

Nutzenbewertung



Gemeinsamer
Bundesausschuss

von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V

**Bewertung von Arzneimitteln für seltene Leiden nach § 35a
Absatz 1 Satz 11 i. V. m. Kapitel 5 § 12 Nr. 1 Satz 2 Verfo**

Wirkstoff: Zanidatamab

Datum der Veröffentlichung: 15. Mai 2026

Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abbildungsverzeichnis	4
Abkürzungsverzeichnis	5
Hintergrund	7
1 Fragestellung	8
2 Auswahl und Methodik der relevanten Studien	9
2.1 Studienbasis für die Nutzenbewertung	9
2.2 Beschreibung der eingeschlossenen Studie	9
2.3 Endpunkte	13
2.3.1 Mortalität	14
2.3.2 Morbidität	14
2.3.3 Lebensqualität	19
2.3.4 Sicherheit	19
2.3.5 Erhebungszeitpunkte	20
2.4 Statistische Methoden	20
2.5 Verzerrungspotential auf Studien- und Endpunktebene	22
2.6 Indirekter Vergleich	22
2.6.1 Zusammenfassende Bewertung der Methodik des indirekten Vergleichs	29
3 Ergebnisse der eingeschlossenen Studie	30
3.1 Studiencharakteristika und Studienmedikation	30
3.2 Mortalität	33
3.3 Morbidität	34
3.4 Lebensqualität	36
3.5 Sicherheit	37
4 Diskussion der Methodik und Ergebnisse	40
4.1 Design und Methodik der Studie	40
4.2 Zulassungsstatus und Zulassungspopulation	41
4.3 Mortalität	42
4.4 Morbidität	42
4.5 Lebensqualität	43
4.6 Sicherheit	43
5 Anforderung an eine qualitätsgesicherte Anwendung	44
6 Zusammenfassung der Nutzenbewertung	45
Referenzen	46
Anhang	47

Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis	9
Tabelle 2: Charakterisierung der Studie HERIZON_BTC.....	10
Tabelle 3: Charakterisierung der Intervention in der Studie HERIZON_BTC	13
Tabelle 4: Zusammenfassung der Endpunktbewertung der Studie HERIZON_BTC.....	13
Tabelle 5: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie HERIZON_BTC	20
Tabelle 6: Allgemeine Angaben; Studie HERIZON_BTC.....	30
Tabelle 7: Charakterisierung der Studienpopulation in der Studie HERIZON_BTC	30
Tabelle 8: Ergebnisse zum Gesamtüberleben in der Studie HERIZON_BTC (Datenschnitt: 28.07.2024); Efficacy Analysis Set.....	33
Tabelle 9: Anteil der Personen mit einer Verschlechterung der EQ-5D-VAS um ≤ 15 Punkte zum Ende der Behandlung; Studie HERIZON_BTC; Efficacy Analysis Set (Datenschnitt: 28.07.2024)	34
Tabelle 10: Anteil der Personen mit einer Verschlechterung der BPI-SF um ≤ 15 % zum Ende der Behandlung; Studie HERIZON_BTC; Efficacy Analysis Set (Datenschnitt: 28.07.2024)	35
Tabelle 11: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungsphase, Studie HERIZON_BTC.....	37
Tabelle 12: UE mit Inzidenz ≥ 10 % in der Studie HERIZON_BTC; Sicherheitspopulation	37
Tabelle 13: Schwere UE (Grad 3 oder höher), die bei ≥ 5 % der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der Studie HERIZON_BTC; Sicherheitspopulation	38
Tabelle 14: SUE, die bei ≥ 5 % der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der Studie HERIZON_BTC; Sicherheitspopulation.....	39
Tabelle 15: Zusammenfassende Darstellung der Ergebnisse der Studie HERIZON_BTC	45
Tabelle 16: Bestätigte objektive Ansprechrate in der Studie HERIZON_BTC (Datenschnitt: 28.07.2024); Efficacy Analysis Set.....	47
Tabelle 17: Herangezogene und verabreichte Interventionen in der externen Vergleichspopulation	47
Tabelle 18: Identifizierte und verwendete Confounder für den Indirekten Vergleich.....	47

Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung des Studienablaufs der Studie HERIZON_BTC	12
Abbildung 2: Kaplan-Meier-Kurve für Zeit bis zum Tod; Studie HERIZON_BTC (Efficacy Analysis Set)	34

Abkürzungsverzeichnis

AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ATT	Average treatment effect on the treated
BTC	Biliary Tract Cancer; Biliäres Karzinom
BPI-SH	Brief Pain Inventory – Short Form
CBR	klinische Nutzenrate
cORR	bestätigtes objektives Ansprechen
CR	Komplettes Ansprechen
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Krankheitskontrollrate
DOR	Dauer des Ansprechens
eCCA	extrahepatisches Cholangiokarzinom
ECOG PS	Eastern Cooperative Oncology Group Performance Score
ENSCCA	European Network for the study of Cholangiocarcinoma
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FHRD	Flatiron Health Research Analytic Database
G-BA	Gemeinsamer Bundesausschuss
GBC	Gallenblasenkarzinom
GKV	Gesetzliche Krankenversicherung
HER2	Human Epidermal Growth Factor Receptor 2
iCCA	intrahepatisches Cholangiokarzinom
IHC	Immunhistochemie
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ICR	Unabhängiges Review-Komitee
ISH	In-situ-Hybridisierung
ITT	Intention to Treat
k. A.	keine Angabe
KI	Konfidenzintervall
KI-Modell	Modell unter Verwendung von Künstlicher Intelligenz
MedDRA	Medical Dictionary for Regulatory Activities
ML	Machine Learning
MW	Mittelwert
N	Anzahl
NLP	Natural Language Processing
PFS	Progressionsfreies Überleben
PR	Partielles Ansprechen
PS-Modell	Propensity-Score-Modell
PT	Preferred Term
pU	pharmazeutischer Unternehmer
SAP	Statistischer Analyseplan

SGB	Sozialgesetzbuch
SITA	Strong Ignorable Treatment Assigment
SLR	systematische Literaturrecherche
SMD	Standardisierte Mittelwertsdifferenz
SMR	Standardized Mortality Ratio
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung des G-BA
ZNS	Zentrales Nervensystem
IVD	In-vitro-Diagnostikum

Hintergrund

Zanidatamab ist zugelassen als Arzneimittel 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. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

Gemäß § 35a Absatz 2 SGB V entscheidet der Gemeinsame Bundesausschuss (G-BA), ob er die Nutzenbewertung selbst durchführt oder das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragt. Ausgehend von der gesetzlichen Vorgabe in § 35a Absatz 1 Satz 11 1. Halbs. SGB V, dass der Zusatznutzen eines Orphan Drug durch die Zulassung als belegt gilt, hat der G-BA in seiner Sitzung vom 15. März 2012 das Verfahren der Nutzenbewertung von Orphan Drugs dahingehend modifiziert, dass bei Orphan Drugs zunächst keine eigenständige Festlegung einer zweckmäßigen Vergleichstherapie mehr durch den G-BA als Grundlage der insoweit allein rechtlich zulässigen Bewertung des Ausmaßes eines gesetzlich zu unterstellenden Zusatznutzens erfolgt. Vielmehr wird ausschließlich auf der Grundlage der Zulassungsstudien das Ausmaß des Zusatznutzens durch den G-BA bewertet.

Bei Orphan Drugs erfolgt eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann, wenn der Umsatz des betreffenden Arzneimittels die Umsatzschwelle gemäß § 35a Absatz 1 Satz 12 SGB V überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt.

Der G-BA bestimmt gemäß Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Zanidatamab zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den European Public Assessment Report (EPAR)), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der in Kapitel 5 § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Zanidatamab in seiner Sitzung am 12. Mai 2026 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 13. Februar 2026 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung (GKV)
- Anzahl der GKV-Patientinnen und -Patienten in der Zielpopulation

Die Nutzenbewertung wird am 15. Mai 2026 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Zanidatamb (Ziihera®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 30 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. Die Bewertung bezieht sich auf folgende Indikation [7]:

Erwachsene mit inoperablem, lokal fortgeschrittenem oder metastasiertem HER2-positivem (IHC 3+) biliärem Karzinom (Biliary Tract Cancer, BTC), die zuvor mit mindestens einer systemischen Therapie behandelt wurden.

Bei infrage kommenden Patientinnen und Patienten muss ein dokumentierter Human-Epidermal-Growth-Factor-Receptor-2-(HER2-)positiver Tumorstatus, definiert als Score von 3+ durch Immunhistochemie, vorliegen. Bewertet durch ein CE-gekennzeichnetes In-vitro-Diagnostikum (IVD), das für den entsprechenden Verwendungszweck validiert ist. Wenn kein CE-gekennzeichnetes IVD zur Verfügung steht, sollte ein alternativer validierter Test verwendet werden.

Gemäß Fachinformation beträgt die zugelassene Dosierung 20 mg/kg und wird als intravenöse Infusion alle zwei Wochen verabreicht bis zur Progression der Erkrankung oder bis zum Auftreten einer inakzeptablen Toxizität.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studiename (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studien zum Wirkstoff				
HERZION_BTC ¹⁾	Ja	Ja	Ja	-
Studien zu externen Kontrollen				
JZP598_515	Ja	Ja	Nein	Siehe Kapitel 2.6

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

Abkürzungen: EPAR: European Public Assessment Report; pU: pharmazeutischer Unternehmer

Zur Nutzenbewertung für Zanidatamab herangezogene Unterlagen

- Dossier des pU zu Zanidatamab [5]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [3]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie HERIZON_BTC [6]
- Unterlagen zum indirekten Vergleich [4]
- Fachinformation zu Zanidatamab [7]
- Zusatzanalysen zum AMNOG-Dossier [8]

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Zanidatamab im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie HERIZON_BTC. Für die Nutzenbewertung wird aus der Studie HERIZON_BTC die post-hoc für das Dossier definierte Auswertungspopulation herangezogen, die ausschließlich Personen mit überexprimierter HER2-Amplifizierung, definiert als Nachweis einer starken Intensität der Membranfärbung der Immunhistochemie (IHC 3+), umfasst. Diese Auswertungspopulation stimmt mit der Zulassungspopulation überein und ist für die Nutzenbewertung relevant.

Die Studie und die Intervention werden in Tabelle 2 und Tabelle 3 charakterisiert.

Tabelle 2: Charakterisierung der Studie HERIZON_BTC

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Eine multizentrische, offene, einarmige Phase-IIb-Studie, um die Anti-Tumor Aktivität von Zanidatamab Monotherapie in Erwachsenen mit HER2-amplifiziertem, inoperablen und fortgeschrittenen/metastasierten biliärem Karzinom (BTC)¹ zu untersuchen.</p> <p>Es wurden zwei Kohorten eingeschlossen:</p> <ul style="list-style-type: none"> • HER2-amplifiziert ermittelt durch ISH und HER2-Überexpression, definiert als IHC 2+ oder IHC 3+ (Kohorte 1); • HER2-amplifiziert ermittelt durch ISH und HER2 IHC 0 oder 1+². <p>Für die Nutzenbewertung relevant ist die Subpopulation der Kohorte 1 mit einer HER2-Überexpression (IHC 3+).</p> <p>Zum Zeitpunkt der Nutzenbewertung ist die Studie abgeschlossen.</p>
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Erwachsene ≥ 18 Jahre. • Histologisch oder zytologisch bestätigtes BTC (GBC, iCCA, eCCA). • Lokal fortgeschritten oder metastasiert und nicht geeignet für eine kurative Resektion, Transplantation oder Ablation. • Erhalt mindestens einer vorherigen Therapielinie mit Gemcitabin-haltiger systemischer Chemotherapie für eine fortgeschrittene Erkrankung und Progredienz nach der letzten vorangegangenen Therapie. <ul style="list-style-type: none"> ◦ Bei Personen, die im Rahmen einer früheren adjuvanten oder neoadjuvanten Behandlung mit Gemcitabin behandelt wurden, wird ein Fortschreiten der Erkrankung < 6 Monate nach der primären chirurgischen Resektion oder dem Abschluss einer Gemcitabin-haltigen adjuvanten Therapie als eine vorherige Therapielinie bei fortgeschrittener Erkrankung betrachtet. • Mindestens eine messbare Zielläsion gemäß RECIST 1.1. • Positiv auf HER2-Amplifikation mittels ISH-Assay in einem Zentrallabor (neue Biopsie oder Archivmaterial). • ECOG PS ≤ 1. • Adäquate hämatologische Funktion <ul style="list-style-type: none"> ◦ definiert als absolute Neutrophilenzahl $\geq 1,5 \times 10^9/L$, Thrombozytenzahl $\geq 75 \times 10^9/L$ (keine Transfusionsunterstützung erforderlich) und Hb ≥ 9 g/dL (Personen mit chronischer Anämie, die durch intermittierende Erythrozytentransfusionen unterstützt werden, sind zugelassen). • Adäquate Leberfunktion: Serumbilirubin $\leq 1,5 \times$ ULN oder $\leq 3 \times$ ULN bei Personen mit Gilbert-Krankheit, AST $\leq 3 \times$ ULN und ALT $\leq 3 \times$ ULN. Bei Personen mit Leberbeteiligung sind AST und ALT $\leq 5,0 \times$ ULN akzeptabel. • Adäquate Herzfunktion, definiert durch eine LVEF ≥ 50 %. • Adäquate Nierenfunktion: GFR ≥ 30 mL/min gemäß der MDRD-Gleichung. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Erhalt einer systemischen Krebstherapie innerhalb von 3 Wochen oder einer Radiotherapie innerhalb von 2 Wochen vor Zanidatamab. • Unbehandelte (symptomatische) Metastasen des ZNS oder Strahlenbehandlung für ZNS-Metastasen innerhalb von 4 Wochen vor Beginn der Studienbehandlung; <ul style="list-style-type: none"> ◦ Behandelte Hirnmetastasen sind zulässig.

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Meningeosis neoplastica. • Unkontrollierte oder aktive hepatobiliäre Erkrankungen, unbehandelte oder andauernde Komplikationen nach laparoskopischen Eingriffen oder Stentimplantation, u. a. aktive Cholangitis, ungelöste biliäre Obstruktion, infiziertes Bilom oder Abszess. • Akute oder chronische Infektionen. • Aktive Hepatitis B oder Hepatitis C <ul style="list-style-type: none"> ◦ Akute oder chronische Hepatitis B (Ausnahme: Personen, die Hepatitis-B-Oberflächenantigen-positiv sind, sind teilnahmeberechtigt, wenn sie eine HBV-DNA von weniger als 500 IU/mL aufweisen); ◦ Infektion mit Hepatitis C (Ausnahme [i] Personen, die keine kurative Virusbehandlung hinter sich haben und nachweislich eine negative Viruslast aufweisen, sind teilnahmeberechtigt; [ii] Personen, die eine kurative Virustherapie \geq 12 Wochen vor der Einschreibung abgeschlossen haben und deren Viruslast negativ ist, sind teilnahmeberechtigt). • QTcF > 470 ms. • Myokardinfarkt oder instabile Angina in der Vorgeschichte innerhalb von 6 Monaten vor der Aufnahme in die Studie; Troponinwerte, die auf einen Herzinfarkt hindeuten, klinisch bedeutsame Herzkrankheiten wie behandlungsbedürftige ventrikuläre Herzrhythmusstörungen, unkontrollierter Bluthochdruck oder eine symptomatische Herzinsuffizienz in der Vorgeschichte. • Akute oder chronische unkontrollierte Pankreatitis oder eine Lebererkrankung der Child-Pugh-Klasse C.
Intervention und Zahl der Patientinnen und Patienten	<p>Gescreent gesamt: N = 131 Eingeschlossen: N = 87 Intervention: N = 87 Kohorte 1: N = 80 IHC 3 + Kohorte 1: N = 62 Kohorte 2: N = 7</p>
Ort und Zeitraum der Durchführung; Datenschnitte	<p>Studienzentren 32 Studienzentren in 9 Ländern (Kanada, Chile, China, Frankreich, Italien, Spanien, Südkorea, Vereinigtes Königreich und USA)</p> <p>Studienzeitraum</p> <ul style="list-style-type: none"> • Erste/r Patient/in, erste Visite: 15. September 2020 • Letzte/r Patient/in, letzte Visite: 11. Juli 2024 <p>Datenschnitt Datenschnitt zum 1. Studienbericht: 10. Oktober 2022 (präspezifiziert) Datenschnitt zum 2. Studienbericht: 28. Juli 2023 Datenschnitt zum 3. Studienbericht: 28. Juli 2024 (final)</p>
Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Bestätigte objektive Ansprechrates nach RECIST v1.1 (cORR) durch IRC.</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Dauer des Ansprechens RECIST v1.1 (DOR) • Anteil der Personen mit einem DOR von \geq 16 Wochen bewertet nach RECIST v.1.1 • Bestes Gesamtansprechen (BOR) • Krankheitskontrolle (DCR) nach RECIST v1.1

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Progressionsfreies Überleben (PFS) nach RECIST v1.1 • Gesamtüberleben • Sicherheit <p>Die Endpunkte zur Bewertung des Ansprechens werden jeweils durch ein IRC und durch das Prüfpersonal bewertet.</p> <p>Explorative Endpunkte</p> <ul style="list-style-type: none"> • BPI • EQ-5D-5L

¹⁾ Biliäres Karzinom wurde in der Studie HERIZON_BTC definiert als Gallenblasenkarzinom (GBC), intrahepatisches (iCCA) sowie extrahepatisches Cholangiokarzinom (eCCA).

²⁾ Die Rekrutierung für die Kohorte 2 wurde im Studienverlauf früh beendet. Da die Studienpopulation der Kohorte 2 nicht der Zulassungspopulation entspricht, wird diese Population nicht weiterbetrachtet.

Abkürzungen: ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; BPI: Brief Pain Inventory; BTC: Biliäres Karzinom; cORR: bestätigte objektive Ansprechrare; DCR: Krankheitskontrollrate; DOR: Dauer des Ansprechens; eCCA: extrahepatisches Cholangiokarzinom; ECOG PS: Eastern Cooperative Oncology Group Performance Score; GBC: Gallenblasenkarzinom, GFR: glomeruläre Filtrationsrate; HER2: Human Epidermal Growth Factor Receptor 2; iCCA: intrahepatisches Cholangiokarzinom; IHC: Immunhistochemie; IRC: Unabhängiges Review-Komitee; ISH: In-situ-Hybridisierung; LVEF: linksventrikuläre Ejektionsfraktion; PFS: Progressionsfreies Überleben; QTcF: frequenzkorrigierte QT-Zeit nach Fridericia; ULN: Upper Limit of Normal; ZNS: Zentrales Nervensystem.

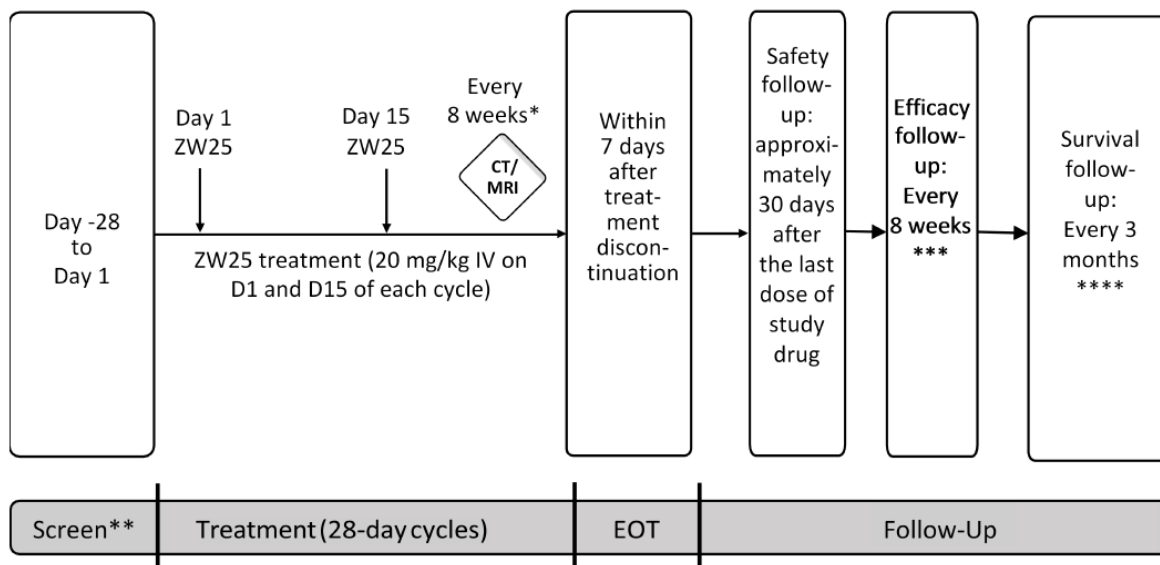


Abbildung 1: Schematische Darstellung des Studienablaufs der Studie HERIZON_BTC

Protokolländerungen

Es wurden drei Änderungen des Originalprotokolls vom 13.02.2020 vorgenommen. Für die Nutzenbewertung wurden keine relevanten Protokolländerungen identifiziert.

Charakterisierung der Intervention

Tabelle 3: Charakterisierung der Intervention in der Studie HERIZON_BTC

Intervention
<p>Im Rahmen der Intervention erfolgte eine prophylaktische Prämedikation (30 bis 60 Minuten vor jeder Zanidatamab-Infusion) mit folgenden Arzneimitteln:</p> <ul style="list-style-type: none"> • Kortikosteroide (Hydrokortison 100 mg i.v. oder Dexamethason 10 mg i.v.). • Antihistamine – Diphenhydramin 50 mg p.o. oder i.v. • Acetaminophen oder Paracetamol 650 bis 1.000 mg p.o. <p>Danach die folgende Gabe der Prüfmedikation</p> <ul style="list-style-type: none"> • Zanidatamab 20 mg/kg, i.v., an Tag 1 und 15 jedes 28-tägigen Zyklus. • Die Gabe der Prüfmedikation erfolgte bis zum Erreichen eines Abbruchereignisses (u. a. klinische Progression, Tod oder UE).
<p>Nicht erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Jegliche Intervention zur Behandlung der unterliegenden Krebserkrankung. • Chinesische oder Herbale Medizin. • Eine unterstützende Radiotherapie für Nicht-Ziel-Läsionen musste durch den Sponsor genehmigt werden. <p>Erlaubte Begleitmedikation</p> <ul style="list-style-type: none"> • Kortikosteroide • Transfusionen • Bisphosphonate • Routineprophylaxe mit Impfstoffen

Abkürzungen: i.v.: intravenös; p.o.: per os; UE: Unerwünschtes Ereignis.

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Alle Endpunkte der eingeschlossenen Studien (siehe Tabelle 2) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle 4 nicht gelistet sind, wurden weder vom pU noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Tabelle 4 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 4: Zusammenfassung der Endpunktbewertung der Studie HERIZON_BTC

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Gesamtüberleben	Mortalität	Ja	Ja
Tumoransprechen ¹⁾²⁾ <ul style="list-style-type: none"> • cORR • CBR • DCR • DOR 	Morbidität	Ja	Nein ³⁾
Progressionsfreies Überleben		Ja	Nein

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Schmerz mittels BPI-SF		Ergänzend	Ja
EQ-5D-VAS		Ergänzend	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Der Endpunkt „Tumoransprechen“ beinhaltet den primären Endpunkt „bestätigte objektive Ansprechrates“.

²⁾ Beinhaltet die Operationalisierungen: bestätigte objektive Ansprechrates (cORR), klinische Nutzenrate (CBR), Krankheitskontrollrate (DCR), Dauer des Ansprechens (DOR).

³⁾ Der primäre Endpunkt „bestätigte objektive Ansprechrates“ wird aus Transparenzgründen im Anhang dargestellt.

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; CBR: klinische Nutzenrate; cORR: bestätigte objektive Ansprechrates; DCR: Krankheitskontrollrate; DOR: Dauer des Ansprechens; pU: pharmazeutischer Unternehmer; VAS: Visuelle Analogskala.

2.3.1 Mortalität

Gesamtüberleben

Der Endpunkt „Gesamtüberleben“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der Endpunkt „Gesamtüberleben“ wurde in der Studie HERIZON_BTC operationalisiert als die Zeit von der ersten Dosis der Studienmedikation bis zum Tod jeglicher Ursache oder bis zum Zeitpunkt des letzten bekannten Überlebens der Personen, die nicht gestorben sind. Personen, die beim letzten Kontakt vor oder zum Datenschnitt nicht verstorben waren, wurden zum Datum des letzten Kontakts zensiert. Nach Progress/Beendigung der Studienmedikation erfolgte die Nachbeobachtung telefonisch alle drei Monate bis zum Tod.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

Validität

Die Erhebung des Gesamtüberlebens in der Studie HERIZON_BTC wird als valide eingeschätzt.

2.3.2 Morbidität

Tumoransprechen

Bestätigte objektive Ansprechrates

Der Endpunkt „Bestätigte objektive Ansprechrates“ wird als nicht patientenrelevant bewertet. Er wird als primärer Endpunkt in der Nutzenbewertung aus Transparenzgründen im Anhang dargestellt.

Operationalisierung

Beschreibung

Die bestätigte objektive Ansprechrates (cORR) wurde definiert als das Erreichen eines kompletten (CR) oder partiellen Ansprechens (PR) erhoben mittels der RECIST-1.1-Kriterien. Das Ansprechen wurde durch Scans basierend auf Computertomographie und/oder Magnetresonanztomographie erhoben. Dazu wurden die Brust, der Abdomen, das Becken sowie weitere Bereiche des Körpers untersucht, bei denen eine Tumormass auf Basis der Patientengeschichte (bspw. Gehirnmassmetastasen) vermutet wurde. Es sollte die gleiche Prozedur für jeden Scan der Personen gewählt werden. Weitere Erhebungen (bspw. Untersuchung der Knochen mittels Nuklearmedizin) durften nach Ermessen des Prüfpersonals durchgeführt werden.

Die Bewertung des Ansprechens erfolgte durch ein Unabhängiges Review-Komitee (ICR) für den primären Endpunkt. Bewertungen durch das Prüfpersonal sollten ebenfalls erfolgen (sekundärer Endpunkt). Wenn ein objektives Ansprechen bei einer Person aufgetreten ist, sollte dies innerhalb von 4 Wochen nach erstmaliger Erhebung bestätigt werden durch das Prüfpersonal. Die Bewertung des Ansprechens durch das Prüfpersonal war auch für darauffolgende jegliche Behandlungsentscheidungen von Bedeutung.

Die Erhebung des Endpunkts erfolgte vom Zeitpunkt der erstmaligen Einnahme der Prüfmedikation bis zu einer dokumentierten Krankheitsprogression, Tod jeglicher Ursache, Rücknahme der Einverständniserklärung oder Einnahme einer nicht Protokoll gemäßen Antikrebsbehandlung.

Bewertung

Die Operationalisierung ist nachvollziehbar. Die Erhebung und Bewertung der bildgebenden Verfahren durch ein unabhängiges Review-Komitee und das Prüfpersonal wird als adäquat eingeschätzt.

Patientenrelevanz

Der primäre Endpunkt „bestätigte objektive Ansprechrates“ basiert auf radiologischen Befunden. Die Beurteilung beruht somit ausschließlich auf bildgebenden Verfahren und berücksichtigt keine von Patientinnen oder Patienten wahrnehmbare Symptomatik. Der Endpunkt wird daher als nicht unmittelbar patientenrelevant bewertet.

Validität

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Tumoransprechen

Der Endpunkt „Tumoransprechen“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht gegebener Patientenrelevanz.

Operationalisierung

Beschreibung

Es liegen weitere Operationalisierungen des Tumoransprechens vor, die auf in den in cORR beschriebenen Methoden zur Erhebung des Ansprechens basieren.

- Klinische Nutzenrate: Erreichen einer stabilen Erkrankung für ≥ 24 Wochen oder eines bestätigten Ansprechens von PR oder CR gemäß RECIST 1.1;
- Krankheitskontrollrate: Definiert als eine stabile Erkrankung oder bestätigte CR oder PR gemäß RECIST 1.1;

- Dauer des Ansprechens: Zeit vom ersten bestätigten objektiven Ansprechen (CR oder PR) bis zum ersten dokumentierten Krankheitsprogress gemäß RECIST 1.1 oder Tod aus jeglicher Ursache;
- Zeit bis zum ersten bestätigten Ansprechen: Zeit von der ersten Dosis der Studienmedikation bis zum frühesten Datum, an dem ein für eine Patientin oder einen Patienten bestätigtes objektives Ansprechen (CR oder PR) erreicht wurde.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der primäre Endpunkt „bestätigte objektive Ansprechrate“ basiert auf radiologischen Befunden. Die Beurteilung beruht somit ausschließlich auf bildgebenden Verfahren und berücksichtigt keine von Patientinnen oder Patienten wahrnehmbare Symptomatik. Der Endpunkt wird daher als nicht unmittelbar patientenrelevant bewertet.

Validität

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Progressionsfreies Überleben

Der Endpunkt „Progressionsfreies Überleben“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht gegebener Patientenrelevanz.

Operationalisierung

Beschreibung

Der Endpunkt Progressionsfreies Überleben (PFS) basiert auf den ICR-Bewertungen gemäß RECIST 1.1 und ist definiert als die Zeit zwischen erster Dosis des Studienmedikaments und dem Zeitpunkt der dokumentierten Krankheitsprogression gemäß RECIST 1.1 oder Tod jeglicher Ursache. Die Erhebung des Ansprechens erfolgt analog zu dem unter cORR beschriebenen Vorgehen.

Zusätzlich wird ein klinischer Progress operationalisiert. Ein klinischer Progress ist definiert als die Verschlechterung oder das Wiederauftreten bereits bestehender Symptome im Zusammenhang mit der zugrundeliegenden Krebserkrankung. Eine Verschlechterung des Eastern Cooperative Oncology Group Performance Score (ECOG PS) kann ebenfalls auf eine Krankheitsprogression hindeuten. Die Krankheitsprogression sollte radiographisch festgehalten werden. Sollte die klinische Krankheitsprogression eindeutig und die Durchführung der radiologischen Untersuchung nicht möglich gewesen sein, konnte auf die radiologische Dokumentation der Krankheitsprogression verzichtet werden. Beim Vorliegen einer eindeutigen Krankheitsprogression ohne radiologische Untersuchung sollte das Prüfpersonal die Gabe der Studienmedikation aufgrund der Krankheitsprogression beenden.

Folgende Zensierungsgründe wurden präspezifiziert:

- andauerndes radiologisches Follow-up;
- Rücknahme der Einwilligung;
- nachfolgende Anti-Krebs-Therapie initiiert;
- ≥ 2 fehlende/nicht auswertbare Bewertungen des Ansprechens;
- keine Post-Baseline-Bewertung des Ansprechens.

Es wurden zwei Sensitivitätsanalysen präspezifiziert. In der Sensitivitätsanalyse 1 wird die „klinische Krankheitsprogression“ als zusätzliches Ereignis definiert. In der Sensitivitätsanalyse 2 wird die nachfolgende Anti-Krebs-Therapie als ein Ereignis operationalisiert. Eine Sensitivitätsanalyse nur unter Einbezug der „klinischen Krankheitsprogression“ konnte nicht identifiziert werden.

Bewertung

Die Operationalisierung ist nachvollziehbar. Einschränkend wird erwähnt, dass die Zensierung aufgrund einer nachfolgenden Anti-Krebs-Therapie nicht dem Treatment-Policy-Ansatz entspricht. Die Sensitivitätsanalyse 1 bezieht die klinische Krankheitsprogression als konkurrierendes Ereignis zur radiologischen Krankheitsprogression ein. Es kann nicht ausgeschlossen werden, dass das Ereignis der radiologischen Krankheitsprogression die Erhebung der klinischen Krankheitsprogression verdeckt. Konkrete Kriterien, anhand derer das Prüfpersonal die klinische Progression feststellen soll, werden nicht näher spezifiziert. Dadurch bleibt bspw. unklar, anhand welcher Symptome eine klinische Progression festgestellt wurde und inwieweit eine systematische Erhebung dieser Symptome erfolgte.

Patientenrelevanz

PFS ist ein kombinierter Endpunkt aus Endpunkten der Kategorien „Mortalität“ und „Morbidität“, wobei die Komponente „Tod“ im patientenrelevanten Endpunkt „Gesamtüberleben“ abgebildet ist. Die Feststellung der radiologischen Krankheitsprogression erfolgte nach den Responsekriterien RECIST 1.1 anhand bildgebender Verfahren ohne Berücksichtigung der für den Patienten oder die Patientin spürbaren Symptomatik. Die Komponente „radiologische Krankheitsprogression“ wird daher als nicht unmittelbar patientenrelevant eingeschätzt und das PFS in der vorliegenden Operationalisierung als nicht patientenrelevant bewertet.

Die Sensitivitätsanalyse unter Einbezug der klinischen Krankheitsprogression kann in der vorliegenden Operationalisierung nicht als patientenrelevant bewertet werden, da diese die radiologische Progression als konkurrierendes Ereignis beinhaltet.

Validität

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Schmerzen mittels Brief Pain Inventory – Short Form

Der Endpunkt „Brief Pain Inventory – Short Form“ (BPI-SF) wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

In der Studie HERIZON_BTC wurde durch den BPI-SF die Intensität der Schmerzen und das Ausmaß der dadurch erlebten Beeinträchtigung im täglichen Leben erfasst. Der Fragebogen umfasst 15 Items. 11 Items werden den Domänen Schmerzintensität (4 Items) und Beeinträchtigung durch den Schmerz (7 Items; Items 9a bis 9g) zugeordnet.

Schmerzintensität

- Item 3: stärkste Schmerzen
- Item 4: geringste Schmerzen
- Item 5: durchschnittliche Schmerzen
- Item 6: Schmerzen im Moment

Beeinträchtigungen durch Schmerzen

- Item 9a: Allgemeine Aktivität
- Item 9b: Stimmung
- Item 9c: Gehvermögen
- Item 9d: Normale Arbeit
- Item 9e: Beziehung zu anderen Menschen
- Item 9f: Schlaf
- Item 9g: Lebensfreude

Die Items werden auf einer Skala von 0 (keine Schmerzen bzw. keine Beeinträchtigung) bis 10 (schlimmste vorstellbare Schmerzen bzw. stärkste Beeinträchtigung) beantwortet. Der Bezugszeitraum umfasst die letzten 24 Stunden. Es werden Mittelwerte der enthaltenen Domänen berechnet. Die Bildung eines Gesamtwerts für den gesamten Fragebogen ist nicht vorgesehen. Jedoch können Gesamtwerte für die Domänen berechnet werden.

Ausführungen zu den statistischen Auswertungen finden sich in Kapitel 2.4.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz und Validität

Schmerzen und die dadurch verursachten Beeinträchtigungen des täglichen Lebens sind patientenrelevant. Die Validität des Messinstruments BPI-SF wurde in vorherigen Nutzenbewertungen bereits bewertet und wird im vorliegenden Anwendungsgebiet als valide eingeschätzt. Es muss beachtet werden, dass der Bezugszeitraum von 24 Stunden vor der Studienvisite nur Aussagen für einen geringen Anteil der Studienzeit erlaubt.

Gesundheitszustand mittels EQ-5D-VAS

Der Endpunkt „EQ-5D-VAS“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die visuelle Analogskala des European Quality of Life 5-Dimension (EQ-5D-VAS) erfasst die Selbsteinschätzung des allgemeinen Gesundheitszustands. Die Studienteilnehmenden bewerteten den Gesundheitszustand auf einer vertikalen Skala, deren Punktwerte von 100 („Denkbar bester Gesundheitszustand“) bis 0 („Denkbar schlechtester Gesundheitszustand“) bezeichnet sind.

Ausführungen zu den statistischen Auswertungen finden sich in Kapitel 2.4.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

Validität

Die EQ-5D-VAS wird als valide und reliabel zur Einschätzung des allgemeinen Gesundheitszustands angesehen.

2.3.3 Lebensqualität

Es liegen keine Endpunkte der Kategorie „Lebensqualität“ vor.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt „Unerwünschte Ereignisse“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

In der Studie HERIZON_BTC wurde ein Unerwünschtes Ereignis (UE) als jegliches medizinisches Ereignis bei einer Person in einer klinischen Studie definiert, die die Intervention erhalten hat, unabhängig von einem kausalen Zusammenhang mit dieser Intervention. Dabei wurden die Kriterien der ICH E2A Guideline sowie der 21 Code of Federal Regulations 312.32 IND Safety Reporting herangezogen.

Die Klassifizierung der UE erfolgte anhand der Medical-Dictionary-for-Regulatory-Activities-(MedDRA-)Version 25 nach Systemorganklasse (SOC) und Preferred Term (PT). Der Schweregrad wurde mittels Common-Terminology-Criteria-for-Adverse-Events-(CTCAE-)Grad festgestellt.

UE wurden bei jeder Visite ab Zyklus 1 Tag 1 bewertet und während der gesamten Studie überwacht und 30 Tage nach der letzten Gabe der Studienmedikation erfasst. Jegliche Veränderung des medizinischen Zustands (bspw. Veränderung der Schwere, Häufigkeit oder Charakter) sollten während der Sicherheitserhebung erhoben werden. Abnormale Laborparameter sollten nicht direkt als UE gewertet werden, außer sie stehen in einem Zusammenhang mit einem klinischen Zeichen, erfordern eine Behandlung, resultieren in einem schwerwiegenden unerwünschten Ereignis (SUE) oder in einer Beendigung der Studie. Bevorzugt sollte das klinische Symptom bei einem abnormalen Laborparameter erhoben werden (bspw. „Anämie“ anstelle von „geringer Hämoglobin-Wert“).

Schwerwiegende unerwünschte Ereignisse (SUE) sind Ereignisse mit folgender Definition:

- fatal oder lebensbedrohlich;
- führen zu persistierender oder signifikanter Behinderung;
- führen zu einer Anomalie/einem Geburtsdefekt;
- medizinisch signifikant, d. h. ein eine Person gefährdendes Ereignis oder ein Ereignis, das einer medizinischen oder chirurgischen Intervention zur Verhinderung der oben aufgelisteten Folgen bedarf;
- Ereignisse, die zu Hospitalisierungen oder zur Verlängerung einer bestehenden Hospitalisierung führen (jegliche geplante Hospitalisierung im Studienkontext wurde nicht als SUE gewertet);
- wichtiges medizinisches Ereignis, welches nach Einschätzung des Prüfpersonals, die Person gefährdet oder eine chirurgische Intervention erfordert, um eines der oben genannten Outcomes zu verhindern.

Die Erhebung der (S)UE beginnt mit Zyklus 1 Tag 1 bis 30 Tage nach der letzten Gabe der Studienmedikation. Alle SUE sollten zusätzlich ab dem Zeitpunkt der Einverständniserklärung erhoben werden. Sollte ein SUE, welches im Zusammenhang mit der Intervention gesehen wird, nach Ende der Sicherheitsperiode auftreten, sollte dieses ebenfalls erhoben werden.

Alle UE (einschließlich SUE) wurden bis zur Rückkehr auf den Ausgangszustand, Stabilisierung, fehlenden klinischen Relevanz oder bis zum Tod, Verlust der Nachverfolgbarkeit oder Widerruf der Einwilligung nachverfolgt.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen. Die Patientenrelevanz von Laborparametern ist unklar.

Validität

Der Endpunkt wurde valide erhoben.

2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 5.

Tabelle 5: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie HERIZON_BTC

Studienvisite (Woche) Endpunkt	Baseline	Tag 28 nach Baseline	8-Wochen-Zyklus Tag 28	Ende der Behandlung	Follow-up
Bestes objektives Ansprechen ¹⁾	x	x	X	x	
Gesamtüberleben	Kontinuierlich				x ²⁾
BPI-SF	x	x	x	x	
EQ-5D-VAS	x	x	x	x	
Unerwünschte Ereignisse	Kontinuierlich				x ³⁾

¹⁾ Der primäre Endpunkt wird aus Transparenzgründen im Anhang dargestellt.

²⁾ Erhebung des Gesamtüberlebens alle drei Monate telefonisch bis zum Tode, Lost to Follow-up, Rücknahme der Einverständniserklärung, Studienabschluss oder Abbruch der Studie durch den Sponsor.

³⁾ Erhebung einmalig 30 Tage nach Ende der Behandlung.

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; VAS: Visuelle Analogskala.

2.4 Statistische Methoden

Für die Nutzenbewertung liegen Zusatzanalysen aus dem Studienbericht vor, die die zulassungsrelevante Auswertungspopulation berücksichtigen.

Analysepopulationen

- Efficacy Analysis Set: Alle Personen, die die Einverständniserklärung unterschrieben haben, die für die Studie geeignet gewesen und die in die Studie aufgenommen worden sind. Dies entspricht dem Intention-to-Treat-(ITT-)Prinzip. Die Analysepopulation bezieht sich auf die Personen mit einem "IHC 3+"-Status.
- Sicherheitspopulation: Alle Personen, die mindestens eine Gabe der Studienmedikation in jeglicher Dosierung erhalten haben.

Datenschnitte

Es konnten 3 Datenschnitte identifiziert werden. Der erste Datenschnitt erfolgte präspezifiziert, nachdem die letzte eingeschlossene Person für 6 Monate nachbeobachtet worden ist. Für den zweiten Datenschnitt konnte keine weitere Begründung identifiziert werden, außer, dass eine längere Nachbeobachtung vorliegt. Dieser Datenschnitt wurde im EPAR berücksichtigt.

Für die Nutzenbewertung relevant ist der finale Datenschnitt vom 28. Juli 2024. Zum Zeitpunkt der Nutzenbewertung befindet sich keine Person mehr in der Nachbeobachtung zum Gesamtüberleben, da die Studie durch den Sponsor beendet worden ist.

Präspezifizierung geplanter und durchgeführter Analysen

Die für die Nutzenbewertung relevanten Auswertungen beziehen sich auf die relevante Teilpopulation der Personen mit "HER2 IHC 3+"-Status. Dafür werden Zusatzanalysen zu den patientenberichteten Instrumenten vorgelegt, die den methodischen Anforderungen der Nutzenbewertung entsprechen.

Fehlende Werte und Daten-Imputation

Es erfolgte keine Daten-Imputation bei fehlenden Werten im BPI-SF und EQ-5D-VAS.

Statistische Auswertung der patientenberichteten Instrumente

BPI-SF

Es liegen folgende post-hoc-Responderauswertungen für die relevante Teilpopulation vor:

- Stärkster Schmerz in den letzten 24 Stunden (Item 3);
- Beeinträchtigung durch Schmerz (Items 9a bis 9g).

Es liegen Responderanalysen und Auswertungen zur mittleren Veränderung für die relevante Teilpopulation als Zusatzanalysen vor. Die Auswertungen beziehen sich auf die Zeitpunkte zur Visite, zum Ende der Behandlung, zum Zeitpunkt des besten objektiven Ansprechens und der niedrigste beobachtete Wert nach Baseline. Die Responderauswertungen liegen zur Verbesserung sowie zur Verschlechterung vor und berücksichtigen die relevante Responsechwelle von 15 %. Für Item 3 wurde ein Responsekriterium von 2 Punkten festgelegt.

Weiterhin liegen post-hoc-Responderauswertungen und MW-Änderungen für die zusammengefassten Items der Schmerzintensität (Items 3 bis 6) in den Zusatzanalysen vor. Für diese liegen Responderauswertungen mit einem Responsekriterium von 1,5 Punkten vor, welches 15% der Skalenspannweite entspricht.

EQ-5D-VAS

Es liegen Responderanalysen und Auswertungen zur mittleren Veränderung für die relevante Teilpopulation vor. Die Auswertungen beziehen sich auf die Zeitpunkte zur Visite, zum Ende der Behandlung, zum Zeitpunkt des besten objektiven Ansprechens und des niedrigsten beobachteten Wertes nach Baseline. Die Responderauswertungen liegen zur Verbesserung sowie zur Verschlechterung vor und berücksichtigen die relevante Responsechwelle von 15 %.

Einschätzung der statistischen Auswertungen

Aufgrund der palliativen Therapiesituation werden Responderanalysen zur Verschlechterung als relevant angesehen und für die Nutzenbewertung herangezogen. Als bewertungsrelevanter Auswertungszeitpunkt wird die Visite zum Ende der Behandlung (EOT) herangezogen.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Da es sich bei der Studie HERIZON_BTC um eine unkontrollierte Studie handelte, wird von einem hohen Verzerrungspotenzial ausgegangen.

2.6 Indirekter Vergleich

Im Rahmen des Dossiers zur Nutzenbewertung wurde ein indirekter Vergleich ohne Brückenkompator unter Einbezug von individuellen Patientendaten durchgeführt. Es wurden Ergebnisse zu den Endpunkten „Gesamtüberleben“ und „Progressionsfreies Überleben (PFS)“ eingereicht. Da der Endpunkt „PFS“ nicht für die Nutzenbewertung herangezogen wird (siehe Abschnitt 2.3.2), werden ausschließlich die Ergebnisse zum „Gesamtüberleben“ bewertet.

Als Informationsquellen zur Durchführung des Indirekten Vergleiches wurde ein Studienprotokoll, ein Studienbericht sowie Appendizes identifiziert. Als bewertungsgegenständlich wird der indirekte Vergleich gegenüber der Zulassungspopulation, die HER2-positiv mit IHC 3+ war, angesehen.

Zuerst werden die vorliegende Datenquelle sowie die Auswahl der Vergleichspopulation beschrieben und danach eingeordnet und bewertet.

Datenquellen

Im Studienprotokoll wurden 2 Datenquellen identifiziert: die Flatiron Health Research Analytic Database (FHRD) und das European Network for the Study of Cholangiocarcinoma (ENSCCA). Beim ENSCCA handelt es sich um ein Register, welches Personen aus 11 europäischen Ländern in 26 partizipierenden Krankenhäusern in einer Beobachtungsstudie nachverfolgt. Daten zu diesem indirekten Vergleich konnten in den eingereichten Unterlagen nicht identifiziert werden. Es konnte ebenfalls keine Begründung identifiziert werden, weshalb der indirekte Vergleich gegenüber den ENSCCA-Daten nicht durchgeführt wurde und nur Daten gegenüber der FHRD zur Nutzenbewertung eingereicht wurden.

Bei der FHRD handelt es sich um von einem Gesundheitsdatendienstleister gesammelte retrospektive Beobachtungsdaten, die durch Anwendung von fortgeschrittenen statistischen Methoden (u. a. Machine Learning) und medizinischer Expertise extrahiert, aggregiert und dargestellt werden. Als vorliegende Datenbasis dienen elektronische Fallakten aus den USA, welche in der stationären Regelversorgung und im akademischen Setting erhoben worden sind. Zur Identifizierung möglicher geeigneter Fallakten wurde ein Natural Language Processing-based Model (NLP) herangezogen. Dabei handelt es sich um ein Subset eines KI-Modells, welches aus Texten Zusammenhänge unter Einbezug von großen Sprachmodellen (Large Language Model und Machine Learning) erkennen kann. Es wurden insgesamt 29.000 elektronische Fallakten identifiziert, die manuell von Personen mit medizinischer Expertise hinsichtlich der Diagnosestellung, Evidenz der HER2-Testung und Initiierung einer Zweitlinientherapie innerhalb der für den indirekten Vergleich operationalisierten Studienperiode (2011-2023) überprüft worden sind. Nach der manuellen Überprüfung (chart-review) wurden 290 Personen identifiziert, die die genannten Kriterien erfüllen.

Innerhalb der FHRD werden strukturiert Daten zu der Demographie, Behandlungssetting, Visite, ausgewählte Labor-/Vitalwerte, Versichertenstatus sowie weitere soziale Determinanten erfasst. Weiterhin werden unstrukturierte Daten erfasst, die nicht regelhaft über die elektronische Fallakte erhoben werden. Die unstrukturierten Daten werden durch das auf Machine-Learning-(ML-)Methoden basierende NLP-Modell erfasst. Als unstrukturierte Daten wurden insbesondere BTC-krankheitsspezifische Charakteristiken operationalisiert, wie der Zeitpunkt der Diagnose, Krankheitsstadium, anatomische Lokalisation, Therapien, HER2-Biomarker Details, Progression und Ansprechen, weitere Malignitäten und Komorbiditäten sowie Metastasen.

In der FHRD werden inhärente Bestandteile der Baseline-Charakteristiken (bspw. Therapielinie, ECOG PS, Krankheitsstadium) synthetisch ermittelt. Die herangezogene Therapielinie wird mittels eines „Therapielinien-selektionsalgorithmus“ hergeleitet. Beim ECOG PS werden die ML- und NLP-Methoden herangezogen, um die Präsenz von möglichen Einschränkungen des Gesundheitszustands zu schätzen. Dies wurde so operationalisiert, weil das ML-NLP-Modell die Abwesenheit von Einschränkungen (bspw. keine Einschränkungen = ECOG PS von 0) nicht erkennen kann. Für die Ermittlung des Krankheitsstadiums konnten keine näheren Informationen identifiziert werden, außer dass dieses aus unstrukturierten Daten hergeleitet wird.

Für die Studie HERIZON_BTC wurde der finale Datenschnitt mit der längsten Nachbeobachtungsdauer vom 28.07.2024 herangezogen.

Bewertung der Datenquelle

Der pU zieht als Datenquelle die FHRD heran. Es konnte keine Begründung für die fehlende Geeignetheit der ENSCCA-Quelle identifiziert werden. Außerdem konnte nicht nachvollzogen werden, inwiefern systematisch nach möglichen Datenquellen im Anwendungsgebiet gesucht worden ist. Eine ergebnisgesteuerte Auswahl der Datenquelle kann daher nicht ausgeschlossen werden.

Zur Erstellung der FHRD-Datenbasis werden fortgeschrittene statistische Methoden (ML, NLP) herangezogen, wobei es sich um ein KI-/ML-basiertes Modell handelt, das strukturierte und unstrukturierte Daten aus elektronischen Fallakten verarbeitet und daraus relevante Variablen extrahiert bzw. ableitet. Dazu liegen Studien zur Validierung der Vorgehensweise vor [1]. Inhärenter Bestandteil von ML- und NLP-Methoden ist die anwendungsspezifische Validierung anhand von Trainingsdatensätzen. Anhand der eingereichten Daten konnte nicht geprüft werden, inwiefern die Modelle geeignet sind, auch unstrukturierte BTC-krankheitsspezifische Informationen korrekt bei der Auswahl der Vergleichspopulation zu berücksichtigen. Es kann auf Basis der vorliegenden Unterlagen nicht davon ausgegangen werden, dass die herangezogenen statistischen Methoden und die manuelle Überprüfung einzelner Variablen (bspw. Krankheitsstadium, Laborparameter), vollständig und hinreichend vergleichbare BTC-krankheitsspezifische Informationen (im Vergleich zur Erhebung in der klinischen Studienkohorte) identifizieren können.

Die BTC-spezifischen Baselinewerte, die mittels ML-NLP-Methoden aus unstrukturierten Daten hergeleitet worden sind, beschreiben möglicherweise nicht den Zustand der Person zum Index-Zeitpunkt, während davon ausgegangen werden kann, dass die strukturierten Daten den Zustand zum Index-Zeitpunkt adäquat beschreiben können.

Der pU verwendet den finalen Datenschnitt der klinischen Studie. Es muss berücksichtigt werden, dass 10 Personen (16,1 %) zum Zeitpunkt der vorzeitigen Beendigung der Studie zensiert worden sind.

Auswahl der Vergleichspopulation

In die Grundpopulation gingen 290 Personen ein.

Die hinreichende Erfüllung von Ein- und Ausschlusskriterien, die im Folgenden beschrieben werden, waren Bedingung für die Aufnahme in die Vergleichspopulation. Diese orientierten sich an denen der Zulassungsstudie HERIZON_BTC. Es wurden 3 Stufen der Ein- und Ausschlusskriterien operationalisiert.

In die 1. Stufe gehen die Personen ein, die eine BTC-Diagnose mit HER2-positiver Überexpression (IHC 3+) sowie eine systemische Zweitlinientherapie erhalten haben. Ausgeschlossen wurden in der 2. Stufe Personen mit NLP-hergeleiteten ECOG PS > 1 sowie Personen mit Hirnmetastasen. Die 2. Stufe wurde vom pU für den indirekten Vergleich herangezogen. In einer weiteren 3. Stufe wurden weitere Ein- und Ausschlusskriterien herangezogen, sofern diese machbar sind.

Von initial 290 Personen insgesamt hatten 196 (67,6 %) ein HER2-negatives Ergebnisse vor Initiierung der Zweitlinientherapie, weitere 20 (6,9 %) einen IHC-Status von 2+ ohne "ISH+"-Validierung vorliegen. Ein weiterer Ausschlussgrund war ein unbekannter BTC-Subtyp bei 30 Personen (10,3 %).

Durch das Vorgehen konnten insgesamt 29 Personen mit einem HER2-positivem IHC-Status von 3+ für die Stufe 1 identifiziert werden. Davon erhielten 12 Personen sowohl in der Erst-, als auch in der Zweitlinientherapie eine Chemotherapie. Weitere 13 Personen erhielten eine HER2-gerichtete Therapie und jeweils 2 Personen eine Immuntherapie oder sonstige Therapie. Sensitivitätsanalysen, die eine adäquate hämatologische oder hepatologische Funktion berücksichtigen (orientiert an den Einschlusskriterien der Zulassungsstudie HERIZON_BTC) führen zu noch geringeren Fallzahlen (n = 8 bei adäquater Leberfunktion und n = 4 bei adäquater hämatologischer Funktion) im Vergleich gegenüber Chemotherapie.

Als externen Kontrollarm zieht der pU den Vergleich der in Stufe 2 selektieren Personen gegenüber Chemotherapie (n = 12) heran. Die verabreichten Therapieregime sind im Anhang dargestellt (Tabelle 17). Für die Personen (n = 12), die Chemotherapie erhalten haben, liegen im Mittel (SD) 10 (12,9) Monate an Nachbeobachtungszeit vor. Dabei ist die Zeit seit Diagnose bis zum Indexdatum im Mittel (SD) 18 (25,5) Monate.

Der Beobachtungszeitraum für die Personen des externen Kontrollarms. beginnt mit der Initiierung der Zweitlinientherapie. Als Zweitlinientherapie wurden mittels des Flatiron-Therapielinienalgorithmus jegliche zweite systemische Therapie definiert, die nach einer ersten systemischen Therapie gegeben worden ist und sich von dieser unterscheidet. Die Therapielinie verändert sich, wenn ein prädefinierter zeitlicher Abstand zwischen der Gabe zwischen zwei Therapieregimen vergangen ist oder eine neue Therapie initiiert worden ist, die kein Bestandteil der vorherigen Therapielinie war. Auch konnte im Gegensatz zur klinischen Studie die Bedingung, dass eine Gemcitabin-haltige Chemotherapie gegeben worden sein musste, nicht operationalisiert werden.

Erhebung von Endpunkten

Das Gesamtüberleben wurde im externen Kontrollarm mittels der elektronischen Fallakte, Sozialversicherungsdaten oder Todesanzeigen hergeleitet und wurde ab dem Zeitpunkt der Initiierung der Zweitlinientherapie erhoben.

Bewertung der verwendeten Vergleichspopulation

Der Selektionsmechanismus aus der Grundpopulation wird beschrieben (von 29.000 identifizierten Personen gehen 12 in den externen Vergleichsarm ein), jedoch kann dieser nicht vollständig nachvollzogen werden. Es kann nicht ausgeschlossen werden, dass fehlende Daten (bspw. durch Einschluss in eine klinische Studie oder Behandlung in einem anderen Setting) Einfluss auf die Datenverfügbarkeit haben und in der Folge den Selektionsmechanismus beeinflussen.

Die Ein- und Ausschlusskriterien wurden in mehreren Stufen operationalisiert. Diese sind an die klinische Studie HERIZON_BTC angelehnt, jedoch konnte eine Vielzahl der Ein- und Ausschlusskriterien nicht operationalisiert oder hergeleitet werden. Es können relevante klinische Parameter zur Therapiezuweisung (bspw. Laborwerte, Infektionen) nicht hinreichend abgebildet werden. Die herangezogenen Ein- und Ausschlusskriterien ECOG PS und Hirnmetastasen, konnten weiterhin nicht vollständig hergeleitet werden, sodass auch hier z. T. fehlende Informationen vorliegen. Die Operationalisierung der adäquaten Leberfunktion als Einschlusskriterium führte zu einem weiteren Fallzahlverlust, sodass davon ausgegangen werden muss, dass 25 % der Personen im externen Kontrollarm anhand der Leberfunktion nicht für die Teilnahme an der klinischen Studie geeignet gewesen wären.

Zusätzlich zu den bereits genannten nicht operationalisierbaren Ein- und Ausschlusskriterien zählen u. a. operative Prozeduren (u. a. biliäre Drainage, laparoskopische Prozeduren) zur Behandlung von potenziellen Komorbiditäten der Grunderkrankungen, die zum biliärem Karzinom führen können (bspw. primäre sklerosierende Cholangitis). Die Komplikationen müssen seit mindestens 2 Wochen vor der ersten Dosis abgeklungen sein und größere Operationen müssen mindestens 4 Wochen vor Studieneintritt stattgefunden haben. Auch durch das Propensity-Score-Modell konnten die fehlenden Ein- und Ausschlusskriterien durch die Confounder-Auswahl nicht berücksichtigt werden.

Während in der klinischen Studie radiologische Erhebungen zu unterschiedlichen Zeitpunkten (bspw. zum Zeitpunkt der Therapiezuweisung) vorgenommen wurden, konnte diese im Rahmen der Regelversorgung nicht durchgeführt werden bzw. nicht durch das ML-NLP-Modell identifiziert werden. Weiterhin können die unstrukturierten Daten der FHRD das Krankheitsstadium zum Zeitpunkt der Therapiezuweisung nicht abbilden, sodass auf Basis der vorliegenden Unterlagen ein Vergleich der Krankheitsstadien zum Zeitpunkt der Therapiezuweisung nicht sinnvoll möglich erscheint. Dies kann auf unterschiedliche Index-Zeitpunkte/Time-Zeros hindeuten.

Weiterhin bestehen Unklarheiten hinsichtlich der Testverfahren zur HER2-Testung. Das Testverfahren für die FHRD-Kohorte unterscheidet sich von dem in der klinischen Studie, bei der der HER2-Status zum Studieneinschluss mittels Zentrallabor zum Studieneinschluss nachgewiesen worden sein musste. Im externen Kontrollarm wurde der HER2-Status mittels unstrukturierter Daten erhoben. Es lagen im Mittel (SD) 11 Monate (9) zwischen erster HER2-Testung und dem Index-Zeitpunkt. Es ist unklar, ob die Therapiezuweisung zur Zweitlinientherapie auf Basis dieser HER2-Testung erfolgte. Auf Basis der vorliegenden Unterlagen kann nicht ausgeschlossen werden, dass eine Diskordanz zwischen Testungen vorlag und es sich bei der Vergleichspopulation um Personen handelt, die falsch-positiv in ihrem Behandlungspfad getestet worden sind. Fehlende Bestätigung einer positiven HER2-Überexpression war einer der häufigsten Gründe für eine Nichtaufnahme in die klinische Studie HERIZON_BTC während des Screenings zur Aufnahme. Die Selektionsmechanismen führen dazu, dass der indirekte Vergleich gegenüber 12 Patientinnen und Patienten durchgeführt wurde.

Es scheinen regionsspezifische Unterschiede in den potenziellen Outcomes vorzuliegen, die im indirekten Vergleich nicht adjustiert werden konnten. Die Region wurde von den Expertinnen und Experten als relevanter Confounder genannt und es finden sich Unterschiede im Gesamtüberleben in einer Subgruppenanalyse der Studie HERIZON_BTC. Ein Großteil der zulassungsrelevanten Teilpopulation besteht aus asiatischen Personen, während der externe Vergleichsarm aus amerikanischen Personen besteht.

Die statistische Positivität der Vergleichspopulation und der klinischen Studie HERIZON_BTC (d. h. die Möglichkeit, dass beide Patientenpopulationen (theoretisch) gleichermaßen für die Therapien bzw. beide Studienpopulationen in Frage kommen und keine Gründe wie bspw. Kontraindikationen gegen eine entsprechende Zuteilung sprechen) ist insgesamt auf Basis der eingereichten Unterlagen nicht ausreichend dargelegt.

Identifikation, Auswahl und Erhebung von Confoundern

Literaturrecherche und Auswahl

Der pU hat eine systematische Literaturrecherche (SLR) nach Confoundern im Anwendungsgebiet durchgeführt. In die SLR wurden insbesondere retrospektive Beobachtungsstudien ab dem Jahr 2013 sowie vereinzelte klinische Studien und prospektive Beobachtungsstudien eingeschlossen. Es wurden 64 Studien identifiziert, die Prognosefaktoren zum Gesamtüberleben enthielten sowie 17 weitere Studien, die die Therapiezuweisung bewerten. Eine Tabelle der identifizierten Confounder sowie deren Relevanzbewertung ist im Anhang dargestellt (Tabelle 18).

Folgende Confounder wurden für den indirekten Vergleich für das Propensity-Score-(PS-)Modell herangezogen:

- Alter bei Initiierung der Zweitlinientherapie;
- Tumorentität;
- Vorerkrankung mit einer chronischen Lebererkrankung;
- Geschlecht.

Bewertung der verwendeten Confounder

Die Durchführung einer SLR mit anschließender Einordnung durch Expertinnen und Experten wird als sachgerecht bewertet. Es bestehen Unsicherheiten bei der Nachvollziehbarkeit der durchgeführten Bewertung der Relevanz der Confounder. Inwiefern eine Zusammenführung von inhaltlich überlappenden Confounder-Variablen durchgeführt worden ist, konnte auf Basis der vorliegenden Unterlagen nicht nachvollzogen werden.

Eine Vielzahl an Confoundern, die als relevant identifiziert und bewertet worden sind (siehe Tabelle 18) wurden nicht in das PS-Modell eingeschlossen. Ob die Relevanz aller durch die SLR identifizierten Confounder bewertet worden ist, konnte anhand der Unterlagen nicht nachvollzogen werden. Es wurden nicht die Confounder, die für das Propensity-Score-Modell von Relevanz sind, herangezogen, sondern die mit einer vorliegenden Datenverfügbarkeit gehen in das PS-Modell ein. Die Auswahl der Confounder wird als nicht sachgerecht bewertet.

Statistische Methoden für den indirekten Vergleich

Eine statistische Adjustierung kann die genannten Limitationen (unklare Time-Zero, inadäquate Operationalisierung der Ein- und Ausschlusskriterien, nicht sachgerechte Confounder-Auswahl) zur Vergleichbarkeit der beiden Populationen nicht beheben und die Ableitung eines Behandlungseffekts erscheint nicht sachgerecht. Im Folgenden wird die statistische Methodik hinter dem indirekten Vergleich dennoch beschrieben und anschließend eingeordnet.

Schätzung des PS-Modells

Zur Erstellung des PS-Modells wurde ein iterativer Prozess durchgeführt. Dabei wurden PS-Modelle auf Basis einer Liste an identifizierten möglichen Confoundern mittels logistischer Regression geschätzt. Sollte das Modell nicht konvergieren, sollte die Operationalisierung der Variable geändert werden oder Variablen entfernt werden, die eine geringe Priorität aufwiesen. Weiterhin wurde eine „Back-up“-Operationalisierung implementiert, sollte das PS-Modell auf Basis der ersten Operationalisierung nicht konvergieren. Im nächsten Schritt sollte die Modellgüte mittels des Gütemaßes zur Balanciertheit der standardisierten Mittelwertsdifferenz (SMD) ($< 0,10$) und der Verteilung der Gewichte und Propensity Scores (PS) bewertet werden. Die Bewertung der Modellgüte wurde herangezogen, um zu beurteilen, ob Anpassungen am PS-Modell notwendig sind. Die Anpassung des PS-Modells sollte unter Einbezug von Expertinnen und Experten durchgeführt werden.

Schätzung eines Behandlungseffektes

Basierend auf den Propensity Scores wurde das Standardized Morbidity Ratio (SMR) berechnet, welches als Gewichtungsmethodik herangezogen wurde. Dazu wird der externe Kontrollarm neu gewichtet, um die Studienkohorte zu repräsentieren. Somit spiegelt der berechnete Propensity Score die Wahrscheinlichkeit wider, eine Person in der klinischen Studie HERIZON_BTC zu sein. Personen der klinischen Studie erhalten ein Gewicht von 1, während das proportionale Gewicht für Personen im Kontrollarm auf dem geschätzten Propensity Score basiert. Durch dieses Verfahren wird die effektive Fallzahl der externen Kontrollgruppe nach Gewichtung der tatsächlichen Fallzahl der klinischen Studie ähneln („Pseudopopulation“).

Zur Berechnung eines Behandlungseffektes wurde eine Cox-Regression durchgeführt und die Annahme der Proportional Hazards sollte laut Studienprotokoll mittels Schoenfeld-Residuen getestet werden. Nachdem die PS-Adjustierung mittels SMR durchgeführt worden ist, wurde das gewichtete Cox-PH-Modell mit dem Behandlungsarm als einzige Kovariate geschätzt. Zur Berechnung der Konfidenzintervalle wurde die Bootstrap-Methode herangezogen. Weiterhin wurde der Unterschied im 12-Monatsüberleben berechnet.

Als Behandlungseffekt wurde der „Average Treatment Effect on the treated“ (ATT) geschätzt.

Modelldiagnostik des PS-Modells mittels SMR-Weights

Zur Bewertung der Güte des Modells wurden die SMD als Maß der Balanciertheit, die geschätzten Gewichte der Kontrollpopulation und die Verteilung der Propensity Scores graphisch dargestellt. Beim eingereichten PS-Modell lagen bei 3 der 4 herangezogenen Confoundern eine SMD von $< 0,10$ vor. Beim Confounder „Vorerkrankung mit einer chronischen Lebererkrankung“ lag eine höhere SMD vor. Laut Studienprotokoll zur Durchführung des indirekten Vergleiches sollten SMR-gewichtete Charakteristiken sowie die dazugehörige SMD berichtet werden. Diese konnten im Studienbericht nicht identifiziert werden.

Das mittlere Gewicht im externen Kontrollarm liegt bei 5,18. Außerdem sind vereinzelte extreme Gewichte (11,84 und 18,84) aufgetreten.

Die Darstellung der Propensity Scores erfolgte durch Histogramme. Dies erschwert die graphische Überprüfung. Die PS-Verteilung deutet auf leicht divergierende Verteilungen vor PS-Adjustierung hin. Nach PS-Adjustierung scheint eine graphische Überlappung der Propensity Scores vorzuliegen.

Bewertung der Confounder-Adjustierung und der statistischen Methoden

Es kann auf Basis der vorliegenden Informationen nicht davon ausgegangen werden, dass es sich beim vorliegenden PS-Modell um ein Modell handelt, welches alle oder hinreichend viele Confounder mit Einfluss auf das Outcome oder Therapiezuweisung enthält (Strong-Ignorable-Treatment-Assignment-(SITA-)Bedingung).

Die verwendeten BTC-krankheitsspezifischen Confounder sind krankheitsbeschreibend. Aufgrund fehlender Daten bzw. der geringen Fallzahl konnte der BTC-krankheitsspezifische Confounder Tumorentität nicht kategoriell operationalisiert werden, sondern musste unter Zusammenlegung zweier Kategorien (Cholangiokarzinom) mit Unterschieden in der Prognose als binäre Variable ins Modell eingehen. Dies geht mit einem Informationsverlust einher. Um eine Vergleichbarkeit der Kohorten zu gewährleisten, hätten weitere Confounder, die beispielsweise die Krankheitsdynamik beschreiben, berücksichtigt werden müssen.

Als Adjustierungsmethode wurde SMR-Gewichtung verwendet mit einem ATT-Schätzer als „Target of Inference“. SMR-Analysen erzeugen eine Pseudopopulation, mit dem Ziel potenzielles Confounding aufzuheben. Die Gewichte für die Analyse mittels SMR basieren auf Propensity Scores, die aufgrund der Unvollständigkeit des PS-Modells nicht sinnvoll interpretierbar sind. Aufgrund der geringen ursprünglichen Fallzahl bekommen einzelne Beobachtungen ein sehr hohes Gewicht, welches trotz der wenigen verwendeten Confounder auf eine Positivitätsverletzung hindeuten kann.

Der pU konnte lediglich eine Fallzahl von 12 Personen für den indirekten Vergleich heranziehen. Dies erscheint bei der Anwendung von Propensity-Score-Methoden unter Einbezug von Gewichtungsverfahren als zu wenig. Die Ableitung eines Behandlungseffektes mittels Cox-Regression als semi-parametrisches Verfahren erscheint vor dem Hintergrund der wenigen aufgetretenen Ereignisse (durch die geringe Fallzahl) im Kontrollarm als nicht sachgerecht. Dabei ist unklar, ob es sich um einen vergleichbaren Beobachtungsbeginn (Index-Datum/Time Zero) der Kohorten handelt. Die Reduzierung des Beobachtungsbeginns auf den Zeitpunkt der Initiierung der Zweitlinientherapie ohne Berücksichtigung von Eignungskriterien zum Zeitpunkt der Therapieinitiierung, erscheint in der Gesamtschau nicht sachgerecht.

Die genannten Limitationen zur Erfüllung der Grundbedingungen (Fallzahl, Vollständigkeit der Confounder, Positivität, „Time Zero“) führen dazu, dass der berechnete Effektschätzer auf Basis eines ATT-Schätzers für Aussagen zur Wirksamkeit nicht sinnvoll interpretierbar ist. Dieser bezieht sich nicht mehr auf die gesamte hergeleitete Vergleichspopulation, sondern auf eine konstruierte Population, die nicht klar beschreibbar ist. Es ist unklar, ob der ATT-Ansatz für den Rahmen einer Nutzenbewertung geeignet ist. Dieser gewichtet die herangezogene Vergleichspopulation der externen Kontrolle so, als ob diese für die Intervention geeignet gewesen wäre. Der geschätzte Effekt bezieht sich auf die Personen, die tatsächlich die Intervention erhalten haben und nicht auf alle Personen, die dafür geeignet gewesen wären (Zielpopulation).

Anhand der vorgelegten Unterlagen kann nicht abschließend beurteilt werden, inwiefern sich die erzeugte Pseudopopulation von der Gesamtpopulation unterscheidet und ob sich die Ergebnisse für die Pseudopopulation auf die Gesamtpopulation übertragen lassen. Als weitere Limitation wird die geringe Stichprobengröße der externen Kontrolle gesehen, welche die Validität von Propensity-Score-Verfahren mittels Gewichtung einschränkt.

Die herangezogenen statistischen Methoden zur Berechnung der Propensity Scores erfordern z. T. sehr starke Annahmen an die Datensituation. Die fehlende Berücksichtigung von relevanten Confoundern und deren eingeschränkte Vergleichbarkeit bezüglich „Time Zero“ sowie nicht dargelegte Positivität können nicht durch komplexe statistische Methoden mit unbegründeten Entscheidungsstrukturen zur Modellwahl aufgewogen werden. Auf Basis der vorliegenden Unterlagen führen die eingereichten Analysen zu dem Modell, welches am besten zu der vorliegenden Datensituation und der herangezogenen Entscheidungsstruktur passt. Eine datengetriebene Modellauswahl kann auf Basis der vorliegenden Unterlagen nicht ausgeschlossen werden. Im Kontext der Nutzenbewertung ist ein verallgemeinerbares Modell von Relevanz.

2.6.1 Zusammenfassende Bewertung der Methodik des indirekten Vergleichs

Der pU legt eine Auswertung auf Basis einer US-Datenbank vor, welche durch Anwendung von ML-NLP-Methoden und medizinischer Expertise erstellt worden ist. Von insgesamt 29.000 identifizierten elektronischen Fallakten gingen 12 Patientinnen und Patienten in den externen Kontrollarm ein. Die Ein- und Ausschlusskriterien der Studie HERIZON_BTC konnten für die externe Vergleichspopulation unzureichend operationalisiert und angewendet werden. Die Fallzahl von 12 Personen erscheint zu gering, um einen validen indirekten Vergleich durchzuführen. Insbesondere unter Einbezug der vom pU herangezogenen Propensity-Score-Methoden und Gewichtungungsverfahren.

Die Unterschiede in den Ein- und Ausschlusskriterien, Daten und Mechanismen für die Aufnahme in die klinische Studie sowie in die Vergleichspopulation begünstigen einen Selektionsbias und resultieren in potenziell systematisch unterschiedlichen Populationen für einen nicht randomisierten Vergleich ohne Brückenkomparator.

Die Validität und der Zeitpunkt der Erhebung von unstrukturierten Daten, welche insbesondere krankheitsspezifische Informationen enthalten und z. T. synthetisch hergeleitet worden sind, sind auf Basis der vorliegenden Informationen unklar. Es kann nicht ausgeschlossen werden, dass systematisch unterschiedlich Index-Zeitpunkte/Time-Zeros für die Intervention und Kontrollpopulationen vorliegen und die Initiierung einer Zweitlinientherapie zu unterschiedlichen Zeitpunkten im Krankheitsverlauf erfolgte.

Der pU legt eine systematische Literaturrecherche nach Confoundern sowie anschließender Einordnung der Relevanz durch Expertinnen und Experten vor. Die Einordnung der Confounder nach Relevanz und möglicher inhaltlicher Zusammenführung kann nicht abschließend nachvollzogen werden. Es wurden nicht alle als relevant bewerteten Confounder ins PS-Modell aufgenommen.

Auf Basis der vorgelegten Datensituation (berichtete Baseline-Charakteristiken, Anwendung der Ein- und Ausschlusskriterien, Confounder-Auswahl, Time-Zero, Berichterstattung des Modells) kann nicht davon ausgegangen werden, dass eine ausreichende statistische Positivität der beiden Populationen für eine adjustierte Analyse vorhanden ist. Die Erfüllung der notwendigen SITA-Bedingung (kein ungemessenes Confounding und bspw. ein PS zwischen 0 und 1) wird als nicht dargelegt angesehen. Der Effekt im Kontrollarm kann auf Basis der vorliegenden Unterlagen nicht als kontrafaktisches Ergebnis (auf Basis einer kontrafaktischen Analyse; Counterfactual-Outcome) für Patientinnen und Patienten des Studienarms gesehen werden.

Die Interpretation eines berechneten Behandlungsunterschieds erscheint nicht sachgerecht. In der Gesamtschau sind die vom pU vorgelegten Ergebnisse nicht für die Quantifizierung eines Zusatznutzens geeignet und werden nicht für die Nutzenbewertung herangezogen.

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Tabelle 6: Allgemeine Angaben; Studie HERIZON_BTC

Studie HERIZON_BTC Allgemeine Angaben	HERZION
Gescreente Personen	131
Kohorte 1 ¹⁾	80
Kohorte 2 ²⁾	7
<u>Zulassungsrelevante Teilpopulation³⁾</u>	
Enrolled-Analysis-Set ⁴⁾ , n (%)	62 (100)
Sicherheitspopulation, n (%)	62 (100)
Abbruch der Einnahme der Studienmedikation, n (%)	
Aufgrund von:	
Radiologischer Progress	50 (80,6)
Klinischer Progress	1 (1,6)
Unerwünschtes Ereignis	2 (3,2)
Tod	1 (1,6)
Lost to Follow-Up	0
Rücknahme der Einverständniserklärung	0
Beendigung der Studie	1 (1,6)
Sonstiges (kein UE)	6 (9,7)
Personen in der Nachbeobachtung zum Gesamtüberleben	0
Erhalt einer Folgetherapie, n (%)	29 (46,8)
Mediane Behandlungsdauer in Monaten (min; max)	7,4 (0,5; 35,4)
Anzahl der Behandlungszyklen: Median (min; max)	8 (1; 35)
Mediane Beobachtungsdauer Monate (min; max)	34 (28; 45)

¹⁾ Personen mit HER2-Amplifikation ermittelt mittels ISH und HER2-Überexpression durch IHC (IHC 2+ und IHC 3+).

²⁾ Personen mit HER2-Amplifikation ermittelt mittels ISH und HER2 IHC 0 oder IHC 1+.

³⁾ Für die Nutzenbewertung relevant ist die Teilpopulation der Personen mit einer HER2+-Überexpression von IHC 3+

⁴⁾ Definition, siehe Kapitel 2.4.

Abkürzungen: IHC: Immunhistochemie; ISH: In-situ-Hybridisierung; UE: Unerwünschtes Ereignis.

Tabelle 7: Charakterisierung der Studienpopulation in der Studie HERIZON_BTC

Studie HERIZON_BTC Charakterisierung der Studienpopulation	Zanidatamab N = 62
<i>Alter (Jahre)</i>	
MW (SD)	62,6 (9,3)
Median (min; max)	64 (38; 79)
<i>Altersgruppe (Jahre), n (%)</i>	
< 65	33 (53,2)
65-74	28 (45,2)
> 75	1 (1,6)

Studie HERIZON_BTC Charakterisierung der Studienpopulation	Zanidatamab N = 62
<i>Geschlecht, n (%)</i> männlich weiblich	 28 (45,2) 34 (54,8)
<i>Abstammung (genetisch), n (%)</i> kaukasisch/weiß asiatisch amerikanisch indigen nicht berichtet/unbekannt	 19 (30,6) 38 (61,3) 1 (1,6) 4 (6,4)
<i>Region, n (%)</i> Nordamerika Asian andere	 15 (24,2) 36 (58,1) 11 (17,7)
<i>ECOG-Performance Score, n (%)</i> 0 1	 20 (32,3) 42 (67,7)
<i>Tumorentität, n (%)</i> Gallenblasenkarzinom intrahepatisches Cholangiokarzinom extrahepatisches Cholangiokarzinom Perihilar Distal	 33 (53,2) 17 (27,4) 12 (19,4) 5 (8,1) 7 (11,3)
<i>Krankheitsstadium zur initialen Diagnose, n (%)</i> I II III IV unbekannt	 2 (3,2) 8 (12,9) 16 (25,8) 34 (54,8) 2 (3,2)
<i>Krankheitsstadium zum Studieneintritt, n (%)</i> I II IIIA IIIB IIIC IV IVA IVB unbekannt	 0 0 1 (1,6) 7 (11,3) 0 21 (33,9) 0 33 (53,2) 0
<i>Einschränkung der Leberfunktion zu Baseline¹⁾, n (%)</i> keine mild moderat schwer	 32 (51,6) 30 (48,4) 0 0
<i>Einschränkung der Nierenfunktion zu Baseline²⁾, n (%)</i> keine mild bis moderat	 20 (32,3) 42 (67,7)
<i>Ergebnis der vorherigen Therapie, n (%)</i> Progress Intolerant	 58 (93,5) 4 (6,5)

Studie HERIZON_BTC Charakterisierung der Studienpopulation	Zanidatamab N = 62
<i>Zeit von Erstdiagnose bis zum fortgeschrittenen oder metastasierenden Krankheitsstadium</i> MW (SD) Median (min; max)	4,62 (10,5) 0 (0; 72,0)
<i>Dokumentierte Hirnmetastasen, n (%)</i> Ja Nein	0 62 (100)
<i>Ergebnis der Testung der HER2-Amplifikation³⁾, n (%)</i> 3+ Anderes	62 (100) 0
<i>Vorherige systemische Anti-Krebs-Therapie, n (%)</i> Ja Nein	62 (100) 0
<i>Anzahl von Therapieregime für metastasierend oder lokal fortgeschrittene Erkrankung⁴</i> MW (SD) Median (min; max)	1,7 (1,2) 1 (1; 8)
<i>Anzahl von Therapieregimen für metastasierend oder lokal fortgeschrittene Erkrankung, n (%)</i> 0 1 2 3 4 5 6 7	0 37 (59,7) 19 (30,6) 3 (4,8) 1 (1,6) 1 (1,6) 0 1 (1,6)
<i>Erhalt von Gemcitabin, n (%)</i> Ja Nein	62 (100) 0
<i>Erhalt folgender Therapieregime⁵⁾, n (%)</i> Gemcitabin/Oxaliplatin Gemcitabin/Cisplatin Gemcitabin/Fluoropyrimidin Gemcitabin/Anderes Gemcitabin-Monotherapie Fluoropyrimidin PD1/PDL1-Inhibitor	10 (16,1) 47 (75,8) 5 (8,1) 1 (1,6) 3 (4,8) 20 (32,3) 16 (25,8)
<i>Vorherige Radiotherapie, n (%)</i> Ja Nein	9 (14,5) 53 (85,5)
<i>Vorherige chirurgische Intervention und kurativer Intention, n (%)</i> Ja Nein	17 (27,4) 45 (72,6)

¹⁾ Nach den Kriterien der Organ Dysfunction Working Group des National Cancer Institute.

- ²⁾ Baseline-Nierenfunktionsstörung gemäß der Cockcroft-Gault-Formel zur Schätzung der Kreatinin-Clearance und der FDA-Richtlinie „Pharmacokinetics in Patients with Impaired Renal Function - Study Design, Data Analysis, and Impact on Dosing and Labeling, September 2020“.
- ³⁾ Alle Personen, die in die Studie eingeschlossen worden sind, waren zum Screeningzeitpunkt ISH+ ermittelt durch Companion Diagnostic und bewertet durch ein Zentrallabor.
- ⁴⁾ Beinhaltet Gemcitabin-basierte Therapien, die im adjuvanten/neoadjuvanten Setting nach einer Operation erhalten worden sind, sofern das Progressionsereignis zur Aufnahme in die klinische Studie innerhalb von 6 Monate stattfand.
- ⁵⁾ Mehrfachnennungen sind möglich.

Abkürzungen: ECOG: Eastern Cooperative Oncology Group; FDA: U.S. Food and Drug Administration; HER2: Human Epidermal Growth Factor Receptor 2; ISH: In-situ-Hybridisierung; MW: Mittelwert; SD: Standardabweichung.

Protokollverletzungen

Anhand der in der Studie HERIZON_BTC dokumentierten Protokollverletzungen ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Begleitmedikation

Anhand der in der Studie HERIZON_BTC dokumentierten Begleitmedikation ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Folgetherapien

Anhand der in der Studie HERIZON_BTC dokumentierten Folgetherapien ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

3.2 Mortalität

Gesamtüberleben

Tabelle 8: Ergebnisse zum Gesamtüberleben in der Studie HERIZON_BTC (Datenschnitt: 28.07.2024); Efficacy Analysis Set

Studie HERIZON_BTC Gesamtüberleben	HERIZON_BTC N = 62
Todesfälle, n (%)	45 (72,6)
Zensierungen, n (%)	
Lebend ¹⁾	10 (16,1)
Lost to Follow-up	2 (3,2)
Rücknahme der Einverständniserklärung	5 (8,1)
Mediane Beobachtungszeit (Monate)	16,41
Mediane Überlebensdauer (Monate) [95%-KI]	18,1 [12,2; 22,9]
Gesamtüberleben, KM-Schätzer (in %) [95%-KI] zu Monat 12	65,0 [51,6; 75,6]

¹⁾ Die Studie wurde nach der letzten Visite der letzten Person beendet, sodass die Studie beendet worden ist, bevor alle Ereignisse auftreten konnten.

Abkürzungen: KI: Konfidenzintervall; KM: Kaplan-Meier

IHC 3+ Cohort 1

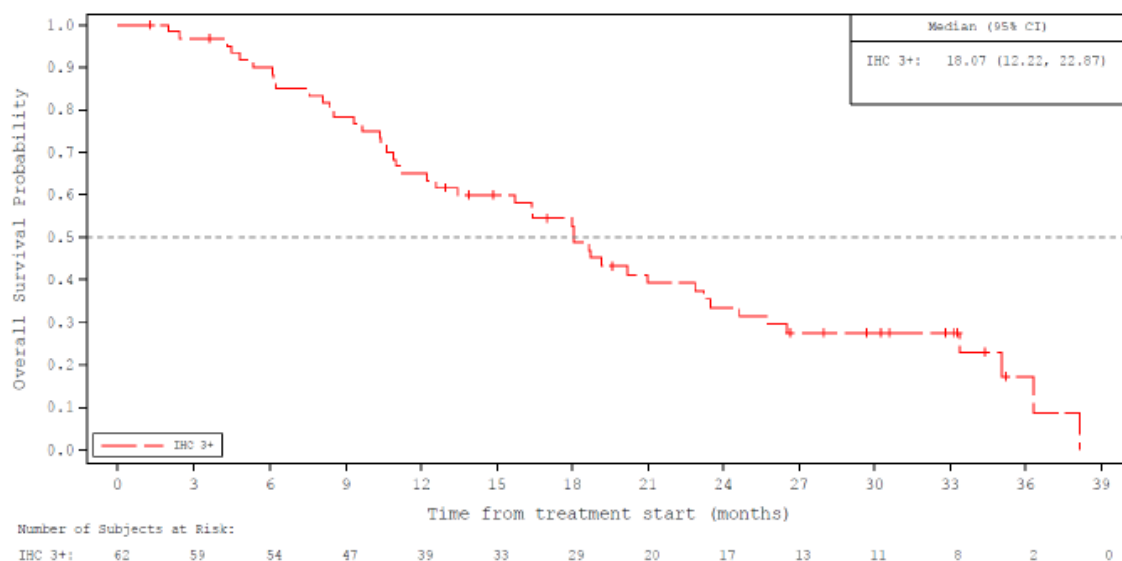


Abbildung 2: Kaplan-Meier-Kurve für Zeit bis zum Tod; Studie HERIZON_BTC (Efficacy Analysis Set)

3.3 Morbidität

Die Endpunkte „EQ-5D-VAS“ sowie „BPI-SF“ werden dargestellt. Im Anhang wird aus Transparenzgründen der primäre Endpunkt „Bestätigte objektive Ansprechrate“ dargestellt.

EQ-5D-VAS

Tabelle 9: Anteil der Personen mit einer Verschlechterung der EQ-5D-VAS um ≤ 15 Punkte zum Ende der Behandlung; Studie HERIZON_BTC; Efficacy Analysis Set (Datenschnitt: 28.07.2024)

Studie HERIZON_BTC EQ-5D-VAS ¹⁾	Zanidatamab N = 62
Baseline	N = 60
MW (SD)	75,2 (18,61)
Median (min; max)	80,0 (30; 100)
Nachbeobachtungszeit in Monaten ²⁾	N = 60
MW (SD)	10,7 (9,4)
Median (min; max)	7,5 (1; 35)
Personen mit einer Erhebung zu EOT, n (%)	44 (71,0)
Personen mit einer Verschlechterung ³⁾ zu EOT, n (%)	9 (14,5)

¹⁾ Werte von 0 bis 100; höhere Werte entsprechen einem besseren Gesundheitszustand.

²⁾ Beinhaltet alle Personen mit mindestens einer Erhebung.

³⁾ Eine Verschlechterung entspricht einer Reduktion des Scores zum Auswertungszeitpunkt im Vergleich zu Baseline um $\leq -15\%$ der Skalenspannweite.

Abkürzungen: EOT: Ende der Behandlung; MW: Mittelwert; SD: Standardabweichung; VAS: Visuelle Analogskala.

Schmerz mittels BPI-SF

Tabelle 10: Anteil der Personen mit einer Verschlechterung der BPI-SF um $\leq 15\%$ zum Ende der Behandlung; Studie HERIZON_BTC; Efficacy Analysis Set (Datenschnitt: 28.07.2024)

Studie HERIZON_BTC BPI-SF ¹⁾	Zanidatamab N = 62
Schmerzen (Mittelwert der Items 3 bis 6)²⁾	
Baseline MW (SD) Median (min; max)	N = 59 1,4 (1,6) 1,0 (0; 8)
Nachbeobachtungszeit in Monaten ³⁾ MW (SD) Median (min; max)	10,7 (9,4) 7,5 (1; 35)
Personen mit einer Erhebung zu EOT, n (%)	43 (69,4)
Personen mit einer Verschlechterung ⁴⁾ zu EOT, n (%)	8 (18,6)
Stärkster Schmerz der letzten 24 Stunden (Item 3)	
Baseline MW (SD) Median (min; max)	N = 59 2,2 (2,44) 1,0 (0; 8)
Nachbeobachtungszeit in Monaten ³⁾ MW (SD) Median (min; max)	N = 60 10,7 (9,44) 7,5 (1; 35)
Personen mit einer Erhebung zu EOT, n (%)	41 (66,1)
Personen mit einer Verschlechterung ⁵⁾ zu EOT, n (%)	8 (12,9)
Beeinträchtigungen durch Schmerz (Items 9a bis 9g)	
Baseline MW (SD) Median (min; max)	N = 59 1,4 (1,98) 0,4 (0; 8)
Nachbeobachtungszeit in Monaten ³⁾ MW (SD) Median (min; max)	N = 60 10,7 (9,4) 7,5 (1; 35)
Personen mit einer Erhebung zu EOT, n (%)	41 (66,1)
Personen mit einer Verschlechterung ⁶⁾ zu EOT, n (%)	19 (30,6)

¹⁾ Schmerz wird auf einer Skala von 0 bis 10 bewertet; höhere Werte entsprechen einem schlimmeren Schmerzempfinden:

²⁾ Besteht aus Fragen zu den stärksten Schmerzen in den letzten 24 Stunden, geringsten Schmerzen in den letzten 24 Stunden, durchschnittlichen Schmerzen und Schmerzen zum jetzigen Zeitpunkt.

³⁾ Beinhaltet alle Personen mit mindestens einer Erhebung.

⁴⁾ Eine Verschlechterung entspricht einer Erhöhung des Scores zum Auswertungszeitpunkt im Vergleich zu Baseline $\geq 15\%$ der Skalenspannweite.

⁵⁾ Eine Verschlechterung entspricht einer Erhöhung des Scores zum Auswertungszeitpunkt im Vergleich zu Baseline ≥ 2 Punkte.

⁶⁾ Eine Verschlechterung entspricht einer Erhöhung des Scores zum Auswertungszeitpunkt im Vergleich zu Baseline $\geq 15\%$ der Skalenspannweite.

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; EOT: Ende der Behandlung; MW: Mittelwert; SD: Standardabweichung.

3.4 Lebensqualität

Es wurden keine Endpunkte der Lebensqualität erhoben.

3.5 Sicherheit

Tabelle 11: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungsphase, Studie HERIZON_BTC

Studie HERIZON_BTC Zusammenfassung der UE <i>Personen mit mindestens einem ...</i>	Zanidatamab N = 62 n (%)
UE (ergänzend dargestellt)	61 (98,4)
UE CTCAE-Grad ≥ 3	40 (64,5)
SUE	33 (53,2)
UE, das zum Abbruch der Studienmedikation führte	2 (3,2)

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Unerwünschte Ereignisse

Tabelle 12: UE mit Inzidenz $\geq 10\%$ in der Studie HERIZON_BTC; Sicherheitspopulation

Studie HERIZON_BTC UE <i>MedDRA-Systemorganklasse Preferred Term</i>	Zanidatamab N = 62 n (%)
Erkrankungen des Gastrointestinaltrakts	50 (80,6)
Diarrhö	33 (53,2)
Abdominalschmerz	12 (19,4)
Übelkeit	13 (21,0)
Erbrechen	11 (17,7)
Schmerzen Oberbauch	7 (11,3)
Untersuchungen	31 (50,0)
Alanin-Aminotransferase erhöht	15 (24,2)
Aspartat-Aminotransferase erhöht	15 (24,2)
Ejektionsfraktion vermindert	9 (14,5)
Vermindertes Gewicht	8 (12,9)
Bilirubin im Blut erhöht	8 (12,9)
Alkalische Phosphatase im Blut erhöht	7 (11,3)
Gamma-Glutamyltransferase erhöht	7 (11,3)
Infektionen und parasitäre Erkrankungen	33 (53,2)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	26 (41,9)
Fieber	10 (16,1)
Ermüdung	9 (14,5)
Asthenie	6 (9,7)
Stoffwechsel- und Ernährungsstörungen	24 (38,7)
Verminderter Appetit	10 (16,1)

Studie HERIZON_BTC UE MedDRA-Systemorganklasse <i>Preferred Term</i>	Zanidatamab N = 62 n (%)
Hypokaliämie	9 (14,5)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	25 (40,3)
Infusions-bedingte Reaktionen	25 (40,3)
Erkrankungen des Blutes und des Lymphsystems	20 (32,3)
Anämie	16 (25,8)
Erkrankungen der Haut und des Unterhautgewebes	21 (33,9)
Juckreiz	10 (16,1)
Leber- und Gallenerkrankungen	13 (21,0)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	18 (29,0)
Erkrankungen des Nervensystems	17 (27,4)
Schwindel	7 (11,3)
Gefäßerkrankungen	16 (25,8)
Hypertonie	10 (16,1)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	11 (17,7)
Herzkrankungen	9 (14,5)

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

Tabelle 13: Schwere UE (Grad 3 oder höher), die bei $\geq 5\%$ der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der Studie HERIZON_BTC; Sicherheitspopulation

Studie HERIZON_BTC Schwere UE MedDRA-Systemorganklasse <i>Preferred Term</i>	Zanidatamab N = 62 n (%)
Erkrankungen des Gastrointestinaltrakts	15 (24,2)
Diarrhö	6 (9,7)
Leber- und Gallenerkrankungen	12 (19,4)
Cholestatische Gelbsucht	3 (4,8)
Gallengangverschluss	3 (4,8)
Untersuchungen	14 (22,6)
Alanin-Aminotransferase erhöht	5 (8,1)
Aspartat-Aminotransferase erhöht	4 (6,5)
Infektionen und parasitäre Erkrankungen	12 (19,4)
Pneumonie	3 (4,8)
Sepsis	2 (3,2)
Erkrankungen des Blutes und des Lymphsystems	11 (17,7)
Anämie	9 (14,5)

Studie HERIZON_BTC Schwere UE <i>MedDRA-Systemorganklasse</i> <i>Preferred Term</i>	Zanidatamab N = 62 n (%)
Stoffwechsel- und Ernährungsstörungen	5 (8,19)
Hypokaliämie	4 (6,5)
Gefäßkrankungen	7 (11,3)
Hypertonie	6 (9,7)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	5 (8,1)

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

Tabelle 14: SUE, die bei $\geq 5\%$ der Patientinnen und Patienten in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term in der Studie HERIZON_BTC; Sicherheitspopulation

Studie HERIZON_BTC SUE <i>MedDRA-Systemorganklasse</i> <i>Preferred Term</i>	Zanidatamab N = 62 n (%)
Leber- und Gallenerkrankungen	12 (19,4)
Cholestatische Gelbsucht	3 (4,8)
Gallengangverschluss	3 (4,8)
Gastrointestinale Erkrankungen	11 (17,7)
Infektionen und parasitäre Erkrankungen	12 (19,4)
Pneumonie	4 (6,5)

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

4.1 Design und Methodik der Studie

Studiendesign

Die Studie HERIZON_BTC war eine einarmige, mulizentrische, internationale offene Phase-IIb-Studie zur Untersuchung der Sicherheit und Wirksamkeit von Zanidatamab in der zulassungskonformen Dosierung bei Erwachsenen mit biliärem Karzinom und einer HER2-Überexpression. Es konnten Personen eingeschlossen werden, die HER2-positiv waren. Eine Überexpression wurde nachgewiesen mittels Immunhistochemie von 2+ und In-situ-Hybridisierung oder einer Überexpression mit IHC 3+. Für die Nutzenbewertung relevant ist die zulassungsrelevante Teilpopulation mit IHC 3+. In die Studie eingeschlossenen wurden Erwachsene mit Cholangiokarzinom (intrahepatisch und extrahepatisch) sowie Gallenblasenkarzinom. Diese musste mindestens eine vorherige systemische Therapie, welche Gemcitabin enthalten musste, für eine fortgeschrittene Erkrankung erhalten haben und eine Krankheitsprogression oder Intoleranz gegenüber der zuletzt erhaltenen Therapie aufweisen. Zum Nachweis der HER2-Überexpression wurde eine Tumorbiopsie durchgeführt, wobei Personen auch im Vorfeld des Screenings ein „Pre-Screening“ hinsichtlich der HER2-Überexpression durchlaufen konnten. Sollten diese Personen eine Krankheitsprogression oder Intoleranz gegenüber ihrer vorherigen Therapie aufweisen, musste die Einverständniserklärung für die weiteren Studienprozeduren ausgefüllt werden. Von 87 Personen (unabhängig des Grades der HER2-Amplifizierung), die das Pre-Screening durchlaufen hatten, wurden 66 in die Studie aufgenommen. Insgesamt durchliefen 131 Personen die Screeningphase. Der häufigste Grund einer fehlenden Eignung zur Studienteilnahme war der fehlende Nachweis einer HER2-Überexpression durch die Untersuchung im Zentrallabor (29,5 %). Weitere Ausschlussgründe waren eine unzureichende Leberfunktion (15,9 %) und ein ECOG PS von > 1 (11,4 %) sowie andere Faktoren, die potenziell Einfluss auf die Sicherheit oder Compliance haben könnten (11,4 %). Der primäre Endpunkt der Studie war die bestätigte objektive Ansprechrate. Die mediane Nachbeobachtung in der Studie HERIZON_BTC beträgt 34 Monate.

Insgesamt wurden 62 Personen in die bewertungsrelevante Population mit HER2-Überexpression mit IHC 3+ in die Studie eingeschlossen. Ein IHC-Wert der Stufe 3+ bedeutet, dass eine vollständige Anfärbung von > 10 % der Tumorzellen nachweisbar ist. Dies bedeutet eine Überexpression des HER2-Proteins.

Zum Zeitpunkt der Nutzenbewertung ist keine Person mehr in der Nachbeobachtung zum Gesamtüberleben und die Studie ist abgeschlossen. Nach Angaben des Sponsors wurden die Studienzentren im Februar 2024 über die beabsichtigte Beendigung der Studie informiert, da die Studienziele als erreicht angesehen wurden. Teilnehmende, die weiterhin von der Behandlung profitierten, konnten in ein Named-Patient-Supply-Programm überführt werden, um Zanidatamab im Rahmen eines Compassionate Use weiter zu erhalten. Die letzte Visite der letzten eingeschlossenen Person (Last Patient Last Visit, LPLV) fand am 11. Juli 2024 statt. Vor dem Hintergrund des Studienprotokolls, das eine Behandlung bis zum Eintreten von Progression, nicht tolerierbarer Toxizität oder anderer Abbruchkriterien vorsah, ist die Beendigung der Studie jedoch nicht primär durch patientenindividuelle Ereignisse bedingt, sondern durch eine sponsor-initiierte Entscheidung nach Erreichen der Studienziele. Dies kann potenziell die Vollständigkeit der Langzeitbeobachtung, insbesondere hinsichtlich des Gesamtüberlebens, einschränken.

Studienpopulation

Bei der Studienpopulation handelt es sich um Erwachsene, die im Median (min; max) 64 (38; 79) Jahre alt waren. Ein Großteil der Patientinnen und Patienten war asiatischer Herkunft (n = 38; 61,3 %) und 19 Personen (30,6 %) waren kaukasischer Herkunft.

Es wurden insgesamt 3 Subtypen der biliären Karzinome in die klinische Studie eingeschlossen. 33 (53,2 %) der Personen hatten ein Gallenblasenkarzinom, 17 (27,4 %) hatten ein intrahepatisches Cholangiokarzinom und 12 (19,4 %) ein extrahepatisches Cholangiokarzinom. Davon waren 5 (8,1 %) perihilär und 7 (11,3 %) distal. Bei 33 (53,2 %) der Patientinnen und Patienten lag zum Studieneintritt ein Krankheitsstadium IVb und bei 21 weiteren (33,9 %) ein Stadium IV vor. In die Studie eingeschlossen wurden Personen mit ECOG PS von 0 oder 1 und bei 42 (67,7 %) der Patientinnen und Patienten lag ein ECOG PS von 1 vor.

Studienmedikation

Alle Personen, die in die Studie eingeschlossen worden sind, erhielten die Studienmedikation. Die Gabe der Studienmedikation erfolgte in der zulassungskonformen Dosierung von 20 mg/kg mittels intravenöser Gabe alle 2 Wochen bis zum Auftreten von nicht akzeptabler Toxizität, Krankheitsprogression (radiographisch oder klinisch) Tod, Loss to Follow-up, Schwangerschaft, Entscheidung des Prüfpersonals oder Rücknahme der Einverständniserklärung. Die mediane Dauer der Behandlung mit Zanidatamab beträgt 7,39 Monate. Der häufigste Grund für den Abbruch der Studienmedikation war die radiologische Progression bei 50 Personen (80,6 %). Weiterhin brachen 6 Personen (9,7 %) die Studie wegen einem Event ab, welches kein UE war, aber nicht näher spezifiziert worden ist.

Zum Zeitpunkt der Nutzenbewertung erhält keine Person mehr die Studienmedikation.

Bewertung der Validität und Aussagekraft der Studie und des VZP

Da es sich bei der Studie HERIZON_BTC um eine unkontrollierte Studie handelt, wird von einem hohen Verzerrungspotenzial ausgegangen.

Indirekter Vergleich

Der vorgelegte indirekte Vergleich wird nicht für die Nutzenbewertung herangezogen. Die Limitationen sind in Kapitel 2.6 beschrieben.

4.2 Zulassungsstatus und Zulassungspopulation

Zulassungsstatus und Anwendungsgebiet

Zanidatamab (Ziihera®) ist laut Fachinformation als Monotherapie zur Behandlung von Erwachsenen mit fortgeschrittenem oder metastasiertem HER2-positivem (IHC 3+) biliärem Karzinom indiziert, die zuvor mindestens eine systemische Therapie erhalten haben.

Patientinnen und Patienten, die mit Zanidatamab behandelt werden, müssen einen dokumentierten HER2-positiven Tumorstatus aufweisen, definiert als Score von 3+, ermittelt durch Immunhistochemie mit einem gemäß dem Verwendungszweck CE-gekennzeichneten In-vitro Diagnostikum.

Die Studie HERIZON_BTC wurde in Erwachsenen mit HER2-Amplifizierung mit einer Überexpression 2+ und 3+ erhoben mittels Immunhistochemie. Die Zulassung erfolgte für Personen mit IHC 3+. Für diese Teilpopulation der klinischen Studie liegen Auswertungen vor, die für die Nutzenbewertung herangezogen werden.

Die Ein- und Ausschlusskriterien legen fest, dass Personen im Vorfeld eine Gemcitabin-haltige systemische Chemotherapie erhalten haben müssen, während das Anwendungsgebiet dies nicht einschränkt. Die S3-Leitlinie zum biliärem Karzinom empfiehlt eine Gemcitabin-haltige Chemotherapie als Erstlinientherapie in der vorliegenden Therapiesituation.

Es liegt ein „conditional approval“ seitens der European Medicines Agency (EMA) vor. Die Zulassungsaufgabe ist die Durchführung einer Phase-III-Studie. Bei dieser klinischen Studie handelt es sich um eine stratifizierte nach HER2-Tumorexpression (IHC 3+ vs. IHC 2+ /ISH+) randomisierte Studie in der Erstlinie gegenüber Standard of Care (SoC) bei HER2-positiven Patientinnen und Patienten. Die Ergebnisse müssen bis September 2029 bei der EMA eingereicht werden.

Übertragbarkeit auf deutschen Versorgungskontext

Die klinische Studie HERIZON_BTC wurde zum Großteil in Asien durchgeführt und 61,2% der behandelten Personen waren asiatischer Herkunft. Der hohe Anteil einer asiatischen Population kann die Übertragbarkeit auf den deutschen Versorgungskontext einschränken. Im deutschen Versorgungskontext wird in der S3-Leitlinie als Erstlinientherapie [2] eine Kombinationstherapie aus Gemcitabin und Cisplatin sowie zusätzlich Durvalumab oder Pembrolizumab empfohlen. Die Umsetzung im Rahmen der klinischen Studie kann auf Basis der vorliegenden Unterlagen nicht abschließend bewertet werden. Es kann jedoch davon ausgegangen werden, dass ein Großteil der Personen keinen PD1-Antikörper während der Erstlinientherapie erhalten haben. Es ist unklar, inwiefern dies die Ergebnisse für Wirksamkeit und Sicherheit in der Zweitlinientherapie beeinflusst.

4.3 Mortalität

In der Endpunktkategorie „Mortalität“ werden in der Nutzenbewertung die Ergebnisse zum Gesamtüberleben berücksichtigt. Die dargestellten Ergebnisse zum Gesamtüberleben sind lediglich deskriptiv, da es sich um eine einarmige Studie handelt.

Bis zum finalen Datenschnitt vom 28.07.2024 verstarben 45 (72,6%) der teilnehmenden Patientinnen und Patienten. Das mediane Überleben beträgt 18,07 Monate (95%-KI: [12,22; 22,87]). 10 Personen waren zum finalen Datenschnitt noch am Leben. Die Studie wurde seitens des pUs beendet und keine Person befindet sich mehr in der Nachbeobachtung zum Gesamtüberleben.

Eine Interpretation und Bewertung der geschätzten Überlebenszeit ist aufgrund der fehlenden Kontrollgruppe nicht möglich. Der Effekt von Zanidatamab auf die Mortalität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.4 Morbidität

Für die Endpunktkategorie „Morbidität“ legt der pU Daten zu den Endpunkten EQ-5D-VAS, BPI-SF sowie zum Tumoransprechen vor. Die detaillierte Bewertung der Endpunkte ist im Kapitel 2.3 abgebildet. Zusammenfassend wurden für die Nutzenbewertung die Ergebnisse der patientenberichteten Instrumente EQ-5D-VAS und BPI-SF herangezogen. Der primäre Endpunkt „bestätigtes objektives Ansprechen“ wird aus Transparenzgründen im Anhang dargestellt.

Es liegen ausreichend hohe Rücklaufquoten von 70 % zum Ende der Behandlung (EOT) vor. Aufgrund der palliativen Therapiesituation werden die Responderanalysen zur Verschlechterung herangezogen. Bei der EQ-5D-VAS lagen bei 9 von 62 Personen (14,5 %) eine Verschlechterung um mindestens 15 % zum Zeitpunkt des Behandlungsendes vor, wobei für ungefähr 30 % der Personen keine Rückläufe vorlagen.

Bei der BPI-SF wurden für 2 Domänen Summenscores berechnet. Bei der Domäne „Schmerzintensität“ lag bei 8 Personen (18,6 %) zum Zeitpunkt des Behandlungsendes eine Verschlechterung um 15 % gegenüber dem Baselinewert vor. Bei der Domäne „Beeinträchtigungen durch Schmerzen“ lag bei 19 Patientinnen und Patienten (30,6 %) zum Zeitpunkt des Behandlungsendes eine Verschlechterung um ≥ 15 % gegenüber dem Baselinewert vor. Für 41 von 62 Personen (66 %) lag eine Erhebung zum Ende der Behandlung vor.

Bei der Interpretation der Ergebnisse muss beachtet werden, dass Personen ohne Erhebung nicht als Personen mit Verschlechterung in die Auswertung eingehen. Es kann aufgrund der vorliegenden Erkrankung nicht davon ausgegangen werden, dass die fehlenden Daten als zufällig fehlend zum relevanten Bewertungszeitpunkt eingestuft werden können.

Da es sich um eine offene, einarmige Studie handelt, sind die Ergebnisse, insbesondere der subjektiven, patientenberichteten Endpunkte, als potenziell hochverzerrt anzusehen. Der Effekt von Zanidatamab auf die Morbidität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.5 Lebensqualität

Es wurden keine Endpunkte der Lebensqualität erhoben.

4.6 Sicherheit

Die Sicherheitserhebung in der Studie HERIZON_BTC begann mit der erstmaligen Gabe der Studienmedikation und endete 30 Tage nach der letzten Gabe der Studienmedikation. Die mediane Behandlungsdauer beträgt 7,39 Monate (min; max: 0,5; 35,4). Die mediane Beobachtungsdauer im Gesamtüberleben beträgt 16,41 Monate. Somit wurden unerwünschte Ereignisse für einen inhärenten Bestandteil der Nachbeobachtung nicht durchgehend erhoben. SUE, die nach der geplanten Sicherheitserhebung auftreten, sollten weiterhin erhoben werden, sofern sie im Zusammenhang mit der Prüfmedikation gesehen werden.

Bei fast allen teilnehmenden Personen trat ein UE auf (98,4 %). Bei 40 Personen (64,5 %) wurde ein UE mit CTCAE-Grad ≥ 3 erfasst. Bei 33 Personen (53,2 %) trat ein SUE auf.

Auf Ebene der Systemorganklassen traten am häufigsten „Erkrankungen des Gastrointestinaltrakts“ (80,6 %) als UE auf. Diese waren ebenfalls die am häufigsten aufgetretenen UE mit CTCAE-Grad ≥ 3 (24,2 %). Die am häufigsten berichteten SUE waren „Leber- und Gallenerkrankungen“ (19,4 %), „Gastrointestinale Erkrankungen“ (17,7 %) und „Infektionen und parasitäre Erkrankungen“ (19,4 %). Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die dargestellten UE eingingen.

Die Sicherheit von Zanidatamab kann aufgrund der fehlenden Kontrollgruppe nicht abschließend beurteilt werden.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Zanidatamab ist zugelassen laut Fachinformation als Monotherapie zur Behandlung von Erwachsenen mit fortgeschrittenem oder metastasiertem HER2-positivem (IHC 3+) biliärem Karzinom, die zuvor mindestens eine systemische Therapie erhalten haben. Die Nutzenbewertung von Zanidatamab basiert auf der zulassungsbegründenden Studie HERIZON_BTC. Es handelt sich bei der Studie um eine einarmige, multizentrische Phase-IIb-Studie zur Untersuchung der Sicherheit und Wirksamkeit.

Weitergehende, über die Angaben in der nachfolgenden zusammenfassenden Tabelle hinausgehende Erläuterungen zur Methodik und zu den Ergebnissen der Endpunkte, finden sich in den jeweiligen Kapiteln 2.3 bis 2.5 und 3.2 bis 3.5 der Nutzenbewertung.

Tabelle 15: Zusammenfassende Darstellung der Ergebnisse der Studie HERIZON_BTC

Studie HERIZON_BTC Darstellung der Ergebnisse	Zanidatamab N = 62	
Mortalität		
<i>Gesamtüberleben</i> Todesfälle, n (%) Mediane Überlebensdauer in Monaten	45 (72,6) 18,07	
<i>Gesamtüberleben, KM-Schätzer (in %) [95%-KI]</i> Zu Studienmonat 12	65,0 [51,6; 75,6]	
Morbidität		
EQ-5D-VAS	<i>Personen mit Erhebung, n (%)</i>	<i>Personen mit Ereignis, n (%)</i>
Verschlechterung der EQ-5D-VAS um ≥ 15 % zum Ende der Behandlung	44 (71,0)	9 (14,5)
BPI-SF	<i>Personen mit Erhebung, n (%)</i>	<i>Personen mit Ereignis, n (%)</i>
Verschlechterung der „Schmerzintensität“ um ≥ 15 % zum Ende der Behandlung	43 (69,4)	8 (18,6)
Verschlechterung der „Beeinträchtigungen durch Schmerzen“ um ≥ 15 % zum Ende der Behandlung	41 (66,1)	19 (30,6)
Sicherheit¹⁾	<i>N = 62</i>	<i>Personen mit Ereignis, n (%)</i>
UE CTCAE-Grad ≥ 3	62	40 (64,5)
SUE	62	33 (53,2)
Therapieabbrüche aufgrund von UE	62	2 (3,2)

¹⁾ Weitere Ergebnisse zur Sicherheit (UEs nach Organsystemen und Einzelereignisse) sind im Ergebniskapitel 3.5 Sicherheit dargestellt

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; KM: Kaplan-Meier; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; VAS: Visuelle Analogskala.

Referenzen

1. **Adamson B, Waskom M, Blarre A, Kelly J, Krismer K, Nemeth S, et al.** Approach to machine learning for extraction of real-world data variables from electronic health records. *Front Pharmacol* 2023;14:1180962.
<https://dx.doi.org/10.3389/fphar.2023.1180962>.
2. **Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Leitlinienprogramm Onkologie, Deutsche Krebsgesellschaft (DKG), Deutsche Krebshilfe.** Diagnostik und Therapie des Hepatozellulären Karzinoms und biliärer Karzinome; S3-Leitlinie, Langversion 5.2 [online]. AWMF-Registernummer 032-053OL. Last updated: 06.2025. Berlin (GER): AWMF; 2025. [Zugriff: 05.05.2026]. URL: https://register.awmf.org/assets/guidelines/032-053OL_S3_Diagnostik-Therapie-Hepatozellulaeres-Karzinom-biliaere-Karzinome_2025-06.pdf.
3. **European Medicines Agency (EMA).** Ziihera: European public assessment report EMEA/H/C/006380/0000 [online]. 25.04.2025. Amsterdam (NED): EMA. [Zugriff: 16.02.2026]. URL: https://www.ema.europa.eu/en/documents/assessment-report/ziihera-epar-public-assessment-report_en.pdf.
4. **Jazz Pharmaceuticals.** Comparative effectiveness of zanidatamab vs. external control for HER2-positive unresectable, locally advanced, or metastatic biliary tract cancer: study JZP598_515; clinical study report [unveröffentlicht]. 25.07.2025.
5. **Jazz Pharmaceuticals.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A: Zanidatamab, Erwachsene mit inoperablem, lokal fortgeschrittenem oder metastasiertem HER2-positivem (IHC 3+) biliärem Karzinom (BTC), die zuvor mit mindestens einer systemischen Therapie behandelt wurden; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 12.02.2026.
6. **Jazz Pharmaceuticals.** A Phase 2b, open-label, single-arm study of Zanidatamab (ZW25) monotherapy in participants with advanced or metastatic HER2-amplified biliary tract cancers: study ZWI-ZW25-203; clinical study report, version 3.0 [unveröffentlicht]. 05.02.2025.
7. **Jazz Pharmaceuticals.** Ziihera 300 mg Pulver für ein Konzentrat zur Herstellung einer Infusionslösung [online]. Dezember 2025. Frankfurt/Main (GER). [Zugriff: 16.02.2026]. URL: <https://www.fachinfo.de/fi/detail/025390/ziihera-300-mg-pulver-fuer-ein-konzentrat-zur-herstellung-einer-infusionsloesung?query=zanidatamab>.
8. **Jazz Pharmaceuticals.** Zusatzanalysen: Studie HERIZON-BTC-01 [unveröffentlicht]. 12.02.2026.

Anhang

Bestätigte objektive Ansprechrate (primärer Endpunkt)

Tabelle 16: Bestätigte objektive Ansprechrate in der Studie HERIZON_BTC (Datenschnitt: 28.07.2024); Efficacy Analysis Set

Bestätigte objektive Ansprechrate Studie HERIZON_BTC	Zanidatamab N = 62
Bewertbare Personen, n (%)	62 (100)
Bestätigtes bestes Ansprechen, n (%)	
CR	3 (4,8)
PR	29 (46,8)
SD	17 (27,4)
Nicht-CR / Nicht-PD	0
PD	13 (21,0)
Bestätigte objektive Ansprechrate (cORR), n (%) [95%-KI]	32 (51,6) [38,6; 64,5]

Abkürzungen: CR: Komplettes Ansprechen; PR: partielles Ansprechen, PD: Progressive Erkrankung; SD: Stabile Erkrankung.

Verabreichte Interventionen im externen Kontrollarm

Tabelle 17: Herangezogene und verabreichte Interventionen in der externen Vergleichspopulation

Verabreichte Interventionen Studie JZP598	Externer Kontrollarm N = 12
Personen mit Chemotherapie, n (%)	12 (100)
Systemische Chemotherapie, n (%)	
Capecitabin, Gemcitabin	1 (8,3)
Carboplatin, Gemcitabin	1 (8,3)
Cisplatin, Gemcitabin	1 (8,3)
Fluorouracil	1 (8,3)
Fluorouracil, Irinotecan, Leucovorin	2 (16,7)
Fluorouracil, Leucovorin, Oxaliplatin	6 (50)

Identifizierte und verwendete Confounder für den Indirekten Vergleich

Ob eine Zusammenführung von den inhaltlich überlappenden Variablen stattgefunden hat, kann auf Basis der vorliegenden Unterlagen nicht abschließend bewertet werden.

Tabelle 18: Identifizierte und verwendete Confounder für den Indirekten Vergleich

Confounder	Operationalisierung	In ≥ 10 Publikationen	Bewertung der Priorität durch Experten	PS-Modell
ECOG PS	Keine (über E/A)	Ja	Hoch	Nein
CA19.9	-	Ja	k. A. ¹⁾	Nein
CEA	-	Ja	k. A. ¹⁾	Nein
Krankheitsstadium ²⁾	Stufe I oder Stufe II/III/IV	Ja	Medium	Nein (unklar) ³⁾

Confounder	Operationalisierung	In ≥ 10 Publikationen	Bewertung der Priorität durch Experten	PS-Modell
NLR	-	Ja	k. A. ¹⁾	Nein
Albumin	-	Ja	k. A. ¹⁾	Nein
Tumorentität	GBC/iCCA/eCCA	Ja	Hoch	Ja
Metastasen	-	Nein	Unklar	Nein
Chemotherapie	-	Nein	Unklar	Nein
Ansprechen auf Chemotherapie	-	Nein	Unklar	Nein
Geschlecht	Frau/Mann	Nein	Mittel	Ja
Alkalinphosphatase	-	Nein	Unklar	Nein
Alter	Kontinuierlich	Nein	Gering	Ja
Zweitlinientherapie	-	Nein	Ja	Nein
Bilirubin	-	Nein	Unklar	Nein
Hämoglobin	-	Nein	Unklar	Nein
PLR	-	Nein	Unklar	Nein
PNI	-	Nein	Unklar	Nein
Vorherige Resektion	-	Nein	Unklar	Nein
Krankheitsstatus	-	Nein	Unklar	Nein
BMI	-	Nein	Unklar	Nein
ERBB2-Alteration	-	Nein	Unklar	Nein
LDH	-	Nein	Unklar	Nein
TNM-Krankheitsstufe	-	Nein	Hoch	Nein
Nennung durch Experten⁴⁾				
Geografie	Keine	Nein	Hoch	Nein
Lebererkrankung	Ja/Nein	Nein	Hoch	Ja

¹⁾ Es wurden Laborparameter als klinisch prognostische Faktoren identifiziert, die aufgrund eingeschränkter Verfügbarkeit im externen Kontrollarm nicht operationalisiert werden konnten.

²⁾ Ins PS-Modell wurde die TNM-Stufe zum Zeitpunkt der Diagnosestellung als möglicher Confounder im Studienprotokoll operationalisiert.

³⁾ Diese Variable wurde als Confounder operationalisiert, ging jedoch nicht ins PS-Modell ein.

⁴⁾ Zum Vorgehen der Experteninterviews konnten keine näheren Informationen identifiziert werden.

Abkürzungen: BMI: Body Mass Index; CA 19.9: Carbohydrat-Antigen; CEA: Carcinoembryonale Antigen; eCCA: extrahepatisches Cholangiokarzinom; ERBB2: Erb-B2 receptor tyrosine kinase 2; GBC: Gallenblasenkarzinom; iCCA: intrahepatisches Cholangiokarzinom; LDH: Laktatdehydrogenase; NLR: Neutrophilen-Lymphozyten Ratio; PLR: Thrombozyten-Lymphozyten Ratio; PNI: Prognostic Nutritional Index; PS-Modell: Propensity-Score-Modell; TNM: Tumor, Node, Metasta.