

Nutzenbewertung

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: Tarlatamab

Datum der Veröffentlichung: 1. Oktober 2026

Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abbildungsverzeichnis	4
Abkürzungsverzeichnis	6
Hintergrund.....	8
1 Fragestellung.....	9
2 Auswahl und Methodik der relevanten Studien	10
2.1 Studienbasis für die Nutzenbewertung	10
2.2 Beschreibung der eingeschlossenen Studie	11
2.3 Endpunkte	16
2.3.1 Mortalität.....	18
2.3.2 Morbidität.....	18
2.3.3 Lebensqualität	24
2.3.4 Sicherheit	25
2.3.5 Erhebungszeitpunkte	26
2.4 Statistische Methoden.....	27
2.5 Verzerrungspotential auf Studien- und Endpunktebene	30
3 Ergebnisse der eingeschlossenen Studie	32
3.1 Studiencharakteristika und Studienmedikation	32
3.2 Mortalität	35
3.3 Morbidität	37
3.4 Lebensqualität.....	39
3.5 Sicherheit	42
4 Diskussion der Methodik und Ergebnisse.....	50
4.1 Design und Methodik der Studie.....	50
4.2 Zulassungsstatus und Zulassungspopulation.....	51
4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit.....	52
4.4 Mortalität	52
4.5 Morbidität	53
4.6 Lebensqualität.....	54
4.7 Sicherheit	54
4.8 Zugrundeliegendes EU-HTA-Verfahren (JCA-MP-2024-17).....	55
5 Anforderung an eine qualitätsgesicherte Anwendung	57
6 Zusammenfassung der Nutzenbewertung	58
Referenzen	60
Anhang	62

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis	10
Tabelle 2:	Charakterisierung der Studie DeLLphi-304	11
Tabelle 3:	Für die Nutzenbewertung relevante Protokolländerungen der Studie DeLLphi-304....	14
Tabelle 4:	Charakterisierung der Intervention	15
Tabelle 5:	Zusammenfassung der Endpunktbewertung der Studie DeLLphi-304	16
Tabelle 6:	Erhebungszeitpunkte der berücksichtigten Endpunkte in der Studie DeLLphi-304 ...	26
Tabelle 7:	Verzerrungspotential der Studie DeLLphi-304.....	30
Tabelle 8:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie DeLLphi-304	31
Tabelle 9:	Allgemeine Angaben; Studie DeLLphi-304	32
Tabelle 10:	Charakterisierung der Studienpopulation; Studie DeLLphi-304	33
Tabelle 11:	Folgetherapien; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	35
Tabelle 12:	Gesamtüberleben; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	36
Tabelle 13:	Gesamtüberleben, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	37
Tabelle 14:	EORTC QLQ-C30 – Symptomskalen, Responderanalyse (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	38
Tabelle 15:	EORTC QLQ-C30 – Symptomskalen, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	39
Tabelle 16:	EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus, Responderanalysen (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025).....	40
Tabelle 17:	EORTC QLQ-C30 – Funktionsskalen, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	41
Tabelle 18:	Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der Behandlungsphase; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	42
Tabelle 19:	UE mit Inzidenz ≥ 10 % bzw. ≥ 10 Personen und ≥ 1 % der Personen in einem Studienarm; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	43
Tabelle 20:	Schwere UE (CTCAE-Grad ≥ 3), die bei ≥ 5 % der Personen bzw. ≥ 10 Personen und ≥ 1 % der Personen in einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	47
Tabelle 21:	SUE, die bei ≥ 5 % der Personen bzw. ≥ 10 Personen und ≥ 1 % der Personen in einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	48
Tabelle 22:	Zusammenfassende Darstellung der Ergebnisse der Studie DeLLphi-304	58

Abbildungsverzeichnis

Abbildung 1: Kaplan-Meier-Kurve für „Gesamtüberleben“; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025).....	36
Abbildung 2: Kaplan-Meier-Kurve für die Symptomskala „Fatigue“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	62
Abbildung 3: Kaplan-Meier-Kurve für die Symptomskala „Übelkeit und Erbrechen“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	63
Abbildung 4: Kaplan-Meier-Kurve für die Symptomskala „Schmerz“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	64
Abbildung 5: Kaplan-Meier-Kurve für die Symptomskala „Dyspnoe“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	65
Abbildung 6: Kaplan-Meier-Kurve für die Symptomskala „Schlaflosigkeit“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	66
Abbildung 7: Kaplan-Meier-Kurve für die Symptomskala „Appetitverlust“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	67
Abbildung 8: Kaplan-Meier-Kurve für die Symptomskala „Obstipation“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	68
Abbildung 9: Kaplan-Meier-Kurve für die Symptomskala „Diarrhö“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025).....	69
Abbildung 10: Kaplan-Meier-Kurve für die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	70
Abbildung 11: Kaplan-Meier-Kurve für die Skala „Körperliche Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	71
Abbildung 12: Kaplan-Meier-Kurve für die Skala „Rollenfunktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	72
Abbildung 13: Kaplan-Meier-Kurve für die Skala „Emotionale Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	73
Abbildung 14: Kaplan-Meier-Kurve für die Skala „Kognitive Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	74
Abbildung 15: Kaplan-Meier-Kurve für die Skala „Soziale Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)	75

Abbildung 16: Kaplan-Meier-Kurve für „Schwere UE (CTCAE-Grad ≥ 3)“ (Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	76
Abbildung 17: Kaplan-Meier-Kurve für „SUE „(Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	76
Abbildung 18: Kaplan-Meier-Kurve für „UE, das zum Abbruch der Studienmedikation führte“ (Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)	77

Abkürzungsverzeichnis

AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
BPI-SF	Brief Pain Inventory – Short Form
BOR	Bestes Gesamtansprechen (Best Overall Response)
CR	Komplettes Ansprechen (Complete Response)
CSR	Klinischer Studienbericht (Clinical Study Report)
CT	Computertomographie
CTCAE	Common Terminology Criteria for Adverse Events
DLL3	Delta-like Ligand 3
ECOG-PS	Eastern Cooperative Oncology Group – Performance Status
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30	EORTC Quality of Life Questionnaire – Core Questionnaire
EORTC QLQ-LC13	EORTC Quality of Life Questionnaire – Lung Cancer Module
EPAR	European Public Assessment Report
EQ-5D-VAS	Visuelle Analogskala des EuroQoL 5-Dimension
ES-SCLC	Kleinzelliges Lungenkarzinom im fortgeschrittenen Stadium (Extensive-Stage Small-Cell Lung Cancer)
FACT-G (GP5)	Functional Assessment of Cancer Therapy – General (Einzelitem GP5)
GBA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
HR	Hazard Ratio
HTA	Health Technology Assessment
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to Treat
JCA	Joint Clinical Assessment
KI	Konfidenzintervall
LTFU	Langzeit-Follow-up
MMRM	Mixed Model for Repeated Measures
MRT	Magnetresonanztomographie
N	Anzahl
OS	Gesamtüberleben (Overall Survival)
PD-(L)1	Programmed Cell Death (Ligand) 1
PFS	Progressionsfreies Überleben (Progression-Free Survival)
PGI-C	Patient Global Impressions scale – Change
PGI-S	Patient Global Impressions scale – Severity
PICO	Population, Intervention, Comparison, Outcome
PR	Partielles Ansprechen (Partial Response)
PRO	Patientenberichteter Endpunkt (Patient-Reported Outcome)
pU	pharmazeutischer Unternehmer
RECIST	Response Evaluation Criteria in Solid Tumors
SAP	Statistischer Analyseplan
SCLC	Kleinzelliges Lungenkarzinom (Small-Cell Lung Cancer)

SFU	Sicherheits-Follow-up
SGB	Sozialgesetzbuch
SUE	Schwerwiegendes unerwünschtes Ereignis
TEAE	Treatment-Emergent Adverse Events
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des GBA

Hintergrund

Tarlatamab 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 (GBA), 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 GBA 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 GBA 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 GBA 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 GBA bestimmt gemäß Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des GBA (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 Tarlatamab zieht der GBA die Zulassung und Unterlagen zur Zulassung (insbesondere den 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 Tarlatamab in seiner Sitzung am 22. September 2026 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 30. Juni 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 1. Oktober 2026 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des GBA (<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 GBA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Tarlatamab (IMDYLLTRA®) 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 GBA 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 [11]:

Behandlung erwachsener Patienten mit kleinzelligem Lungenkarzinom im fortgeschrittenen Stadium (extensive-stage small cell lung cancer, ES-SCLC), die nach Krankheitsprogression während oder nach einer Erstlinienbehandlung basierend auf platinhaltiger Chemotherapie eine systemische Therapie benötigen.

Gemäß Fachinformation umfasst die zugelassene Dosierung eine Anfangsdosis von 1 mg an Tag 1, gefolgt von 10 mg an den Tagen 8 und 15 sowie danach alle 2 Wochen. Die Behandlung erfolgt als Monotherapie. Patienten sollten bis zum Fortschreiten der Erkrankung oder bis zum Auftreten einer nicht tolerierbaren Toxizität behandelt werden [11].

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studiename	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studie zum Wirkstoff Tarlatamab				
DeLLphi-304 ¹⁾	Ja	Ja	Ja	-

¹⁾ Zulassungsrelevante Studie gemäß EPAR [13].

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

Die für die Nutzenbewertung herangezogene Studie entspricht derjenigen Studie, auf der die Einschätzungen des pU zum Zusatznutzen basieren.

Zur Nutzenbewertung für Tarlatamab herangezogene Unterlagen

- Dossier des pU zu Tarlatamab [1, 2]
- EU-Dossier / Joint Clinical Assessment (JCA)-Dossier (JCA-MP-2024-17) des pU zu Tarlatamab [8, 9, 10]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [13]
- Klinischer Studienbericht (CSR) inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie DeLLphi-304 [3, 4, 5, 6]
- Fachinformation zu Tarlatamab [11]

Berücksichtigung des JCA-Berichts

Der JCA-Bericht zu Tarlatamab (JCA-MP-2024-17) war zu Beginn des Nutzenbewertungsverfahrens noch nicht veröffentlicht und konnte daher entsprechend § 7 Absatz 1 Nr. 2 Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) für die vorliegende Nutzenbewertung nicht berücksichtigt werden.

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Tarlatamab im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie DeLLphi-304 (NCT05740566). Die Studie und die Intervention werden in Tabelle 2 bis Tabelle 4 charakterisiert.

Tabelle 2: Charakterisierung der Studie DeLLphi-304

Charakteristikum	Beschreibung
Design und Studienablauf	Design Multizentrische, randomisierte, Open-label-Phase-III-Studie¹⁾ bei erwachsenen Patienten mit SCLC und Krankheitsprogression während oder nach einer vorangegangenen platinbasierten Chemotherapie von Tarlatamab vs. Chemotherapie (Topotecan, Lurbinectedin oder Amrubicin)²⁾. Die Randomisierung im 1:1-Verhältnis erfolgte stratifiziert nach³⁾: • Vorherige Exposition mit PD-(L)1-Inhibitoren (ja vs. nein) • Chemotherapie-freies Intervall⁴⁾ (< 90 Tage, ≥ 90 und < 180 Tage, ≥ 180 Tage) • Vorliegen von (früheren oder aktuellen) Hirnmetastasen (ja vs. nein) • SoC (Topotecan/Amrubicin vs. Lurbinectedin) Studienphasen • <u>Screeningphase</u> (bis zu 21 Tage) • <u>Randomisierte Behandlungsphase</u>: Tarlatamab in 28-Tage-Zyklen und SoC in 21-Tage-Zyklen. • <u>Sicherheits-Follow-up (SFU)</u>: Nach dauerhafter Beendigung der Studienmedikation aufgrund jeglicher Ursache wird eine SFU-Visite etwa 60 (+ 5)⁵⁾ Tage nach der letzten Dosis der Studienmedikation durchgeführt. • <u>Langzeit-Follow-up (LTFU)</u>: Das LTFU schließt an die SFU-Visite an. Die Personen wurden mittels Telefon, Besuch in der Klinik oder der Auswertung von Krankenakten nachverfolgt, um den Überlebensstatus sowie nachfolgende Krebstherapien zu erfassen. Nach der primären Analyse war es Personen, die ursprünglich zum SoC-Arm randomisiert wurden (inkl. derer im LTFU), ohne bisherige Exposition mit Tarlatamab möglich auf eine Behandlung mit Tarlatamab zu wechseln. Voraussetzung hierfür war eine Krankheitsprogression unter der aktuellen Therapie.
Population	Wesentliche Einschlusskriterien • ≥ 18 Jahre. • Histologisch bzw. zytologisch bestätigtes SCLC mit nachgewiesener Progression oder Rezidiv. • Krankheitsprogression oder Rezidiv nach platinbasierter Therapie: ◦ Dokumentierte erste Krankheitsprogression nach oder während einer platinbasierten, systemischen Chemotherapie im fortgeschrittenen Stadium (ES) oder begrenzten Stadium (LS). ◦ Personen mit Rezidiv nach der Behandlung für LS. ◦ Personen mit Rezidiv nach Behandlung mit Platin-Etoposid nach Resektion des SCLC. ◦ In Ländern, in welchen die SoC-Erstlinien-Systemtherapie des ES-SCLC eine platinbasierte Chemotherapie in Kombination mit einem PD-(L)1-Inhibitor umfasst, besteht die Voraussetzung, dass die Personen im Rahmen ihrer Erstlinien-Systemtherapie nicht auf einen PD-(L)1-Inhibitor angesprochen haben oder für eine Therapie mit einem PD-(L)1-Inhibitor nicht infrage kommen. • Messbare Erkrankung (gemäß RECIST) innerhalb der (bis zu) 21-tägigen Screeningphase. • ECOG-PS 0 oder 1.

Charakteristikum	Beschreibung
	• Minimale Lebenserwartung von 12 Wochen. • Adäquate Organfunktion (hämatologische, Koagulations-, Nieren-, Pulmonal- und Herzfunktion; gemäß CSP). Wesentliche Ausschlusskriterien • Symptomatische Metastasen des ZNS; ausgenommen: ◦ Asymptomatische Hirnmetastasen ◦ Abschluss der Strahlentherapie oder Operation mind. 2 Wochen vor Gabe der ersten Dosis der Studienmedikation (stereotaktische Radiochirurgie mind. 7 Tage vor Gabe der ersten Dosis der Studienmedikation). ◦ Jegliche ZNS-Erkrankung war klinisch stabil, die Person hatte seit mind. 5 Tagen keine Steroide zur Behandlung der ZNS-Erkrankung mehr eingenommen (außer die Steroidgabe war aus einem anderen Grund als der ZNS-Erkrankung indiziert) und die Person hatte mind. 14 Tage vor der ersten Dosis der Studienbehandlung keine Antiepileptika mehr eingenommen oder eine stabile Dosis davon erhalten. ◦ Zulässig: Personen mit unbehandelten Hirnmetastasen, die asymptomatisch waren und keine Behandlung mit Kortikosteroiden oder eine lokale Therapie benötigten. • Leptomeningeale Erkrankung. • Vorherige Behandlung mit Immun-Checkpoint-Inhibitoren, die zu schwerwiegenden oder lebensbedrohlichen immunvermittelten unerwünschten Ereignissen führten. • Aktive Autoimmunerkrankung, die in den letzten 2 Jahren eine systemische Behandlung (mit Ausnahme einer Substitutionstherapie) erforderlich gemacht hat, oder sonstige Erkrankungen, die während der Teilnahme an der Studie eine immunsuppressive Therapie erfordern. • Organtransplantierte Personen (solide Organe). • Andere maligne Erkrankungen in den vergangenen 24 Monaten (Ausnahmen gemäß CSP). • Myokardinfarkt und/oder symptomatische Herzinsuffizienz (NYHA-Klasse > II) innerhalb von 12 Monaten vor der ersten Dosis der Studienbehandlung. • Arterielle Thrombose innerhalb der letzten 12 Monate vor der ersten Dosis der Studienbehandlung. • Aktive oder zurückliegende Virusinfektion (Ausnahmen gemäß CSP). • Systemische Kortikosteroid- oder andere immunsuppressive Therapie innerhalb von 7 Tagen vor der ersten Dosis der Studienbehandlung (Ausnahmen gemäß CSP). • Akute und/oder unkontrollierte aktive systemische Infektion innerhalb von 7 Tagen vor Gabe der ersten Dosis der Studienmedikation (gemäß Angaben im CSP). • Interstitielle Lungenerkrankung oder aktive, nicht-infektiöse Pneumonitis.
Intervention und Zahl der Patientinnen und Patienten	Gescreent gesamt: N = 688 Randomisiert/Eingeschlossen: N = 509 Intervention: N = 254 Kontrolle: N = 255⁶⁾
Ort und Zeitraum der Durchführung; Datenschnitte	Studienzentren 166 Studienzentren in China, Israel, Japan, Korea, Malaysia, Singapur, Taiwan, Österreich, Belgien, Tschechische Republik, Dänemark, Frankreich, Deutschland, Griechenland, Ungarn, Italien, Niederlande, Polen, Portugal, Rumänien, Spanien, Schweiz, Türkei, Vereinigtes Königreich Großbritannien und Nordirland, Kanada, USA, Argentinien, Australien, Brasilien und Mexiko. Studienzeitraum • Erste Person erste Visite: 31.05.2023 • Einschluss letzte Person in die Studie: 30.07.2024; mit diesem Zeitpunkt war die Rekrutierung abgeschlossen.

Charakteristikum	Beschreibung
	Studienende: Die Studie ist zum Zeitpunkt der Nutzenbewertung nach Angaben in Modul 4 des Dossiers noch nicht abgeschlossen. Die finale Analyse erfolgt nach Angaben im CSP, wenn die Rekrutierung abgeschlossen ist und alle Patientinnen und Patienten die Studie einschließlich des LTFU abgeschlossen haben. Nach Angaben im CSR wurde die Interimsanalyse als primäre Analyse der Studie DeLLphi-304 deklariert, da bereits statistische Signifikanz für den primären Endpunkt erreicht wurde. Laut Studienregister ist das geschätzte Studienende der 26.03.2028 [7]. Datenschnitte Datenschnitt: 29.01.2025 mit Data Snapshot vom 13.03.2025 (überprüfte und bereinigte Patientendaten vom Datenschnitt 29.01.2025); bei diesem Datenschnitt handelt es sich um den Zeitpunkt der geplanten Interimsanalyse des primären Endpunkts Gesamtüberleben. Die Interimsanalyse wurde als primäre Analyse genutzt, da die prädefinierten Erfolgskriterien zu diesem Zeitpunkt bereits erreicht wurden.
Endpunkte gemäß Studienprotokoll	Primärer Endpunkt Gesamtüberleben Sekundäre Endpunkte - Progressionsfreies Überleben - Tumorsprechen (Objektive Ansprechrate, Krankheitskontrollrate, Dauer des Ansprechens) - Patientenberichtete Endpunkte (EORTC QLQ-C30, EORTC QLQ-LC13, PRO-CTCAE⁷⁾, FACT-G, EQ-5D-5L, BPI-SF, PGI-C, PGI-S) - Unerwünschte Ereignisse

¹⁾ Es handelte sich um eine nicht-verblindete Studie. Trotz des offenen Studiendesigns erfolgten keine Datenauswertungen von primären und sekundären Endpunkten vor der Interims- bzw. primären Analyse durch das Studienteam.

²⁾ Die Wahl der Studienmedikation im Chemotherapie-Arm war abhängig von den jeweiligen lokalen Empfehlungen. Weitere Informationen hierzu sind Tabelle 4 der vorliegenden Nutzenbewertung zu entnehmen. Die Entscheidung, welche Chemotherapie verabreicht wird, wurde vor der Randomisierung gemäß den lokalen Bestimmungen getroffen.

³⁾ Der Ausschuss für Humanarzneimittel (CHMP) der Europäischen Arzneimittelagentur (EMA) kritisierte die Verwendung des Vorliegens von Hirnmetastasen, aufgrund der geringen prognostischen Relevanz dieses Faktors in der Zweitlinientherapie, sowie die Art der Chemotherapie (Wirkstoff) anstatt der Region als Stratifikationsfaktoren.

⁴⁾ Anzahl an Tagen, die zwischen der letzten Dosis der vorangegangenen Chemotherapie und dem Zeitpunkt des ersten Verdachts auf ein Wiederauftreten oder Fortschreiten der Erkrankung (klinisch oder radiologisch) verstrichen sind.

⁵⁾ Nach Angabe im Studienprotokoll war eine Erweiterung des Zeitfensters um 5 Tage möglich. Das SFU konnte demnach bis 65 Tage nach der letzten Dosis der Studienmedikation stattfinden.

⁶⁾ In der Kontrollgruppe wurden die folgenden Chemotherapeutika verabreicht: Topotecan (n = 185), Amrubicin (n = 23), Lurbinectedin (n = 47).

⁷⁾ Der pU stellt den Endpunkt „PRO-CTCAE“ im Anhang zu Modul 4 dar, zieht den Endpunkt allerdings nicht in der Zusammenfassung der Ergebnisse bzw. zur Ableitung des Zusatznutzens in der Endpunktkategorie „Sicherheit“ heran. Daher erfolgt keine weitere Berücksichtigung in der vorliegenden Nutzenbewertung.

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; CSP: Klinisches Studienprotokoll; CSR: Klinischer Studienbericht; ECOG-PS: Eastern Cooperative Oncology Group – Performance Status; EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; EORTC QLQ-LC13: EORTC Quality of Life Questionnaire – Lung Cancer Module; EQ-5D-5L: EuroQoL 5-Dimension 5-Level; ES-SCLC: Kleinzelliges Lungenkarzinom im fortgeschrittenen Stadium; FACT-G: Functional Assessment of Cancer Therapy – General; LTFU: Langzeit-Follow-up; NYHA: New York Heart Association; PD-(L)1: Programmed Cell Death (Ligand) 1; PGI-C: Patient Global Impressions scale – Change; PGI-S: Patient Global Impressions scale – Severity; PRO-CTCAE: Patient-Reported Outcomes – Common Terminology Criteria for Adverse Events; pU: pharmazeutischer Unternehmer; RECIST: Response Evaluation Criteria in Solid Tumors; SCLC: Kleinzelliges Lungenkarzinom; SFU: Sicherheits-Follow-up; SoC: Standard of Care; ZNS: Zentrales Nervensystem.

Protokolländerungen

Es wurden 5 Änderungen des Originalprotokolls vom 09.12.2022 vorgenommen, davon 4 vor Erstellung des CSR vom 27.05.2025.

Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie DeLLphi-304

Amendment	Wesentliche Änderungen
Amendment 2 vom 06.09.2023 (Anzahl bis dahin eingeschlossene Personen: k. A.)	- Die statistischen Annahmen der Studie DeLLphi-304 wurden hinsichtlich der Studienpopulation überarbeitet. Statt 700 einzuschließenden Personen wurde die Anzahl auf 490 reduziert, basierend auf den Ergebnissen einer Phase-II-Studie. - Eine der beiden geplanten Interimsanalysen wurden entfernt (Analyse des Gesamtüberlebens bei 60 % der präspezifizierten Ereignisse).
Amendment 3 vom 11.12.2023 (Anzahl bis dahin eingeschlossene Personen: k. A.)	Keine für die Nutzenbewertung als relevant erachteten Änderungen.
Amendment 4 vom 19.07.2024 (Anzahl bis dahin eingeschlossene Personen: k. A.)	- Der Zeitraum des SFU wurde von 42 (+ 5) Tage auf 60 (+ 5) Tage erweitert. - Der Zeitraum der Bildgebung und Bewertung des Tumors wurde auf ± 7 Tage erweitert. - Die Erhebung von Endpunkten alle 6 Wochen (u. a. PRO) wurde um ein Zeitfenster von ± 7 Tagen ergänzt.
Amendment 5 vom 04.09.2025 (Anzahl bis dahin eingeschlossene Personen: k. A.)	Erstellt nach dem für die vorliegende Nutzenbewertung relevanten CSR vom 27.05.2025. Die Änderungen sind für die vorliegende Nutzenbewertung daher nicht relevant.

Abkürzungen: CSR: Klinischer Studienbericht; k. A.: keine Angabe; PRO: Patientenberichteter Endpunkt; SFU: Sicherheits-Follow-up.

Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention

Intervention	Kontrolle
Tarlatamab • Verabreichung von Tarlatamab als intravenöse Infusion über 60 ($\pm$ 10) Minuten mit anschließendem Slow Bolus Flush. • Dosierung und Gabe der Studienmedikation: ◦ 1 mg Tarlatamab zur Dosis-Titration (Step-Dose¹⁾) an Zyklus 1 Tag 1 ◦ 10 mg Zieldosis ab Zyklus 1 Tag 8, Tag 15 und danach alle 2 Wochen ab Zyklus 2 • 28-Tage-Zyklen <u>Prämedikation in Zyklus 1</u> Dexamethason 8 mg intravenös (oder eine äquivalente Dosis eines anderen Kortikosteroids) innerhalb 1 Stunde vor Gabe von Tarlatamab (Zyklus 1 Tag 1 und Tag 8). <u>Postmedikation in Zyklus 1</u> Intravenöse Hydrations-Kochsalzlösung zur Prophylaxe (1 l über 2–4 Stunden) nach der Gabe von Tarlatamab (Zyklus 1 Tag 1 und Tag 8).	Chemotherapie (SoC) Die Gabe der Chemotherapien erfolgte in 21-Tage-Zyklen. Es erfolgte die Gabe eines der 3 folgenden Chemotherapeutika, in Abhängigkeit von lokalen Bestimmungen, mit unterschiedlicher Dosierung: <u>Lurbinectedin</u>²⁾ • Dosierung: 3,2 mg/m² • Tag 1 jedes 21-Tage-Zyklus • Intravenöse Infusion über 60 ($\pm$ 10) Minuten <u>Topotecan</u>³⁾ • Dosierung: 1,5 mg/m² als intravenöse Infusion über 30 ($\pm$ 5) Minuten • Ausnahme China: 1,2 mg/m² als intravenöse Infusion oder 2,3 mg/m²/Tag (oral) an den Tagen 1, 2, 3, 4 und 5 eines jeden 21-Tage-Zyklus <u>Amrubicin</u>⁴⁾ • Dosierung: 40 mg/m² als intravenöse Infusion über 5–15 Minuten an den Tagen 1–3 eines jeden 21-Tage-Zyklus • Nach Maßgabe des Prüfpersonals war die Gabe von G-CSF erlaubt
Ende der Studienbehandlung Die Verabreichung von Tarlatamab bzw. der Chemotherapie erfolgte bis zur radiologischen Krankheitsprogression (gemäß RECIST), zum Auftreten inakzeptabler Toxizität, zum Widerruf der Einwilligung, zum Tod oder zum Studienende, je nachdem, was zuerst eintritt. Unter bestimmten Voraussetzungen war eine Fortführung der Verabreichung der Studienmedikation nach radiologischer Progression möglich. Nicht erlaubte Vor- und Begleitmedikationen • Vorherige Therapie mit Tarlatamab oder mit selektiven DLL3-Inhibitoren oder einer der Standardtherapien (SoC) zur Chemotherapie, die im Rahmen dieser Studie zum Einsatz kommen. • Mehr als eine vorherige systemische Therapie für SCLC. • Krebstherapie (Chemo-, Immun- oder zielgerichtete Therapie während oder innerhalb von 21 Tagen vor Gabe der ersten Dosis der Studienmedikation). • Größere operative Eingriffe innerhalb von 28 Tagen vor der ersten Dosis der Studienmedikation. • Radiotherapie⁵⁾. Erlaubte Begleitmedikationen⁶⁾ • Rescue-Therapie für das Zytokin-Freisetzungssyndrom, einschließlich Kortikosteroide, Tocilizumab oder Siltuximab. • Gemäß CSP erforderliche prophylaktische Dexamethason-Gabe (s. oben) sowie etwaige Antiemetika. • Niedrigdosis-Kortikosteroide (Prednison $\leq$ 10 mg pro Tag oder Äquivalent). • Adjuvante Hormontherapie bei reseziertem Brustkrebs.	

¹⁾ Die Step-Dose wurde zur Reduktion des Risikos des Zytokin-Freisetzungssyndrom verabreicht.

²⁾ Die Verabreichung von Lurbinectedin erfolgte in Kanada, Australien, Korea, Singapur und den USA.

³⁾ Die Verabreichung von Topotecan erfolgte in allen Ländern mit Ausnahme von Japan. Für Abweichungen in China s. Angaben in Tabelle 4.

⁴⁾ Die Verabreichung von Amrubicin erfolgte nur in Japan.

⁵⁾ Eine Ausnahme kann die Radiotherapie zur Symptomkontrolle darstellen (etwa für Knochen- oder Hirnmetastasen). Die Therapie sollte den Brustkorbbereich ausschließen und muss mind. 7 Tage vor der nächsten Tarlatamab-Dosis abgeschlossen sein.

⁶⁾ Begleitmedikationen (einschließlich verschreibungspflichtiger und rezeptfreier Medikamente, pflanzlicher Heilmittel und alternativer Therapien) waren bis zum Ende des SFU zu erheben.

Abkürzungen: CSP: Klinisches Studienprotokoll; DLL3: Delta-like Ligand 3; G-CSF: Granulozyten-Kolonie-stimulierender Faktor; RECIST: Response Evaluation Criteria in Solid Tumors; SCLC: Kleinzelliges Lungenkarzinom; SFU: Sicherheits-Follow-up; SoC: Standard of Care.

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das deutsche Dossier, das europäische JCA-Dossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Alle Endpunkte der eingeschlossenen Studien (siehe Tabelle 2) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle nicht gelistet sind, wurden weder vom pU im deutschen Dossier noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Hinsichtlich des europäischen JCA-Dossiers wurden nur diejenigen Endpunkte berücksichtigt, von denen Ergebnisse vorliegen, die die Fragestellung der Nutzenbewertung gemäß Arzneimittelmarktneuordnungsgesetz (AMNOG) abbilden und vom pU zum Nachweis des Zusatznutzens herangezogen wurden. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie DeLLphi-304

Endpunkt	Kategorie	Ergebnisse im EU-Dossier zum JCA ¹⁾	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Gesamtüberleben ²⁾	Mortalität	Ja	Ja	Ja
Progressionsfreies Überleben	Morbidität	Ja	Ja	Nein ³⁾
Tumoransprechen • Objektive Ansprechrates • Dauer des Ansprechens • Krankheitskontrollrate	Morbidität	Ja	Ja	Nein ³⁾
Allgemeiner Gesundheitszustand mittels EQ-5D-VAS	Morbidität	Ja ⁴⁾	Ja	Nein ⁵⁾
EORTC QLQ-C30 (Symptomskalen)	Morbidität	Ja ⁶⁾	Ja	Ja
EORTC QLQ-LC13	Morbidität	Ja ⁶⁾	Ja	Nein ⁵⁾
PGI-C	Morbidität	Nein ⁷⁾	Ja	Nein ⁵⁾
PGI-S	Morbidität	Nein ⁷⁾	Ja	Nein ⁵⁾
Schmerzen mittels BPI-SF	Morbidität	Nein ⁷⁾	Ja	Nein ⁵⁾
Nebenwirkung der Therapie mittels GP5 (FACT-G)	Morbidität	Nein ⁷⁾	Ja	Nein ³⁾
Gesundheitsbezogene Lebensqualität (EORTC QLQ-C30)	Lebensqualität	Ja ⁸⁾	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja ⁹⁾	Ja	Ja

- ¹⁾ Im Rahmen des JCA-Verfahrens zu Tarlatamab (JCA-MP-2024-17) wurden Daten für insgesamt 7 PICO-Fragestellungen (4 Populationen) eingereicht. In Appendix C [10] befinden sich Auswertungen zur ITT-Population der Studie DeLLphi-304 für den Vergleich von Tarlatamab und SoC. Der Spalteneintrag „Ja“ bezieht sich auf die Darstellung von Ergebnissen für die Gesamtpopulation (ITT-Population) der Studie DeLLphi-304 im JCA-Dossier, die für die nationale Bewertung von Tarlatamab relevant ist. Der Spalteneintrag „Nein“ bezieht sich auf eine fehlende Darstellung von Ergebnissen für die ITT-Population.
- ²⁾ Primärer Endpunkt der Studie DeLLphi-304.
- ³⁾ Der Endpunkt wird aufgrund der fehlenden direkten Patientenrelevanz bzw. nicht-gegebener Validität nicht in der vorliegenden Nutzenbewertung berücksichtigt. Weitere Informationen sind Kapitel 2.3.2 der vorliegenden Nutzenbewertung zu entnehmen.
- ⁴⁾ Ergebnisse zur EQ-5D-VAS wurden in Appendix C des EU-Dossiers für die ITT-Population der Studie DeLLphi-304 dargestellt. Für den Endpunkt wurden kontinuierliche Analysen (Veränderung gegenüber Baseline zu Woche 19) vorgelegt. Kontinuierliche Auswertungen zu PRO werden in der vorliegenden Nutzenbewertung als nicht relevant bewertet.
- ⁵⁾ Aufgrund der Erhebungs- und Datensituation erfolgt keine Berücksichtigung dieser Endpunkte in der vorliegenden Nutzenbewertung. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.
- ⁶⁾ Mit Appendix C [10] reichte der pU im EU-Dossier Auswertungen für „Dyspnoe“, „Husten“ und „Brustschmerzen“ sowie für die restlichen Symptomskalen ein. Das Item „Dyspnoe“ des EORTC QLQ-C30 wurde gemeinsam mit den Items im EORTC QLQ-LC13 als zusammengesetzter Endpunkt ausgewertet. Ferner erfolgte eine Darstellung der Dyspnoe-Symptomskala des EORTC QLQ-LC13 sowie der restlichen Symptomskalen. Als Auswertungsstrategie wurde der Unterschied von Baseline zu Woche 19 nach Baseline genutzt. Kontinuierliche Auswertungen zu PRO sind in der vorliegenden Nutzenbewertung nicht relevant.
- ⁷⁾ Keine Darstellung von Ergebnissen in Appendix C des EU-Dossiers für die ITT-Population der Studie DeLLphi-304.
- ⁸⁾ Mit Appendix C [10] reichte der pU im EU-Dossier Auswertungen für die körperliche Funktionsskala und die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ sowie für die restlichen Funktionsskalen ein. Als Auswertungsstrategie wurde der Unterschied von Baseline bis Woche 18 nach Baseline genutzt. Kontinuierliche Auswertungen zu PRO sind in der vorliegenden Nutzenbewertung nicht relevant.
- ⁹⁾ Mit Appendix C [10] des EU-Dossiers reichte der pU deskriptive Sicherheitsdaten zu UE-Gesamtraten für den Vergleich von Tarlatamab gegenüber SoC ein.

Abkürzungen: BPI-SF: Brief Pain Inventory – Short Form; EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; EORTC QLQ-LC13: EORTC Quality of Life Questionnaire – Lung Cancer Module; EQ-5D-VAS: Visuelle Analogskala des EuroQoL 5-Dimension; FACT-G: Functional Assessment of Cancer Therapy – General; ITT: Intention to Treat; JCA: Joint Clinical Assessment; PGI-C: Patient Global Impressions scale – Change; PGI-S: Patient Global Impressions scale – Severity; PICO: Population, Intervention, Comparison, Outcome; PRO: Patientenberichteter Endpunkt; pU: pharmazeutischer Unternehmer; SoC: Standard of Care.

2.3.1 Mortalität

Gesamtüberleben

Der Endpunkt „Gesamtüberleben“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Das Gesamtüberleben (OS) diente als primärer Endpunkt der Studie DeLLphi-304. Der Endpunkt wurde als Zeit in Monaten vom Datum der Randomisierung bis zum Tod aufgrund jeglicher Ursache folgendermaßen operationalisiert:

$$OS = (\text{Datum Tod} - \text{Datum Randomisierung} + 1) \times 12 / 365,25$$

Die Erfassung erfolgte kontinuierlich über den gesamten Studienverlauf und damit auch über die randomisierte Behandlungsphase hinweg, einschließlich Sicherheits-Follow-up (SFU) und Langezeit-Follow-up (LTFU). Während des LTFU nach der SFU-Visite, erfolgte eine Erhebung des Überlebensstatus alle 12 (± 2) Wochen für maximal 3 Jahre nach Rekrutierung der letzten Person in die Studie oder bis 1 Jahr nach der letzten Behandlung einer Person.

Patientinnen und Patienten, die zum Zeitpunkt des Datenschnitts noch am Leben waren, wurden zum letzten bekannten Zeitpunkt zensiert. Wenn dieses Datum nach dem Datenschnitt lag, wurden die Personen zum Datum des letzten Kontakts bzw. zum Datum des Datenschnitts zensiert.

Bewertung

Die Operationalisierung des Endpunkts „Gesamtüberleben“ ist nachvollziehbar.

Patientenrelevanz

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

Validität

Die Erhebung des Endpunkts „Gesamtüberleben“ in der Studie DeLLphi-304 wird als valide erachtet.

2.3.2 Morbidität

Progressionsfreies Überleben

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

Operationalisierung

Beschreibung

Die Erhebung des Endpunkts „Progressionsfreies Überleben“ (PFS) erfolgte aufgrund des offenen Studiendesigns unverblindet. Das progressionsfreie Überleben war in der Studie DeLLphi-304 definiert als Zeit zwischen Randomisierung bis zur Krankheitsprogression oder Tod, je nachdem, was zuerst eintrat:

$$PFS = (\text{Datum Krankheitsprogression oder Tod} - \text{Datum Randomisierung} + 1) \times 12 / 365,25.$$

Die Krankheitsprogression wurde gemäß der „Response Evaluation Criteria in Solid Tumors“ (RECIST, Version 1.1) festgestellt. Die Beurteilung erfolgte mittels bildgebender Verfahren (MRT oder CT). Zur Reduktion der Strahlenbelastung sollte, wenn möglich, ein Niedrigdosis-CT verwendet werden.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Die Bestimmung des progressionsfreien Überlebens erfolgte anhand RECIST. Die Beurteilung beruht somit ausschließlich auf bildgebenden Verfahren und berücksichtigt keine von Patientinnen oder Patienten unmittelbar wahrnehmbare Symptomatik, etwa mittels klinischer Progressionskriterien. Daher ist die Operationalisierung des Endpunkts „Progressionsfreies Überleben“ nicht direkt patientenrelevant.

Validität

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

Tumoransprechen

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

Operationalisierung

Beschreibung

Der Endpunkt „Tumoransprechen“ wird in der Studie DeLLphi-304 mittels „Objektive Anprechrates“, „Krankheitskontrollrate“ und „Dauer des Ansprechens“ operationalisiert.

Die objektive Anprechrates ist der Anteil der Patientinnen und Patienten, die als bestes Gesamtansprechen (BOR) ein bestätigtes komplettes Ansprechen (CR) oder ein bestätigtes partielles Ansprechen (PR) erreichen.

Die Krankheitskontrollrate ist als der Anteil der Patientinnen und Patienten mit objektiver Anprechrates (CR oder PR) oder einer für mindestens 6 Wochen gemessenen stabilen Erkrankung gemäß RECIST und Prüfpersonal definiert.

Die Dauer des Ansprechens ist definiert als die Zeit vom ersten Nachweis einer CR oder PR (was zuerst dokumentiert wird), bis zum Eintreten einer Krankheitsprogression oder zum Tod aufgrund jeglicher Ursache, je nachdem, was zuerst eintritt. Die Dauer des Ansprechens wird nur für Patientinnen und Patienten berechnet, die ein bestätigtes BOR von PR oder CR erreichten.

Die Beurteilung des Tumors erfolgt durch den Prüfpersonal entsprechend RECIST auf Basis der Befunde bildgebender Verfahren (CT oder MRT). Zur Reduktion der Strahlenexposition wird ein Niedrigdosis-CT empfohlen. Die Erhebungszeitpunkte hatten ein Abweichungsfenster von ± 7 Tagen.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Der Endpunkt „Tumoransprechen“ 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.

Allgemeiner Gesundheitszustand mittels EQ-5D-VAS

Der Endpunkt „EQ-5D-VAS“ wird in der Nutzenbewertung nicht berücksichtigt. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.

Operationalisierung

Beschreibung

Die „Visuelle Analogskala des EuroQoL 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 schlechter Gesundheitszustand“) bezeichnet sind. Der Bezugszeitraum der EQ-5D-VAS ist der Tag der Erhebung. In der Studie DeLLphi-304 wurde die Fragebogen-version EQ-5D-5L genutzt.

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.

Krankheitssymptomatik

Die Krankheitssymptomatik wurde mittels EORTC QLQ-C30 und EORTC QLQ-LC13 sowie PGI-C, PGI-S und BPI-SF erhoben.

EORTC QLQ-C30 – Symptomskalen

Der Endpunkt „EORTC QLQ-C30 – Symptomskalen“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der „Quality of Life Questionnaire – Core Questionnaire“ (QLQ-C30) der „European Organisation for Research and Treatment of Cancer“ (EORTC) ist ein generisches Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität und der Symptome bei an Krebs erkrankten Personen. Der Fragebogen gliedert sich in:

- 5 Funktionsskalen
 - Körperliche Funktion (5 Items)
 - Rollenfunktion (2 Items)
 - Emotionale Funktion (4 Items)
 - Kognitive Funktion (2 Items)
 - Soziale Funktion (2 Items)
- 3 Symptomskalen
 - Fatigue (3 Items)
 - Schmerz (2 Items)
 - Übelkeit und Erbrechen (2 Items)
- 1 globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ (2 Items)

Zusätzlich umfasst der Fragebogen einzelne Items zu Symptomen (Dyspnoe, Appetitverlust, Schlaflosigkeit, Obstipation, Diarrhö) und zu finanziellen Beeinträchtigungen.

Die Fragen werden auf 4- bzw. 7-Punkte-Skalen beantwortet, Bezugszeitraum ist die vergangene Woche. Gemäß Manual der EORTC werden der Gesamtscore und die Scores für Funktions- und Symptomskalen sowie die Einzelskalen auf eine Skala von 0 bis 100 Punkten transformiert.

Höhere Werte auf den Funktionsskalen und der globalen Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ bedeuten eine bessere Funktion bzw. Gesundheit/Lebensqualität. Höhere Werte auf den Symptomskalen bzw. den einzelnen Symptom-Items bedeuten eine schwerere Symptomatik.

Bewertung

Die Operationalisierung ist nachvollziehbar. Schmerzen werden unter anderem mit dem BPI-SF sowie dem EORTC QLQ-LC13 erhoben. Das Symptom „Dyspnoe“ wird darüber hinaus mit dem PGI-S, PGI-C sowie dem EORTC QLQ-LC13 erhoben.

Patientenrelevanz

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

Validität

Beim EORTC QLQ-C30 handelt es sich um ein etabliertes und validiertes generisches Messinstrument zur Untersuchung der Symptome und gesundheitsbezogenen Lebensqualität bei an Krebs erkrankten Personen.

EORTC QLQ-LC13

Der Endpunkt „EORTC QLQ-LC13“ wird in der Nutzenbewertung nicht berücksichtigt. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.

Operationalisierung

Beschreibung

Der „Quality of Life Questionnaire – Lung Cancer Module“ (QLQ-LC13) der EORTC ist ein krankheits-spezifisches Instrument für Lungenkrebserkrankungen, welches sowohl krankheitsspezifische Symptome (z. B. Husten, Hämoptyse, Dyspnoe und Schmerz) als auch behandlungsbedingte Symptome (z. B. Alopezie, wunder Mund, Dysphagie) umfasst. Es handelt sich um eine modulare Erweiterung des EORTC QLQ-C30.

Der Fragebogen beinhaltet 13 lungenkrebsspezifische Items und gliedert sich in

- 1 Symptomskala „Dyspnoe“ (3 Items; die Auswertung dieser Symptomskala sollte nur erfolgen, wenn alle 3 Items beantwortet wurden)
- 9 einzelne Symptom-Items
 - Husten
 - Hämoptyse
 - wunder Mund
 - Dysphagie
 - periphere Neuropathie
 - Alopezie
 - Schmerzen in der Brust
 - Schmerzen im Arm oder in der Schulter
 - Schmerzen an anderen Stellen des Körpers
- 1 Item zur Erfassung der Schmerzlinderung durch Schmerzmedikation
Frage nach Einnahme von Schmerzmedikation in der vergangenen Woche (ja – nein). Falls diese Frage positiv beantwortet wird, soll auf einer 4-stufigen Skala beantwortet werden, wie sehr die Medikation geholfen hat.

Die Fragen werden (mit Ausnahme der Frage nach Schmerzen in der vergangenen Woche) auf einer 4-Punkte-Skala beantwortet, der Bezugszeitraum ist die vergangene Woche. Gemäß Manual werden die Scores für die Symptomskala „Dyspnoe“ sowie die einzelnen Symptom-Items – ähnlich wie beim EORTC QLQ-C30 – auf eine Skala von 0 bis 100 Punkten transformiert. Höhere Werte auf der Symptomskala bzw. den einzelnen Symptom-Items bedeuten eine schwerere Symptomatik. Die Auswertung des Items zur Schmerzlinderung durch Schmerzmedikation ist optional, hier bedeuten höhere Werte ein höheres Maß an Schmerzreduktion.

Bewertung

Die Operationalisierung ist nachvollziehbar. Das Symptom „Husten“ wird unter anderem bereits mit dem PGI-S und PGI-C erhoben. Die Symptome „Schmerzen“ und „Dyspnoe“ werden zudem bereits mit dem PGI-S und PGI-C sowie dem EORTC QLQ-C30 und BPI-SF erfasst. Nach Angaben im PRO-SAP wird zudem „Dyspnoe“ als zusammengesetzter Endpunkt (bestehend aus einem Item des EORTC QLQ-C30 sowie den 3 Items des EORTC QLQ-LC13) ausgewertet. Im Rahmen des Dossiers werden zwei getrennte Auswertungen für den EORTC QLQ-C30 und den EORTC QLQ-LC13 vorgelegt, wodurch eine entsprechende Doppelerfassung (Multi-Item vs. Single-Item) auch zwischen den beiden EORTC-Fragebögen anzumerken ist.

Patientenrelevanz

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

Validität

Die Validität des Endpunkts „EORTC QLQ-LC13“ wurde in vorherigen Nutzenbewertungsverfahren bereits in vergleichbaren Anwendungsgebieten bewertet, die Validierung des Endpunkts ist auf das vorliegende Anwendungsgebiet übertragbar.

PGI-C

Der Endpunkt „PGI-C“ wird in der Nutzenbewertung nicht berücksichtigt. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.

Operationalisierung

Beschreibung

In der Studie DeLLphi-304 wird mittels des „Patient Global Impressions scale – Change“ (PGI-C) erfragt, wie sich 3 spezifische Symptome (Husten, Dyspnoe und Brustschmerzen) global seit Initiierung der Therapie verändert haben. Für jedes der Symptome wurde ein eigenes Fragebogen-Item genutzt. Die Antwortoptionen wurden 5-stufig operationalisiert: „viel besser“, „ein bisschen besser“, „etwa gleich“, „ein bisschen schlechter“, „viel schlechter“.

Bewertung

Die Operationalisierung ist nachvollziehbar. Bezugnehmend auf die Abfrage der krankheits-spezifischen Symptomatik (Husten, Dyspnoe und Brustschmerzen) ist eine Doppelerfassung zum PGI-S anzumerken.

Patientenrelevanz und Validität

Die Einschätzung der Patientinnen und Patienten zur Veränderung der jeweils eigenen erkrankungs-spezifischen Symptomatik ist grundsätzlich als patientenrelevant zu bewerten. Methodisch werden dabei i. d. R. Instrumente bevorzugt, die sowohl zu Baseline- als auch zu Post-Baseline-Studienvisiten erhoben werden. Veränderungseinschätzungen gegenüber Baseline ohne eigenständige Baseline-Erhebung können demgegenüber unter Umständen weniger aussagekräftige Ergebnisse liefern. Es gibt Hinweise darauf, dass bei Einschätzungen von Veränderungen größtenteils der gegenwärtige Gesundheitszustand erfasst wird. Dieser Effekt wird umso größer, je länger das Erinnerungsintervall ist [15, 16, 17]. Die mit dem PGI-C erfassten Symptome „Husten“,

„Dyspnoe“ und „Brustschmerzen“ werden gleichzeitig durch den PGI-S erfasst, für den die Veränderung zu Baseline auf Basis von tatsächlichen Erhebungen erfasst werden.

PGI-S

Der Endpunkt „PGI-S“ wird in der Nutzenbewertung nicht berücksichtigt. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.

Operationalisierung

Beschreibung

Der „Patient Global Impressions scale – Severity;“ (PGI-S) dient zur Erfassung der Symptomstärke von 3 Symptomen aus Perspektive der Patientinnen und Patienten: Husten, Dyspnoe und Brustschmerzen. Die Operationalisierung erfolgte mittels 3 Einzelitems. Der Bezugszeitraum ist die vergangene Woche. Für jedes der Symptome wurde ein eigenes Fragebogen-Item genutzt. Die Antwortoptionen wurden 4-stufig operationalisiert: „keine“, „mild“, „moderat“, „stark“.

Bewertung

Die Operationalisierung ist nachvollziehbar. Bezüglich der spezifischen Symptome „Husten“, „Dyspnoe“ und „Brustschmerzen“ ist eine Doppelerfassung mit dem EORTC QLQ-LC13 und dem PGI-C anzumerken. Zudem werden Schmerzen mit dem vollumfänglichen Fragebogen BPI-SF erhoben.

Patientenrelevanz

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

Validität

Die Validität des Endpunkts „PGI-S“ hinsichtlich der allgemeinen Symptome wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und als valide eingeschätzt.

Schmerzen mittels BPI-SF

Der Endpunkt „BPI-SF“ wird in der Nutzenbewertung nicht berücksichtigt. Weitere Informationen sind Kapitel 2.4 der vorliegenden Nutzenbewertung zu entnehmen.

Operationalisierung

Beschreibung

Der Fragebogen „Brief Pain Inventory – Short Form“ (BPI-SF) umfasst 15 Items, von denen 11 in die beiden Domänen „Schmerzintensität“ (4 Items) und „Beeinträchtigung durch den Schmerz“ (7 Items) aufgeteilt werden können.

Der Bezugszeitraum umfasst die letzten 24 Stunden. Die Items werden auf einer 11-Punkte-Skala von 0 („keine Schmerzen“ bzw. „keine Beeinträchtigung“) bis 10 („schlimmste vorstellbare Schmerzen“ bzw. „stärkste Beeinträchtigung“) beantwortet.

Bewertung

Die Operationalisierung ist nachvollziehbar. Der pU legt mit dem Dossier Auswertungen für die Einzelitems vor. Für die vorliegende Nutzenbewertung sind die Gesamtwerte (Fragen 3–6 des BPI-SF sowie die daraus ermittelte mittlere Schmerzintensität und die Beeinträchtigung durch den Schmerz innerhalb der letzten 24 Stunden (Fragen 9 A–G, einzeln und als Gesamtwert)) relevant, da die Einzelitems in den mittleren Gesamtwerten enthalten sind. Schmerzen werden durch weitere Instrumente innerhalb der Studie DeLLphi-304 erfasst, darunter der EORTC QLQ-C30, EORTC QLQ-LC13 sowie PGI-S/PGI-C. Der BPI-SF erfasst das Konstrukt „Schmerz“ vollumfänglich und bezieht dabei sowohl die Symptomatik als auch die dadurch entstehende Beeinträchtigung mit ein.

Patientenrelevanz

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

Validität

Die Validität des Endpunkts „BPI-SF“ wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt.

Nebenwirkung der Therapie mittels „FACT-G GP5

Der Endpunkt „FACT-G GP5“ wird in der Nutzenbewertung aufgrund nicht-gegebener Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Der „Functional Assessment of Cancer Therapy – General“ (FACT-G) ist ein Fragebogen mit 27 Items, mit dem 4 Domänen gesundheitsbezogener Lebensqualität gemessen werden sollen: physisches, soziales, emotionales und funktionales Wohlbefinden.

Im Rahmen der Studie DeLLphi-304 wird nur das Einzelitem „GP5“ erhoben: „I am bothered by side effects of treatment“ (Deutsch: „Die Nebenwirkungen der Behandlung beeinträchtigen mich.“). Dieses wird mit einer 5-Punkt-Likert-Skala von „überhaupt nicht“ bis „sehr“ beantwortet. Das Einzelitem stammt aus der Subskala „Physisches Wohlbefinden“.

Bewertung

Die Operationalisierung ist noch vollziehbar.

Patientenrelevanz

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

Validität

Die Validität des gesamten FACT-G wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird als valide eingeschätzt. Allerdings konnten keine Untersuchungen identifiziert werden, in denen nur das Item „GP5“ einzeln auf seine inhaltliche Validität sowie die weiteren psychometrischen Gütekriterien hin untersucht wurde. Dementsprechend ist unklar, ob mit dem Endpunkt „FACT-G GP5“ das gesamte Nebenwirkungsspektrum einer Therapie valide erfasst werden kann.

Die Erhebung der Belastung durch Nebenwirkungen der Therapie über ein Einzelitem ist aus inhaltlichen Erwägungen nicht sinnvoll. Für die einzelne Patientin oder den einzelnen Patienten ist häufig nicht zu unterscheiden, ob sich die Belastung im Einzelfall auf Nebenwirkungen der Therapie zurückführen lässt oder auf eine andere Ursache, wie die Symptomatik der Grunderkrankung. Eine einheitliche Erhebung der Belastung durch Nebenwirkungen der Therapie ist aus diesem Grund nicht möglich. Die vom pU vorgelegten Auswertungen auf Basis der Einzelfrage zur Belastung durch Nebenwirkungen der Therapie sind für die Nutzenbewertung demnach nicht geeignet.

2.3.3 Lebensqualität

EORTC-QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus

Die Funktionsskalen und die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 werden in der Nutzenbewertung berücksichtigt.

Beschreibung der Operationalisierung sowie Bewertung der Validität und Patientenrelevanz siehe Kapitel 2.3.2 (Abschnitt „EORTC QLQ-C30 – Symptomskalen“).

2.3.4 Sicherheit

Unerwünschte Ereignisse

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

Operationalisierung

Beschreibung

Ein unerwünschtes Ereignis (UE) war definiert als jegliches unerwünschte medizinische Ereignis (Erkrankung, Symptom oder abnormaler Laborbefund), unabhängig von einem kausalen Bezug zur Studienmedikation. In die Auswertung wurden ausschließlich „Treatment-Emergent Adverse Events“ (TEAE) einbezogen, d. h. Ereignisse, die mit Beginn oder nach der ersten Dosis jeglicher Studienmedikation auftraten.

Die Erfassung von UE erfolgt bei allen Personen, die mindestens eine Dosis der Studienmedikation erhalten haben. Es wurden diejenigen Ereignisse als UE definiert, die nach Gabe der ersten Dosis der Studienmedikation (bzw. auch ein Nicht-Prüfpräparat oder Hilfspräparat), im SFU oder bis 60 Tage (bzw. bis zu 65 Tage) nach Gabe der letzten Dosis der Studienmedikation (bzw. auch ein Nicht-Prüfpräparat oder Hilfspräparat, je nachdem, was später lag) auftrat. Die Erhebung im SFU wurde auch durchgeführt, wenn zwischenzeitlich eine Folgetherapie initiiert wurde. Wenn es einer Person nicht möglich ist an der SFU-Visite teilzunehmen, sollte zumindest eine Kontaktaufnahme mittels Telefon, Besuch in der Klinik oder Auswertung der Krankenakte erfolgen, um den Überlebensstatus sowie den Beginn von Folgetherapien zu erfassen.

In der Studie DeLLphi-304 wurden UE nach Angaben in Modul 4 gemäß „Medical Dictionary for Regulatory Activities“ (MedDRA, Version 27.1) nach Systemorganklasse und Preferred Term kodiert; im SAP wurde spezifiziert, dass Version 25.0 oder höher genutzt werden sollte. Der Schweregrad der UE wurde anhand der „Common Terminology Criteria for Adverse Events“ (CTCAE, Version 5.0) eingestuft.

Schwerwiegende unerwünschte Ereignisse (SUE) waren definiert als jegliches unerwünschte medizinische Ereignis, das

- tödlich war;
- unmittelbar lebensbedrohlich war;
- eine Hospitalisierung oder eine Verlängerung dieser erforderlich machte;
- zu einer dauerhaften oder erheblichen Behinderung/Arbeitsunfähigkeit führte;
- eine angeborene Anomalie / einen Geburtsfehler darstellte.

Für SUE erfolgte ein erweitertes Follow-up (LTFU) von bis zu 3 Jahren nach Rekrutierung der letzten Person oder bis zu 1 Jahr nach der letzten Behandlung einer Person, je nachdem, welcher Zeitpunkt später lag.

Nach Angaben in Modul 4 sowie im SAP gingen erkrankungsbezogene Ereignisse nicht in die Erhebung der UE ein. Dies waren Ereignisse, die direkt mit der Grunderkrankung des Lungenkarzinoms bzw. dessen Progression zusammenhingen. Entsprechende Preferred Terms, die mit der Grunderkrankung in Verbindung stehen, wurden aus der TEAE-Analyse ausgeschlossen. In Modul 4 wurden die folgenden Ereignisse als erkrankungsbezogene Ereignisse definiert: Krankheitsprogression; kleinzelliges Lungenkarzinom (SCLC); ES-SCLC; metastasiertes SCLC; kleinzelliges Karzinom; metastasiertes Lungenkarzinom; Metastasen im zentralen Nervensystem (ZNS); Metastasen in den Hirnhäuten; Metastasen in der Hypophyse; Metastasen im Rückenmark; Metastasen in der Wirbelsäule.

Bewertung

Die Operationalisierung des Endpunkts ist im Grundsatz nachvollziehbar. Die Nutzung von TEAE kann grundsätzlich als sachgerecht bewertet werden. Der Zeitpunkt der Erhebung ab Gabe der ersten Dosis der Studienmedikation ist als adäquat zu bewerten, da zwischen Randomisierung und Initiierung der Therapie maximal 3 Kalendertage vergehen sollten.

Im Studienprotokoll erfolgt keine abschließende, sondern nur eine exemplarische Darstellung von UE (Preferred Terms), die aufgrund ihrer direkten Assoziation mit der Grunderkrankung „Lungenkrebs“ nicht in die TEAE-Analyse eingehen. Mit Modul 4 werden diese erst post hoc näher spezifiziert. Es ist unklar, wie die Auswahl der Preferred Terms erfolgte. Die Auswahl erscheint plausibel, um Progressionsereignisse abzubilden, jedoch werden Symptome der Grunderkrankung damit nicht erfasst. Es kann daher nicht ausgeschlossen werden, dass weiterhin Ereignisse der Grunderkrankung in den Ergebnissen zu den UE enthalten sind.

Patientenrelevanz

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

Validität

Die Erhebung des Endpunkts wird als valide erachtet.

2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 6.

Tabelle 6: Erhebungszeitpunkte der berücksichtigten Endpunkte in der Studie DeLLphi-304

Studienvisite Endpunkt	Screening (21 Tage)	Randomisierte Behandlungsphase		Follow-up-Phase	
		Studien- behandlung ¹⁾	EOT	SFU ²⁾	LTFU ³⁾
Gesamtüberleben ⁴⁾		x ⁵⁾			
EORTC QLQ-C30	x	x ⁶⁾	x	x	
Unerwünschte Ereignisse		x ⁷⁾			x ⁸⁾

¹⁾ Während der randomisierten Behandlungsphase erfolgt nach Tag 1 von Zyklus 1 eine Erhebung der patientenberichteten Endpunkte alle 6 Wochen (± 7 Tage). Eine Ausnahme stellt der EORTC QLQ-C30 dar.

²⁾ Nach dauerhafter Beendigung der Studienmedikation aufgrund jeglicher Ursache wird eine SFU-Visite etwa 60 (+ 5) Tage nach der letzten Dosis der Studienmedikation durchgeführt; unabhängig davon, ob zwischenzeitlich eine Folgetherapie (Anit-Krebstherapie) initiiert wurde.

³⁾ Während des LTFU nach der SFU-Visite erfolgte eine Erhebung des Überlebensstatus alle 12 (± 2) Wochen für maximal 3 Jahre nach Rekrutierung der letzten Person in die Studie oder bis 1 Jahr nach der letzten Behandlung einer Person, je nachdem, welcher Zeitpunkt später lag.

⁴⁾ Primärer Endpunkt der Studie DeLLphi-304.

⁵⁾ Kontinuierliche Erhebung.

⁶⁾ Wöchentliche Erhebung des EORTC QLQ-C30 in der randomisierten Behandlungsphase bis Zyklus 4 Tag 1 (Taratamab) bzw. die ersten 12 Wochen (Chemotherapie).

⁷⁾ Kontinuierliche Erhebung.

⁸⁾ SUE, die während des LTFU aufgetreten sind, wurden weiterhin gemeldet.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; EOT: End of Treatment; LTFU: Langzeit-Follow-up; SFU: Sicherheits-Follow-up; SUE: Schwerwiegendes unerwünschtes Ereignis.

2.4 Statistische Methoden

Grundlage der Auswertung der Studie DeLLphi-304 ist der SAP vom 10.05.2023 mit insgesamt 5 Änderungen, zuletzt am 01.10.2025. Nur 4 der Änderungen sind für die vorliegende Nutzenbewertung relevant, da die 5. Änderung nach dem relevanten Datenschnitt durchgeführt wurde.

Für die Auswertung von patientenberichteten Endpunkten (PRO) lag ein separater SAP (Version 2.0) vom 16.01.2025 vor; die ursprüngliche Version des PRO-SAP ist vom 25.10.2024. Damit wurde dieser erst nach Einschluss der letzten Person in die Studie (30.07.2024) und somit im späteren Studienverlauf erstellt.

Analysepopulationen

- Intention-to-Treat (ITT)-Population: Alle randomisierten Personen. Die Analyse basierend auf der ITT-Population erfolgt entsprechend der randomisierten Zuordnung zu einem der beiden Behandlungsarme. Die ITT-Analysen werden für die primären und sekundären Wirksamkeitsendpunkte genutzt.
- Sicherheitspopulation: Alle Personen, die mindestens eine Dosis der Studienmedikation erhalten haben. Die Analysen erfolgen basierend darauf, welche Studienmedikation die Personen tatsächlich erhalten haben.

Datenschnitte

Der für die vorliegende Nutzenbewertung zugrundeliegende Datenschnitt basiert auf der ersten geplanten Interimsanalyse zum Gesamtüberleben vom 29.01.2025. Aufgrund des Erreichens vorab festgelegter Kriterien diente diese als primäre Analyse der Studie DeLLphi-304.

Präspezifizierte Subgruppenanalysen

Neben den Stratifikationsfaktoren, die zu Randomisierungszwecken genutzt wurden, wurden folgende Subgruppen für den primären sowie ausgewählte sekundäre Endpunkte präspezifiziert:

- Region (Nordamerika und Europa vs. Rest der Welt)
- Alter (< 65 Jahre vs. ≥ 65 Jahre)
- Geschlecht (männlich vs. weiblich)
- Ethnischer Hintergrund (weiß, schwarz, asiatisch, andere)

Neben den zuvor genannten Subgruppen wurden zusätzlich Subgruppenanalysen für „Krankheitsstadium zu Screening“ durchgeführt. Subgruppen wurden bei unzureichender Anzahl an Personen (< 10) in einer Subgruppe entweder kombiniert, wenn möglich, oder andernfalls von der Analyse ausgeschlossen. Laut SAP konnten zudem Biomarker-Daten in zusätzliche explorative Subgruppen- oder multivariate Analysen einbezogen werden.

Zugehörige Interaktionsterme reicht der pU mit Modul 4 ein, wobei p-Werte unter 0,05 als signifikant anzusehend sind. Die zweiseitigen Interaktions-p-Werte der Time-to-Event-Analysen basieren auf einem Cox-Regressionsmodell mit Merkmal zu Baseline, Behandlung sowie Interaktionsterm (Merkmal x Behandlung) als Kovariaten. Subgruppen mit signifikantem Interaktionsterm werden in der Nutzenbewertung berücksichtigt, insofern diese als relevant erachtet werden.

Präspezifizierung geplanter und durchgeführter Analysen

Im SAP wurden eine Interimsanalyse, primäre Analyse und finale Analyse präspezifiziert. Entsprechend des Datenschnitts vom 29.01.2025 ist die Interimsanalyse, die als primäre Analyse genutzt wurde, für die vorliegende Nutzenbewertung maßgeblich. Die vergleichenden Auswertungen basieren auf einem Vergleich des Tarlatamab-Arms vs. Chemotherapie, unabhängig vom eingesetzten Wirkstoff (Topotecan, Amrubicin oder Lurbinectedin).

Gesamtüberleben

Die ursprünglich geplante primäre Analyse des primären Endpunkts „Gesamtüberleben“ war ereignisgesteuert und für den Zeitpunkt vorgesehen, zu dem insgesamt 345 Todesfälle beobachtet wurden. Diese Anzahl bezieht sich kumulativ auf beide Behandlungsgruppen.

Eine vorgelagerte Interimsanalyse sollte bei 259 OS-Ereignissen erfolgen, was 75 % der für die finale OS-Analyse vorgesehenen Ereignisse entspricht. Wenn die präspezifizierte OS-Wirksamkeitsschwelle früh erreicht wird, dient die Interimsanalyse als primäre Analyse des Gesamtüberlebens. Die OS-Analyse zum Zeitpunkt der Interims- und primären Analyse wird nach der Lan-DeMets-Alpha-Spending-Funktion mit einem O'Brien-Fleming-Ansatz adjustiert. Für den Fall, dass die Interimsanalyse des Endpunkts „Gesamtüberleben“ statistisch signifikant war, erfolgte die hierarchische Testung der sekundären Endpunkte („Progressionsfreies Überleben“ und PRO).

Primäre inferenzstatistische Vergleiche des Gesamtüberlebens zwischen dem Tarlatamab- und dem Chemotherapie-Arm erfolgen basierend auf der ITT-Population. Es wurde ein zweiseitiger, stratifizierter Log-Rank-Test genutzt; p-Werte von 0,05 wurden als signifikant angesehen. Laut SAP werden die Faktoren der stratifizierten Randomisierung genutzt. Nach Angaben im CSR wurde ausschließlich nach „Chemotherapie-freies Intervall“ und „Vorliegen von Hirnmetastasen“ stratifiziert, um eine Überstratifizierung zu vermeiden. Damit wurden 2 der 4 Randomisierungsstrata genutzt. Es wurde keine Begründung für die Auswahl der beiden Faktoren vorgelegt.

Das Hazard Ratio (HR) als Effektmaß sowie das zugehörige 95%-Konfidenzintervall (KI) wurde mittels eines stratifizierten Cox-Proportional-Hazard-Modells geschätzt. Der Median, die Quartile sowie das zugehörige 95%-KI werden mittels Kaplan-Meier-Methode berechnet. Die Follow-up-Dauer wird ebenfalls mittels Kaplan-Meier-Methode geschätzt.

Patientenberichtete Endpunkte

PRO wurden als sekundäre Endpunkte der Studie DeLLphi-304 ausgewertet. Die Analysen basieren auf der ITT-Population und Rücklaufquoten (Compliance Rates bzw. Completion Rates basierend auf den Personen, für die eine Erhebung per Studienprotokoll vorgesehen war, bzw. der ITT-Population) wurden für jede Visite, zu welcher ein Endpunkt erhoben wurde, in den Studienunterlagen dargestellt.

Der pU legt mit Modul 4 post hoc durchgeführte Responderanalysen für die Zeit bis zur ersten Verschlechterung operationalisiert als Ereigniszeitanalysen (Time-to-Event) vor. Statistische Outputs (außer der Subgruppenanalysen in Anhang G) wurden zu diesen Post-hoc-Auswertungen nicht durch den pU eingereicht. Für die Endpunkte „EORTC QLQ-C30“ und „EORTC QLQ-LC13“ wurde die adäquate Responderschwelle von 10 Punkten genutzt. Für die PRO „EQ-5D-VAS“, „PGI-S“, „PGI-C“ und „BPI-SF“ wurde 15 % der Skalenspannweite als Responderschwelle genutzt bzw. eine 15-Punkte-Schwelle nach entsprechender Transformation auf eine Skala von 100. Die verwendeten Responderschwellen sind als adäquat anzusehen. Mit dem PRO-SAP hatte der pU Analysen für die Zeit bis zur Verschlechterung für einzelne Skalen präspezifiziert, allerdings mit abweichenden Responderschwellen. Diese sind für die vorliegende Nutzenbewertung nicht maßgeblich.

Die Responderanalysen für die „Zeit bis zur ersten Verschlechterung“ werden mittels einer Überlebenszeitanalyse (Kaplan-Meier-Methode) ausgewertet. HR als Effektmaß sowie das zugehörige 95%-KI wurden mittels eines stratifizierten Cox-Proportional-Hazard-Modells geschätzt. In Modul 4 wurden für die PRO die folgenden Zensierungsregeln dargestellt; diese waren deckungsgleich mit den Zensierungsregeln des separaten PRO-SAP:

- Der Tod wurde als Ereignis gewertet, wenn keine Verschlechterung im Vergleich zu Studienbeginn beobachtet werden konnte, aber die Person vor oder an der letzten PRO-Beurteilung verstorben sind.
- Keine Verschlechterung im Vergleich zu Studienbeginn und Personen verstarben nicht vor oder an der letzten PRO-Beurteilung: Zensierung zum Zeitpunkt der zuletzt verfügbaren PRO-Erhebung.
- Keine Verschlechterung im Vergleich zu Studienbeginn und Personen verstarben nach 2 oder mehr aufeinanderfolgenden fehlenden PRO-Erhebungen: Zensierung zum Zeitpunkt der zuletzt verfügbaren PRO-Erhebung.

Der pU beschreibt, dass für die vorliegende Nutzenbewertung der Tod nicht als Ereignis gewertet wurde. Unter dieser Annahme können die genutzten Zensierungsregeln als angemessen bewertet werden.

Die Rücklaufquoten der PRO nehmen über die Behandlungszyklen rapide ab. Zu späteren Erhebungszeitpunkten liegen stark divergierende Rückläufe zwischen den Studienarmen vor, mit deutlich niedrigeren Rückläufen im Vergleichsarm. Durch den erwarteten Krankheitsverlauf und nur einer geplanten Erhebung nach Abbruch der Studienmedikation kann eine informative Zensierung nicht ausgeschlossen werden.

Der Endpunkt „EORTC QLQ-C30“ wird über die ersten 12 Wochen wöchentlich erhoben, wodurch Verschlechterungsereignisse engmaschig erfasst werden. Die verbleibenden PRO werden in der Studie DeLLphi-304 nur alle 6 Wochen nach Beginn der Therapie erhoben. Entsprechend der Beschreibung in Kapitel 2.3.2 der vorliegenden Nutzenbewertung sind die verbleibenden PRO (EORTC QLQ-LC13, EQ-5D-VAS, BPI-SF, PGI-S, PGI-C) grundsätzlich geeignet, die subjektive Morbidität zu erheben. Allerdings ist nicht auszuschließen, dass in der Zeitspanne, in der die Endpunkte nicht erhoben werden, Verschlechterungsereignisse stattfinden, die aufgrund der Erhebungssituation nicht erfasst werden (6 Wochen zwischen den Erhebungen). Für diese PRO liegen die Rücklaufquoten meist bereits zum ersten bzw. zweiten Post-Baseline-Erhebungszeitpunkt bei < 70 %, bezogen auf die ITT-Population. Aufgrund der Erhebungs- und Datensituation ergeben sich wesentliche Limitationen, sodass diese PRO nicht in der vorliegenden Nutzenbewertung berücksichtigt werden.

Ferner reicht der pU mit Anhang G zu Modul 4 kontinuierliche Auswertungen mittels „Mixed Model for Repeated Measures“ (MMRM) basierend auf der ITT-Population ein. Initial hatte der pU kontinuierliche Auswertungen zur Veränderung von Baseline zu Woche 18 mittels MMRM bzw. „Generalized Linear Mixed Model for Repeated Measures“ präspezifiziert. Diese sind aufgrund der Datensituation für die vorliegende Nutzenbewertung nicht relevant.

Sicherheit

Deskriptive Auswertungen zur Sicherheit wurden sowohl für Tarlatamab vs. Chemotherapie als auch für die einzelnen Kontrollbehandlungsgruppen Topotecan, Lurbinectedin und Amrubicin durchgeführt. Die Analysen basieren auf der Sicherheitspopulation der Studie DeLLphi-304. In der vorliegenden Nutzenbewertung wird, entsprechend der Einreichung mit Modul 4, die gesamte Kontrollgruppe ohne Differenzierung nach Art der Chemotherapie betrachtet, da nur hierfür vergleichende Auswertungen vorliegen.

Nach Angaben in Modul 4 erfolgte die Auswertung des Endpunkts „Unerwünschte Ereignisse“ anhand einer Überlebenszeitanalyse basierend auf der Kaplan-Meier-Methode. Als Effektschätzer dient HR, allerdings wird nicht spezifiziert, welches Modell zur Ableitung des HR genutzt wurde. Analog zu den anderen Ereigniszeitanalysen wird davon ausgegangen, dass ein Cox-Proportional-Hazard-Modell verwendet wurde. Die 95%-KI basieren auf der geschätzten Varianz der log-log-Transformation der Kaplan-Meier-Schätzung. Der Gruppenvergleich basierte auf einem Log-Rank-

Test. Dieser wurde unstratifiziert durchgeführt. Nach Angaben in Modul 4 wurden Personen, bei denen zum Zeitpunkt des Datenschnitts (29.01.2025) kein Ereignis beobachtet wurde, frühestens zum Zeitpunkt der Rücknahme der Einwilligung, Lost-to-Follow-up, Behandlungsende + 60 Tage (+5 Tage), dem Datum des Datenschnitts oder dem Todesdatum zensiert. Das Vorgehen der Zensierung lässt sich als adäquat bewerten. Im Rahmen des CSR erfolgte ausschließlich eine deskriptive Darstellung der UE entsprechend der Präspezifizierung durch den SAP. Die Ereigniszeitanalysen, die mit dem nationalen Dossier eingereicht wurden, werden daher als Post-Hoc-Analysen eingeordnet.

Aufgrund unterschiedlicher medianer Beobachtungsdauern für UE zwischen Tarlatamab-Arm (6,31 Monate (min; max: 0,1; 17,3) und Chemotherapie-Arm (4,42 Monate (min; max: 0,1; 15,5) werden die Ereigniszeitanalysen mit dem vergleichenden Effektmaß HR als adäquat erachtet.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 7: Verzerrungspotential der Studie DeLLphi-304

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
DeLLphi-304	Ja	Ja	Nein ¹⁾	Nein ¹⁾	Nein	Nein	Niedrig ²⁾

¹⁾ Bei der Studie DeLLphi-304 handelte es sich um eine Open-label-Studie. Infolgedessen waren weder die Teilnehmenden der Studie noch das Studienpersonal verblindet. Trotz des offenen Studiendesigns erfolgten keine Datenauswertungen von primären und sekundären Endpunkten vor der Interims- bzw. primären Analyse durch das Studienteam.

²⁾ Der Einfluss der fehlenden Verblindung wird auf der Endpunktebene individuell berücksichtigt.

Das Verzerrungspotential auf Studienebene wird als niedrig eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 8 dargestellt.

Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie DeLLphi-304

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
Gesamtüberleben	Nein ¹⁾	Ja	Nein	Nein	Niedrig
Morbidität / Lebensqualität					
EORTC QLQ-C30	Nein ²⁾	Nein ³⁾	Nein	Ja ⁴⁾⁵⁾	Hoch
Sicherheit					
Schwere UE (Gesamtrate, SOC/PT)	Nein ⁶⁾	Nein ⁷⁾	Nein	Nein	Hoch
SUE (Gesamtrate, SOC/PT)	Nein ⁶⁾	Nein ⁷⁾	Nein	Nein	Hoch
Abbruch wegen UE (Gesamtrate)	Nein ⁸⁾	Nein	Nein	Nein	Hoch ⁹⁾
UE, jeglichen Schweregrads (SOC/PT)	Nein ⁸⁾	Nein ⁷⁾	Nein	Nein	Hoch

¹⁾ Bei der Studie DeLLphi-304 handelt es sich um eine Studie im Open-label-Design. Bei dem Endpunkt „Gesamtüberleben“ handelt es sich um einen objektiven Endpunkt, wodurch der Einfluss auf die Erhebung und Auswertung des Endpunkts als gering eingeschätzt wird.

²⁾ Die fehlende Verblindung kann die subjektive Einschätzung von Symptomen und Lebensqualität beeinflussen.

³⁾ Es ist eine mögliche informative Zensierung durch Studienabbrüche zu beachten. Eine regelmäßige Erhebung erfolgte nur unter Behandlung mit der Studienmedikation. Nach Studienabbruch erfolgte nur eine einmalige Erhebung von PRO zu Tag 60 (+ 5). Die Ergebnisse ermöglichen nur Aussagen für die Population unter Behandlung, jedoch nicht für die ITT-Population.

⁴⁾ Die PRO werden in beiden Behandlungsarmen zu gleichen Zeitpunkten erhoben. Allerdings unterscheiden sich die Zeitpunkte der Verabreichung der Studienmedikation bzw. die Dauer der Zyklen. Entsprechend erfolgt eine ungleiche Abbildung der Therapiebelastung in den beiden Studienarmen, was die Erhebung von Verschlechterungsereignissen potentiell beeinflussen kann.

⁵⁾ Bereits frühzeitig im Studienverlauf ist der Anteil fehlender Werte hoch (< 70% dauerhaft ab Woche 5 im Kontrollarm; < 70 % dauerhaft ab Woche 8 im Tarlatamab-Arm). Ferner sind differierende Rücklaufquoten zwischen den beiden Studienarmen zu beachten (Unterschied > 5 % bereits zu Woche 2).

⁶⁾ Bei der Studie DeLLphi-304 handelt es sich um eine Studie im Open-Label-Design. Bei den schweren UE und SUE ist aufgrund von objektiven Bewertungskriterien von einem geringen Verzerrungspotential auszugehen.

⁷⁾ Das Ende der Behandlungsphase bei einem Progress und damit der Eintritt in die zeitlich begrenzte Nachbeobachtungsphase kann zu einer verkürzten Erhebung (60 (+ 5) Tage) und einer Verzerrung durch informative Zensierung führen. Ferner ist in diesem Kontext die ungleiche Behandlungsdauer zwischen den beiden Behandlungsarmen zu beachten. Die Ergebnisse ermöglichen nur Aussagen für die Population unter Behandlung, jedoch nicht für die ITT-Population.

⁸⁾ Aufgrund der fehlenden Verblindung ist ein möglicher Einfluss auf die Meldung von UE sowie ein möglicher Einfluss auf den Abbruch der Studienbehandlung zu beachten.

⁹⁾ Aufgrund des Open-label-Designs der Studie DeLLphi-304 wird das Verzerrungspotential als hoch bewertet, da die fehlende Verblindung potentiell zur subjektiven Entscheidung beitragen könnte die Therapie abzubrechen.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; ITT: Intention to Treat; PRO: Patientenberichteter Endpunkt; PT: Preferred Term; SOC: Systemorganklasse; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

3 Ergebnisse der eingeschlossenen Studie

Es werden die Ergebnisse der randomisierten Behandlungsphase der Studie DeLLphi-304 zum Datenschnitt vom 29.01.2025 (basierend auf dem CSR vom 27.05.2025) dargestellt. Ergebnisse, die der pU post hoc ausgewertet hat, wurden Modul 4 bzw. Anhang 4-G des eingereichten Dossiers entnommen.

3.1 Studiencharakteristika und Studienmedikation

Insgesamt wurden 688 Personen gescreent, von denen 509 in die Studie eingeschlossen und randomisiert wurden: 254 in den Tarlatamab-Arm, 255 in den Chemotherapie-Arm. Die mediane Behandlungsdauer betrug zum eingereichten Datenschnitt 4,19 Monate im Tarlatamab-Arm und 2,51 Monate im Chemotherapie-Arm. Weitere Angaben sind Tabelle 9 zu entnehmen.

Tabelle 9: Allgemeine Angaben; Studie DeLLphi-304

Studie DeLLphi-304 Allgemeine Angaben	Tarlatamab N = 254	Chemotherapie N = 255
ITT-Population ¹⁾ , n (%)	254	255
Sicherheitspopulation ¹⁾ , n (%)	252	244
Personen mit einer laufenden Studienteilnahme, n (%)	143 (56,3)	102 (40,0)
Personen, die randomisiert wurden, aber keine Dosis der Studienmedikation erhalten haben, n (%)	2 (0,8)	11 (4,3)
Personen, die die Studienteilnahme abgeschlossen haben, n (%)	0 (0)	0 (0)
Abbruch der Einnahme der Studienmedikation, n (%)	183 (72,0)	225 (88,2)
Aufgrund von:		
UE	5 (2,0)	15 (5,9)
Krankheitsprogression	152 (59,8)	159 (62,4)
Lost to Follow-up	0 (0)	0 (0)
Tod	16 (6,3)	19 (7,5)
Protokollverletzung	0 (0)	0 (0)
Entscheidung der Person	9 (3,5)	17 (6,7)
Erfordernis einer Alternativtherapie	0 (0)	4 (1,6)
Anderer Grund	1 (0,4)	11 (4,3)
Personen, die die Einnahme der Studienmedikation nach initialer Krankheitsprogression fortgeführt haben ²⁾	77 (30,6)	k. A. ³⁾
Personen, die die Studienteilnahme abgebrochen haben, n (%)	111 (43,7)	153 (60,0)
Aufgrund von:		
Tod	111 (43,7)	151 (59,2)
Entscheidung Patientin/Patient	0 (0)	2 (0,8)
Mediane Behandlungsdauer, in Monaten (min; max)	4,19 (0,02; 17,2)	2,51 (0,02; 14,7)
<i>Behandlungsdauer, n (%⁴⁾)</i>		
≥ 3 Monate	149 (59,1)	109 (44,7)
≥ 6 Monate	104 (41,3)	46 (18,9)
≥ 9 Monate	58 (23,0)	17 (7,0)
≥ 12 Monate	26 (10,3)	3 (1,2)
Anzahl der Behandlungszyklen ⁵⁾ , Median (min; max)	5,0 (2,0; 10,0)	4,0 (2,0; 8,0)
Mediane Beobachtungsdauer ⁶⁾ , in Monaten (min; max)	11,2 (0,1; 17,4)	11,7 (0,1; 18,5)

¹⁾ Definition, siehe Kapitel 2.4.

²⁾ Unter bestimmten Voraussetzungen war eine Fortführung der Verabreichung der Studienmedikation (Tarlatamab bzw. Chemotherapie) nach radiologischer Progression möglich. Die prozentuale Angabe bezieht sich auf die Sicherheitspopulation.

³⁾ Für den Kontrollarm erfolgten hierzu keine Angaben in den Studienunterlagen.

⁴⁾ Relative Häufigkeiten zur Exposition mit der Studienmedikation beziehen sich auf die Sicherheitspopulation der Studie DeLLphi-304.

⁵⁾ Die Prüfintervention Tarlatamab wurde in 28-Tage-Zyklen verabreicht. Die Chemotherapie im Kontrollarm wurde in 21-Tage-Zyklen verabreicht. Im Chemotherapie-Arm erhielten die Personen mehrheitlich Topotecan (n = 185), gefolgt von Lurbinectedin (n = 47) und Amrubicin (n = 23).

⁶⁾ Die mediane Beobachtungsdauer bezieht sich auf den Endpunkt „Gesamtüberleben“.

Abkürzungen: ITT: Intention to Treat; k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Tabelle 10: Charakterisierung der Studienpopulation; Studie DeLLphi-304

Studie DeLLphi-304 Charakterisierung der Studienpopulation	Tarlatamab N = 254	Chemotherapie N = 255
<i>Alter (Jahre)</i> MW (SD) Median (min; max)	63,6 (9,4) 64,0 (20; 86)	64,2 (9,2) 66,0 (26; 84)
<i>Altersgruppe (Jahre), n (%)</i> 18 bis 64 65 bis 74 75 bis 84 ≥ 85	129 (50,8) 95 (37,4) 28 (11,0) 2 (0,8)	115 (45,1) 115 (45,1) 25 (9,8) 0 (0)
<i>Geschlecht, n (%)</i> männlich weiblich	182 (71,7) 72 (28,3)	169 (66,3) 86 (33,7)
<i>Abstammung (genetisch), n (%)</i> asiatisch schwarz/afroamerikanisch kaukasisch/weiß hispanisch oder Latino nicht hispanisch oder Latino Ureinwohner Amerikas und Alaskas Andere Fehlend	97 (38,2) 2 (0,8) 152 (59,8) 12 (4,7) 140 (55,1) 1 (0,4) 1 (0,4) 1 (0,4)	107 (42,0) 3 (1,2) 139 (54,5) 11 (4,3) 128 (50,2) 1 (0,4) 3 (1,2) 2 (0,8)
<i>Region, n (%)</i> Nordamerika Europa ¹⁾ Asien Rest der Welt	13 (5,1) 127 (50,0) 97 (38,2) 17 (6,7)	15 (5,9) 113 (44,3) 110 (43,1) 17 (6,7)
<i>ECOG-PS, n (%)</i> 0 1 2	83 (32,7) 169 (66,5) 2 (0,8)	80 (31,4) 173 (67,8) 2 (0,8)
<i>Rauchhistorie, n (%)</i> Nie Raucher Aktueller Raucher Ehemaliger Raucher	23 (9,1) 54 (21,3) 177 (69,7)	31 (12,2) 51 (20,0) 173 (67,8)
<i>Vorthherapie zu Studienbeginn (Anzahl), n (%)</i> 1 ≥ 2	249 (98,0) 5 (2,0)	248 (97,3) 7 (2,7)

Studie DeLLphi-304 Charakterisierung der Studienpopulation	Tarlatamab N = 254	Chemotherapie N = 255
<i>Vorthherapie mit einem PD-(L)1-Inhibitor²⁾, n (%)</i>		
ja	180 (70,9)	180 (70,6)
nein	74 (29,1)	75 (29,4)
<i>Vorherige Radiotherapie für aktuelle Krebserkrankung, n (%)</i>		
ja	159 (62,6)	161 (63,1)
nein	95 (37,4)	94 (36,1)
<i>Vorherige Operation für aktuelle Krebserkrankung, n (%)</i>		
ja	26 (10,2)	26 (10,2)
nein	228 (89,9)	229 (89,8)
<i>Krankheitsstadium (Screening), n (%)</i>		
Stadium 0	0 (0)	0 (0)
Stadium I	0 (0)	0 (0)
Stadium II	0 (0)	2 (0,8)
Stadium III	21 (8,3)	18 (7,1)
Stadium IV	233 (91,7)	235 (92,2)
<i>Metastasierung (Baseline), n (%)</i>		
ja	227 (89,4)	246 (92,5)
nein	27 (10,6)	19 (7,5)
<i>Chemotherapie-freies Intervall²⁾³⁾, n (%)</i>		
< 90 Tage	109 (42,9)	114 (44,7)
≥ 90 Tage	145 (57,1)	141 (55,3)
≥ 90 Tage und < 180 Tage	85 (33,5)	78 (30,6)
≥ 180 Tage	60 (23,6)	63 (24,7)
<i>Hirnmetastasen (vorherige oder aktuelle, Baseline)²⁾, n (%)</i>		
ja	113 (44,5)	115 (45,1)
nein	141 (55,5)	140 (54,9)
<i>Lebermetastasen (Baseline), n (%)</i>		
ja	84 (33,1)	95 (37,3)
nein	170 (66,9)	160 (62,7)
<i>DLL3, n (%)</i>		
fehlend	37 (14,6)	41 (16,1)
positiv ⁴⁾	207 (95,4)	198 (92,5)
DLL3-Schwellenwerte		
< 75% bei 2+ und 3+ Farbintensität	119 (46,9)	114 (44,7)
≥ 75% bei 2+ und 3+ Farbintensität	98 (38,6)	100 (39,2)
< 25% bei 2+ und 3+ Farbintensität	52 (20,5)	52 (20,4)
≥ 25% bei 2+ und 3+ Farbintensität	165 (65,0)	162 (63,5)
<i>EQ-5D-VAS zu Baseline</i>		
n (%)	247 (97,2)	250 (98,0)
MW (SD)	73,9 (17,2)	73,8 (18,1)
Median (min; max)	k. A.	k. A.

¹⁾ Darunter 11 Personen (2,2 % der gesamten Studienpopulation) aus Deutschland.

²⁾ Stratifizierungsfaktor der Studie DeLLphi-304.

³⁾ Anzahl an Tagen, welche zwischen der letzten Dosis der vorangegangenen Chemotherapie und dem Zeitpunkt des ersten Verdachts auf ein Wiederauftreten oder Fortschreiten der Erkrankung (klinisch oder radiologisch) verstrichen sind.

⁴⁾ N bezieht sich auf Personen mit nicht-fehlendem DLL3-Status (217 im Tarlatamab-Arm, 214 im Chemotherapie-Arm).

Abkürzungen: DLL3: Delta-like Ligand 3; ECOG-PS: Eastern Cooperative Oncology Group – Performance Status; EQ-5D-VAS: Visuelle Analogskala des EuroQoL 5-Dimension; k. A.: keine Angabe; MW: Mittelwert; PD-(L)1: Programmed Cell Death (Ligand) 1; SD: Standardabweichung.

Protokollverletzungen

Anhand der in der Studie DeLLphi-304 dokumentierten Protokollverletzungen ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Begleitmedikation

In der Studie DeLLphi-304 erhielten 251 Personen (99,6 %) im Tarlatamab- und 242 (99,2 %) im Chemotherapie-Arm eine Begleitmedikation. Anhand der in der Studie DeLLphi-304 dokumentierten Begleitmedikation ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Folgetherapien

Angaben zu Folgetherapien waren den statistischen Outputs in Anhang 4-G des Dossiers zu entnehmen. Die relativen Angaben beziehen sich auf die Sicherheitspopulation der Studie DeLLphi-304. Nach Abbruch der Studienmedikation haben 112 Personen (44,4 %) im Tarlatamab-Arm und 125 (51,2 %) im Chemotherapie-Arm mindestens eine Folgetherapie erhalten. Der Anteil an anschließenden Folgetherapien war somit im Kontrollarm höher als im Interventionsarm. Es liegen keine Daten zur kumulierten Anzahl an Folgetherapien vor. Die häufigste erste Folgetherapie war eine Mono-Chemotherapie in beiden Behandlungsarmen (40,5 % im Tarlatamab-Arm bzw. 36,5 % im Chemotherapie-Arm, bezogen auf die ITT-Population), wobei Carboplatin den häufigsten Wirkstoff darstellte (14,3 % bzw. 11,5 %). Weitere Angaben sind Tabelle 11 zu entnehmen.

Tabelle 11: Folgetherapien; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 Folgetherapie	Tarlatamab N = 252 n (%)	Chemotherapie N = 244 n (%)
Erhalt einer Folgetherapie	112 (44,4)	125 (51,2)
<i>Erste Folgetherapie nach Behandlung mit der Studienmedikation</i>		
Antikörper-Wirkstoff-Konjugat	4 (1,6)	4 (1,6)
Anti-DLL3-Therapie	0 (0)	10 (4,1) ²⁾
Chemotherapie ¹⁾	1 (0,4)	1 (0,4)
Monochemotherapie	102 (40,5)	89 (36,5)
PD-(L)1-Checkpoint-Inhibitor	12 (4,8)	13 (5,3)
Strahlentherapie	1 (0,4)	3 (1,2)
Tyrosinkinase-Inhibitor	5 (2,0)	18 (7,4)
Andere	4 (1,6)	4 (1,6)

¹⁾ Kombinationschemotherapie (Cisplatin + Irinotecan bzw. Carboplatin + Irinotecan).

²⁾ 9 Personen (3,7 %) im Chemotherapie-Arm haben Tarlatamab als erste Folgetherapie erhalten.

Abkürzungen: DLL3: Delta-like Ligand 3; PD-(L)1: Programmed Cell Death (Ligand) 1.

3.2 Mortalität

Gesamtüberleben

In Tabelle 12 und Abbildung 1 werden die Ergebnisse zum Gesamtüberleben für die ITT-Population zum Datenschnitt vom 29.01.2025 dargestellt. Die mediane Follow-up-Dauer für den Endpunkt „Gesamtüberleben“ betrug 11,2 Monate (min; max: 0,1; 17,4) im Tarlatamab-Arm und 11,7 Monate (min; max: 0,1; 18,5) im Chemotherapie-Arm. Insgesamt wurden bis zum 29.01.2025 263 Todesfälle in der Studie DeLLphi-304 beobachtet; im Tarlatamab-Arm waren es 111 Todesfälle (43,7 %), im Chemotherapie-Arm 152 (59,6 %). Der Effektschätzer (HR: 0,60 (95%-KI: [0,47; 0,77])) war statistisch signifikant ($p < 0,001$).

Tabelle 12: Gesamtüberleben; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 Gesamtüberleben	Tarlatamab N = 254	Chemotherapie N = 255
Ereignis, n (%)	111 (43,7)	152 (59,6)
Zensierungen ¹⁾ , n (%)	143 (56,3)	103 (40,4)
Mediane Beobachtungsdauer, Monate (min; max)	11,2 (0,1; 17,4)	11,7 (0,1; 18,5)
Medianes Überleben, Monate [95%-KI]	13,6 [11,1; n. a.]	8,3 [7,0; 10,2]
HR [95%-KI] ²⁾ p-Wert ³⁾	0,60 [0,47; 0,77] < 0,001	

¹⁾ Alle Zensierungen im Tarlatamab-Arm (n = 143) und n = 102 Zensierungen im Chemotherapie-Arm waren auf einen positiven Überlebensstatus zum Zeitpunkt des Datenschnitts zurückzuführen. 1 Person im Chemotherapie-Arm wurde aufgrund eines Widerrufs der Einwilligung zensiert.

²⁾ Das HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell. Nach Angaben im CSR wurden 2 der 4 Stratifikationsfaktoren genutzt: „Chemotherapie-freies Intervall (< 90 Tage vs. ≥ 90 Tage)“ und „Vorliegen von Hirnmetastasen (ja vs. nein)“.

³⁾ Der p-Wert basiert auf einem stratifizierten Log-Rank-Test sowie einem zweiseitigen Testverfahren. Der Log-Rank-Test wurde mit einem Signifikanzniveau von 0,05 durchgeführt. Nach Angaben im CSR wurden 2 der 4 Stratifikationsfaktoren genutzt: „Chemotherapie-freies Intervall (< 90 Tage vs. ≥ 90 Tage)“ und „Vorliegen von Hirnmetastasen (ja vs. nein)“.

Abkürzungen: CSR: Klinischer Studienbericht; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall; n. a.: nicht anwendbar.

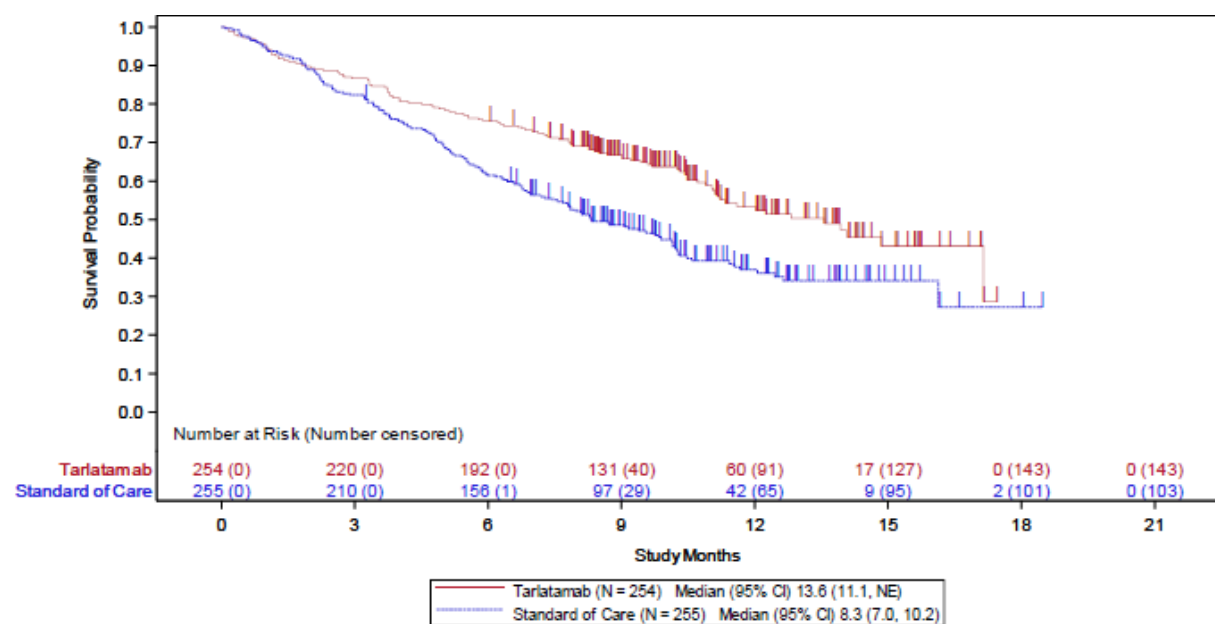


Abbildung 1: Kaplan-Meier-Kurve für „Gesamtüberleben“;
Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Subgruppenanalysen Gesamtüberleben

Für die vorliegende Nutzenbewertung relevante Subgruppeneffekte, mit mindestens einem signifikanten Effektschätzer in den jeweiligen Subgruppen, werden in Tabelle 13 dargestellt. Es zeigt sich ein statistisch signifikanter ($p = 0,024$) Interaktionsterm für „Vorliegen von (früheren oder aktuellen) Hirnmetastasen“.

*Tabelle 13: Gesamtüberleben, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population
(Datenschnitt: 29.01.2025)*

Studie DeLLphi-304 Gesamtüberleben, Subgruppenanalysen		Tarlatamab N = 254	Chemotherapie N = 255	HR [95%-KI] ¹⁾ p-Wert ²⁾
Gesamtüberleben – Vorliegen von (früheren oder aktuellen) Hirnmetastasen ($p = 0,024$)³⁾				
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Ja	N	113 (100)	115 (100)	0,45 [0,31; 0,65] < 0,001
	Ereignisse (%)	47 (41,6)	78 (67,8)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	14,1 [10,9; n. a.]	7,5 [5,7; 8,7]	
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Nein	N	141 (100)	140 (100)	0,81 [0,58; 1,13] 0,22
	Ereignisse (%)	64 (45,4)	74 (52,9)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	12,3 [10,4; n. a.]	10,3 [7,9; n. a.]	

¹⁾ Die Schätzung des HR und 95%-KI erfolgte mit einem unstratifizierten Cox-Proportional-Hazards-Modell.

²⁾ Der p-Wert basiert auf dem unstratifizierten Log-Rank-Test.

³⁾ p-Wert des Subgruppen-Interaktionstests ($\alpha = 0,05$).

⁴⁾ Nach Angaben in Modul 4 handelt es sich um die mediane Zeit in Monaten. Aus Plausibilitätsgründen wird angenommen, dass es sich um die mediane Zeit in Wochen handelt.

⁵⁾ Die Schätzung der medianen Zeit erfolgte mittels Kaplan-Meier-Methode. Die 95%-KI basieren auf einer log-log-Transformation gemäß Brookmeyer und Crowley.

Abkürzungen: HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall; n. a.: nicht anwendbar.

3.3 Morbidität

Krankheitssymptomatik mittels EORTC QLQ-C30 – Symptomskalen

Die Responderanalysen „Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte“ im Rahmen der Symptomskalen des EORTC QLQ-C30 sind Tabelle 14 zu entnehmen; im Anhang finden sich die dazugehörigen Kaplan-Meier-Kurven (Abbildung 2 bis Abbildung 9).

Bei der Responderanalyse „Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte“ zeigte sich für das Symptom „Diarrhö“ ein signifikanter Unterschied zwischen den beiden Behandlungsgruppen zugunsten von Tarlatamab. Ferner zeigte sich ein statistisch signifikanter Unterschied für das Symptom „Appetitverlust“ zuungunsten von Tarlatamab.

Tabelle 14: EORTC QLQ-C30 – Symptomskalen, Responderanalyse (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 EORTC QLQ-C30 – Symptomskalen	Tarlatamab N = 254	Chemotherapie N = 255	HR [95%-KI] ¹⁾ ; p-Wert ²⁾
Fatigue			
Ereignisse (%) ³⁾	194 (76,4)	188 (73,7)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	2,3 [2,0; 2,4]	3,0 [2,3; 3,3]	1,17 [0,96; 1,43]; 0,13
Übelkeit und Erbrechen			
Ereignisse (%) ³⁾	119 (46,9)	137 (53,7)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	18,4 [9,1; 47,8]	7,1 [5,3; 10,3]	0,79 [0,62; 1,02]; 0,07
Schmerz			
Ereignisse (%) ³⁾	156 (61,4)	163 (63,9)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	4,6 [3,3; 6,8]	5,7 [4,3; 7,0]	0,97 [0,78; 1,21]; 0,78
Dyspnoe			
Ereignisse (%) ³⁾	124 (48,8)	138 (54,1)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	18,1 [6,8; 30,9]	7,3 [5,3; 9,5]	0,79 [0,62; 1,01]; 0,06
Schlaflosigkeit			
Ereignisse (%) ³⁾	127 (50,0)	132 (51,8)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	18,1 [7,5; 30,2]	10,0 [7,4; 11,7]	0,88 [0,68; 1,12]; 0,29
Appetitverlust			
Ereignisse (%) ³⁾	171 (67,3)	151 (59,2)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	3,3 [2,4; 4,6]	5,1 [4,3; 6,5]	1,27 [1,02; 1,59]; 0,04
Obstipation			
Ereignisse (%) ³⁾	148 (58,3)	141 (55,3)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	7,4 [5,3; 10,1]	6,8 [5,1; 9,3]	0,97 [0,77; 1,22]; 0,80
Diarrhö			
Ereignisse (%) ³⁾	70 (27,6)	90 (35,3)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	n. a. [61,9; n. a.]	30,5 [18,1; n. a.]	0,67 [0,49; 0,92]; 0,01

¹⁾ Die Berechnung des HR und 95 %-KI erfolgte anhand eines stratifizierten Cox-Proportional-Hazard-Modells.

²⁾ Der p-Wert basiert auf dem stratifizierten Log-Rank-Test. Als Stratifizierungsfaktoren dienten „Chemotherapie-freies Intervall (< 90 Tage vs. ≥ 90 Tage)“ sowie „Vorliegen von Hirnmetastasen (ja vs. nein)“. Es wird angenommen, dass der Log-Rank-Test zweiseitig mit einem Signifikanzniveau von 0,05 durchgeführt wurde.

³⁾ Es konnten keine Angaben zu Zensierungsgründen (Anzahl) aus den Unterlagen entnommen werden.

⁴⁾ Nach Angaben in Modul 4 handelt es sich um die mediane Zeit in Monaten. Aus Plausibilitätsgründen wird angenommen, dass es sich um die mediane Zeit in Wochen handelt.

⁵⁾ Die mediane Zeit bis zum Ereignis wurde mittels Kaplan-Meier-Methode geschätzt; die 95%-KI basieren auf einer log-log-Transformation gemäß Brookmeyer und Crowley.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall; n. a.: nicht anwendbar.

Subgruppenanalysen EORTC QLQ-C30

Es liegen signifikante Interaktionsterme ($p < 0,05$) für die Symptomskalen des EORTC QLQ-C30 vor. In Tabelle 15 wird die Subgruppenanalyse für das Symptom „Schmerz“ dargestellt. Weitere Subgruppenanalysen werden für die Symptomskalen des EORTC QLQ-C30 in der vorliegende Nutzenbewertung aufgrund fehlender Relevanz nicht dargestellt.

Tabelle 15: EORTC QLQ-C30 – Symptomskalen, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 EORTC QLQ-C30 – Symptomskalen, Subgruppenanalysen		Tarlatamab N = 252	Chemotherapie N = 254	HR [95%-KI] ¹⁾ p-Wert ²⁾
Schmerz – Vorliegen von (früheren oder aktuellen) Hirnmetastasen ($p < 0,001$)³⁾				
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Ja	N	113	115	0,61 [0,44; 0,86] 0,004
	Ereignisse (%)	63 (55,8)	77 (67,0)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	9,4 [3,6; 24,2]	4,3 [3,3; 6,3]	
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Nein	N	141	140	1,37 [1,02; 1,84] 0,04
	Ereignisse (%)	93 (66,0)	86 (61,4)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	3,4 [2,4; 5,0]	7,1 [4,3; 10,7]	

¹⁾ Die Schätzung des HR und 95%-KI erfolgte mit einem unstratifizierten Cox-Proportional-Hazards-Modell.

²⁾ Der p-Wert basiert auf dem unstratifizierten Log-Rank-Test.

³⁾ p-Wert des Subgruppen-Interaktionstests ($\alpha = 0,05$).

⁴⁾ Nach Angaben in Modul 4 handelt es sich um die mediane Zeit in Monaten. Aus Plausibilitätsgründen wird angenommen, dass es sich um die mediane Zeit in Wochen handelt.

⁵⁾ Die Schätzung der medianen Zeit erfolgte mittels Kaplan-Meier-Methode. Die 95%-KI basieren auf einer log-log-Transformation gemäß Brookmeyer und Crowley.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall.

3.4 Lebensqualität

EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus

Die Responderanalysen „Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte“ im Rahmen der Funktionsskalen und der globalen Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 sind Tabelle 16 zu entnehmen; im Anhang finden sich die dazugehörigen Kaplan-Meier-Kurven (Abbildung 10 bis Abbildung 15). Bei der Responderanalyse zeigten sich keine signifikanten Unterschiede zwischen den beiden Behandlungsarmen.

Tabelle 16: EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus, Responderanalysen (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus	Tarlatamab N = 254	Chemotherapie N = 255	HR [95%-KI] ¹⁾ ; p-Wert ²⁾
Globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“			
Ereignisse (%) ³⁾	172 (67,7)	172 (67,5)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	2,6 [2,3; 3,7]	4,1 [3,3; 5,1]	1,07 [0,86; 1,32]; 0,57
Kognitive Funktion			
Ereignisse (%) ³⁾	139 (54,7)	141 (55,3)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	7,7 [4,4; 12,2]	6,4 [5,3; 9,8]	0,96 [0,75; 1,21]; 0,68
Emotionale Funktion			
Ereignisse (%) ³⁾	112 (44,1)	125 (49,0)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	25,6 [11,5; n. a.]	9,1 [6,3; 24,3]	0,78 [0,60; 1,01]; 0,06
Körperliche Funktion			
Ereignisse (%) ³⁾	154 (60,6)	159 (62,4)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	4,3 [3,4; 6,0]	5,4 [4,1; 7,3]	0,98 [0,78; 1,23]; 0,86
Rollenfunktion			
Ereignisse (%) ³⁾	172 (67,7)	176 (69,0)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	2,6 [2,3; 4,0]	4,1 [3,3; 4,6]	1,02 [0,82; 1,26]; 0,87
Soziale Funktion			
Ereignisse (%) ³⁾	147 (57,9)	159 (62,4)	
Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	3,3 [2,4; 6,5]	4,6 [3,6; 5,6]	0,95 [0,76; 1,20]; 0,65

¹⁾ Die Berechnung des HR und 95%-KI erfolgte anhand eines stratifizierten Cox-Proportional-Hazard-Modells.

²⁾ Der p-Wert basiert auf dem stratifizierten Log-Rank-Test. Als Stratifizierungsfaktoren dienten „Chemotherapie-freies Intervall (< 90 Tage vs. ≥ 90 Tage)“ sowie „Vorliegen von Hirnmetastasen (ja vs. nein)“. Es wird angenommen, dass der Log-Rank-Test zweiseitig mit einem Signifikanzniveau von 0,05 durchgeführt wurde.

³⁾ Es konnten keine Angaben zu Zensierungsgründen (Anzahl) aus den Unterlagen entnommen werden.

⁴⁾ Nach Angaben in Modul 4 handelt es sich um die mediane Zeit in Monaten. Aus Plausibilitätsgründen wird angenommen, dass es sich um die mediane Zeit in Wochen handelt.

⁵⁾ Die mediane Zeit bis zum Ereignis wurde mittels Kaplan-Meier-Methode geschätzt; die 95%-KI basieren auf einer log-log-Transformation gemäß Brookmeyer und Crowley.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall; n. a.: nicht anwendbar.

Subgruppenanalysen EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus

Es liegen signifikante Interaktionsterme ($p < 0,05$) für die Funktionsskalen / den globalen Gesundheitsstatus des EORTC QLQ-C30 vor. In Tabelle 17 wird die Subgruppenanalyse für die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ dargestellt. Subgruppenanalysen für die Funktionsskalen des EORTC QLQ-C30 werden in der vorliegenden Nutzenbewertung aufgrund fehlender Relevanz nicht dargestellt.

Tabelle 17: EORTC QLQ-C30 – Funktionsskalen, Subgruppenanalysen; Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 EORTC QLQ-C30 – Funktionsskalen, Subgruppenanalysen		Tarlatamab N = 252	Chemotherapie N = 254	HR [95%-KI] ¹⁾ p-Wert ²⁾
Globaler Gesundheitsstatus – Vorliegen von (früheren oder aktuellen) Hirnmetastasen ($p = 0,02$)³⁾				
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Ja	N	113	115	0,80 [0,58; 1,11] 0,18
	Ereignisse (%)	70 (61,9)	75 (65,2)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	3,0 [2,3; 6,4]	3,3 [2,4; 4,8]	
Vorliegen von (früheren oder aktuellen) Hirnmetastasen: Nein	N	141	140	1,33 [1,01; 1,76] 0,04
	Ereignisse (%)	102 (72,3)	97 (69,3)	
	Mediane Zeit (Wochen ⁴⁾) [95%-KI] ⁵⁾	2,6 [2,3; 3,6]	4,4 [3,3; 6,3]	

¹⁾ Die Schätzung des HR und 95%-KI erfolgte mit einem unstratifizierten Cox-Proportional-Hazards-Modell.

²⁾ Der p-Wert basiert auf dem unstratifizierten Log-Rank-Test.

³⁾ p-Wert des Subgruppen-Interaktionstests ($\alpha = 0,05$).

⁴⁾ Nach Angaben in Modul 4 handelt es sich um die mediane Zeit in Monaten. Aus Plausibilitätsgründen wird angenommen, dass es sich um die mediane Zeit in Wochen handelt.

⁵⁾ Die Schätzung der medianen Zeit erfolgte mittels Kaplan-Meier-Methode. Die 95%-KI basieren auf einer log-log-Transformation gemäß Brookmeyer und Crowley.

Abkürzungen: EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall.

3.5 Sicherheit

Die medianen Beobachtungszeiten für UE unterscheiden zwischen den beiden Behandlungsarmen. Im Tarlatamab-Arm war diese mit 6,31 Monaten (min; max: 0,1; 17,3) länger als im Chemotherapie-Arm mit 4,42 Monaten (min; max: 0,1; 15,5). Ebenso unterscheidet sich die mediane Behandlungsdauer in beiden Studienarmen: Im Tarlatamab-Arm betrug sie 4,19 Monate, im Chemotherapie-Arm 2,51 Monate.

Damit liegen Unterschiede zwischen den Studienarmen vor, die bei der Interpretation der beobachteten Häufigkeiten und Hazard Ratios berücksichtigt werden müssen.

Der pU legte TEAE-Analysen unter Ausschluss von Ereignissen vor, die mit einer Progression des SCLC zusammenhängen; die Ergebnisse sind Tabelle 18 bis Tabelle 21 zu entnehmen. In Abbildung 16 bis Abbildung 18 (im Anhang der vorliegenden Nutzenbewertung) finden sich die Kaplan-Meier-Kurven zu schweren UE (CTCAE-Grad ≥ 3), SUE und UE, das zum Abbruch der Studienmedikation führte.

Tabelle 18: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der Behandlungsphase; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 Zusammenfassung der UE	Tarlatamab N = 252¹⁾ <i>Personen mit Ereignis, n (%)</i> <i>Mediane Zeit bis zum Ereignis, Monate [95%-KI]²⁾</i>	Chemotherapie N = 244¹⁾ <i>Personen mit Ereignis, n (%)</i> <i>Mediane Zeit bis zum Ereignis, Monate [95%-KI]²⁾</i>	HR [95%-KI] p-Wert³⁾
<i>Personen mit mindestens einem ...</i>			
UE (ergänzend dargestellt)	249 (98,8) 0,1 [n. a.; n. a.]	243 (99,6) 0,3 [0,2; 0,3]	-
schweren UE (CTCAE-Grad ≥ 3)	136 (54,0) 5,0 [3,3; 9,3]	195 (79,9) 1,1 [0,7; 1,5]	0,46 [0,37; 0,58] < 0,001
SUE	129 (51,2) 5,1 [3,6; n. a.]	125 (51,2) 4,6 [3,4; 7,0]	1,03 [0,81; 1,33] 0,83
UE, das zum Abbruch der Studienmedikation führte	13 (5,2) n. a. [n. a.; n. a.]	30 (12,3) n. a. [11,5; n. a.]	0,37 [0,19; 0,72] 0,002

¹⁾ Die Sicherheitspopulation umfasst alle Personen, die zu einem Behandlungsarm randomisiert wurden und mind. eine Medikationsdosis erhalten haben.

²⁾ Das 95%-KI basiert auf der geschätzten Varianz der log-log-Transformation der Kaplan-Meier-Schätzung.

³⁾ Der zweiseitige p-Wert basiert auf einem unstratifizierten Log-Rank-Test; p-Werte unter 0,05 werden als statistisch signifikant angesehen.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; HR: Hazard Ratio; KI: Konfidenzintervall; n. a.: nicht anwendbar; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; - : nicht berechenbar.

Unerwünschte Ereignisse

Tabelle 19: UE mit Inzidenz $\geq 10\%$ bzw. ≥ 10 Personen und $\geq 1\%$ der Personen in einem Studienarm; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 UE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Erkrankungen des Blutes und des Lymphsystems	114 (45,2) 8,7 [4,6; n. a.]	80 (73,8) 0,7 [0,7; 1,2]	0,38 [0,30; 0,49] < 0,001
Anämie	78 (31,0) n. a. [10,7; n. a.]	156 (63,9) 1,5 [1,3; 2,1]	0,30 [0,23; 0,40] < 0,001
Neutropenie	27 (10,7) n. a. [n. a.; n. a.]	76 (31,1) n. a. [10,2; n. a.]	0,26 [0,17; 0,41] < 0,001
Leukopenie	24 (9,5) n. a. [n. a.; n. a.]	54 (22,1) n. a. [n. a.; n. a.]	0,37 [0,23; 0,60] < 0,001
Lymphopenie	23 (9,1) n. a. [n. a.; n. a.]	10 (4,1) n. a. [n. a.; n. a.]	2,19 [1,04; 4,61] 0,03
Thrombozytopenie	14 (5,6) n. a. [n. a.; n. a.]	62 (25,4) n. a. [n. a.; n. a.]	0,18 [0,10; 0,32] < 0,001
Febrile Neutropenie	5 (2,0) n. a. [n. a.; n. a.]	28 (11,5) n. a. [n. a.; n. a.]	0,13 [0,05; 0,35] < 0,001
Herzerkrankungen	32 (12,7) n. a. [n. a.; n. a.]	32 (13,1) n. a. [n. a.; n. a.]	0,85 [0,52; 1,39] 0,52
Endokrine Erkrankung	21 (8,3) n. a. [n. a.; n. a.]	9 (3,7) n. a. [n. a.; n. a.]	2,07 [0,94; 4,53] 0,06
Augenerkrankungen	5 (2,0) n. a. [n. a.; n. a.]	15 (6,1) n. a. [n. a.; n. a.]	0,25 [0,09; 0,69] 0,004
Erkrankungen des Gastrointestinaltrakts	142 (56,3) 3,3 [1,5; 6,6]	155 (63,5) 1,6 [0,9; 2,5]	0,77 [0,61; 0,97] 0,03
Obstipation	72 (28,6) n. a. [n. a.; n. a.]	54 (22,1) n. a. [10,8; n. a.]	1,19 [0,83; 1,69] 0,35
Übelkeit	61 (24,2) n. a. [n. a.; n. a.]	78 (32,0) n. a. [n. a.; n. a.]	0,70 [0,50; 0,98] 0,037
Dyspepsie	13 (5,2) n. a. [n. a.; n. a.]	2 (0,8) n. a. [n. a.; n. a.]	5,96 [1,34; 26,54] 0,01
Abdominalschmerzen	19 (7,5) n. a. [n. a.; n. a.]	10 (4,1) n. a. [n. a.; n. a.]	1,66 [0,77; 3,56] 0,19
Stomatitis	8 (3,2) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	0,63 [0,23; 1,58] 0,32
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	161 (63,9) 1,2 [0,5; 2,0]	154 (63,1) 2,1 [1,4; 2,9]	1,07 [0,85; 1,33] 0,62
Fatigue	72 (28,6) n. a. [n. a.; n. a.]	74 (30,3) n. a. [n. a.; n. a.]	0,87 [0,63; 1,21] 0,41
Fieber	69 (27,4) n. a. [n. a.; n. a.]	27 (11,1) n. a. [n. a.; n. a.]	2,74 [1,75; 4,28] < 0,001

Studie DeLLphi-304 UE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Asthenie	28 (11,1) n. a. [n. a.; n. a.]	36 (14,8) n. a. [n. a.; n. a.]	0,68 [0,41; 1,13] 0,13
Brustschmerzen	15 (6,0) n. a. [n. a.; n. a.]	12 (4,9) n. a. [n. a.; n. a.]	1,14 [0,53; 2,44] 0,73
Peripheres Ödem	5 (2,0) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	0,4 [0,14; 1,16] 0,08
Leber- und Gallenerkrankungen	10 (4,0) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	0,68 [0,28; 1,63] 0,39
Erkrankungen des Immunsystems	142 (56,3) 0,3 [0,3; 0,5]	4 (1,6) n. a. [n. a.; n. a.]	50,7 [18,8; 137,0] < 0,001
Zytokin-Freisetzungssyndrom	142 (56,3) 0,3 [0,3; 0,5]	3 (1,2) n. a. [n. a.; n. a.]	67,7 [21,6; 212,5] < 0,001
Infektionen und parasitäre Erkrankungen	91 (36,1) n. a. [9,3; n. a.]	95 (38,9) 7,0 [5,2; n. a.]	0,82 [0,61; 1,10] 0,18
Pneumonie	23 (9,1) n. a. [n. a.; n. a.]	32 (13,1) n. a. [n. a.; n. a.]	0,57 [0,33; 0,98] 0,04
Harnwegsinfektion	16 (6,3) n. a. [n. a.; n. a.]	8 (3,3) n. a. [12,2; n. a.]	1,61 [0,68; 3,80] 0,28
Infektion der oberen Atemwege	11 (4,4) n. a. [n. a.; n. a.]	9 (3,7) n. a. [11,4; n. a.]	0,88 [0,36; 2,16] 0,77
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	14 (5,6) n. a. [n. a.; n. a.]	17 (7,0) n. a. [n. a.; n. a.]	0,65 [0,32; 1,33] 0,23
Untersuchungen	112 (44,4) 10,2 [5,4; n. a.]	100 (41,0) 7,6 [4,7; n. a.]	1,02 [0,78; 1,34] 0,88
Leukozytenzahl vermindert	24 (9,5) n. a. [n. a.; n. a.]	32 (13,1) n. a. [n. a.; n. a.]	0,61 [0,36; 1,04] 0,07
Gewicht vermindert	31 (12,3) n. a. [n. a.; n. a.]	16 (6,6) n. a. [n. a.; n. a.]	1,84 [1,00; 3,37] 0,045
Neutrophilenzahl vermindert	20 (7,9) n. a. [n. a.; n. a.]	42 (17,2) n. a. [n. a.; n. a.]	0,37 [0,21; 0,63] < 0,001
Thrombozytenzahl vermindert	10 (4,0) n. a. [n. a.; n. a.]	44 (18,0) n. a. [n. a.; n. a.]	0,18 [0,09; 0,36] < 0,001
Lymphozytenzahl vermindert	15 (6,0) n. a. [n. a.; n. a.]	12 (4,9) n. a. [n. a.; n. a.]	1,14 [0,53; 2,45] 0,73
Alanin-Aminotransferase erhöht	19 (7,5) n. a. [n. a.; n. a.]	14 (5,7) n. a. [n. a.; n. a.]	1,24 [0,62; 2,47] 0,56
Aspartat-Aminotransferase erhöht	19 (7,5) n. a. [n. a.; n. a.]	8 (3,3) n. a. [n. a.; n. a.]	2,12 [0,92; 4,87] 0,07

Studie DeLLphi-304 UE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Stoffwechsel- und Ernährungsstörungen	146 (57,9) 2,8 [1,4; 4,6]	112 (45,9) 5,1 [3,6; n. a.]	1,40 [1,09; 1,79] 0,01
Appetit vermindert	89 (35,3) n. a. [13,4; n. a.]	54 (22,1) n. a. [n. a.; n. a.]	1,62 [1,15; 2,27] 0,01
Hyponatriämie	43 (17,1) n. a. [n. a.; n. a.]	24 (9,8) n. a. [n. a.; n. a.]	1,69 [1,02; 2,78] 0,04
Hypokaliämie	24 (9,5) n. a. [n. a.; n. a.]	19 (7,8) n. a. [n. a.; n. a.]	1,17 [0,64; 2,15] 0,61
Hypoalbuminämie	22 (8,7) n. a. [n. a.; n. a.]	14 (5,7) n. a. [n. a.; n. a.]	1,46 [0,74; 2,86] 0,27
Hypomagnesiämie	17 (6,7) n. a. [n. a.; n. a.]	8 (3,3) n. a. [n. a.; n. a.]	1,90 [0,81; 4,41] 0,13
Hyperglykämie	12 (4,8) n. a. [n. a.; n. a.]	4 (1,6) n. a. [n. a.; n. a.]	2,56 [0,82; 7,98] 0,10
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	74 (29,4) 14,6 [14,1; n. a.]	61 (25,0) n. a. [9,0; n. a.]	1,05 [0,74; 1,48] 0,81
Rückenschmerzen	22 (8,7) n. a. [n. a.; n. a.]	17 (7,0) n. a. [n. a.; n. a.]	1,11 [0,59; 2,10] 0,75
Arthralgie	19 (7,5) n. a. [n. a.; n. a.]	17 (7,0) n. a. [n. a.; n. a.]	0,86 [0,44; 1,67] 0,65
Myalgie	16 (6,3) n. a. [n. a.; n. a.]	6 (2,5) n. a. [n. a.; n. a.]	2,55 [0,99; 6,52] 0,04
Schmerzen in den Extremitäten	11 (4,4) n. a. [n. a.; n. a.]	10 (4,1) n. a. [n. a.; n. a.]	0,85 [0,35; 2,06] 0,72
Muskelschwäche	2 (0,8) n. a. [n. a.; n. a.]	0 (4,1) n. a. [n. a.; n. a.]	0,15 [0,03; 0,67] 0,004
Erkrankungen des Nervensystems	134 (53,2) 3,3 [1,4; 5,9]	73 (29,9) n. a. [7,4; n. a.]	2,21 [1,66; 2,95] < 0,001
Dysgeusie	61 (24,2) n. a. [n. a.; n. a.]	4 (1,6) n. a. [n. a.; n. a.]	16,2 [5,89; 44,6] < 0,001
Kopfschmerzen	38 (15,1) n. a. [n. a.; n. a.]	21 (8,6) n. a. [n. a.; n. a.]	1,72 [1,01; 2,94] 0,044
Schwindel	24 (9,5) n. a. [n. a.; n. a.]	25 (10,2) n. a. [n. a.; n. a.]	0,87 [0,50; 1,53] 0,64
Immuneffektorzell-assoziiertes Neurotoxizitätssyndrom	15 (6,0) n. a. [n. a.; n. a.]	2 (0,8) n. a. [n. a.; n. a.]	7,03 [1,60; 30,8] 0,003
Psychiatrische Erkrankungen	30 (11,9) n. a. [n. a.; n. a.]	32 (13,1) n. a. [n. a.; n. a.]	0,83 [0,50; 1,36] 0,45
Schlaflosigkeit	18 (7,1) n. a. [n. a.; n. a.]	21 (8,6) n. a. [n. a.; n. a.]	0,76 [0,40; 1,42] 0,39

Studie DeLLphi-304 UE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Erkrankungen der Nieren und Harnwege	31 (12,3) n. a. [n. a.; n. a.]	13 (5,3) n. a. [n. a.; n. a.]	2,14 [1,12; 4,11] 0,02
Erkrankungen der Atemwege, des Brustraums und Mediastinum	91 (36,1) 12,9 [11,1; n. a.]	87 (35,7) n. a. [7,5; n. a.]	0,88 [0,65; 1,18] 0,38
Husten	31 (12,3) n. a. [n. a.; n. a.]	34 (13,9) n. a. [n. a.; n. a.]	0,75 [0,46; 1,23] 0,26
Dyspnoe	18 (7,1) n. a. [n. a.; n. a.]	28 (11,5) n. a. [n. a.; n. a.]	0,57 [0,31; 1,04] 0,06
Husten mit Auswurf	11 (4,4) n. a. [n. a.; n. a.]	15 (6,1) n. a. [n. a.; n. a.]	0,52 [0,24; 1,16] 0,11
Erkrankungen der Haut und des Unterhautgewebes	56 (22,2) n. a. [n. a.; n. a.]	52 (21,3) n. a. [n. a.; n. a.]	0,91 [0,62; 1,34] 0,64
Pruritus	22 (8,7) n. a. [n. a.; n. a.]	9 (3,7) n. a. [n. a.; n. a.]	2,14 [0,98; 4,67] 0,051
Ausschlag	21 (8,3) n. a. [n. a.; n. a.]	12 (4,9) n. a. [n. a.; n. a.]	1,55 [0,76; 3,17] 0,22
Alopezie	8 (3,2) n. a. [n. a.; n. a.]	19 (7,8) n. a. [n. a.; n. a.]	0,34 [0,15; 0,77] 0,007
Gefäßerkrankungen	34 (13,5) n. a. [n. a.; n. a.]	27 (11,1) n. a. [n. a.; n. a.]	1,17 [0,70; 1,94] 0,57
Hypertonie	14 (5,6) n. a. [n. a.; n. a.]	4 (1,6) n. a. [n. a.; n. a.]	3,26 [1,07; 9,95] 0,029
Hypotension	11 (4,4) n. a. [n. a.; n. a.]	6 (2,5) n. a. [n. a.; n. a.]	1,62 [0,59; 4,42] 0,34

¹⁾ Die Kodierung erfolgte nach Angaben im CSR nach MedDRA-Version 27.1.

²⁾ Die Sicherheitspopulation umfasst alle Personen, die zu einem Behandlungsarm randomisiert wurden und mind. eine Medikationsdosis erhalten haben.

³⁾ Das 95%-KI basiert auf der geschätzten Varianz der log-log-Transformation der Kaplan-Meier-Schätzung.

⁴⁾ Der zweiseitige p-Wert basiert auf einem unstratifizierten Log-Rank-Test; p-Werte unter 0,05 werden als statistisch signifikant angesehen.

Abkürzungen: CSR: Klinischer Studienbericht; HR: Hazard Ratio; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; n. a.: nicht anwendbar; UE: Unerwünschtes Ereignis.

Schwere unerwünschte Ereignisse (CTCAE-Grad ≥ 3)

Tabelle 20: Schwere UE (CTCAE-Grad ≥ 3), die bei $\geq 5\%$ der Personen bzw. ≥ 10 Personen und $\geq 1\%$ der Personen in einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term; Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

Studie DeLLphi-304 Schwere UE (CTCAE-Grad ≥ 3) MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Erkrankungen des Blutes und des Lymphsystems	40 (15,9) n. a. [n. a.; n. a.]	117 (48,0) 4,8 [2,9; 8,3]	0,22 [0,15; 0,32] < 0,001
Anämie	11 (4,4) n. a. [n. a.; n. a.]	70 (28,7) n. a. [n. a.; n. a.]	0,11 [0,06; 0,21] < 0,001
Neutropenie	15 (6,0) n. a. [n. a.; n. a.]	57 (23,4) n. a. [10,2; n. a.]	0,21 [0,12; 0,37] < 0,001
Leukopenie	4 (1,6) n. a. [n. a.; n. a.]	34 (13,9) n. a. [n. a.; n. a.]	0,10 [0,04; 0,29] < 0,001
Thrombozytopenie	2 (0,8) n. a. [n. a.; n. a.]	28 (11,5) n. a. [n. a.; n. a.]	0,06 [0,01; 0,25] < 0,001
Febrile Neutropenie	5 (2,0) n. a. [n. a.; n. a.]	28 (11,5) n. a. [n. a.; n. a.]	0,13 [0,05; 0,35] < 0,001
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	25 (9,9) n. a. [n. a.; n. a.]	31 (12,7) n. a. [n. a.; n. a.]	0,67 [0,39; 1,14] 0,13
Fatigue	9 (3,6) n. a. [n. a.; n. a.]	17 (7,0) n. a. [n. a.; n. a.]	0,47 [0,21; 1,05] 0,06
Infektionen und parasitäre Erkrankungen	32 (12,7) n. a. [n. a.; n. a.]	42 (17,2) n. a. [n. a.; n. a.]	0,62 [0,39; 0,99] 0,04
Pneumonie	14 (5,6) n. a. [n. a.; n. a.]	20 (8,2) n. a. [n. a.; n. a.]	0,56 [0,28; 1,11] 0,09
Untersuchungen	33 (13,1) n. a. [n. a.; n. a.]	54 (22,1) n. a. [n. a.; n. a.]	0,48 [0,31; 0,74] < 0,001
Leukozytenzahl vermindert	5 (2,0) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	0,36 [0,12; 1,04] 0,049
Neutrophilenzahl vermindert	10 (4,0) n. a. [n. a.; n. a.]	28 (11,5) n. a. [n. a.; n. a.]	0,28 [0,13; 0,58] < 0,001
Thrombozytenzahl vermindert	1 (0,4) n. a. [n. a.; n. a.]	20 (8,2) n. a. [n. a.; n. a.]	0,04 [0,01; 0,33] < 0,001
Stoffwechsel- und Ernährungsstörungen	27 (10,7) n. a. [n. a.; n. a.]	33 (13,5) n. a. [n. a.; n. a.]	0,69 [0,41; 1,15] 0,15
Hyponatriämie	12 (4,8) n. a. [n. a.; n. a.]	13 (5,3) n. a. [n. a.; n. a.]	0,78 [0,35; 1,72] 0,54
Erkrankungen des Nervensystems	8 (3,2) n. a. [n. a.; n. a.]	14 (5,7) n. a. [n. a.; n. a.]	0,48 [0,20; 1,14] 0,09
Erkrankungen der Atemwege, des Brustraums und Mediastinums	20 (7,9) n. a. [n. a.; n. a.]	16 (6,6) n. a. [n. a.; n. a.]	1,06 [0,54; 2,05] 0,87

¹⁾ Die Kodierung erfolgte nach Angaben im CSR nach MedDRA-Version 27.1.

²⁾ Die Sicherheitspopulation umfasst alle Personen, die zu einem Behandlungsarm randomisiert wurden und mind. eine Medikationsdosis erhalten haben.

³⁾ Das 95%-KI basiert auf der geschätzten Varianz der log-log-Transformation der Kaplan-Meier-Schätzung.

⁴⁾ Der zweiseitige p-Wert basiert auf einem unstratifizierten Log-Rank-Test; p-Werte unter 0,05 werden als statistisch signifikant angesehen.

Abkürzungen: CSR: Klinischer Studienbericht; CTCAE: Common Terminology Criteria for Adverse Events; HR: Hazard Ratio; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; n. a.: nicht anwendbar; UE: Unerwünschtes Ereignis.

Schwerwiegende unerwünschte Ereignisse

*Tabelle 21: SUE, die bei ≥ 5 % der Personen bzw. ≥ 10 Personen und ≥ 1 % der Personen in einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term;
Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)*

Studie DeLLphi-304 SUE MedDRA-Systemorganklasse ¹⁾ Preferred Term	Tarlatamab N = 252 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	Chemotherapie N = 244 ²⁾ Personen mit Ereignis, n (%) Mediane Zeit bis zum Ereignis, Monate [95%-KI] ³⁾	HR [95%-KI] p-Wert ⁴⁾
Erkrankungen des Blutes und des Lymphsystems	6 (2,4) n. a. [n. a.; n. a.]	45 (18,4) n. a. [n. a.; n. a.]	0,10 [0,04; 0,24] < 0,001
Febrile Neutropenie	5 (2,0) n. a. [n. a.; n. a.]	24 (9,8) n. a. [n. a.; n. a.]	0,16 [0,06; 0,42] < 0,001
Anämie	1 (0,4) n. a. [n. a.; n. a.]	10 (4,1) n. a. [n. a.; n. a.]	0,09 [0,01; 0,70] 0,004
Thrombozytopenie	0 (0,0) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	n. a. [n. a.; n. a.] < 0,001
Erkrankungen des Gastrointestinaltrakts	13 (5,2) n. a. [n. a.; n. a.]	5 (2,0) n. a. [n. a.; n. a.]	2,32 [0,82; 6,53] 0,10
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	25 (9,9) n. a. [n. a.; n. a.]	6 (2,5) n. a. [n. a.; n. a.]	3,87 [1,58; 9,46] 0,001
Fieber	14 (5,6) n. a. [n. a.; n. a.]	0 (0,0) n. a. [n. a.; n. a.]	n. a. [n. a.; n. a.] < 0,001
Erkrankungen des Immunsystems	43 (17,1) n. a. [n. a.; n. a.]	1 (0,4) n. a. [n. a.; n. a.]	45,7 [6,29; 331,6] < 0,001
Zytokin-Freisetzungssyndrom	43 (17,1) n. a. [n. a.; n. a.]	1 (0,4) n. a. [n. a.; n. a.]	45,7 [6,29; 331,6] < 0,001
Infektionen und parasitäre Erkrankungen	28 (11,1) n. a. [n. a.; n. a.]	43 (17,6) n. a. [n. a.; n. a.]	0,53 [0,32; 0,85] 0,008
Pneumonie	10 (4,0) n. a. [n. a.; n. a.]	21 (8,6) n. a. [n. a.; n. a.]	0,38 [0,18; 0,82] 0,01
Erkrankungen des Nervensystems	16 (6,3) n. a. [n. a.; n. a.]	10 (4,1) n. a. [n. a.; n. a.]	1,46 [0,66; 3,23] 0,35
Erkrankungen der Atemwege, des Brustraums und Mediastinums	18 (7,1) n. a. [n. a.; n. a.]	11 (4,5) n. a. [n. a.; n. a.]	1,39 [0,65; 2,97] 0,39

¹⁾ Die Kodierung erfolgte nach Angaben im CSR nach MedDRA-Version 27.1.

²⁾ Die Sicherheitspopulation umfasst alle Personen, die zu einem Behandlungsarm randomisiert wurden und mind. eine Medikationsdosis erhalten haben.

³⁾ Das 95%-KI basiert auf der geschätzten Varianz der log-log-Transformation der Kaplan-Meier-Schätzung.

⁴⁾ Der zweiseitige p-Wert basiert auf einem unstratifizierten Log-Rank-Test. P-Werte unter 0,05 werden als statistisch signifikant angesehen.

Abkürzungen: CSR: Klinischer Studienbericht; HR: Hazard Ratio; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; n. a.: nicht anwendbar; SUE: Schwerwiegendes unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

Für die vorliegende Nutzenbewertung wurde vom pU sowohl das deutsche Dossier zur Nutzenbewertung als auch das europäische JCA-Dossier vorgelegt. Im deutschen Dossier verweist der pU auf Angaben und Auswertungen, die bereits im europäischen JCA-Dossier vorgelegt wurden. Weitere Angaben zu den Inhalten des JCA-Dossiers, die in der vorliegenden Nutzenbewertung herangezogen werden konnten, finden sich in Kapitel 4.8.

Da der JCA-Bericht zu Tarlatamab (JCA-MP-2024-17) zu Beginn des deutschen Nutzenbewertungsverfahrens noch nicht veröffentlicht war, konnte dieser für die vorliegende Nutzenbewertung nicht berücksichtigt werden.

4.1 Design und Methodik der Studie

Studiendesign

Die vorliegende Nutzenbewertung stützt sich auf die laufende, pivotale Studie DeLLphi-304, eine multizentrische, randomisierte, kontrollierte, Open-label-Phase-III-Studie bei erwachsenen Patientinnen und Patienten mit SCLC und Krankheitsprogression während oder nach einer vorangegangenen platinbasierten Chemotherapie. Die Randomisierung erfolgte stratifiziert nach vorheriger Exposition mit PD-(L)1-Inhibitoren, Chemotherapie-freies Intervall, Vorliegen von Hirnmetastasen und Art der Chemotherapie (Wirkstoff) im 1:1-Verhältnis. Der primäre Wirksamkeitsendpunkt der Studie DeLLphi-304 war „Gesamtüberleben“.

Nach einer bis zu 21-tägigen Screeningphase folgte die randomisierte Behandlungsphase mit Tarlatamab (in 28-Tage-Zyklen) bzw. einer Chemotherapie (Topotecan, Lurbinectedin oder Amrubicin, je nach lokalen Bestimmungen; in 21-Tage-Zyklen). Die Behandlung erfolgte bis zur radiologischen Krankheitsprogression (gemäß RECIST), zum Auftreten inakzeptabler Toxizität, zum Widerruf der Einwilligung, zum Tod oder zum Studienende, je nachdem, was zuerst eintrat. Unter bestimmten Voraussetzungen war eine Fortführung der Verabreichung der Studienmedikation nach radiologischer Progression möglich; im Tarlatamab-Arm war dies bei 77 Personen (30,6 %) der Fall. Nach dauerhafter Beendigung der Studienmedikation aufgrund jeglicher Ursache erfolgt ein SFU 60 (+ 5) Tage nach der letzten Dosis der Studienmedikation. Im anschließenden LTFU erfolgt die Nachverfolgung des Überlebensstatus (alle 12 ($\pm$ 2) Wochen) sowie nachfolgender Krebstherapien.

Nach Angaben im CSR wurde die Interimsanalyse als primäre Analyse der Studie DeLLphi-304 deklariert, da bereits statistische Signifikanz für den primären Endpunkt (Gesamtüberleben) erreicht wurde. Der zugehörige Datenschnitt vom 29.01.2025, der auch im Rahmen der Zulassung herangezogen wurde, wird in der vorliegenden Nutzenbewertung dargestellt. Die Studienteilnahme hatten zum Zeitpunkt des Datenschnitts 111 Personen (43,7 %) im Tarlatamab-Arm (alle aufgrund von Tod) und 153 Personen (60,0 %) im Kontrollarm (davon 151 aufgrund von Tod) abgebrochen.

Nach Abbruch der Studienmedikation haben 112 Personen (44,4 %) im Tarlatamab-Arm und 125 (51,2 %) im Chemotherapie-Arm mindestens eine Folgetherapie erhalten. Die häufigste erste Folgetherapie war eine Mono-Chemotherapie in beiden Behandlungsarmen (40,5 % im Tarlatamab-Arm bzw. 36,5 % im Chemotherapie-Arm), wobei Carboplatin am häufigsten eingesetzt wurde (14,3 % bzw. 11,5 %).

Studienpopulation

In die Studie wurden Personen $\geq$ 18 Jahre sowie mit histologisch bzw. zytologisch bestätigtem SCLC mit nachgewiesener Progression oder Rezidiv nach platinbasierter Therapie eingeschlossen. Die Studienpopulation bestand aus 509 Personen, wovon 254 in den Tarlatamab-Arm und 255 in den Chemotherapie-Arm randomisiert wurden. Im Chemotherapie-Arm erhielten die Personen

mehrheitlich Topotecan (n = 185), gefolgt von Lurbinectedin (n = 47) und Amrubicin (n = 23). Die Sicherheitspopulation umfasste 496 Personen, da 2 Personen im Tarlatamab-Arm und 11 Personen im Chemotherapie-Arm keine Dosis der zugewiesenen Studienmedikation erhalten haben.

Die teilnehmenden Personen waren im Median 64 bzw. 66 Jahre alt. In beiden Studienarmen nahmen mehr männliche als weibliche Personen teil. Die Geschlechterverteilung ist kongruent mit der Epidemiologie der Erkrankung. Anzumerken ist, dass die Patientinnen und Patienten per Einschlusskriterium fast ausschließlich einen guten ECOG-PS (0 oder 1) hatten. Die Baseline-Charakteristika waren größtenteils ausgeglichen zwischen den beiden Studienarmen.

Studienmedikation

Im Rahmen der Studie DeLLphi-304 erhielten Personen im Interventionsarm Tarlatamab in 28-Tage-Zyklen mit einer Dosierung von 10 mg, verabreicht als intravenöse Infusion. Es war eine Dosis-Titration von 1 mg an Tag 1 von Zyklus 1 vorgesehen. Diese diente zur Reduktion des Risikos eines Zytokin-Freisetzungssyndroms.

Personen im Kontrollarm erhielten in 21-Tage-Zyklen ein Chemotherapeutikum. Je nach lokalen Bestimmungen handelte es sich hierbei um Topotecan (in allen Ländern mit Ausnahme von Japan), Lurbinectedin (in Kanada, Australien, Korea, Singapur, USA) oder Amrubicin (nur in Japan).

Die mediane Behandlungsdauer für Tarlatamab lag bei 4,19 Monaten und war damit länger als die mediane Behandlungsdauer im Kontrollarm (2,51 Monate). Der Anteil Personen mit einer Behandlungsdauer ≥ 3 Monate war im Tarlatamab-Arm etwa 15 % höher als im Chemotherapie-Arm (44,7 % vs. 59,1 %). Zum Zeitpunkt des Datenschnitts hatten 183 Personen (72,0 %) im Tarlatamab-Arm und 225 Personen (88,2 %) im Chemotherapie-Arm die Einnahme der Studienmedikation abgebrochen, die meisten aufgrund einer Krankheitsprogression (152 (59,5 %) bzw. 159 (62,4 %) Personen). Damit lag der Anteil der Personen, die die Studienmedikation abgebrochen haben, im Kontrollarm um mehr als 16 % höher als im Tarlatamab-Arm.

4.2 Zulassungsstatus und Zulassungspopulation

Zulassungsstatus und Anwendungsgebiet

Tarlatamab (IMDYLLTRA®) ist als Monotherapie zugelassen für die Behandlung erwachsener Patienten mit kleinzelligem Lungenkarzinom im fortgeschrittenen Stadium (ES-SCLC), die nach Krankheitsprogression während oder nach einer Erstlinienbehandlung basierend auf platinhaltiger Chemotherapie eine systemische Therapie benötigen. Die Anwendung in der Studie DeLLphi-304 erfolgte gemäß dem empfohlenen Dosierungsschema von Tarlatamab: Anfangsdosis von 1 mg an Tag 1 Zyklus 1, gefolgt von 10 mg an den Tagen 8 und 15 sowie anschließend 10 mg alle 2 Wochen. Ferner erhielten die Personen im Tarlatamab-Arm eine Prämedikation mit Dexamethason sowie eine Postmedikation mit einer Hydrations-Kochsalzlösung, was den Empfehlungen der Fachinformation für die Tage 1 und 8 in Zyklus 1 entspricht [11]. Nach Angaben in der Fachinformation wird keine Dosisreduktion von Tarlatamab empfohlen, was mit dem Protokoll der Studie DeLLphi-304 übereinstimmt.

Darüber hinaus ist anzumerken, dass in der Studie DeLLphi-304 ein relevanter Anteil (30,6 %) der Patientinnen und Patienten im Tarlatamab-Arm nach initialer radiologischer Progression weiterhin die Studienmedikation erhalten hat. Dies lässt sich als nicht konform mit der Fachinformation bewerten, da die Personen demnach bis zum Fortschreiten der Erkrankung oder bis zum Auftreten einer nicht tolerierbaren Toxizität behandelt werden sollen [11]. Zudem ist darauf hinzuweisen, dass 8,3 % der Personen im Tarlatamab-Arm sowie 7,1 % im Chemotherapie-Arm ein Krankheitsstadium III während des Screenings aufwiesen. Außerdem lag bei 10,6 % bzw. 7,5 % der Personen keine Metastasierung zu Baseline vor.

Übertragbarkeit auf deutschen Versorgungskontext

Die Studie DeLLphi-304 wurde in 166 Studienzentren in den Ländern China, Israel, Japan, Korea, Malaysia, Singapur, Taiwan, Österreich, Belgien, Tschechische Republik, Dänemark, Frankreich, Deutschland, Griechenland, Ungarn, Italien, Niederlande, Polen, Portugal, Rumänien, Spanien, Schweiz, Türkei, Vereinigtes Königreich Großbritannien und Nordirland, Kanada, USA, Argentinien, Australien, Brasilien und Mexiko durchgeführt. Insgesamt wurden 240 Personen (47,2 %) in Europa randomisiert, 207 Personen (40,7 %) in Asien, 28 Personen (5,5 %) in Nordamerika sowie 34 Personen im Rest der Welt (hierunter 12 (2,4 %) in Australien). Nur 11 Personen (2,2 %) wurden in Deutschland randomisiert.

Mit der globalen Durchführung der Studie sind Einschränkungen der Ergebnissicherheit mit Blick auf die jeweils eingesetzten Komparatoren zu beachten. In der EU handelt es sich bei Topotecan um die einzige Therapieoption, die für die Behandlung des SCLC in der Zweitlinie zugelassen ist [12]. Kürzlich wurde Lurbinectedin in Kombination mit Atezolizumab in der EU zugelassen, allerdings zur Erhaltungstherapie bei erwachsenen Patienten mit ES-SCLC, was somit einem abweichenden Anwendungsgebiet entspricht [14]. Entsprechend wurden 72,5 % (185 Personen) des Kontrollarms gemäß der Zulassung behandelt bzw. erhielten 27,5 % (70 Personen) keine in Deutschland/Europa zugelassene Behandlung. Mit Modul 4 legte der pU keine vergleichenden Auswertungen für Tarlatamab vs. Topotecan vor. Damit ist bei den Ergebnissen eine mögliche Verzerrung (Unter- bzw. Überschätzung) durch die nicht zugelassenen Therapieoptionen zu beachten. Ferner könnte die Einschlussvoraussetzung eines ECOG-PS von 0 oder 1 die Übertragbarkeit der Ergebnisse einschränken. Darüber hinaus war eine Behandlung über die Krankheitsprogression hinaus möglich.

Insgesamt ist die Übertragbarkeit auf den deutschen Versorgungskontext als eingeschränkt zu bewerten.

4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit

Zusammenfassend zeigt die Studie DeLLphi-304, aufgrund ihres randomisiert-kontrollierten Studiendesigns, ein niedriges Verzerrungspotential auf Studienebene. Die Ergebnissicherheit wird allerdings eingeschränkt durch die kurze Nachbeobachtungsdauer nach Behandlungsabbruch, was die langfristige Erfassung von Morbidität, Lebensqualität und Sicherheit erschwert.

Auf Endpunktebene, insbesondere für die patientenberichteten Endpunkte und „Unerwünschte Ereignisse“, wird das Verzerrungspotential jedoch als hoch eingeschätzt. Gründe hierfür sind:

- Nicht-verblindetes Studiendesign.
- Unterschiedliche Behandlungszyklen im Interventions- und Kontrollarm führen zu einer ungleichen Abbildung der Therapiebelastung.
- Die Rücklaufquoten sinken im Studienverlauf frühzeitig.
- Im weiteren Studienverlauf sind differierende Rücklaufquoten zwischen den beiden Studienarmen zu beachten.

In der Studie DeLLphi-304 wurden 70 Personen (27,5 % des Kontrollarms) mit einem nicht in Europa zugelassenen Komparator behandelt.

4.4 Mortalität

Im Tarlatamab-Arm traten zum vorliegenden Datenschnitt (29.01.2025) weniger Todesfälle auf als im Chemotherapie-Arm (111 vs. 152 Ereignisse). Für den Endpunkt „Gesamtüberleben“ liegt ein statistisch signifikanter Unterschied zugunsten von Tarlatamab gegenüber Chemotherapie vor (HR: 0,60 (95%-KI: [0,47; 0,77])). In der Kaplan-Meier-Kurve zeigt sich, dass die Kurven bereits zu Monat 2 auseinanderlaufen und der Unterschied zwischen den Behandlungsarmen über den

gesamten Studienzeitraum bestehen bleibt. Aufgrund der objektiven Operationalisierung des Endpunkts „Gesamtüberleben“ wurde das Verzerrungspotential als niedrig bewertet.

Die Studie DeLLphi-304 ist noch nicht abgeschlossen. Es ist anzumerken, dass 56,3 % der Personen im Tarlatamab-Arm sowie 40,0 % im Chemotherapie-Arm zum vorliegenden Datenschnitt noch an der Studie teilnehmen. Der vorliegende Datenschnitt basiert auf der Interimsanalyse, die als primäre Analyse genutzt wurde. Entsprechend ist eine Änderung im Effektschätzer zum Gesamtüberleben – im Sinne der Erhebung im Rahmen des LTFU – mit einem späteren Datenschnitt zur finalen Analyse noch möglich.

Limitierend ist anzumerken, dass ein relevanter Anteil (30,6 %) der Patientinnen und Patienten im Tarlatamab-Arm nach initialer radiologischer Progression weiterhin die Studienmedikation erhalten hat. Dies entspricht nicht der Behandlung gemäß der Fachinformation [11]. Es bleibt unklar, ob das Ausmaß der Entscheidung die Therapie fortzuführen durch das offene Studiendesign beeinflusst wurde und zu dem beobachteten Unterschied zwischen den beiden Behandlungsgruppen beigetragen hat. Informationen zur Fortführung der Therapie im Chemotherapie-Arm fehlen in den Studienunterlagen.

4.5 Morbidität

In der Endpunktkategorie „Morbidität“ wurden in der vorliegenden Nutzenbewertung die Symptomskalen des EORTC QLQ-C30 herangezogen. Weitere PRO wurden aufgrund der Erhebungs- und Datensituation nicht in der Nutzenbewertung herangezogen. Die detaillierte Bewertung der Endpunkte ist in Kapitel 2.3.2 abgebildet. Weiterführende Informationen sind Kapitel 2.4 zu entnehmen. Der pU legte mit Modul 4 für die Symptomskalen des EORTC QLQ-C30 Responderanalysen für die Zeit bis zur Verschlechterung mit einer adäquaten Responseschwelle von 10 Punkten vor.

Für den EORTC QLQ-C30 erfolgte die Erhebung für die ersten 12 Wochen (entspricht den ersten 3 bzw. 4 Therapiezyklen) engmaschig im wöchentlichen Rhythmus.

Für den EORTC QLQ-C30 zeigten sich in der Symptomskala „Appetitverlust“ signifikante Unterschiede zwischen den Behandlungsarmen zuungunsten von Tarlatamab (HR: 1,27 (95%-KI: [1,02; 1,59])). Ein Verschlechterungsereignis wurde bei 171 Personen (67,3 %) im Tarlatamab-Arm und bei 151 (59,2 %) im Chemotherapie-Arm beobachtet, wobei die mediane Zeit bis zur Verschlechterung im Interventionsarm kürzer war als im Kontrollarm (3,3 vs. 5,1 Monate). In der Symptomskala „Diarrhö“ zeigten sich hingegen signifikante Unterschiede zugunsten von Tarlatamab (HR: 0,67 (95%-KI: [0,49; 0,92])). Ein Verschlechterungsereignis wurde bei 70 Personen (27,6 %) im Tarlatamab-Arm sowie bei 90 (35,3 %) im Chemotherapie-Arm beobachtet. Für weitere Symptomskalen des EORTC QLQ-C30 zeigten sich keine signifikanten Unterschiede zwischen den Behandlungsarmen.

Bei der Interpretation der beobachteten Ergebnisse ist eine relevante Einschränkung der Ergebnis-sicherheit zu beachten. Das Verzerrungspotential des Endpunkts wird als „hoch“ bewertet. Dies ist neben dem Open-label-Design der Studie auf die schnell sinkenden Rücklaufquoten sowie differierenden Rücklaufquoten zwischen den beiden Behandlungsarmen zurückzuführen. Außerdem ist auf eine unterschiedliche Abbildung der Therapiebelastung durch unterschiedliche Therapiezyklen in den Behandlungsarmen bei gleichen Erhebungszeitpunkten hinzuweisen, was zu einer ungleichen Erhebung von Verschlechterungsereignissen führen kann. Dies ist insbesondere bei akuten Symptomen zu beachten. Ferner ist aufgrund der verkürzten Nachbeobachtungsdauer nach Behandlungs-/Studienabbruch eine potentielle Verzerrung durch informative Zensierung zu beachten.

4.6 Lebensqualität

Für die Endpunktkategorie „Lebensqualität“ legte der pU Ereigniszeitanalysen (Zeit bis zur Verschlechterung) für das patientenberichtete Instrument EORTC QLQ-C30 vor. Für die Funktions-skalen sowie den globalen Gesundheitsstatus zeigten sich keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Analog zu den Morbiditätspunkten (s. vorheriges Kapitel) sind relevante Einschränkungen der Ergebnissicherheit, zusammenhängend mit dem hohen Verzerrungspotential, zu beachten.

4.7 Sicherheit

Die Erfassung von UE erfolgt bei allen Personen, die mindestens eine Dosis der Studienmedikation erhalten haben. Es wurden diejenigen Ereignisse als UE definiert, die nach Gabe der ersten Dosis der Studienmedikation (bzw. auch ein Nicht-Prüfpräparat oder Hilfspräparat), im SFU oder bis 60 Tage (bzw. bis zu 65 Tage) nach Gabe der letzten Dosis der Studienmedikation (bzw. auch ein Nicht-Prüfpräparat oder Hilfspräparat, je nachdem, was später lag) auftrat. Der pU gibt an, Ereignisse der Grunderkrankung nicht in die TEAE-Analysen einzubeziehen. Die ausgewählten Preferred Terms erscheinen plausibel, um Progressionsereignisse abzubilden, jedoch werden Symptome der Grunderkrankung damit nicht erfasst. Folglich kann daher nicht ausgeschlossen werden, dass weiterhin Ereignisse der Grunderkrankung in den Ergebnissen enthalten sind.

Zwischen den beiden Behandlungsarmen gab es Unterschiede in der Behandlungsdauer: Im Tarlatamab-Arm betrug diese im Median 4,19 Monate, wohingegen diese im Chemotherapie-Arm mit einem Median von 2,51 Monaten kürzer war. Dieser Aspekt wurde mithilfe der eingereichten Ereigniszeitanalysen berücksichtigt. Dennoch ist dies bei der Interpretation der Ereignisraten zu beachten.

Auf Ebene der Gesamtraten wurden signifikante Unterschiede zwischen den beiden Behandlungsarmen für „Schwere UE (CTCAE-Grad ≥ 3)“ sowie „UE, das zum Abbruch der Studienmedikation führte“ identifiziert. Für schwere UE (CTCAE-Grad ≥ 3) lag der Effektschätzer HR bei 0,46 (95%-KI: [0,37; 0,58]). Damit wurde ein signifikanter Unterschied ($p < 0,001$) zugunsten von Tarlatamab identifiziert; im Tarlatamab-Arm traten bei 136 Personen (54,0 %) entsprechende Ereignisse auf, wohingegen es im Chemotherapie-Arm 195 Personen (79,9 %) waren. Ebenso war der Anteil von Personen, bei denen ein UE zum Abbruch der Therapie führte im Tarlatamab-Arm, geringer: 13 (5,2 %) im Tarlatamab-Arm vs. 30 (12,3 %) im Chemotherapie-Arm (HR: 0,37 (95%-KI: [0,19; 0,72]); $p = 0,002$).

Bei UE jeglichen Schweregrads mit Inzidenz ≥ 10 % bzw. ≥ 10 Personen und ≥ 1 % der Personen in einem Studienarm ließen sich signifikante Unterschiede identifizieren; auf Ebene der Systemorganklassen waren diese: Erkrankungen des Blutes und des Lymphsystems (HR: 0,38 (95%-KI: [0,30; 0,49])), Augenerkrankungen (HR: 0,25 (95%-KI: [0,09; 0,69])), Erkrankungen des Gastrointestinaltrakts (HR: 0,77 (95%-KI: [0,61; 0,97])), Erkrankungen des Immunsystems (HR: 50,7 (95%-KI: [18,9; 137,0])), Stoffwechsel- und Ernährungsstörungen (HR: 1,40 (95%-KI: [1,09; 1,79])), Erkrankungen des Nervensystems (HR: 2,21 (95%-KI: [1,66; 2,95])) sowie Erkrankungen der Nieren und Harnwege (HR: 2,14 (95%-KI: [1,12; 4,11])). Entsprechend zeigten sich in der Gesamtschau Effekte zugunsten und zuungunsten von Tarlatamab. Auffällig war insbesondere der große Effekt zuungunsten von Tarlatamab hinsichtlich der Systemorganklasse „Erkrankungen des Immunsystems“. Dies war insbesondere bedingt durch das Auftreten des Zytokin-Freisetzungssyndroms: Hiervon waren mehr als die Hälfte der Personen im Tarlatamab-Arm ($n = 142$ (56,3 %)) und lediglich 3 Personen (1,2 %) im Chemotherapie-Arm betroffen. Dieser Aspekt wurde bereits im Rahmen der Zulassung diskutiert [13]. Diese Immunreaktion steht dabei im Zusammenhang mit der Wirkweise von Tarlatamab als DLL3-gerichteter CD3 T-Zell-Engager.

Auf Ebene der schweren UE (CTCAE-Grad ≥ 3) wurden bei den Systemorganklassen „Erkrankungen des Blutes und des Lymphsystems“ (HR: 0,22 (95%-KI: [0,15; 0,32])), „Infektionen und parasitäre Erkrankungen“ (HR: 0,62 (95%-KI: [0,39; 0,99])) sowie „Untersuchungen“ (HR: 0,48 (95%-KI: [0,31; 0,74])) signifikante Unterschiede ($p < 0,05$) zugunsten von Tarlatamab beobachtet.

Auf Ebene der SUE lagen signifikante Unterschiede zugunsten von Tarlatamab für „Erkrankungen des Blutes und des Lymphsystems“ (HR: 0,10 (95%-KI: [0,04; 0,24])) sowie „Infektionen und parasitäre Erkrankungen“ (HR: 0,53 (95%-KI: [0,32; 0,85])) vor und zuungunsten von Tarlatamab für „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (HR: 3,87 (95%-KI: [1,58; 9,46])) sowie „Erkrankungen des Immunsystems“ (HR: 45,7 (95%-KI: [6,29; 331,6])). Bei „Erkrankungen des Immunsystems“ waren (analog zu jeglichen UE) alle Ereignisse auf das Zytokin-Freisetzungssyndrom zurückzuführen, für das im Tarlatamab-Arm 43 Personen (17,1 %) ein SUE aufwiesen und im Chemotherapie-Arm 1 Person (0,4 %). Auf Ebene der Gesamtrate der SUE war kein statistisch signifikanter Unterschied zwischen Tarlatamab (129 Personen (51,2 %)) und Chemotherapie (125 Personen (51,2 %)) zu beobachten. Mit Blick auf die Kaplan-Meier-Kurve (Abbildung 18) ist darauf hinzuweisen, dass sich die Kurven in den Monaten 3 bis 5 überkreuzen und die Kaplan-Meier-Kurve für Tarlatamab nachfolgend über der Kurve der Chemotherapie verläuft. Dies kann ein Hinweis auf eine Verletzung der Proportional-Hazard-Annahme sein. Entsprechend könnte die Größe des Effektschätzers über die Zeit variieren. Aufgrund der vergleichbaren Anzahl von Personen mit Ereignis in beiden Behandlungsarmen – trotz der längeren Beobachtungsdauer im Tarlatamab-Arm (6,31 vs. 4,42 Monate) – wird über die gesamte Behandlung hinweg von keinem Nachteil für Tarlatamab ausgegangen.

Im Rahmen der Sicherheit ist die limitierte Follow-up-Dauer von 60 (+5) Tagen nach Behandlungsende sowie die kurze Expositionsdauer mit Tarlatamab anzumerken. Dadurch besteht das Risiko einer informativen Zensurierung. Hierdurch ergeben sich Unsicherheiten in der Langzeitbewertung des Wirkstoffs Tarlatamab. Insgesamt ist bei der Interpretation der Ergebnisse in der Endpunktkategorie „Sicherheit“ das hohe Verzerrungspotential zu berücksichtigen, was die Ergebnissicherheit einschränkt. Die fehlende Verblindung könnte einen Einfluss auf die Meldung von UE jeglichen Schweregrads sowie die Entscheidung für den Abbruch der Studienbehandlung gehabt haben.

4.8 Zugrundeliegendes EU-HTA-Verfahren (JCA-MP-2024-17)

Grundlage für die Nutzenbewertung nach § 35a SGB V ist das nationale Dossier. In Modul 4 verweist der pU zum Teil auf das EU-Dossier im zugrundeliegenden EU-HTA-Verfahren. Hierunter fallen unter anderem die Charakterisierung der eingeschlossenen Studie DeLLphi-304, die Charakterisierung der Interventionen sowie die Operationalisierung einzelner Endpunkte. Im nationalen Dossier wird zu den in der vorliegenden Nutzenbewertung herangezogenen Endpunkten auf keine Ergebnisdarstellung im EU-Dossier verwiesen.

Für das nationale Nutzenbewertungsverfahren wurde auf Daten und Auswertungen zurückgegriffen, die im europäischen EU-HTA-Verfahren vorgelegt worden sind. Aus dem EU-Dossier wurden Daten aus Kapitel 4 und 5 sowie Appendix A1 und C berücksichtigt, die folgende Inhalte betreffen:

- Studiendesign und Studiencharakteristika
- Charakterisierung der Interventionen
- Studienpopulation/Patientencharakteristika
- Operationalisierung von Endpunkten
- Statistische Methoden

In Tabelle 5 der vorliegenden Nutzenbewertung wurde weiterhin gekennzeichnet, welche Ergebnisse bereits im EU-Dossier zum JCA dargestellt worden sind. Dabei ist anzumerken, dass mit dem EU-Dossier zum JCA Ergebnisse für die 7 definierten PICO-Fragestellungen für 4 unter-

schiedliche Populationen eingereicht wurden. Die Angaben im EU-Hauptdossier beziehen sich auf Ergebnisse zu den einzelnen PICO-Fragestellungen. Gegenstand des vorgelegten deutschen Dossiers war die abweichende, breiter gefasste Population gemäß zugelassenem Anwendungsgebiet „Erwachsene mit kleinzelligem Lungenkarzinom im fortgeschrittenen Stadium (extensive stage small-cell lung cancer, ES-SCLC), die nach Krankheitsprogression während oder nach einer Erstlinienbehandlung basierend auf platinhaltiger Chemotherapie eine systemische Therapie benötigen“, für welche die gesamte pivotale Studie DeLLphi-304 eingereicht wurde. Mit Appendix C wurden durch den pU Ergebnisse basierend auf der ITT-Population eingereicht. Die Analyse zum Gesamtüberleben ist dabei im JCA-Dossier (Appendix C) und Modul 4 des nationalen Dossiers identisch. Aufgrund der Auswertungsstrategie der PRO wurden keine Ergebnisse aus dem EU-Dossier für die in der vorliegenden Nutzenbewertung berücksichtigten Endpunkte extrahiert. Ferner wurden im EU-Dossier nur deskriptive Sicherheitsdaten auf Ebene der Gesamtraten dargestellt. Folglich wurden die entsprechenden Auswertungen berücksichtigt, die mit Modul 4 eingereicht wurden.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Tarlatamab (IMDYLLTRA®) ist zugelassen als Monotherapie zur Behandlung erwachsener Patienten mit kleinzelligem Lungenkarzinom im fortgeschrittenen Stadium (ES-SCLC), die nach Krankheitsprogression während oder nach einer Erstlinienbehandlung basierend auf platinhaltiger Chemotherapie eine systemische Therapie benötigen. Die Nutzenbewertung von Tarlatamab basiert auf der zulassungsbegründenden Studie DeLLphi-304, eine multizentrische, randomisierte, kontrollierte, Open-label-Phase-III-Studie bei erwachsenen Patientinnen und Patienten mit SCLC und Krankheitsprogression während oder nach einer vorangegangenen platinbasierten Chemotherapie.

Die Ergebnisse der in der Nutzenbewertung berücksichtigten Endpunkte der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Tarlatamab ein statistisch signifikant positiver Effekt bei hohem oder unklarem (↑) oder niedrigem (↑↑) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem (↓) oder niedrigem (↓↓) Verzerrungspotential bzw. kein Unterschied (↔) gezeigt werden konnte.

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

Tabelle 22: Zusammenfassende Darstellung der Ergebnisse der Studie DeLLphi-304

Studie DeLLphi-304 Darstellung der Ergebnisse	Tarlatamab N = 254		Chemotherapie N = 255		Tarlatamab vs. Chemotherapie	Effekt
Mortalität ¹⁾						
	N	Personen mit Ereignis, n (%) Medianes Überleben, Monate [95%-KI]	N	Personen mit Ereignis, n (%) Medianes Überleben, Monate [95%-KI]	HR [95%-KI] p-Wert	
Gesamtüberleben	254	111 (43,7) 13,6 [11,1; n. a.]	255	152 (59,6) 8,3 [7,0; 10,2]	0,60 [0,47; 0,77] < 0,001	↑↑
Morbidität ¹⁾						
EORTC QLQ-C30 – Symptomskalen	N	Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Wochen [95%-KI]	N	Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Wochen [95%-KI]	HR [95%-KI] p-Wert	
Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte						
Fatigue	254	194 (76,4) 2,3 [2,0; 2,4]	255	188 (73,7) 3,0 [2,3; 3,3]	1,17 [0,96; 1,43] 0,13	↔
Übelkeit und Erbrechen	254	119 (46,9) 18,4 [9,1; 47,8]	255	137 (53,7) 7,1 [5,3; 10,3]	0,79 [0,62; 1,02] 0,07	↔
Schmerz	254	156 (61,4) 4,6 [3,3; 6,8]	255	163 (63,9) 5,7 [4,3; 7,0]	0,97 [0,78; 1,21] 0,78	↔
Dyspnoe	254	124 (48,8) 18,1 [6,8; 30,9]	255	138 (54,1) 7,3 [5,3; 9,5]	0,79 [0,62; 1,01] 0,06	↔
Schlaflosigkeit	254	127 (50,0) 18,1 [7,5; 30,2]	255	132(51,8) 10,0 [7,4; 11,7]	0,88 [0,68; 1,12] 0,29	↔
Appetitverlust	254	171 (67,3) 3,3 [2,4; 4,6]	255	151 (59,2) 5,1 [4,3; 6,5]	1,27 [1,02; 1,59] 0,04	↓

Studie DeLLphi-304 Darstellung der Ergebnisse	Tarlatamab N = 254		Chemotherapie N = 255		Tarlatamab vs. Chemotherapie	Effekt
Obstipation	254	148 (58,3)7,4 [5,3; 10,1]	255	141 (55,3) 6,8 [5,1; 9,3]	0,97 [0,77; 1,22] 0,80	↔
Diarrhö	254	70 (27,6) n. a. [61,9; n. a.]	255	90 (35,3) 30,5 [18,1; n. a.]	0,67 [0,49; 0,92] 0,01	↑
Lebensqualität¹⁾						
EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus	<i>N</i>	<i>Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Wochen [95%-KI]</i>	<i>N</i>	<i>Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Wochen [95%-KI]</i>	<i>HR [95%-KI] p-Wert</i>	
Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte						
Globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“	254	172 (67,7) 2,6 [2,3; 3,7]	255	4,1 [3,3; 5,1] 172 (67,5)	1,07 [0,86; 1,32] 0,57	↔
Kognitive Funktion	254	139 (54,7) 7,7 [4,4; 12,2]	255	141 (55,3) 6,4 [5,3; 9,8]	0,96 [0,75; 1,21] 0,68	↔
Emotionale Funktion	254	112 (44,1) 25,6 [11,5; n. a.]	255	125 (49,0) 9,1 [6,3; 24,3]	0,78 [0,60; 1,01] 0,06	↔
Körperliche Funktion	254	154 (60,6) 4,3 [3,4; 6,0]	255	159 (62,4) 5,4 [4,1; 7,3]	0,98 [0,78; 1,23] 0,86	↔
Rollenfunktion	254	172 (67,7) 2,6 [2,3; 4,0]	255	176 (69,0) 4,1 [3,3; 4,6]	1,02 [0,82; 1,26] 0,87	↔
Soziale Funktion	254	147 (57,9) 3,3 [2,4; 6,5]	255	159 (62,4) 4,6 [3,6; 5,6]	0,95 [0,76; 1,20] 0,65	↔
Sicherheit²⁾³⁾						
Unerwünschte Ereignisse	<i>N</i>	<i>Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Monate [95%-KI]</i>	<i>N</i>	<i>Personen mit Ereignis, n (%) Mediane Zeit bis Ereignis, Monate [95%-KI]</i>	<i>HR [95%-KI] p-Wert</i>	
Schwere UE (CTCAE-Grad ≥ 3)	252	136 (54,0) 5,0 [3,3; 9,3]	244	195 (79,9) 1,1 [0,7; 1,5]	0,46 [0,37; 0,58] < 0,001	↑
SUE	252	129 (51,2) 5,1 [3,6; n. a.]	244	125 (51,2) 4,6 [3,4; 7,0]	1,03 [0,81; 1,33] 0,83	↔
UE, das zum Abbruch der Studien- medikation führte	252	13 (5,2) n. a. [n. a.; n. a.]	244	30 (12,3) n. a. [11,5; n. a.]	0,37 [0,19; 0,72] 0,002	↑

¹⁾ Die Analysen wurden basierend auf der ITT-Population durchgeführt.

²⁾ Weitere Ergebnisse zur Sicherheit (UE nach Organsystemen und Einzelereignisse) sind in Kapitel 3.5 dargestellt.

³⁾ Die Sicherheitspopulation umfasst alle Personen, die zu einem Behandlungsarm randomisiert wurden und mind. eine Medikationsdosis erhalten haben.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; EORTC QLQ-C30: EORTC Quality of Life Questionnaire – Core Questionnaire; HR: Hazard Ratio; ITT: Intention to Treat; KI: Konfidenzintervall; n. a.: nicht anwendbar; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **Amgen.** Dossier zur Nutzenbewertung - Modul 4 A: Tarlatamab (Imdylltra), Anhang 4-G: Ergänzende Daten zur Studie DeLLphi-304 [unveröffentlicht]. 13.05.2026.
2. **Amgen.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A: Tarlatamab (Imdylltra), Erwachsene mit ES-SCLC, die nach Krankheitsprogression während oder nach einer Erstlinienbehandlung basierend auf platinhaltiger Chemotherapie eine systemische Therapie benötigen; medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 30.06.2026.
3. **Amgen.** A randomized, open-label, phase 3 study of tarlatamab compared with standard of care in subjects with relapsed small cell lung cancer after platinum-based first-line chemotherapy (DeLLphi-304); clinical study report [unveröffentlicht]. 2025.
4. **Amgen.** A randomized, open-label, phase 3 study of tarlatamab compared with standard of care in subjects with relapsed small cell lung cancer after platinum-based first-line chemotherapy (DeLLphi-304); statistical analysis plan [unveröffentlicht]. 2025.
5. **Amgen.** A randomized, open-label, phase 3 study of tarlatamab compared with standard of care in subjects with relapsed small cell lung cancer after platinum-based first-line chemotherapy (DeLLphi-304); study protocol [unveröffentlicht]. 2025.
6. **Amgen.** A randomized, open-label, phase 3 study of tarlatamab compared with standard of care in subjects with relapsed small cell lung cancer after platinum-based first-line chemotherapy (DeLLphi-304); supplemental statistical analysis plan: key secondary and secondary patient reported outcome analysis [unveröffentlicht]. 2025.
7. **Amgen.** Study comparing tarlatamab with standard of care chemotherapy in relapsed small cell lung cancer (DeLLphi-304) [online]. NCT05740566. In: ClinicalTrials.gov. Last updated: 05.01.2026. [Zugriff: 07.09.2026]. URL: <https://clinicaltrials.gov/study/NCT05740566>.
8. **Amgen Europe.** Dossier of the Joint Clinical Assessment (JCA) of a medicinal product (MP): tarlatamab (imdylltra); JCA-MP-2024-17 [unveröffentlicht]. 06.07.2026.
9. **Amgen Europe.** Dossier of the Joint Clinical Assessment (JCA) of a medicinal product (MP): tarlatamab (imdylltra); JCA-MP-2024-17, appendix A.1: tables listing and information on methods of all studies included in the JCA [unveröffentlicht]. 06.07.2026.
10. **Amgen Europe.** Dossier of the Joint Clinical Assessment (JCA) of a medicinal product (MP): tarlatamab (imdylltra); JCA-MP-2024-17, appendix C: tarlatamab vs SOC (DeLLphi-304 ITT population) [unveröffentlicht]. 06.07.2026.
11. **Amgen Europe.** Imdylltra 1 mg Pulver zur Herstellung eines Konzentrats und Lösung zur Herstellung einer Infusionslösung; Imdylltra 10 mg Pulver zur Herstellung eines Konzentrats und Lösung zur Herstellung einer Infusionslösung [online]. 05.2026. Frankfurt/Main (GER). [Zugriff: 07.09.2026]. URL: <https://www.fachinfo.de/fi/pdf/025769/imdylltra-r-1-mg-10-mg-pulver-zur-herstellung-eines-konzentrats-und-loesung-zur-herstellung-einer-infusionsloesung>.
12. **European Medicines Agency (EMA).** Hycamtin: Zusammenfassung der Merkmale des Arzneimittels (EPAR Product information) [online]. Amsterdam (NED): EMA; 2026. [Zugriff: 07.09.2026]. URL: https://www.ema.europa.eu/de/documents/product-information/hycamtin-epar-product-information_de.pdf.
13. **European Medicines Agency (EMA).** Imdylltra (tarlatamab): European public assessment report EMEA/H/C/006451/0000 [online]. 26.03.2026. Amsterdam (NED): EMA. [Zugriff: 07.09.2026]. URL: https://www.ema.europa.eu/en/documents/assessment-report/imdylltra-epar-public-assessment-report_en.pdf.

14. **European Medicines Agency (EMA).** Zepzelca: Zusammenfassung der Merkmale des Arzneimittels (EPAR Product information) [online]. Amsterdam (NED): EMA; 2026. [Zugriff: 07.09.2026]. URL: https://www.ema.europa.eu/de/documents/product-information/zepzelca-epar-product-information_de.pdf.
15. **Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ.** Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol* 2010;63(7):760-766 e761. <https://dx.doi.org/10.1016/j.jclinepi.2009.09.009>.
16. **Nixon A, Doll H, Kerr C, Burge R, Naegeli AN.** Interpreting change from patient reported outcome (PRO) endpoints: patient global ratings of concept versus patient global ratings of change, a case study among osteoporosis patients. *Health Qual Life Outcomes* 2016;14:25. <https://dx.doi.org/10.1186/s12955-016-0427-5>.
17. **Norman GR, Stratford P, Regehr G.** Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 1997;50(8):869-879. [https://dx.doi.org/10.1016/s0895-4356\(97\)00097-8](https://dx.doi.org/10.1016/s0895-4356(97)00097-8).

Anhang

EORTC QLQ-C30 – Symptomskalen: Kaplan-Meier-Kurven

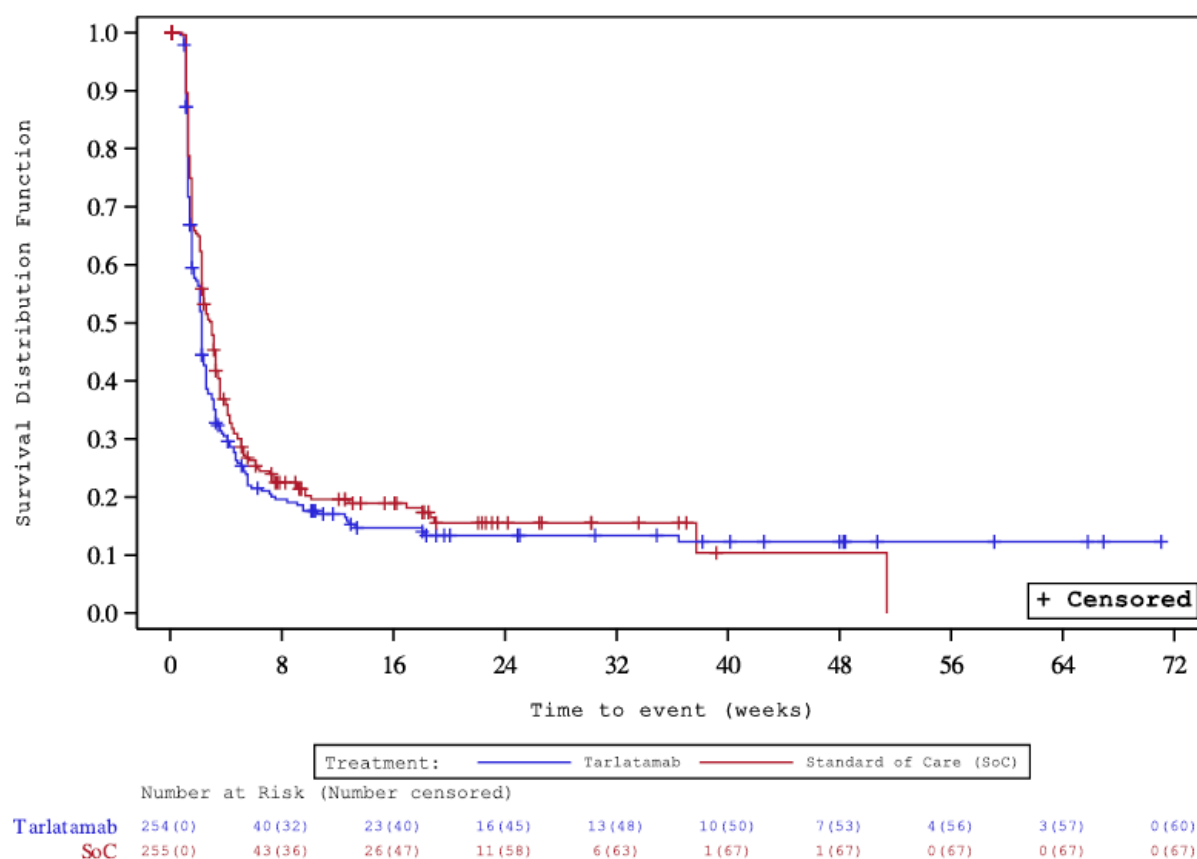


Abbildung 2: Kaplan-Meier-Kurve für die Symptomskala „Fatigue“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

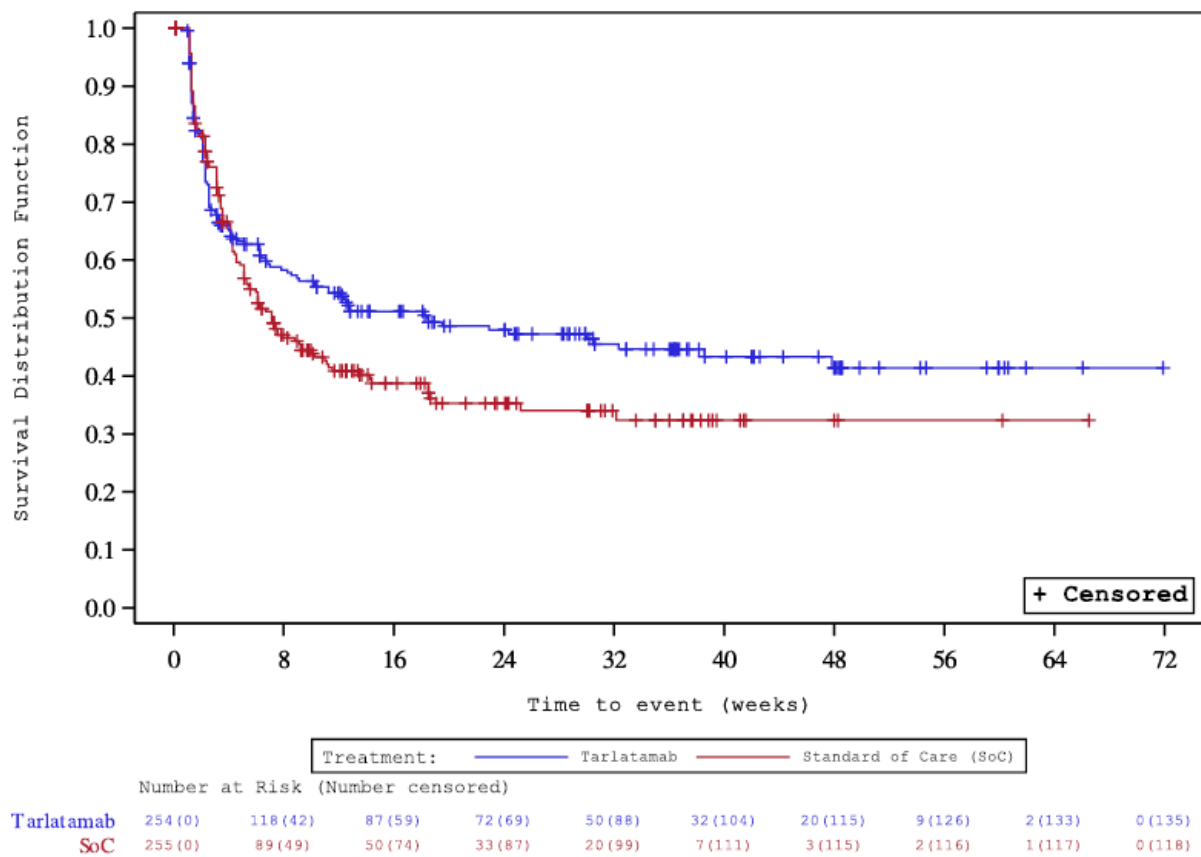


Abbildung 3: Kaplan-Meier-Kurve für die Symptomskala „Übelkeit und Erbrechen“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

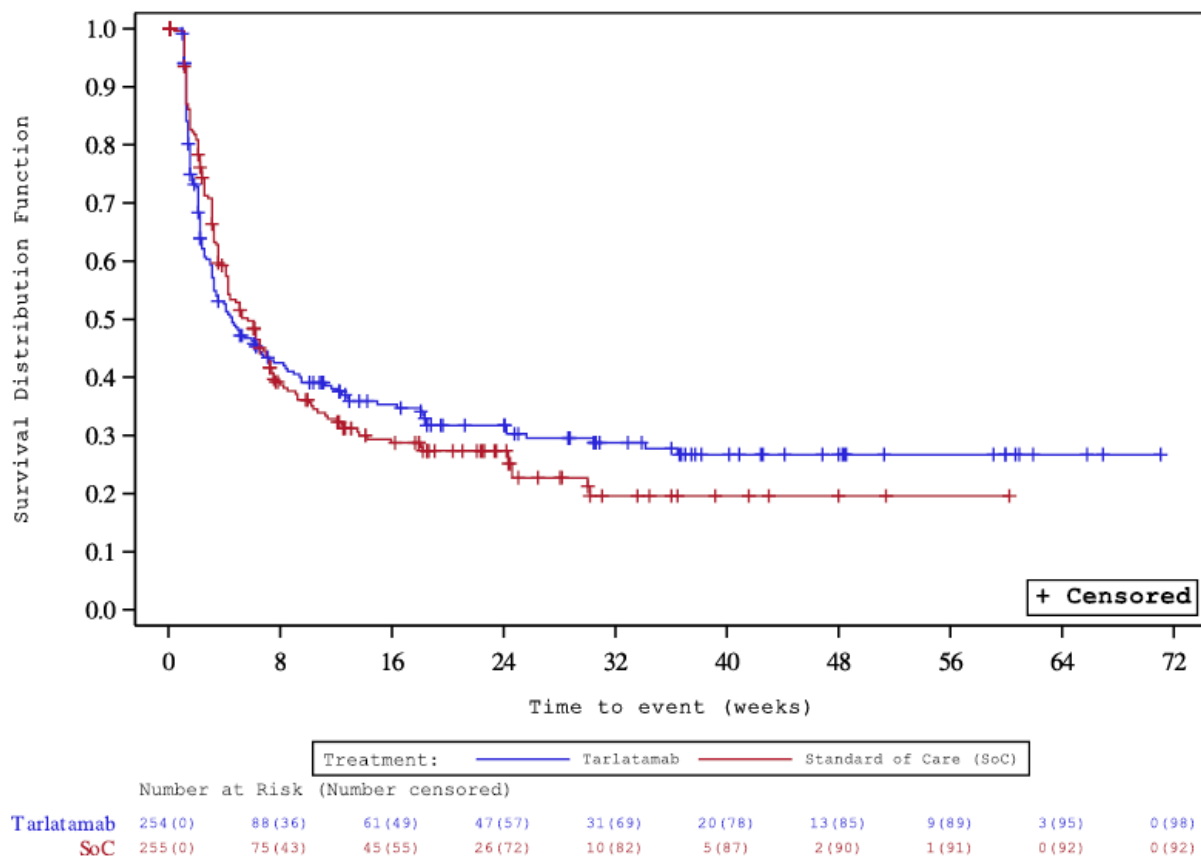


Abbildung 4: Kaplan-Meier-Kurve für die Symptomskala „Schmerz“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

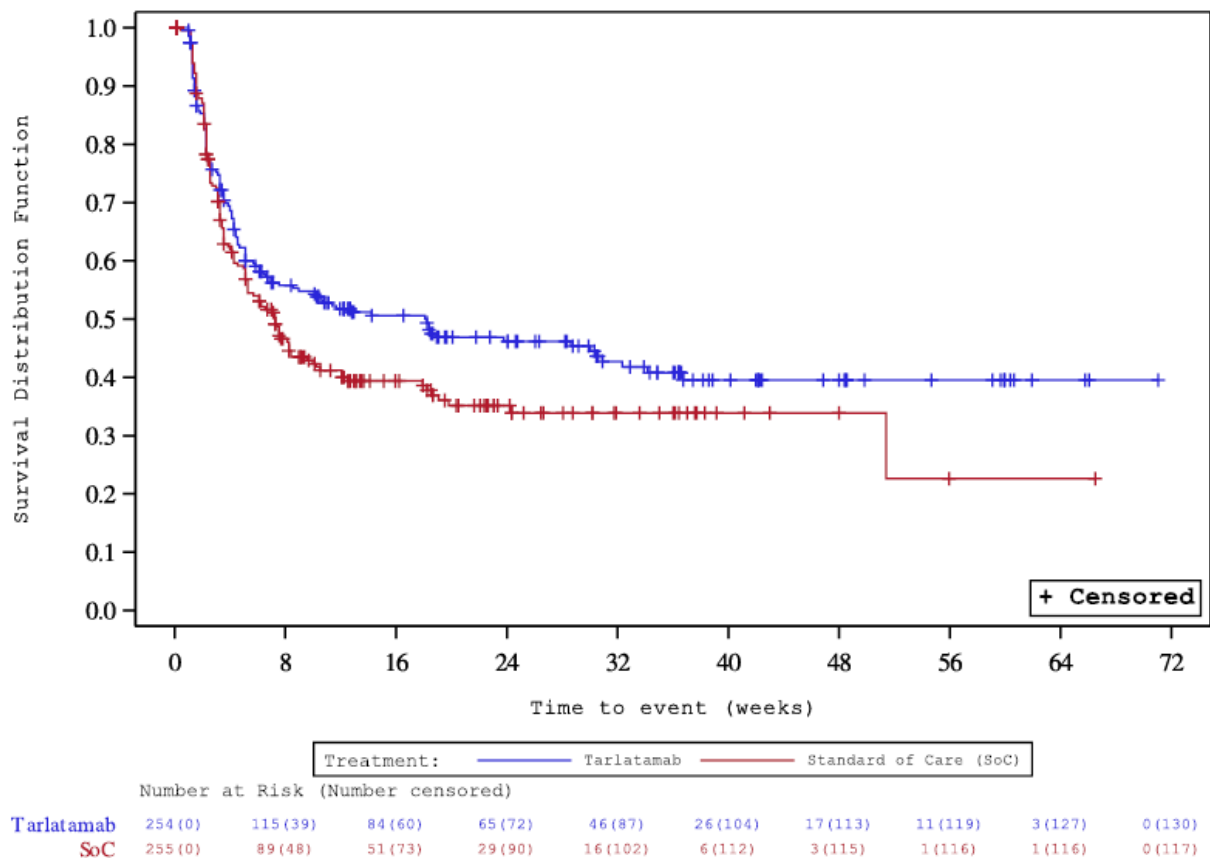


Abbildung 5: Kaplan-Meier-Kurve für die Symptomskala „Dyspnoe“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

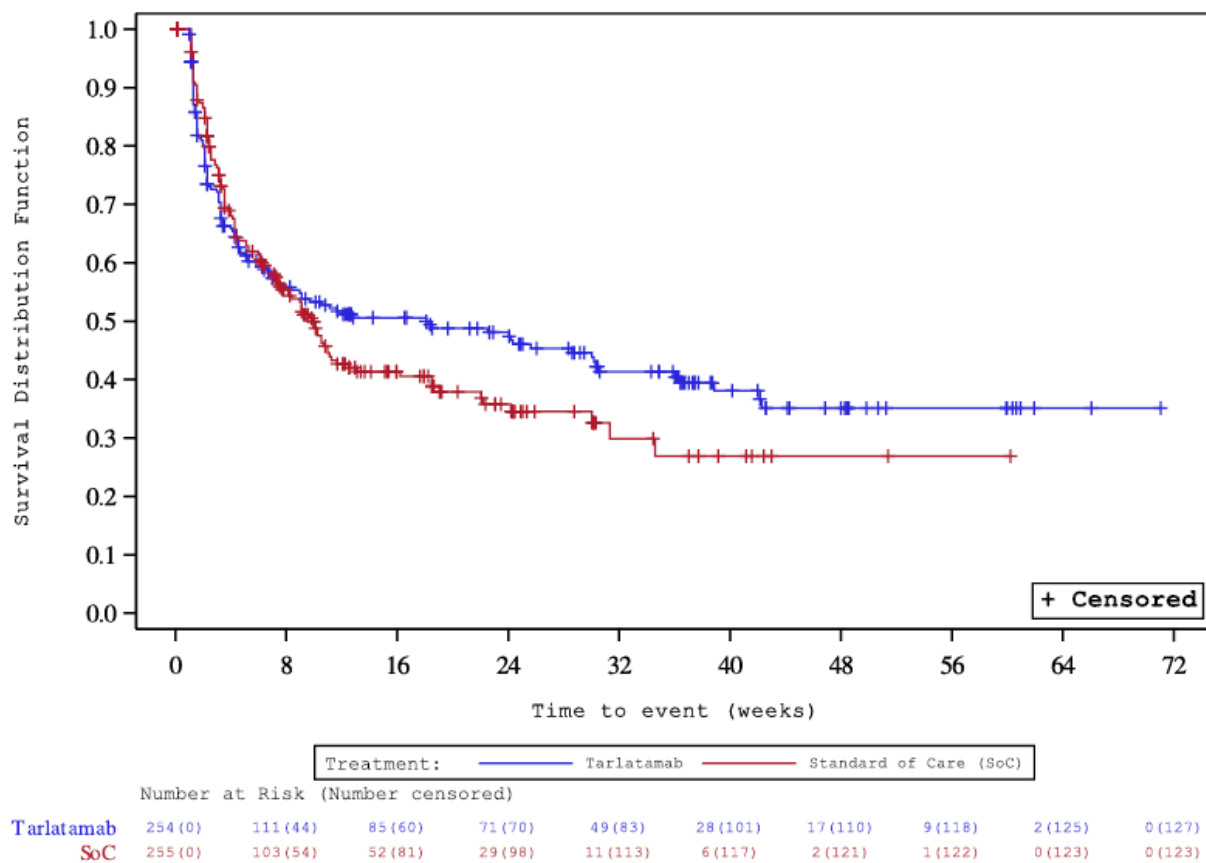


Abbildung 6: Kaplan-Meier-Kurve für die Symptomskala „Schlaflosigkeit“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

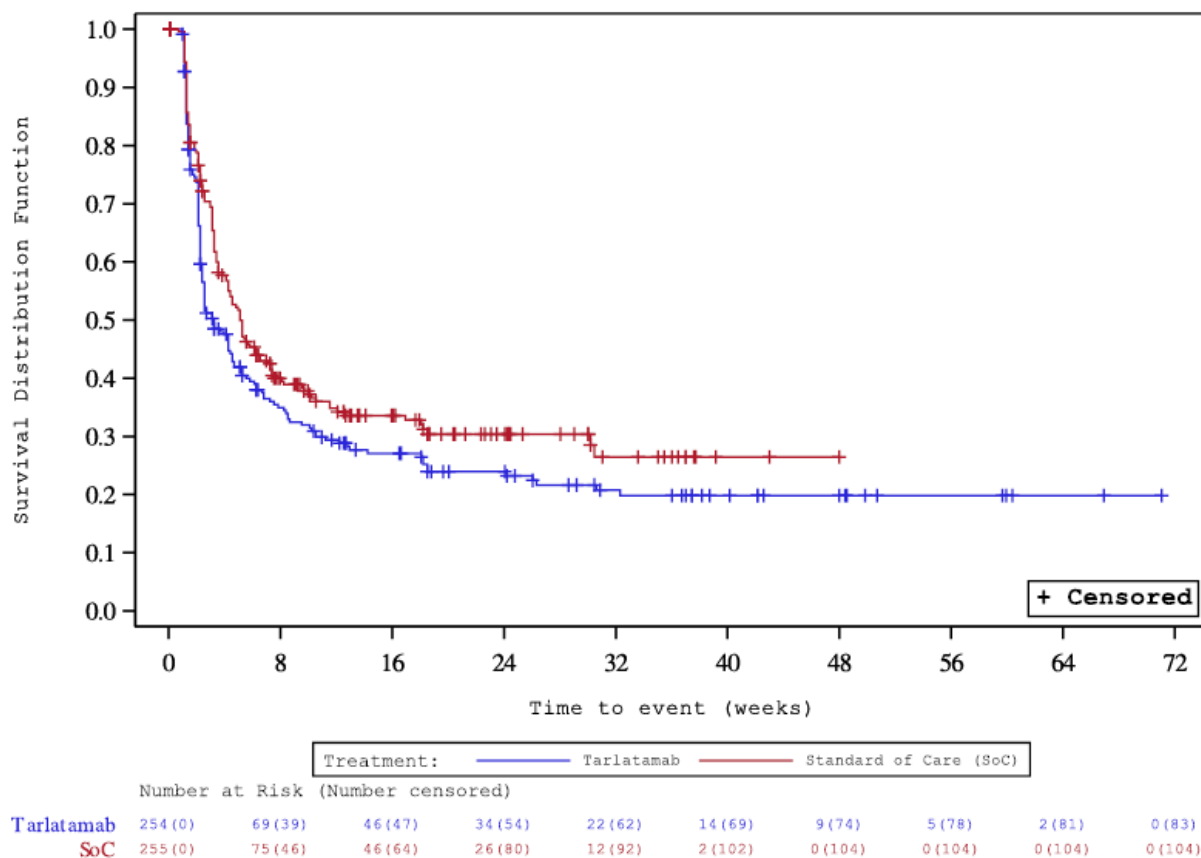


Abbildung 7: Kaplan-Meier-Kurve für die Symptomskala „Appetitverlust“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

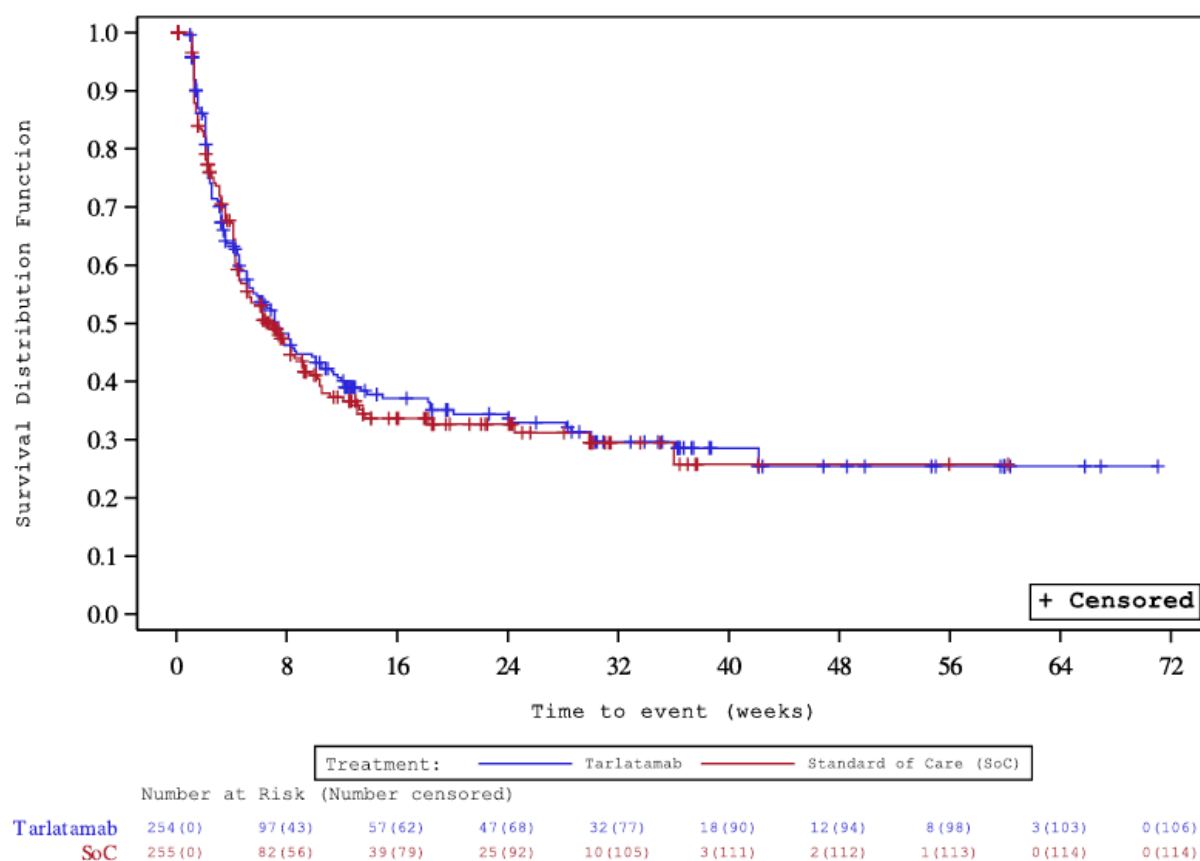


Abbildung 8: Kaplan-Meier-Kurve für die Symptomskala „Obstipation“ des EORTC QLQ-C30
 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
 Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

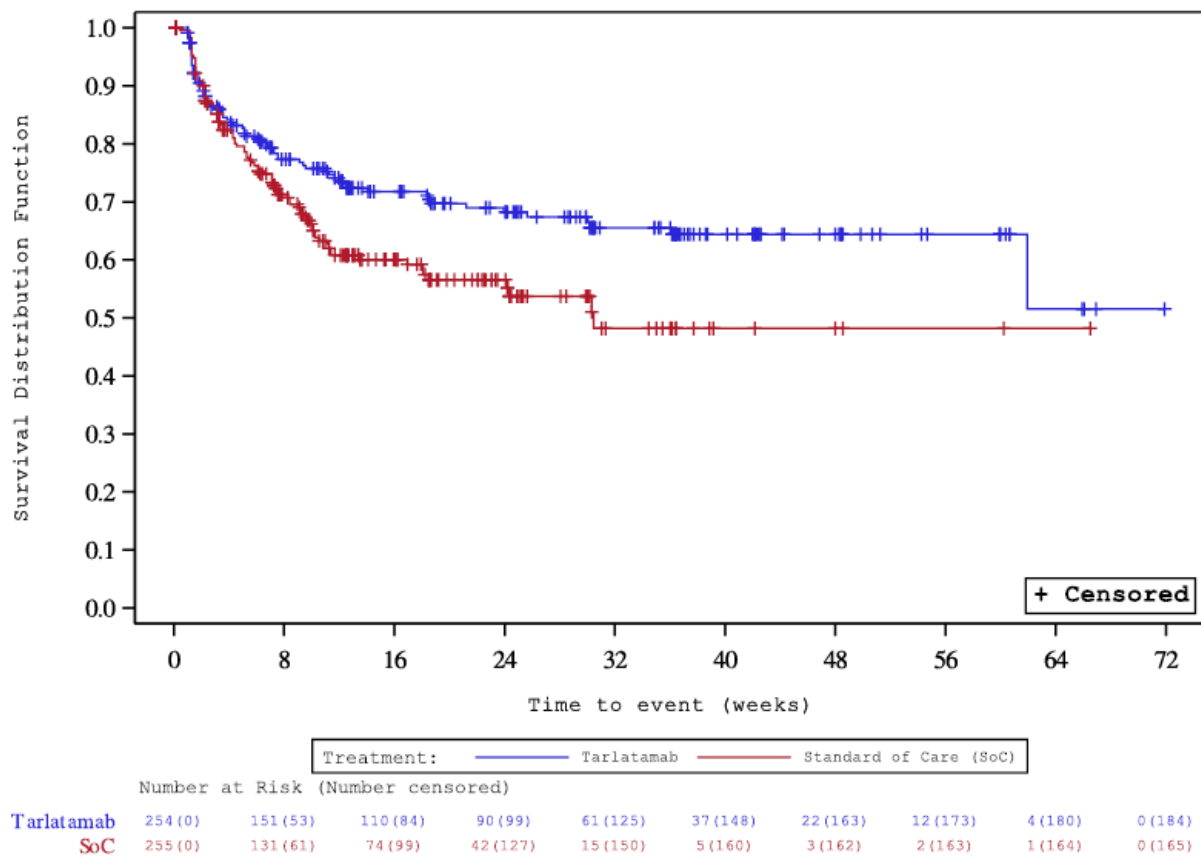


Abbildung 9: Kaplan-Meier-Kurve für die Symptomskala „Diarrhö“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304; ITT-Population (Datenschnitt: 29.01.2025)

EORTC QLQ-C30 – Funktionsskalen und globaler Gesundheitsstatus: Kaplan-Meier-Kurven

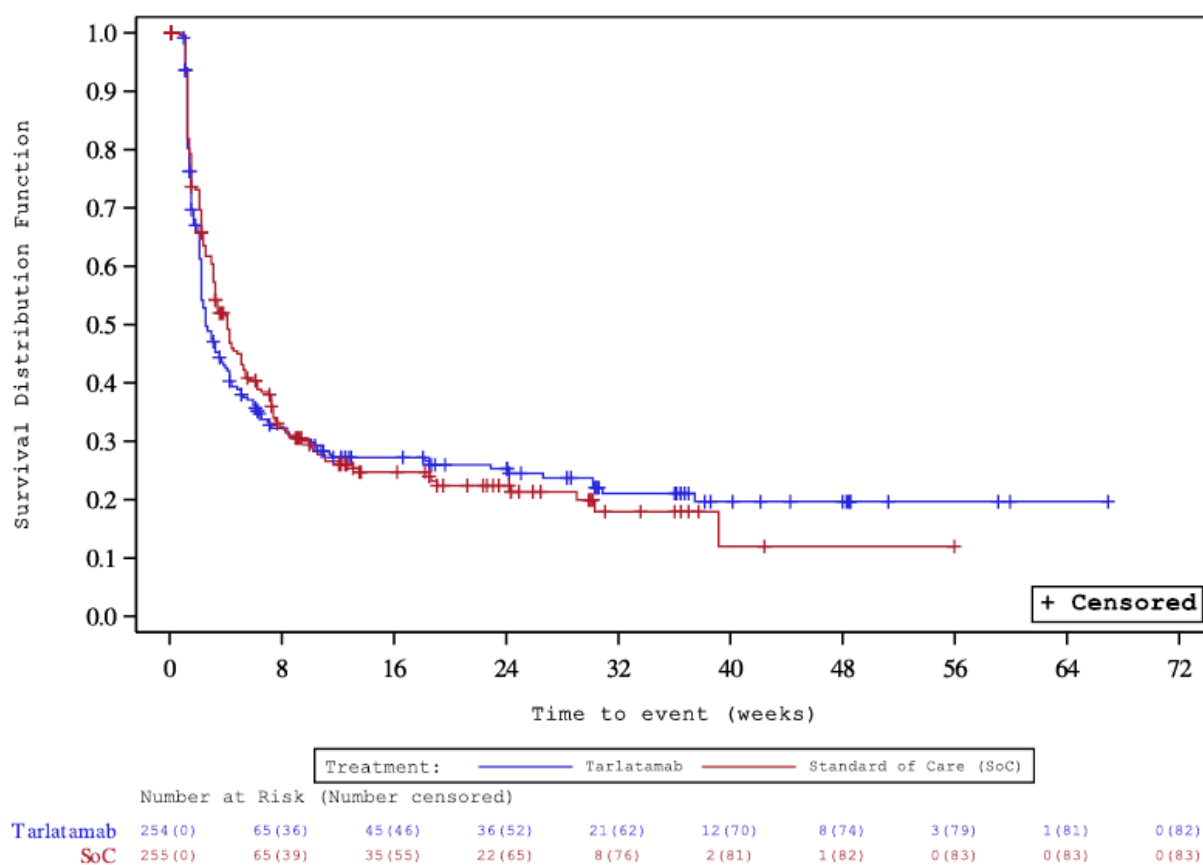


Abbildung 10: Kaplan-Meier-Kurve für die globale Skala „Allgemeiner Gesundheitsstatus / Lebensqualität“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

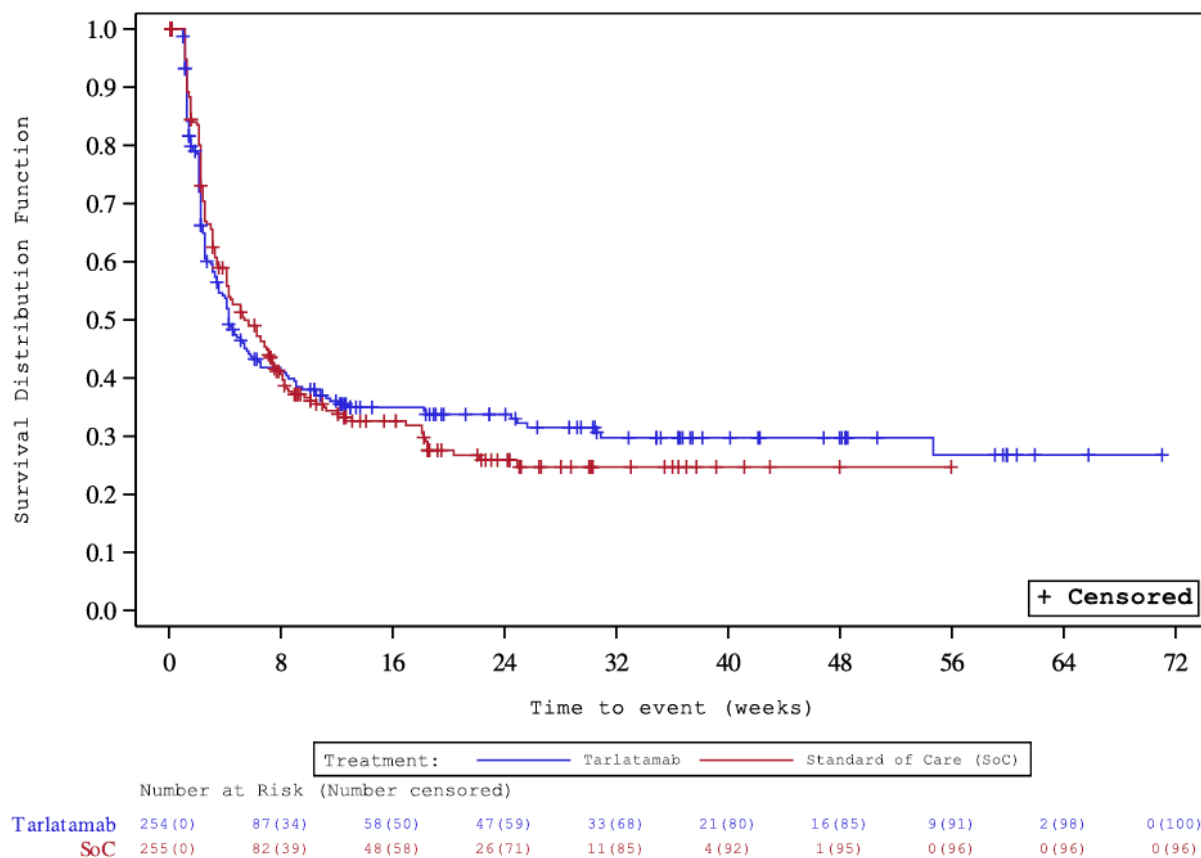


Abbildung 11: Kaplan-Meier-Kurve für die Skala „Körperliche Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

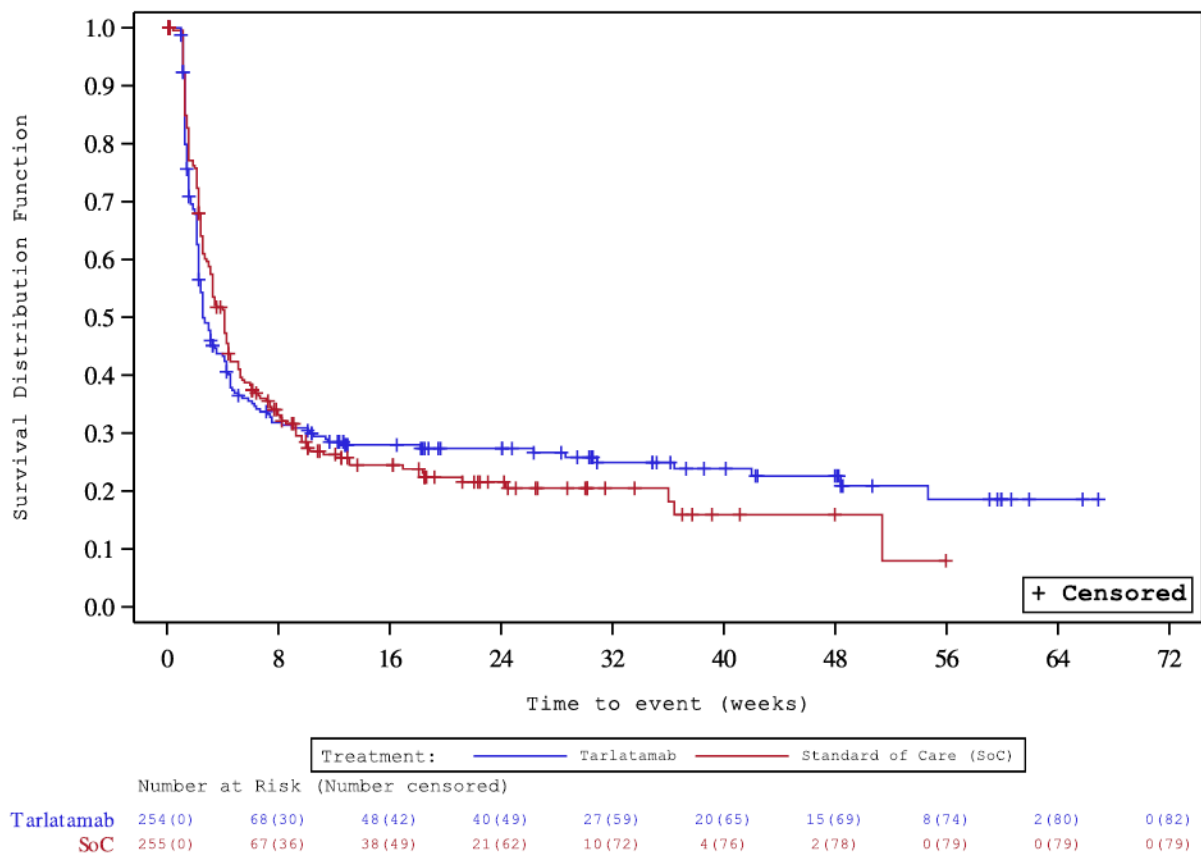


Abbildung 12: Kaplan-Meier-Kurve für die Skala „Rollenfunktion“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

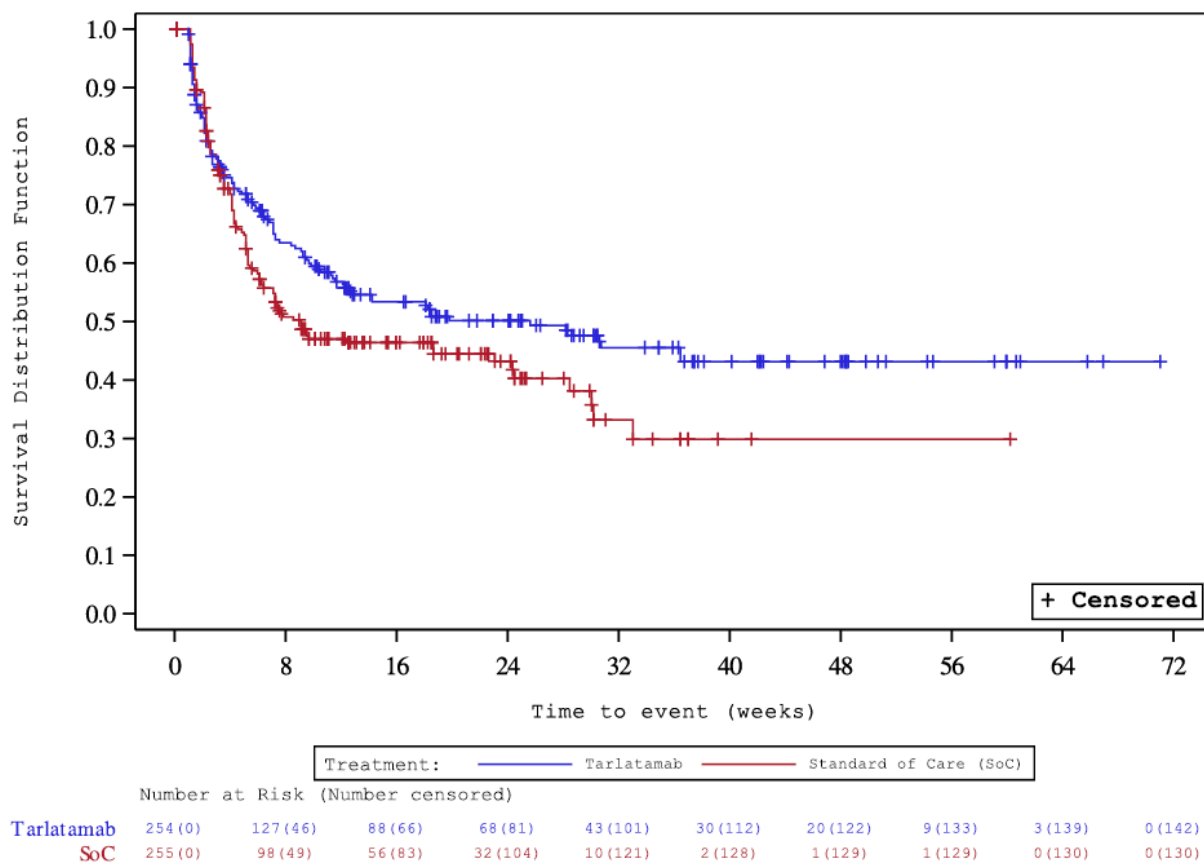


Abbildung 13: Kaplan-Meier-Kurve für die Skala „Emotionale Funktion“ des EORTC QLQ-C30
(Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte);
Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

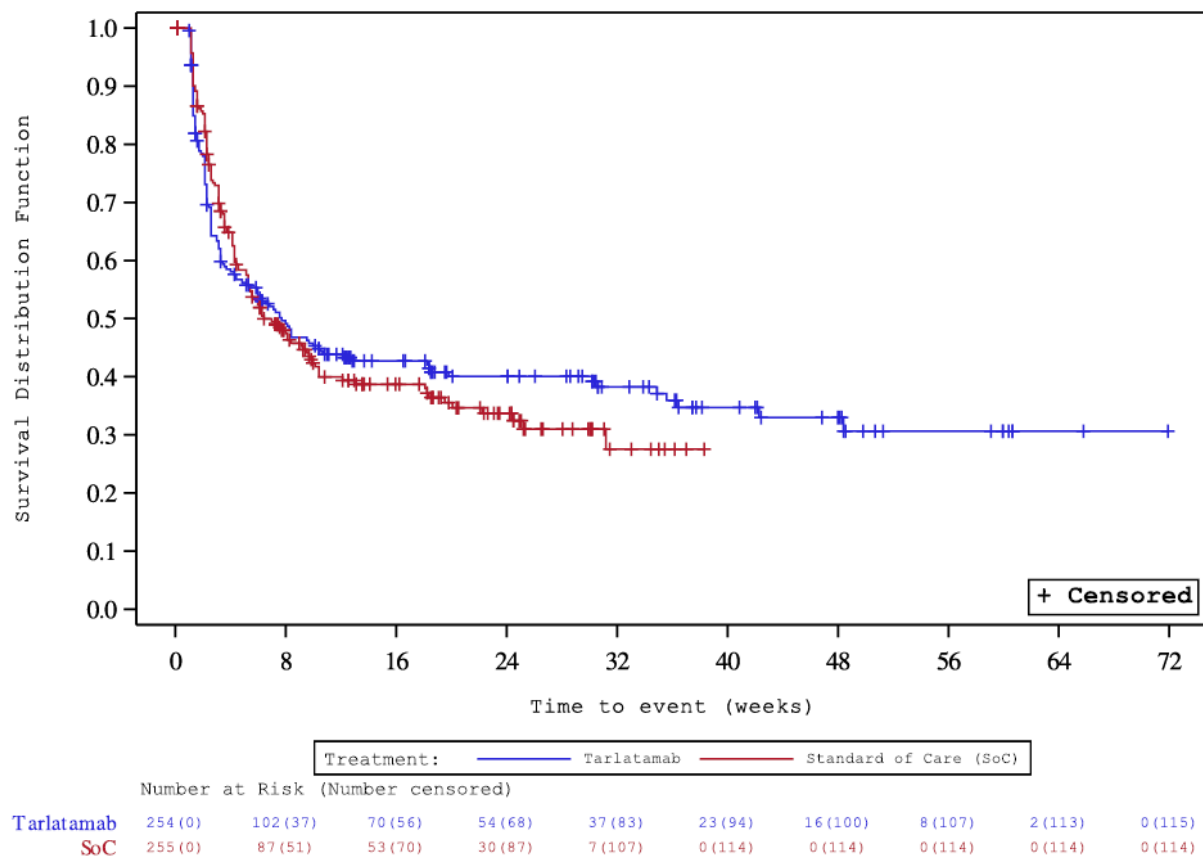


Abbildung 14: Kaplan-Meier-Kurve für die Skala „Kognitive Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

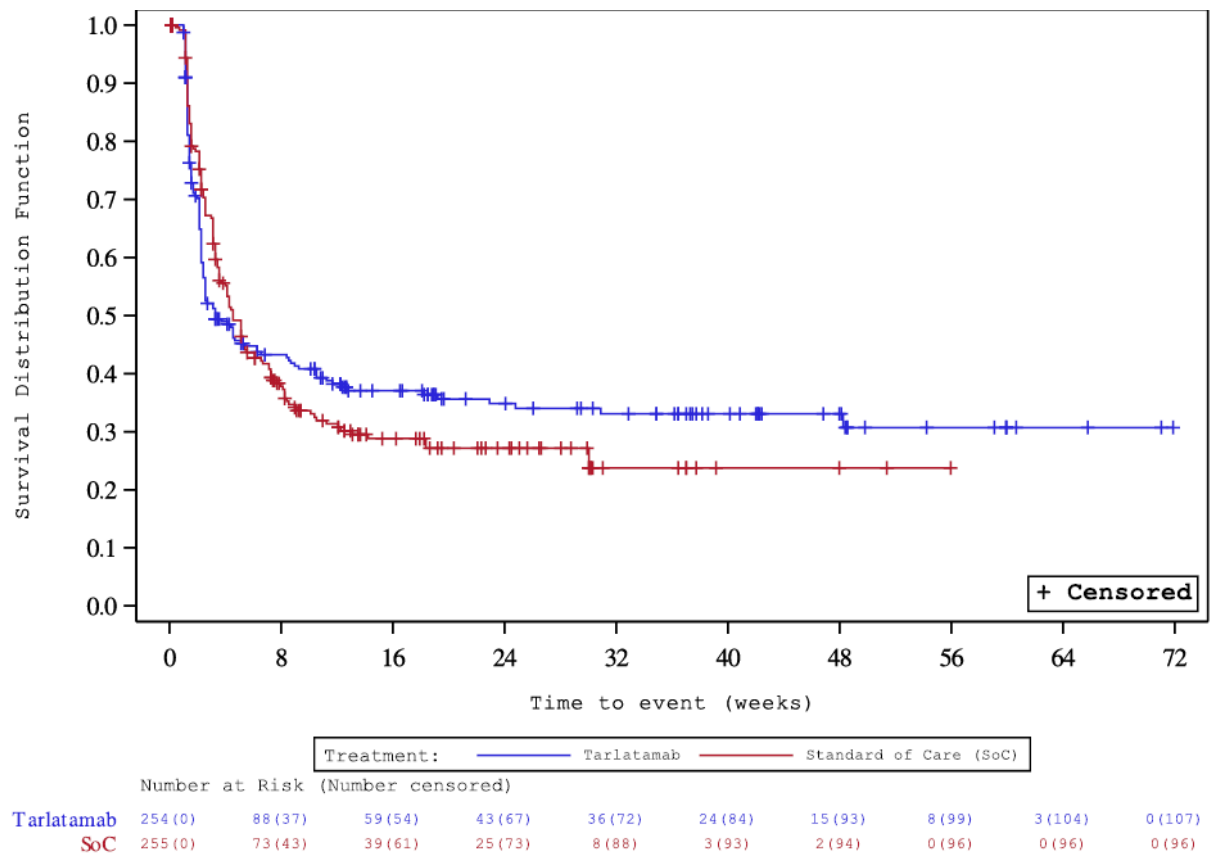


Abbildung 15: Kaplan-Meier-Kurve für die Skala „Soziale Funktion“ des EORTC QLQ-C30 (Zeit bis zur ersten Verschlechterung um ≥ 10 Punkte); Studie DeLLphi-304, ITT-Population (Datenschnitt: 29.01.2025)

Unerwünschte Ereignisse (Gesamtraten): Kaplan-Meier-Kurven

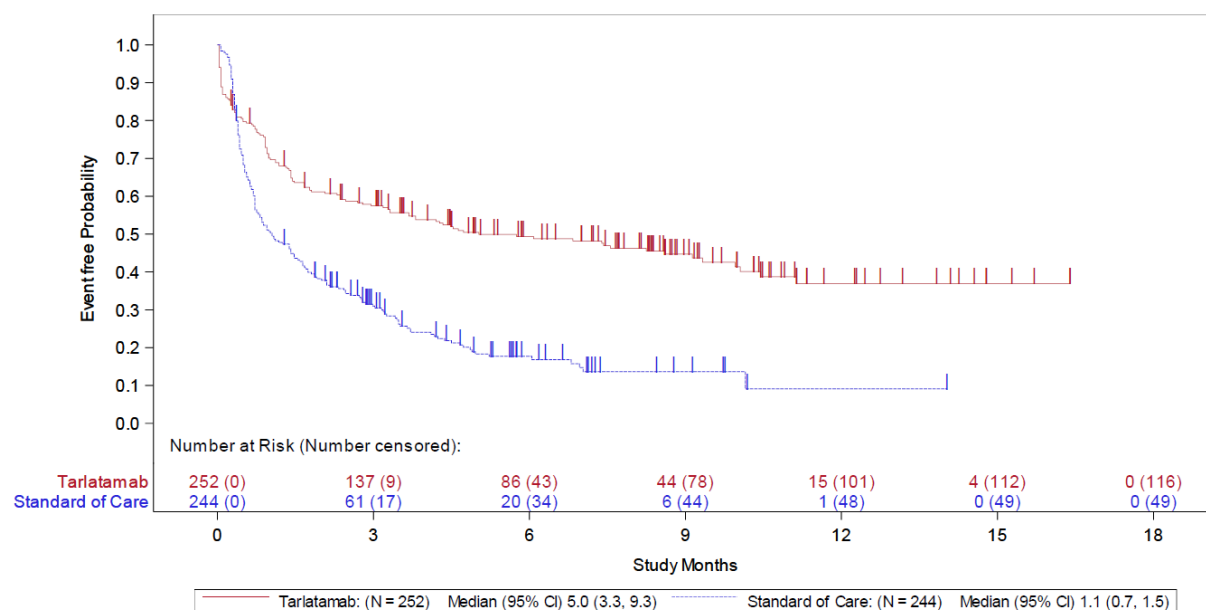


Abbildung 16: Kaplan-Meier-Kurve für „Schwere UE (CTCAE-Grad ≥ 3)“ (Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

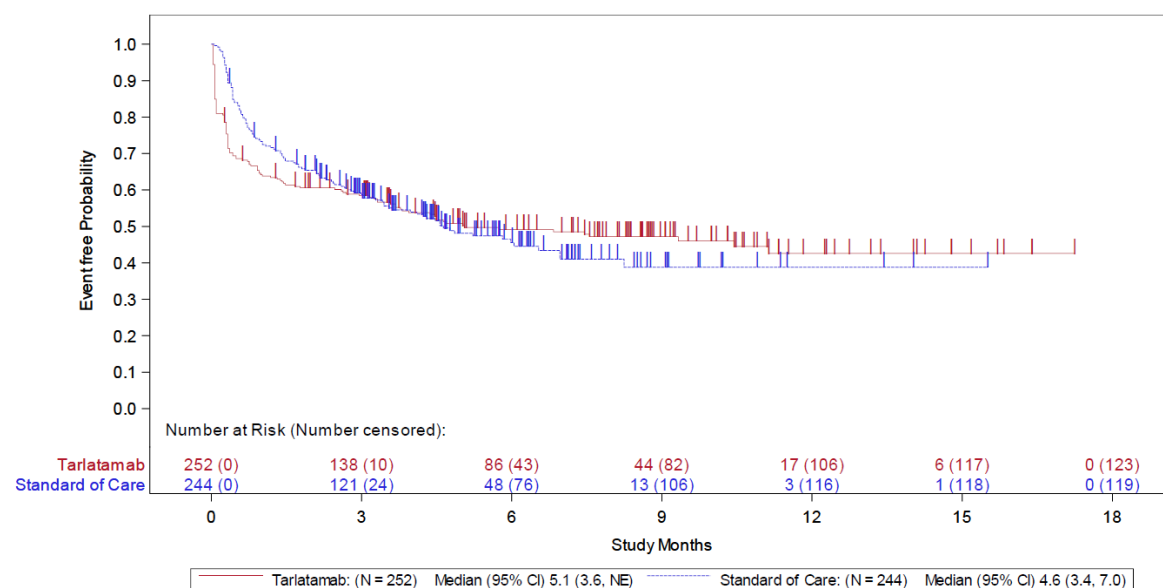


Abbildung 17: Kaplan-Meier-Kurve für „SUE“ (Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)

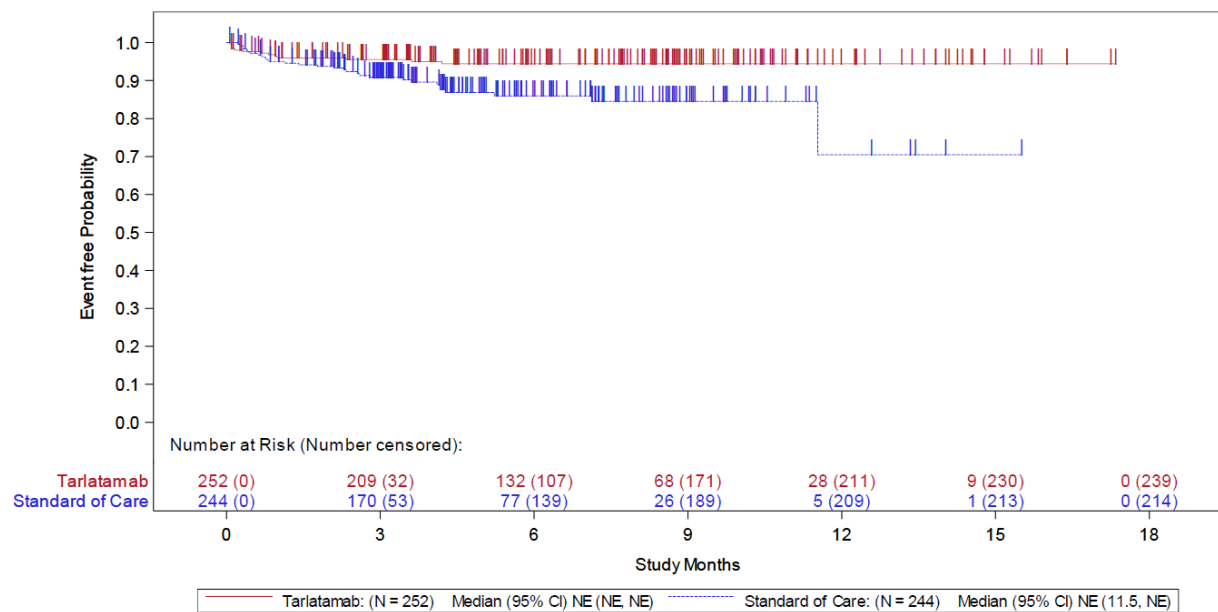


Abbildung 18: Kaplan-Meier-Kurve für „UE, das zum Abbruch der Studienmedikation führte“ (Zeit bis zum ersten Auftreten); Studie DeLLphi-304, Sicherheitspopulation (Datenschnitt: 29.01.2025)