapie mit Regorafenib kam.

Analyse der Fläche unter der Kurve

Die Analyse der Zeit-adjustierten Fläche unter der Kurve für den EORTC QLQ-C30 wird in Tabelle 4-41 gezeigt. Die Wirkung auf die Lebensqualität der Patienten war ähnlich in den Gruppen Regorafenib + BSC und Placebo + BSC (Überlappung der Konfidenzintervalle). Unterschiede bis fünf Punkte gelten als klinisch nicht relevant; somit ist das Ergebnis in der Betrachtung der Fläche unter der Kurve (-1,19 Punkte) als ebenfalls nicht klinisch relevant anzusehen.

Tabelle 4-41: Ergebnisse für Behandlungseffekte in der Gesundheitsbezogenen Lebensqualität anhand EORTC QLQ-Gesamtskalen der Zeit-adjustierten Fläche unter der Kurve (ITT)

Studie	Behandlungsarm	n	Mittelwert der Fläche (95 %-KI)	Differenz: Regorafenib – Placebo (95 %-KI)
CORRECT 14387	Regorafenib + BSC	505	56,93 (54,79; 59,08)	-1,19 (-3,13; 0,75)
	Placebo + BSC	255	58,13 (55,72; 60,53)	

(Quelle: CORRECT 14387 Studienbericht: Tabelle 14.2/32)

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie herangezogen wird.

4.3.1.3.7 Endpunkt Verträglichkeit: 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 4-42: Operationalisierung von unerwünschten Ereignissen

Studie	Operationalisierung
CORRECT 14387	<p>Alle Patienten, die mindestens eine Dosis Studienmedikation erhalten haben, wurden in die Analyse eingeschlossen (Sicherheitspopulation). Patienten wurden in die Analyse eingeschlossen wie behandelt, unabhängig von der Randomisierung.</p> <p>Unerwünschte Ereignisse (UE), einschließlich Labor-UE, wurden nach dem MedDRA® Kodiersystem (Version 8.1 oder höher) kodiert und nach den Toxizitätskriterien CTCAE (Common Terminology Criteria for Adverse Events) Version 3.0 des NCI (National Cancer Institute) in Grad 1-5 eingestuft (wobei ein höherer Grad einen höheren Schweregrad bedeutet).</p> <p>Ein UE war eine Reaktion, eine Nebenwirkung oder ein anderes unerwünschtes medizinisches Ereignis, unabhängig von einer Verbindung zur Studienmedikation, das zu jeder Zeit vom Beginn der Verabreichung der Studienmedikation in Zyklus 1 bis zur letzten Studiervisite oder bis zum Studienabbruch auftrat. Ein UE konnte jegliches ungünstige und unbeabsichtigte Zeichen (einschließlich abnormer Laborergebnisse), Symptome oder eine Erkrankung sein, egal ob in kausalem Zusammenhang mit dem medizinischen Produkt stehend oder nicht.</p> <p>Ein schwerwiegendes unerwünschtes Ereignis (SUE) ist durch die International Conference of Harmonization (ICH) definiert als jegliches ungünstiges medizinisches Ereignis, das bei jeglicher Dosierung zum Tod führt, lebensbedrohlich ist, einen stationären Krankenhausaufenthalt oder die Verlängerung eines bestehenden Krankenhausaufenthaltes erforderlich machte zu einer andauernden oder erheblichen Behinderung/Beeinträchtigung oder einer angeborenen Anomalie oder zu einem Geburtsschaden führt.</p> <p>UE/SUE wurden ab der ersten Behandlung zu allen im Prüfplan vorgesehenen Visiten bis 30 Tage nach Einnahme der letzten Dosis der Studienmedikation erfasst, ebenfalls bei Abbruch der Behandlung (welche zu jeder geplanten oder ungeplanten Visite erfolgen konnte).</p> <p>Eine Nachbeobachtung der UE war für die folgenden 30 Tage nach der letzten Dosis erforderlich, um jegliche neue oder fortlaufende medikamentenbezogene UE oder SUE zu bestimmen, ungeachtet eines Zusammenhangs mit dem Medikament. Die Nachbeobachtung durch die Zentren konnte telefonisch erfolgen und musste in den Quelldokumenten verzeichnet sein.</p> <p>Ein Überblick der aufgetretenen UE/SUE wurde in Inzidenztabelle dargestellt, zusammengefasst nach Häufigkeit (mindestens 10 % bei UE, mindestens 2 % bei SUE) oder Grad (insbesondere Grad 3, 4 oder 5).</p>

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

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

Tabelle 4-43: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse 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
CORRECT 14387	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial des Endpunktes Verträglichkeit wird als niedrig eingestuft. Die Untersucher und Erheber der Endpunkte waren gegenüber der Behandlung der Patienten verblindet, die Auswertung wurde auf der Basis der Sicherheitspopulation durchgeführt und alle Ergebnisse wurden so berichtet, wie sie im SAP und Prüfplan zur Studie vordefiniert waren.

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

Inzidenz von UE

Tabelle 4-44: Ergebnisse für unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel mit abgeleitetem relativem Risiko

CORRECT	Regorafenib + BSC	Placebo + BSC	Relatives Risiko (95 %-KI)^a
Sicherheitspopulation	n=500 n (%)	n=253 n (%)	
Patienten mit UE	498 (99,6)	245 (96,8)	Placebo + BSC vs. Regorafenib + BSC 0,97 (0,95; 0,99)
CTCAE Grad =1	12 (2,4)	45 (17,8)	Regorafenib + BSC vs. Placebo + BSC 0,13 (0,007; 0,25)
CTCAE Grad =2	96 (19,2)	76 (30,0)	Regorafenib + BSC vs. Placebo + BSC 0,64 (0,49; 0,83)
CTCAE Grad =3	280 (56,0)	67 (26,5)	Placebo + BSC vs. Regorafenib + BSC 0,47 (0,38; 0,59)
CTCAE Grad =4	43 (8,6)	20 (7,9)	Placebo + BSC vs. Regorafenib + BSC 0,92 (0,55; 1,53)
CTCAE Grad =5	67 (13,4)	37 (14,6)	Regorafenib + BSC vs. Placebo + BSC 0,92 (0,63; 1,33)
Patienten mit SUE	219 (43,8)	100 (39,5)	Placebo + BSC vs. Regorafenib + BSC 0,90 (0,75; 1,08)
Patienten mit UE, die zum Studienabbruch geführt haben	88 (17,6)	32 (12,6)	Placebo + BSC vs. Regorafenib + BSC 0,72 (0,49; 1,05)
Patienten mit UE, die eine Dosismodifikation erfuhren	333 (66,6)	57 (22,5)	Placebo + BSC vs. Regorafenib + BSC 0,34 (0,27; 0,43)
Die Punktschätzer basieren auf den beobachteten Raten a: Post-hoc Analyse: 14387_ZA1 (Quelle: CORRECT Studienbericht Text Tabelle 10-2 und SAS-Output Tabelle 14.3.1/1)			

Tabelle 4-45: Überblick über UE nach CTCAE-Grad und Zusammenfassung von UE-Kategorien in der CORRECT-Studie in RCT mit dem zu bewertenden Arzneimittel und abgeleiteter Odds Ratio

CORRECT	Regorafenib + BSC	Placebo + BSC	Regorafenib vs. Placebo
Sicherheitspopulation	n=500 n (%)	n=253 n (%)	Odds ratio (95 %-KI) p-Wert^a
Patienten mit UE	498 (99,6)	245 (96,8)	8,13 (1,71; 38,6) 0,0083
CTCAE-Grad 1	12 (2,4)	45 (17,8)	0,11 (0,06; 0,22) <0,0001
CTCAE-Grad 2	96 (19,2)	76 (30,0)	0,55 (0,39; 0,78) 0,0009
CTCAE-Grad 3	280 (56,0)	67 (26,5)	3,53 (2,54; 4,92) <0,0001
CTCAE-Grad 4	43 (8,6)	20 (7,9)	1,10 (0,63; 1,91) 0,7450
CTCAE-Grad 5	67 (13,4)	37 (14,6)	0,90 (0,59; 1,39) 0,6456
CTCAE Grad ≥ 3	390 (78,0)	124 (49,0)	3,69 (2,67; 5,10) <0,0001
CTCAE Grad ≥ 4	110 (22,0)	57 (22,5)	0,97 (0,68; 1,39) 0,8684
Patienten mit SUE	219 (43,8)	100 (39,5)	1,19 (0,88; 1,62) 0,2624
Patienten mit UE, die zum Studienabbruch geführt haben	88 (17,6)	32 (12,6)	1,48 (0,95; 2,28) 0,0809
Patienten mit UE, die eine Dosismodifikation erfuhren	333 (66,6)	57 (22,5)	6,86 (4,84; 9,72) <0,0001
Die Punktschätzer basieren auf den beobachteten Raten a – Post-hoc Analyse: 14387_ZA1 (Quellen: CORRECT 14387 Studienbericht Text Tabelle 10-2 und SAS-Output Tabelle 14.3.1/1)			

Mindestens ein UE wurde während der gesamten Studie von den meisten Patienten der Sicherheitspopulation (SAF) berichtet: 498 (99,6 %) in der Regorafenib + BSC-Gruppe und 245 (96,8 %) in der Placebo + BSC-Gruppe.

Aus der Berechnung der relativen Risiken in den einzelnen UE-Kategorien (Tabelle 4-44) lässt sich konstatieren, dass es in der Placebo + BSC-Gruppe signifikant weniger Patienten mit einem UE (jeglichen Grades), mit einer Grad 3-Toxizität sowie bei Patienten, die eine Dosismodifikation in Zusammenhang mit einem UE erfuhren, gab. Demgegenüber gab es

signifikant weniger Patienten in der Regorafenib + BSC-Gruppe mit einer Grad 1 bzw. Grad 2-Toxizität.

Bei den Patienten mit einer Grad 4- bzw. Grad 5-Toxizität, Patienten mit einem SUE sowie bei den Patienten, deren Studienbehandlung in Zusammenhang mit einem UE dauerhaft abgebrochen wurde, gab es keine signifikanten Unterschiede.

Dies wird bestätigt durch die Analyse auf der Basis der berechneten Odds Ratios (Tabelle 4-45): Für Regorafenib + BSC wurden signifikant mehr Patienten mit mindestens einem UE (OR: 8,13, 95 %-KI 1,71; 38,6, $p=0,0083$) und mit mindestens einem UE mit CTCAE Grad ≥ 3 als unter Placebo + BSC-Behandlung berichtet (OR: 3,69, 95 %-KI 2,67; 5,10, $p<0,0001$). Die Anzahl der Patienten mit UE mit CTCAE Grad ≥ 4 , sowie Patienten mit mindestens einem SUE oder einem UE, dass zum Studienabbruch geführt hat, war unter Regorafenib + BSC nicht signifikant höher als unter Placebo + BSC.

Während UE vom Grad 1 und 2 signifikant häufiger in der Placebo + BSC-Gruppe registriert wurden (Grad 1: 17,8 % vs. 2,4 %; Grad 2: 30,0 % vs. 19,2 %), waren UE vom Grad 3 signifikant häufiger in der Regorafenib + BSC-Gruppe (56,0 % vs. 26,5 %). Dies ist im Wesentlichen auf das substanzspezifische Nebenwirkungsprofil von Regorafenib zurückzuführen, bei dem spezifische Nebenwirkungen oft den Grad 3 erreichen, aber selten darüber hinausgehen. Entsprechend war die Inzidenz von UE der Grade 4 und 5 ähnlich in der Regorafenib + BSC-Gruppe wie in der Placebo + BSC-Gruppe und nicht signifikant unterschiedlich (8,6 % vs. 7,9 % bzw. 13,4 % vs. 14,6 %). Generell wurden die Unterschiede in den UE durch die Inzidenzen im Grad 3 bedingt, während die Unterschiede in den Graden 4 und 5 zwischen den beiden Gruppen klein, statistisch nicht signifikant und klinisch nicht relevant sind.

Obwohl signifikant mehr Patienten im Regorafenib + BSC-Arm Dosismodifikationen aufgrund eines UE hatten (66,6 % vs. 22,5 %), war der Unterschied in den UE, die zu einer dauernden Beendigung der Therapie führten, relativ klein (17,6 % vs. 12,6 %) und statistisch nicht signifikant. Dies weist darauf hin, dass die meisten UE mit Dosismodifikationen beherrscht werden können und nicht zur dauernden Therapiebeendigung führten. Die durchschnittlich applizierte Tagesdosis von 147 mg /Tag ist nah an der Standarddosierung von 160 mg/Tag, so dass man hier schlussfolgern kann, dass der überwiegende Teil der Dosismodifikationen gering und vorübergehend war.

Tabelle 4-46 zeigt alle UE jeglichen Grades (nach CTCAE und schwerstem aufgetretenen Grad), die in mindestens 10 % eines Armes aufgetreten sind.

Tabelle 4-46: Studie 14387 (CORRECT) - Inzidenzen von unerwünschten Ereignissen (alle Grade), die in mind. 10 % aller Patienten in einer der Behandlungsgruppen auftraten (SAF)

Studie	Unerwünschte Ereignisse in mind. 10 % aller Patienten (SAF)			
	Kategorie	Regorafenib + BSC (N=500) n (%)	Placebo + BSC (N=253) n (%)	Odds ratio Regorafenib vs. Placebo (95 %-KI) p-Wert [#]
CORRECT 14387	Jegliches UE ^a	498 (99,6)	245 (96,8)	8,13 (1,71; 38,57) 0,0083
	Blut/Knochenmark	150 (30,0)	37 (14,6)	2,50 (1,68; 3,72) <0,0001
	Anämie	72 (14,4)	30 (11,9)	1,25 (0,79; 1,97) 0,3364
	Thrombozytopenie	78 (15,6)	6 (2,4)	7,61 (3,27; 17,71) <0,0001
	Herz-Kreislauf	165 (33,0)	29 (11,5)	3,80 (2,48; 5,84) <0,0001
	Hypertonie	152 (30,4)	20 (7,9)	5,09 (3,10; 8,35) <0,0001
	Allgemein	423 (84,6)	159 (62,8)	3,25 (2,28; 4,62) <0,0001
	Konstitutionelle Symptome – andere (Spezifizierung)	53 (10,6)	33 (13,0)	0,79 (0,50; 1,26) 0,3201
	Fatigue	317 (63,4)	117 (46,2)	2,01 (1,48; 2,74) <0,0001
	Fieber	142 (28,4)	39 (15,4)	2,18 (1,47; 3,22) 0,0001
	Gewichtsverlust	160 (32,0)	28 (11,1)	3,78 (2,45; 5,84) <0,0001
	Dermatologie/Haut	356 (71,2)	58 (22,9)	8,31 (5,85; 11,81) <0,0001
	Trockene Haut	50 (10,0)	10 (4,0)	2,70 (1,35; 5,42) 0,0052
	Hand-Fuß-Haut Reaktion	235 (47,0)	19 (7,5)	10,9 (6,63; 18,00) <0,0001
	Exanthem	145 (29,0)	13 (5,1)	7,54 (4,18; 13,61) <0,0001
	Gastrointestinaltrakt	416 (83,2)	166 (65,6)	2,60 (1,83; 3,68) <0,0001
	Anorexie	234 (46,8)	72 (28,5)	2,21 (1,60; 3,06)

			<0,0001
Obstipation	119 (23,8)	48 (19,0)	1,33 (0,92; 1,94) 0,1328
Diarrhoe	214 (42,8)	43 (17,0)	3,65 (2,52; 5,31) <0,0001
Mucositis (funktionelle /symptomatisch), Mundhöhle	144 (28,8)	12 (4,7)	8,12 (4,41; 14,97) <0,0001
Übelkeit	112 (22,4)	55 (21,7)	1,04 (0,72; 1,50) 0,8372
Erbrechen	80 (16,0)	41 (16,2)	0,99 (0,65; 1,49) 0,9421
Blutungen	102 (20,4)	17 (6,7)	3,56 (2,08; 6,09) <0,0001
Infektionen	126 (25,2)	36 (14,2)	2,03 (1,35; 3,05) 0,0006
Lymphatisches System	61 (12,2)	18 (7,1)	1,81 (1,05; 3,14) 0,0335
Metabolisch/Laborwerte	228 (45,6)	62 (24,5)	2,58 (1,85; 3,61) <0,0001
Bilirubin (Hyperbilirubinämie)	100 (20,0)	24 (9,5)	2,39 (1,49; 3,83) 0,0003
Neurologie	135 (27,0)	63 (24,9)	1,12 (0,79; 1,58) 0,5367
Neuropathie: sensorisch	50 (10,0)	25 (9,9)	1,01 (0,61; 1,68) 0,9591
Schmerzen	328 (65,6)	134 (53,0)	1,69 (1,24; 2,31) 0,0008
Bauchschmerzen	122 (24,4)	47 (18,6)	1,42 (0,97; 2,06) 0,0713
Rückenschmerzen	75 (10,05)	26 (10,3)	1,54 (0,96; 2,48) 0,0740
Kopfschmerzen	52 (10,4)	18 (7,1)	1,52 (0,87; 2,65) 0,1448
Lunge/Obere Atemwege	252 (50,4)	70 (27,7)	2,66 (1,92; 3,68) <0,0001
Husten	55 (11,0)	28 (11,1)	0,99 (0,61; 1,61) 0,9778
Atemnot (Kurzatmigkeit)	91 (18,2)	33 (13,0)	1,48 (0,96; 2,28) 0,0729

Stimmveränderungen n	160 (32,0)	16 (6,3)	6,97 (4,06; 11,96) <0,0001
a – Patienten, bei denen ein UE mehrmals auftrat, wurde das UE nur einmal mit dem schwersten Grad gezählt.			
# Post-hoc Analyse: Odds ratios, 95 %-KI und p-Wert berechnet auf Basis des Wald-Tests (Software: SAS 9.2 (SAS Institute, Cary, NC, USA); 14387 ZA1			
(Quelle: CORRECT 14387 Studienbericht: Tabelle 10-3 und Tabelle 14.3.1/3)			

Die Inzidenz der folgenden Einzel-UE (jeglicher Grad, nach Nomenklatur der CTCAE-Klassifikation) war signifikant und um mindestens zehn Prozentpunkte höher in der Regorafenib + BSC-Gruppe als in der Placebo + BSC-Gruppe:

- Fatigue (63,4 % vs. 46,2 %)
- Hand-Fuß-Hautreaktion (47,0 % vs. 7,5 %)
- Anorexie (46,8 % vs. 28,5 %)
- Diarrhoe (42,8 % vs. 17,0 %)
- Gewichtsverlust (32,0 % vs. 11,1 %)
- Stimmveränderungen (32,0 % vs. 6,3 %)
- Hypertonie (30,4 % vs. 7,9 %)
- Exanthem (29,0 % vs. 5,1 %)
- Mukositis (funktionelle/symptomatische), Mundhöhle (28,8 % vs. 4,7 %)
- Fieber (28,4 % vs. 15,4 %)
- Hyperbilirubinämie (20,0 % vs. 9,5 %)
- Thrombozytopenie (15,6 % vs. 2,4 %)
- Blutungen (alle Organsysteme) (20,4 % vs. 6,7 %)
- Infektionen (alle Organsysteme) (25,2 % vs. 14,2 %)

Inzidenz von UE vom Grad 3, 4 und 5

Die Inzidenz der UE von Grad 3 und 4, die in mindestens 2 % der Patienten eines Therapiearmes auftraten, sind in Tabelle 4-47 zusammengefasst.

Tabelle 4-47: Inzidenzen von UE von CTCAE-Grad 3 und 4, die in mdst. 2 % aller Patienten in einer der Behandlungsgruppen in RCT mit dem zu bewertenden Arzneimittel auftraten

Studie	UE von CTCAE-Grade 3 und 4 in mdst. 2 % der Patienten				
	Kategorie NCI CTCAE-Terminus	Regorafenib+ BSC (N=500) n (%)		Placebo + BSC (N=253) n (%)	
		Grad 3	Grad 4	Grad 3	Grad 4
CORRECT 14387	Jegliches UE	280 (56,0)	43 (8,6)	67 (26,5)	20 (7,9)
	Blut/Knochenmark	49 (9,8)	6 (1,2)	8 (3,2)	1 (0,4)
	Anämie	27 (5,4)	2 (0,4)	8 (3,2)	1 (0,4)

Thrombozytopenie	17 (3,4)	2 (0,4)	1 (0,4)	0
Herz-Kreislauf	40 (8,0)	2 (0,4)	3 (1,2)	0
Hypertonie	38 (7,6)	0	2 (0,8)	0
Allgemein	87 (17,4)	8 (1,6)	27 (10,7)	8 (3,2)
Fatigue	75 (15,0)	2 (0,4)	21 (8,3)	5 (2,0)
Konstitutionelle Symptome – Andere (Spezifizierung)	11 (2,2)	5 (1,0)	10 (4,0)	3 (1,2)
Gastrointestinaltrakt	101 (20,2)	5 (1,0)	30 (11,9)	3 (1,2)
Anorexie	23 (4,6)	0	11 (4,3)	0
Dehydratation	10 (2,0)	0	6 (2,4)	0
Diarrhoe	41 (8,2)	1 (0,2)	5 (2,0)	0
Mucositis (funktionelle/symptomatisch), Mundhöhle	16 (3,2)	0	0	0
Obstruktion, GI, Dünndarm	3 (0,6)	0	5 (2,0)	0
Hepatobiliär/Pankreas	12 (2,4)	1 (0,2)	4 (1,6)	0
Infektionen	32 (6,4)	8 (1,6)	13 (5,1)	1 (0,4)
Muskuloskeletal/ Weichteile	8 (1,6)	3 (0,6)	5 (2,0)	0
Metabolisch/Laborveränderungen	100 (20,0)	15 (3,0)	22 (8,7)	9 (3,6)
ALP	11 (2,2)	0	4 (1,6)	0
AST	12 (2,4)	0	3 (1,2)	0
Bilirubin (Hyperbilirubinämie)	34 (6,8)	4 (0,8)	11 (4,3)	5 (2,0)
Lipase	15 (3,0)	6 (1,2)	1 (0,4)	1 (0,4)
Hypokalämie	13 (2,6)	0	1 (0,4)	0
Hyponaträmie	18 (3,6)	1 (0,2)	4 (1,6)	0
Hypophosphatämie	23 (4,6)	0	1 (0,4)	0
Neurologie	19 (3,8)	3 (0,6)	10 (4,0)	2 (0,8)
Schmerz	49 (9,8)	2 (0,4)	17 (6,7)	2 (0,8)
Bauchschmerzen	24 (4,8)	0	5 (2,0)	1 (0,4)
Lunge/Obere Atemwege	11 (2,2)	3 (0,6)	12 (4,7)	0
Atemnot (Kurzatmigkeit)	7 (1,4)	3 (0,6)	8 (3,2)	0
Niere/Harnwege	12 (2,4)	0	9 (3,6)	1 (0,4)
Dermatologie/Haut	112 (22,4)	1 (0,2)	2 (0,8)	1 (0,4)
Hand-Fuß	83 (16,6)	0	1 (0,4)	0

Hautreaktion				
Exanthem	29 (5,8)	0	1 (0,4)	0

a - Patienten, bei denen ein UE mehrmals auftrat, wurde das UE nur einmal mit dem schwersten Grad gezählt.
ALP: Alkalische Phosphatase; AST: Aspartataminotransferase
(Quelle: CORRECT 14387 Studienbericht: Tabelle 10-4 und Tabelle 14.3.1/3)

Die Unterschiede in der Inzidenz der UE von Grad 3 und 4 zwischen den zwei Gruppen war hauptsächlich durch die Inzidenz der UE vom Grad 3 bedingt, während es bei den UE vom Grad 4 nur geringe Unterschiede gab. Die Inzidenz der UE vom Grad 3 war höher in der Regorafenib + BSC-Gruppe im Vergleich zur Placebo + BSC-Gruppe (56,0 % vs. 26,5 %). UE vom Grad 4 wurden in beiden Gruppen mit ähnlicher Inzidenz berichtet: 8,6 % in der Regorafenib + BSC-Gruppe und 7,9 % in der Placebo + BSC-Gruppe.

Die häufigsten UE vom Grad 3, die eine höhere Inzidenz (> fünf Prozentpunkte) in der Regorafenib + BSC- im Vergleich zur Placebo + BSC-Gruppe aufweisen, sind die folgenden:

- Hand-Fuß-Hautreaktion (16,6 % vs. 0,4 %)
- Fatigue (15,0 % vs. 8,3 %)
- Diarrhoe (8,2 % vs. 2,0 %)
- Hypertonie (7,6 % vs. 0,8 %)
- Exanthem (5,8 % vs. 0,4 %)

Die Inzidenz der häufigsten (≥ 1 % der Patienten in einer der Therapiearme) UE vom Grad 4 war in beiden Armen ähnlich häufig. Die folgenden UE vom Grad 4 wurden in ≥ 1 % der Patienten in einer der Therapiearme berichtet (Regorafenib + BSC vs. Placebo + BSC):

- Hyperlipasämie (1,2 % vs. 0,4 %)
- Konstitutionelle Symptome, Andere (1,0 % vs. 1,2 %)
- Hyperbilirubinämie (0,8 % vs. 2,0 %)
- Fatigue (0,4 % vs. 2,0 %)

Todesfälle durch UE

Die Inzidenzen der UE, die zum Tod führten, (Grad 5) sind in der Tabelle 4-48 zusammengefasst.

Tabelle 4-48: Inzidenzen von UE, die zum Tode führten, (Grad 5), die bei mdst. einem Patienten in RCT mit dem zu bewertenden Arzneimittel auftraten

Studie	UE mit Grad 5 mit $n \geq 1$			
CORRECT 14387	Kategorie NCI CTCAE-Terminus	Regorafenib + BSC (N=500)	Placebo + BSC (N=253)	Odds ratio
		n (%)	n (%)	Regorafenib vs. Placebo (95 %-KI) p-Wert ^a

Jegliches UE	67 (13,4)	37 (14,6)	0,90 (0,59; 1,39) 0,6456
Herz-Kreislauf	1 (0,2)	1 (0,4)	0,51 (0,03; 8,11) 0,6295
Myokardischämie/-infarkt	1 (0,2)	1 (0,4)	0,51 (0,03; 8,11) 0,6295
Allgemein	20 (4,0)	12 (4,7)	0,84 (0,40; 1,74) 0,6334
Fieber	1 (0,2)	0	1,01 (0,03; 30,27) 0,9945
Konstitutionelle Symptome – Andere (Spezifizierung)	19 (3,8)	12 (4,7)	0,79 (0,38; 1,66) 0,5391
Tod	20 (4,0)	9 (3,6)	1,13 (0,51; 2,52) 0,7657
Tod nicht mit einem CTCAE-Terminus assoziiert, Krankheitsprogression	14 (2,8)	6 (2,4)	1,19 (0,45; 3,12) 0,7301
Tod nicht mit einem CTCAE-Terminus assoziiert, Multiorganversagen	4 (0,8)	1 (0,4)	2,03 (0,23; 18,28) 0,5269
Tod nicht mit einem CTCAE-Terminus assoziiert, plötzlicher Tod	2 (0,4)	2 (0,8)	0,50 (0,07; 3,60) 0,4945
Dermatologie/Haut	0	1 (0,4)	0,25 (0,01; 7,54) 0,4270
Dermatologie – Andere (Spezifizierung)	0	1 (0,4)	0,25 (0,01; 7,54) 0,4270
Gastrointestinaltrakt	1 (0,2)	5 (2,0)	0,10 (0,01; 0,86) 0,0355
Ileus	0	1 (0,4)	0,25 (0,01; 7,54) 0,4270
Obstruktion, Kolon	1 (0,2)	1 (0,4)	0,51 (0,03; 8,11) 0,6295
Obstruktion, Dünndarm	0	2 (0,8)	0,13 (0,01; 2,80) 0,1900
GI – Andere (Spezifizierung)	0	1 (0,4)	0,25 (0,01; 7,54) 0,4270
Blutungen	4 (0,8)	0	4,07 (0,21; 77,33) 0,3498
Blutung, Abdomen	1 (0,2)	0	1,01 (0,03; 30,27)

			0,9945
Blutung, peranal	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Blutung – Andere (Spezifizierung)	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Blutung pulmonal, Bronchus	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Blutung, vaginal	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Hepatobiliär/Pankreas	8 (1,6)	3 (1,2)	1,36 (0,36; 5,15)
			0,6559
Leberdysfunktion	8 (1,6)	3 (1,2)	1,36 (0,36; 5,15)
			0,6559
Infektion	3 (0,6)	3 (1,2)	0,50 (0,10; 2,51)
			0,4021
Infektion (klinisch dokumentiert), Lunge (Pneumonie)	3 (0,6)	2 (0,8)	0,76 (0,13; 4,56)
			0,7617
Infektion – Andere (Spezifizierung)	0	1 (0,4)	0,25 (0,01; 7,54)
			0,4270
Neurologie	2 (0,4)	0	2,03 (0,09; 45,14)
			0,6551
Enzephalopathie	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Neurology – Andere (Spezifizierung)	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Lunge / Obere Atemwege	8 (1,6)	3 (1,2)	1,36 (0,36; 5,15)
			0,6559
ARDS	1 (0,2)	1 (0,4)	0,51 (0,03; 8,11)
			0,6295
Pneumothorax	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Dyspnoe (Kurzatmigkeit)	6 (1,2)	2 (0,8)	1,52 (0,31; 7,61)
			0,6073
Nieren/Harnwege	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Nierenversagen	1 (0,2)	0	1,01 (0,03; 30,27)
			0,9945
Vaskulär	1 (0,2)	1 (0,4)	0,51 (0,03; 8,11)
			0,6295
Vaskulär – Andere (Spezifizierung)	1 (0,2)	0	1,01 (0,03; 30,27)

			0,9945
Thrombose/Embolie	0	1 (0,4)	0,25 (0,01; 7,54)
			0,4270
a – Post-hoc Analyse: Odds ratios, 95 %-KI und p-Wert berechnet auf Basis des Wald-Tests (Software: SAS 9.2 (SAS Institute, Cary, NC, USA); Zero counts were replaced by 0.5 to calculate the effect size. ARDS: Acute respiratory distress syndrome (Quelle: CORRECT 14387 Studienbericht: Tabelle 10-5 und Tabelle 14.3.1/3; 14387 ZA1)			

UE vom Grad 5 waren in beiden Therapiegruppen ähnlich häufig und nicht signifikant unterschiedlich: 13,4 % in der Regorafenib + BSC-Gruppe und 14,6 % in der Placebo + BSC-Gruppe.

Die folgenden UE vom Grad 5 nach den *CTCAE-Kategorien* wurden in ≥ 1 % der Patienten in einem der Therapiearme berichtet (Regorafenib + BSC vs. Placebo + BSC):

- Konstitutionelle Symptome (4,0 % vs. 4,7 %)
- Tod (4,0 % vs. 3,6 %)
- Hepatobiliär/Pankreas (1,6 % vs. 1,2 %)
- Pulmonal / oberer Atemtrakt (1,6 % vs. 1,2 %)
- Infektion (0,6 % vs. 1,2 %)
- Gastrointestinal (0,2 % vs. 2,0 %)

Die folgenden UE vom Grad 5 nach den *NCI CTCAE-Termini* wurden in ≥ 1 % der Patienten in einem der Therapiearme berichtet (Regorafenib + BSC vs. Placebo + BSC):

- Konstitutionelle Symptome, Andere (3,8 % vs. 4,7 %)
- Tod, nicht mit einem CTCAE-Terminus assoziiert, Krankheitsprogression, nicht anders spezifiziert (2,8 % vs. 2,4 %)
- Leberdysfunktion (1,6 % vs. 1,2 %)
- Dyspnoe (1,2 % vs. 0,8 %)

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.

Es wurden keine Metaanalysen durchgeführt, da für die Zusatznutzenbewertung nur eine Studie 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.

4.3.1.3.8 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.

In der CORRECT-Studie wurden die *a priori* im statistischen Analyseplan definierten Subgruppenanalysen durchgeführt, um den Einfluss potentieller Effektmodifikatoren auf den Behandlungseffekt bezüglich des Wirksamkeitsendpunktes des medianen Gesamt- bzw. des medianen progressionsfreien Überlebens zu untersuchen. Für alle anderen beschriebenen Endpunkte wurden die Subgruppenanalysen *post hoc* durchgeführt.

Subgruppenanalysen zum Gesamtüberleben und progressionsfreiem Überleben

Die deskriptive Statistik, die Schätzungen der HR (95 %-KI) sowie die Ergebnisse der Interaktionstests der jeweiligen Subgruppen in den einzelnen Kategorien für OS und PFS finden sich in Tabelle 4-49 und Tabelle 4-50. Eine graphische Darstellung in Form von Forest-Plots erfolgt in Abbildung 4-5 und Abbildung 4-6.

Tabelle 4-49: Subgruppenanalyse für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel

Studie	Subgruppenanalyse für OS (ITT)								
	Subgruppenvariable	N	Anzahl Ereignisse	HR (Regorafenib/Placebo)		Median (Tage)		Interaktionstest p-Wert ^a	
Schätzung				95 %-KI	Reg + BSC	Pla + BSC			
CORRECT 14387	Geschlecht							0,893231	
	Männlich	464	259	0,773	(0,599; 0,998)	204	165		
	Weiblich	296	173	0,751	(0,552; 1,022)	184	146		
	Alter								0,405053

<65 Jahre	475	274	0,716	(0,561; 0,914)	204	150	
≥65 Jahre	285	158	0,856	(0,614; 1,193)	181	169	
Ethnie							0,875678
Kaukasier	593	350	0,757	(0,609; 0,940)	190	149	
Asiatisch	111	48	0,793	(0,435; 1,447)	200	212	
ECOG zum Beginn							0,749761
0	411	204	0,702	(0,530; 0,929)	260	212	
1	349	228	0,773	(0,586; 1,020)	138	109	
Zeit seit der Diagnose der metastatischen Erkrankung							0,877144
<18 Monate	140	90	0,816	(0,532; 1,251)	141	124	
≥18 Monate	620	342	0,760	(0,609; 0,948)	204	165	
Region ^b							0,924554
1	632	377	0,768	(0,622; 0,948)	184	149	
2	104	45	0,790	(0,429; 1,456)	200	212	
3	24	10	0,694	(0,195; 2,466)	-	-	
Vorherige Anti- VEGF-Therapie							-
Ja	760	432	0,767	(0,630; 0,933)	196	151	
Vorherige antitumoröse Therapie							0,560518
FOIB	375	214	0,825	(0,625; 1,089)	187	162	
FOIBE	385	218	0,710	(0,538; 0,938)	205	150	
Anzahl der Therapielinien							0,448184
≤3	301	175	0,709	(0,521; 0,967)	190	127	
>3	459	257	0,804	(0,624; 1,038)	198	169	
Anzahl der Therapielinien bei oder nach Diagnose der metastatischen Erkrankung							0,890635
≤3	395	220	0,788	(0,599; 1,038)	204	151	
>3	365	212	0,747	(0,564; 0,988)	191	154	
KRAS Mutation							0,236344
Nein	299	204	0,653	(0,476; 0,895)	219	150	
Ja	430	228	0,867	(0,670; 1,123)	187	154	
Primärlokalisierung							0,179753

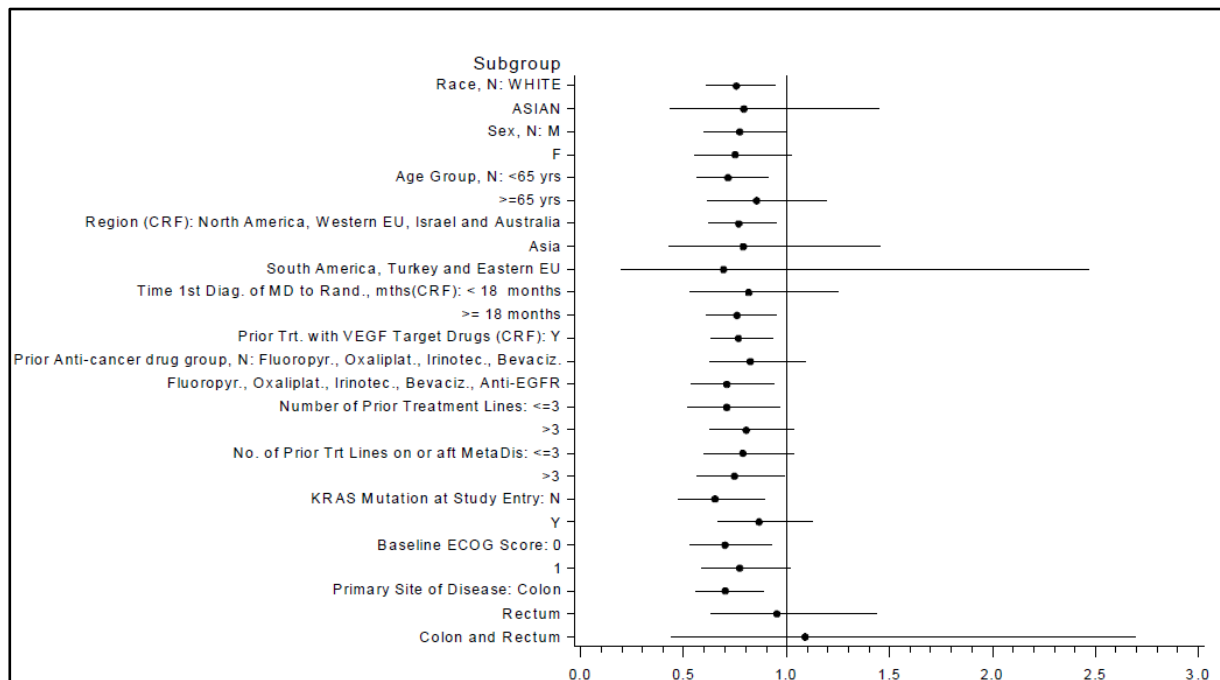
der Erkrankung							
Kolon	495	301	0,703	(0,557; 0,887)	184	140	
Rektum	220	109	0,953	(0,633; 1,436)	246	237	
Kolon und Rektum	44	22	1,091	(0,441; 2,697)	208	220	

a – Der p-Wert des Interaktionstests vergleicht die beiden in der jeweiligen kategorie genannten Subgruppen. Dieser p-Wert ist zweiseitig.

b – Region 1: Noramerika, Westeuropa, Israel und Australien. Region 2: Asien. Region 3: Südamerika Türukei und Osteuropa.

FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper

(Quelle: CORRECT 14387 Studienbericht: Tabelle 9-16, Tabelle 14.2/5 und 14387_ZA4: Tabelle 16.1.9.2/1)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiater, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diag. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastasierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 Studienbericht: Grafik 16.1.9.2/10)

Abbildung 4-5: Forest-Plot der Subgruppenanalyse für Gesamtüberleben (ITT) in RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-50: Subgruppenanalyse für PFS in RCT mit dem zu bewertenden Arzneimittel

Studie		Subgruppenanalyse für PFS (ITT)						
CORRECT 14387	Subgruppen- variable	N	An- zahl Ereig- nisse	HR (Regorafenib/Placebo)		Reg + BSC	Pla + BSC	Interak- tionstest p-Wert ^a
				Schät- zung	95 %-KI			
Geschlecht								0,262147
	Männlich	464	405	0,536	(0,434; 0,662)	62	53	
	Weiblich	296	266	0,438	(0,337; 0,570)	57	51	
Alter								0,014384
	<65 Jahre	475	428	0,418	(0,340; 0,514)	59	51	
	≥65 Jahre	285	243	0,651	(0,496; 0,855)	58	54	
Ethnie								0,770724
	Kaukasier	593	527	0,503	(0,418; 0,605)	59	52	
	Asiatisch	111	94	0,444	(0,285; 0,690)	59	51	
ECOG zum Beginn								0,268617
	0	411	359	0,444	(0,355; 0,556)	70	53	
	1	349	312	0,569	(0,447; 0,723)	57	50	
Zeit von der Diagnose der metastatischen Erkrankung								0,423393
	<18 Monate	140	132	0,582	(0,405; 0,838)	56	51	
	≥18 Monate	620	539	0,484	(0,403; 0,581)	62	52	
Region ^b								0,727581
	1	632	564	0,500	(0,418; 0,599)	59	52	
	2	104	90	0,433	(0,277; 0,679)	61	51	
	3	24	17	0,576	(0,199; 1,664)	106	56	
Vorherige Anti- VEGF-Therapie								-
	Ja	760	671	0,501	(0,425; 0,590)	59	52	
Vorherige antitumoröse Therapie								0,947180
	FOIB	375	336	0,514	(0,409; 0,646)	56	51	
	FOIBE	385	335	0,497	(0,393; 0,629)	63	53	
Anzahl der Therapielinien								0,740498
	≤3	301	278	0,523	(0,404; 0,676)	56	51	

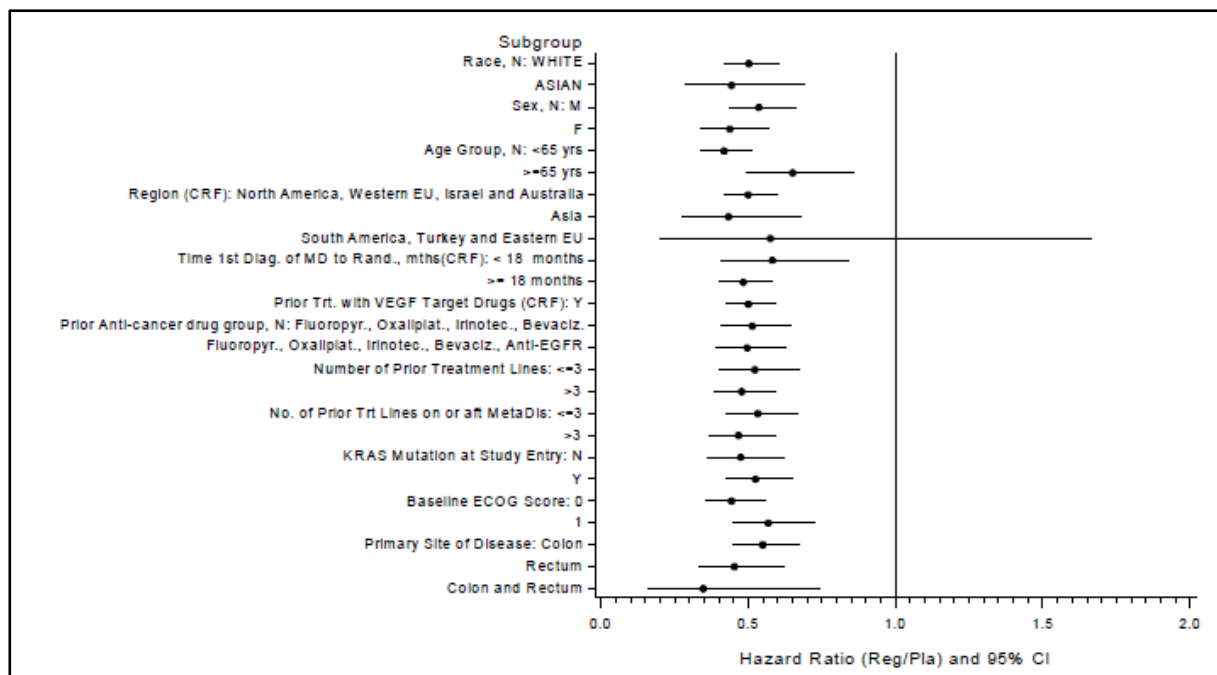
>3	459	393	0,478	(0,387; 0,592)	70	53
Anzahl der Therapielinien bei oder nach Diagnose der metastatischen Erkrankung						
						0,542853
≤3	395	358	0,533	(0,426; 0,667)	56	51
>3	365	313	0,467	(0,368; 0,594)	70	53
KRAS-Mutation						
						0,853983
Nein	299	261	0,475	(0,362; 0,623)	59	53
Ja	430	385	0,525	(0,425; 0,649)	58	51
Primärlokalisation der Erkrankung						
						0,426121
Kolon	495	442	0,550	(0,450; 0,671)	57	52
Rektum	220	190	0,454	(0,332; 0,620)	99	52
Kolon und Rektum	44	38	0,348	(0,163; 0,745)	107	53

a – Der p-Wert des Interaktionstests vergleicht die beiden in der jeweiligen kategorie genannten Subgruppen. Dieser p-Wert ist zweiseitig.

b – Region 1: Nordamerika, Westeuropa, Israel und Australien. Region 2: Asien. Region 3: Südamerika, Türkei und Osteuropa.

FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper

(Quelle: CORRECT 14387 Studienbericht: Tabelle 9-17, Tabelle 14.2/11 und 14387_ZA4: Tabelle 16.1.9.2/2)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiate, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diag. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastasierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 Studienbericht: Grafik 16.1.9.2/10)

Abbildung 4-6: Forest-Plot der Subgruppenanalyse für PFS (ITT) in RCT mit dem zu bewertenden Arzneimittel

Die Subgruppenanalysen für OS und PFS zeigen einen konsistenten Trend zugunsten von Regorafenib + BSC gegenüber Placebo + BSC für die Subgruppen Geschlecht, Alter, Ethnie, ECOG bei Randomisierung, Zeit seit der Diagnose der metastasierten Erkrankung, Region, vorherige Anti-VEGF-Therapie, vorherige Therapie mit Fluoropyrimidinen, Oxaliplatin, Irinotecan, Bevacizumab (FOIB) bzw. Therapie mit Fluoropyrimidinen, Oxaliplatin, Irinotecan, Bevacizumab und einem Anti-EGFR-Antikörper (FOIBE), Anzahl der vorangegangenen Therapielinien sowie KRAS-Status. Für die Subgruppe „primäre Lokalisation der Erkrankung Kolon+Rektum“ favorisieren die Ergebnisse für das Gesamtüberleben Placebo + BSC (HR: 1,091, 95 %-KI: 0,441 bis 2,697). Dies basiert auf einer sehr geringen Ereignisrate. Die Ergebnisse für das progressionsfreie Überleben favorisieren wiederum Regorafenib + BSC mit einer HR von 0,348 (95 %-KI: 0,163 bis 0,745).

Die Ergebnisse der Interaktionstests in der Analyse von OS und PFS für die jeweiligen Subgruppen sind als explorativ anzusehen und die p-Werte sind nicht auf multiples Testen adjustiert. Somit sollten die p-Werte mit Vorsicht interpretiert werden und insbesondere ein $p < 0,05$ sollte nicht als statistisch signifikant angesehen werden. Vielmehr sollte ein solcher im Kontext der biologischen Plausibilität als interessantes Signal interpretiert werden. Wenn multiples Testen berücksichtigt wird, müssten die p-Werte $< 0,0045$ sein, damit eine 5 %ige Irrtumswahrscheinlichkeit pro Testung (basierend auf elf Subgruppenkategorien und je elf Tests für OS und PFS) gewahrt bleibt. Keiner der Subgruppen-Interaktionstests kann als signifikant angesehen werden, als interessantes Signal gilt lediglich die Analyse der Altersgruppen im PFS ($p = 0,014384$). Da es jedoch an einer biologischen Rationale für eine unterschiedliche Wirksamkeit von Regorafenib in diesen beiden Altersgruppen mangelt, handelt es sich wahrscheinlich um einen Zufallsbefund. Wenn die Signifikanzschwelle für multiples Testen angewendet würde, wäre auch dieser Wert nicht signifikant.

Zusätzliche Subgruppenanalysen zum Gesamtüberleben zum Zeitpunkt des Cross-overs (data cut-off 13. November 2011)

Da zwischen *data cut-off* für die zweite geplante Zwischenanalyse (21. Juli 2011) und der Auswertung bzw. Kommunikation der Ergebnisse an die Prüfzentren mehrere Monate vergingen, wurde mit den Zulassungsbehörden vereinbart, eine weitere deskriptive Analyse nur zum Gesamtüberleben zum Zeitpunkt unmittelbar vor Beginn des Cross-overs durchzuführen. Das Cut-off-Datum dafür war der 13. November 2011. Diese Analyse basierte

auf 566 der ursprünglich 582 geplanten Todesfälle (97 %). Die Ergebnisse sind in Tabelle 4-51 dargestellt. Eine graphische Darstellung in Form von Forest Plots erfolgt in Abbildung 4-7.

Tabelle 4-51: Subgruppenanalyse für Gesamtüberleben zum Zeitpunkt des Cross-overs (*data cut-off* 13.11.2011)

Studie								
CORRECT 14387	Subgruppen- variable	n	An- zahl Ereig- nisse	Hazard Ratio (Regorafenib/Placebo)		Median (Tage)		Interak- tionstest p-Wert ^a
				Schät- zung	95 %-KI	Reg + BSC	Pla + BSC	
	Geschlecht							0,815355
	Männlich	464	340	0,806	(0,643; 1,009)	197	163	
	Weiblich	296	226	0,762	(0,581; 1,001)	181	146	
	Alter							0,511530
	<65 Jahre	475	365	0,753	(0,608; 0,932)	197	150	
	≥65 Jahre	285	201	0,862	(0,640; 1,161)	189	169	
	Ethnie							0,311064
	Kaukasier	593	447	0,768	(0,633; 0,932)	190	149	
	Asiatisch	111	74	0,994	(0,599; 1,648)	200	181	
	ECOG zum Beginn							0,657008
	0	410	284	0,777	(0,610; 0,989)	251	196	
	1	350	282	0,745	(0,580; 0,956)	140	109	
	Zeit seit der Diagnose der metastatischen Erkrankung							0,796016
	<18 Monate	140	117	0,841	(0,574; 1,233)	141	124	
	≥18 Monate	620	449	0,780	(0,642; 0,947)	204	165	
	Region ^b							0,245459
	1	632	482	0,750	(0,622; 0,904)	184	150	
	2	104	70	1,014	(0,608; 1,691)	206	181	
	Vorherige antitumoröse Therapie							0,730566
	FOIB	375	278	0,821	(0,643; 1,048)	179	154	
	FOIBE	385	288	0,758	(0,593; 0,970)	208	150	
	Anzahl der							0,512542

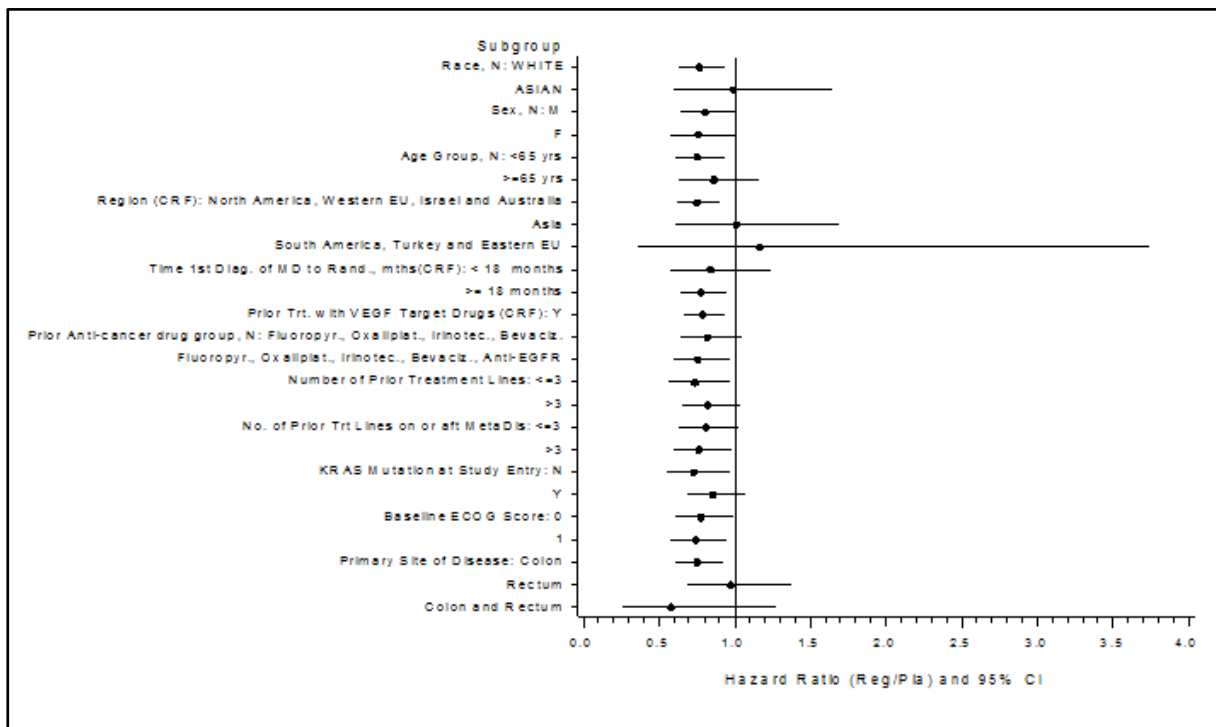
Therapielinien							
≤3	301	233	0,739	(0,563; 0,969)	192	127	
>3	459	333	0,823	(0,657; 1,032)	196	169	
Anzahl der Therapielinien bei oder nach Diagnose der metastatischen Erkrankung							0,755168
≤3	395	295	0,812	(0,639; 1,032)	204	151	
>3	365	271	0,766	(0,596; 0,984)	181	162	
KRAS Mutation							0,443817
Nein	299	224	0,732	(0,554; 0,966)	214	150	
Ja	430	319	0,858	(0,683; 1,077)	179	154	
Primärlokalisation der Erkrankung							0,201359
Kolon	495	387	0,751	(0,610; 0,925)	181	140	
Rektum	220	153	0,973	(0,688; 1,375)	211	218	

a – Der p-Wert des Interaktionstests vergleicht die beiden in der jeweiligen Kategorie genannten Subgruppen. Dieser p-Wert ist zweiseitig.

b - Region 1: Nordamerika, Westeuropa, Israel und Australien. Region 2: Asien. Region

FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper

(Quelle: CORRECT 14387 ZA3: Tabelle 16.1.9.2/1)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiate, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diang. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastasierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 ZA3: Grafik 16.1.9.2/1)

Abbildung 4-7: Subgruppenanalyse für Gesamtüberleben zum Zeitpunkt des Cross-overs (data cut-off 13.11.2011) – Darstellung der Ergebnisse im Forest-Plot

Auch diese zusätzliche Analyse zu einem späteren *data cut-off* bestätigt die konsistente Wirkung von Regorafenib über die untersuchten Subgruppen hinweg. Die durchgeführten Interaktionstests geben auch hier keinen Hinweis auf bestimmte Subgruppeneffekte, so dass die Gesamtpopulation für die Wirksamkeit von Regorafenib betrachtet werden kann.

Subgruppenanalysen zu unerwünschten Ereignissen in der CORRECT-Studie

Auf der Basis eines logistischen Regressionsmodells, das für die Studie 14387 angepasst wurde, wurden Subgruppenanalysen der Sicherheitspopulation hinsichtlich des Auftretens von UE, SUE und UE, die zur dauerhaften Beendigung der Studientherapie führten, durchgeführt (siehe Tabelle 4-52 und Abbildung 4-8, Tabelle 4-53 und Abbildung 4-9, Tabelle 4-54 und Abbildung 4-10).

Tabelle 4–52: Subgruppenanalyse für unerwünschte Ereignisse (UE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 - SAF-Population

Kovariable	Kategorie	Inzidenz von UE unter		Odds Ratio (95 %-KI)	p-Wert des Interaktionstests
		Regorafenib 160 mg	Placebo		
Ethnie, n (%)	Kaukasier	388/389 (99,7)	195/200 (97,5)	9,948 (1,154; 85,733)	ND ^a
	Asiatisch	74/74 (100,0)	31/34 (91,2)	212062 (0,000; 64E150)	
Geschlecht, n (%)	Männlich	305/307 (99,3)	149/152 (98,0)	3,071 (0,508; 18,575)	ND
	Weiblich	193/193 (100)	96/101 (95,0)	503455 (0,000; 16E195)	
Altersgruppe, n (%)	<65 Jahre	305/307 (99,3)	157/164 (95,7)	6,799 (1,396; 33,117)	ND
	≥65 Jahre	193/193 (100)	88/89 (98,9)	70731 (0,000; 52E156)	
Region, n (%)	Nordamerika, Westeuropa, Israel, Australien	415/417 (99,5)	208/211 (98,6)	2,993 (0,496; 18,051)	ND
	Asien	67/67 (100)	31/34 (91,2)	198549 (0,000; 18E153)	
Zeit seit Diagnose der metastasierten Erkrankung bis Randomisierung, n (%)	<18 Monate	88/89 (98,9)	49/49 (100,0)	0 (0,000; 57E123)	ND
	≥18 Monate	410/411 (99,8)	196/204 (96,1)	16,735 (2,079; 134,73)	
Vorherige Anti-VEGF Therapie, n (%)	Ja	498/500 (99,6)	245/253 (96,8)	8,131 (1,714; 38,576)	-
Vorherige Antitumorthherapie, n (%)	FOIB	237/238 (99,6)	130/134 (97,0)	7,292 (0,807; 65,926)	0,890331
	FOIBE	261/262 (99,6)	115/119 (96,6)	9,078 (1,004; 82,116)	
Anzahl vorheriger Therapielinien	≤3	201/202 (99,5)	94/98 (95,9)	8,553 (0,943; 77,574)	0,958068
	>3	297/298 (99,7)	151/155 (97,4)	7,868 (0,872; 71,007)	
Anzahl vorheriger Therapielinien seit Diagnose des metastasierten Stadiums, n (%)	≤3	254/256 (99,2)	130/135 (96,3)	4,885 (0,935; 25,520)	ND
	>3	244/244 (100,0)	115/118 (97,5)	189390 (0,000; 13E151)	
KRAS-Mutation bei Studieneintritt, n (%)	Nein	202/203 (99,5)	89/92 (96,7)	6,809 (0,699; 66,362)	0,869932

	Ja	269/270 (99,6)	152/157 (96,8)	8,848 (1,024; 76,432)	
ECOG Status bei Studieneintritt, n (%)	0	263/263 (100,0)	137/144 (95,1)	488436 (0,000; 94E166)	ND
	1	235/237 (99,2)	108/109 (99,1)	1,088 (0,098; 12,128)	
Lokalisation des Primärtumors, n (%)	Kolon	317/318 (99,7)	165/171 (96,5)	11,525 (1,376; 96,512)	0,571058
	Rektum	150/151 (99,3)	66/68 (97,1)	4,545 (0,405; 51,008)	

Die Odds Ratio wurde berechnet in Bezug auf Regorafenib gegenüber Placebo.

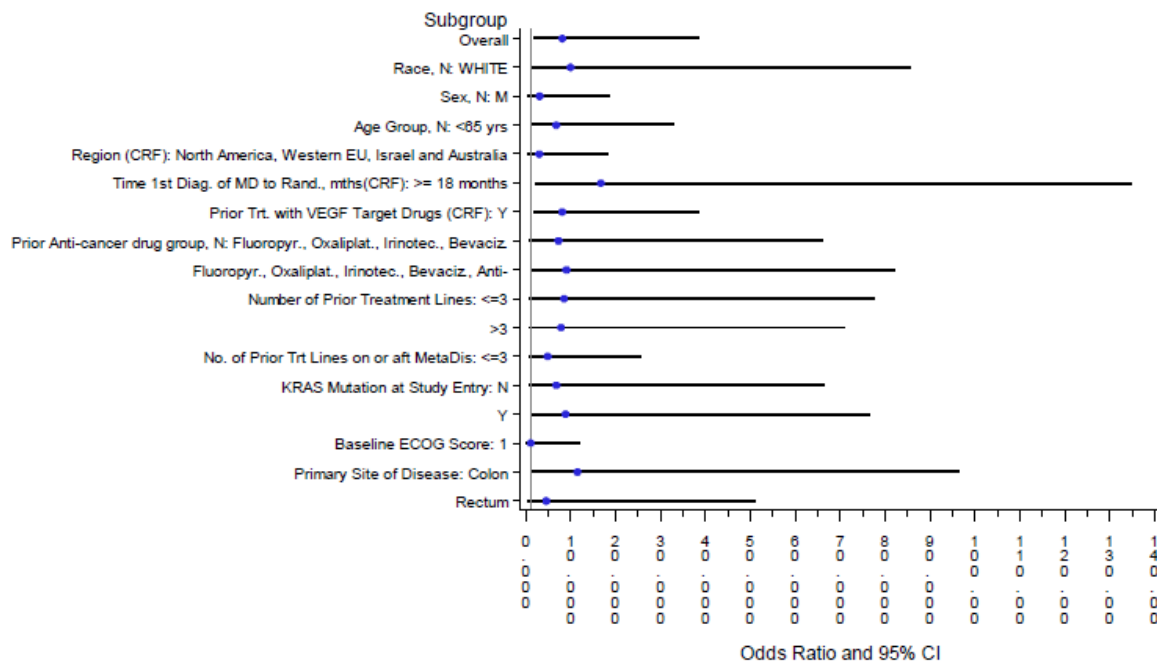
Der p-Wert des Interaktionstests wurde anhand eines logistischen Regressionsmodells, das die Therapiegruppe, die jeweilige Subgruppe und die Behandlung nach der Studieninteraktion als Kovariate, stratifiziert nach der Studie, beinhaltet, berechnet.

Wenn ein Minimum von 5 Ereignissen sowohl in der Regorafenib- als auch in der Placebo-Gruppe beobachtet wurde, wurde die Odds Ratio berechnet.

a – ND – Es wurde kein p-Wert des Interaktionstests berechnet, da in mindestens einer der Kategorien der Subgruppe kein Ereignis beobachtet wurde oder ein Ereignis trat in 100 % der Fälle in einer Subgruppe auf.

FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper

(Quelle: 14387_ZA5: Tabelle 16.1.9.2/3)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiatische, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diag. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastatisierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 ZA5: Grafik 16.1.9.2/2)

Abbildung 4-8: Subgruppenanalyse für unerwünschte Ereignisse (UE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF-Population, Forest-Plot-Darstellung mit logarithmischer Darstellung der x-Achse

Tabelle 4–53: Subgruppenanalyse für schwerwiegende unerwünschte Ereignisse (SUE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 - SAF-Population

Kovariable	Kategorie	Inzidenz von SUE unter		Odds Ratio (95 %-KI)	p-Wert des Interaktionstests
		Regorafenib 160 mg	Placebo		
Ethnie, n (%)	Kaukasier	176/389 (45,2)	80/200 (40,0)	1,239 (0,877; 1,752)	0,467014
	Asiatisch	24/74 (32,4)	12/34 (35,3)	0,880 (0,374; 2,070)	
Geschlecht, n (%)	Männlich	122/307 (39,7)	59/152 (38,8)	1,039 (0,698; 1,548)	0,272379
	Weiblich	97/193 (50,3)	41/101 (40,6)	1,479 (0,908; 2,407)	
Altersgruppe, n (%)	< 65 Jahre	137/307 (44,6)	64/164 (39,0)	1,259 (0,856; 1,853)	0,653715
	≥ 65 Jahre	82/193 (42,5)	36/89 (40,4)	1,087 (0,653; 1,812)	
Region, n (%)	Nordamerika, Westeuropa, Israel, Australien	196/417 (47,0)	87/211 (41,2)	1,264 (0,904; 1,767)	0,249061
	Asien	19/67 (28,4)	12/34 (35,3)	0,726 (0,301; 1,752)	
Zeit seit Diagnose der metastasierten Erkrankung bis Randomisierung, n (%)	< 18 Monate	43/89 (48,3)	24/49 (49,0)	0,974 (0,485; 1,956)	0,514399
	≥ 18 Monate	176/411 (42,8)	76/204 (37,3)	1,261 (0,894; 1,780)	
Vorherige Anti- VEGF Therapie, n (%)	Ja	219/500 (43,8)	100/253 (39,5)	1,192 (0,877; 1,622)	-
Vorherige Antitumorthérapie, n (%)	FOIB	96/238 (40,3)	54/134 (40,3)	1,002 (0,651; 1,542)	0,283350
	FOIBE	123/262 (46,9)	46/119 (38,7)	1,404 (0,903; 2,184)	
Anzahl vorheriger Therapielinien, n (%)	≤ 3	85/202 (42,1)	42/98 (42,9)	0,969 (0,595; 1,578)	0,283877
	> 3	134/298 (45,0)	58/155 (37,4)	1,366 (0,918, 2,033)	
Anzahl vorheriger Therapielinien seit Diagnose des metastasierten Stadiums, n (%)	≤ 3	110/256 (43,0)	56/135 (41,5)	1,063 (0,697; 1,621)	0,436867
	> 3	109/244 (44,7)	44/118 (37,3)	1,358 (0,865; 2,131)	
KRAS-Mutation bei	Nein	84/203	36/92	1,098	0,834546

Studieneintritt, n (%)		(41,4)	(39,1)	(0,664; 1,816)	
	Ja	119/270 (44,1)	63/157 (40,1)	1,176 (0,789; 1,753)	
ECOG Status bei Studieneintritt, n (%)	0	97/263 (36,9)	43/144 (29,9)	1,372 (0,887; 2,123)	0,277009
	1	122/237 (51,5)	57/109 (52,3)	0,968 (0,615; 1,524)	
Lokalisation des Primärtumors, n (%)	Kolon	143/318 (45,0)	73/171 (42,7)	1,097 (0,754; 1,596)	0,406339
	Rektum	65/151 (43,0)	23/68 (33,8)	1,479 (0,814; 2,686)	

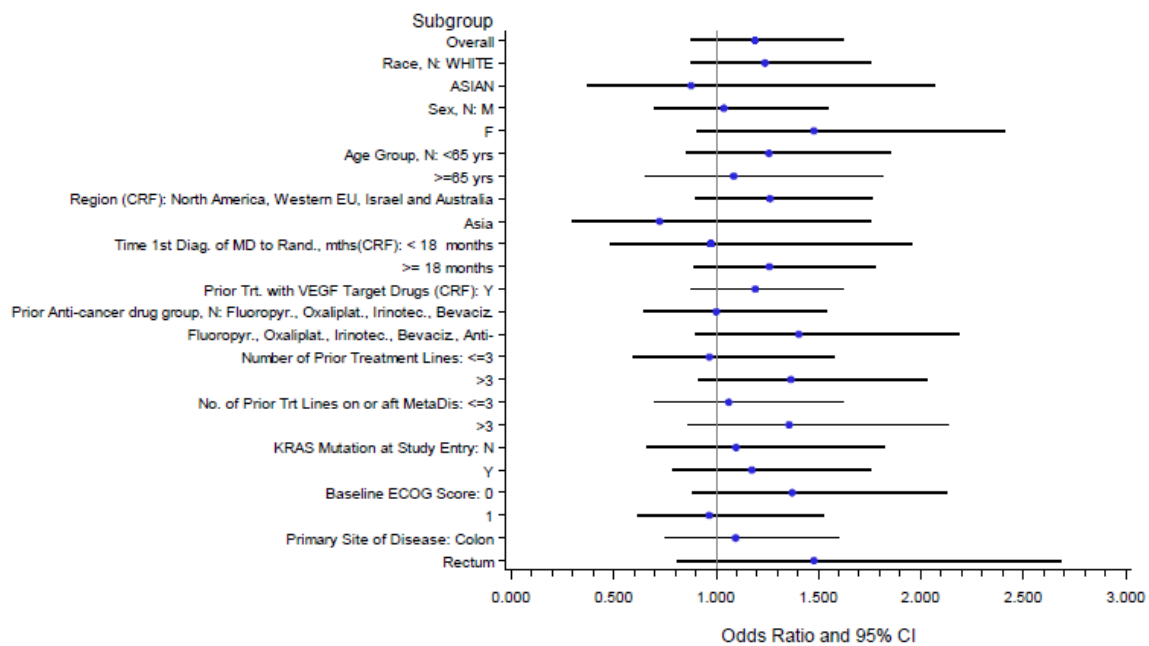
Die Odds Ratio wurde berechnet in Bezug auf Regorafenib gegenüber Placebo.

Der p-Wert des Interaktionstests wurde anhand eines logistischen Regressionsmodells, das die Therapiegruppe, die jeweilige Subgruppe und die Behandlung nach der Studieninteraktion als Kovariate, stratifiziert nach der Studie, beinhaltete, berechnet.

Wenn ein Minimum von 5 Ereignissen sowohl in der Regorafenib- als auch in der Placebo-Gruppe beobachtet wurde, wurde die Odds Ratio berechnet.

FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper

(Quelle: 14387_ZA5: Tabelle 16.1.9.2/4)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiater, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diag. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastasierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 ZA5: Grafik 16.1.9.2/3)

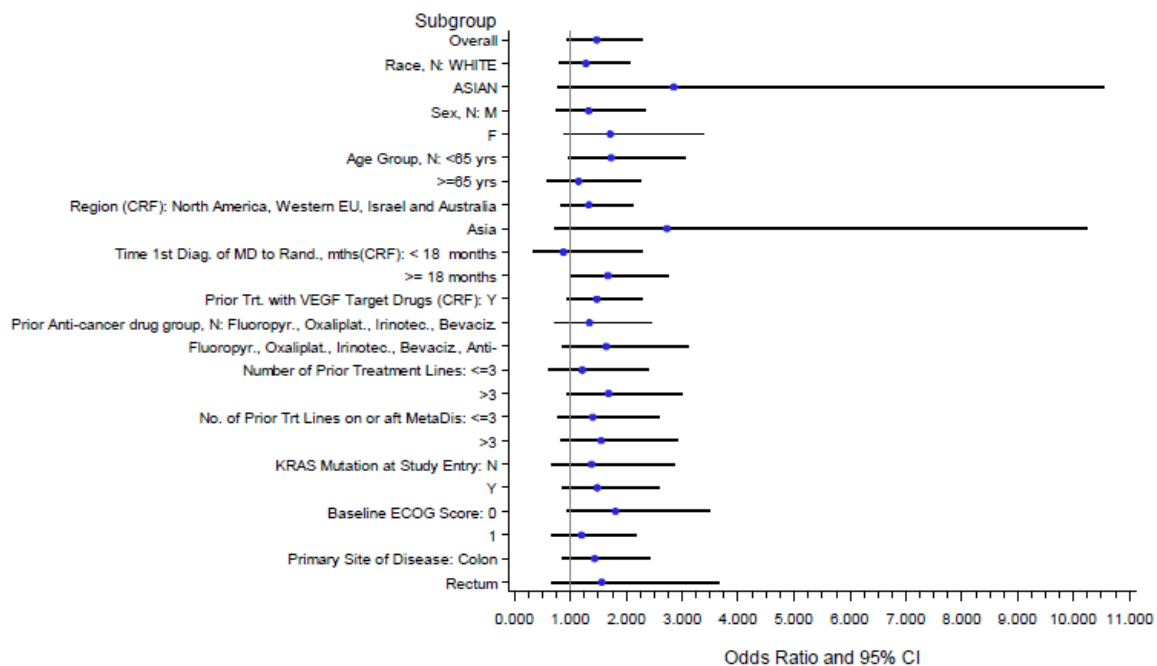
Abbildung 4-9: Subgruppenanalyse für schwerwiegende unerwünschte Ereignisse (SUE) – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF-Population, Forest-Plot-Darstellung mit logarithmischer Darstellung der x-Achse

Tabelle 4–54: Subgruppenanalyse für unerwünschte Ereignisse (UE), die zur dauerhaften beendigung der Studienmedikation führten – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF Population

Kovariable	Kategorie	Inzidenz von SUE unter		Odds Ratio (95 %-KI)	p-Wert des Interaktionstests
		Regorafenib 160 mg n (%)	Placebo n (%)		
Ethnie	Kaukasier	67/389 (17,2)	28/200 (14,0)	1,278 (0,792; 2,062)	0,258990
	Asiatisch	16/74 (21,6)	3/34 (8,8)	2,850 (0,771; 10,542)	
Geschlecht	Männlich	49/307 (16,0)	19/152 (12,5)	1,329 (0,752; 2,350)	0,574328
	Weiblich	39/193 (20,2)	13/101 (12,9)	1,714 (0,868; 3,384)	
Altersgruppe	< 65 Jahre	54/307 (17,6)	18/164 (11,0)	1,731 (0,978; 3,064)	0,362147
	≥ 65 Jahre	34/193 (17,6)	14/89 (15,7)	1,146 (0,580; 2,262)	
Region	Nordamerika, Westeuropa, Israel, Australien	73/417 (17,5)	29/211 (13,7)	1,332 (0,835; 2,123)	0,316152
	Asien	14/67 (20,9)	3/34 (8,8)	2,729 (0,727; 10,252)	
Zeit seit Diagnose der metastasierten Erkrankung bis Randomisierung	< 18 Monate	13/89 (14,6)	8/49 (16,3)	0,877 (0,336; 2,288)	0,239888
	≥ 18 Monate	75/411 (18,2)	24/204 (11,8)	1,674 (1,021; 2,744)	
Vorherige Anti-VEGF Therapie	Ja	88/500 (17,6)	32/253 (12,6)	1,475 (0,953; 2,282)	-
Vorherige Antitumorthherapie	FOIB	41/238 (17,2)	18/134 (13,4)	1,341 (0,736; 2,443)	0,653811
	FOIBE	47/262 (17,9)	14/119 (11,8)	1,640 (0,864; 3,112)	
Anzahl vorheriger Therapielinien	≤3	34/202 (16,8)	14/98 (14,3)	1,214 (0,618; 2,385)	0,468891
	>3	54/298 (18,1)	18/155 (11,6)	1,684 (0,950; 2,987)	
Anzahl vorheriger Therapielinien seit Diagnose des metastasierten	≤3	43/256 (16,8)	17/135 (12,6)	1,401 (0,765; 2,565)	0,817805
	>3	45/244 (18,4)	15/118 (12,7)	1,553 (0,826; 2,918)	

Stadiums					
KRAS-Mutation bei Studieneintritt	Nein	32/203 (15,8)	11/92 (12,0)	1,378 (0,661; 2,871)	0,878160
	Ja	48/270 (17,8)	20/157 (12,7)	1,481 (0,843; 2,601)	
ECOG Status bei Studieneintritt	0	40/263 (15,2)	13/144 (9,0)	1,808 (0,932; 3,504)	0,367286
	1	48/237 (20,3)	19/109 (17,4)	1,203 (0,668; 2,165)	
Lokalisation des Primärtumors	Kolon	58/318 (18,2)	23/171 (13,5)	1,435 (0,850; 2,422)	0,870240
	Rektum	26/151 (17,2)	8/68 (11,8)	1,560 (0,667; 3,650)	

Die Odds Ratio wurde berechnet in Bezug auf Regorafenib gegenüber Placebo.
 Der p-Wert des Interaktionstests wurde anhand eines logistischen Regressionsmodells, das die Therapiegruppe, die jeweilige Subgruppe und die Behandlung nach der Studieninteraktion als Kovariate, stratifiziert nach der Studie, beinhaltet, berechnet. Wenn ein Minimum von fünf Ereignissen sowohl in der Regorafenib- als auch in der Placebo-Gruppe beobachtet wurde, wurde die Odds Ratio berechnet.
 FOIB: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab; FOIBE: Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab, EGFR-Antikörper
 (Quelle: 14387_ZA5: Tabelle 16.1.9.2/5)



Übersetzung: Subgroup = Subgruppe, Race = Rasse, White = Weißer, Asian = Asiate, Sex = Geschlecht, M = Männlich, F = Weiblich, Age group = Altersgruppe, Region = Region, North America = Nordamerika, Western Europe = Westeuropa, Australia = Australien, Asia = Asien, South America = Südamerika, Turkey = Türkei, Eastern Europe = Osteuropa, Time 1st Diag. of MD to Rand. months (CRF) = Zeit bis zur ersten Diagnose der metastasierten Erkrankung in Monaten (laut CRF), Prior Trt. with VEGF Target Drugs (CRF) = Vorherige anti-VEGF-Therapie, Prior Anti-cancer drug group = Vorherige antitumoröse Therapie, Number of prior treatment lines = Anzahl der Therapielinien, No. of prior Trt Lines on or aft MetaDis = Anzahl der Therapielinien bei der oder nach Diagnose der metastatischen Erkrankung, KRAS Mutation at study entry = KRAS-Mutation, Baseline ECOG score = ECOG zum Beginn, Primary site of disease = Primärlokation der Erkrankung (Quelle: CORRECT 14387 ZA5: Grafik 16.1.9.2/4)

Abbildung 4-10: Subgruppenanalyse für unerwünschte Ereignisse, die zur dauerhaften Beendigung der Studienmedikation führten – Ergebnisse eines logistischen Regressionsmodells angepasst für Studie 14387 – SAF-Population, Forest-Plot-Darstellung mit logarithmischer Darstellung der x-Achse

Weder bei den aufgetretenen UE, bei den SUE noch bei den UE, die zu einer dauerhaften Beendigung der Studienmedikation führten, ergaben die durchgeführten Interaktionsteste einen Hinweis dafür, dass es signifikant mehr Ereignisse im Regorafenib- oder Placebo-Arm gab.

Biomarkeranalysen in der CORRECT-Studie

In der CORRECT-Studie wurden drei genetische (KRAS, PIK3CA und BRAF) und 15 nicht-genetische Biomarker (ANG-2, IL-6, IL-8, PIGF, VEGFR-1, TIE-1, VEGF-A, VEGF-C, VEGF-D, VEGFA- 121, BMP-7, VWF, M-CSF, SDF-1 und TIMP-2) analysiert. Da sowohl Ergebnisse präklinischer Studien als auch die von Phase-I- und Phase-II-Studien keine starken Hinweise auf einen prädiktiven Biomarker in Bezug auf die Wirkung von Regorafenib lieferten, war die Zulassungsstudie CORRECT nicht spezifisch dafür gewertet, Biomarkerfragen zu adressieren. Die Biomarkeranalysen in dieser Studie wurden retrospektiv durchgeführt, und werden als explorativ und Hypothesen-generierend angesehen. Somit wurden in den Analysen die p-Werte nicht in Bezug auf multiples Testen korrigiert, um nicht eine potentiell wichtige Korrelation zu übersehen. Daher werden p-Werte $<0,05$ formal nicht als statistisch signifikant angesehen, sondern sollten als potentiell interessantes Signal betrachtet werden.

Genetische Biomarkeranalysen

In der CORRECT-Studie wurden viele Patienten vor dem Einschluss in die Studie anhand archivierten Tumormaterials auf das Vorliegen der KRAS-Mutation getestet. Diese präexistenten „historischen“ KRAS-Daten wurden für die Subgruppenanalysen, so wie im Studienbericht beschrieben, verwendet. Neue genetische Biomarkeranalysen, die durchgeführt wurden, waren die Bestimmung des Mutationsstatus von drei Protoonkogenen, die häufig mit einem kolorektalen Karzinom assoziiert sind: KRAS, PI3KCA und BRAF. Der Mutationsstatus jedes dieser drei Gene wurde mittels eines hochsensitiven Assays (BEAMing) (48) in der DNA, isoliert sowohl aus einer Plasmaprobe bei Studieneinschluss als auch aus archiviertem Tumormaterial, bestimmt.

Wie im Studienbericht aufgeführt, waren vorhandene/historische Daten zum KRAS-Mutationsstatus von 96 % aller randomisierten Patienten verfügbar, von denen 59 % die KRAS-Mutation und 41 % den KRAS-Wildtyp aufwiesen. Korrelationsanalysen zeigen einen starken Trend eines Überlebensvorteils von Regorafenib im Vergleich zu Placebo sowohl in der Subgruppe der KRAS-Wildtyp-Patienten (Regorafenib/Placebo HR: 0,653 (95 %-KI: 0,476 bis 0,895)) als auch in der Subgruppe der KRAS-mutierten Patienten (Regorafenib/Placebo HR: 0,867 (95 %-KI: 0,670 bis 1,123)). Der Trend zu einem kleineren Überlebensvorteil in der Subgruppe der KRAS-mutierten Patienten kann zumindest teilweise

auf einen höheren Anteil nachfolgender antitumoröser Therapien in der KRAS-mutierten im Vergleich zur KRAS-Wildtyp-Subgruppe erklärt werden (33,1 % vs. 24,9 %). Auch ist hier zu beachten, dass diese Studie gepowert war, um einen Überlebensvorteil in der gesamten Studienpopulation zu entdecken, nicht in Subpopulationen mit bestenfalls der Hälfte der gesamten Studienpopulation. Der Interaktionstest zwischen beiden Gruppen in Bezug auf das Gesamtüberleben war ebenfalls nicht signifikant ($p=0,236344$). In einer exploratorischen *ad hoc* Analyse des Gesamtüberlebens bei nur mit Placebo behandelten Patienten, war der KRAS-Status kein prognostischer Faktor für das Gesamtüberleben (49).

Ähnliche Subgruppenanalysen für PFS anhand des historischen KRAS-Status zeigten ebenfalls einen starken Trend zugunsten der mit Regorafenib behandelten Patienten sowohl mit KRAS-Wildtyp als auch mit KRAS-Mutation (KRAS-WT: HR 0,475, 95 %-KI: 0,362; 0,623; KRAS-Mutation: HR 0,525, 95 %-KI: 0,425; 0,649). Der Interaktionstest zwischen beiden Gruppen in Bezug auf das PFS war ebenfalls nicht signifikant ($p=0,853983$). In einer exploratorischen *ad hoc* Analyse des progressionsfreien Überlebens bei nur mit Placebo behandelten Patienten, war der KRAS-Status ebenfalls kein prognostischer Faktor für PFS (49).

Zusätzlich zur Analyse der historischen KRAS-Daten wurde der KRAS-Mutationsstatus nochmals neu bestimmt mittels BEAMing-Technologie in DNA, die aus frisch gewonnenem Plasma zu Studieneinschluss isoliert wurde. Der Mutationsstatus aus DNA zu Beginn der Therapie könnte ein genaueres und „Real Time“-Abbild des KRAS-Mutationsstatus liefern als die Daten aus archiviertem Tumormaterial, das Monate bis Jahre vor Studienbeginn gewonnen wurde (50).

Der KRAS-Mutationsstatus aus Plasma-DNA, gewonnen zu Studienbeginn, konnte von 66 % der randomisierten Patienten bestimmt werden, von denen 69 % die KRAS-Mutation und 31 % den KRAS-Wildtyp aufwiesen. Die höhere Frequenz der KRAS-Mutation in der DNA isoliert aus frisch gewonnenem Plasma (69 %) im Vergleich zu archiviertem Tumormaterial (59 %) kann zumindest teilweise damit erklärt werden, dass die Tumore einem Selektionsdruck durch die Exposition gegenüber Anti-EGFR-Antikörpern bei KRAS-Wildtyp-Patienten und der nachfolgenden Akquirierung der KRAS-Mutation ausgesetzt waren (51, 52). Korrelationsanalysen zeigen einen Trend eines Überlebensvorteils von Regorafenib im Vergleich zu Placebo sowohl in der Subgruppe der KRAS-Wildtyp-Patienten (Regorafenib/Placebo HR: 0,67 (95 %-KI: 0,41 bis 1,08)) als auch in der Subgruppe der KRAS-mutierten Patienten (Regorafenib/Placebo HR: 0,81 (95 %-KI: 0,61 bis 1,09)). Während diese Ergebnisse ähnlich denen der historischen KRAS-Daten sind, war das Ergebnis des Interaktionstestes der *de novo* KRAS-Subgruppen mit einem p-Wert von 0,561 nicht signifikant, was darauf hindeutet, dass die klinische Wirksamkeit von Regorafenib (vs. Placebo) nicht vom KRAS-Mutationsstatus abhängig ist. Ebenso waren die p-Werte der Interaktionstests bezüglich des KRAS-Mutationsstatus, bestimmt mittels BEAMing-Technologie aus dem archivierten Tumormaterial, nicht signifikant.

Der PI3KCA-Mutationsstatus wurde ebenfalls in der DNA, die zu Studienbeginn aus einer Plasmprobe isoliert wurde, bestimmt. Der PI3KCA-Mutationsstatus aus Plasma-DNA konnte

von 66 % der randomisierten Patienten bestimmt werden, von denen 17 % die PI3KCA-Mutation und 83 % den PI3KCA-Wildtyp aufwiesen. Korrelationsanalysen zeigen einen Trend eines Überlebensvorteils von Regorafenib im Vergleich zu Placebo sowohl in der Subgruppe der PI3KCA-Wildtyp-Patienten (Regorafenib/Placebo HR: 0,75 (95 %-KI: 0,57 bis 0,99)) als auch in der Subgruppe der PI3KCA-mutierten Patienten (Regorafenib/Placebo HR: 0,84 (95 %-KI: 0,47 bis 1,50)). Der p-Wert des Interaktionstestes zwischen diesen beiden Subgruppen betrug 0,723, was darauf hinweist, dass es bei der klinischen Wirksamkeit von Regorafenib (vs. Placebo) keinen signifikanten Unterschied in Bezug auf den PI3KCA-Mutationsstatus gibt. Ebenso waren die p-Werte der Interaktionstests bezüglich des PI3KCA-Mutationsstatus, bestimmt mittels BEAMing-Technologie aus dem archivierten Tumormaterial, nicht signifikant.

Da KRAS- und PI3KCA-Mutationen im selben Tumor koexistieren können, wurden Subgruppenanalysen auch basierend auf verschiedenen Kombinationen des KRAS- und PI3KCA-Mutationsstatus, bestimmt mittels BEAMing-Technologie von DNA aus frischem Plasma, durchgeführt. In der Subgruppe der KRAS- und PI3KCA-mutierten Patienten betrug die Regorafenib/Placebo HR 0,71 (95 %-KI: 0,37 bis 1,35). In der KRAS-mutierten und PI3KCA-Wildtyp-Subgruppe betrug die Regorafenib/Placebo HR 0,84 (95 %-KI: 0,61 bis 1,16). In der KRAS- und PI3KCA-Wildtyp-Subgruppe betrug die Regorafenib/Placebo-HR 0,57 (95 %-KI: 0,34 bis 0,96). In der KRAS-Wildtyp und PI3KCA-Mutations-Subgruppe waren zu wenig Patienten, um eine aussagekräftige Korrelationsanalyse durchführen zu können.

Der BRAF-Mutationsstatus aus Plasma-DNA konnte von 66 % der randomisierten Patienten bestimmt werden, von denen 3,4 % eine BRAF-Mutation und 96,6 % den BRAF-Wildtyp aufwiesen. Aufgrund der geringen Anzahl BRAF-mutierter Patienten wurden keine Korrelationsanalysen in Bezug auf den BRAF-Mutationsstatus durchgeführt.

Nicht-genetische Biomarkeranalysen

Die nicht-genetischen Biomarkeranalysen in der CORRECT-Studie umfassten die quantitative Bestimmung von 15 verschiedenen Plasmaproteinen zu Studienbeginn; viele von ihnen sind mit einer Angiogenese assoziiert (ANG-2, TIE-1, VEGF-A, VEGFA-121, VEGF-C, VEGF-D, PIGF, VEGFR1, IL-6, IL-8), aber auch solche, die eine bekannte oder hypothetische Rolle in der Pathogenese des kolorektalen Karzinoms spielen (SDF-1, BMP-7, M-CSF, TIMP-2 und von-Willebrand-Faktor (VWF)). Eine Reihe dieser Proteine werden entweder direkt von Regorafenib inhibiert oder interagieren mit Proteinen oder Signalübertragungswegen, die von Regorafenib gehemmt werden (VEGFR1, VEGF-A, VEGFA-121, VEGF-C, VEGF-D, PIGF, TIE-1) (53).

Korrelationsanalysen des Gesamtüberlebens von Subgruppen mit hohen oder niedrigen medianen Plasmakonzentrationen der untersuchten Proteine ergaben, dass keines der untersuchten Proteine prädiktiv für die klinische Wirksamkeit von Regorafenib (vs. Placebo) war, d. h. keiner der p-Werte in Interaktionstests erreichte einen Wert von $p < 0,05$. Wenn jedoch die Gruppeneinteilung in „hoch“ und „niedrig“ anhand eines Cut-off-Wertes aus

einem „Best-Fit-Modell“ neu definiert wurde, erschienen die Ausgangskonzentrationen von TIE-1 für die klinische Aktivität von Regorafenib prädiktiv zu sein, obwohl Regorafenib einen günstigen Trend sowohl in der Subgruppe mit einem hohen als auch einem niedrigen Ausgangswert zeigte. Hohe TIE-1-Konzentrationen korrelierten mit einem grösseren Regorafenib-Vorteil als niedrige Konzentrationen dieses Proteins (Regorafenib/Placebo HR für OS bei Patienten mit niedrigem TIE-1: 0,87 (95 %-KI: 0,64 bis 1,20); Regorafenib/Placebo-HR für OS bei Patienten mit hohem TIE-1: 0,56 (95 %-KI: 0,41 bis 0,77); Interaktionstest p-Wert: 0,035). Zusätzlich erschienen eine Reihe von Proteinen prognostisch für das Überleben in der Placebogruppe zu sein (z. B. TIE-1, ANG-2, IL-6 und IL-8).

Korrelationsanalysen des PFS im Vergleich von Subgruppen mit hohen oder niedrigen Proteinkonzentrationen basierend auf den Medianwerten ergaben, dass die Konzentrationen von VWF prädiktiv für die klinische Aktivität von Regorafenib erschienen, obwohl Regorafenib einen positiven Trend sowohl in der Subgruppe mit hohen als auch niedrigen Konzentrationen dieses Proteins zeigte. Niedrige VWF-Konzentrationen korrelierten mit einem grösseren Regorafenibvorteil als hohe VWF-Konzentrationen (Regorafenib/Placebo-HR für PFS bei Patienten mit niedrigem VWF: 0,39 (95 %-KI: 0,30 bis 0,51); bei Patienten mit hohem VWF: 0,60 (95 %-KI: 0,46 bis 0,78); Interaktionstest p-Wert: 0,020).

Bemerkenswerterweise korrelierte bei der Analyse des PFS die TIE-1-Konzentration nicht mit dem Regorafenibvorteil, und umgekehrt korrelierte in der OS-Analyse die VWF-Konzentration nicht mit der Regorafenibaktivität. Die Ursache dieser Diskrepanz zwischen OS und PFS bleibt unklar, eröffnet jedoch die Erklärungsmöglichkeit, dass diese Ergebnisse zufällige Falsch-Positive sind und keine robusten Prädiktoren für eine Regorafenib-Wirkung. Zusätzlich wurden multiple Schwellenwerte für die Eingruppierung in „niedrig“ bzw. „hoch“ für jedes der untersuchten Proteine getestet, was die Wahrscheinlichkeit für falsch-positive Befunde weiter erhöht.

Zusammenfassung der Biomarkeranalysen

In *de novo* durchgeführten genetischen Analysen wurde der Mutationsstatus von KRAS, PI3KCA und BRAF analysiert. Korrelationsanalysen zeigten einen Trend zugunsten von Regorafenib (vs. Placebo) in allen Biomarker-Subgruppen anhand des KRAS- und PI3KCA-Mutationsstatus. Für eine aussagekräftige Subgruppenanalyse anhand des BRAF-Mutationsstatus waren zu wenig Patienten vorhanden. Während sowohl aus den *de novo* als auch aus den aus archiviertem Tumormaterial durchgeführten KRAS-Mutationsanalysen hervorging, dass KRAS-Wildtyp-Patienten einen grösseren Vorteil von der Regorafenibtherapie erfuhren als die KRAS-mutierten Patienten, ergaben Interaktionstests zwischen diesen Subgruppen keinen statistisch signifikanten Unterschied. Diese Daten deuten darauf hin, dass weder der KRAS- noch der PI3KCA-Mutationsstatus nützliche Biomarker wären, um Patienten, die wahrscheinlich von einer Regorafenib-Therapie profitieren würden, zu selektieren.

Nicht-genetische Biomarkeranalysen haben Plasmaproteine identifiziert, die für einen Vorteil von Regorafenib in Bezug auf OS (TIE-1) oder PFS (VWF) prädiktiv sein könnten, obwohl Regorafenib einen positiven Trend sowohl bei niedrigen als auch bei hohen Konzentrationen dieser Proteine aufwies. Somit weisen diese Daten darauf hin, dass TIE-1 und VWF keine nützlichen Biomarker sein würden, um Patienten, die wahrscheinlich von einer Regorafenibtherapie profitieren würden, zu selektieren.

In die Pathogenese und Krankheitsprogression des kolorektalen Karzinoms sind multiple Signalübertragungswege involviert. Diese umfassen z. B. die Ligand/Rezeptor-Signalwege EGF/EGFR, VEGF/VEGFR, PDGF/PDGFR, FGF/FGFR-Weg, sowie die intrazellulären Signalwege RAS-RAF-MEK-ERK und PI3K-PTEN-AKT-mTOR. Es wurde bspw. gezeigt, dass mehrere kompensatorische Signalwege während einer Therapie mit Bevacizumab + FOLFIRI aktiviert werden. Aufgrund der Komplexität des spezifischen Therapiesettings in der CORRECT-Studie sowie des Multikinasehemmungsprofils von Regorafenib wurde nicht erwartet, dass ein einziger Biomarker für das Ansprechen auf Regorafenib prädiktiv sein würde. Vielmehr dürfte die positive Wirkung von Regorafenib gerade durch die Multikinasehemmung in dieser fortgeschrittenen Patientenpopulation bedingt sein. Nicht überraschend ist auch, dass für andere Multikinasehemmer (z. B. Sorafenib, Sunitinib) bisher ebenfalls kein prädiktiver Biomarker identifiziert werden konnte, der robust und reproduzierbar in der klinischen Praxis einsetzbar wäre (49, 53).

Zusammenfassend kann konstatiert werden, dass, obwohl 18 Biomarker (drei genetische, 15 Plasmaproteine), die mechanistisch bei der Regorafenibwirkung und/oder einer Angiogenese und/oder der Tumorphathogenese eine Rolle spielen, priorisiert und in der CORRECT-Studie analysiert wurden, kein spezifischer Biomarker identifiziert werden konnte, der für die Patientenselektion nutzbar wäre, da ein positiver Trend für alle Biomarkersubgruppen in Bezug auf die Regorafenibwirkung gefunden wurde.

4.3.1.3.9 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Der folgende Abschnitt fasst die Ergebnisse der CORRECT-Studie in Bezug auf den primären Endpunkt Gesamtüberleben aus zwei separaten Analysen für die zwei Behandlungsarme Regorafenib + BSC und Placebo + BSC zusammen.

Der 21. Juli 2011 war Cut-off-Datum für die zweite geplante Zwischenanalyse, deren Ergebnisse dazu führten, dass die Studie am 13. November 2011 vorzeitig entblindet und den mit Placebo + BSC behandelten Patienten angeboten wurde, mit Regorafenib + BSC

weiterbehandelt zu werden (Cross-over). Diese Daten stellen auch gleichzeitig die Endanalyse dar, die für die Zulassung des Präparates diente. Zum Zeitpunkt vor dem tatsächlich durchgeführten Cross-over wurde auch das Cut-off-Datum für eine zusätzliche Analyse der Daten zum Gesamtüberleben gelegt, da dieser Zeitpunkt die maximale Dauer der verblindeten Studienbehandlung darstellt und noch keine Verwässerungseffekte durch ein Cross-over vorhanden waren.

Basis der Analyse mit *data cut-off* 21. Juli 2011 waren 432 Todesfälle: 275 in der Regorafenib + BSC-Gruppe (von 505 Patienten, = 54,5 %) und 157 in der Placebo + BSC-Gruppe 157 (von 255 Patienten, = 61,6 %). Die zusätzliche Analyse mit *data cut-off* 13. November 2011 hatte 566 Todesfälle zur Grundlage: 369 in der Regorafenib + BSC-Gruppe (von 505 Patienten, = 73,1 %) und 197 in der Placebo + BSC-Gruppe (von 255 Patienten, = 77,3 %).

Die Behandlungscompliance war in beiden Behandlungsgruppen hoch. Patienten in der Regorafenib + BSC-Gruppe erhielten 78,9 % der geplanten Dosis; Patienten in der Placebo + BSC-Gruppe 90,2 %. Die durchschnittliche (\pm STD) tägliche Dosis betrug 147,2 (\pm 18,6) mg in der Regorafenib + BSC-Gruppe und 159,3 (\pm 4,9) mg in der Placebo + BSC-Gruppe.

Für den primären Endpunkt **Gesamtüberleben (OS)** zeigte sich zum Zeitpunkt der zweiten Zwischenanalyse (=Endanalyse) für den Behandlungsarm Regorafenib + BSC eine signifikante Verlängerung ($p=0,005178$) der medianen Gesamtüberlebenszeit im Vergleich zum Behandlungsarm Placebo + BSC: Die geschätzte HR betrug 0,774 (95 %-KI: 0,636 bis 0,942) und somit eine Risikoreduktion um 22,6 % gegenüber Placebo + BSC oder eine Verlängerung der Überlebenszeit gegenüber Placebo + BSC um 29,2 %. Das vordefinierte alpha (Rate falsch-positiver Ergebnisse) für diese Analyse betrug 0,009279. Die mediane Gesamtüberlebenszeit betrug 196 Tage (95 %-KI: 178 bis 222 Tage) in der Regorafenib + BSC-Gruppe und 151 Tage (95 %-KI: 134 bis 177 Tage) in der Placebo + BSC-Gruppe.

Die zusätzlich durchgeführte Analyse zum Zeitpunkt unmittelbar vor einem Cross-over bestätigte den Vorteil von Regorafenib in Bezug auf das Gesamtüberleben. Die geschätzte HR betrug 0,790 (95 %-KI: 0,0664 bis 0,0939) und somit eine Risikoreduktion um 21 % gegenüber Placebo + BSC oder eine Verlängerung der Überlebenszeit gegenüber Placebo + BSC um 27 %. Die mediane Gesamtüberlebenszeit in dieser zusätzlichen Analyse betrug 194 Tage (95 %-KI: 177 bis 214 Tage) in der Regorafenib + BSC-Gruppe und 152 Tage (95 %-KI: 134 bis 178 Tage) in der Placebo + BSC-Gruppe.

Sekundäre Wirksamkeitsvariable war u. a. das **progressionsfreie Überleben** Die Analyse des PFS basierte auf der Beurteilung des Prüfarztes gemäß RECIST-Kriterien Version 1.1, wobei die erste Beurteilung nach der Randomisierung ca. 56 Tage nach der Randomisierung stattfand. Konsistent mit den Ergebnissen für das Gesamtüberleben zeigte sich im PFS ein statistisch signifikanter Vorteil zugunsten der Patienten, die in die Gruppe Regorafenib + BSC randomisiert wurden ($p<0,000001$). Die geschätzte Hazard Ratio (Risiko der Progression oder

des Todes vor einer Progression mit Regorafenib + BSC vs. Placebo + BSC) betrug 0,494 (95 %-KI: 0,419 bis 0,582), was einer Reduktion des Risikos um 50,6 % für die Behandlung mit Regorafenib + BSC gegenüber Placebo + BSC entspricht. Insgesamt hatten 671 von 760 Patienten (88,3 %) ein Ereignis (Krankheitsprogression oder Tod): 430 von 505 Patienten (85,1 %) unter Regorafenib + BSC und 241 von 255 Patienten (94,5 %) unter Placebo + BSC. Das mediane PFS war signifikant länger ($p < 0,000001$) für Patienten in der Regorafenib + BSC-Gruppe (59 Tage, 95 %-KI: 57 bis 65 Tage) als für Patienten in der Placebo + BSC-Gruppe (52 Tage, 95 %-KI: 51 bis 53 Tage).

Sensitivitätsanalysen ergaben konsistente und unterstützende Ergebnisse für das Gesamtüberleben (stratifiziert nach Randomisierung auf der Grundlage von IVRS sowie unstratifiziert) sowie das progressionsfreie Überleben (stratifiziert nach Randomisierung auf der Grundlage von IVRS, unstratifiziert, sowie Analysen, die alle Tumorbeurteilungen in der Follow-up-Periode berücksichtigten, und Analysen, die das Datum der Initiierung einer neuen Behandlung in der Nachbeobachtungsperiode als Ereignisdatum benutzten).

Vorteile für Regorafenib + BSC gegenüber Placebo + BSC wurden auch in den **Subgruppenanalysen** für OS und PFS gezeigt (inklusive in den für die Therapie des kolorektalen Karzinoms als wichtig angesehenen Subgruppen nach dem KRAS-Status und der Vortherapie mit FOIB/FOIBE).

Der beobachtete Vorteil im Gesamt- und progressionsfreien Überleben ist signifikant und klinisch bedeutsam in einer Patientenpopulation, bei der alle verfügbaren Therapien versagt haben.

Weitere sekundäre Wirksamkeitsvariablen waren die Tumoransprechrage, die Krankheitskontrollrate und die Dauer der Tumorstabilisierung.

Die **Objektive Tumoransprechrage** (CR + PR) in der ITT-Population betrug 1,0 % in der Regorafenib + BSC-Gruppe und 0,4 % in der Placebo + BSC-Gruppe, mit einer Intergruppendifferenz von -0,60 % (95 %-KI: -1,74 % bis 0,53 %), was statistisch nicht signifikant war (einseitiges $p = 0,188432$; präspezifiziertes α 0,025). Komplette Remissionen wurden nicht beobachtet, sodass die Ansprechrage auf bestätigten partiellen Remissionen basierte (fünf in der Regorafenib + BSC-Gruppe, eine in der Placebo + BSC-Gruppe).

Die **Krankheitskontrollrate** (CR + PR + SD) war statistisch signifikant höher in der Regorafenib + BSC-Gruppe (41,0 %; 95 %-KI: 36,7 % bis 45,4 %; einseitiges $p < 0,000001$) im Vergleich zur Placebo + BSC-Gruppe (14,9 %; 95 %-KI: 10,8 % bis 19,9 %). Eine Krankheitsstabilisierung wurde für die Berechnung der Krankheitskontrollrate nur dann herangezogen, wenn die Beurteilung mindestens sechs Wochen nach der Randomisierung erfolgte.

Tertiäre Wirksamkeitsvariablen waren die Dauer des Ansprechens und der Tumorstabilisierung sowie die gesundheitsbezogene Lebensqualität anhand des Instruments EORTC QLQ-C30 und EQ-5D.

Die mediane **Dauer der Tumorstabilisierung** betrug 60 Tage (95 %-KI: 57 bis 70) in der Regorafenib + BSC-Gruppe und 52 Tage (95 % KI: 51 bis 53) in der Placebo + BSC-Gruppe. Der geringe Unterschied ist hauptsächlich durch die primär progredienten Patienten in der Regorafenib + BSC-Gruppe (250 (=49,5 %) bzw. der Placebo + BSC-Gruppe) (204 (=80,0 %) bedingt. Die HR von 0,471 weist jedoch auch hier auf den Nutzen dieses Parameters hin.

Die mittleren Veränderungen vom Ausgangswert im globalen Gesundheitsstatus des EORTC QLQ-C30 und den fünf funktionalen Dimensionen zeigte eine Verschlechterung der **gesundheitsbezogenen Lebensqualität** der Patienten in ähnlicher Größenordnung in der Regorafenib + BSC-Gruppe und der Placebo + BSC-Gruppe (Ausnahme ist das Submodul „Soziale Funktionalität, in dem schlechtere Scorewerte bei den mit Regorafenib-behandelten Patienten registriert wurden). Bei der Zeit-adjustierten Berechnung der Fläche unter der Kurve bei diesem HRQoL-Instrument war der Unterschied nicht signifikant (Differenz: -1.19 (95 % KI: -3,13; 0,75). Somit ergab sich kein Unterschied in der Lebensqualität zwischen den beiden Gruppen.

Die Gesamtinzidenz **unerwünschter Ereignisse** in den beiden Behandlungsgruppen war in der Größenordnung vergleichbar, wobei die Mehrzahl der Patienten in beiden Gruppen jeweils mindestens ein UE hatte (99,6 % in der Regorafenib + BSC-Gruppe bzw. 96,8 % in der Placebo + BSC-Gruppe). Die in den absoluten Inzidenzraten vergleichbare Häufigkeit ist relativ gesehen jedoch signifikant unterschiedlich. Jedoch wird bei der Betrachtung des Relativen Risikos bei oberen Werten des KI >0,90 nach der Methodik des IQWiG trotz formaler statistischer Signifikanz dies als klinisch nicht relevant gewertet.

Insgesamt waren Unterschiede in UE hauptsächlich bedingt durch Unterschiede in den Ereignissen vom Grad 1 bis 3, wobei UE vom Grad 1 und 2 signifikant häufiger in der Placebo + BSC-Gruppe und UE vom Grad 3 signifikant häufiger in der Regorafenib + BSC-Gruppe auftraten. Die Unterschiede in den Graden 4 und 5 waren zwischen den beiden Behandlungsgruppen klein, statistisch nicht signifikant und klinisch bedeutungslos. Ebenso gab es keine statistisch signifikanten Unterschiede bei den schwerwiegenden UE und bei den UE, die zu einer dauerhaften Beendigung der Studienmedikation führten.

Die Rate an dauerhafter Beendigung der Therapie aufgrund von UE und unabhängig von einer Krankheitsprogression war niedrig und zwischen den Behandlungsgruppen nicht signifikant unterschiedlich (17,6 % vs. 12,6 %). Von diesen Patienten beendeten 41 (8,2%) im Regorafenib + BSC-Arm und 3 (1,2%) im Placebo + BSC-Arm dauerhaft die Studientherapie, da das jeweilige UE als in ursächlichem Zusammenhang mit der Studienmedikation stehend angesehen wurde. Dies deutet darauf hin, dass die meisten Toxizitäten durch Dosismodifikationen kontrolliert werden konnten. Die Inzidenz an Ereignissen vom Grad 3, 4 und 5, die zu einer dauerhaften Therapiebeendigung führten, war in beiden Gruppen ähnlich.

Da Placebo + BSC per se nicht auf den Tumor wirkt und somit keine Nebenwirkungen einer antitumorösen Therapie auftreten, kann die höhere Inzidenzrate von UE, die durch die Studienmedikation bedingt angesehen wurden, den Nutzen von Regorafenib nicht infrage stellen, da für einen patientenrelevanten und zweckmäßigen Vergleich der Verträglichkeit Substanzen herangezogen werden sollten, die im Anwendungsgebiet das gleiche Therapieziel, nämlich die Verlängerung des Gesamtüberlebens, haben. Da solche Therapien derzeit nicht existieren, wird das Nutzen-Risiko-Verhältnis von Regorafenib als positiv gewertet. Darüber hinaus lag kein signifikanter Unterschied in der Lebensqualität der beiden Gruppen vor.

Zusammenfassend zeigen die Daten zur Sicherheit, dass Regorafenib gut toleriert wurde. Die unerwünschten Ereignisse, die mit Regorafenib assoziiert waren, waren in den meisten Fällen handhabbar. Die Rate dauerhafter Therapiebeendigungen aufgrund unerwünschter Ereignisse war gering, da die meisten durch Dosismodifikationen beherrscht werden konnten. Dies wird u.a. deutlich an der mittleren Tagesdosis von 147,2 mg, die nur wenig unter der normalen Dosierung von 160 mg pro Tag lag. Das Sicherheitsprofil ist insgesamt ähnlich dem, das aus früheren Phase-I- und Phase-II-Studien bekannt ist, und entspricht zum Teil auch dem Nebenwirkungsprofil der Population von Patienten mit einem metastasierten kolorektalen Karzinom. Basierend auf diesen Daten wird eingeschätzt, dass potentielle Vorteile potentielle Sicherheitsrisiken bei Patienten, die das Arzneimittel erhalten, deutlich überwiegen, insbesondere, da die Lebensqualität der Patienten der Regorafenib + BSC-Gruppe in der globalen Analyse der Zeit-adjustierten Fläche unter der Kurve nicht unterschiedlich ist.

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 4-55: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend.					

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 4-56: 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>
Nicht zutreffend.					

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 4-57: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
	Nicht zutreffend.

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 4-58: 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
	Nicht zutreffend					

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 4-59: 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)
	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 4.3.1.3.8.

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 4-60: 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		
Nicht zutreffend.						

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-61: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
Nicht zutreffend.	

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 4-62: 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
Nicht zutreffend.				

1) 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 4.3.1.3.8.

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

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche

- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren 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 4-63: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
1	Nicht zutreffend.

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 4.3.1.3.8.

Nicht zutreffend.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

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

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Die CORRECT-Studie entspricht nach der in der Verfo-GBA dargestellten Evidenzhierarchie der Evidenzstufe Ib (2). Die Studienqualität ist als sehr gut anzusehen, da der primäre Endpunkt (Gesamtüberleben) subjektiven Einflüssen entzogen ist. Das Gesamtüberleben (Overall Survival) ist zudem ein patientenrelevanter Endpunkt und i. d. R. ein herausragendes Therapieziel (27). Verzerrungspotentiale liegen für die Studie nicht vor, da grundlegende Kriterien wie Verblindung, adäquate Randomisierungsmethoden und eine Sicherstellung der verdeckten Zuteilung der Patienten zu den Studienarmen erfüllt sind. Der primäre Endpunkt wurde vorab mit Zulassungsbehörden anhand des zu erreichenden und als klinisch relevant anzusehenden Unterschieds in Bezug auf die Hazard Ratio definiert und in der Fallzahlplanung adäquat berücksichtigt. Die statistischen Methoden entsprechen den Methoden einer Überlebenszeitanalyse und werden mit den entsprechenden Effektschätzern dargestellt (Hazard Ratio, Kaplan-Meier-Kurve). Die Ergebnisdarstellung gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet werden. Für alle Ergebnisschätzer sind die zugehörigen Konfidenzintervalle dargestellt, auch für nicht signifikante Ergebnisse. Auch wurden Sensitivitätsanalysen für die Endpunkte Gesamtüberleben und progressionsfreies Überleben durchgeführt (erweitertes Cox-Modell), um die Robustheit der Ergebnisse zu überprüfen. Alle Subgruppenanalysen wurden vorab definiert und gemäß den festgelegten Termini ausgewertet und umfassend dargestellt. Eine zusätzliche Analyse des

Gesamtüberlebens unmittelbar vor dem Beginn des Cross-overs wurde auf Anfrage der Zulassungsbehörden durchgeführt und ist entsprechend gekennzeichnet und dargestellt.

Die Überprüfung der genannten Kriterien bei den Operationalisierungen, Analysen und der Ergebnisdarstellung der untersuchten patientenrelevanten Endpunkte in Bezug auf die Mortalität (OS), Mortalität/Morbidität (PFS), und Morbidität (Objektive Tumorsprechrate, Krankheitskontrollrate und Dauer der Tumorstabilisierung) zeigen ebenfalls ein niedriges Verzerrungspotential.

Der Endpunkt Verträglichkeit wurde vorab definiert und als sekundärer Endpunkt nicht in der Fallzahlplanung berücksichtigt. Der Endpunkt wurde deskriptiv ausgewertet wie zuvor im statistischen Analyseplan beschrieben. Die Ergebnisdarstellung gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet werden. Für alle Ergebnisschätzer sind die zugehörigen Konfidenzintervalle dargestellt, auch für nicht signifikante Ergebnisse. Subgruppenanalysen wurden für den Endpunkt Verträglichkeit ebenfalls durchgeführt.

Da das Studiendesign einer doppelblinden, randomisierten kontrollierten Studie entspricht, kein Verzerrungspotential besteht und mit dem Gesamtüberleben ein patientenrelevanter primärer Endpunkt abgebildet wird, kann von einer hohen Ergebnissicherheit ausgegangen werden. Mit einer umfassenden Recherche in Studienregistern als auch einer Sichtung der Studien des pharmazeutischen Unternehmers wurde gewährleistet, dass jedwede adäquate Evidenz in die Beurteilung eines Zusatznutzens einfließt und sichergestellt ist, dass keine abgebrochenen Studien vorliegen, die einen Einfluss auf die Evidenz haben könnten.

Auch das Vorliegen nur einer Phase-III-Studie für die Beurteilung eines Zusatznutzens von Regorafenib schränkt diese hohe Ergebnissicherheit nicht ein, da die besonderen Anforderungen an die Studie und deren Ergebnisse wie von der EMA (54) und vom IQWiG (26) für einen Beleg gestellt, erfüllt sind. Die interne Validität ist wie in Modul 4 dargestellt gegeben, die externe Validität wird durch das Studiendesign der CORRECT-Studie, insbesondere mit der Vergleichstherapie Best Supportive Care, und der ebenfalls recht breit gefassten Einschlusskriterien für die Studienpopulation gewährleistet und ermöglicht so die Übertragbarkeit der Ergebnisse in den klinischen Alltag.

Begründung für die Ergebnissicherheit „Beleg“

Die „Allgemeine Methoden Version 4.0“ des IQWiG fordern für die Aussage „Beleg“ mindestens zwei Studien mit mehrheitlich hoher Ergebnissicherheit und einem „gleichgerichteten“ Effekt (26). Im selben Abschnitt heißt es jedoch auch: „Soll aus lediglich einer Studie im Ausnahmefall ein Nutzenbeleg abgeleitet werden, so sind an eine solche Studie und ihre Ergebnisse besondere Anforderungen zu stellen“. Als Referenz wird hier das Papier des Committee for Proprietary Medicinal Products (CPMP) der European Medicines Agency (EMA) „Points to consider on application with 1. Meta-analyses; 2. One pivotal study“ (15). In diesem Papier sind Kriterien für Zulassungen auf der Basis nur einer Phase-III-Studie genannt, die im Folgenden aufgeführt und mit der Studie 14387 CORRECT hinsichtlich der Erfüllung dieser Kriterien verglichen werden.

Folgende Kriterien sollten vorliegen, wenn die Zulassung auf der Basis nur einer Phase III-Studie erfolgen soll.

- Interne Validität: Es sollte keinen Hinweis auf einen potentiellen Bias geben.
- Externe Validität: Die Studienpopulation sollte so sein, dass man von ihr auf die zu behandelnde Patientenpopulation extrapolieren kann.
- Klinische Relevanz: Die Größe des Therapieeffektes sollte so bedeutsam sein, dass sie klinisch relevant ist.
- Grad der statistischen Signifikanz: Die statistische Evidenz sollte grösser als das gewöhnliche Signifikanzniveau von $p < 0,05$ sein, zusammen mit einer präzisen Schätzung des Therapieeffektes, d. h. ein enges Konfidenzintervall. Das geforderte Signifikanzniveau ist abhängig von Faktoren wie der therapeutischen Indikation, dem primären Endpunkt, der Menge unterstützender Daten und ob die zusätzlichen Analysen, die eine Konsistenz zum Hauptergebnis zeigen, vorab geplant waren.
- Datenqualität.
- Interne Konsistenz: Ähnliche Effekte sind in den präspezifizierten Subgruppen gezeigt. Alle wichtigen Endpunkte zeigen ähnliche Befunde.
- Zentreffekte: Keines der Studienzentren sollte das Gesamtergebnis dominieren, weder in Bezug auf die Anzahl der Patienten noch auf die Größe des Behandlungseffektes.
- Die Plausibilität der getesteten Hypothese.

Die CORRECT-Studie erfüllt all diese Kriterien.

Folgende Kriterien sollten **nicht** vorliegen, wenn die Zulassung auf Basis nur einer Phase-III-Studie erfolgen soll.

- Fehlen einer pharmakologischen Rationale (unbekannter Wirkmechanismus)
- Neues pharmakologisches Prinzip
- Die Daten aus Phase I- und II-Studien sind limitiert oder wenig überzeugend.
- Therapiegebiet, auf dem es in der Vergangenheit negative Studien oder Studien gab, die vorherige anscheinend überzeugende Daten nicht bestätigen konnten
- Notwendigkeit, die Wirksamkeit und/oder Verträglichkeit bei verschiedenen Subpopulationen, mit verschiedenen Komedikationen oder anderen Interventionen, oder mit verschiedenen Vergleichstherapien nachzuweisen

- Andere Fragen, die in verschiedenen Phase-III-Studien untersucht werden müssten

Bei der CORRECT Studie liegen diese Kriterien nicht vor. Im Ergebnis dieser Analyse wird die Ergebnissicherheit „Beleg“ für den festgestellten Zusatznutzen reklamiert.

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.

Als Zusammenfassung der Ergebnisdarstellung im Modul 4.3.1 wird für die einzelnen Kategorien der patientenrelevanten Endpunkte in der Tabelle 4-64 und Tabelle 4-65 der jeweilige Zusatznutzen aus Sicht des pharmazeutischen Unternehmers dargestellt.

Tabelle 4-64: Regorafenib + BSC vs. Placebo + BSC - Ausmaß des Zusatznutzens auf Endpunktebene

		Effektschätzer (95 %-KI) Ereignisanteil Regorafenib + BSC vs. Placebo + BSC p-Wert (wenn vorhanden) Wahrscheinlichkeit	Ableitung des Ausmaßes
Mortalität	Gesamtmortalität (OS)	ITT (Zeitschnitt 21.07.2011) HR 0,77 (0,64; 0,94) Medianes Gesamtüberleben: 196 Tage vs. 151 Tage p=0,0052 Wahrscheinlichkeit: Beleg	Endpunktkategorie: Überlebenszeit Zusatznutzen Ausmaß : beträchtlich
	Gesamtmortalität (OS) [Zusätzliche Analyse (zum Cross-over- Zeitpunkt)]	ITT (Zeitschnitt 13.11.2011) HR 0,79 (0,66; 0,94) Medianes Gesamtüberleben: 194 Tage vs. 152 Tage p=0,0038 Wahrscheinlichkeit: Beleg	Endpunktkategorie: Überlebenszeit Zusatznutzen Ausmaß : beträchtlich
Mortalität/Morbidität	PFS	ITT (Zeitschnitt 21.07.2011) HR 0,49 (0,42; 0,58) Medianes PFS: 59 Tage vs. 52 Tage p<0,000001 Wahrscheinlichkeit: Beleg	Endpunktkategorie: Überlebenszeit Zusatznutzen Ausmaß : erheblich
Morbidität	Objektive Tumoransprechrage (ORR)	ITT Differenz: -0,60 (-1,74; 0,53) P=0,1884	Endpunktkategorie: Inzidenzen Zusatznutzen Ausmaß : Kein Zusatznutzen
	Krankheitskontrollrate (DCR)	ITT Differenz: -25,94 (-32,06; -19,82) DCR pro Gruppe: 41,0 % vs. 14,9 % p<0,000001 Wahrscheinlichkeit: Beleg	Endpunktkategorie: Inzidenzen Zusatznutzen Ausmaß : beträchtlich
	Dauer der Tumorstabilisierung	ITT HR 0,471 (0,398; 0,558) Mediane Dauer der Tumorstabilisierung: 60 Tage vs. 52 Tage	Endpunktkategorie: Inzidenzen Zusatznutzen Ausmaß: beträchtlich

		Wahrscheinlichkeit: Beleg	
Gesundheits- bezogene Lebens- qualität	EORTC QLQ-C30 Zeit-adjustierte Fläche unter der Kurve	ITT Differenz: -1,19 (-3,13; 0,75)	Endpunktkategorie: Inzidenzen Zusatznutzen Ausmaß: Kein Zusatznutzen
Verträglichkeit	Unerwünschte Ereignisse	Sicherheitspopulation Placebo vs. Regorafenib RR 0,97 (0,95; 0,99) 96,84 % vs. 99,60 %	Kein größerer Schaden belegt
	Schwerwiegende unerwünschte Ereignisse	Sicherheitspopulation Placebo vs. Regorafenib RR 0,90 (0,75; 1,08) 39,53 % vs. 43,80 %	Kein größerer Schaden belegt
	Unerwünschte Ereignisse CTCAE Grad 1	Sicherheitspopulation Regorafenib vs. Placebo RR 0,13 (0,07; 0,25) 2,40 % vs. 17,79 % Wahrscheinlichkeit: Beleg	Geringerer Schaden Ausmaß: beträchtlich
	Unerwünschte Ereignisse CTCAE Grad 2	Sicherheitspopulation Regorafenib vs. Placebo RR 0,64 (0,49; 0,83) 19,20 % vs. 30,04 %	Geringerer Schaden Ausmaß: gering
	Unerwünschte Ereignisse CTCAE Grad 3	Sicherheitspopulation Placebo vs. Regorafenib RR 0,47 (0,38; 0,59) 26,48 % vs. 56,00 % Wahrscheinlichkeit: Beleg	Größerer Schaden Ausmaß: erheblich
	Unerwünschte Ereignisse CTCAE Grad 4	Sicherheitspopulation Placebo vs. Regorafenib RR 0,92 (0,55; 1,53) 8,60 % vs. 7,91 %)	Kein größerer Schaden belegt
	Unerwünschte Ereignisse CTCAE Grad 5	Sicherheitspopulation Regorafenib vs. Placebo RR 0,92 (0,63; 1,33) 13,40 % vs. 14,62 %	Kein größerer Schaden belegt
	Unerwünschte Ereignisse, die zum Studienabbruch geführt haben	Sicherheitspopulation Placebo vs. Regorafenib RR 0,72 (0,49; 1,05) 12,65 % vs. 17,60 %	Kein größerer Schaden belegt

Tabelle 4-65: Begleitende Resultate zur Gesamtaussage zum Zusatznutzen

Zusatznutzen	Größerer Schaden
Beleg für einen erheblichen Zusatznutzen <ul style="list-style-type: none"> • PFS 	Beleg für einen größeren Schaden <ul style="list-style-type: none"> • UE Grad 3 (Ausmaß: erheblich)
Beleg für einen beträchtlichen Zusatznutzen <ul style="list-style-type: none"> • OS (beide Auswertzeitpunkte) • DCR • Dauer der Tumorstabilisierung 	
Beleg für einen geringeren Schaden <ul style="list-style-type: none"> • UE Grad 1 (Ausmaß: beträchtlich) • UE Grad 2 (Ausmaß: gering) 	

Der Zusatznutzen leitet sich aus den Endpunkten Gesamtmortalität, PFS, DCR und Dauer der Tumorstabilisierung ab, die allesamt einen Beleg für einen beträchtlichen (OS, DCR, Dauer der Tumorstabilisierung) oder erheblichen (PFS) Zusatznutzen darstellen sowie aus dem Beleg für einen geringeren Schaden bei den UE vom Grad 1 und Grad 2. Demgegenüber wurde für UE vom Grad 3 ein größerer Schaden mit Ausmaß erheblich gefunden. Für den Endpunkt gesundheitsbezogene Lebensqualität wurde kein Zusatznutzen festgestellt.

In der Gesamtschau ergibt sich damit für Regorafenib zur Behandlung von Patienten mit metastasiertem kolorektalem Karzinom für die Gesamtpopulation ein **Beleg für einen beträchtlichen Zusatznutzen**.

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 4-66: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit metastasiertem Kolorektalkarzinom (KRK), die zuvor mit verfügbaren Therapien behandelt wurden oder die für diese nicht geeignet sind. Diese Therapien umfassen Fluoropyrimidin-basierte Chemotherapie, eine Anti-VEGF-Therapie und eine Anti-EGFR-Therapie.	Beträchtlich Wahrscheinlichkeit: Beleg

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 zutreffend.

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 4-67: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Nicht zutreffend.	

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

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

Nicht zutreffend.

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

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

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

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

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

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

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

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

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

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.

Nicht zutreffend.

4.6 Liste der eingeschlossenen Studien

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

- Studie 14387 (CORRECT) – *Clinical Study Report (CSR) No. A53306*, 19. März 2012 – Dokument M4A_14387_Studienbericht.pdf
- Studie 14387 (CORRECT) – Integrated Clinical Study Protocol – Dokument M4A_14387_Studienprotokoll.pdf
- Studie 14387 (CORRECT) – post hoc Analyse zu Relativen Risiken und Odds Ratios von Verträglichkeitsparametern – Dokument 14387_ZA1.pdf
- Studie 14387 (CORRECT) – Germany Reimbursement Statistical Analysis 3 – Dokument 14387_ZA2.pdf
- Studie 14387 (CORRECT) – OS update2 based on 13NOV2011 data cutoff – Dokument 14387_ZA3.pdf
- Studie 14387 (CORRECT) – Ad-hoc Germany Reimbursement Statistical Analysis – Dokumente 14387_ZA4.pdf
- Studie 14387 (CORRECT) – Germany Reimbursement Statistical Analysis 2 – Dokument 14387_ZA5.pdf
- Studie 14387 (CORRECT) – Descriptive OS update based on 13Nov2011 data cutoff – Dokument 14387_ZA6.pdf

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. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV Beratungsanforderung 2012-B-026. Berlin: 2012.
2. Gemeinsamer Bundesausschuss (GBA). 5. Kapitel der Verfahrensordnung. Bewertung des Nutzens von Arzneimitteln nach § 35a SGB V. . Available from: <http://www.g-ba.de>.
3. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. Epub 2010/03/25.
4. Zhuang SH, Xiu L, Elsayed YA. Overall survival: a gold standard in search of a surrogate: the value of progression-free survival and time to progression as end points of drug efficacy. *Cancer J*. 2009;15(5):395-400. Epub 2009/10/15.

5. Piedbois P, Miller Crosswell J. Surrogate endpoints for overall survival in advanced colorectal cancer: a clinician's perspective. *Stat Methods Med Res.* 2008;17(5):519-27. Epub 2008/02/21.
6. Hurvitz S, Mathias S, Doan J, Crosby R. The value of progression-free survival from the patient perspective: An online survey of women with metastatic breast cancer in the United States. *Cancer Research.* 2011;71(24 (Suppl. 3)):P5-17-08.
7. Siena S, Peeters M, Van Cutsem E, Humblet Y, Conte P, Bajetta E, et al. Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. *British journal of cancer.* 2007;97(11):1469-74. Epub 2007/11/28.
8. Thong MS, Mols F, Coebergh JW, Roukema JA, van de Poll-Franse LV. The impact of disease progression on perceived health status and quality of life of long-term cancer survivors. *Journal of cancer survivorship : research and practice.* 2009;3(3):164-73. Epub 2009/06/27.
9. European Medicines Agency. Methodological Considerations for using Progression-Free Survival (PFS) as primary endpoint in confirmatory trials for Registration. In: EMA, editor. London2008.
10. Food and Drug Administration. Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. In: FDA, editor. Rockville, MD2007.
11. Health Canada. Issues Analysis Summary: The use of Progression-Free Survival as the Efficacy Endpoint for Approval of Targeted and Chemotherapeutic Agents for Advanced Cancer. In: Canada H, editor. Ottawa2007.
12. Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. *Nature reviews Clinical oncology.* 2012;9(1):41-7. Epub 2011/10/20.
13. Matsuyama R, Reddy S, Smith TJ. Why do patients choose chemotherapy near the end of life? A review of the perspective of those facing death from cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2006;24(21):3490-6. Epub 2006/07/20.
14. Victorson D, Soni M, Cella D. Metaanalysis of the correlation between radiographic tumor response and patient-reported outcomes. *Cancer.* 2006;106(3):494-504. Epub 2005/12/15.
15. European Medicines Agency. Point to Consider on Application with 1. Meta-analyses; 2. One Pivotal Study. In: CPMP, editor. London2001.
16. Grothey A, Van Cutsem E, Sobrero A, Siena S, Falcone A, Ychou M, et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet.* 2013;381(9863):303-12. Epub 2012/11/28.
17. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. *Meta-Analysis Group in Cancer. Lancet.* 2000;356(9227):373-8. Epub 2000/09/06.
18. Johnson KR, Ringland C, Stokes BJ, Anthony DM, Freemantle N, Irs A, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. *Lancet Oncol.* 2006;7(9):741-6. Epub 2006/09/02.

19. Cousin S, Taieb S, Penel N. A paradigm shift in tumour response evaluation of targeted therapy: the assessment of novel drugs in exploratory clinical trials. *Curr Opin Oncol.* 2012;24(3):338-44. Epub 2012/03/16.
20. Goodwin R, Giaccone G, Calvert H, Lobbezoo M, Eisenhauer EA. Targeted agents: how to select the winners in preclinical and early clinical studies? *Eur J Cancer.* 2012;48(2):170-8. Epub 2011/11/19.
21. Nishino M, Jagannathan JP, Krajewski KM, O'Regan K, Hatabu H, Shapiro G, et al. Personalized tumor response assessment in the era of molecular medicine: cancer-specific and therapy-specific response criteria to complement pitfalls of RECIST. *AJR Am J Roentgenol.* 2012;198(4):737-45. Epub 2012/03/28.
22. Tunariu N, Kaye SB, Desouza NM. Functional imaging: what evidence is there for its utility in clinical trials of targeted therapies? *British journal of cancer.* 2012;106(4):619-28. Epub 2012/01/28.
23. Bayer Vital GmbH. Fachinformation Stivarga 40mg Filmtabletten. 2013;September 2013.
24. M4A_Informationsbeschaffung_Suchstrategien-Studienregister_Anhang4-B.pdf.
25. M4A_Informationsbeschaffung_Studienregister-ausgeschlossene-Studien_Anhang4-B.pdf.
26. IQWiG. Allgemeine Methoden, Version 4.0. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); 2011.
27. IQWiG. Aussagekraft von Surrogatendpunkten in der Onkologie. Version 1.0 ed2011.
28. Chirila C, Odom D, Devercelli G, Khan S, Sherif BN, Kaye JA, et al. Meta-analysis of the association between progression-free survival and overall survival in metastatic colorectal cancer. *Int J Colorectal Dis.* 2012;27(5):623-34. Epub 2011/11/15.
29. Giessen C, Laubender RP, Ankerst DP, Stintzing S, Modest DP, Mansmann U, et al. Progression-free survival as a surrogate endpoint for median overall survival in metastatic colorectal cancer: literature-based analysis from 50 randomized first-line trials. *Clin Cancer Res.* 2013;19(1):225-35. Epub 2012/11/15.
30. Sidhu R, Rong A, Dahlberg S. Evaluation of progression-free survival as a surrogate endpoint for survival in chemotherapy and targeted agent metastatic colorectal cancer trials. *Clin Cancer Res.* 2013;19(5):969-76. Epub 2013/01/11.
31. Breitbart W, Rosenfeld B, Pessin H, Kaim M, Funesti-Esch J, Galietta M, et al. Depression, hopelessness, and desire for hastened death in terminally ill patients with cancer. *JAMA.* 2000;284(22):2907-11. Epub 2001/01/09.
32. Catt S, Blanchard M, Addington-Hall J, Zis M, Blizard R, King M. Older adults' attitudes to death, palliative treatment and hospice care. *Palliat Med.* 2005;19(5):402-10. Epub 2005/08/23.
33. Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ, et al. Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public. *BMJ.* 1990;300(6737):1458-60. Epub 1990/06/02.
34. Institut für Angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH. Kolorektales Karzinom - Abschlussbericht. 2011 [02.05.2012]; Available from: http://www.sqg.de/sqg/downloads/Entwicklung/Abschlussberichte/KRK/Abschlussbericht_Kolorektales_Karzinom.pdf
35. Lee P, Weerasuriya DK, Lavori PW, Quon A, Hara W, Maxim PG, et al. Metabolic tumor burden predicts for disease progression and death in lung cancer. *Int J Radiat Oncol Biol Phys.* 2007;69(2):328-33. Epub 2007/09/18.

36. Schmoll HJ, Arnold D. Kolorektales Karzinom. In: Schmoll, Höffken, Possinger, editors. Kompendium Internistische Onkologie. 4 ed. Heidelberg: Springer-Verlag; 2006. p. 3787-937.
37. 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. Epub 2007/11/21.
38. Tang PA, Bentzen SM, Chen EX, Siu LL. Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2007;25(29):4562-8. Epub 2007/09/19.
39. Clinicaltrials.gov. Patients With Metastatic Colorectal Cancer Treated With Regorafenib or Placebo After Failure of Standard Therapy - NCT01103323. National Institutes of Health; 2013 [cited 2013]; Available from: <http://www.clinicaltrials.gov/ct2/show/NCT01103323?term=01103323&rank=1>.
40. Bayer Healthcare. Studienbericht Studie CORRECT 14387 (Clinical Study Report No. A53306, BAY 73-4506 / 14387). 2012.
41. Bayer Healthcare. Studienprotokoll Studie CORRECT 14387 (Integrated Clinical Study Protocol, BAY 73-4506 / 14387, version 4.0). 2011.
42. van Cutsem Eea. PHASE 3 CORRECT TRIAL OF REGORAFENIB IN METASTATIC COLORECTAL CANCER (MCRC): OVERALL SURVIVAL UPDATE. *Ann Oncol*. 2012;23(Suppl. 9):LBA 18.
43. Osoba D, Aaronson N, Zee B, Sprangers M, te Velde A. Modification of the EORTC QLQ-C30 (version 2.0) based on content validity and reliability testing in large samples of patients with cancer. The Study Group on Quality of Life of the EORTC and the Symptom Control and Quality of Life Committees of the NCI of Canada Clinical Trials Group. *Qual Life Res*. 1997;6(2):103-8. Epub 1997/03/01.
44. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 1998;16(1):139-44. Epub 1998/01/24.
45. Fayers P, Bottomley A. Quality of life research within the EORTC-the EORTC QLQ-C30. European Organisation for Research and Treatment of Cancer. *Eur J Cancer*. 2002;38 Suppl 4:S125-33. Epub 2002/02/23.
46. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res*. 1996;5(6):555-67. Epub 1996/12/01.
47. Deutschnoff G, Friedrich C, Thiem U, Voigtmann R, Pientka L. Lebensqualität in der Onkologie - Status Quo und Ausblick. *Der Onkologe*. 2005;2(11):164-72. Epub 2005/01/28.
48. Diehl F, Li M, He Y, Kinzler KW, Vogelstein B, Dressman D. BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. *Nat Methods*. 2006;3(7):551-9. Epub 2006/06/23.
49. Jeffers M vCE, Sobrero AF. Mutational analysis of biomarker samples from the CORRECT study: Correlating mutation status with clinical response to regorafenib. ASCO Gastrointestinal Cancer Conference 2013.

50. Higgins MJ, Jelovac D, Barnathan E, Blair B, Slater S, Powers P, et al. Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood. *Clin Cancer Res.* 2012;18(12):3462-9. Epub 2012/03/17.
51. Diaz LA, Jr., Williams RT, Wu J, Kinde I, Hecht JR, Berlin J, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature.* 2012;486(7404):537-40. Epub 2012/06/23.
52. Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D, et al. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature.* 2012;486(7404):532-6. Epub 2012/06/23.
53. Lenz HJ. Analysis of plasma protein biomarkers from the CORRECT phase III study of regorafenib for metastatic colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2013;31(15 (Suppl.)).
54. European Medicines Agency. Guideline on the Evaluation of Anticancer Medicinal Products in Man. . London2005.
55. Clinicaltrials.gov. Asian Subjects With Metastatic Colorectal Cancer Treated With Regorafenib or Placebo After Failure of Standard Therapy (CONCUR) - NCT01584830. National Institutes of Health; 2013 [cited 2013]; Available from: <http://www.clinicaltrials.gov/ct2/show/NCT01584830?term=nct01584830&rank=1>.

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

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.

Da eine Zulassung von Regorafenib weltweit noch nicht länger als zwölf Monate besteht und keine indirekten Vergleiche herangezogen werden, ist nach den Kriterien der Verfo-GBA eine bibliographische Literaturrecherche in Literaturdatenbanken nicht notwendig und wurde aus diesem Grund nicht durchgeführt.

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

Suchstrategie in den Studienregistern

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/home
Datum der Suche	13.08.2013
Suchstrategie	Stivarga OR Regorafenib OR BAY 73-4506 OR BAY73-4506 [Search]
Treffer	33

Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	13.08.2013
Suchstrategie	Stivarga OR Regorafenib OR BAY73-4506 [Search]
Treffer	9

Studienregister	Klinische Prüfungen - Pharmnet.Bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	13.08.2013
Suchstrategie	Stivarga [Textfelder] ODER Regorafenib [Textfelder] ODER BAY 73-4506 [Textfelder] ODER BAY73-4506 [Textfelder]
Treffer	7

Studienregister	International Clinical Trial Registry Platform
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	13.08.2013
Suchstrategie	Stivarga OR Regorafenib OR BAY 73-4506 OR BAY73-4506 [Search]
Treffer	89 Treffer für 35 Studien

Studien aus der Suche in den Studienregistern

Die Datenbankabfrage über clinicaltrials.gov, EU Clinical Trials Register, klinische Prüfungen Pharmnet.Bund und International Clinical Trials Registry Platform ergab 138 Treffer für 84 Studien. Von 84 Studien wurden 48 anhand ihrer NCT-Studienbezeichnung und ihrer Titel als Dubletten eingestuft. Die 36 identifizierten Studien und die verwendeten Kriterien für den Studienausschluss sind in Tabelle 4-68 dargestellt. Die Begründung für den Studienausschluss erfolgte anhand der prädefinierten Einschlusskriterien für Studien zur Bewertung des medizinischen Nutzens und Zusatznutzens von Regorafenib (vgl. Tabelle 4-7).

Tabelle 4-68 (Anhang): Abgeschlossene, laufende und abgebrochene Studien aus der Studienregisterrecherche und Kriterien für Studienausschluss

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
1	ACTRN12612000239864	A randomised phase II double-blind study to evaluate the effect of Regorafenib or placebo on progression-free survival and objective response rate in refractory advanced oesophago-gastric cancer (AOGC)	Nicht rekrutierend	ja (E1: Magenkrebs, E4: nicht rekrutierend)	http://www.anzctr.org.au/ACTRN12612000239864.aspx	ICTRP
2	EUCTR2013-000236-94-ES	Phase II study of Regorafenib as a single agent for first-line treatment of patients with metastatic colorectal cancer (MCRC) who are fragile and/or not candidates for polychemotherapy	Laufend	ja (E1: Patienten, die gebrechlich oder nicht für Polychemotherapie geeignet sind, E4: einarmig, laufend)	https://www.clinicaltrialsregister.eu/ctr-search/search?query=2013-000236-94+	ICTRP EU Clinical Trials Register
3	JPRN-UMIN000010888	Multidisciplinary team approach in management of Regorafenib toxicities for patients with metastatic colorectal cancer: Multi-institutional joint research in Shizuoka	Rekrutierend	ja (E4: rekrutierend, keine RCT)	http://www.umin.ac.jp/ctr/index.htm	ICTRP
4	NCT00664326	A Phase II Uncontrolled Study of BAY73-4506 in Previously Untreated Patients With Metastatic or Unresectable Renal Cell Cancer (RCC)	Aktiv, nicht rekrutierend	ja (E1: Nierenkrebs, E4: unkontrollierte Studie)	http://ClinicalTrials.gov/show/NCT00664326	Clinicaltrials.gov ICTRP
5	NCT00934882	Phase I, Open-label, Non-placebo Controlled Study to Determine the Safety, Pharmacokinetics, and	Abgeschlossen	ja (E2: Regorafenib in Kombination mit mFOLFOX)	http://ClinicalTrials.gov/show/NCT00934882	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		Pharmacodynamics of BAY73-4506 in Combination With mFOLFOX6 or FOLFIRI as First or Second Line Therapy in Patients With Metastatic Colorectal.		6, E4: Phase I, E5: Regorafenib in Kombination mit FOLFIRI)		
6	NCT00960258	Phase I, Uncontrolled, Open-label, Non-randomized Study to Investigate Pharmacokinetics and Safety of BAY73-4506 in Asian (Japanese) Patients With Advanced, Refractory Solid Tumors	Aktiv, nicht rekrutierend	ja (E1: Patienten mit fortgeschrittenen, refraktären Tumorerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT00960258	Clinicaltrials.gov ICTRP
7	NCT01002378	A Phase I, Randomized, Open Label, 3-Way Cross-Over Study to Determine the Effect of a High-Fat Breakfast, a Low-Fat Breakfast and Fasting State on the Pharmacokinetics of a Single Oral Dose of 160 mg Regorafenib (BAY73-4506) in Healthy Volunteers	Abgeschlossen	ja (E1: gesunde Freiwillige, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01002378	Clinicaltrials.gov ICTRP
8	NCT01003015	An Uncontrolled Open Label Multicenter Phase II Safety Study of BAY73-4506 in Patients With Hepatocellular Carcinoma (HCC)	Abgeschlossen	ja (E1: hepatozelluläres Karzinom, E4: unkontrollierte Studie)	http://ClinicalTrials.gov/show/NCT01003015	Clinicaltrials.gov ICTRP PharmNet.bund
9	NCT01068769	A Non-randomized, Open Label, Multi-center Phase II Study Evaluating the Efficacy and Safety of Regorafenib in Patients With Metastatic and/or	Aktiv, nicht rekrutierend	ja (E1: gastrointestinaler Stromatumor, E4: nicht randomisiert, aktiv)	http://ClinicalTrials.gov/show/NCT01068769	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		Unresectable Gastrointestinal Stromal Tumor (GIST), Resistant or Intolerant to at Least Imatinib and Sunitinib				
10	NCT01096030	Uncontrolled, Open-label, Non-randomized, Phase 1 Study to Investigate the Pharmacokinetics, Safety, Tolerability, and Efficacy of BAY73-4506 in Chinese Patients With Advanced, Refractory Solid Tumors	Abgeschlossen	ja (E1: Patienten mit fortgeschrittenen, refraktären Tumorerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01096030	Clinicaltrials.gov ICTRP
11	NCT01103323	A Randomized, Double-blind, Placebo-controlled Phase III Study of Regorafenib Plus BSC Versus Placebo Plus BSC in Patients With Metastatic Colorectal Cancer (CRC) Who Have Progressed After Standard Therapy	Aktiv, nicht rekrutierend	nein	http://ClinicalTrials.gov/show/NCT01103323	Clinicaltrials.gov PharmNet.b und EU Clinical Trials Register
12	NCT01117623	Open Label, Phase I Study to Determine the Safety, Tolerability, Maximum Tolerated Dose, Pharmacokinetics, and Biomarker Status of BAY73-4506 in Patients With Advanced Malignancies	Aktiv, nicht rekrutierend	ja (E1: Patienten mit fortgeschrittenen Tumorerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01117623	Clinicaltrials.gov ICTRP
13	NCT01187615	An Open-label, Multi-center, Non-randomized Phase Ib Study to Investigate the Safety, Efficacy, and Pharmacokinetics of BAY 73-4506	Abgebrochen	ja (E1: nicht-kleinzellige Lungenkarzinom, E4: Phase Ib)	http://ClinicalTrials.gov/show/NCT01187615	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		"Regorafenib", Administered in Combination With Pemetrexed and Cisplatin in Patients With Advanced Nonsquamous Non-Small Cell Lung Cancer				
14	NCT01189903	Clinical Evaluation - A Phase IIA Proof of Concept Study of Regorafenib (Bayer 73-4506) in Biopsy-amenable Asian Colorectal Cancer Patients	Rekrutierend	ja (E4: rekrutierend)	http://ClinicalTrials.gov/show/NCT01189903	Clinicaltrials.gov ICTRP
15	NCT01265979	Whole Body Diffusion-weighted MRI (WB-DWI) for Early Prediction and Evaluation of Therapy Response in Patients With Advanced Metastatic Gastrointestinal Stromal Tumors (GIST) Treated With Regorafenib.	Abgebrochen	ja (E1: gastrointestinaler Stromatumor, E4: keine RCT)	http://ClinicalTrials.gov/show/NCT01265979	Clinicaltrials.gov ICTRP
16	NCT01271712	A Randomized, Double-blind, Placebo-controlled Phase III Study of Regorafenib Plus Best Supportive Care Versus Placebo Plus Best Supportive Care for Subjects With Metastatic and/or Unresectable Gastrointestinal Stromal Tumors (GIST) Whose Disease Has Progressed Despite Prior Treatment With at Least Imatinib and Sunitinib	Aktiv, nicht rekrutierend	ja (E1: gastrointestinaler Stromatumor)	http://ClinicalTrials.gov/show/NCT01271712	Clinicaltrials.gov ICTRP PharmNet.bund

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
17	NCT01287598	A Phase I, Non-randomized Open-label Study to Evaluate the Effect of BAY73-4506 (Regorafenib) on Probe Substrates of CYP 2C9 (Warfarin), 2C19 (Omeprazole) and 3A4 (Midazolam) in a Cocktail Approach (Group A) and on a Probe Substrate of CYP 2C8 (Rosiglitazone, Group B) in Patients With Advanced Solid Tumors	Aktiv, nicht rekrutierend	ja (E1: Patienten mit fortgeschrittenen Tumorerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01287598	Clinicaltrials.gov ICTRP
18	NCT01289821	An Uncontrolled, Open-label, Phase II Study in Subjects With Metastatic Adenocarcinoma of the Colon or Rectum Who Are Receiving First Line Chemotherapy With mFOLFOX6 (Oxaliplatin/ Folinic Acid/5-fluorouracil [5-FU]) in Combination With Regorafenib	Aktiv, nicht rekrutierend	ja (E1: Patienten, die mFOLFOX6 in Kombination mit Regorafenib erhalten haben, E4: unkontrollierte Studie)	http://ClinicalTrials.gov/show/NCT01289821	Clinicaltrials.gov EU Clinical Trials Register ICTRP PharmNet.bund
19	NCT01298570	Multi-Center, Randomized, Placebo-Controlled Phase II Study of Regorafenib in Combination With FOLFIRI Versus Placebo With FOLFIRI as Second-Line Therapy in Patients With Metastatic Colorectal Cancer	Rekrutierend	ja (E2: Regorafenib in Kombination mit FOLFIRI, E4: rekrutierend, E5: Placebo in Kombination mit FOLFIRI)	http://ClinicalTrials.gov/show/NCT01298570	Clinicaltrials.gov ICTRP
20	NCT01318265	A Phase I, Non-Randomized, Open Label Study to Determine the Effect	Abgeschlossen	ja (E1: gesunde Freiwillige, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01318265	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		of Ketoconazole on the Pharmacokinetics of a Single Oral Dose of Regorafenib (BAY73-4506) in Healthy Volunteers				
21	NCT01322438	A Phase 1, Non-Randomized, Open Label Study to Determine the Effect of Rifampin on the Pharmacokinetics of a Single Oral Dose of 160 mg Regorafenib (BAY73-4506) in Healthy Volunteers	Abgeschlossen	ja (E1: gesunde Freiwillige, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01322438	Clinicaltrials.gov ICTRP
22	NCT01339104	An Open-label, Non-randomized Phase I Study of Regorafenib (BAY73-4506) to Evaluate Cardiovascular Safety Parameters, Tolerability, Pharmacokinetics, and Anti-tumor Activity in Patients With Advanced Solid Tumors	Aktiv, nicht rekrutierend	ja (E1: Patienten mit fortgeschrittenen Tumorerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01339104	Clinicaltrials.gov ICTRP
23	NCT01466504	A Study of the Effect of Sorafenib or Regorafenib on p63 Expression and Keratinocyte Differentiation in Human Skin	Abgebrochen	ja (E1: Hauterkrankungen, E4: keine RCT)	http://ClinicalTrials.gov/show/NCT01466504	Clinicaltrials.gov
24	NCT01538680	An Open-label Phase IIIb Study of Regorafenib in Patients With Metastatic Colorectal Cancer (CRC) Who Have Progressed After Standard Therapy	Verfügbar	ja (E5: einarmig)	http://ClinicalTrials.gov/show/NCT01538680	Clinicaltrials.gov PharmNet.bund
25	NCT01584830	A Randomized, Double-blind, Placebo-controlled Phase III Study of Regorafenib Plus	Aktiv, nicht rekrutierend	ja (E4: voraussichtliches Studienende in 2014)	http://ClinicalTrials.gov/show/NCT01584830	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		Best Supportive Care (BSC) Versus Placebo Plus BSC in Asian Subjects With Metastatic Colorectal Cancer (CRC) Who Have Progressed After Standard Therapy				
26	NCT01646593	An Open-label Expanded Access Program of Regorafenib in Patients With Gastrointestinal Stromal Tumors (GIST) After Disease Progression on or Intolerance to Imatinib and Sunitinib	Verfügbar	ja (E1: gastrointestinaler Stromatumor, E4: keine RCT)	http://ClinicalTrials.gov/show/NCT01646593	Clinicaltrials.gov
27	NCT01689376	Guidelines for Treatment With Regorafenib in Patients With Gastrointestinal Stromal Tumors (GIST) After Disease Progression on or Intolerance to Imatinib and Sunitinib (Managed Access Program)	Verfügbar	ja (E1: gastrointestinaler Stromatumor, E4: keine RCT)	http://ClinicalTrials.gov/show/NCT01689376	Clinicaltrials.gov ICTRP
28	NCT01774344	A Randomized, Double Blind, Placebo Controlled, Multicenter Phase III Study of Regorafenib in Patients With Hepatocellular Carcinoma (HCC) After Sorafenib	Rekrutierend	ja (E1: hepatozelluläres Karzinom, E4: rekrutierend)	http://ClinicalTrials.gov/show/NCT01774344	Clinicaltrials.gov EU Clinical Trials Register ICTRP PharmNet.bund
29	NCT01786538	Randomized Phase III Study of Oxaliplatin, Fluorouracil and Leucovorin (FOLFOX) With or Without Regorafenib in Patients With Metastatic Colorectal	Abgebrochen	ja (E2: Patienten, die eine Progression nach der Therapie mit Irinotecan und	http://ClinicalTrials.gov/show/NCT01786538	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
		Cancer Progressed After First-line Irinotecan Plus Fluoropyrimidines		Fluoropyrimidinen aufweisen, E2: Regorafenib in Kombination mit weiteren Therapien)		
30	NCT01843400	Drug Use Investigation of Regorafenib/ STIVARGA for Unresectable, Metastatic or Recurrent Colorectal Cancer	Rekrutierend	ja (E4: rekrutierend)	http://ClinicalTrials.gov/show/NCT01843400	Clinicaltrials.gov ICTRP
31	NCT01853046	A Phase I, Multi-center, Non-randomized, Open Label, Parallel-group Study Evaluating the Pharmacokinetics and Safety of Regorafenib (BAY 73-4506) in Cancer Subjects With Severe Renal Impairment Compared to a Control Group	Rekrutierend	ja (E1: Nierenerkrankungen, E4: Phase I)	http://ClinicalTrials.gov/show/NCT01853046	Clinicaltrials.gov ICTRP
32	NCT01853319	An Open-label Phase III Study of Regorafenib in Patients With Metastatic Colorectal Cancer (mCRC) Who Have Progressed After Standard Therapy	Noch nicht rekrutierend	ja (E4: noch nicht rekrutierend)	http://ClinicalTrials.gov/show/NCT01853319	Clinicaltrials.gov ICTRP
33	NCT01875380	A Phase II Study of Single-agent Regorafenib in the First Line Treatment of Frail and/or Unfit for Polychemotherapy Patients With Metastatic Colorectal Cancer (mCRC)	Rekrutierend	ja (E1: Patienten, die gebrechlich und/oder für Polychemotherapie nicht geeignet sind, E4:	http://ClinicalTrials.gov/show/NCT01875380	Clinicaltrials.gov ICTRP

	Studienbezeichnung	Titel	Status	Ausschlussgrund (ja/nein) (Grund)	URL	Identifikationsort
				rekrutierend)		
34	NCT01896856	Phase I Study of SGI-110 Combined With Irinotecan Followed by a Randomized Phase II Study of SGI-110 Combined With Irinotecan Versus Regorafenib in Previously Treated Metastatic Colorectal Cancer Patients	Noch nicht rekrutierend	ja (E4: noch nicht rekrutierend)	http://ClinicalTrials.gov/show/NCT01896856	Clinicaltrials.gov ICTRP
35	NCT01900743	Activity and Safety of Regorafenib in Patients With Metastatic Soft Tissue Sarcoma Previously Treated With Anthracycline-based Chemotherapy : a Multinational, Randomized, Phase II, Placebo-controlled Trial	Rekrutierend	ja (E1: Weichteilsarkome, E4: rekrutierend)	http://ClinicalTrials.gov/show/NCT01900743	Clinicaltrials.gov EU Clinical Trials Register ICTRP
36	NCT01913639	Phase II Study of FOLFOX Plus Regorafenib in Patients With Unresectable or Metastatic Esophagogastric Cancer	Rekrutierend	ja (E2: regorafenib in Kombination mit FOLFOX, E4: rekrutierend)	http://ClinicalTrials.gov/show/NCT01913639	Clinicaltrials.gov

**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 wurde keine bibliographische Recherche durchgeführt.

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 4-69 (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, da keine Studie, die im Abschnitt 4.3.1.1 identifiziert wurde, abgebrochen wurde.

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 4-70 (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						
CONCUR 15808	Ja (Asien/China)	Ja	Nein	Nein	Ja (55)	Nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
<p>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.</p>						

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 4-71 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 4-71 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-71 (Anhang): Studiendesign und -methodik für Studie CORRECT

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Vergleich der Wirksamkeit und Sicherheit von Regorafenib bei Patienten mit einem metastasierten Kolorektalkarzinom nach Progression von Standard-Vortherapien, die entweder mit Regorafenib plus Best-Supportive-Care (BSC) oder Placebo plus BSC behandelt wurden. Nachzuweisen ist, ob ein Vergleich zwischen der Interventionsgruppe Regorafenib + BSC vs. Placebo + BSC bezüglich des Gesamtüberlebens einen Unterschied zugunsten der Regorafenib + BSC Gruppe zeigt.
-	Methoden	
3	Studiendesign	Prospektive, randomisierte, zweiarmige, doppelblinde, kontrollierte, multizentrische Phase III Studie.
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase III doppelblinde, 2:1 randomisierte, multizentrische Studie mit zwei parallelen Gruppen (Regorafenib + BSC, Placebo + BSC)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Während des Studienverlaufs gab es drei Protokolländerungen. Keine hatte relevante Änderungen der Methodik der Studie zum Inhalt.
4	Probanden / Patienten	Patienten mit einem metastasierten Kolorektalkarzinom (Stadium IV), die alle zugelassenen Standardtherapien erhalten hatten und unter diesen eine Progression erlitten oder die darunter eine intolerable Toxizität entwickelten, konnten in die CORRECT-Studie aufgenommen werden. Zur gleichmäßigen Verteilung von Faktoren, die das Ergebnis möglicherweise beeinflussen können, wurden die Patienten, die in die Studie aufgenommen wurden, auf der Basis vorangegangener Anti-VEGF-Therapie (ja/nein), der Zeit seit der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate), sowie nach der geographischen Region (Region 1: Nordamerika, Westeuropa, Israel, Australien vs. Region 2: Asien vs. Region 3: Südamerika, Türkei, Osteuropa) stratifiziert.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> <ol style="list-style-type: none"> 1. Unterschriebene Einverständniserklärung vor der ersten studien-spezifischen Maßnahme. Patienten müssen die Einverständniserklärung verstehen können und diese unterzeichnen wollen. 2. Männer oder Frauen ≥ 18 Jahre. 3. Histologische oder zytologische Dokumentation eines Adenokarzinoms des Kolons oder Rektums. Alle anderen histologischen Typen sind ausgeschlossen. 4. Patienten mit einem metastasierten Kolorektalkarzinom (Stadium IV). 5. Progression während oder binnen drei Monaten nach der letzten Gabe einer zugelassenen Standardtherapie, die Fluoropyrimidine, Oxaliplatin, Irinotecan, Bevacizumab und

Item ^a	Charakteristikum	Studieninformation
		<p>Cetuximab oder Panitumumab (wenn KRAS WT vorliegt) umfassen muss. Patienten, die mit Oxaliplatin in einem adjuvanten Setting behandelt wurden, sollten während der Behandlung oder binnen sechs Monaten nach Beendigung der Behandlung eine Progression erlitten haben. Patienten, die nach sechs Monaten nach Beendigung einer adjuvanten Oxaliplatin-haltigen Therapie progredient waren, müssen erneut mit einer Oxaliplatin-basierten Therapie behandelt worden sein, um eingeschlossen werden zu können. Patienten, die eine Standardtherapie wegen intolerabler Toxizität vor einer Progression beendeten und bei denen eine erneute Therapie mit der gleichen Substanz nicht indiziert war, konnten ebenfalls in die Studie aufgenommen werden. Patienten mit einem unbekanntem KRAS-Status zum Screening mussten eine vorangegangene anti-EGFR-Therapie gehabt haben.</p> <ol style="list-style-type: none"> 6. Patienten mit einem mCRC mit messbaren oder nicht-messbaren Tumorerkrankungen, entsprechend RECIST-Kriterien, Version 1.1 7. Eastern Cooperative Oncology Group (ECOG) Performance Status ≤ 1 8. Lebenserwartung von mindestens drei Monaten 9. Frauen in gebärfähigem Zustand und Männer müssen zustimmen, eine adäquate Verhütungsmaßnahme zu verwenden von der Unterzeichnung der Einwilligungserklärung bis mindestens drei Monate nach dem Ende der Einnahme der Studienmedikation. 10. Adäquate Knochenmarks-, Leber- und Nierenfunktion, bestimmt anhand der folgenden Laborparameter innerhalb von sieben Tagen vor Beginn der Studienbehandlung: <ol style="list-style-type: none"> a. Gesamtbilirubin $\leq 1,5$fache des oberen Normwertes (ONW) b. Alaninaminotransferase (ALT) und Aspartataminotransferase (AST) $\leq 2,5$ fache des ONW c. Amylase und Lipase $\leq 1,5$fache des ONW d. Serumkreatinin $\leq 1,5$fache des ONW e. Glomeruläre Filtrationsrate (GFR) ≥ 30ml/min/1,73m² entsprechend der verkürzten MDRD-Formel f. International Normalized Ratio (INR) / Partielle Thromboplastinzeit (PTT) $\leq 1,5$fache des ONW. g. Thrombozytenzahl ≥ 100.000/mm³, Hämoglobin ≥ 9g/dl, Absolute Neutrophilenzahl ≥ 1.500/mm³. Transfusionen, um die genannten Werte zu erreichen, waren nicht erlaubt. h. Alkalische Phosphatase (ALP) $\leq 2,5$fache des ONW (≤ 5fache des ONW bei Patienten mit Lebermetastasen). <p><u>Ausschlusskriterien</u></p> <ol style="list-style-type: none"> 1. Vorangegangene Therapie mit Regorafenib 2. Vorherige Behandlung innerhalb dieser Studie. 3. Vorangegangene oder gleichzeitig bestehende andere Tumorerkrankung innerhalb von fünf Jahren vor der Randomisierung außer kurativ behandeltes carcinoma-in-situ der Zervix, nicht-melanomartige Hauttumoren und oberflächliches Harnblasenkarzinom

Item ^a	Charakteristikum	Studieninformation
		<ol style="list-style-type: none"> 4. Extended-Field Radiotherapie innerhalb von vier Wochen oder Limited-Field Radiotherapie innerhalb von zwei Wochen vor der Randomisierung. Patienten müssen sich von allen therapiebedingten Nebenwirkungen erholt haben. Die bestrahlte Läsion muss eine Progression aufweisen, wenn sie die einzige Tumorkläsion war. 5. Große chirurgische Eingriffe, offene Biopsien, oder signifikante Traumata innerhalb von 28 Tagen vor Beginn der Studienbehandlung. 6. Schwangere oder stillende Frauen. Frauen mit gebärfähigem Potential müssen einen negativen Schwangerschaftstest innerhalb von sieben Tagen aufweisen. 7. Chronische Herzinsuffizienz \geq NYHA-Klasse II 8. Instabile Angina pectoris, neu aufgetretene Angina oder Myokardinfarkt innerhalb von sechs Monaten vor Beginn der Studienbehandlung 9. Herzrhythmusstörungen, die einer antiarrhythmischen Therapie bedürfen (Betablocker und Digoxin sind erlaubt) 10. Unkontrollierte Hypertonie (systolischer Blutdruck $>$ 150 mmHg oder diastolischer Blutdruck $>$ 90 mmHg trotz optimaler Behandlung) 11. Patienten mit Phäochromozytom 12. Pleuraerguß oder Aszites die zur Atembehinderung führen (\geq CTCAE-Grad 2 in Dyspnoe) 13. Arterielle oder venöse Thrombosen oder Embolien, wie z. B. ischämischer Hirninfarkt (inkl. einer transitorisch-ischämischen Attacke), tiefe Beinvenenthrombosen oder Lungenembolien innerhalb von sechs Monaten vor Beginn der Studienbehandlung 14. Bestehende Infektion $>$ Grad 2 CTCAE Version 3.0 15. Bekannte HIV-Infektion 16. Aktive Hepatitis B oder C, oder chronische Hepatitis B oder C, die einer Behandlung mit antiviralen Substanzen bedarf 17. Patienten mit Epilepsie, die medikamentös behandlungsbedürftig sind 18. Anamnese von oder bestehende Hirnmetastasen 19. Z. n. Organtransplantation 20. Patienten mit Anamnese oder bestehender Blutungsdiathese 21. Blutungen \geq Grad 3 CTCAE innerhalb von vier Wochen vor Beginn der Studienbehandlung 22. Nicht-heilende Wunden, Ulcera oder Knochenfrakturen 23. Nierenversagen mit Erfordernis einer Hämo- oder Peritonealdialyse 24. Dehydratation \geq Grad 1 CTCAE Version 3.0 25. Substanzabusus, medizinische, psychologische oder soziale Umstände, die mit der Studienteilnahme oder der Interpretation der Ergebnisse interferieren würden 26. Bekannte Hypersensitivität gegenüber der Studienmedikation, der Klasse der Studienmedikation oder den Hilfsstoffen 27. Jedwede Erkrankung oder medizinischer Begleitumstand, der instabil ist und die Sicherheit des Patienten oder seine

Item ^a	Charakteristikum	Studieninformation
		<p>Compliance innerhalb der Studie gefährden könnte</p> <p>28. Interstitielle Lungenerkrankung mit bestehenden Symptomen zur Zeit der Unterzeichnung der Einverständniserklärung</p> <p>29. Persistierende Proteinurie von CTCAE-Grad 3 oder höher (> 3,5 g/24 Stunden, gemessen anhand der Protein//Kreatinin-Ratio in einer beliebigen Urinprobe)</p> <p>30. Patienten, die orale Medikation nicht schlucken können</p> <p>31. Jedwedes Malabsorptionssyndrom</p> <p>32. Enge Beziehungen zum Prüfzentrum, z. B. naher Verwandter des Prüfarztes, abhängige Person (z. B. Angestellter oder Student des Prüfzentrums, der Zugang zu den Studienunterlagen und zum CRF hätte)</p> <p>33. Bestehende Toxizität grösser als CTCAE-Grad 1, die einer vorangegangenen Therapie zugeordnet wird, ausgenommen Alopezie und Oxaliplatin-induzierte Neurotoxizität ≤ Grad 2</p> <p>34. Systemische antitumoröse Therapie einschließlich zytotoxischer Therapie, Signaltransduktionshemmer, Immuntherapie und Hormontherapie innerhalb dieser Studie oder von 4 Wochen (bzw. sechs Wochen für Mitomycin C) vor Beginn der Studienbehandlung</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>Multizentrisch</p> <p>Aktive Zentren in folgenden Ländern (Anzahl der Zentren): China (1), Japan (19), Australien (5), Belgien (6), Kanada (5), Frankreich (9), Deutschland (15), Israel (5), Italien (9), Niederlande (2), Spanien (8), Schweiz (1), USA (17), Tschechien (2), Ungarn (1).</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p><u>Interventionsgruppe:</u> BSC + Regorafenib 160 mg (4 x 40 mg-Tabletten) einmal täglich für drei Wochen, gefolgt von einer Woche Pause (= vier-wöchiger Zyklus)</p> <p><u>Kontrollgruppe:</u> BSC + Placebo vier identische Tabletten einmal täglich für drei Wochen, gefolgt von einer Woche Pause (= vier-wöchiger Zyklus)</p>
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><u>Primäres Zielkriterium:</u> Vergleich des Gesamtüberlebens von Patienten, die mit BSC + Regorafenib und BSC + Placebo behandelt wurden.</p> <p>Zeitraum: Ab der Randomisierung bis zum Sterbezeitpunkt unabhängig von der Todesursache. Eine Zensurierung fand zum letzten Zeitpunkt statt, an dem sicher war, dass der Patient lebt. Wenn ein Patient für die Beobachtung verloren ging und kein Kontakt nach der Randomisierung stattfand, wurde er zum Tag 1 zensiert.</p> <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> a) Progressionsfreies Überleben (PFS), definiert als Zeit (in Tagen) von der Randomisierung bis zur ersten beobachteten Krankheitsprogression (radiologisch oder klinisch) oder dem Tod jedweder Ursache, wenn der Tod vor der Progression dokumentiert wurde. b) Objektive Tumoransprechrates (ORR), definiert als prozentualer Anteil der Patienten, die als bestes Gesamtansprechen eine komplette (CR) oder partielle

Item ^a	Charakteristikum	Studieninformation
		<p>Remission (PR) aufwiesen. Ein bestes Gesamtansprechen wurde für alle Patienten mittels der RECIST-Kriterien, Version 1.1 definiert. Patienten, deren bestes Gesamtansprechen keine CR oder PR war, und Patienten, die keine Post-Baseline Tumorbeurteilung hatten, wurden für diese Analyse als Nonresponder gewertet. Radiologische Tumorbeurteilungen fanden zur Baseline-Erhebung sowie alle acht Wochen während der Behandlungsphase, sowie ggf. zum Ende der Behandlung, statt. In Fällen, in denen keine radiologische Beurteilung möglich war, wurde die Progression klinisch diagnostiziert. Eine klinische Progression basierte auf der Beurteilung des Prüfarztes.</p> <p>c) Krankheitskontrollrate (DCR), definiert als Prozentzahl an Patienten, deren bestes Ansprechen nicht PD (d. h. CR, PR, oder stabile Erkrankung) war.</p> <p><u>Tertiäre Zielkriterien:</u></p> <p>a) Dauer des Ansprechens, definiert als Zeit (in Tagen) vom ersten dokumentierten Ansprechen mit einer PR oder CR, was immer auch früher auftrat, bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression eintrat).</p> <p>b) Dauer der stabilen Erkrankung, definiert als Zeit (in Tagen) von der Randomisierung bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression eintrat). Dieses Kriterium wurde nur für die Patienten ausgewertet, die keine CR oder PR als bestes Gesamtansprechen aufwiesen.</p> <p>c) Patientenberichtete Endpunkte (PRO): Gesundheitsbezogene Lebensqualität wurde mittels der Fragebögen EORTC QLQ-C30 und EQ-5D gemessen.</p> <p>d) Pharmakokinetik von Regorafenib, um die Konzentrationen von Regorafenib und seiner Metaboliten M-2 (BAY 75-7495), M-4 (BAY 75-1098) und M-5 (BAY 81-8752) zu bestimmen.</p> <p>e) Biomarker: Genetische und Proteinbiomarker, analysiert in Plasma und Vollblutproben sowie aus archivierten diagnostischen Tumorproben (mit separater Einwilligungserklärung für genetische Untersuchungen).</p> <p><u>Verträglichkeitskriterien:</u></p> <p>Die Sicherheit wurde anhand von körperlichen Untersuchungen, Untersuchung aller Organsysteme bzw. relevanter Organsysteme, Vitalzeichen (Herzfrequenz, Blutdruck, Körpertemperatur), Gewicht und Körpergröße, EKG, Echokardiographie oder MUGA, Laboruntersuchungen sowie der Registrierung von unerwünschten Ereignissen bis 30 Tagen nach Beendigung der Studienmedikation analysiert.</p> <p>Die National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) Version 3.0 wurde für die Beurteilung der Toxizität und bei Berichten zu schwerwiegenden unerwünschten Ereignissen verwendet.</p> <p>Ein unabhängiges Data Monitoring Committee (DMC) wurde für diese Studie gegründet, um Sicherheits- und Wirksamkeitsdaten während der Studie zu evaluieren.</p>

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Fallzahlberechnung basiert auf dem primären Zielkriterium Gesamtüberleben. Die Studie ist konzipiert, um mit 90 % Power einen 33,3%igen Anstieg im medianen Gesamtüberleben (d. h. eine Hazard Ratio von 0,75 von Regorafenib gegenüber Placebo) zu zeigen. Unter der Annahme eines einseitigen Gesamt-Alphas von 0,025, einer Power von 90 %, einer Randomisierung im Verhältnis 2:1 zwischen Regorafenib und Placebo und zwei formalen Zwischenanalysen für OS während der Studie unter Verwendung des Verbrauchs des alpha-Fehlers nach O'Brien-Fleming werden insgesamt 582 Todesfälle für die Endanalyse benötigt. Wenn die Grenzen zur Beendigung der Studie in Bezug auf OS bei den Zwischenanalysen nicht überschritten werden, werden die OS-Daten als reif angesehen, wenn mindestens 582 Todesfälle aufgetreten sind. Es wurde projiziert, dass 582 Todesfälle nach ungefähr 32 Monaten auftreten, unter der Annahme, dass 690 Patienten mit einer Rekrutierungsgeschwindigkeit von 30 Patienten pro Monat nach einer anfänglichen vier-monatigen Ramp-up Periode, einer Drop-out-Rate von 3 %, einer exponentiellen Verteilung der OS-Zeiten und 4,5 bzw. 6,0 Monaten OS in der Placebo- bzw. Regorafenib-Gruppe, rekrutiert werden. Ca. 690 Patienten sollten randomisiert werden (460 Patienten im Regorafenibarm, und 230 Patienten im Placeboarm der Studie).</p> <p>Die Fallzahl wurde mittels der Software East Version 5 berechnet.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Zwei formale Zwischenanalysen waren während der Studie geplant, und zwar wenn ca. 30 % und 70 % der geplanten Gesamtanzahl an benötigten Todesfällen eingetreten waren. Die erste Zwischenanalyse (bei ca. 30 % der geplanten Gesamtereigniszahl) diente lediglich als Analyse auf Vergeblichkeit (Futility). Die zweite Zwischenanalyse (bei ca. 70 % der geplanten Gesamtereigniszahl) wurde zur Analyse auf Wirksamkeit und Futility durchgeführt. Grenzen zur Beendigung der Studie wurden für die Zwischenanalysen auf der Basis der bis zum Data- cut-off tatsächlich eingetretenen Ereigniszahlen berechnet. Eine Lan-Demets Alpha-Verbrauchsfunktion bestimmte den Grenzwert zur vorzeitigen Beendigung wegen Wirksamkeit, so dass die Gesamtrate für einen falsch positiven Ausgang, alpha, $\leq 0,025$ (einseitig) war. Ähnlich den Grenzwerten zur Beendigung wegen Wirksamkeit waren separat Grenzwerte für die Zwischenanalysen zur Beendigung wegen Futility berechnet worden. Die Grenzwerte für Futility basierten auf der Ablehnung einer wahren Hazard Ratio (HR) von 0,7502 oder weniger (korrespondierend zu einer beabsichtigten 33,3%igen oder mehr Erhöhung im medianen Gesamtüberleben gegenüber Placebo). Da die Grenzwerte für Wirksamkeit und Futility voneinander unabhängig waren, würde eine Nicht-Adhärenz der Futility-Grenzwerte nicht die Gesamt-Falschpositivenrate, alpha, auf über 0,025 (einseitig) erhöhen. Die Empfehlung zur Fortsetzung der Studie war sowohl von den Grenzwerten für die formalen Zwischenanalysen als auch von den Analysen zur Sicherheit anhand der Analysen aller Sicherheitsdaten entsprechend der DMC Charter abhängig.</p>

Item ^a	Charakteristikum	Studieninformation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Am Ende der Screeningperiode wurden geeignete Patienten zufällig im Verhältnis 2:1 in einer verblindeten Art und Weise dem Arm Regorafenib + BSC oder dem Arm Placebo+ BSC zugeteilt. Um dies zu gewährleisten, wurde eine vom Bayer Randomisierungsmanager durch ein Computerprogramm erstellte Randomisierungsliste erzeugt, und zwar mit einem unizentrischen Randomisierungsschema. Die Randomisierungsnummer für jeden einzelnen Patienten erfährt der Prüfarzt durch ein Interaktives Voice Response System (IVRS).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde stratifiziert nach: <ul style="list-style-type: none"> - Vorgegangener Anti-VEGF-Therapie (ja/nein), - Zeit von der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate) - Geographische Region 1 (Nordamerika, Westeuropa, Israel, und Australien) vs. Region 2 (Asien) vs. Region 3 (Südamerika, Türkei und Osteuropa). Um eine ausgewogene Repräsentation der drei Regionen zu gewährleisten, sollten in der Region 2 (Asien) nicht mehr als 250 Patienten randomisiert werden. Die Stratifizierung erfolgte ebenfalls über das IVRS-System und dem darin hinterlegten Randomisierungs- und Stratifizierungsschema.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Geheimhaltung der Gruppenzugehörigkeit war durch das zentralisierte Randomisierungsverfahren und durch die zentrale Vergabe der Patientennummern über Computer (IVRS) gewährleistet.
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?	Der Prüfarzt oder sein Beauftragter ruft das IVRS nachdem die Erfüllung aller Ein- und Ausschlusskriterien bestätigt wurden, an. Das IVRS vergibt eine Randomisierungsnummer für den Patienten, die dazu benutzt wurde, den Patienten einem Behandlungsarm zuzuordnen. Patienten müssen innerhalb von sieben Tagen nach der Randomisierung mit der Studienbehandlung beginnen. Wenn dies nicht der Fall ist, muss der Sponsor unter Angabe des Grundes warum die Behandlung nicht begonnen wurde, benachrichtigt 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	Patienten wurden randomisiert, um Regorafenib oder ein identisch aussehendes Placebo in einer doppelblinden Art und Weise zu erhalten, d. h. weder der Prüfarzt, Sponsor, noch der Patient wissen, welcher Behandlungsmethode der Patient zugeteilt wurde. Regorafenib und Placebo waren identisch im Aussehen, um die Verblindung zu gewährleisten. Patienten nahmen vier 40 mg-Tabletten einmal pro Tag oral ein. Um hier die Verblindung zu sichern, wurde die Studienmedikation (Regorafenib oder Placebo)

Item ^a	Charakteristikum	Studieninformation
	die Verblindung vorgenommen?	mit einer nur einmal vorkommenden Packungsnummer versehen, die auf jede Flasche aufgedruckt wurde und durch das IVRS dem jeweiligen Patienten zugeordnet war.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Interventionen waren identisch.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p><u>ITT-Population:</u> alle Patienten, die randomisiert und den Behandlungsgruppen zugeteilt wurden.</p> <p><u>Safety Population:</u> alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben.</p> <p><u>Gesamtüberleben:</u> Der Vergleich des Gesamtüberlebens zwischen den beiden Behandlungsgruppen wurde mittels eines stratifizierten log-Rank-Tests durchgeführt, wobei die Stratifizierung wie bei der Randomisation nach erfolgter Vortherapie mit VEGF-basierten Medikamenten (ja/nein), der Zeit von der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate vs. < 18 Monate) und der geographischen Region 1 (Nordamerika, Westeuropa, Israel, Australien) vs. Region 2 (Asien) vs. Region 3 (Südamerika, Türkei, Osteuropa) erfolgte. Um eine balancierte Repräsentation der drei Regionen zu gewährleisten, sollte die Region 2 nicht mehr als 250 Patienten randomisieren. Zusätzlich wurde die HR (Regorafenib + BSC / Placebo + BSC) für das Gesamtüberleben und ihr 95 %iges Konfidenzintervall mit einem Cox-Modell berechnet, stratifiziert nach den o. g. Faktoren. Kaplan-Meier (KM)-Schätzungen für das Gesamtüberleben und KM-Überlebenskurven wurden für jede der Behandlungsgruppen dargestellt. Die KM-Schätzungen zu den Zeitpunkten drei Monate, sechs Monate, usw. zusammen mit ihren 95 %-KI sowie die Differenzen dieser Schätzungen zwischen der Regorafenib + BSC und der Placebo + BSC-Gruppe wurden ebenfalls berechnet.</p> <p><u>Progressionsfreies Überleben:</u> Das progressionsfreie Überleben (PFS) wurde definiert als Zeit (Tage) von der Randomisierung bis zum Datum der ersten beobachteten Progression (radiologisch oder klinisch) oder dem Tod jedweder Ursache, wenn der Tod vor der Progression eintrat. Das Datum der radiologischen Tumorbeurteilung wurde für diese Berechnung benutzt. Wenn die Tumorbeurteilung länger als einen Tag benötigte (z. B. Thorax- und Abdomen-CT wurden an unterschiedlichen Tagen im selben Beobachtungszyklus durchgeführt), wurde das frühere Datum für die Berechnung des PFS benutzt. Für Patienten ohne eine dokumentierte klinische oder radiologische Progression oder Tod zum Zeitpunkt der Analyse wurde das PFS am Tag der letzten Tumorevaluierung zensiert. Das PFS für Patienten, die nach der Baseline keine Tumorevaluierung hatten aber noch am Leben waren, wurde am Tag 1 zensiert. Wenn eine Progression nach zwei aufeinander folgenden nicht durchgeführten oder nicht-auswertbaren Tumorevaluierungen</p>

Item ^a	Charakteristikum	Studieninformation
		<p>erfolgte (d. h. Progression später als das Datum der letzten auswertbaren Tumorevaluierung + 16 Wochen + 1 Woche), wurde das PFS am Tag der letzten auswertbaren Tumorevaluierung vor den beiden nicht durchgeführten Evaluierungen zensiert.</p> <p>Tod ohne Progression wurde als PFS-Ereignis nur dann gewertet, wenn er innerhalb der 16+1 Woche nach letzter auswertbarer Tumorevaluierung auftrat. Wenn er später auftrat, wurde das PFS zum Tag der letzten auswertbaren Tumorevaluierung zensiert.</p> <p>Für Patienten, die die Studie ohne Progression beendeten, wurde das PFS am Tag der letzten auswertbaren Tumorevaluierung zensiert, es sei denn, der Patient verstarb innerhalb von 16 + 1 Woche nach der letzten auswertbaren Tumorevaluierung. In diesem Fall wurde der Tod als PFS-Ereignis gewertet.</p> <p>Für Patienten, die die Antitumorbehandlung auf etwas anderes als die Studienmedikation vor dem Beobachten einer Progression wechselten, wurde das PFS am Tag der letzten Tumorevaluierung vor dem Therapiewechsel zensiert, bzw. am Tag der letzten Tumorevaluierung unter Studienbehandlung, wenn der Patient in der Nachbeobachtungsperiode die Therapie wechselte.</p> <p>Fehlende oder nicht-auswertbare Tumorevaluierungen wurden bei der Berechnung des PFS nicht berücksichtigt. Jegliche Anstrengung wurde unternommen, um die radiologische Beurteilung für eine Tumorprogression zu erhalten. In Fällen, in denen Patienten aufgrund der Verschlechterung ihres Zustandes keine radiologische Tumorbeurteilung unterlaufen konnten, wurde das Datum der klinischen Progression für die Bestimmung des Zeitpunkts der Progression herangezogen.</p> <p>Für die Analyse des PFS wurden die beiden Gruppen mittels eines log-Rank-Tests, stratifiziert nach denselben Faktoren wie bei der Analyse des primären Endpunkts, verglichen. Die HR (Regorafenib + BSC / Placebo + BSC) und das zugehörige 95 %-KI wurden berechnet. Kaplan-Meier (KM)-Schätzungen und KM-Kurven wurden für jede Behandlungsgruppe erstellt. Die KM-Schätzungen zu den Zeitpunkten zwei Monate, vier Monate usw. wurden zusammen mit ihren zugehörigen 95 %-KI sowie den Differenzen dieser Schätzungen zwischen den beiden Gruppen Regorafenib + BSC und Placebo + BSC berechnet.</p> <p><u>Objektive Tumoransprechrates:</u></p> <p>Die Objektive Tumoransprechrates war definiert als die Prozentzahl der Patienten mit CR oder PR. Ein bestes Gesamtansprechen wurde für alle Patienten anhand der RECIST-Kriterien, Version 1.1 definiert. Patienten ohne CR oder PR, sowie Patienten, die ohne eine Tumorevaluierung die Studie beendeten, wurden für diese als Non-Responder gewertet.</p> <p><u>Krankheitskontrollrate:</u></p> <p>Die Krankheitskontrollrate (DCR) war definiert als Prozentzahl der Patienten, deren bestes Ansprechen nicht PD war (d. h. also CR, PR oder SD). Tumorevaluierungen mit SD als Ansprechen, die vor sechs Wochen nach Randomisierung erfolgten, wurden nicht mit einbezogen.</p> <p>Die Objektive Tumoransprechrates und die DCR wurden zwischen den Behandlungsgruppen mittels Cochran-Mantel-Haenszel-Test verglichen, der für dieselben Stratifizierungsfaktoren wie für den primären Endpunkt adjustiert war. Schätzungen mit ihren 95 %-KI</p>

Item ^a	Charakteristikum	Studieninformation
		<p>wurden für jede Behandlungsgruppe berechnet. Die Unterschiede in den objektiven Ansprechraten zwischen der Regorafenib – und Placebogruppe mit ihren zugehörigen 95 %-KI wurden ebenfalls berechnet.</p> <p>Zusammenfassende Statistiken für alle Kategorien des besten Ansprechens wurden zusätzlich dargestellt: CR, PR, SD, PD nach radiologischer Beurteilung, PD nach klinischer Beurteilung. Häufigkeiten und Prozentzahlen mit exakten 95 %-KI wurden angegeben.</p> <p>Wenn ein Patient keine Tumorevaluierung nach der Baseline aufwies, d. h. das beste Ansprechen fehlte, wurde der Patient als nicht auswertbar erachtet und nicht in den Numerator einbezogen, der Patient wurde jedoch in den Denominator für die Berechnung der objektiven Tumoransprechraten, der Krankheitskontrollrate sowie der besten Ansprechraten (CR, PR, SD, PD) einbezogen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Tertiäre Zielvariablen:</p> <p>Dauer des Ansprechens:</p> <p>Die Dauer des Ansprechens (für PR und CR) wurde definiert als Zeit (Tage) von der ersten dokumentierten CR oder PR, welches auch immer zuerst auftrat, bis zur Krankheitsprogression oder dem Tod (wenn der Tod vor der Progression eintrat). Die Dauer des Ansprechens wurde nur für Responder, d.h. Patienten mit CR oder PR, definiert. Die Daten der eigentlichen Tumorevaluierungen wurden für diese Berechnung benutzt. Die Dauer des Ansprechens wurde für Patienten, die keine Progression hatten oder nicht verstorben sind, zum Zeitpunkt ihrer letzten Tumorevaluierung zensiert.</p> <p>Da die Responder keine randomisierte Gruppe darstellen, wurden keine statistischen Tests durchgeführt. Die Analyse der Dauer des Ansprechens war deskriptiv. KM-Schätzungen und Verteilungskurven wurden für jede Gruppe dargestellt.</p> <p>Dauer der Tumorstabilisierung:</p> <p>Die Dauer der Tumorstabilisierung war definiert als Zeit (Tage) von der Randomisierung bis zur Progression oder dem Tod (wenn der Tod vor der Progression eintrat). Diese Variable wurde nur für Patienten berechnet, die keine PR oder CR als bestes Ansprechen aufwiesen. Auch wurden Patienten, die nach der Baseline keine Tumorevaluierung aufwiesen, ausgeschlossen. Die Daten der eigentlichen Tumorevaluierung wurden für diese Analyse herangezogen. Die Dauer der Tumorstabilisierung wurde für Patienten, die keine Progression hatten oder nicht verstorben sind, zum Zeitpunkt ihrer letzten Tumorevaluierung zensiert.</p> <p>Gesundheitsbezogene Lebensqualität:</p> <p>Die gesundheitsbezogene Lebensqualität wurde mittels EORTC QLQ-C30 sowie EQ-5D erhoben. Der EORTC QLQ-C30 beinhaltet fünf funktionale Skalen (körperliche, Rollen-, emotionale, soziale und kognitive Funktion), drei Symptomskalen (Fatigue, Schmerz, Übelkeit/Erbrechen), eine Skala zum Gesundheitsstatus, eine Anzahl einzelner Items, die zusätzliche Symptome erfassen (Dyspnoe, Schlafstörungen, Obstipation, Diarrhoe), sowie den finanziellen Einfluss. Der EQ-5D misst fünf Gesundheitsdimensionen: Mobilität, Selbstversorgung, gewöhnliche Tätigkeiten, Schmerz und</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Angst/Depression. Der Fragebogen enthält auch eine EQ-VAS, die den selbst-empfundene Gesundheitszustand der Befragten anhand einer vertikalen graduierten visuellen Analogskala misst.</p> <p>Die Daten zur gesundheitsbezogenen Lebensqualität mittels EORTC QLQ-C30 und EQ-5D wurden analysiert, um die Unterschiede zur gesundheitsbezogenen Lebensqualität und der Gesundheitsverlaufswerte zwischen den Behandlungsarmen anhand der Zeit-adjustierten Fläche unter der Kurve (AUC) mit allen erhältlichen Daten zu erfassen.</p> <p>Deskriptive zusammenfassende Statistiken werden für jeden der Endpunkte der gesundheitsbezogenen Lebensqualität zu jedem Zeitpunkt ihrer Erhebung und ihrer Änderung vom Baseline per Behandlungsgruppe präsentiert. Die Häufigkeit und Gründe für fehlende Messungen zur gesundheitsbezogenen Lebensqualität werden per Behandlungsgruppe zusammengefasst.</p> <p>Primäre Analyse der gesundheitsbezogenen Lebensqualität: Ein Kovarianz-Analysemodell (ANCOVA) wurde verwendet, um die Zeit-adjustierte AUC zwischen den Behandlungsgruppen unter Verwendung des Baseline-Scores und Stratifizierungsfaktoren zu vergleichen. Schätzungen der Mittelwerte, Standardfehler und 95 %-KI wurden für jede Behandlungsgruppe und für die Unterschiede zwischen den Behandlungsgruppen berechnet. Die Annahmen des ANCOVA-Modells wurden überprüft. Die Behandlung nach Kovariatinteraktionen wurde in den ANCOVA-Modellen geprüft und die Konsistenz der Behandlungseffekte auf die gesundheitsbezogene Lebensqualität über verschiedene Subgruppen hinweg, die über die Kovariate des Modells definiert wurden, wurde ermittelt. Im Fall, dass die Annahmen des ANCOVA-Modells nicht erfüllt wurden (z. B. Normalität und homogene Varianz der Irrtumstermini, Gleichheit für verschiedene Behandlungsregressionslinien), konnte eine Rank-Analyse der Kovarianz benutzt werden.</p> <p>Sekundäre Analyse der gesundheitsbezogenen Lebensqualität: Sensitivitätsanalysen mit verschiedenen Anrechnungsmethoden für fehlende Beurteilungen konnten durchgeführt werden. Zusätzliche explorative Analysen konnten bei Bedarf mit linearen gemischten Effektmodellen durchgeführt werden, um die Effekte der Behandlung, der Zeit oder anderer Kovariate auf die Endpunkte unter der Annahme, dass die fehlenden Daten zufällig verteilt waren, zu explorieren.</p> <p><u>Sensitivitätsanalysen:</u></p> <p>Sensitivitätsanalysen wurden wie folgt durchgeführt:</p> <ul style="list-style-type: none"> - OS und PFS wurden mittels eines unstratifizierten log-Rank-Tests getestet. - Patienten, die die Studie nicht aufgrund einer PD oder des Todes beendeten, begannen eine Nachbeobachtungsperiode und Tumorevaluierungen wurden solange gesammelt (wenn erhältlich), bis die Patienten eine PD aufwiesen. Bei der Sensitivitätsanalyse des PFS wurden alle erhältlichen Tumorevaluierungen, auch aus der Nachbeobachtungsperiode, einbezogen. <p><u>Subgruppenanalysen:</u></p> <p>Deskriptive Statistiken und HR-Schätzungen mit 95 %-KI für OS und PFS wurden in folgenden Kategorien für folgende Variablen</p>

Item ^a	Charakteristikum	Studieninformation
		<p>gebildet, vorausgesetzt, es gab eine genügend große Anzahl von Ereignissen innerhalb der Subgruppe über die behandlungsarme hinweg:</p> <ul style="list-style-type: none"> - Demographische Kriterien wie Ethnische Gruppe, Geschlecht, Altersgruppe (<65 Jahre, ≥ 65 Jahre) - Region: Region 1 (Nordamerika, Westeuropa, Israel, Australien), Region 2 (Asien), Region 3 (Südamerika, Türkei, Osteuropa) - Zeit seit der Diagnosestellung der metastasierten Erkrankung (≥ 18 Monate und < 18 Monate) - Vorgegangene systemische Antitumorthherapie: <ul style="list-style-type: none"> - Vorgegangene anti-VEGF-Therapie (ja/nein) - Vorgegangene Antitumorthherapie nach folgenden Kategorien: Fluoropyrimidine, Oxaliplatin, Irinotacen, Bevacizumab (FOIB); Fluoropyrimidine, Oxaliplatin, Irinotacen, Bevacizumab, Anti-EGFR-Antikörper (FOIBE) - Anzahl der Therapielinien (≤3, >3) - Anzahl der Therapielinien bei oder nach Diagnose der metastatischen Erkrankung (≤3, >3) - Historischer KRAS-Status (anhand der Information des Prüfarztes im CRF) - Andere wichtige Baseline-Charakteristika der Tumorerkrankung (z. B. ECOG Performance Status: 0 und 1) <p>Wenn wichtige Effekte in Subgruppen gefunden wurden, konnten Interaktionsanalysen zwischen der Behandlung und den Subgruppen für wichtige Effekte in Subgruppen durchgeführt werden.</p> <p>Verträglichkeit:</p> <p>Verträglichkeitsvariablen waren folgende: Ausmaß der Exposition, unerwünschte Ereignisse, Laborwertveränderungen (Hämatologie, klinische Chemie und Urinanalyse), Veränderungen in den Vitalparametern (Blutdruck, Herzfrequenz, Atemfrequenz, Temperatur) und im EKG.</p>
-	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) Regorafenib + BSC: N=505 Placebo + BSC: N=255 Entspricht der ITT-Population b) Regorafenib + BSC: N=500 Placebo + BSC: N=253 Entspricht der Safety-Population c) wie a)

Item ^a	Charakteristikum	Studieninformation
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Regorafenib + BSC: fünf Patienten erhielten nach der Randomisierung keine Studienmedikation: <ul style="list-style-type: none"> - Patient 140010006: Einwilligung vom Patienten zurückgezogen - Patient 200070005: UE durch klinische PD - Patient 200090002: UE durch klinische PD - Patient 220070006: UE nicht durch klinische PD - Patient 280050010: ECOG Performance Status>1 nach Randomisierung Placebo + BSC: zwei Patienten erhielten nach der Randomisierung keine Studienmedikation: <ul style="list-style-type: none"> - Patient 200080002: UE durch klinische PD - Patient 240040006: UE durch klinische PD
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Randomisierung der Patienten erfolgte vom 19.Mai 2010 bis zum 22.März 2011. Abschluss der Datenerhebung: 21.Juli 2011 (Cut-off Stichtag für zweite Zwischenanalyse) 13. November 2011: Beginn der Möglichkeit für Placebo-Patienten, auf Regorafenib zu wechseln (Cross-over) – Stichtag für eine weitere Analyse zum Gesamtüberleben
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde beendet, nachdem der primäre Endpunkt erreicht und analysiert wurde und alle Patienten die Behandlung innerhalb der Studie beendet haben.
a: nach CONSORT 2010 (3)		

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

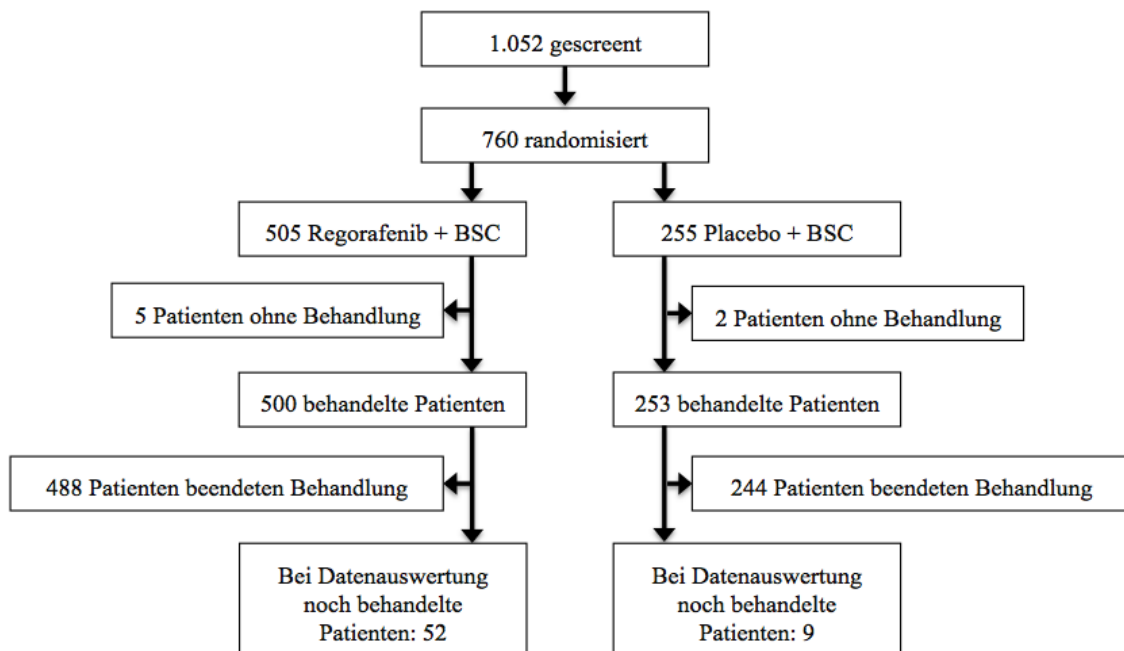


Abbildung 4-11: Patientenfluss CORRECT Studie Datenauswertungspunkt 1

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 4-72 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 14387 CORRECT

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. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: CORRECT 14387

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Klinischer Studienreport der Studie CORRECT 14387	CSR 14387

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:

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:

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:

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 Endpunkts, 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: __ Gesamtmortalität __

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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 (PFS)

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

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

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante 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:

Der Endpunkt wurde verblindet erhoben und objektiv evaluiert sowie adäquat auf Basis der ITT-Population ausgewertet.

Endpunkt: ___ Objektive Tumoransprechrates (ORR) ___

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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:

Der Endpunkt wurde verblindet erhoben und objektiv evaluiert sowie adäquat auf Basis der ITT-Population ausgewertet.

Endpunkt: ___Krankheitskontrollrate (DCR)___

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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:

Der Endpunkt wurde verblindet erhoben und objektiv evaluiert sowie adäquat auf Basis der ITT-Population ausgewertet.

Endpunkt: ___Dauer der Tumorstabilisierung___

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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:

Die Analyse der Dauer der Tumorstabilisierung basiert auf einer mITT-Population, die mit 242/255 (94,9 %) in der Placebo+BSC Gruppe und 470/505 (93,1 %) in der Regorafenib + BSC Gruppe um mehr als 5 % von der ITT-Population abweicht. Berechnet wurde dieser Endpunkt nur für die Patienten, die als beste Therapieantwort keinen kompletten oder partiellen Response gezeigt haben. Diese Bedingung ist aber notwendig, um die Dauer der Tumorstabilisierung zu berechnen. Diese Definition galt für beide Behandlungsgruppen in der Studie.

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:

Der Endpunkt wurde verblindet erhoben und objektiv evaluiert. Allerdings wurde als Analysepopulation eine modifizierte ITT-Population verwendet, die um mehr als 5 % von der ITT-Population abweicht. Berechnet wurde dieser Endpunkt nur für die Patienten, die als beste Therapieantwort keinen kompletten oder partiellen Response gezeigt haben. Diese Bedingung ist aber notwendig, um die Dauer der Tumorstabilisierung zu berechnen. Diese Definition galt für beide Behandlungsgruppen in der Studie. Daher wurde das Verzerrungspotential als niedrig eingestuft.

Endpunkt: ___ **Gesundheitsbezogene Lebensqualität anhand EORTC QLQ C30** ___

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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:

Von den 760 randomisierten Patienten hatten 697 (92 %) analysierbare Daten anhand des EORTC QLQ-C30-Fragebogens zur Baseline-Visite; 565 (74 %) hatten analysierbare Daten zum Zyklus 2, 256 (34 %) zum Zyklus 3, 206 (27 %) zum Zyklus 4, und 412 (54 %) zum jeweiligen Ende der Therapie. Ab Zyklus 5 wurden die Patientenzahlen pro Gruppe zu klein, um verwertbare Analysen der Lebensqualität durchzuführen. Die abnehmende Zahl der Fragebogen ausfüllender Patienten ist nicht ungewöhnlich, da im Verlauf der Studie die Ausfüllraten bedingt durch verschiedene Gründe, hauptsächlich durch Beendigung der Studie oder Eintritt des Todes, im allgemeinen rückläufig sind. Damit gab es eine Abweichung von der adäquaten Umsetzung des ITT-Prinzips, da für mehr als fünf Prozent der ITT-Population keine Daten vorlagen. Der häufigste Grund, warum die Fragebogen den Patienten sowohl in der Regorafenib- als auch in der Placebogruppe nicht ausgehändigt wurden, waren „Institution Errors“ sowie „Other“. Generell gab es jedoch zwischen den beiden Behandlungsgruppen keine deutlichen Unterschiede in der Anzahl der Patienten, die die Fragebogen zur Lebensqualität nicht ausgefüllt haben

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:

Die Analyse des EORTC QLQ-C30 basierte auf einer mITT-Population, die um mehr als 5 % von der ITT abwich. Es wurden Gründe für das Fehlen der Daten zur Lebensqualität beschrieben, eine Verzerrung der Ergebnisse ist aber nicht auszuschliessen. Sensitivitätsanalysen lagen nicht vor.

Endpunkt: ___ Unerwünschte Ereignisse (Verträglichkeit) ___

1. Verblindung der Endpunkterheber

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

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

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

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:

Der Endpunkt wurde verblindet erhoben und objektiv evaluiert sowie adäquat auf Basis der Sicherheitspopulation ausgewertet, die mit 253/255 (99,2 %) in der Placebo+BSC Gruppe und 500/505 (99,0 %) in der Regorafenib+BSC-Gruppe, um weniger als 5 % von der ITT-Population abweicht.
