

Nutzenbewertung



Gemeinsamer
Bundesausschuss

**von Arzneimitteln mit neuen Wirkstoffen nach
§ 35a SGB V**

**Bewertung von Arzneimitteln für seltene Leiden
nach § 35a Absatz 1 Satz 11 i.V.m.**

5. Kapitel § 12 Nr. 1 Satz 2 Verfo

Wirkstoff: Selumetinib

Datum der Veröffentlichung: 15. November 2021

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abkürzungsverzeichnis	5
Hintergrund.....	7
1 Fragestellung.....	8
2 Auswahl und Methodik der relevanten Studien.....	9
2.1 Studienbasis für die Nutzenbewertung	9
2.2 Beschreibung der eingeschlossenen Studien.....	13
2.3 Endpunkte.....	23
2.3.1 Mortalität.....	24
2.3.2 Morbidität.....	25
2.3.3 Lebensqualität.....	51
2.3.4 Sicherheit	54
2.3.5 Übersicht der Erhebungszeitpunkte.....	55
2.4 Statistische Methoden	56
2.5 Verzerrungspotential auf Studien- und Endpunktebene	57
3 Ergebnisse der eingeschlossenen Studien.....	58
3.1 Studiencharakteristika und Studienmedikation	58
3.2 Mortalität.....	63
3.3 Morbidität.....	63
3.4 Lebensqualität	77
3.5 Sicherheit.....	79
4 Diskussion der Methodik und Ergebnisse.....	87
4.1 Zulassungsstatus und Zulassungspopulation von Selumetinib	87
4.2 Design und Methodik der Studie	87
4.3 Mortalität	90
4.4 Morbidität	90
4.5 Lebensqualität	93
4.6 Sicherheit.....	94
5 Anforderung an eine qualitätsgesicherte Anwendung	95
6 Zusammenfassung der Nutzenbewertung	96
Referenzen	101
Anhang	106

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis	9
Tabelle 2:	Charakterisierung der pivotalen Studie SPRINT	13
Tabelle 3:	Für die Nutzenbewertung relevante Protokolländerungen der Phase I der Studie SPRINT	18
Tabelle 4:	Für die Nutzenbewertung relevante Protokolländerungen der Phase II der Studie SPRINT	20
Tabelle 5:	Charakterisierung der Intervention Selumetinib der Studie SPRINT.....	21
Tabelle 6:	Zusammenfassung der Endpunktbewertung.....	23
Tabelle 7:	Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Phase I der Studie SPRINT	55
Tabelle 8:	Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Phase II der Studie SPRINT	55
Tabelle 9:	Allgemeine Angaben zum Studienverlauf der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitte: 29.06.2018 und 29.03.2019)	58
Tabelle 10:	Charakterisierung der Studienpopulation der Studie SPRINT; FAS Phase I und Phase II, Stratum 1 (Datenschnitt: 29.06.2018)	59
Tabelle 11:	Angaben zur Exposition mit der Studienmedikation Selumetinib; FAS der Studie SPRINT Phase I und Phase II, Stratum 1 (Datenschnitte 29.06.2018 und 29.03.2019).....	61
Tabelle 12:	Begleitmedikation mit einer Häufigkeit > 20 %; FAS der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt: 29.06.2018).....	62
Tabelle 13:	Responderanalyse NRS-11 für den schlimmsten Schmerz; FAS der Studie SPRINT, Phase II, Stratum 1 im Alter von 8 bis 18 Jahren (Datenschnitt 29.06.2018)	64
Tabelle 14:	Responderanalysen zur Schmerzlinderung für „Schmerz des Ziel-PN“; FAS der Studie SPRINT, Phase II, Stratum 1 im Alter von 8 bis 18 Jahren (Datenschnitt 29.06.2018)	65
Tabelle 15:	Veränderungen im Grooved Pegboard Test; FAS ab 5 Jahren mit einem PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)	66
Tabelle 16:	Veränderungen der PROMIS-Skalen „Mobilität“ und „Obere Extremitäten“; FAS im Alter von 8 bis 18 Jahren mit Motor-PN-bezogener Morbidität der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)	67
Tabelle 17:	Responderanalysen für Verbesserungen und Verschlechterungen der Sehschärfe; FAS mit PN der Augenhöhle der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	68
Tabelle 18:	Responderanalysen der Symptomatik; FAS der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	69
Tabelle 19:	Responderanalyse zum GIC; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)	76

Tabelle 20:	Responderanalyse für den PedsQL; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	78
Tabelle 21:	Veränderungen des PedsQL-Gesamtwerts; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	79
Tabelle 22:	Unerwünschte Ereignisse; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	79
Tabelle 23:	UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	80
Tabelle 24:	UE CTCAE-Grad 3 oder höher mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	84
Tabelle 25:	SUE mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	85
Tabelle 26:	UE von besonderem Interesse und zugeordnete Preferred Terms mit Inzidenz ≥ 5 %; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	85
Tabelle 27:	Zusammenfassende Darstellung der Ergebnisse; FAS der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018).....	96
Tabelle 28:	Veränderungen der 6-Minuten-Gehstrecke; FAS ab 5 Jahren mit PN der unteren Extremitäten, Kompression des Rückenmarks und PN der oberen Atemwege der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)..	106

Abkürzungsverzeichnis

6MWT	6-Minuten-Gehtest (Six-Minute Walk Test)
AHI	Apnoe-Hypopnoe-Index
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ATS	Amblyopia Treatment Study
BOR	Bestes Objektives Ansprechen (Best Objective Response)
CTCAE	Common Terminology Criteria for Adverse Events
DIF	Differential Item Functioning
DVQ	Dysfunctional Voiding Questionnaire
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FAS	Full Analysis Set
FDA	U. S. Food and Drug Administration
FEV	Einsekundenkapazität (Forced Expiratory Volume)
FI	Fachinformation
G-BA	Gemeinsamer Bundesausschuss
GIC	Global Impression of Change
GKV	Gesetzliche Krankenversicherung
ICC	Intraklassen-Korrelationskoeffizient (Intraclass Correlation Coefficient)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Item-Response-Theorie
ITT	Intention to Treat
KI	Konfidenzintervall
logMAR	Logarithm of the Minimum Angle of Resolution
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference
MMRM	Mixed Effect Model Repeat Measures
MMT	Manueller Muskeltest
MPNST	Maligner peripherer Nervenscheidentumor (Malignant Peripheral Nerve Sheath Tumor)
MRC	Medical Research Council
MRT	Magnetresonanztomographie
NCI	National Cancer Institute
NCI POB	National Cancer Institute – Pediatric Oncology Branch
NF1	Neurofibromatose Typ 1
NRS-11	Numerical Rating Scale (Schmerzskala)

ORR	Objektive Ansprechrate (Objective Response Rate)
PD	Progressive Erkrankung (Progressive Disease)
PedsQL	Pediatric Quality of Life Inventory
PFS	Progressionsfreies Überleben (Progression-Free Survival)
PII	Pain Interference Index
PII-P	Pain Interference Index – Parent-Report
PN	Plexiformes Neurofibrom
PR	Partielles Ansprechen (Partial Response)
PRO	Patient-Reported Outcomes
PROMIS	Patient-Reported Outcomes Measurement Information System
pU	pharmazeutischer Unternehmer
R ₂₀	Atemwegswiderstand bei 20 Hertz
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RECIST	Response Evaluation Criteria in Solid Tumors
REiNS	Response Evaluation in Neurofibromatosis and Schwannomatosis
SAP	Statistischer Analyseplan
SGB	Sozialgesetzbuch
SUE	Schwerwiegendes unerwünschtes Ereignis
TTP	Zeit bis zur Progression (Time to Progression)
TTR	Zeit bis zum Ansprechen (Time to Response)
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA

Hintergrund

Selumetinib ist zugelassen als Arzneimittel zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

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

Dementsprechend hat der G-BA in seiner Sitzung am 15. März 2012 den mit Beschluss vom 1. August 2011 erteilten Auftrag an das IQWiG zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen gemäß § 35a Absatz 2 SGB V in der Weise abgeändert, dass bei Orphan Drugs eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erfolgt, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt (vgl. § 35a Absatz 1 Satz 12 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Selumetinib zieht der G-BA 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 im 5. Kapitel § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Selumetinib in seiner Sitzung am 9. November 2021 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 18. August 2021 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 15. November 2021 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Selumetinib (Koselugo®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. In die Bewertung werden gemäß Zulassungsstatus Personen in folgender Indikation eingeschlossen [5]:

Kinder ab 3 Jahren und Jugendliche zur Behandlung von symptomatischen, inoperablen plexiformen Neurofibromen (PN) bei Neurofibromatose Typ 1 (NF1)

Gemäß Fachinformation (FI) [5] beträgt die zugelassene Dosierung 25 mg/m² Körperoberfläche, zweimal täglich oral eingenommen (alle 12 Stunden). Die Behandlung mit Selumetinib sollte so lange fortgesetzt werden, wie ein klinischer Nutzen beobachtet wird bzw. bis zum Auftreten einer Progression der plexiformen Neurofibrome oder einer inakzeptablen Toxizität [5].

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Table 1: Übersicht über die Studienbasis

Studienname / -nummer	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Studien zum Wirkstoff				
SPRINT ¹⁾	Ja	Ja	Ja	
Baldo et al., 2020 [10]	Nein	Ja	Nein	Studie liegt nur in Form einer Publikation vor. Detaillierte Angaben zu Ein- und Ausschlusskriterien, Patientenfluss, Begleittherapien, Endpunkterhebung und -operationalisierung sowie eine statistische Auswertung der Ergebnisse fehlen. Studienprotokoll, SAP und Eintrag in ein Studienregister wurden nicht identifiziert, sodass nicht von einem präspezifizierten Vorgehen ausgegangen werden kann. Es wurden nur wenige Personen in die Studie eingeschlossen (n = 9).
Espírito Santo et al., 2020 [17]	Nein	Ja	Nein	Studie liegt nur in Form einer Publikation vor. Detaillierte Angaben zu Ein- und Ausschlusskriterien, Patientenfluss, Begleittherapien, Endpunkterhebung und -operationalisierung sowie eine statistische Auswertung der Ergebnisse fehlen. Studienprotokoll, SAP und Eintrag in ein Studienregister wurden nicht identifiziert, sodass nicht von einem präspezifizierten Vorgehen ausgegangen werden kann.
D1346C00013 ²⁾	Ja	Nein	Nein	Es liegen noch keine Ergebnisse vor.
D1346C00011 ³⁾	Ja	Nein	Nein	Es liegen noch keine Ergebnisse vor.
Studien zu externen Kontrollen				
NCI-08-C-0079	Teilweise ⁴⁾	Ja	Nein	<ul style="list-style-type: none"> Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht nachgewiesen (Unterschiede in Ein- und Ausschlusskriterien und einigen Baseline-Charakteristika wie Alter und Tumolvolumen zu Baseline). Endpunkte für indirekten Vergleich nicht patientenrelevant. Mangelnde Informationsgrundlage, z. B. keine Informationen zur Operationalisierung einer Progression, detaillierte Informationen zu Patientenfluss und Baseline-Charakteristika fehlen.

Studienname / -nummer	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Placebo-Arm der Studie 01-C-0222 [54]	Nein ⁵⁾	Ja	Nein	<ul style="list-style-type: none"> • Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht nachgewiesen (Gemäß Einschlusskriterien wiesen alle Teilnehmenden eine progressive Erkrankung auf. In der SPRINT-Studie lag der Anteil mit progressiver Erkrankung bei 42 %). • Endpunkte für indirekten Vergleich nicht patientenrelevant. • Mangelnde Informationsgrundlage, z. B. detaillierte Informationen zu Baseline-Charakteristika fehlen.
Fisher et al., 2008 [19]	Nein	Ja	Nein	<ul style="list-style-type: none"> • Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht nachgewiesen (keine Angaben zu Baseline-Charakteristika außer Alter und Geschlecht, Unterschiede in Studiendesign, Einschlusskriterien, Datenerhebung). • Datengrundlage unzureichend beschrieben⁶⁾. • Endpunkte für indirekten Vergleich nicht patientenrelevant.
Nguyen et al., 2012 [37]	Nein	Ja	Nein	<ul style="list-style-type: none"> • Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht gegeben (Einschluss von Kindern und Erwachsenen ohne separate Analysen, weitere Unterschiede in Einschlusskriterien, keine Angabe zu Baseline-Charakteristika, Unterschiede in Studiendesign und Datenerhebung). • Datengrundlage unzureichend beschrieben⁶⁾. • Endpunkte für indirekten Vergleich nicht patientenrelevant.
Nguyen et al., 2013 [38]	Nein	Ja	Nein	<ul style="list-style-type: none"> • Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht nachgewiesen (Einschluss postoperativ, weitere Unterschiede in Ein- und Ausschlusskriterien, Unterschiede in Baseline-Charakteristika z. B. hinsichtlich Tumolvolumen zu Baseline, Unterschiede in Studiendesign, Datenerhebung). • Datengrundlage unzureichend beschrieben⁶⁾. • Endpunkte für indirekten Vergleich nicht patientenrelevant.

Studiename / -nummer	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
Well et al., 2021 [52]	Nein	Ja	Nein	<ul style="list-style-type: none"> • Hinreichende Vergleichbarkeit mit Population der SPRINT-Studie nicht nachgewiesen (Unterschiede in Ein- und Ausschlusskriterien, keine Angabe zu Baseline-Charakteristika außer Alter und Angabe zu Genmutation, Unterschiede in Studiendesign und Datenerhebung). • Datengrundlage unzureichend beschrieben⁶⁾. • Endpunkte für indirekten Vergleich nicht patientenrelevant.

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ Bei der Studie D1346C00013 handelt es sich um eine laufende Phase-I-Studie bei japanischen Patientinnen und Patienten.

³⁾ Bei der Studie D1346C00011 handelt es sich um eine laufende Phase-I-Studie bei chinesischen Patientinnen und Patienten.

⁴⁾ Es wurde ein Studienprotokoll und ein Bericht über den indirekten Vergleich mit Studie NCI-08-C-0079 eingereicht. Zudem liegen Informationen zum indirekten Vergleich im Studienbericht der SPRINT-Studie vor. Ein Studienbericht oder ausführlicher Ergebnisbericht der Studie NCI-08-C-0079 wurde nicht vorgelegt.

⁵⁾ Indirekter Vergleich mit Placebo-Arm der Studie 01-C-0222 ist Teil des Studienberichts der SPRINT-Studie. Studienprotokoll, SAP und Studienbericht zur Studie 01-C-002 wurden nicht eingereicht.

⁶⁾ Grundlage bildet eine Publikation. Studienprotokoll, SAP sowie eine detaillierte Ergebnisdarstellung wurden nicht eingereicht. Dadurch fehlen bspw. detaillierte Angaben zur Studienpopulation (z. B. Einschlusskriterien, Patientenfluss, Baseline-Charakteristika vor), Studiendurchführung (Erhebungsmethodik, Erhebungszeitpunkte) und zur Operationalisierung der Endpunkte.

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

Aus dem vom pU vorgelegten Studienpool kann zur Nutzenbewertung lediglich die zulassungsbegründende Studie SPRINT herangezogen werden (Tabelle 1). Die vom pU vorgelegten weiteren Daten aus dem „Compassionate Use Program“ basieren als Datenquelle lediglich auf Publikationen [10,17]. Studienprotokoll, statistischer Analyseplan (SAP) sowie eine detaillierte Ergebnisdarstellung wurden nicht eingereicht. Dadurch fehlen bspw. detaillierte Angaben zu Studienpopulation (z. B. Einschlusskriterien, Patientenfluss, Baseline-Charakteristika), Studiendurchführung (Erhebungsmethodik, Erhebungszeitpunkte), Datengrundlage (Vollständigkeit, Vollzähligkeit), Operationalisierung der Endpunkte und angewendete statistische Methoden. Für die Studien existieren ferner keine Studienregistereinträge. Diese Daten stellen damit keine geeignete Grundlage für die Nutzenbewertung dar.

Daneben legt der pU einen indirekten Vergleich mit externen Kontrollstudien [19,36,37,38,52,54] vor. Für 4 Studien [19,37,38,52] basiert der Vergleich auf publizierten Studien, welche aber zum Teil große Unterschiede hinsichtlich der eingeschlossenen Studienpopulation aufweisen und für die mangels ausführlicher Informationen zu Baseline-Charakteristika keine hinreichende Vergleichbarkeit mit der Studienpopulation der Studie SPRINT nachgewiesen werden kann. Für keine der 4 Studien wurden ein Studienprotokoll, ein SAP sowie eine detaillierte Ergebnisdarstellung eingereicht, sodass relevante Angaben zum Nachweis der Vergleichbarkeit der Studienpopulationen und Endpunkte fehlen. Daneben führt der pU einen indirekten Vergleich mit einer prospektiven Beobachtungsstudie (NCI-08-C-0079) zum natürlichen Verlauf der NF1-Erkrankung und mit dem Placebo-Arm der RCT 01-C-0222 durch. Die Studie zum natürlichen Verlauf dient als Rahmenprotokoll für das klinische Studienprogramm des National Cancer Institute – Pediatric Oncology Branch

(NCI POB), welches auch die Studie SPRINT koordiniert hat. Die eingeschlossene Studienpopulation wird in einer der Publikationen zur Studie als wenig repräsentativ für die allgemeine NF1-Population beschrieben, sondern umfasst eher Patientinnen und Patienten, die aufgrund der Schwere ihrer Erkrankung um eine Behandlung ihrer PN ersuchen [2]. Für diese Studie wurde ein Studienprotokoll sowie ein Bericht über eine mittels Propensity Score adjustierte Analyse für den indirekten Vergleich vorgelegt. Die Analyse wird auch im Studienbericht der Phase II der SPRINT-Studie beschrieben. Dennoch fehlen Angaben bspw. zur Operationalisierung der erhobenen Endpunkte. Zudem gibt es Unterschiede in den Ein- und Ausschlusskriterien (z. B. keine Notwendigkeit einer PN-bezogenen Morbidität; keine Kriterien bezüglich Leber-, Lungen-, Nieren- und Herzfunktion und Infektionen; kein Ausschluss von Personen mit Gliom oder malignem peripheren Nervenscheidentumor (MPNST), welche eine Behandlung mit Chemo- oder Strahlentherapie erfordern), sodass die Patientenpopulation insgesamt breiter ist als die der SPRINT-Studie. Auch die Studienpopulation des Placebo-Arms der Studie 01-C-0222 unterscheidet sich wesentlich von der der SPRINT-Studie, da in die Studie 01-C-0222 bei ansonsten sehr ähnlichen Einschlusskriterien nur Personen mit einer progressiven Erkrankung ($\geq 20\%$ Zunahme des Volumens der PN, oder $\geq 13\%$ Zunahme in zweidimensionalen Messungen, oder $\geq 6\%$ Zunahme in eindimensionaler Messung über die 2 letzten volumetrischen Magnetresonanztomographie (MRT)-Scans oder innerhalb eines Jahres vor Einschluss in die Studie) eingeschlossen wurden. In der Studie SPRINT lag der Anteil an Personen mit progressiver Erkrankung ($\geq 20\%$ Zunahme des PN-Volumens innerhalb von 12 bis 15 Monaten vor Einschluss) bei 42 %. Weitere Unterschiede zwischen den Studienpopulationen zu Baseline zeigen sich bezüglich demographischer Faktoren wie Alter (Teilnehmende in der Studie SPRINT sind ca. 2 Jahre älter als in Kontrollstudien) und bezüglich Krankheitscharakteristika (Tumorvolumen in der Studie SPRINT ca. 1,5 x größer als in externen Kontrollstudien). Da die für den indirekten Vergleich herangezogenen Endpunkte „Tumorvolumen“ und „Progressionsfreies Überleben“ zu dem als nicht patientenrelevant erachtet werden (siehe Abschnitt 2.3.2), werden der indirekte Vergleich und die dafür eingereichten externen Kontrollstudien nicht für die Nutzenbewertung herangezogen.

Zur Nutzenbewertung für Selumetinib herangezogene Studien und Daten

- Herstellerdossier zu Selumetinib [4]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [36]
- Studienbericht inkl. Studienprotokoll und SAP der Studie SPRINT (NCT01362803) [6,7] sowie 90 days Update Safety Report [8]

2.2 Beschreibung der eingeschlossenen Studien

Die Unterlagen zur Nutzenbewertung für Selumetinib basieren auf der Zulassungsstudie SPRINT (NCT01362803). Die Studie und die Intervention werden in den Tabellen 2–5 charakterisiert.

Tabelle 2: Charakterisierung der pivotalen Studie SPRINT

Charakteristikum	Beschreibung
Design und Studienablauf	<p>Die Studie SPRINT beinhaltet eine Phase I und eine Phase II.</p> <p>Phase I</p> <p>Ziel der Phase I war die Bestimmung der maximal tolerierten Dosis von Selumetinib bei Kindern und Jugendlichen mit NF1 und inoperablen PN. Die Phase I bestand aus einer offenen, einarmigen, dosisescalierenden, multizentrischen Studie mit Selumetinib bei pädiatrischen Patientinnen und Patienten mit NF1 und inoperablen PN.</p> <p>Phase II</p> <p>Ziel der Phase II war die Untersuchung von Sicherheit und Wirksamkeit von Selumetinib bei Kindern und Jugendlichen mit NF1 und inoperablen PN. Auch dieser Studienphase lag ein offenes, einarmiges, multizentrisches Studiendesign zugrunde. Für die Phase II wurden die Patientinnen und Patienten in eines von zwei Strata aufgenommen, basierend darauf, ob zum Zeitpunkt ihres Studieneinschlusses bereits eine PN-bezogene Morbidität vorlag (Stratum 1) oder keine signifikante klinische Morbidität jedoch das Potential für eine solche bestand (Stratum 2). Als PN-bezogene Morbidität galten durch PN hervorgerufene Schmerzen, Entstellung oder Funktionseinschränkungen (Sehverlust, Störung der Gesichtsmotorik, Hörverlust, Probleme beim Schlucken, Sprachbehinderung, Blockierung der Atemwege, Beeinträchtigung der Atmung, Störung der Blasenfunktion, Störung der Darmfunktion, Muskelschwäche, Einschränkung des Bewegungsumfanges, sensorische Störungen, andere).</p> <p>In beiden Phasen erfolgte die Behandlung in Zyklen von 28 Tagen ohne Einnahmepausen zwischen den Zyklen. Teilnehmende mit Krankheitsprogression innerhalb von 1,5 Jahren vor Studieneintritt erhielten das Studienmedikament so lange sie keine schweren Nebenwirkungen hatten und sich die Krankheit nicht verschlimmerte. Bei Teilnehmenden ohne Krankheitsprogression in den 1,5 Jahren vor Studieneintritt und ohne Ansprechen auf die Behandlung mit Selumetinib (Reduktion des Tumolvolumen um $\geq 20\%$) sollte die Behandlung nach 2 Jahren beendet werden. Die Patientinnen und Patienten sollten für insgesamt 7 Jahre nach Beginn der Behandlung mit Selumetinib oder 5 Jahre nach Beendigung der Behandlung mit Selumetinib in der Studie SPRINT nachverfolgt werden, je nachdem was länger ist.</p>
Population	<p>Einschlusskriterien</p> <ul style="list-style-type: none"> • Alter in Phase I: Zwischen ≥ 3 und ≤ 18 Jahren. • Alter in Phase II: zwischen ≥ 2 und ≤ 18 Jahren + Körperoberfläche von $\geq 0,55$ m². • Fähigkeit ganze Kapseln zu schlucken. • Diagnostizierte NF1 mit mindestens einem inoperablem PN (PN kann nicht vollständig chirurgisch entfernt werden, ohne das Risiko einer erheblichen Morbidität aufgrund der Umhüllung oder Lokalisation in der Nähe lebenswichtiger Strukturen, einer Invasivität oder einer hohen Vaskularität). • PN verursacht eine signifikante Morbidität (Stratum 1) oder hat das Potential eine signifikante Morbidität zu verursachen (Stratum 2), bspw. (jedoch nicht ausschließlich) in Form von folgenden Veränderungen: <ul style="list-style-type: none"> ○ Kopf- und Halsläsionen, die eine Kompression der Atemwege oder großer Gefäße verursachen (können); ○ Paraspinal-Läsionen, die eine Myelopathie verursachen (können); ○ Läsionen des Plexus brachialis (Armgeflecht) oder Plexus lumbalis (Lendenervengeflecht), die zu einer Nervenkompression und einem Funktionsverlust führen (können);

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> ○ Läsionen, die zu schwerwiegenden Deformierungen (z. B. Augenläsionen) oder zu deutlichen Entstellungen führen (können); ○ Läsionen der Extremitäten, die eine Hypertrophie der Gliedmaßen oder einen Funktionsverlust verursachen (können); ○ Schmerzhaftes Läsionen. ● PN war definiert als ein Neurofibrom, welches entlang eines Nervs gewachsen ist und mehrere Faszikel und Verzweigungen betreffen kann. Ein spinale PN umfasst 2 oder mehr Ebenen mit Verbindungen zwischen den Ebenen oder erstreckt sich entlang des Nervs. Zusätzlich zu einem PN müssen alle Teilnehmenden entweder einen positiven Befund in einem Gentest für eine NF1 haben oder mindestens ein anderes der hier aufgeführten diagnostischen Kriterien für eine NF1 aufweisen: <ul style="list-style-type: none"> ○ Sechs oder mehr Café-au-lait-Flecken ($\geq 0,5$ cm bei präpubertären oder $\geq 1,5$ cm bei post-pubertären Patientinnen und Patienten); ○ Sommersprossen in Achselhöhle oder Leiste, ○ Optikusgliom, ○ zwei oder mehr Lisch-Knötchen; ○ eine ausgeprägte Knochenveränderung (Dysplasie des Keilbeins oder Dysplasie/Ausdünnung der Röhrenknochenrinde); ○ einen Verwandten ersten Grades mit NF1. ● Messbare Erkrankung <ul style="list-style-type: none"> ○ Mindestens ein messbares PN definiert als eine Läsion von mindestens 3 cm in einer Dimension. ○ Teilnahme nach einer Resektion des PN möglich, sofern das PN unvollständig reseziert wurde und messbar ist. ○ Phase II: Messbarkeit und Eignung für eine volumetrische MRT-Untersuchung der Zielläsion. Als Zielläsion wird das klinisch relevanteste PN definiert, welches für eine volumetrische MRT-Untersuchung geeignet sein muss. ● Vorherige Therapie <ul style="list-style-type: none"> ○ Mindestens 4 Wochen zwischen Absetzen einer medikamentösen Therapie für PN und Einschluss in die Studie. ○ Erholung von akuten, toxischen Effekten der Vortherapien auf \leq Grad 1 vor Einschluss in die Studie. ○ Wachstumsfaktoren, die die Anzahl oder Funktion der Thrombozyten oder der weißen Blutkörperchen unterstützen, dürfen nicht innerhalb der letzten 7 Tage vor der Aufnahme in die Studie verabreicht worden sein. ○ Vor der Aufnahme in die Studie müssen mindestens 6 Wochen seit Erhalt einer Strahlentherapie vergangen sein. ○ Es müssen mindestens 4 Wochen seit einer Operation verstrichen sein, wobei eine gute Wundheilung nachgewiesen werden muss. ● Patientinnen und Patienten > 16 Jahre müssen einen Karnofsky-Index von ≥ 70 %, Kinder und Jugendliche ≤ 16 Jahre einen Lansky-Index von ≥ 70 % aufweisen. Personen, die aufgrund einer Lähmung unabhängig von ihrem PN an den Rollstuhl gebunden sind, sollten als gehfähig betrachtet werden, wenn sie sich aufrecht in ihrem Rollstuhl befinden. Ebenso werden Personen mit einer eingeschränkten Mobilität zusätzlich zur Notwendigkeit mechanischer Unterstützung (wie bspw. eine Tracheotomie oder CPAP bei einem PN der Atemwege), für den Zweck dieser Studie als gehfähig betrachtet. ● Hämatologische Funktion <ul style="list-style-type: none"> ○ Absolute Neutrophilenzahl $\geq 1500/\mu\text{l}$. ○ Hämoglobinwert ≥ 9 g/dl. ○ Thrombozytenzahl $\geq 100.000/\mu\text{l}$. ● Leberfunktion <ul style="list-style-type: none"> ○ Bilirubin-Wert $\leq 1,5$-fachen ULN für ihre Altersgruppe mit Ausnahme von Personen mit Gilbert-Syndrom. ○ AST/ALT ≤ 3-fach ULN.

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Nierenfunktion: Kreatinin-Clearance oder radioisotope GFR von $\geq 60 \text{ ml/min/1,73 m}^2$ oder normaler Serumkreatinin-Wert entsprechend der Altersgruppe: <ul style="list-style-type: none"> ○ 0,8 mg/dl bei ≤ 5 Jahren ○ 1,0 mg/dl bei $5 < \text{Alter} \leq 10$ Jahren ○ 1,2 mg/dl bei $10 < \text{Alter} \leq 15$ Jahren ○ 1,5 mg/dl bei > 15 • Herzfunktion <ul style="list-style-type: none"> ○ Normale Ejektionsrate (anhand von Echokardiografie oder Herz-MRT) von $\geq 53 \%$ (bei der Angabe eines Bereiches, wird der obere Wert des Bereiches verwendet). ○ QTC oder QTcF von ≤ 450 msec. • Blutdruck ≤ 95. Perzentil für Alter, Größe und Geschlecht. Ein akzeptabler Blutdruck kann mit Medikamenten zur Behandlung einer Hypertonie erreicht werden. • Bereitschaft, eine übermäßige UV-Exposition zu vermeiden und adäquate Sonnenschutzmittel im Falle einer Exposition zu nutzen. • Bereitschaft, den Verzehr von Grapefruits oder Bitterorangen während der Studie zu vermeiden (sowie andere Produkte, die diese Früchte enthalten wie z. B. Säfte oder Marmeladen), da diese die Verstoffwechslung von Selumetinib beeinflussen können. <p>Ausschlusskriterien</p> <ul style="list-style-type: none"> • Phase I: Patientinnen und Patienten, bei denen die Notwendigkeit eines chirurgischen Eingriffs innerhalb der ersten drei Zyklen (3 Monate) absehbar ist. • Einnahme eines Prüfpräparats innerhalb der letzten 30 Tage. • Laufende Strahlentherapie, Chemotherapie, gegen den Tumor gerichtete Hormontherapie, Immuntherapie oder Therapie mit einem Biopharmazeutikum. • Schwere oder unkontrollierte systemische Erkrankung, aktive Infektion, aktive Blutungsneigung oder Nierentransplantation, Hepatitis B, Hepatitis C oder HIV. Patientinnen und Patienten mit HIV, die einen adäquaten CD4-Wert vorweisen und keine antiretrovirale Medikamente benötigen, können eingeschlossen werden. • Hartnäckige Übelkeit und Erbrechen, chronische gastrointestinale Erkrankung (z. B. entzündliche Darmerkrankungen) oder signifikante Darmresektion, die eine adäquate Absorption verhindern würde. • Personen, die zuvor mit Selumetinib oder einem anderen spezifischen MEK1/2-Inhibitor behandelt wurden (es sei denn die Probandin / der Proband erfüllt die Kriterien für eine erneute Behandlung). • Personen mit dem Nachweis eines Optikusglioms, eines malignen Glioms, eines MPNST oder einer anderen Krebserkrankung, die eine Behandlung mit Chemo- oder Strahlentherapie nötig machte. • Personen, bei denen Vitamin E von mehr als 100 % der empfohlenen Tagesdosis supplementiert wird. Jedes Vitamin-E-haltige Multivitaminpräparat muss vor Beginn der Therapie abgesetzt werden. • Herzfunktion <ul style="list-style-type: none"> ○ Bekannte vererbte Koronarerkrankung. ○ Symptomatische Herzinsuffizienz (nach NYHA frühere oder aktuelle Kardiomyopathie Klasse II–IV oder schwere Herzklappenerkrankung). ○ Frühere oder aktuelle Kardiomyopathie. ○ Schwere Herzklappenerkrankung. ○ Vorgeschichte von Vorhofflimmern. • Ophthalmologische Erkrankungen <ul style="list-style-type: none"> ○ Aktuelle oder frühere Vorgeschichte einer zentralen serösen Netzhautschädigung. ○ Aktueller Netzhautvenenverschluss oder Netzhautvenenverschluss in der Vorgeschichte. ○ Bekannter IOD > 21 mmHg (oder ULN angepasst an Alter) oder unkontrolliertes Glaukom (unabhängig vom IOD). Personen mit bekanntem Glaukom und erhöhtem IOD, die kein bedeutendes Sehvermögen haben (nur Lichtwahrnehmung oder keine Lichtwahrnehmung) und die keine Schmerzen im Zusammenhang mit

Charakteristikum	Beschreibung
	<p>dem Glaukom verspüren, können nach Diskussion mit der Studienleitung in Frage kommen.</p> <ul style="list-style-type: none"> ○ Personen, die eine andere signifikante Anomalie bei der augenärztlichen Untersuchung aufweisen, sollten durch die Studienleitung auf eine mögliche Eignung für die Teilnahme geprüft werden. ○ Augenärztliche Befunde, die auf ein langjähriges Gliom der Sehnervenbahn zurückzuführen sind (wie z. B. Sehverlust, Blässe des Sehnervs oder Strabismus) oder ein langjährig bestehendes orbito-temporales PN werden NICHT als signifikante Abnormität für die Zwecke der Studie betrachtet. <ul style="list-style-type: none"> • Personen mit größeren Operationen innerhalb von mindestens 4 Wochen vor Beginn der Studienbehandlung. Ausgenommen sind chirurgische Eingriffe für Gefäßzugänge. • Personen mit jeglicher Form bestehender chronischer Toxizität mit CTCAE-Grad ≥ 2 aus vorheriger Anti-NF1-Therapie. Ausgenommen hierbei ist eine Alopezie.
Intervention und Zahl der Patientinnen und Patienten	<p>Phase I</p> <p>Zur Anzahl an Personen, die für die Studie gescreent wurden, liegen keine Angaben vor.</p> <ul style="list-style-type: none"> • Selumetinib 25 mg/m² BID, N = 6; alle zum Datenschnitt vom 29.06.2018 weiter unter Beobachtung. • Selumetinib 20 mg/m² BID (nicht FI-konform), N = 12; davon 8 zum Datenschnitt vom 29.06.2018 weiter unter Beobachtung. • Selumetinib 30 mg/m² BID (nicht FI-konform), N = 6; davon 2 zum Datenschnitt vom 29.06.2018 weiter unter Beobachtung. <p>Phase II</p> <p>Zur Anzahl an Personen, die für die Studie gescreent wurden, liegen keine Angaben vor.</p> <ul style="list-style-type: none"> • Selumetinib 25 mg/m² BID: Stratum 1: N = 50 – davon 46 bzw. 44 zum DS vom 29.06.2018 bzw. 29.03.2019 weiter unter Beobachtung; Stratum 2: k. A.
Ort und Zeitraum der Durchführung	<p>Ort</p> <p>4 Zentren in den USA.</p> <p>Zeitraum</p> <p>September 2011 – laufend</p> <p>Phase I</p> <p>Erste/r Patient/in eingeschlossen: 21.09.2011 Letzte/r Patient/in eingeschlossen: 27.02.2014 Datenschnitte:</p> <ul style="list-style-type: none"> • 04.01.2016 (Daten veröffentlicht in Publikation Dombi et al., 2016 [14]) • 29.06.2018 (Studienbericht liegt vor) • 27.02.2021 (keine Daten eingereicht) <p>Phase II</p> <p>Erste/r Patient/in eingeschlossen: 12.08.2015 Letzte/r Patient/in eingeschlossen: 22.08.2016 Datenschnitte:</p> <ul style="list-style-type: none"> • Primäre Analyse: 29.06.2018 (Studienbericht liegt vor) • 29.03.2019 (nur Sicherheit) • 31.03.2021 (keine Daten eingereicht)
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte	<p>Phase I</p> <p><u>Primäre Endpunkte</u></p> <ul style="list-style-type: none"> • Bestimmung der maximal tolerierbaren Dosis, der erweiterten Verträglichkeit und der empfohlene Phase-II-Dosis. • Plasma-Pharmakokinetik

Charakteristikum	Beschreibung
gemäß Studienprotokoll	<p><u>Sekundäre Endpunkte</u></p> <ul style="list-style-type: none"> • Wachstumsrate des PN. • Adhärenz • Toxizität <p>Phase II</p> <p><u>Primärer Endpunkt</u></p> <ul style="list-style-type: none"> • Objektive Ansprechrate (ORR) in Stratum 1. <p><u>Sekundäre Endpunkte</u></p> <ul style="list-style-type: none"> • Ansprechrate in der Gesamtpopulation der behandelten Personen und in Subgruppen mit „typischem PN“ vs. „nodulärem PN“ vs. „solitär nodulärem PN“. • Langzeitverträglichkeit und Sicherheit. • Dauer des Ansprechens (DOR). • Knochenmineraldichte bei Patientinnen und Patienten am NCI mit beeinträchtigter Knochenmineraldichte zum Zeitpunkt der Aufnahme in die Studie. • Schmerz mittels NRS-11, PII, Pain Medication Survey. • Lebensqualität mittels PedsQL. • Körperliche Funktionsfähigkeit mittels 6MWT. • Funktionelle Beeinträchtigung sekundär durch PN <ul style="list-style-type: none"> ○ PN der Augenhöhle: funktionale Untersuchung (Sehfähigkeit); ○ Atemwegs-PN: funktionale Untersuchung (Schlafstudien, Lungenfunktion); ○ Motor-PN: Kraft und Bewegungsfreiheit, Bewertung der Beinlänge, Grooved Pegboard Test, PROMIS (Mobilität und obere Extremitäten); ○ Darm-/Blasen-PN: Fragebogen zur Blasen- und Darmfunktion; ○ PN, das andere Morbidität verursacht: PN-spezifische Funktionsbewertung, in Absprache mit Studienteam. • Entstellung mittels fotografischer oder Videoauswertung. • Wachstumsrate des PN mittels MRT-gestützter Analyse. • TTP und PFS bei progressivem PN. • Vergleich des PFS für progressive PN mit dem Placebo-Arm der Studie 01-C-0222. • Pharmakokinetik. • Veränderungen in Zell- und Zytokinprofilen und Korrelation mit dem Ansprechen. • Bei Personen, bei denen ein Tumor der Sehnervenbahn oder ein anderes Gliom vorliegt, welches keine Behandlung mit Chemo- oder Strahlentherapie erfordert: Bewertung der Wirkung von Selumetinib auf Veränderungen der Größe eines Tumors der Sehnervenbahn oder eines anderen Glioms. • Bewertung der Wirkung von Selumetinib auf den MAP-Kinase-Weg (ERK-Phosphorylierung) in peripheren mononukleäre Blutzellen.
Subgruppenanalysen	<p>In SAP und Protokoll wurden keine Subgruppenanalysen präspezifiziert, außer die des sekundären Endpunkts der Phase II: Ansprechrate in Subgruppen mit „typischem PN“ vs. „nodulärem PN“ vs. „solitär nodulärem PN“.</p> <p>Für einzelne Endpunkte wurden funktionelle Subgruppen gebildet:</p> <ul style="list-style-type: none"> • Grooved Pedboard Test: Subgruppen für unilateral PN und Betrachtung des Effekts der betroffenen vs. der nicht betroffenen Hand sowie eine Gruppe mit bilateraler PN. • 6MWT: Personen mit PN-der unteren Extremitäten und Kompression des Rückenmarks (Gruppe 1) und PN der oberen Atemwege (Gruppe 2).

Abkürzungen: 6MWT: 6-Minuten-Gehtest; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; BID: zweimal täglich; CPAP: Continuous Positive Airway Pressure; CTCAE: Common Terminology Criteria for Adverse Events; FI: Fachinformation; IOD: Intraokulärer Druck; k. A.: keine Angabe; NCI: National Cancer Institute; NF1: Neurofibromatose Typ 1; NRS-11: Numerical Rating Scale; MPNST: Maligner peripherer Nervenscheidentumor; MRT: Magnetresonanztomographie; NYHA: New York Heart Association; PedsQL: Pediatric Quality of Life Inventory; PFS: Progressionsfreies Überleben; PII: Pain Interference Index; PN: Plexiformes Neurofibrom; PROMIS: Patient-Reported Outcomes Measurement Information System; QTC: Frequenzkorrigiertes QT-Intervall; QTcF: Frequenzkorrigiertes QT-Intervall nach Friderica; SAP: Statistischer Analyseplan; TTP: Zeit bis zur Progression; ULN: Upper Limit of Normal.

Für die Nutzenbewertungen werden der Studienarm der Phase I mit der FI-konformen Dosierung von 25 mg/m² sowie Stratum 1 der Phase II der Studie SPRINT herangezogen (Tabelle 2). Stratum 2 der Phase II umfasst asymptomatische Personen, welche nicht vom Anwendungsgebiet von Selumetinib umfasst sind und für die keine Studienunterlagen vorliegen. Zur sprachlichen Vereinfachung bezeichnet Phase II im Folgenden daher immer (nur) das Stratum 1 der Phase II.

Für Phase I der Studie SPRINT wird der Datenschnitt der primären Analyse vom 29.06.2018 in der Nutzenbewertung dargestellt. Für Phase II wird ebenfalls der Datenschnitt der primären Analyse vom 29.06.2018 dargestellt. Für die Endpunktkategorie „Sicherheit“, inklusive Todesfälle, liegen neuere Daten eines durch die U. S. Food and Drug Administration (FDA) getriggerten Datenschnitts vom 29.03.2019 vor, welche für diese Endpunkte zur Nutzenbewertung ergänzend herangezogen werden. Ein weiterer neuerer Datenschnitt wurde von der European Medicines Agency (EMA) im Zusammenhang mit der bedingten Zulassung von Selumetinib beauftragt (Phase I: 27.02.2021; Phase II: 31.03.2021). Der pU gibt an, dass ihm für diese Datenschnitte zum Zeitpunkt der Dossier-einreichung noch keine Daten bzw. Studienberichte vorlagen.

Protokolländerungen

Es wurden insgesamt 16 Änderungen des Originalprotokolls vom 06.04.2011 vorgenommen. Protokolländerungen innerhalb des Rekrutierungszeitraums der Phase I betrafen u. a. die Altersgrenze und erlaubten den Einschluss deutlich jüngerer Kinder (ab 3 Jahren) als anfangs geplant (ab 13 Jahren). Neben Änderungen an den Ein- und Ausschlusskriterien betrafen weitere Änderungen die Kriterien für Dosisanpassungen und die Wiederaufnahme der Therapie nach Dosisunterbrechungen, mit dem Ziel eine möglichst lange Behandlungsdauer zu ermöglichen. Zwischen Abschluss der Rekrutierung der Phase I und Beginn der Rekrutierung für Phase II wurden gehäuft Änderungen an den Ein- und Ausschlusskriterien vorgenommen (siehe Tabelle 4), sodass für Phase II geringfügig andere Kriterien galten als für Phase I.

Für den SAP werden insgesamt 5 Versionen gelistet. Version 1.0 ist auf den 28.09.2016 datiert. Die aktuellste und für den Studienbericht relevante Version 5.0 wurde erst nach dem Datenschnitt (vom 29.06.2018) am 21.01.2019 fertiggestellt. Es werden jedoch im Studienbericht nur wenig relevante Änderungen nach Verfügbarkeit der finalen, validierten Ergebnisse gelistet.

Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Phase I der Studie SPRINT

Amendment-version	Wesentliche Protokolländerungen der Phase I der Studie SPRINT
Version A (24.10.2011)	Überarbeitung der "Comprehensive Adverse Events and Potential Risk List" bezüglich Reporting der auf die Behandlung mit Selumetinib zurückzuführenden UE.
Version B (30.01.2012)	Änderung des Einschlusskriteriums für Alter, sodass Kinder ab 12 Jahren eingeschlossen werden konnten.
Version C (30.04.2012)	<ul style="list-style-type: none"> Studie wurde als multizentrische Studie erweitert (NCI POB koordinierendes Zentrum). Behandlungsdauer wurde flexibilisiert, sodass die Behandlung im Interesse der Patientinnen und Patienten auch früher beendet werden konnte.
Version D (16.07.2012)	<ul style="list-style-type: none"> Änderung des Einschlusskriteriums für Alter, sodass Kinder ab 3 Jahren eingeschlossen werden konnten, wenn sie in der Lage waren ganze Kapseln (zur Einnahme der Studienmedikation) zu schlucken.
Version E (30.12.2012)	<ul style="list-style-type: none"> Änderung des Einschlusskriteriums bezüglich Vortherapie, sodass 6 Wochen zwischen Einschluss und einer vorangegangenen Strahlentherapie liegen mussten. Update BSC mit Behandlungsoptionen für zu erwartende Standardtoxizitäten.

Amendment- version	Wesentliche Protokolländerungen der Phase I der Studie SPRINT
	<ul style="list-style-type: none"> • Aktualisierung MRT-Erfassungszeitpunkte auf Studienbeginn, vor den Zyklen 6, 11 und dann jeweils nach 6 Zyklen, solange die Patientin / der Patient in der Studie verbleibt.
Version F (31.02.2013)	Es wurde die Möglichkeit ergänzt, die Dosis ein zweites Mal zu reduzieren.
Version G (26.08.2013)	<ul style="list-style-type: none"> • Einschlusskriterium für die Herzfunktion wurde hinzugefügt, „normale Ejektionsfraktion (Echokardiografie) $\geq 55\%$ (wenn ein Bereich angegeben ist, wird der obere Wert verwendet)“. • Kardiale Toxizität wurde als Dosis-limitierende Toxizität hinzugefügt, Bedingungen für eine Weiterbehandlung mit Selumetinib wurden festgelegt und Details zur Behandlung von asymptomatischen und symptomatischen Verminderungen der LVEF beschrieben. • Studienbegleitende Auswertungen wurden modifiziert, um eine Echokardiografie-Überwachung bei Personen zu ermöglichen, die die Therapie wegen kardialer Toxizität abbrechen und die Selumetinib nach Erholung der LVEF wieder aufnehmen.
Version H (05.01.2015)	<ul style="list-style-type: none"> • \geq Grad 1 Katarakt wurde als Ausschlusskriterium gestrichen und wurde auch nicht länger als Dosis-limitierende Toxizität betrachtet, stattdessen sollten Katarakte des Grades 3 von Fall zu Fall in Absprache mit der Studienleitung und der Ophthalmologie berücksichtigt werden. • Dosisstufe von 25 mg/m² BID wurde in das Dosierungsschema aufgenommen.
Version I (06.05.2015)	<ul style="list-style-type: none"> • Definition von Toxizität, die eine Dosisänderungen erfordern, wurde aktualisiert und UE wurden aufgenommen: Kardiale Toxizität, Harnwegsobstruktion Grad 2 (nicht tumorbezogen), Netzhautablösung Grad 2, Pneumonitis Grad 2, Netzhautvenenthrombose oder jede nicht tolerierbare Selumetinib-bezogene Toxizität Grad 2 von ≥ 7 Tagen, die nicht durch Standardmaßnahmen kontrolliert werden können. • Neue Dosismodifikation hinzugefügt, die ein 5-Tage-on-/2-Tage-off-Schema beinhaltet. • Es wurde festgelegt, dass bei Personen, bei denen eine PR auftritt und bei denen die Ziel-PN für eine Operation zugänglich wird, die Selumetinib-Therapie nach der Entfernung fortgesetzt werden kann, sofern das Ziel-PN noch messbar ist. • Personen, die bereits aus der Behandlung genommen wurden und bei denen sich eine progrediente Erkrankung entwickelt hat wurde ein Neustart der Selumetinib-Behandlung ermöglicht.
Version J (28.12.2015)	Keine für die Nutzenbewertung relevanten Änderungen.
Version K (26.04.2016)	Relevante Änderungen siehe Phase II.
Version L (10.08.2016)	Relevante Änderungen siehe Phase II.
Version M (28.09.2016)	Relevante Änderungen siehe Phase II.
Version N (18.06.2018)	Relevante Änderungen siehe Phase II.
Version O (14.11.2018, nach 1. DS)	Relevante Änderungen siehe Phase II.
Version P (14.01.2019)	Relevante Änderungen siehe Phase II.

Abkürzungen: BID: zweimal täglich; BSC: Best Supportive Care; DS: Datenschnitt; LVEF: linksventrikuläre Ejektionsfraktion; MRT: Magnetresonanztomographie; NCI POB: National Cancer Institute – Pediatric Oncology Branch; PN: Plexiformes Neurofibrom; PR: Partielles Ansprechen; UE: Unerwünschte Ereignisse.

Tabelle 4: Für die Nutzenbewertung relevante Protokolländerungen der Phase II der Studie SPRINT

Amendment- version	Wesentliche Protokolländerungen der Phase I der Studie SPRINT
Version H (05.01.2015 ¹⁾)	<ul style="list-style-type: none"> • \geq Grad 1 Katarakt wurde als Ausschlusskriterium gestrichen und wurde auch nicht länger als Dosis-limitierende Toxizität betrachtet, stattdessen sollten Katarakte des Grades 3 von Fall zu Fall in Absprache mit dem Studienleiter und der Ophthalmologie berücksichtigt werden. • Dosisstufe von 25 mg/m² BID wurde in das Dosierungsschema aufgenommen.
Version I (06.05.2015 ¹⁾)	<ul style="list-style-type: none"> • Aktualisierung der Einschlusskriterien <ul style="list-style-type: none"> ○ Aufnahme einer Definition für inoperable PN. ○ Verwendung zulässiger vorheriger Prüfsubstanzen oder biologischer Therapie und vorheriger Operationen. ○ Anforderung für minimal erforderliche absolute Neutrophilenzahl wurde von $\geq 1.000/\mu\text{l}$ auf $\geq 1.500/\mu\text{l}$ geändert. ○ Anforderung für die Leberfunktion wurde von „ALT innerhalb $\leq 1,5 \times$ oberer Grenzwert der Norm“ in „AST/ALT innerhalb $\leq 3 \times$ oberer Grenzwert der Norm“ geändert. ○ Aktualisierung des Einschlusskriteriums 9 (Herzfunktion). Die normale Ejektionsfraktion wurde gemäß NIH-Untergrenze von 55 auf 53 % (oder den institutionellen Normalwert) aktualisiert und QTcF ≤ 450 msec wurde hinzugefügt. ○ Details zum Ausschluss von Personen mit Hepatitis B, Hepatitis C oder HIV hinzugefügt. ○ Vorherige Behandlung mit MEK1/2-Inhibitoren wurde ausgeschlossen (zuvor war nur Vorbehandlung mit Selumetinib ausgeschlossen). • Hinzufügen mehrerer Einschlusskriterien bzgl. Blutdruck, Vermeidung von Sonnenexposition und Verzehr von Grapefruit und Bitterorangen. • Hinzufügen von Ausschlusskriterien bezüglich <ul style="list-style-type: none"> ○ refraktärer Übelkeit und Erbrechen, chronische gastrointestinaler Erkrankungen (z. B. entzündliche Darmerkrankungen) oder signifikanter Darmresektionen; ○ Personen, die trotz antihypertensiver Therapie zur Blutdruckkontrolle keinen ausreichenden Blutdruck erreichen; ○ kardialer Funktion; ○ ophthalmologischer Bedingungen; ○ klinischem Urteil der Prüferin / des Prüfartzes; ○ Vermeidung von zusätzlichen Medikamenten, die nicht zur Studie gehören und die mit den Studienmedikamenten interferieren können; ○ Überempfindlichkeit gegen Selumetinib oder Hilfsstoffe; ○ kürzlich durchgeführter größere Operation; ○ ungelöster chronischer Toxizität durch vorherige Anti-NF1-Therapie.
Version J (28.12.2015 ¹⁾)	Keine für die Nutzenbewertung relevanten Änderungen.
Version K (26.04.2016)	Überarbeitung der „Comprehensive Adverse Events and Potential Risk List“ bezüglich Reporting der auf die Behandlung mit Selumetinib zurückzuführenden UE.
Version L (10.08.2016)	<ul style="list-style-type: none"> • Personen, die nach Dosisreduktion aufgrund von Toxizitäten und erneutem Tumorstadium auf eine 5-Tage-/2-Tage Dosierung mit der Ursprungsdosis gewechselt waren, wurde die kontinuierliche Dosierung mit der Ursprungsdosis erlaubt. • Die Kriterien für das Absetzen der Protokolltherapie wurden zur Vollständigkeit der Sicherheitsberichterstattung aktualisiert, um klarzustellen, dass vor der Dokumentation des Absetzens von der Protokolltherapie Anstrengungen unternommen werden müssen, damit alle Personen etwa 30 Tage nach der letzten Dosis der Studientherapie eine Sicherheitsuntersuchung absolvieren.
Version M (28.09.2016)	<ul style="list-style-type: none"> • Kriterien für den Beginn nachfolgender Behandlungszyklen wurden aktualisiert, um Personen, bei denen eine Gewichtszunahme \geq Grad 3 auftrat, die Möglichkeit zu geben, nach Ermessen der Prüferin / des Prüfartzes mit einer reduzierten Dosis bei \leq Grad 2 erneut zu beginnen.

Amendment-version	Wesentliche Protokolländerungen der Phase I der Studie SPRINT
	<ul style="list-style-type: none"> • Aktualisierung des Zeitraums für die Screening-MRT-Untersuchung von innerhalb von 4 bis 6 auf innerhalb von 6 Wochen vor Aufnahme in die Studie geändert. • Anhang VIII des Studienprotokolls wurde aktualisiert, um klarzustellen, dass Personen, die eine Tracheostomie hatten, welche die, durch das PN verursachte, Atemwegsobstruktion umging, keine Schlafstudienauswertung benötigten. Personen, die eine CPAP- oder BiPAP-Unterstützung benötigten, sollten eine Baseline-Schlafstudie in ihrer häuslichen Umgebung durchführen lassen. • Anhang VIII des Studienprotokolls (Atemwegsfunktion) wurde aktualisiert, um klarzustellen, dass Personen mit einer Tracheostomie keine Lungenfunktionstests durchführen müssen.
Version N) (18.06.2018)	Überarbeitung der "Comprehensive Adverse Events and Potential Risk List".
Version O (14.11.2018, nach 1. DS)	<ul style="list-style-type: none"> • Klarstellung, dass das volumetrische MRT, das zur Bestimmung der Eignung für eine erneute Behandlung verwendet wird, als neuer Baseline-Scan für nachfolgende volumetrische Vergleiche verwendet werden sollte, anstelle des volumetrischen Baseline-MRT bei Studieneinschluss. • Aktualisierung der Abschnitte zu den Kriterien für das Absetzen der Protokolltherapie und den Kriterien für das Verlassen der Studie, damit die Patientinnen und Patienten länger in Behandlung bleiben können.
Version P (14.01.2019)	Überarbeitung der "Comprehensive Adverse Events and Potential Risk List".

¹⁾ Protokolländerung nach Abschluss der Rekrutierung für Phase I und vor Beginn der Phase II.

Abkürzungen: ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; BID: zweimal täglich; BiPAP: Bilevel Positive Airway Pressure; CPAP: Continuous Positive Airway Pressure; DS: Datenschnitt; MRT: Magnetresonanztomographie; NF1: Neurofibromatose Typ 1; NIH: National Institute of Health; PN: Plexiformes Neurofibrom; QTcF: Frequenzkorrigiertes QT-Intervall nach Friderica; UE: Unerwünschtes Ereignis.

Charakterisierung der Intervention

Die in der Fachinformation empfohlene Dosierung von 25 mg/m² Körperoberfläche wurde in 6 Personen der Phase I und in allen Personen der Phase II eingesetzt (Tabelle 2). Diese Population wird für die Nutzenbewertung herangezogen. Detaillierte Angaben zur Behandlung mit Selumetinib in der für die Nutzenbewertung relevanten Dosierung sind in Tabelle 5 aufgeführt.

Tabelle 5: Charakterisierung der Intervention Selumetinib der Studie SPRINT

Intervention
<p>Die Dosis betrug 25 mg/m² Körperoberfläche 2 x täglich. Die Dosis wurde auf die nächste 10mg-Stufe gerundet und bei einer Körperoberfläche von $\geq 2 \text{ m}^2$ gedeckelt. Die Einnahme sollte ungefähr alle 12 Stunden und auf leeren Magen erfolgen, ohne Essen oder Trinken 2 h vor und 1 h nach der Einnahme.</p> <p>Die Behandlung erfolgte in Zyklen von 28 Tagen ohne Einnahmepausen zwischen den Zyklen. Nach Beendigung eines Behandlungszyklus war die Weiterbehandlung mit der gleichen Dosis möglich, wenn keine Dosis-limitierenden Toxizitäten (Phase I) oder Toxizitäten, die eine Anpassung der Dosierung notwendig machten (Phase II) aufgetreten waren (siehe unten). Bei Personen, bei denen ebensolche Toxizitäten aufgetreten sind, mussten sich diese auf \leq Grad 1 zurückgebildet haben, bevor ein neuer Zyklus begonnen werden konnte. Falls innerhalb von 21 Tagen nach Beendigung des Zyklus kein Rückgang der Toxizität erfolgte, musste die Therapie beendet werden, außer wenn die/der Teilnehmende einen klaren klinischen Nutzen (partielles Ansprechen oder stabile Erkrankung bei Personen mit progressiver Erkrankung zum Zeitpunkt des Studieneinschlusses oder Verbesserung von Symptomen oder Funktionen) durch die Therapie hatte und eine Erholung innerhalb von 3 Monaten nach Beendigung der Therapie erfolgte. Außerdem konnte bei einer Gewichtszunahme \geq Grad 3 die Therapie im Ermessen des ärztlichen Personals fortgesetzt werden, wenn sich die Gewichtszunahme auf \leq Grad 2 zurückgebildet hatte.</p>

Intervention

Toxizitäten, die eine Anpassung der Dosierung notwendig machten

Veränderungen der Dosierungen waren notwendig bei jeglichen UE \geq Grad 3, welche möglicherweise, wahrscheinlich oder definitiv durch Selumetinib verursacht wurden. Es wurden Ausnahmen für verschiedene Laborwertabweichungen, Hautauschlag und gastrointestinale Toxizität definiert, die entweder nach Absetzen der Behandlung innerhalb von 2 Wochen zurückgingen oder durch Supplementation bzw. supportive Interventionen behandelbar waren. Für Gewichtszunahme und Katarakte wurden ebenfalls Ausnahmeregelungen definiert. Weitere UE, welche eine Änderung der Dosierung erforderten, waren:

- Kardiale Toxizität.
- Grad 2 Harnwegsobstruktion, welche mindestens möglicherweise mit der Einnahme von Selumetinib in Verbindung steht (und klar nicht durch einen Tumor verursacht wurde).
- Netzhautablösung Grad 2.
- Pneumonie Grad 2.
- Jegliche auf die Behandlung mit Selumetinib zurückzuführende Toxizität des Grades 2, welche länger als 7 Tage anhält, welche nicht toleriert wird und nicht durch supportive Standardmaßnahmen kontrolliert werden kann.
- Netzhautvenenthrombose (bei Auftreten permanente Beendigung der Therapie).

Dosisanpassungen

Nach Pausieren der Therapie aufgrund von Toxizitäten und Rückbildung der Toxizitäten konnte eine Wiederaufnahme der Therapie mit einer um 25 bis 33 % reduzierten Dosis erfolgen. Dosisreduktionen waren zweimal möglich bevor die Therapie aufgrund von Toxizitäten beendet werden musste. Eine Erhöhung auf die ursprüngliche Dosis sollte nicht erfolgen. Bei Wachstum des Tumorumfanges um ≥ 10 % und keinem Erreichen der Kriterien für einen Progress, und wenn die Selumetinib-Dosis gut vertragen wurde, war eine Wiederaufnahme der Therapie mit der ursprünglichen Dosierung für 5 Tage und der reduzierten Dosis für 2 Tage möglich. Sollte trotz Erhöhung der Dosis ein weiteres Wachstum des Tumors erfolgen und die 5+2-Dosierung gut vertragen werden, konnte die Dosierung an allen Tagen auf die Anfangsdosis erhöht werden.

Dauer der Behandlung

- Für Teilnehmende mit einer dokumentierten Krankheitsprogression in den 1,5 Jahren vor Einschluss in die Studie (definiert als ≥ 20 % Zunahme des Tumorumfanges (oder ≥ 13 % Zunahme des Produkts der längsten 2 perpendikularen Durchmesser des Tumors oder ≥ 6 % Zunahme des längsten Durchmessers) gab es keine Limitation der Behandlungsdauer. Die Therapie konnte durch das ärztliche Prüfpersonal im Interesse der Patientin / des Patienten beendet werden.
- Bei Teilnehmenden ohne Krankheitsprogression in den 1,5 Jahren vor Studieneintritt und ohne Ansprechen (Reduktion des Tumorumfanges um ≥ 20 %) sollte die Behandlung nach 2 Jahren beendet werden.
 - Sollte in den Follow-up-Untersuchungen im Anschluss an die 2-jährige Therapie ein Progress festgestellt werden (Wachstum ≥ 15 % im Vergleich zum Behandlungsende innerhalb von ungefähr 2 Jahren), konnte die Therapie mit Selumetinib wiederaufgenommen werden.
 - Bei Personen mit Ansprechen gemäß Bildgebung wurde die Behandlungsdauer nicht begrenzt, außer bei darauffolgender Krankheitsprogression oder Erreichen anderer Kriterien, die eine Beendigung der Therapie notwendig machen.
- Bei Teilnehmenden mit einem partiellen Ansprechen, in denen nach der Therapie mit Selumetinib der Tumor durch eine Operation teilweise entfernt werden kann, konnte die Therapie nach der Operation fortgesetzt werden, wenn der Tumor nach der Operation weiter messbar war.

Erlaubte Begleitmedikation

- Geeignete Antibiotika, Unterstützung mit Blutprodukten, Antiemetika, allgemeine supportive Maßnahmen. Für die Behandlung spezifischer Nebenwirkungen wurden Leitlinien mit Behandlungsoptionen formuliert.
- Einnahme von Kortikosteroiden zur Kontrolle der NF1-Symptomatik oder aus anderen Gründen.
- Die Einnahme der Begleitmedikation sollte über die Studiendauer möglichst unverändert bleiben.

Intervention
Nichterlaubte Begleitmedikationen <ul style="list-style-type: none"> • Andere Chemotherapien, Strahlentherapie, Immuntherapien, biologische Therapien, hämatopoetische Wachstumsfaktoren oder andere Prüfsubstanzen. • Vitamin-E-Dosen größer als 100 % der täglich empfohlenen Dosis, Vitamin E in Form von Supplementen (auch Multivitamin) sollen nicht eingenommen werden, da Vitamin E bereits in Selumetinib-Kapsel enthalten ist. • Einnahme von Selumetinib mit Vorsicht bei Personen, die ebenfalls Kumarin-haltige Antikoagulantien, z. B. Warfarin, einnehmen. • Personen sollen Medikamente vermeiden, die die Aktivität der hepatischen, mikrosomalen Isoenzyme CYP1A2, CYP2C19 und CYP3A4 beeinflussen. Welche Medikamente das betraf, wurde auf einer Liste festgehalten, die regelmäßig aktualisiert wurde.

Abkürzungen: NF1: Neurofibromatose Typ 1; UE: Unerwünschte Ereignis.

2.3 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden Herstellerdossier, zugehöriger Studienbericht, Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Tabelle 6 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 6: Zusammenfassung der Endpunktbewertung

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle	Mortalität	Ja	Ja
Ansprechen Objektive Ansprechrates (ORR) ¹ Bestes objektives Ansprechen (BOR) Zeit bis zum Ansprechen (TTR) Dauer des Ansprechens (DOR)	Morbidität	Ja	Nein
Verringerung des Tumolvolumens	Morbidität	Ja	Nein
Progressionsfreies Überleben (PFS)	Morbidität	Ja	Nein
Globalbeurteilung der klinischen Veränderung mittels Global Impression of Change (GIC)	Morbidität	Ja	Ja
Schmerz	Morbidität		
Numerical Rating Scale (NRS-11)	Morbidität	Ja	Ja
Pain Interference Index (PII)	Morbidität	Ja	Nein
Motorfunktion	Morbidität		
Manueller Muskeltest	Morbidität	Ja	Nein
Greifkraft und Kraft des Schlüsselgriffs	Morbidität	Ja	Nein
Bewegungsumfang der Gelenke	Morbidität	Ja	Nein
Grooved Pegboard Test	Morbidität	Ja	Ja
PROMIS (Mobilität und obere Extremitäten)	Morbidität	Ja	Ja
Unterschiede in der Beinlänge	Morbidität	Ja	Nein

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Ophtamologische Endpunkte			
Sehschärfe	Morbidität	Ja	Ja ²⁾
Proptosis / Exophthalmus	Morbidität	Ja	Nein
Atemfunktion			
Einsekundenkapazität (FEV _{1 / 0,75})	Morbidität	Ja	Nein
Atemwegswiderstand (R ₂₀)	Morbidität	Ja	Nein
Apnoe-Hypopnoe-Index (AHI)	Morbidität	Ja	Nein
Darm- und Blasenfunktion mittels Dysfunctional Voiding Questionnaire (DVQ)	Morbidität	Ja	Nein
Symptomatik mittels Symptomcheckliste	Morbidität	Ja	Ja
Gehfähigkeit mittels 6-Minuten-Gehtest (6MWT)	Morbidität	Ja	Ergänzend
Pediatric Quality of Life Inventory (PedsQL)	Lebensqualität	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

¹⁾ Primärer Endpunkt.

²⁾ Gemessen mittels HOTV-Sehtafel.

Abkürzungen: PROMIS: Patient-Reported Outcomes Measurement Information System; pU: pharmazeutischer Unternehmer.

2.3.1 Mortalität

Todesfälle

Der Endpunkt „Todesfälle“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Todesfälle wurde im Rahmen der Sicherheit erfasst. Patientinnen und Patienten sollten für insgesamt 7 Jahre nach Beginn der Behandlung mit Selumetinib oder 5 Jahre nach Beendigung der Behandlung mit Selumetinib nachverfolgt werden, je nachdem was länger ist. In Protokoll oder SAP wurden keine Analysen bezüglich des Gesamtüberlebens präspezifiziert.

Bewertung

Es erfolgte keine Erhebung des Gesamtüberlebens über die Teilnahme an Studienvisiten hinaus.

Patientenrelevanz

Entsprechend § 2 Satz 3 Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Gesamtüberleben ist demzufolge patientenrelevant.

Validität

Es wird davon ausgegangen, dass die Erfassung von Todesfällen als Sicherheitsendpunkt valide erfolgte.

2.3.2 Morbidität

Ansprechen

Für den Endpunkt „Ansprechen“ wurden separate Analysen für „Objektive Ansprechrate“ (ORR), „Bestes Objektives Ansprechen“ (BOR), „Zeit bis zum Ansprechen“ (TTR) und „Dauer des Ansprechens“ durchgeführt. Der Endpunkt „Ansprechen“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz.

Operationalisierung

Beschreibung

Die Erhebung zum Ansprechen erfolgte mittels 3D-MRT nach einem festgelegten MRT-Protokoll. Die volumetrische Analyse der MRT-Aufnahmen erfolgte zentral am NCI POB. Zusätzlich wurde im SAP vom 21.09.2019 festgelegt, dass die 3D-MRT unabhängig gemäß den modifizierten REiNS-Kriterien und den modifizierten RECIST 1.1 durch einen externen Dienstleister bewertet werden. Zudem wurden Regeln festgelegt, wie damit umgegangen werden sollte, wenn die Zielläsion Post-Baseline nur teilweise messbar war. Die Erhebungszeitpunkte unterschieden sich leicht zwischen den Studienphasen (siehe Tabelle 7 bzw. Tabelle 8). Zu Baseline sollten alle messbaren PN mittels MRT erfasst werden. Das klinisch relevanteste, inoperable PN, welches mittels volumetrischer 3D-MRT-Messung erfasst werden konnte, wurde als Zielläsion ausgewählt. Daneben konnten maximal 2 weitere, inoperable PN als zusätzliche Nicht-Zielläsionen für weitere 3D-MRT-Untersuchungen ausgewählt werden. Bei mehr als 3 inoperablen PN wurden die 3 PN mit der höchsten klinischen Relevanz ausgewählt.

Für Phase I wurden folgende Responsekriterien gewählt:

- Responsive Disease: $\geq 20\%$ ige Reduktion der Summe des Volumens aller Zielläsionen über ≥ 4 Wochen.
- Stabile Erkrankung: $< 20\%$ ige Reduktion und $< 20\%$ ige Erhöhung der Summe des Volumens aller Zielläsionen über ≥ 4 Wochen.
- Progrediente Erkrankung: (PD): $\geq 20\%$ ige Zunahme des Volumens von mindestens einem des Ziel-PN im Vergleich zum vor Beginn der aktuellen Behandlungsphase gemessenen Volumen. Das Auftreten einzelner, neuer subkutaner Neurofibrome gilt nicht als Krankheitsprogression. Eine Verschlimmerung bestehender Symptome oder das Auftreten neuer Symptome, die länger als 7 Tage andauern und von denen man annimmt, dass sie eindeutig mit dem PN zusammenhängen, sollten durch eine Wiederholung der MRT-Untersuchung bewertet werden. Patientinnen und Patienten sollten nicht allein aufgrund neuer oder verstärkter Symptome als Personen mit fortschreitender Erkrankung eingestuft werden, ohne dass der Fall mit dem ärztlichen Prüfpersonal besprochen wurde.

In Phase II galten folgende, leicht abweichende Responsekriterien:

- Vollständiges Ansprechen: Verschwinden der Zielläsion.
- Partielles Ansprechen (PR): Verringerung des Volumens der Ziel-PN um 20 % oder mehr im Vergleich zum Ausgangswert. Das PR gilt als unbestätigt beim ersten Nachweis, bestätigt, wenn es innerhalb von 3 bis 6 Monaten erneut beobachtet wird, und anhaltend, wenn das Ansprechen für 6 Monate oder länger anhält.
- Stabile Erkrankung: Unzureichende Volumenveränderung, um entweder als PR oder PD eingestuft zu werden.
- Progrediente Erkrankung (PD): Zunahme des Volumens der Ziel-PN um 20 % oder mehr im Vergleich zum Ausgangswert oder zum Zeitpunkt des besten Ansprechens nach Nachweis

einer PR. Das Auftreten neuer PN (mit Ausnahme von einzelnen, neuen subkutanen Neurofibromen) oder eine eindeutige Progression bestehender klinisch relevanter Nicht-Ziel-PN wird ebenfalls als PD betrachtet. Eine Verschlimmerung bestehender Symptome oder das Auftreten neuer Symptome, die länger als 7 Tage andauern und von denen man annimmt, dass sie eindeutig mit dem PN zusammenhängen, sollten durch eine Wiederholung der MRT-Untersuchung bewertet werden. Patientinnen und Patienten sollten nicht allein aufgrund neuer oder verstärkter Symptome als Personen mit fortschreitender Erkrankung eingestuft werden, ohne dass der Fall mit dem ärztlichen Prüfpersonal besprochen wurde.

ORR (primärer Endpunkt) wurde definiert als der prozentuale Anteil der Patientinnen und Patienten der Sicherheitspopulation des Stratum 1 der Phase II, welche ein bestätigtes oder partielles Ansprechen erreichen. Es wurden Daten bis zur Progression oder bis zum letzten evaluierbaren Messzeitpunkt herangezogen.

Als BOR wird das beste Ansprechen auf die Behandlung gewertet, definiert als die maximale Schrumpfung des Tumors im Vergleich zum Ausgangswert.

Die TTR-Analyse wurde im SAP, jedoch nicht im Studienprotokoll spezifiziert. „Zeit bis zum Ansprechen“ war definiert als Zeit vom ersten Zyklus der Studienbehandlung bis zur MRT-Untersuchung vor einem neuen Behandlungszyklus, bei dem ein vollständiges Ansprechen oder ein unbestätigtes partielles Ansprechen (sofern dieses im Nachgang bestätigt wird) festgestellt wird.

„Dauer des Ansprechens“ wird für Personen mit vollständigem und teilweise Ansprechen berechnet als die Zeit vom ersten dokumentierten Ansprechen bis zur dokumentierten Progression oder bis zum Tod jedweder Ursache. Die Zensierungen erfolgten wie für PFS.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Es werden keine konkreten Kriterien zur Auswahl des Ziel-PN beschrieben. Die Operationalisierungen des Ansprechens unterscheiden sich zwischen Phase I und Phase II, was die Vergleichbarkeit der Ergebnisse einschränkt.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Ansprechen verbunden mit einer für die Patientin / den Patienten spürbaren Abnahme von Krankheitssymptomen ist für die Nutzenbewertung grundsätzlich patientenrelevant. Der pU beschreibt im Dossier, dass durch das Volumen, welches die PN einnehmen, je nach Lokalisation eine Vielzahl von Symptomen ausgelöst werden können, welche die Patientinnen und Patienten im Alltag sehr stark körperlich einschränken und psychisch belasten können. Gemäß Einschlusskriterien des Stratum 1 der Phase II der Studie SPRINT weisen alle Teilnehmenden eine PN-bezogene Morbidität, wie Schmerzen, Entstellungen oder Funktionseinschränkungen auf. Vom pU wird auf die durch PN verursachten massiven Entstellungen der Patientinnen und Patienten verwiesen, die insbesondere durch oberflächlich sitzende PN verursacht werden. Die Symptomatik findet bei der Bewertung des Ansprechens jedoch keine Berücksichtigung. Zudem ist unklar, ob das Erreichen des (partiellen) Ansprechens im Sinne einer Volumenreduktion des Ziel-PN von 20 % oder mehr zwangsläufig mit einer klinisch relevanten Verbesserung der Symptomatik einhergeht. Das Ansprechen wird durch die radiologisch festgestellte Volumenänderung einer zu Baseline festgelegten Zielläsion definiert.

Weitere PN werden in Phase II (und damit der Mehrheit der Studienteilnehmenden) nicht berücksichtigt. Bspw. bedeutet ein vollständiges Ansprechen nicht, dass die Patientin / der Patient keine PN mehr aufweist, sondern lediglich, dass die Zielläsion verschwunden ist. Weitere PN könnten demzufolge weiterhin vorliegen und mit einer für die Patientin / den Patienten spürbaren Symptomatik einhergehen. Ebenfalls bleiben nicht volumetrisch messbare PN bei der Bewertung des Ansprechens unberücksichtigt, können jedoch zur Morbidität beitragen. PN können komplexe Formen annehmen und netzartig und verdrängend wachsen. Nicht volumetrisch messbare PN können wesentlich zum klinischen Bild beitragen, werden in der vorliegenden Operationalisierung jedoch nicht berücksichtigt. ORR und BOR werden aus diesen Gründen nicht als unmittelbar patientenrelevant erachtet. Die mit einem Ansprechen einhergehende mögliche Verbesserung der Symptomatik sollte sich in anderen Endpunkten der Kategorien „Morbidität“ und „Lebensqualität“ widerspiegeln. ORR und BOR stellen zudem keine validierten Surrogate eines patientenrelevanten Endpunkts in der vorliegenden vom Anwendungsgebiet umfassten Population dar. Demzufolge können auch TTR und „Dauer des Ansprechens“ nicht als patientenrelevant gewertet werden.

Validität

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

Veränderung des Tumorumfanges

Der Endpunkt „Verringerung des Tumorumfanges“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz.

Operationalisierung

Beschreibung

Der Endpunkt wird als Volumenänderung des als Zielläsion definierten PN, gemessen mittels 3D-MRT, definiert. Details zu Messung und Auswertung sind für den Endpunkt „Ansprechen“ beschrieben. Laut SAP wird die absolute Änderung und die prozentuale Änderung vom jeweiligen Erhebungszeitpunkt im Vergleich zu Baseline berechnet sowie die größte Reduktion des Tumorumfanges bzw. minimalste Vergrößerung deskriptiv berichtet. Im Dossier werden zusätzlich Responderanalysen für eine Verringerung des Tumorumfanges um mindestens 20, 40, 60 und 80 % dargestellt. Eine Präspezifizierung dieser ResponderSchwellen konnte in den Studienunterlagen nicht identifiziert werden.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Es gelten die gleichen Einschränkungen der Patientenrelevanz wie für den Endpunkt „Ansprechen“ dargelegt. Der Endpunkt wird insgesamt als nicht unmittelbar patientenrelevant erachtet, da die auf Bildgebung beruhende Messung des Tumorumfanges Änderungen in der Symptomatik nicht berücksichtigt und weiterhin nicht alle (klinisch relevanten) PN in der Betrachtung berücksichtigt werden.

Validität

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

Progressionsfreies Überleben

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

Operationalisierung

Beschreibung

Progressionsfreies Überleben (PFS) war definiert als die Zeit vom 1. Zyklus der Studienbehandlung bis zur MRT-Bewertung, in der eine Progression oder der Tod durch jegliche Ursache festgestellt wurde, unabhängig davon, ob die Patientin / der Patient die Studientherapie abbrach oder eine andere NF1-PN-Behandlung (nach Absetzen der Studienbehandlung) vor der Progression erhielt.

Patientinnen und Patienten, die zum Zeitpunkt der Analyse weder eine Krankheitsprogression erfahren hatten noch verstorben waren, wurden zum Zeitpunkt ihrer letzten auswertbaren MRT-Untersuchung zensiert. Wenn die Patientin / der Patient nach zwei oder mehr verpassten MRT-Untersuchungen eine Krankheitsprogression erfuhr oder verstarb, wurde sie/er zum Zeitpunkt der letzten auswertbaren MRT-Bewertung zensiert. Falls für eine Patientin / einen Patienten keine auswertbaren MRT-Bewertungen oder keine Baseline-Daten vorlagen, wurde sie/er zum 1. Tag des 1. Zyklus der Studienbehandlung zensiert, es sei denn, die Patientin / der Patient verstarb innerhalb der ersten zwei Behandlungszyklen. Die Progression war definiert als eine mindestens 20%ige Zunahme des Volumens der Zielläsion, gemessen anhand volumetrischer MRT-Aufnahmen. Das Auftreten neuer PN (mit Ausnahme von einzelnen, neuen subkutanen Neurofibromen) oder eine eindeutige Progression bestehender klinisch relevanter Nicht-Ziel-PN wird ebenfalls als PD betrachtet. Eine Verschlimmerung bestehender Symptome oder das Auftreten neuer Symptome, die länger als 7 Tage andauern und von denen man annimmt, dass sie eindeutig mit dem PN zusammenhängen, sollte durch eine Wiederholung der MRT-Untersuchung bewertet werden. Patientinnen und Patienten sollten nicht allein aufgrund neuer oder verstärkter Symptome als Personen mit fortschreitender Erkrankung eingestuft werden, ohne dass der Fall mit dem ärztlichen Prüfpersonal besprochen wurde.

Zusätzlich zum PFS wurde die Zeit bis zur Progression (TTP) erfasst. Diese war definiert als Zeit vom 1. Zyklus der Studienbehandlung bis zur MRT-Untersuchung, in der eine objektive Progression festgestellt wurde, unabhängig davon, ob die Patientin / der Patient die Studientherapie abbrach oder eine andere NF1-PN-Behandlung (nach Absetzen der Studienbehandlung) vor der Progression erhielt.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

PFS stellt eine Kombination aus Mortalitäts- und Morbiditätsendpunkten dar. Tod, ungeachtet der zugrunde liegenden Todesursache, ist ein Teilaspekt, der im Endpunkt „Todesfälle“ abgebildet ist; Gesamtmortalität wird als patientenrelevant erachtet. Ein Progress festgestellt durch bildgebende Verfahren ohne Berücksichtigung der für die Patientin / den Patienten spürbaren Symptomatik wird als nicht patientenrelevant erachtet. Der Endpunkt „Progressionsfreies Überleben“ wird insgesamt nicht als patientenrelevant bewertet. TTP bildet nur den Aspekt des Progresses ab und ist ebenfalls als nicht patientenrelevant anzusehen.

Validität

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

Globalbeurteilung der klinischen Veränderung

Der Endpunkt „Globalbeurteilung der klinischen Veränderung“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Globalbeurteilung der klinischen Veränderung wurde mittels Global Impression of Change (GIC) erfasst. Der GIC ist eine 1-Item-Skala zur Bewertung der klinischen Bedeutung von Veränderungen der Schmerzintensität oder anderer Erkrankungen. In der SPRINT-Studie wurde eine angepasste Version des GIC mit 3 Items verwendet. Gemessen wurde die Veränderung der Tumorschmerzen, der Gesamtschmerzen sowie der PN-bedingten Morbidität gegenüber dem Zeitpunkt vor Einnahme der Studienmedikation. Die Bewertung erfolgte auf einer 7-Punkte-Skala: 1 = sehr viel besser; 2 = viel besser; 3 = wenig besser; 4 = keine Veränderung; 5 = wenig schlechter; 6 = viel schlechter; 7 = sehr viel schlechter.

Bei Kindern ab 8 Jahren erfolgte eine Selbsteinschätzung, bei Kindern ≥ 5 Jahre erfolgte eine Fremdeinschätzung durch die Eltern/Betreuer.

Responder mit einer bedeutenden Verbesserung im GIC waren definiert als die Antwortkategorien "sehr viel verbessert" und "viel verbessert".

Die Erhebung des GIC erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung zum Therapieende und in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung ist nachvollziehbar. Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die klinische Veränderung der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die patientenberichtete Erfassung der Globalbeurteilung der klinischen Veränderung wird als patientenrelevant erachtet. Eine durch die Eltern berichtete klinische Veränderung kann für eine bestimmte Patientenpopulation als patientenrelevant erachtet werden, z. B. für jüngere oder kognitiv beeinträchtigte Kinder, die nicht in der Lage sind, den Fragebogen selbst auszufüllen. Dies trifft in der vorliegenden Studie nur auf eine Teilpopulation zu.

Validität

Der GIC dient generell zur Einschätzung des allgemeinen Funktionsstatus einer Person. Bei der vorliegenden Version des Instruments handelt es sich um eine angepasste Version, mit der eine Bewertung der klinischen Bedeutung von Veränderungen der Tumorschmerzen, Gesamtschmerzen sowie PN-bedingter Morbidität durchgeführt wird.

Die GIC-Skala gilt als standardisiertes, etabliertes Instrument. Die Skala wird von der Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) als zentrales Ergebnismaß zur globalen Verbesserung unter Behandlung für klinische Studien mit chronischen Schmerzen empfohlen und wurde in mehreren Forschungsstudien zu chronischen Schmerzen eingesetzt [15]. Angaben zur Validierung der Skala als krankheitsspezifisches, schmerz-bewertendes Instrument im Allgemeinen und für Kinder im Speziellen liegen nicht vor.

Darüber hinaus wird die Erfassung von Veränderungen mittels GIC in der Forschung kritisch diskutiert, da die Befragten bei der rückwirkenden Einschätzung einer Veränderung in der Lage sein müssen, sowohl ihren gegenwärtigen Zustand als auch ihren Ausgangszustand zu bewerten und einen Unterschied im Sinne einer Verbesserung, Verschlechterung oder keiner Veränderung zu quantifizieren. Größtenteils werde bei dieser Abschätzung allerdings der gegenwertige Gesundheitszustand erfasst. Dieser Effekt sei umso größer, je länger das Erinnerungsintervall ist. Zudem seien Erkrankungen mit hoher Symptomvariabilität besonders anfällig für diese Form der Verzerrung (Recall Bias) [23,26,39,40]. Diese Limitationen erscheinen auch für die vorliegende Erhebung relevant und sind bei der Interpretation der Werte zu berücksichtigen.

Die Responderdefinition von Verbesserung („sehr viel besser“ und „viel besser“) wird als adäquat angesehen. Eine Responderdefinition, welche die Verschlechterung im GIC abbildet, liegt nicht vor.

Schmerz

Numerical Rating Scale für Schmerz

Der Endpunkt „Numerical Rating Scale“ für Schmerz wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Numerical Rating Scale (NRS-11) ist eine 11-stufige numerische Ratingskala zur Selbstbeurteilung der Intensität der auftretenden Schmerzen. Kinder ab einem Alter von 8 Jahren beurteilten die Intensität ihrer schlimmsten Schmerzen in der letzten Woche auf der NRS-11 von 0 „keine Schmerzen“ bis 10 „schlimmste vorstellbare Schmerzen“. Laut SAP wurden insgesamt 4 Fragen zur Schmerzintensität mittels NRS-11 erhoben. Es wurde der Schmerz bezogen auf die durch das ärztliche Prüfpersonal festgelegte Zielläsion (nur Version 2) sowie für Versionen 1 und 2 einer selbst durch die Patientin / den Patienten gewähltes wichtiges PN, der Gesamttumorschmerz und andere Schmerzen (als Beispiel werden hier Kopf- oder Rückenschmerzen genannt) bewertet. Von den Patientinnen und Patienten, die bei der Baseline-Untersuchung Version 1 des Fragebogens ausfüllten, wurden nur jene in die Ergebnisanalyse bezüglich des Schmerzes der Zielläsion einbezogen, deren selbst-gewähltes PN mit der durch das ärztliche Prüfpersonal bestimmten Zielläsion übereinstimmte.

Es wurden Responderanalysen mit einer in SAP 5.0 festgelegten Responderschwelle von 2 Punkten durchgeführt. Zusätzlich wird für das Dossier eine post hoc definierte Analyse für eine Verbesserung um mindestens 2 Punkte im gesamten Studienverlauf dargestellt.

Zudem wurden 2 Definitionen für Schmerzlinderung formuliert:

1. Für Personen mit einem Score $\geq X$ Punkte zu Baseline: Schmerzlinderung entspricht einer Abnahme von $\geq X$ Punkten ohne Erhöhung des Schmerzmittel-Scores.
2. Für Personen mit einem Score $\geq X$ Punkte zu Baseline: Schmerzlinderung entspricht einer Abnahme von $\geq X$ Punkten ohne Erhöhung des Schmerzmittel-Scores; ODER
Für Personen mit einem Score $< X$ Punkte und einem Schmerzmittel-Score ≥ 2 Punkte zu Baseline: Schmerzlinderung entspricht einem stabilen oder verminderten NRS-11-Score bei gleichzeitiger Abnahme des Schmerzmittel-Scores um eine Stufe.

X soll in beiden Definitionen einem Wert von 2, 1 oder des in der Studie abgeleiteten Clinically Meaningfully Threshold (CMT) entsprechen.

Der Schmerzmittel-Score war in Anlehnung an das Stufenschema der Weltgesundheitsorganisation (WHO) zum Schmerzmitteleinsatz folgendermaßen definiert:

- Score 0: Kein Schmerzmittel
- Score 1: Nicht-Opioidanalgetika
- Score 2: Niederpotente Opioidanalgetika
- Score 3: Hochpotente Opioidanalgetika und neuropathische Schmerzmittel

Der Schmerzmittel-Score wurde anhand des Schmerzmitteleinsatz in den letzten 7 Tagen vor Erhebung der NRS-11 gebildet. Bei der Einnahme mehrerer Schmerzmedikamente wurde der höchste Score-Wert herangezogen.

Als Zeit bis zur Schmerzlinderung wurde der Zeitraum von der ersten Einnahme der Studienmedikation bis zur ersten beobachteten Schmerzlinderung berechnet. Personen ohne Schmerzminderung wurden zum Zeitpunkt der letzten NRS-11-Erhebung zensiert. Bei Definition 1 wurden zudem Teilnehmende mit einem Schmerzmittel-Score < 2 zu Baseline zensiert.

Die Erhebung des Schmerzes erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung ist größtenteils nachvollziehbar. Aufgrund des Bezugszeitraums von 7 Tagen ermöglicht die Operationalisierung keine Aussagen darüber, inwieweit eine dauerhafte Schmerzlinderung erreicht wird. Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht den Schmerz der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Das Auftreten von Schmerzen und deren Intensität ist patientenrelevant.

Validität

Eine NRS ist ein etabliertes Verfahren zur Messung der Symptomatik wie bspw. Schmerz bei Erwachsenen. Untersuchungen bei Kindern zeigten, dass die NRS ebenfalls gute psychometrische Eigenschaften [9,34] aufweist. Eine gute Test-Retest-Reliabilität und Änderungssensitivität zeigte sich bei der mündlichen Administration der NRS bei Kindern im Alter von 8 bis 17 Jahren [9]. Eine Arbeitsgruppe der REiNS International Collaboration empfiehlt den Einsatz der NRS-11 bei Kindern mit NF1 ab 8 Jahren [55]. Die NRS wird insgesamt als grundsätzlich reliabel und valide zur Messung der Schmerzintensität bewertet. Innerhalb der SPRINT-Studie wurde mittels NRS der auf die Zielläsion bezogene Schmerz, der Schmerz einer durch die Patientin / den Patienten selbst gewählten wichtigen PN, der Gesamt tumorschmerz und anderer Schmerz erfragt. Für diese unterschiedlichen Aspekte des Schmerzes liegen keine separaten Validierungen (für Kinder und Jugendliche mit NF1) vor und es ist unklar, ob die Teilnehmenden die unterschiedlichen Aspekte des Schmerzes voneinander differenzieren können. Der auf die Zielläsion bezogene Schmerz wurde je nach Fragebogen-

version mit unterschiedlichen Fragen erhoben. Zudem liegen für Teilnehmende, die in Fragebogen-version 1 eine andere als die Zielläsion zur Bewertung der Schmerzintensität auswählten, keine Angaben vor. Dies kann zu Einschränkungen der Validität führen. Die Responderschwelle von 2 Punkten (entsprechend > 15 % der Skalenspannweite) wird als adäquat angesehen. Es werden die präspezifizierten Analysen für die Nutzenbewertung herangezogen.

Pain Interference Index

Der Endpunkt „Pain Interference Index“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Validität.

Operationalisierung

Beschreibung

Eine durch Schmerzen hervorgerufene Beeinträchtigung wurde mittels Pain Interference Index (PII) erhoben. Kinder und Jugendliche ab 8 Jahren beantworteten den PII selbst. Eltern von Kindern und Jugendlichen im Alter von 5 bis 18 Jahren füllten die Eltern-/Betreuerversion (Pain Interference Index – Parent-Report (PII-P)) aus. In beiden Versionen (PII, PII-P) wird anhand von 6 Items erfragt, in welchem Umfang vorhandene Schmerzen innerhalb der letzten Woche einen Einfluss auf Schule, Freizeitaktivitäten, Freunde treffen, Laune, körperliche Aktivitäten und Schlafen hatten. Die Befragten beantworten die Items auf einer 7-Punkte-Likert-Skala von 0 (gar nicht) bis 6 (komplett). Es wird ein Gesamtwert als Mittelwert über alle Items gebildet.

Es sollten Responderanalysen für den PII mit einer aus der Studie abgeleiteten Responderschwelle (ankerbasiert: anhand GIC für Tumorschmerz und andere Schmerzen; verteilungsbasiert: halbe Standardabweichung der Baseline-Werte) für Verbesserungen, Verschlechterungen und keine Änderungen durchgeführt werden. Als Resonderschwelle wurde 0,75 für den PII (entspricht 12,5 % der Skalenspannweite) und 1,78 für den PII-P (entspricht 29,7 % der Skalenspannweite) im Studienbericht verwendet. Im Dossier werden Post-hoc-Responderanalysen mit einer Responderschwelle von 15 % der Skalenspannweite (0,9 Punkte) dargestellt.

Bewertung

Die Operationalisierung ist größtenteils nachvollziehbar.

Die in der Studie SPRINT eingesetzte Version des PII/PII-P unterscheidet sich zum Teil von der Version, welche in der Literatur beschrieben und in den Validierungsstudien eingesetzt wurde. Zum einen bestehen Unterschiede im Bezugszeitraum, der bei 7 Tagen in der Studie SPRINT und bei 2 Wochen in der publizierten Version liegt. Weiterhin sind einige Fragen anders formuliert als in der publizierten Version; z. B. Studie SPRINT: 1. Has your pain made it hard for you to pay attention (for example, do classwork, homework, read)?, 6. Has your pain made it hard for you to sleep? vs. Martin et al (2015) [29]: Has Your Pain: 1. Made it difficult for you to do schoolwork?, 6. Affected your sleep?

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die durch Schmerzen hervorgerufene Beeinträchtigung der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Das Auftreten von Schmerzen und der Einfluss von Schmerzen auf Alltagsaktivitäten berichtet durch die Betroffenen ist patientenrelevant.

Validität

Der PII wurde ursprünglich in Schweden entwickelt. Grundlage für die Entwicklung waren die „Interference Scale“ des Multidimensional Pain Inventory (MPI) und die „Pain Interference Items“ des Brief Pain Inventory (BPI) [53]; diese wurden auf die Bedürfnisse von Jugendlichen angepasst. Ausführliche Angaben zur Fragebogenentwicklung und inwieweit Betroffene in die Entwicklung involviert waren liegen nicht vor. Die einzelnen Items erscheinen augenscheinlich inhaltsvalide. Der Fragebogen wurde ins Englische übersetzt und hinsichtlich interner Konsistenz und Konstruktvalidität in 60 Personen im Alter von 6 bis 24 Jahren, davon 31 mit NF1, untersucht [29]. Die Analyse zeigte eine hohe interne Konsistenz (Cronbach's alpha: 0,84 (Patientinnen und Patienten); 0,94 (Mütter); 0,96 (Väter)) [29]. Zudem wurden moderate Korrelationen der Patientenversion mit der Eltern-/Betreuerversion ($r = 0,43-0,48$), hohe Korrelationen des PII mit dem Modified Brief Pain Inventory (MBPI) ($r = 0,81$) und einer visuellen Analogskala für Schmerzen ($r = 0,54$) gezeigt. Die Werte der Patientenversion waren insgesamt niedriger als die der Eltern-/Betreuerversion, was vorwiegend auf höhere Werte bezüglich des Einflusses der Schmerzen auf die Laune und den Schlaf bei der Bewertung durch die Eltern zurückzuführen war [29]. Known-Groups-Validität wurde anhand von Unterschieden im PII bzw. PII-P zwischen Personen gezeigt, welche verschriebene Schmerzmittel einnahmen und denen die diese nicht einnahmen [29]. Personen mit milder NF1-Erkrankung vs. moderater bzw. schwerer Erkrankung hatten einen signifikant höheren PII [29]. PII-P war ebenfalls höher für Personen mit moderater bzw. schwerer Erkrankung im Vergleich zu denen mit einer milden Erkrankung, allerdings war der Unterschied statistisch nicht signifikant [29]. Die Untersuchungen weisen insgesamt auf eine ausreichende Konstruktvalidität hin. Longitudinale Untersuchungen bezüglich Test-Retest-Reliabilität und Änderungssensitivität wurden vom pU nicht vorgelegt und konnten auch nicht identifiziert werden. Eine kleine Interventionsstudie ($n = 14$) zeigt eine signifikante Abnahme des PII-P, aber nicht der selbstberichteten PII durch eine Akzeptanz- und Commitment-Therapie [30]. Dies kann als erster Hinweis auf Änderungssensitivität für die Eltern-/Betreuerversion des PII gewertet werden, ist zum Nachweis allerdings nicht ausreichend. Wie oben beschrieben unterscheidet sich die in den Validierungsstudien eingesetzte Version des PII von der in der SPRINT-Studie eingesetzten hinsichtlich Bezugszeitraum und Formulierung einiger Fragen. Dies kann zu anderen psychometrischen Eigenschaften führen, als den in den Validierungsstudien. In der Gesamtschau ist die Validität der in der SPRINT-Studie eingesetzten Version des PII nicht ausreichend belegt, um für die Nutzenbewertung herangezogen werden zu können.

Motorfunktion

Manueller Muskeltest

Der Endpunkt „Manueller Muskeltest“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Validität.

Operationalisierung

Beschreibung

Bei allen Patientinnen und Patienten mit PN, die eine motorische Dysfunktion, Schwäche oder eine Kompression des Rückenmarks verursachen oder verursachen können, sollte die Kraft aller Muskelgruppen untersucht werden. Diese Bewertung/Messung sollte nach Möglichkeit durch ein Mitglied der Rehabilitationsmedizin, der Physiotherapie oder der Physikalischen Therapie durchgeführt und idealerweise bei jedem Besuch von der-/demselben Untersuchenden vorgenommen werden.

Die Kraft der Muskelgruppen wurde mittels manueller Muskeltests (MMT) nach der Methodik des Medical Research Councils (MRC) [33] erhoben. Jede durch ein PN beeinträchtigte Muskelgruppe wurde hierbei, basierend auf der anatomischen Lage des PN, einem Körperquadranten der Patientin / des Patienten zugewiesen:

- Obere oder untere Extremitäten
- Unilateral oder Bilateral

Der Gesamtwert des MMT entspricht dem Mittelwert der Kraft sämtlicher Muskeln, die im gleichen Körperquadranten wie das PN liegen. Anhand der Skala des MRC wird nach einem festgelegten Protokoll die jeweilige Muskelkraft auf einer 6-stufigen Skala bewertet [33].

- 0 – Keine Muskelkontraktion
- 1 – Sicht- oder tastbare Kontraktion
- 2 – Bewegung ohne Einfluss der Schwerkraft möglich
- 3 – Bewegung gegen Schwerkraft möglich
- 4 – Bewegung gegen Schwerkraft und Widerstand möglich
- 5 – Normale Muskelkraft

Die Skalenwerte konnten durch Verwendung eines \pm um 0,3 Punkte auf- bzw. abgewertet werden. Insgesamt konnten 17 Muskeln bzw. Muskelgruppen in sitzender oder liegender (Bauch, Rücken, Seite) Position untersucht werden. Zu Baseline sollten alle Muskeln untersucht werden und in den Follow-up-Untersuchungen nur die betroffenen bzw. potentiell betroffenen Muskelgruppen. Zum Vergleich sollte auch die kontralaterale Seite untersucht werden. Wenn die kontralaterale Seite auch potentiell betroffen war, sollte eine dritte Extremität getestet werden, um die Gesamtstärke zu bestimmen. Die Muskeln, welche in den Follow-up-Untersuchungen evaluiert werden sollten, wurden durch das Studienteam basierend auf der Lokalisation des PN und den Ergebnissen der Baseline-Untersuchung festgelegt.

Bewertung

Zur Durchführung des Tests wird im Studienprotokoll lediglich auf die Originalpublikation des MRC verwiesen. Die exakte Testdurchführung innerhalb der Studie wird nicht näher beschrieben, sodass nicht davon ausgegangen werden kann, dass eine Schulung der Untersuchenden durchgeführt wurde. Die Auf- bzw. Abwertung des Scores um \pm 0,3 Punkte, ändert die Skala im Vergleich zu Literatur.

Weiterhin ist zu berücksichtigen, dass der Test nur bei einem Teil der Patientenpopulation (Personen mit PN, die eine motorische Dysfunktion, Schwäche oder eine Kompression des Rückenmarks verursachen oder verursachen können) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie ein PN in diesem Bereich entwickelte wurden somit nicht erfasst.

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Muskelkraft der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Der MMT erscheint patientenrelevant bezüglich der Verbesserung des Gesundheitszustands.

Validität

Der MMT wurde entwickelt, um Untersuchungen zu Nervenverletzungen bei Kriegsversehrten durchzuführen [33]. Der pU legt als Nachweis der Validität und Reliabilität die Publikation einer Studie bei erwachsenen, kritisch kranken Patientinnen und Patienten auf Intensivstationen vor [11]. Die Übertragbarkeit der Studienergebnisse auf Kinder und Jugendliche mit NF1 wird angezweifelt. Untersuchungen bei Kindern und Jugendlichen (n = 33) auf einer Kinderintensivstation zeigten, dass die Untersuchung bei Kindern nicht immer möglich sind, z. B. bei jüngeren Kindern, aufgrund kognitiver Einschränkungen, fehlender Kooperation oder Schmerzen [42]. Eine Untersuchung an 12 Kindern mit Duchenne-Muskeldystrophie kam zu dem Schluss, dass regelmäßige Schulungen der Untersuchenden notwendig sind, um eine ausreichende Reliabilität (ICC > 0,75) zu erhalten [16]. In der oben genannten Studie auf der Kinderintensivstation wurden Schulungen der Untersuchenden durchgeführt und gute Kennwerte für die Test-Retest-Reliabilität erhalten [42]. Aus den Studienunterlagen geht nicht hervor, dass eine solche Schulung innerhalb der Studie erfolgte. Es ist daher unklar, inwieweit die Muskelkraft in der Studie mittels MMT objektiv erfasst werden konnte. Innerhalb der Studie sollte möglichst die/derselbe Untersuchende die Teilnehmenden zu allen Zeitpunkten untersuchen. Dadurch werden Unterschiede aufgrund unterschiedlicher Testdurchführung innerhalb eines Individuums minimiert. Zwischen den Individuen können bei unterschiedlichen Untersuchenden große Variationen auftreten. In den Untersuchungen zur Reliabilität wurde jedoch nie die in der SPRINT-Studie verwendete Skala mit Auf- und Abwertung jeder Stufe um 0,3 verwendet, bei der durch eine feinere Skalierung eine reliable Erhebung erschwert wird. Die Übertragbarkeit der Studienergebnisse bezüglich Reliabilität ist somit fraglich. Insgesamt bestehen Zweifel, inwieweit in der Studie SPRINT eine reliable Erhebung des MMT erfolgte. Die Ergebnisse des MMT werden daher nicht für die Nutzenbewertung herangezogen.

Greifkraft und Kraft des Schlüsselgriffs

Der Endpunkt „Greifkraft und Kraft des Schlüsselgriffs“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund fehlender Angaben zur Testdurchführung.

Operationalisierung

Beschreibung

Aus dem Studienprotokoll geht hervor, dass Greifkraft und Kraft des Schlüsselgriffs (grip and (key) pinch strength) in Kilogramm bzw. Pounds bestimmt werden sollte. Wie die Bestimmung erfolgen sollte, geht aus dem Studienprotokoll nicht hervor. Im Dossier wird beschrieben, dass die Bestimmung mittels Dynamometer erfolgte. Weitere Angaben liegen nicht vor.

Die Greifkraft und die Kraft des Schlüsselgriffs wurde bei Kindern und Jugendlichen mit PN-assoziiert Morbidität der Motorfunktion gemessen. Ob eine PN-assoziierte Morbidität der Motorfunktion vorlag, entschied das ärztliche Prüfpersonal bei Studienbeginn auf Basis der Lage der PN und der vorhandenen Symptome.

Die Analyse erfolgt auf Basis von Veränderungen gegenüber Baseline in Messeinheit.

Bewertung

Die Operationalisierung ist nicht nachvollziehbar, da keine Angaben zur Messmethodik vorliegen. Es ist zu berücksichtigen, dass der Test nur bei einem Teil der Patientenpopulation (Personen mit PN-assoziiert Morbidität der Motorfunktion) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie ein PN in diesem Bereich entwickelte, wurden somit nicht erfasst.

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Greifkraft und Kraft des Schlüsselgriffs der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Einschränkungen der Greifkraft und der Kraft des Schlüsselgriffs können zu erheblichen Einschränkungen der Handmotorik führen. Der Endpunkt wird als patientenrelevant erachtet.

Validität

Die Validität kann nicht beurteilt werden, da keine Angaben zur Durchführung der Messung vorliegen. Generell ist bei ausreichender Standardisierung eine reliable Messung möglich [32]. Die Greifkraft steigt bei Kindern mit zunehmenden Alter an [32], sodass die Beurteilung der Messwerte ohne Bezug zu altersstandardisierten Normwerten, wie in der SPRINT-Studie durchgeführt, keine verlässlichen Aussagen bezüglich der Wirksamkeit der Intervention in einer unkontrollierten Studie zulässt. Der Endpunkt wird daher nicht zur Nutzenbewertung herangezogen.

Bewegungsumfang der Gelenke

Der Endpunkt „Bewegungsumfang der Gelenke“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund fehlender Angaben zur Testdurchführung.

Operationalisierung

Beschreibung

Bei allen Personen mit PN, die eine motorische Dysfunktion, Schwäche oder eine Kompression des Rückenmarks verursachen oder verursachen können, sollte der Bewegungsumfang aller Gelenke untersucht werden. Diese Bewertung/Messung sollte nach Möglichkeit durch ein Mitglied der Rehabilitationsmedizin, der Physiotherapie oder der Physikalischen Therapie durchgeführt und idealerweise bei jedem Besuch von der-/demselben Untersuchenden vorgenommen werden.

Zu Baseline sollten alle festgelegten Gelenke/Bewegungen untersucht werden und in den Follow-up-Untersuchungen nur die betroffenen bzw. potentiell betroffenen Gelenke. Für alle Gelenke/Bewegungen wurde der Bewegungsumfang in Grad angegeben. Zum Vergleich sollte auch die kontralaterale Seite untersucht werden. Die Gelenke, welche in den Follow-up-Untersuchungen evaluiert werden sollten, wurden durch das Studienteam basierend auf der Lokalisation des PN und den Ergebnissen der Baseline-Untersuchung festgelegt und folgendermaßen kategorisiert: Hals / Obere Extremitäten / Untere Extremitäten / Andere.

Angaben zur Durchführung und Standardisierung der Untersuchung des Bewegungsumfangs konnten in den Studienunterlagen nicht identifiziert werden.

Bewertung

Die Operationalisierung ist nicht nachvollziehbar, da keine Angaben zur Durchführung der Messung vorliegen.

Es ist zu berücksichtigen, dass die Untersuchung nur bei einem Teil der Patientenpopulation (Personen mit PN, die eine motorische Dysfunktion, Schwäche oder eine Kompression des Rückenmarks verursachen oder verursachen können) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie ein PN in diesem Bereich entwickelten wurden somit nicht erfasst.

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht den Bewegungsumfang der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Starke Einschränkungen des Bewegungsumfanges einzelner Gelenke können zu Funktionseinschränkung führen. Der Endpunkt wird daher als patientenrelevant erachtet.

Validität

Die Validität kann nicht beurteilt werden, da keine Angaben zur Durchführung der Messung vorliegen. Ohne Standardisierung kann nicht von Reliabilität ausgegangen werden und eine reliable Messung ist zur validen Erhebung notwendig [20].

Unterschied in der Beinlänge

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

Operationalisierung

Beschreibung

Bei Personen mit PN im lumbosakralen Plexus oder darunter sollte die Beinlänge jedes Beins gemessen werden, um die Unterschiede in der Beinlänge zu untersuchen. Die Messungen erfolgten liegend mit Hüft- und Knieextension von der Spina iliaca anterior superior zum Malleolus medialis des Knöchels. Zudem wurde erfasst, ob bei Vorliegen eines Unterschieds in der Beinlänge ein Beckenschiefstand diagnostizierbar war. Die Messungen sollten zu den Studienvisiten idealerweise durch dieselben Personen durchgeführt werden.

Die Auswertung erfolgt auf Basis der durchschnittlichen Änderung im Vergleich zu Baseline.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Es ist zu berücksichtigen, dass der Test nur bei einem Teil der Patientenpopulation (Personen mit PN im lumbosakralen Plexus oder darunter) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie ein PN in diesem Bereich entwickelte, wurden somit nicht erfasst.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Eine ausgeprägte Diskrepanz in der Beinlänge kann zu einem Beckenschiefstand sowie einer Ausgleichsskoliose der Wirbelsäule führen und mit Schmerzen, Verspannungen und einem erhöhten Verschleiß einhergehen. Geringe Diskrepanzen sind dagegen ohne nennenswerten Krankheitswert. In der vorliegenden Operationalisierung kann zwischen morbiditätsrelevanten und nicht-relevanten Veränderungen nicht unterschieden werden. Die Veränderung des Unterschieds in der Beinlänge wird daher nicht als per se patientenrelevant erachtet.

Validität

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

Grooved Pegboard Test

Der Endpunkt „Grooved Pegboard Test“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Alle Kinder und Jugendlichen ab einem Alter von 5 Jahren mit einem PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks sollten sich dem Grooved Pegboard Test unterziehen.

Bei diesem Test müssen 25 schlüsselartige Stifte (pegs) in Löcher mit zufällig angeordneten Schlitzern auf der Stecktafel eingesetzt werden. Die Stifte müssen hierbei gedreht werden, damit sie in das Loch passen. Erhoben wird die Zeit bis zur Vervollständigung der Stecktafel. Kinder im Alter von 5 bis 8 Jahren mussten nur die ersten beiden Reihen der Stecktafel (entsprechend 10 Stifte) vervollständigen. Zur Auswertung wurde ein anhand altersabhängiger Normwerte aus dem Manual (Mittelwert und Standardabweichung) standardisierter Z-Score herangezogen sowie die Anzahl der fallengelassenen Stifte. Es wurden Veränderungen dieser Variablen im Vergleich zu Baseline berechnet. Der Test wurde jeweils einmal mit der dominanten Hand und einmal mit der nicht-dominanten Hand durchgeführt. Zusätzlich zu der Unterscheidung nach der Händigkeit wurde auch jeweils die Lokalisation des PN berücksichtigt und die Ergebnisse jeweils einmal für die Hand auf der Seite des PN sowie für die Hand auf der anderen Seite dargestellt. Falls eine bilaterale PN-Beeinträchtigung vorhanden war, wurde zwischen der mehr beeinträchtigten und der weniger beeinträchtigten Hand unterschieden.

Im Studienprotokoll sind standardisierte Anweisungen für die Untersuchenden formuliert.

Die Erhebung des Grooved Pegboard Tests erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Es ist zu berücksichtigen, dass der Test nur bei einem Teil der Patientenpopulation (Personen mit PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie ein PN in diesem Bereich entwickelte, wurden somit nicht erfasst. Die Erhebung erfolgte weiterhin nur inner-

halb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht den Grooved Pegboard Test der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Der Grooved Pegboard Test misst die manuelle Geschicklichkeit und die Auge-Hand-Koordination und wird daher als patientenrelevant angesehen.

Validität

Eine Untersuchung der Test-Retest-Reliabilität des Grooved Pegboard Tests in 53 Personen (Alter 3 bis 85 Jahre) im Abstand von 2 Wochen zeigte eine moderate bis hohe Test-Retest-Reliabilität (ICC [95%-KI] = 0,91 [0,84; 0,95] für die rechte bzw. 0,85 [0,81; 0,89] für die linke Hand) [51]. Weiterhin wurden im Rahmen von Untersuchungen zur Known-Groups-Validität Unterschiede zwischen Altersgruppen, Fingerfertigkeit gemäß Bruininks-Oseretsky-Test-Klassifizierung sowie Purdue Pegboard Test [51] festgestellt. Prinzipiell erscheint eine reliable und valide Messung bei Durchführung gemäß Manual und Studienprotokoll möglich, da ein hoher Grad an Standardisierung der Messung erfolgt.

Anhand welcher Normwerte die Berechnung der Z-Scores erfolgte wird in Studienprotokoll und SAP nicht beschrieben. Laut Angaben in Modul 4 wurden altersabhängige Normwerte aus dem Handbuch für die Berechnung der Z-Scores herangezogen. Im vorgelegten Handbuch [27] werden jedoch ausschließlich alters- und geschlechtsspezifische Normwerte für Kinder unter 9 Jahren gelistet. Für Kinder und Jugendliche ab 9 Jahren liegen sowohl alters- und geschlechtsspezifische Normwerte als auch ausschließlich altersspezifische Normwerte vor. Es ist daher unklar, welche Normwerte für die Auswertung tatsächlich verwendet wurden. Angaben zur Normstichprobe wie bspw. Stichprobengröße und Gesundheitszustand konnten nicht identifiziert werden. Die Repräsentativität der Normstichprobe kann daher nicht beurteilt werden.

PROMIS – Körperliche Funktionsfähigkeit Mobilität / Obere Extremitäten

Der Endpunkt „PROMIS – Körperliche Funktionsfähigkeit Mobilität / Obere Extremitäten“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Erhebung der Fragebögen des Patient-Reported Outcomes Measurement Information System (PROMIS) für körperliche Funktionsfähigkeit erfolgte in allen Personen mit PN, welche die Motorfunktion beeinträchtigten.

PROMIS ist ein System, das aus domänenspezifischen Instrumenten zur selbst- und fremdberichteten Beurteilung des Wohlbefindens besteht. Der domänenspezifische Ansatz basiert auf der Idee, dass Gesundheitsmerkmale wie Schmerz und körperliche Funktion nicht auf eine bestimmte Krankheit beschränkt sind. Die PROMIS-Fragebögen stellen somit generische Instrumente dar. Die Domänen sind als Itembänke organisiert, aus denen individuell Items ausgewählt werden können. Die pädiatrischen Itembänke (für Kinder im Alter von 5 bis 17 Jahren) umfassen generische Instrumente zur Erfassung von Funktionen, Symptomen, Verhaltensweisen und Gefühlen. Es stehen

sowohl Instrumente zur Selbsteinschätzung (für Kinder ab 8 Jahren) als auch zur Fremdbeurteilung (Eltern-/Betreuerversion für Kinder zwischen 5 und 17 Jahren) zur Verfügung.

Die PROMIS-Entwicklungsgruppe schlägt für jede Itembank eine Kurzform vor. Innerhalb der SPRINT-Studie erfolgte die Erhebung der pädiatrischen PROMIS-Itembank „Körperliche Funktionsfähigkeit“ und den dafür erstellten Kurzformen (jeweils 8a, Version 1.0) für die Dimension „Mobilität“ und „Obere Extremitäten“. Kinder im Alter von 8 bis 18 Jahren füllten eigene Fragebögen aus. Eltern von Kindern im Alter von 5 bis 18 Jahren füllten die Eltern-/Betreuerversion aus. Die Kurzform „Mobilität“ bezieht sich auf einfache Bewegung (bspw. Aufstehen, Beine bewegen, auf Zehenspitzen stehen, Treppensteigen ohne Festhalten) und fragt nach der Möglichkeit Sport mit gleichaltrigen Kindern zu treiben, dabei mithalten zu können und körperliche Aktivitäten auszuführen, die man möchte. Die Beantwortung erfolgte mithilfe einer 5-Punkte-Likert-Skala (4–0 Punkte) je Item: ohne Schwierigkeiten / mit etwas Schwierigkeiten / mit einigen Schwierigkeiten / mit viel Schwierigkeiten / nicht in der Lage, es zu tun. Zudem wird nach der Verwendung von Gehhilfen und eines Rollstuhls mit folgenden Antwortmöglichkeiten gefragt: keine Nutzung / ein bisschen / manchmal / meistens / immer. Die Kurzform „Obere Extremitäten“ erfragt unter anderem die Möglichkeiten Schuhe und T-Shirt anzuziehen, Knöpfe zu schließen, schwere Türen zu öffnen, aufzuschließen, mit Stift oder Bleistift zu schreiben. Auch hier erfolgt die Beantwortung auf der oben beschriebenen 5-Punkte-Likert Skala. Die Eltern-/Betreuerversionen erfragen die gleichen Aktivitäten jeweils in Bezug auf das eigene Kind. Der Bezugszeitraum umfasst jeweils die letzten 7 Tage.

Das Scoring wird in Studienprotokoll und SAP nicht beschrieben. Im Manual wird das „Response Pattern Scoring“ als genaueste Methode des Scorings beschrieben, welches Scores automatisch nach Eingabe in ein Berechnungsprogramm ausgibt und dabei die Scores auf Item-Ebene berechnet und besser mit fehlenden Werten umgehen kann. Daneben ist auch ein Scoring basierend auf Berechnung eines rohen Summenscores und die Umwandlung des Scores in T-Scores gemäß publizierter Tabellen möglich. Diese Scoreberechnung ist nur möglich, wenn keine fehlenden Werte auftreten. Gemäß dieser Tabelle weisen sowohl der Mobilitätsscore als auch der Score für die oberen Extremitäten in Version 1.0 einen Wertebereich der Rohwerte von 0 bis 32 für die Eltern-/Betreuerversion und ebenfalls 0–32 Punkte für die Kinderversion der Kurzform „Mobilität“ auf. Die Kinderversion der Kurzform für „Obere Extremitäten“ weist gemäß Manual einen Wertebereich von 0 bis 30 Punkten auf. Wie sich diese maximale Gesamtpunktzahl aus den im Protokoll abgedruckten Fragebögen mit 9 bzw. 10 Fragen ergibt, kann nach Aufaddieren der maximal erreichbaren Punktzahl nicht nachvollzogen werden. Es werden sowohl Rohwerte als auch T-Scores mit einem Mittelwert von 50 und einer Standardabweichung von 10 in der SPRINT-Studie berechnet. Als Referenzpopulation wird eine US-amerikanische Allgemeinbevölkerung verwendet. Höhere Werte beschreiben eine bessere körperliche Funktion. Für Respondernanalysen soll laut SAP eine Responderschwelle gemäß Thissen et al. (2016) [44] auf Basis der T-Scores verwendet werden. Daneben soll eine MID (Minimal Important Difference) innerhalb der SPRINT-Studie mit dem GIC für tumor-bezogene Morbidität abgeleitet und verwendet werden. Im Dossier wird post hoc die Anzahl der Patientinnen und Patienten mit einer Verbesserung im Rohwert des PROMIS im Studienverlauf im Vergleich zu Baseline um $\geq 4,8$ Punkte (entsprechend 15 % der Skalenspannweite von 0 bis 32 Punkten) dargestellt. Im SAP waren ebenfalls Analysen kontinuierlicher Daten mittels „Mixed Effect Model Repeat Measures“ (MMRM) geplant.

Die Erhebung der PROMIS-Skalen erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Das Scoring der Fragebögen ist nicht nachvollziehbar beschrieben. Es besteht eine Diskrepanz zwischen der sich aus den im Protokoll abgedruckten Fragebögen ergebenden erreichbaren Maximalpunktzahl für die Rohwerte und der im Manual angegeben.

Weiterhin ist zu berücksichtigen, dass der Fragebogen nur bei einem Teil der Patientenpopulation (Personen mit PN-bezogenen Störungen der Motorfunktion zu Studienbeginn) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie eine Störung der Motorfunktion entwickelte, wurden somit nicht erfasst. Die Erhebung erfolgte weiterhin nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die körperliche Funktionsfähigkeit der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die patientenberichtete Erfassung der körperlichen Funktionsfähigkeit mittels PROMIS wird als patientenrelevant erachtet. Ein durch die Eltern berichtete körperliche Funktionsfähigkeit mittels PROMIS kann für eine bestimmte Patientenpopulation als patientenrelevant erachtet werden, z. B. für jüngere oder kognitiv beeinträchtigte Kinder, die nicht in der Lage sind, den Fragebogen selbst auszufüllen. Diese trifft in der vorliegenden Studie nur auf eine Teilpopulation zu.

Validität

Die Itembänke und die Kurzformen wurden basierend auf der Item-Response-Theorie (IRT) entwickelt [13] [25]. Basierend auf den Angaben zur Entwicklung wird von Inhaltsvalidität und Konstruktvalidität ausgegangen [13,25]. Known-Groups-Validität wurde für verschiedene chronische Erkrankungen (Asthma, Krebs, chronische Nierenerkrankung, Sichelzellanämie) und bezüglich Hospitalisierung innerhalb der letzten 6 Monate (ja/nein) gezeigt [12]. Keine Gruppenunterschiede wurden für Kinder mit unterschiedlicher Ausprägung rheumatischer Erkrankungen und für die Kurzversion „Obere Extremitäten“ für Kinder mit Übergewicht festgestellt [12]. Untersuchungen der konvergenten Validität konnten nicht identifiziert werden. Test-Retest-Reliabilität wurde für die beiden Kurzformen anhand einer Stichprobe von 54 Kindern mit Messung im Abstand von 2 Wochen gezeigt [47]. Mittels Scale Judgement Method und IRT-Analyse wurde für die pädiatrischen Kurzformen „Mobilität“, „Depressive Symptome“, „Pain Interference“ und „Fatigue“ eine MID von 2 durch Kliniker und 3 durch Jugendliche und Eltern abgeleitet [44]. Dabei beurteilten Jugendliche, Eltern und Kliniker jeweils 6 Testvignetten für jede Skala zu 2 unterschiedlichen Zeitpunkten, ob es sich dabei um eine wichtige Veränderung handelt. Als MID wurde der Wert verwendet, bei dem 50 % der beurteilenden Personen von einer wichtigen Veränderung ausgegangen sind. Dies stellt kein etabliertes Verfahren zur Ableitung einer MID dar und die durchgeführte Untersuchung weist verschiedene Limitationen auf (z. B. kleine Anzahl ausschließlich fiktive Fallvignetten zur Ableitung der MID). Daten zur Änderungssensitivität konnten nicht identifiziert werden.

Instrumente zur Fremdeinschätzung (Elternfragebogen/Proxy-Versionen)

Basierend auf dem initialen pädiatrischen Itempool aus 293 Items wurden Proxy-Versionen der pädiatrischen PROMIS-Itembänke entwickelt, die eine Fremdeinschätzung der Kinder durch ihre

Eltern ermöglichen sollen. Zunächst wurden die Items der Selbsteinschätzung durch Umformulierungen in Proxy-Items konvertiert und kognitive Interviews zum Verständnis der Items mit 25 Eltern von Kindern zwischen 5 und 17 Jahren durchgeführt. Anschließend wurden die 293 Items bei Eltern ($n = 1.980$) gesunder und chronisch kranker Kinder zwischen 5 und 17 Jahren getestet. Dafür wurden die Items auf 9 verschiedene Test-Fragebögen verteilt, wobei jedes Item in 3 Fragebögen enthalten war. Jedes Item wurde von mindestens 428 Eltern beantwortet. Bei einem Großteil der Eltern ($n = 1.548$) waren die Kinder ≥ 8 Jahre alt; diese Kinder beantworteten parallel die Items der Selbstauffüllerversion [25].

Zur Untersuchung der Proxy-Itembänke wurden zunächst die Daten der Eltern von Kindern ab 8 Jahren sowie die Daten der Selbsteinschätzung dieser Kinder verwendet [49]. Konfirmatorische Faktoren- und IRT-Analysen zur Skalendimensionalität und lokalen Abhängigkeit sowie IRT-Analysen zum Differential Item Functioning (DIF) wurden durchgeführt. Bei der finalen Kalibrierung der eindimensionalen Domänen wurden die Skalen der Proxy-Version auf die Metrik der Kinder-Items angepasst. Ebenso wie für die Itembänke zur Selbsteinschätzung wurden eindimensionale Kurzformen für die einzelnen Itembänke der Eltern-/Betreuerversionen entwickelt, wobei im Ergebnis die Proxy-Kurzformen jeweils ein Teil-Set der Kinderversionen darstellten.

Zweidimensionale IRT-Modelle wurden eingesetzt, um die Korrelation zwischen Eltern- und Kinderversion zu schätzen. Es zeigte sich eine schwache bis moderate Übereinstimmung zwischen Fremd- und Selbsteinschätzung (Korrelationen zwischen 0,4 und 0,7). Diese Ergebnisse werden von der Autorengruppe als konsistent mit den Angaben in der Literatur beschrieben, wonach Selbsteinschätzungen der Lebensqualität durch die Kinder und Fremdeinschätzungen durch die Eltern je nach Domäne voneinander abweichen können [49]. Angaben zur Änderungssensitivität der Proxy-Itembänke wurden nicht identifiziert.

Zur Untersuchung der Proxy-Versionen für Eltern von Kindern zwischen 5 und 7 Jahren wurden IRT-Analysen zum DIF durchgeführt, um zu überprüfen, ob die Items bei Eltern von jüngeren Kindern dieselben unbeobachteten Konstrukte auf die gleiche Weise messen wie bei Eltern von älteren Kindern [50]. Dazu wurde das DIF zwischen den Eltern-Antworten der Altersgruppen 5–7 und 8–17 Jahre untersucht. Datengrundlage war die Stichprobe der 1.980 Eltern gesunder und chronisch kranker Kinder zwischen 5 und 17 Jahren (s. o. [25]). Items, die DIF aufwiesen, wurden einem 7-köpfigen Experten-Panel vorgelegt; dieses entschied im Konsens unter Berücksichtigung der DIF-Analysen und der Inhalte der Items, ob die entsprechenden Items von der Proxy-Version für Eltern von Kindern zwischen 5 und 7 Jahren gestrichen wurden oder mit einem Hinweis, diese Items nicht für den Vergleich von jüngeren und älteren Kindern zu verwenden, beibehalten wurden. Im Ergebnis wurden die meisten Items auch in der Proxy-Version für Eltern von jüngeren Kindern berücksichtigt. Zu beachten ist, dass ein Teil der Items aufgrund unzureichender Stichprobengröße nicht in die Analyse einfluss [25].

Die beiden PROMIS-Kurzversionen „Mobilität“ und „Obere Extremitäten“ werden als ausreichend valide und reliabel eingeschätzt. Die im Dossier dargestellten Responderanalysen beruhen auf einer Responderchwelle entsprechend 15 % der Skalenspannweite von 0 bis 32 Punkten. Diese Skalenspannweite ergibt sich jedoch nicht aus den im Studienprotokoll abgedruckten Fragebogen-versionen. Die Responderanalysen werden daher nicht für die Nutzenbewertung herangezogen und anstelle dessen die kontinuierlichen Analysen dargestellt.

Ophtalmologische Endpunkte

Sehschärfe

Der Endpunkt „Sehschärfe“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Ermittlung des Fernvisus erfolgte für alle Teilnehmenden der SPRINT-Studie mit orbitalen PN. Die Bestimmung sollte mittels HOTV- oder Teller-Acuity-Sehtesttafeln erfolgen. Beim HOTV-Test müssen die Buchstaben H, O, T und V in abnehmender Größe erkannt werden. Es sollte das HOTV-ATS-Protokoll der Pediatric Eye Disease Investigator Group (PEDIG) mit einem Abstand von 3 Metern verwendet werden. Bei Kindern, die zu jung waren, um einen HOTV-Test zuverlässig durchzuführen, wurde die Sehschärfe mittels Teller-Acuity-Sehtesttafeln gemessen. Dabei muss ein Streifenmuster von einer grauen Oberfläche unterschieden werden. Die Streifen variieren in Breite und im Abstand. Je enger die Streifen nebeneinanderliegen, desto schwieriger wird deren visuelle Wahrnehmung. Beginnend mit einem niederfrequenten Gitter (grobe Streifen), betrachtet der Beobachter das Blickverhalten des Kindes. Basierend auf einer Vielzahl von Hinweisen (z. B. Fixation, Augenbewegungen, Kopfbewegungen oder Zeigen), beurteilt der Beobachter, ob das Kind das Gitter auf jeder Karte sehen kann. Der Abstand zu den Sehtesttafeln sollte 55 cm betragen. Zu jedem Testzeitpunkt sollte das gleiche Verfahren angewendet werden.

Die Sehschärfe wurde beim HOTV-Test in logMAR (Logarithm of the Minimum Angle of Resolution) angegeben und bei den Teller-Acuity-Sehtesttafeln zunächst in Zyklen/cm aufgezeichnet und anschließend in logMAR umgerechnet. Der logMAR ist dabei der Logarithmus des minimalen Auflösungswinkels. Ein abnehmender logMAR-Wert entspricht einer Verbesserung der Sehschärfe. Die Sehschärfe sollte für das Alter zum Zeitpunkt des Tests korrigiert und eine alters-basierte Sehschärfe berechnet werden. Im Protokoll wird eine Veränderung von $\geq 0,2$ logMAR als signifikant betrachtet. Die Erhebung der Sehschärfe erfolgte nur in Phase II (siehe Tabelle 8).

Bewertung

Die Operationalisierung ist nachvollziehbar. Die genaue Durchführung des Sehtests mittels Teller-Acuity-Tafeln ist unklar, da kein Testprotokoll und keine Details zur Testdurchführung im Studienprotokoll beschrieben werden. Teller-Acuity-Sehtesttafeln gibt es in verschiedenen Ausführungen. Auch die Art der verwendeten Tafeln wird nicht beschrieben.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die Sehschärfe wird als unmittelbar patientenrelevant angesehen.

Validität

HOTV-Test

Es wurden vom pU keine Studien zu Validität und Herleitung der MID vorgelegt. In einer Querschnittsstudie zur Untersuchung der Vergleichbarkeit und Prüfbarkeit von Sehschärfetests wurden bei australischen Vorschulkindern im Alter von mindestens 24 Monaten ($n = 1.746$) das HOTV-ATS-Protokoll verwendet. Bei Kindern ab 60 Monaten ($n = 576$) wurde die auf logMAR basierende HOTV-

Sehtafel mit der auf logMAR basierenden ETDRS-Sehtafel (CSV 1000) verglichen. Es wurde die Prüfbarkeit untersucht, definiert als die Fähigkeit den Sehschärfetest mit beiden Augen beim ersten Mal vollständig abzuschließen. Die Prüfbarkeit betrug mindestens 93 % beim HOTV-ATS-Protokoll bei Kindern ab 42 Monaten ($n = 352$). Bei Kindern ab 60 Monaten war die Prüfbarkeit mit dem HOTV-Test im Vergleich zur ETDRS-Sehtafel höher (HOTV: 94 % vs. ETDRS: 59 %) [28]. Es wurde eine gute Test-Retest-Reliabilität bzw. Interrater-Reliabilität des HOTV-ATS-Protokolls bei Kindern mit einem mittleren Alter von $5,3 \pm 1,1$ Jahren festgestellt [24]. In 93 % der 176 mit dem HOTV-ATS-Protokoll untersuchten Augen (88 Personen) lag der Retest-Wert innerhalb von 0,1 logMAR-Einheiten des initialen Werts [24].

Teller-Acuity-Sehtesttafeln

Das REiNS Visual Outcomes Committee empfiehlt primär den Einsatz von Teller-Acuity-Sehtesttafeln zur Messung der Sehschärfe bei Personen mit NF1 und Optikusgliomen, da er bereits ab einem Alter von 6 Monaten eingesetzt werden kann und als reliabel gilt [18]. Es liegt eine Reihe von Studien zu Test-Retest-Reliabilität und Interrater-Reliabilität vor, welche zeigen, dass bei entsprechend standardisierter Testdurchführung, Erfahrung und Schulung reliable Ergebnisse erhalten werden können [22,31,43]. Für die SPRINT-Studie liegen jedoch keine Angaben zur Testdurchführung vor.

Responderschwelle

Gemäß REiNS Visual Outcomes Committee existiert keine validierte Definition einer klinisch signifikanten Veränderung für die Sehschärfe [18]. Es wird eine Änderung um 0,2 logMAR-Einheiten oder größer als Responderschwelle empfohlen, da dies ungefähr der doppelten Variation entspricht, die zwischen unterschiedlichen Testdurchführenden und Testsessions beobachtet wird. Eine Bestätigung der klinisch signifikanten Veränderungen zu einem späteren Zeitpunkt im Studienverlauf wird empfohlen.

Zusammenfassung

Die Messung der Sehschärfe mittels HOTV wird als valide und reliabel erachtet und für die Nutzenbewertung herangezogen. Für die mittels Teller-Acuity-Sehtesttafeln gemessene Sehschärfe sind die Informationen zur Operationalisierung nicht ausreichend und die Durchführung/Standardisierung des Tests ist nicht nachvollziehbar. Diese wird daher nicht zur Nutzenbewertung herangezogen. Die gewählte Responderschwelle von 0,2 logMAR erscheint plausibel.

Proptosis / Exophthalmus

Der Endpunkt „Proptosis / Exophthalmus“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund unklarer Patientenrelevanz.

Operationalisierung

Beschreibung

Exophthalmometrische Messungen der Proptosis sollten während der gesamten Studiendauer mit demselben Gerät für jede Person durchgeführt werden. Die Messungen wurden dreimal bei jeder Studienvisite durchgeführt und anschließend gemittelt. Der Typ des Exophthalmometers und die Basis zwischen den Augenhöhlenrändern sollten bei jeder Messung und jeder Studienvisite gleich sein, es sei denn, das Kind hatte zwischen den Messzeitpunkten ein deutliches Wachstum des Gesichts. Die Messung des Ausmaßes der Proptosis wird in mm angegeben. Für die Auswertung, werden Veränderungen der Proptosis > 2 mm als signifikant angesehen.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Das REiNS Visual Outcomes Committee beschreibt bezüglich Proptosis, dass die funktionelle Auswirkung einer Änderung in der Proptosis unklar ist [18]. Zudem ist unklar, welche Relevanz mögliche mit einer Proptosis einhergehende Entstellungen im Anwendungsgebiet haben.

Validität

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

Atemfunktion

Einsekundenkapazität

Der Endpunkt „Einsekundenkapazität“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund fehlender Validität als Surrogat für patientenrelevante Endpunkte.

Operationalisierung

Beschreibung

Die Einsekundenkapazität ($FEV_{1/0,75}$) wird mittels Spirometrie (Lungenfunktionstest) bei allen Patientinnen und Patienten mit PN der Atemwege (obere Atemwege / extrathorakal und untere Atemwege / intrathorakal) erhoben. Bei Vorschulkindern wird $FEV_{0,75}$ anstelle von FEV_1 verwendet. Responderanalysen waren geplant für eine absolute Veränderung um $\geq 12\%$ (Verbesserung) bzw. $\geq -12\%$ (Verschlechterung) und keine Veränderung im Vergleich zu Baseline. Auf Grundlage des REiNS-Komitees für klinische Studien wurde seitens pU eine Zunahme des $FEV_{1/0,75}$ -Werts um 12% im Vergleich zu Baseline als klinisch relevante Verbesserung beschrieben [41].

Bewertung

Die Operationalisierung ist weitgehend nachvollziehbar. Es liegen keine Angaben zur standardisierten Durchführung des Tests vor.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die Einsekundenkapazität ist kein unmittelbar patientenrelevanter Endpunkt. Die Validität der Einsekundenkapazität als Surrogat patientenrelevanter Morbiditätsendpunkte wird weder vom pU hergeleitet noch durch entsprechende Literatur ausreichend gestützt.

Validität

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

Atemwegswiderstand (R₂₀)

Der Endpunkt „Atemwegswiderstand (R₂₀)“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund fehlender Validität als Surrogat für patientenrelevante Endpunkte.

Operationalisierung

Beschreibung

Der Atemwegswiderstand bei 20 Hertz (R₂₀) wurde mittels Impulsoszillometrie bei allen Patientinnen und Patienten mit PN der Atemwege (obere Atemwege / extrathorakal und untere Atemwege / intrathorakal) erhoben. Responderanalysen waren geplant für eine absolute Veränderung um $\geq -20\%$ (Verbesserung) bzw. $\geq 20\%$ (Verschlechterung) und keine Veränderung im Vergleich zu Baseline. Auf Grundlage des REINS-Komitees für klinische Studien wurde seitens pU eine Abnahme des R₂₀-Werts im Vergleich zu Baseline als klinisch relevante Verbesserung beschrieben.

Bewertung

Die Operationalisierung ist weitgehend nachvollziehbar. Es liegen keine Angaben zur standardisierten Durchführung des Tests vor.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Der Atemwegswiderstand ist kein unmittelbar patientenrelevanter Endpunkt. Die Validität des Atemwegswiderstandes als Surrogat patientenrelevanter Morbiditätsendpunkte wird weder vom pU hergeleitet noch durch entsprechende Literatur ausreichend gestützt.

Validität

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

Apnoe-Hypopnoe-Index

Der Endpunkt „Apnoe-Hypopnoe-Index“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund fehlender Validität als Surrogat für patientenrelevante Endpunkte.

Operationalisierung

Beschreibung

Patientinnen und Patienten mit PN der Atemwege (obere Atemwege / extrathorakal und untere Atemwege / intrathorakal) sollten sich einer Schlafuntersuchung unterziehen. Diese war nicht notwendig für Patientinnen und Patienten mit einer Tracheostomie, welche die Obstruktion der Atemwege umgeht. Teilnehmende mit CPAP (Continuous Positive Airway Pressure) oder BiPAP (Bilevel Positive Airway Pressure) sollten die Schlafuntersuchung mit ihren Einstellungen von zu Hause durchführen.

Der Apnoe-Hypopnoe-Index (AHI) erfasst die Anzahl der Schlafstörungen (Apnoe- und Hypopnoe-Ereignisse) pro Stunde im Schlaf (i. d. R. 8 Stunden oder mehr). Patientinnen und Patienten wurden bei der Studienvsiste als schlafgestört eingestuft (ja/nein), wenn der AHI > 1 betrug. Für Personen mit einem AHI > 5 zu Baseline wurde eine klinisch relevante Verbesserung in Abhängigkeit vom Ausgangswert bestimmt [41].

Der Endpunkt wurde zu Baseline vor Behandlungsbeginn untersucht. Bei normalem Ergebnis sollte die Untersuchung zur Studienvsiste vor Zyklus 13 erneut erhoben werden und danach nur, wenn

klinisch indiziert, zum Monitoring. Bei unnormalen/auffälligen Ergebnissen sollte die Erhebung zur Studiervisite vor Zyklus 5, 9, 13 und danach jährlich erhoben werden.

Bewertung

Die Operationalisierung ist weitgehend nachvollziehbar. Es liegen keine Angaben zur standardisierten Durchführung des Tests vor.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Der AHI ist kein unmittelbar patientenrelevanter Endpunkt. Die Validität des AHI als Surrogat patientenrelevanter Morbiditätsendpunkte wird weder vom pU hergeleitet noch durch entsprechende Literatur ausreichend gestützt.

Validität

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

Darm- und Blasenfunktion mittels Dysfunctional Voiding Questionnaire

Der Dysfunctional Voiding Questionnaire wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Validität.

Operationalisierung

Beschreibung

Die Darm- und Blasenfunktion wurde mittels Dysfunctional Voiding Questionnaire (DVQ) erhoben. Der Fragebogen sollte bei allen Testpersonen mit Blasen- oder Darmstörung eingesetzt werden. Ob eine PN-bezogene Störung der Darm- und Blasenfunktion vorliegt wurde zu Studienbeginn durch das Prüfpersonal diagnostiziert. Falls die Testperson dazu in der Lage war, erfolgte die Beantwortung durch Selbstbewertung. Andernfalls war eine Fremdbewertung durch ein Elternteil oder die betreuende Person vorgesehen. In Modul 4 führt der pU aus, dass der Fragebogen von allen Testpersonen ab 8 Jahren selbst berichtet wurde. Beim DVQ handelt es sich um einen Symptomfragebogen zur Erfassung des Schweregrads von Darm- und Blasenfunktionsstörungen, bestehend aus 13 Symptom-Items und einem Feedback-Item zur Bewertung des Fragebogens. Die Beantwortung der Items erfolgt auf einer Likert-Skala mit 5 Punkten. Aus den 13 Symptom-Items wird ein Gesamtwert gebildet, der von 0 bis 52 Punkten reicht, wobei höhere Werte einer höheren Symptomlast entsprechen. Ein Gesamtwert von ≥ 11 Punkten wird vom pU als Schwelle zur Klassifikation einer Darm- und Blasenfunktionsstörung angegeben.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es ist zu berücksichtigen, dass der Fragebogen nur bei einem Teil der Patientenpopulation (Personen mit PN-bezogenen Störungen der Darm- und Blasenfunktion zu Studienbeginn) eingesetzt wurde. Testpersonen, bei denen sich im Laufe der Studie eine Störung der Darm- und Blasenfunktion entwickelte, wurden somit nicht erfasst. Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Darm- und Blasenfunktion der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Störungen der Darm- und Blasenfunktion sind patientenrelevant und treten im Krankheitsbild der PN auf.

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Validität

Darm- und Blasenfunktionsstörungen wurden in der Studie SPRINT über den DVQ erhoben, einem diagnostischen Fragebogen zur Erfassung eines dysfunktionalen Ausscheidungssyndroms mit insgesamt 14 Items. Der Fragebogen wurde zur Selbstbewertung bei Personen ab einschließlich 9 Jahren entwickelt und ist in der Entwicklungsstudie auch als Vancouver Nonneurogenic Lower Urinary Tract Dysfunction / Dysfunctional Elimination Syndrome Questionnaire benannt [1]. Bei jüngeren Personen können Eltern die Bewertung vornehmen. Inwiefern sich Selbst- und Fremdbewertung bei dem Fragebogen strukturell unterscheiden ist unklar. Der pU nimmt getrennte Auswertungen vor.

Entwickelt wurde der Fragebogen über eine Kombination von Literaturrecherche, Expertenmeinung und Patienten-/Familienbefragungen. Einbezogen wurden Personen mit nicht-neurogenen Dysfunktionen des unteren Harntrakts und dysfunktionalen Ausscheidungssyndromen [1]. Die Validierung des Fragebogens erfolgte ebenfalls anhand von Personen dieser Indikationen (n = 62) und 50 Kontrollpersonen. Dabei zeigte sich eine gute Test-Retest-Reliabilität. Die interne Konsistenz war jedoch mit einem Cronbach's alpha von 0,445 nur moderat ausgeprägt. Über eine Faktorenanalyse wurden darüber hinaus auch 4 unterschiedliche Domänen des Fragebogens identifiziert. In der Folge wurden diese Domänen jedoch nicht in die Auswertung des Fragebogens integriert. Bezüglich der Validität liegen ausschließlich Untersuchungen unter Einbezug der Kontrollen vor. Erwartungsgemäß waren die Gesamtwerte der betroffenen Personen deutlich höher als die Gesamtwerte der Kontrollen und über eine ROC-Kurve wurde ein Gesamtwert von 11 Punkten als optimaler Cutoff-Wert für die Diagnose von Dysfunktionen des unteren Harntrakts und dysfunktionalen Ausscheidungssyndromen ermittelt. Validierungsdaten zu weiteren Instrumenten oder unter Einbezug von unterschiedlichen Krankheits-Schweregraden liegen nicht vor. Auch Untersuchungen, ob der Fragebogen Veränderungen valide abbilden kann, konnten nicht identifiziert werden. Somit ist fraglich, ob sich der DVQ zur Evaluierung in klinischen Studien eignet oder ob es sich hierbei in erster Linie um ein Instrument zur Diagnostizierung von Störungen der Darm- und Blasenfunktion handelt. In Verbindung mit den moderaten Werten bei der internen Konsistenz und der Nicht-Berücksichtigung der identifizierten Domänenstruktur wird das Instrument als nicht ausreichend valide zur Bewertung von Störungen der Darm- und Blasenfunktion eingestuft.

Symptomatik mittels Symptomcheckliste

Der Endpunkt „Symptomatik mittels Symptomcheckliste“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Mithilfe der Symptomcheckliste wird das Ausmaß von 36 Symptomen innerhalb der letzten 2 Wochen erfragt. Die Symptomcheckliste sollte von der teilnehmenden Person selbst ausgefüllt werden, sofern diese in der Lage dazu war. Falls nicht sollten Eltern bzw. Erziehungsberechtigte die Checkliste ausfüllen. Die Symptomcheckliste sollte zu allen Zeitpunkten von der gleichen Person ausgefüllt werden. Die Beantwortung erfolgte auf einer 5-stufigen Likert-artigen Skala (mit den Antwortkategorien: überhaupt nicht / ein wenig / etwas / ziemlich viel / viel) auf die Frage, wieviel die Symptome ein Problem für die befragten Person darstellten. Folgende Symptome wurden abgefragt: Fatigue/ Müdigkeit; Schlafprobleme; Verringerter Appetit; Erhöhter Appetit; Kopfschmerzen; Sehstörungen; Vermindertes Hörvermögen; Wunde Stellen im Mund; Schwierigkeiten beim Schlucken; Würgen; Schnarchen; Häufiges Aufwachen in der Nacht; Husten; Keuchen; Schwierigkeiten beim Atmen; Schmerzen in der Brust; Herzklopfen/-flattern; Kurzatmigkeit bei Belastung; Kurzatmigkeit in Ruhe; Anschwellen der Füße/Hände; Unterleibsschmerzen; Sodbrennen; Übelkeit; Erbrechen; Diarrhö; Verstopfungen; Stuhlinkontinenz; Schmerzen beim Urinieren; Erhöhte Harnfrequenz / Erhöhter Harndrang; Schwierigkeiten beim Beginn des Urinierens; Harninkontinenz; Schwäche; Muskelschmerzen; Schwindel; Taubheitsgefühl; Kribbeln.

Im Dossier werden Häufigkeiten für die einzelnen Antwortkategorien dargestellt sowie Responderanalysen für eine Verbesserung um eine Antwortkategorie.

Die Erhebung der Symptomatik erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung ist größtenteils nachvollziehbar. Es ist unklar, anhand welcher Kriterien entschieden wurde, ob die Teilnehmenden selbst die Symptomcheckliste ausfüllten oder eine Checkliste von Eltern/Betreuern ausgefüllt wurde.

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Symptomatik der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Das Auftreten der oben genannten Symptome ist patientenrelevant.

Validität

In den Studienunterlagen wurden keine Informationen zu Entwicklung, Item-Auswahl und Validierung der Symptomcheckliste identifiziert. Es ist daher bspw. unklar, inwieweit die abgefragten Symptome alle wesentlichen Symptome der Erkrankung abbilden. Zudem ist unklar, ob die Formulierung der Items und Antwortkategorien eindeutig verständlich für die Teilnehmenden und Eltern war und ob es

Unterschiede im Antwortverhalten zwischen Teilnehmenden und Eltern gab. Die Symptomcheckliste wird dennoch für die Nutzenbewertung herangezogen. Die Responderdefinition für eine Veränderung um eine Antwortkategorie entspricht > 15 % der Skalenspannweite und kann damit für die Nutzenbewertung herangezogen werden.

Gehfähigkeit mittels 6-Minuten-Gehtest

Der Endpunkt „Gehfähigkeit mittels 6-Minuten-Gehtest“ wird in der Nutzenbewertung ergänzend im Anhang dargestellt.

Operationalisierung

Beschreibung

Die Gehfähigkeit wurde mittels 6-Minuten-Gehtest (6MWT) anhand der zurückgelegten Strecke in Metern bei allen Patientinnen und Patienten ab 5 Jahren mit PN der unteren Extremitäten, mit einer Rückenmarkskompression, mit PN der Atemwege und mit PN-assoziierten Morbiditäten der Motorfunktion erhoben. Die Bewertung sollte, wenn möglich, durch die Rehabilitationsmedizin, die Physiotherapie oder Physiotherapie erfolgen. Die Tests wurden zu ausgewählten Studienvisiten jeweils vor der Behandlung bzw. den Therapiezyklen sowie am Ende der Therapie durchgeführt. Neben der Gehstrecke wurden die prozentual erreichte Normalstrecke mit Referenzwerten der Publikation von Geiger et al. [21] sowie die Gehgeschwindigkeit (m/Min.) erhoben.

Als klinisch relevante Veränderung hat der pU ein Schwellenwert mit einer absoluten Veränderung von ≥ 30 m (Verbesserung / Verschlechterung / keine Veränderung) präspezifiziert.

Die Erhebung der Gehfähigkeit erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar. Angaben zur standardisierten Durchführung des Tests wurden nicht identifiziert. Zudem bleibt unklar, ob das Prüfpersonal für die Durchführung des Tests geschult wurde.

Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Gehfähigkeit der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die Gehfähigkeit erhoben mittels 6MWT als eine zentrale körperliche Funktion des alltäglichen Lebens ist geeignet zur Erfassung der körperlichen Belastbarkeit und stellt einen patientenrelevanten Endpunkt dar.

Validität

Der 6MWT wurde entwickelt zur Messung der funktionalen körperlichen Fähigkeiten bzw. körperlichen Fitness [3]. Es handelt sich um ein standardisiertes und etabliertes Testverfahren, das zu Diagnostik und Verlaufskontrolle bei einer Vielzahl von Indikationen zum Einsatz kommt. Bei Erfassung und Erhebung des Endpunkts sind mehrere potentielle Einflussfaktoren bzw. Quellen für

Variationen der Testergebnisse des 6MWT zu berücksichtigen. Hierzu zählen zum einen Patientencharakteristika (Alter, Körpergröße und -gewicht, Geschlecht) als auch äußere Faktoren wie Motivation, Schulung des Studienpersonals und Erfahrung der Testpersonen mit dem 6MWT. Wird eine Veränderung der körperlichen Fähigkeiten durch den 6MWT ermittelt, sollten mögliche Einflussfaktoren konstant gehalten werden, um eine potentielle Veränderung der Gehstrecke auf den zu untersuchenden Faktor, in diesem Fall auf die Studienmedikation, zurückführen zu können.

Es wird grundsätzlich davon ausgegangen, dass bei Einhaltung der Vorgaben der Richtlinie der American Thoracic Society valide Testergebnisse erzielt werden können. Die Reproduzierbarkeit der Testergebnisse des 6MWT ist bei qualitätssichernden Maßnahmen auch bei Kindern hoch [35]. Zu diesen Maßnahmen gehören z. B. Schulungen des Studienpersonals, standardisierte Formulierungen der Ermutigungen nach bestimmten Zeiteinheiten, Messung durch dieselbe Person, ein bis zwei Probe-Testläufe. Zur Durchführung des Tests und Einhaltung der Maßnahmen der American Thoracic Society finden sich in den vorliegenden Studienunterlagen keine Angaben. Eine abschließende Beurteilung zur standardisierten Durchführung des Tests sowie zur Validität der Testergebnisse ist nicht möglich. Die Ergebnisse werden daher nur ergänzend dargestellt.

Der vom pU vorgeschlagene Normwertvergleich (Prozentsatz der vorhergesagten Normalstrecke) wird als nicht valide gesehen, da für die Ableitung der Referenzwerte die Testprozedur der American Thoracic Society modifiziert wurde und ein Handlauftrad als Motivation genutzt wurde und unklar ist, ob eine ähnliche Testdurchführung auch innerhalb der SPRINT-Studie stattfand [21,35].

Als klinisch relevante Veränderung wird in der SPRINT-Studie eine Veränderung von ≥ 30 m präspezifiziert. Bezug genommen wird hierfür auf eine Untersuchung zur Enzymersatztherapie mit Velmerase alpha bei Kindern und Erwachsenen mit Alpha-Mannosidose. In der Publikation findet sich keine Herleitung der Relevanzschwelle für den 6MWT. Eine Übertragbarkeit auf das vorliegende Anwendungsgebiet erscheint nicht sachgerecht. Allgemein ist davon auszugehen, dass eine Schwelle für klinisch relevante Veränderungen im 6MWT erkrankungs- bzw. populationsspezifisch abzuleiten ist, wobei sowohl Art der Erkrankung, Erkrankungsschwere als auch individuelle Eigenschaften wie z. B. Alter und Geschlecht einen relevanten Einfluss auf die klinische Relevanzschwelle haben können. Aufgrund der genannten Einschränkungen werden die verwendeten Relevanzschwellen nicht für die Nutzenbewertung herangezogen.

Die ergänzende Darstellung erfolgt daher für die rohen Distanzwerte der 6-Minuten-Gehstrecke. Aufgrund des einarmigen Studiendesign können bei dieser Darstellungsweise Therapie- und Alters-effekte auf die zurückgelegte Distanz nicht voneinander unterschieden werden.

2.3.3 Lebensqualität

Pediatric Quality of Life Inventory

Der Endpunkt „Pediatric Quality of Life Inventory“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Erfassung der gesundheitsbezogenen Lebensqualität erfolgte mittels Pediatric Quality of Life Inventory (PedsQL) 4.0 [48]. Das Instrument umfasst multidimensionale Skalen zur Selbst- und Fremdeinschätzung für Kinder und Eltern zur Beurteilung der Lebensqualität bei Kindern, Jugendlichen und jungen Erwachsenen im Alter von 2 bis 25 Jahren. Der PedsQL ist eine standardisierte pädiatrische Lebensqualitätsskala, die sowohl generische als auch krankheitsspezifische Module umfasst. Sie besteht aus einer Kernskala zur globalen Lebensqualität (23 Items), die 4 Subskalen enthält:

- Körperliche Funktion (8 Items),
- Emotionale Funktion (5 Items),
- Soziale Funktion (5 Items),
- Schulische Funktion (5 Items).

Jedes Item wird auf einer 5-stufigen Likert-Skala bewertet (0 = nie ein Problem; 1 = fast nie ein Problem; 2 = manchmal ein Problem; 3 = oft ein Problem; 4 = fast immer ein Problem). Die Items werden umgekehrt bewertet und linear in eine Skala von 0 bis 100 transformiert. Höhere Werte verweisen auf eine bessere Lebensqualität. Ein Gesamtskalenwert berechnet sich aus der Summe aller Items geteilt durch die Anzahl der beantworteten Items auf allen Skalen. Sofern mehr als 50 % der Items in der Skala unbeantwortet sind, wird der Skalenwert nicht berechnet.

Verfügbar sind verschiedene Versionen des PedsQL für die Selbstauskunft gestaffelt nach Alter der Patientinnen/Patienten von 5 bis 25 Jahren (Kleinkind: 5–7, Kind: 8–12, Jugendliche/r: 13–18, junge/r Erwachsene/r: 18–25 Jahre) und für die Fremdbewertung durch die Eltern, ebenfalls gestaffelt nach Alter der Patientinnen/Patienten von 2 bis 18 Jahren (Kleinkind: 2–4, junges Kind: 5–7, Kind: 8–12, Jugendliche/r: 13–18 Jahre). Für die SPRINT-Studie fand eine Selbsteinschätzung der Lebensqualität durch Kinder im Alter von 8 bis 18 Jahren statt, eine Fremdeinschätzung durch Eltern oder Betreuer (Fremdbewertung) bei Kindern im Alter von 2 bis 18 Jahren. In der Studie SPRINT wurde die Akutversion des PedsQL mit einem Bezugszeitraum von 7 Tage verwendet.

Eine Responderschwelle sollte innerhalb der Studie mithilfe des GIC (Tumor-bezogene Morbidität jeweils für korrespondierende Selbst- oder Fremdbewertung) abgeleitet werden. Darüber hinaus werden aus der Literatur abgeleitete [46] klinisch relevante Sensitivitätsschwellenwerte pro Subskala und für den Gesamtwert berichtet:

PedsQL-Skalen	Selbstbewertet (Kind)	Fremdbewertet (Eltern)
Gesamtwert	4,36	4,50
Körperliche Funktion	6,66	6,92
Emotionale Funktion	8,94	7,79
Soziale Funktion	8,36	8,98
Schulische Funktion	9,12	9,67

Es soll die Anzahl an Personen mit Verbesserungen und Verschlechterungen um den jeweiligen Schwellenwert sowie ohne Veränderung berechnet werden. Zudem werden Veränderungen über die Zeit bis Monat 13 mittels MMRM-Analyse berechnet. Die in den Studienunterlagen beschriebenen klinischen Relevanzschwellen werden vom pU nicht für das Nutzendossier ausgewertet. Der pU stellt Responderanalysen für Patientinnen und Patienten mit einer Verbesserung des PedsQL im Vergleich zu Baseline um 15 Punkte auf der transformierten Skala dar, und bezieht sich dabei auf das im IQWiG-Methodenpapier 6.0 empfohlene Responsekriterium von 15 % der Skalenspannweite von 0 bis 100.

Die Erhebung des PedsQL erfolgte nur in Phase II (siehe Tabelle 8). Es erfolgte keine Erhebung nach Beendigung der Studienmedikation in der Langzeitnachbeobachtung.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar. Die Erhebung erfolgte nur innerhalb von Patientinnen und Patienten unter Selumetinib-Therapie und nicht nach Behandlungsabbruch. Die gewonnenen Daten spiegeln dadurch nicht die Lebensqualität der in die Studie eingeschlossenen gesamten Studienpopulation wider.

Es ist unklar, warum die Akutversion des PedsQL mit einem Bezugszeitraum von 7 Tagen eingesetzt wurde und nicht die Standardversion mit einem Bezugszeitraum von einem Monat.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Die patientenberichtete Erfassung der Lebensqualität mittels PedsQL wird als patientenrelevant erachtet. Ein durch die Eltern berichtete Lebensqualität mittels PedsQL kann für eine bestimmte Patientenpopulation, z. B. jüngere oder kognitiv beeinträchtigte Kinder, die nicht in der Lage sind, den Fragebogen selbst auszufüllen, als patientenrelevant erachtet werden. Dies trifft in der vorliegenden Studie nur auf eine Teilpopulation zu.

Validität

Der PedsQL ist ein etabliertes und standardisiertes Instrument zur Erfassung der Lebensqualität bei pädiatrischen Populationen mit chronischen Erkrankungen. Das Instrument wurde an einer Population US-amerikanischer Kinder entwickelt und untersucht. Es bietet die Möglichkeit der Fremdbewertung durch Eltern für 2- bis 18-jährige Kinder und Jugendliche sowie die Möglichkeit der Selbstbewertung für Kinder, Jugendliche und junge Erwachsene von 5 bis 25 Jahren. In der vorliegenden SPRINT-Studie wurden selbstbewertete Fragebögen für Kinder im Alter von 8 bis 18 Jahren erhoben sowie durch Eltern/Erziehungsberechtigte fremdbewertete Fragebögen von Kindern im Alter 2–18 Jahre. Vorliegende Untersuchungen zum PedsQL 4.0 verweisen auf eine gute Reliabilität und Validität des Instruments. Die interne Konsistenz wurde an einer umfangreichen Stichprobe mit ca. 8.800 gesunden und ca. 850 chronisch kranken Kindern untersucht. Die Test-Retest-Reliabilität bei Kindern (7–12 Jahre) wird mit 0,86 angegeben. Die interne Validität wurde an einer Stichprobe mit ca. 5.500 Kindern (ca. 570 davon chronisch krank) untersucht. Es konnte gezeigt werden, dass Kinder mit chronischen Erkrankungen konsistent eine niedrigere Lebensqualität berichten als gesunde Kinder.

In der Literatur wird für die Selbsteinschätzung durch das Kind eine Veränderung des Gesamtskalenwerts um 4,4 Punkte als klinisch relevante Veränderung beschrieben, für die Fremdbewertung durch die Eltern eine Veränderung des Gesamtskalenwerts um 4,5 Punkte [46]. Die MID wurde für alle verfügbaren Versionen des PedsQL gemeinsam für Selbstangaben durch Kinder und Jugendliche verteilungsbasiert als SEM (Standardfehler) abgeleitet. Angaben zur Herleitung der klinischen Relevanzschwelle mittels ankerbasierter Verfahren finden sich in der Publikation nicht. Die Ermittlung der klinischen Relevanzschwelle basiert auf Querschnittsdaten, was grundsätzlich als nicht ausreichend zur Ableitung klinischer Relevanzschwellen für Veränderungen bei Patientinnen und Patienten angesehen wird. Die beschriebenen Sensitivitätsschwellenwerte entsprechen den Angaben in der Publikation. Der pU hat für das Dossier zudem Post-hoc-Analysen mit einer 15%igen Responseschwelle der Skalenspannweite durchgeführt. Diese werden in der Nutzenbewertung berücksichtigt.

2.3.4 Sicherheit

Unerwünschte Ereignisse

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

Operationalisierung

Beschreibung

Unerwünschte Ereignisse (UE) sind definiert als alle ungünstigen und unbeabsichtigten Zeichen (einschließlich eines abnormalen Laborbefundes), Symptome oder Krankheit, die zeitlich mit der Anwendung einer medizinischen Behandlung oder eines Verfahrens in Verbindung gebracht werden; unabhängig davon, ob ein Zusammenhang mit der Behandlung oder dem Verfahren besteht (Zuordnung "nicht verwandt", "unwahrscheinlich", "möglich", "wahrscheinlich" oder "definitiv").

Alle UE, einschließlich klinisch signifikanter abnormaler Laborbefunde, werden unabhängig vom Schweregrad bis zu ihrem Abklingen oder ihrer Stabilisierung nachverfolgt. UE wurden ab dem Zeitpunkt der ersten Dosis der Studienmedikation erfasst und sollten, je nachdem, was zuerst eintrat, bis zu 30 Tage nach der letzten Dosis des Studienmedikaments oder bis zu einem Studienabbruch gemeldet werden. UE, die als behandlungsbedingt, erwartet, anhaltend, aber bis 30 Tage nach Behandlungsende nicht abgeklungen waren (z. B. Alopezie), werden nach Ablauf der 30-Tage-Frist nicht weiterverfolgt.

Schwerwiegende unerwünschte Ereignisse (SUE) sind alle unerwünschten Ereignisse, die zu folgenden Ereignissen führen:

- Tod,
- lebensbedrohliches unerwünschtes Ereignis,
- stationärer Krankenhausaufenthalt oder Verlängerung eines bestehenden Krankenhausaufenthalts,
- eine anhaltende oder erhebliche Behinderung/Invalidität,
- eine kongenitale Anomalie/Geburtsfehler und
- andere medizinisch bedeutsame Ereignisse.

Die Kodierung der UE erfolgte mittels Medical Dictionary for Regulatory Activities (MedDRA), Version 21. UE wurden zudem nach Systemorganklasse und Preferred Term zusammengefasst. Die Einteilung des Schweregrads der UE basiert auf den Common Terminology Criteria for Adverse Events (CTCAE), Version 4.0, des National Cancer Institute.

Unerwünschte Ereignisse von besonderem Interesse waren in den Studienunterlagen nicht prä-spezifiziert. Folgende UE von besonderem Interesse wurden erhoben:

- Ausschlag, akneiform
- Ausschlag, nicht akneiform
- Erkrankungen des Nagels
- Effekte der Entzündung der oralen Mundschleimhaut
- Effekte einer Leukopenie
- Effekte einer Thrombozytopenie
- Effekte einer Erythropenie
- Die Netzhaut betreffende Effekte
- Effekte der Herzinsuffizienz (verminderte Ejektionsfraktion, andere kardiale Ereignisse)
- Muskelbezogene Effekte
- Physeale Dysplasie

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar.

Patientenrelevanz

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Das Auftreten von UE ist patientenrelevant. Die Patientenrelevanz von Laborparametern ist unklar.

Validität

Der Erhebung der UE wird als valide angesehen.

2.3.5 Übersicht der Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte in Phase I bzw. Phase II findet sich in Tabelle 7 bzw. Tabelle 8.

Tabelle 7: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Phase I der Studie SPRINT

Studienvisite Endpunkt	Baseline	Vor den Zyklen		Alle 6 Zyklen	Therapieende
		6	11		
Todesfälle	Erhebung erfolgte im Rahmen der Sicherheit				
Unerwünschte Ereignisse	Kontinuierliche Erhebung bis 30 Tage nach Beendigung der Einnahme der Studienmedikation				

Tabelle 8: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Phase II der Studie SPRINT

Studienvisite Endpunkt	Base- line	Vor den Zyklen							Alle 6 Zyklen	Alle 12 Zyklen	Therapie- ende	Lang- zeit- FUP
		3	5	9	13	17	21	25				
Schmerz	X	X	X	X	X	-	-	X	-	X	X	-
Grooved Pegboard Test	X		X	X	X	-	-	X	-	X	X	-
PROMIS – Körperliche Funktionsfähigkeit	X	X	X	X	X	-	-	X	-	X	X	-
Sehschärfe mittels HOTV-Sehtesttafel	X		X	X	X	-	-	X	-	X	X	wenn klinisch indiziert
Symptomcheckliste	X	X	X	X	X	X	X	X	X	-	X	-
6-Minuten-Gehstrecke ¹⁾	X		X	X	X	-	-	X	-	X	X	-
GIC	X	X	X	X	X	-	-	X	-	X ²⁾	-	
PedsQL	X	X	X	X	X	-	-	X	-	X	X	-

Studienvisite Endpunkt	Base- line	Vor den Zyklen						Alle 6 Zyklen	Alle 12 Zyklen	Therapie- ende	Lang- zeit- FUP
		3	5	9	13	17	21				
Todesfälle	Erhebung erfolgte im Rahmen der Sicherheit bzw. über Teilnahme an Studienvisiten										
Unerwünschte Ereignisse	Kontinuierliche Erhebung bis 30 Tage nach Beendigung der Einnahme der Studienmedikation										

¹⁾ Endpunkt wird ergänzend dargestellt.

²⁾ Erhebung bis 3 Jahre nach Beginn der Einnahme der Studienmedikation.

Abkürzungen: FUP: Follow-up; GIC: Global Impression of Change; PedsQL: Pediatric Quality of Life Inventory; PROMIS: Patient-Reported Outcomes Measurement Information System.

2.4 Statistische Methoden

Analysepopulationen

Intention-to-Treat (ITT)-Population: Eine Definition für die ITT-Population wird im Studienprotokoll und SAP nicht genannt. Laut Modul 4 entspricht die ITT-Population für Phase II Stratum 1 dem FAS.

Full Analysis Set (FAS) / Sicherheitspopulation (n = 50): Alle Teilnehmenden, die mindestens eine Dosis Selumetinib erhalten haben.

Für verschiedene Morbiditätsendpunkte und deren Auswertungen werden Teilpopulationen definiert. Diese sind in Kapitel 2.3.2. beschrieben.

Datenschnitte

Phase I:

- 04.01.2016 (Daten veröffentlicht in Publikation Dombi et al., 2016 [14])
- 29.06.2018
- 27.02.2021 (keine Daten eingereicht)

Phase II:

- Primäre Analyse: 29.06.2018
- 29.03.2019 (nur Daten zur Sicherheit)
- 31.03.2021 (keine Daten eingereicht)

Für die Nutzenbewertung werden die Datenschnitte der primären Analyse vom 29.06.2018 herangezogen. Der für die Studienberichte (Phase I: 08.08.2019; Phase II: 28.08.2019) relevante SAP (Version 5.0) wurde am 21.01.2019 und damit deutlich nach Durchführung des Datenschnitts am 29.06.2018 finalisiert. Für die Endpunktkategorie „Sicherheit“, inklusive Todesfälle, liegen neuere Daten eines durch die FDA getriggerten Datenschnitts vom 29.03.2019 vor, welche ergänzend für diese Endpunkte zur Nutzenbewertung herangezogen werden.

Präspezifizierung geplanter und durchgeführter Analysen

Aufgrund des einarmigen Studiendesigns wurden größtenteils deskriptive Auswertungen geplant. Für einige Analysen der PRO (Patient-Reported Outcomes) und funktionalen Endpunkte wurden MMRM-Analysen geplant. Als primäres Ziel gilt die Analyse von Baseline bis zur Studienvisite vor Zyklus 13. In diese Analysen werden alle Teilnehmenden mit einem Baseline-Wert einbezogen. Es wurden nur gemessene Werte in der Analyse berücksichtigt. Die Analyse soll zu allen Zeitpunkten erfolgen, bei

denen für 10 oder mehr Teilnehmende Daten erhoben wurden. Die abhängige Variable im Modell ist die Veränderung von Baseline zu den Erhebungszeitpunkten zu Monat 3, 5, 9 und 13. Als unabhängige Variablen werden der Zeitpunkt der Studienvisite, Baseline-Wert, Alter, Anzahl an Morbiditäten zu Baseline und ein Interaktionsterm aus Baseline-Wert und Studienvisite berücksichtigt. Das Modell nutzt eine unstrukturierte Kovarianzmatrix für die wiederholten Messungen innerhalb einer Person. Sollte dadurch keine Konvergenz erreicht werden, sollten alternative Strukturen in folgender Reihenfolge verwendet werden: „spatial power“, „Gaussian“, „spherical“ und „compound symmetry“. Die Modell-Annahmen sollen durch Analyse der Residuen überprüft werden und falls notwendig eine Transformation der Rohdaten vorgenommen werden.

Fehlende Werte und Daten-Imputation

Fehlende Werte wurden in der Regel nicht imputiert.

Für einige Analysen der PRO und funktionalen Endpunkte wurden MMRM-Analysen mit Missing-at-Random-Annahme durchgeführt. Sollten Analysen der fehlenden Werte auf eine Abweichung von der Missing-at-Random-Annahme hinweisen, so sollte der Einfluss der fehlenden Werte auf die Analyse untersucht werden.

Subgruppenanalysen

Systematische Subgruppenanalysen für bestimmte Merkmale, die über alle Endpunkte hinweg durchgeführt werden, waren laut SAP nicht geplant. Im Dossier (Modul 4) werden Subgruppenanalysen für Geschlecht und Status des Ziel-PN (progressiv, nicht-progressiv, unbekannt) dargestellt. Aufgrund des insgesamt kleinen Stichprobenumfangs und fehlender Tests auf Interaktionen werden diese Subgruppenanalysen nicht zur Nutzenwertung herangezogen.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Aufgrund des einarmigen Studiendesigns wird das Verzerrungspotential als hoch bewertet.

3 Ergebnisse der eingeschlossenen Studien

3.1 Studiencharakteristika und Studienmedikation

Allgemeine Angaben zu Patientenpopulationen und zum Patientenfluss für die verschiedenen Studienphasen und Datenschnitte mit Relevanz für die Nutzenbewertung sind in Tabelle 9 dargestellt.

In den Studienarm mit der FI-konformen Dosierung der Phase I wurden insgesamt 6 Personen eingeschlossen, welche alle bis zum Datenschnitt vom 29.06.2018 behandelt wurden. Die mediane Behandlungsdauer in Phase I betrug 1.597 Tage (\approx 52,5 Monaten oder 4,4 Jahren).

In Phase II wurden insgesamt 50 Personen eingeschlossen. Innerhalb Phase II brachen bis zum Datenschnitt vom 29.03.2019 20 Personen (40 %) die Behandlung mit Selumetinib ab, 12 % aufgrund von UE, 10 % aufgrund einer Krankheitsprogression, 8 % aufgrund der Entscheidung des ärztlichen Prüfpersonals, 4 % weil sie die Behandlungsperiode abgeschlossen hatten, 4 % wegen der Rücknahme der Einverständniserklärung und 2 % wegen Protokollverletzungen. Für 2 Teilnehmende (4 %) wird eine Wiederaufnahme der Behandlung nach Abbruch der Einnahme der Studienmedikation berichtet. Die mediane Behandlungsdauer in Phase II betrug 1.027,5 Tage (\approx 33,8 Monaten oder 2,8 Jahren) für den Datenschnitt vom 29.03.2019. Angaben zur Beobachtungsdauer für Phase II konnten in den Studienunterlagen nicht identifiziert werden. Bis zum Datenschnitt vom 29.03.2019 brachen 6 Teilnehmende (12 %) ihre Studienteilnahme in der Phase II ab, davon 3 aufgrund des Entzugs der Einverständniserklärung, 2 aus anderen Gründen und 1 Person wird als Lost to Follow-up verzeichnet. Folglich befanden sich zum Zeitpunkt des Datenschnitts 30 Personen in der Behandlungsphase und 14 Personen, die ihre Behandlung abgebrochen hatten, befanden sich in der Nachbeobachtungsphase der Phase II der Studie SPRINT.

Tabelle 9: Allgemeine Angaben zum Studienverlauf der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitte: 29.06.2018 und 29.03.2019)

Studie SPRINT	Phase I	Phase II	Phase II
Datenschnitt	29.06.2018	29.06.2018	29.03.2019
ITT-Population ¹⁾ , n (%)	6 (100)	50 (100)	50 (100)
FAS ¹⁾ , n (%)	6 (100)	50 (100)	50 (100)
Sicherheitspopulation ¹⁾ , n (%)	6 (100)	50 (100)	50 (100)
Abbruch der Einnahme der Studienmedikation, n (%)	0	16 (32)	20 (40)
UE		6 (12)	6 (12)
Krankheitsprogression		3 (6)	5 (10)
Entscheidung des ärztlichen Prüfpersonals		3 (6)	4 (8)
Tod		0	0
Protokollverletzung		1 (2)	1 (2)
Entzug der Einverständniserklärung		1 (2)	2 (4)
Abschluss der Behandlungsperiode		2 (4)	2 (4)
Personen mit Wiederaufnahme der Behandlung nach Abbruch der Einnahme der Studienmedikation, n (%)	0	0	2 (4)
Abbruch der Studienteilnahme, n (%)	0	4 (8)	6 (12)
Lost to Follow-up		1 (2)	1 (2)
Entzug der Einverständniserklärung		2 (4)	3 (6)
andere		1 (2)	2 (4)
Mediane Behandlungsdauer ²⁾ , Tage (min; max)	1.597 (1.583; 1.617)	801,5 (28; 1053)	1.027,5 (28; 1326)
Mediane Beobachtungsdauer, Tage (min; max)	k. A.	k. A.	k. A.

- ¹⁾ Definition, siehe Kapitel 2.4 Statistische Methoden.
²⁾ Datum der letzten Dosis – Datum der ersten Dosis +1.

Abkürzungen: FAS: Full Analysis Set; ITT: Intention to Treat; k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Demographische und krankheitsspezifische Charakteristika zu Baseline sind in Tabelle 10 zusammengefasst. Das mediane Alter zu Baseline lag bei 6,8 Jahren in Phase I und bei 10,2 Jahren in Phase II. Die Studienpopulation war überwiegend europäischer Abstammung. Zwischen Diagnose der NF1-Erkrankung und Beginn der Behandlung mit Selumetinib lagen im Median ca. 8 Jahre. Zwischen Diagnose des PN und Beginn der Behandlung mit Selumetinib lagen im Median ca. 6 Jahre.

Innerhalb der Studie SPRINT wurde das klinisch relevanteste, inoperable PN, welches mittels volumetrischer 3D-MRT-Messung erfasst werden konnte, als Ziel-PN definiert. Daneben konnten maximal 2 weitere, inoperable PN als zusätzliche Nicht-Zielläsionen für weitere 3D-MRT-Untersuchungen ausgewählt werden. Angaben zur Anzahl weiterer PN außer dem Ziel-PN konnten in den Studienunterlagen nicht identifiziert werden. Für Phase II wird lediglich im Studienbericht angegeben, dass keine Nicht-Ziel-PN definiert wurden. Die dem Ziel-PN zugeschriebenen Morbiditäten in Phase II waren „Entstellung“ (88 %), „Motordysfunktion“ (66 %), „Blockierung der Atemwege“ (32 %), „Störung der Darm- und Blasenfunktion“ (20 %), „Einschränkung des Sehens“ (20 %) und „Andere Dysfunktion“ (24 %); für Phase I liegen keine Angaben dazu vor. 52 % der Studienteilnehmenden der Phase II der Studie SPRINT gaben an Schmerzen aufgrund der Ziel-PN zu haben. Das Ziel-PN verursachte dabei 1 bis 4 Morbiditäten (Median = 3).

Tabelle 10: Charakterisierung der Studienpopulation der Studie SPRINT; FAS Phase I und Phase II, Stratum 1 (Datenschnitt: 29.06.2018)

Studie SPRINT Charakterisierung der Studienpopulation	Phase I N = 6	Phase II N = 50
<i>Alter (Jahre)</i>		
MW (SD)	7,82 (4,1)	10,3 (3,9)
Median (min; max)	6,8 (3,0; 14,1)	10,2 (3,5; 17,4)
<i>Geschlecht, n (%)</i>		
männlich	3 (50,0)	30 (60,0)
weiblich	3 (50,0)	20 (40,0)
<i>Abstammung (genetisch) n (%)</i>		
europäisch	6 (100)	42 (84,0)
afrikanisch	0	4 (8,0)
asiatisch	0	1 (2,0)
unbekannt	0	3 (6,0)
<i>Karnofsky Performance Status n (%)</i>	0	n = 3
90 – Fähig zu normaler Aktivität, kaum oder geringe Symptome		1 (33,3)
80 – Normale Aktivität mit Anstrengung möglich, deutliche Symptome		2 (66,7)
<i>Lansky Performance Status, n (%)</i>	n = 6	n = 47
100 – Volle Aktivität, normal	2 (33,3)	6 (12,8)
90 – Kleine Einschränkungen in körperlich schweren Aktivitäten	2 (33,3)	24 (51,1)
80 – Aktiv, aber schnelle Ermüdung	1 (16,7)	13 (27,7)
70 – Stärkere Einschränkung, weniger Zeit für spielerische Aktivitäten	1 (16,7)	4 (8,5)

Studie SPRINT Charakterisierung der Studienpopulation	Phase I N = 6	Phase II N = 50
<i>NF1-Diagnosekriterien, n (%)</i>		
6 oder mehr Café-au-lait-Flecken	6 (100)	49 (98,0)
Freckling in der Achselhöhle oder in der Leistenengegend	2 (33,3)	42 (84,0)
Optikusgliom	4 (66,7)	22 (44,0)
2 oder mehr Lisch-Knötchen	1 (16,7)	19 (38,0)
Typische Knochenveränderungen	0	12 (24,0)
Verwandte/r ersten Grades mit NF1	2 (33,3)	15 (30,0)
Andere Manifestation	2 (33,3)	32 (64,0)
<i>Krankheitsdauer (Jahre), Median (min; max)</i>		
Zeit ab Diagnose von NF1 bis Beginn der Behandlung	n = 4 8,3 (2,5; 13,6)	n = 48 8,0 (2,0; 16,5)
Zeit ab Diagnose des PN bis Beginn der Behandlung	n = 5 6,4 (4,7; 13,6)	n = 45 6,3 (0,7; 16,5)
<i>Klassifikation des Ziel-PN, n (%)</i>		
typisch	k. A.	45 (90,0)
nodulär		4 (8,0)
einzelnen nodulär		1 (2,0)
<i>Status des Ziel-PN, n (%)</i>		
progressiv	k. A.	21 (42,0)
nicht-Progressiv		15 (30,0)
unbekannt		14 (28,0)
<i>Ziel-PN zugeschriebene Morbidität, n (%)</i>		
Orbital (Sehen)		10 (20,0)
Entstellung		44 (88,0)
Atemwege	k. A.	16 (32,0)
Motordysfunktion		33 (66,0)
Darm- oder Blasendysfunktion		10 (20,0)
Andere Dysfunktion		12 (24,0)
<i>Schmerzen aufgrund des Ziel-PN, n (%)</i>		
ja	k. A.	26 (52,0)
nein		22 (44,0)
fehlend		2 (4,0)
<i>Anzahl an durch PN verursachte Morbiditäten</i>		
Median (min; max)	k. A.	3 (1;4)
<i>Lokalisation des Ziel-PN, n (%)</i>		
Kopf	-	9 (18,0)
Kopf/Hals	-	8 (16,0)
Hals	1 (17)	-
Hals/Rumpf	-	12 (24,0)
Rumpf	1 (17)	5 (10,0)
Rumpf/Extremitäten	-	12 (24,0)
Extremitäten	-	4 (8,0)
Auge	1 (17)	-
Kopf – Gesicht oder Hals – unspezifiziert	2 (33)	-
Pelvis	1 (17)	-
<i>Volumen des Ziel-PN (ml)</i>		
MW (SD)	1.131 (1.058)	837 (925)
Median (min; max)	1.076 (29; 2.238)	488 (6; 3.820)
<i>Teilnehmende mit vorangegangener Therapie der PN, n (%)</i>		
Medikamentöse Therapie	6 (100)	39 (78,0)
Operation	5 (83,3)	31 (62,0)
Strahlentherapie	3 (50,0)	28 (56,0)
	0	1 (2,0)

Studie SPRINT Charakterisierung der Studienpopulation	Phase I N = 6	Phase II N = 50
Anzahl an PN Median (min; max)	k. A.	k. A.
Anzahl an Nicht-Ziel-PN Median (min; max)	k. A.	0

Abkürzungen: FAS: Full Analysis Set; k. A.: keine Angabe; MW: Mittelwert; NF1: Neurofibromatose Typ 1; PN: Plexiformes Neurofibrom; SD: Standardabweichung.

Exposition mit der Studienmedikation

Angaben zur Exposition mit Selumetinib in der Studie SPRINT sind in Tabelle 11 zusammengefasst. Selumetinib wurde im Median an 944 Tagen in Phase II (Datenschnitt 29.03.2019) und an 1.489 Tagen in Phase I eingenommen. Für alle Studienteilnehmenden wurden Dosisunterbrechungen registriert. Ca. 80 % der Studienpopulation (n = 5 in Phase I, n = 41 in Phase II) wiesen ein UE auf, welches zu einer Dosisunterbrechung führte. Angaben zur Anzahl an Teilnehmenden mit einer Dosisreduktionen bzw. Dosiserhöhung konnten in den Studienunterlagen nicht identifiziert werden. In Phase I wiesen 3 Teilnehmende (50%) ein UE auf, das zu einer Dosisreduktion führte und in Phase II 13 Teilnehmende (26 %).

Tabelle 11: Angaben zur Exposition mit der Studienmedikation Selumetinib; FAS der Studie SPRINT Phase I und Phase II, Stratum 1 (Datenschnitte 29.06.2018 und 29.03.2019)

Studie SPRINT Exposition mit der Studienmedikation	Phase I N = 6	Phase II N = 50	Phase II N = 50
Datenschnitt	29.06.2018	29.06.2018	29.03.2019
<i>Dauer der Exposition (Tage)¹⁾</i>			
MW (SD)	1.473 (55)	670 (241)	825 (340)
Median (min; max)	1.489 (1.389; 1.539)	722 (26; 1022)	944 (26; 1290)
<i>Anzahl an Zyklen</i>			
MW (SD)	k. A.	k. A.	k. A.
Median (min; max)	k. A.	k. A.	k. A.
<i>Dosisanpassungen, n (%)</i>			
Unterbrechung	6 (100)	50 (100)	50 (100)
Erhöhung	k. A.	k. A.	k. A.
Reduktion	k. A.	k. A.	k. A.
<i>Durchschnittliche Dosis (mg/Tag)</i>			
MW (SD)	k. A.	k. A.	k. A.
Median (min; max)	k. A.	k. A.	k. A.

¹⁾ Summe der Tage, an denen die Studienmedikation eingenommen wurde.

Abkürzungen: FAS: Full Analysis Set; k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung.

Begleitmedikation

Angaben zur Begleitmedikation sind in Tabelle 12 dargestellt.

Tabelle 12: Begleitmedikation mit einer Häufigkeit > 20 %; FAS der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt: 29.06.2018)

Studie SPRINT Begleitmedikation mit einer Häufigkeit > 20 %; ATC-Klasse	Phase I N = 6 n (%)	Phase II N = 50 n (%)
Serotonin-5HT3-Rezeptorantagonisten	3 (50,0)	31 (62,0)
Anilide	3 (50,0)	31 (62,0)
Andere Antibiotika zur topischen Anwendung	4 (66,7)	29 (58,0)
Propionsäurederivate	-	25 (50,0)
Antiinfektiva zur Behandlung der Akne	3 (50,0)	24 (48,0)
Cephalosporine der 1. Generation	4 (66,7)	22 (44,0)
Piperazin-Derivate	-	20 (40,0)
Corticosteroide, mittelstark wirksam (Gruppe II)	5 (83,3)	19 (38,0)
Penicilline mit erweitertem Wirkungsspektrum	3 (50,0)	19 (38,0)
Andere Antihistaminika zur systemischen Anwendung	3 (50,0)	18 (36,0)
Vitamin D und Analoga	-	16 (32,0)
Andere Dermatika	2 (33,3)	16 (32,0)
Imidazol- und Triazol-Derivate	3 (50,0)	15 (30,0)
Vaseline und Fett-haltige Mittel	-	15 (30,0)
Zentral wirkende Sympathomimetika	3 (50,0)	15 (30,0)
Aminoalkylether	2 (33,3)	15 (30,0)
Osmotisch wirkende Laxanzien	2 (33,3)	14 (28,0)
Corticosteroide, schwach wirksam (Gruppe I)	3 (50,0)	14 (28,0)
Andere Analgetika und Antipyretika	2 (33,3)	14 (28,0)
Tetracykline	-	13 (26,0)
Kombinationen von Penicillinen, inkl. Beta-Lactamase-Inhibitoren	3 (50,0)	13 (26,0)
Selektive Beta2-Adrenozeptoragonisten	-	13 (26,0)
Corticosteroide, stark wirksam (Gruppe III)	-	12 (24,0)
Peroxide	-	11 (22,0)
Melatoninrezeptoragonisten	3 (50,0)	11 (22,0)
Histamin-H2-Rezeptorantagonisten	2 (33,3)	10 (20,0)
Lincosamide	2 (33,3)	10 (20,0)
Kombination von Sulfonamiden und Trimethoprim, inkl. Derivaten	2 (33,3)	-
Impfstoffe	3 (50,0)	-
Antipropulsiva	2 (33,3)	-

Studie SPRINT Begleitmedikation mit einer Häufigkeit > 20 %; <i>ATC-Klasse</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Antiinfektiva	2 (33,3)	-
Corticosteroide und Antiinfektiva in Kombination	2 (33,3)	-
Cephalosporine der 3. Generation	2 (33,3)	-
Zink	2 (33,3)	-
Harnstoffmittel	2 (33,3)	-
Expectorantien	2 (33,3)	-
Neuraminidase-Inhibitoren	2 (33,3)	-

Abkürzungen: ATC: Anatomisch-Therapeutisch-Chemisch (Klassifikationssystem); FAS: Full Analysis Set.

Folgetherapien

Angaben zu möglichen Folgetherapie konnten in den Studienunterlagen nicht identifiziert werden.

3.2 Mortalität

Es wurde kein Todesfall in Phase I bis zum Datenschnitt vom 29.06.2018 und kein Todesfall in Phase II bis zum Datenschnitt vom 29.03.2019 berichtet.

3.3 Morbidität

Schmerz

Responderanalysen für eine Verbesserung bzw. Verschlechterung um 2 Punkte des mittels NRS-11 gemessenen schlimmsten Schmerzes innerhalb der letzten 7 Tage sind in Tabelle 13 für die (Teil-) Population der 8- bis 18-Jährigen in der Studie SPRINT dargestellt. Die Darstellung erfolgt bis zum Erhebungszeitpunkt vor Zyklus 13. Zum nächsten Erhebungszeitpunkt vor Zyklus 25 lag die Rücklaufquote für alle Schmerz-Operationalisierungen (Schmerz Patienten- bestimmtes PN; Schmerz Ziel-PN; Gesamttumorschmerz; Andere Schmerzen) unter 70 % und die Ergebnisse werden nicht berichtet. Für „Andere Schmerzen“ wurde zu keinem Erhebungszeitpunkt nach Baseline eine Rücklaufquote von über 70 % erreicht. Der Grund dafür ist unklar.

Tabelle 13: Responderanalyse NRS-11 für den schlimmsten Schmerz; FAS der Studie SPRINT, Phase II, Stratum 1 im Alter von 8 bis 18 Jahren
(Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalyse NRS-11 ¹⁾	Baseline N = 34		Vor Zyklus 3 N = 34		Vor Zyklus 5 N = 34		Vor Zyklus 9 N = 34		Vor Zyklus 13 N = 34	
	n	MW (SD) Median (min; max)	n	n (% [95%-KI] ²⁾	n	n (% [95%-KI] ²⁾	n	n (% [95%-KI] ²⁾	n	n (% [95%-KI] ²⁾
Schmerz Patienten-bestimmtes PN Verbesserung um ≥ 2 Punkte ³⁾ Verschlechterung um ≥ 2 Punkte ⁴⁾	33	3,4 (3,2) 3,0 (0,0; 10;0)	31	13 (41,9 [24,5; 60,9]) 1 (3,2 [0,1; 16,7])	31	20 (64,5 [45,4; 80,8]) 2 (6,5 [0,8; 21,4])	31	18 (58,1 [39,1; 75,5]) 0 (0 [0; 11,2])	29	17 (58,6 [38,9; 76,5]) 0 (0 [0; 11,9])
Schmerz Ziel-PN Verbesserung um ≥ 2 Punkte ⁵⁾ Verschlechterung um ≥ 2 Punkte ⁶⁾	26	3,2 (3,1) 2,5 (0,0; 10;0)	25	9 (36,0 [18,0; 57,5]) 1 (4,0 [0,1; 20,4])	25	12 (48,0 [27,8; 68,7]) 2 (8,0 [1,0; 26,0])	25	13 (52,0 [31,3; 72,2]) 0 (0 [0; 13,7])	24	12 (50,0 [29,1; 70,9]) 0 (0 [0; 14,2])
Gesamtumerschmerz Verbesserung um ≥ 2 Punkte ⁷⁾ Verschlechterung um ≥ 2 Punkte ⁸⁾	28	3,0 (3,1) 2,0 (0,0; 10;0)	25	10 (40,0 [21,1; 61,3]) 2 (8,0 [1,0; 26,0])	24	11 (45,8 [25,6; 67,2]) 3 (12,5 [2,7; 32,4])	26	12 (46,2 [26,6; 66,6]) 0 (0 [0; 13,2])	23	- ⁹⁾
Andere Schmerzen Verbesserung um ≥ 2 Punkte ¹⁰⁾ Verschlechterung um ≥ 2 Punkte	29	3,9 (2,8) 5,0 (0,0; 8;0)	22	- ⁹⁾	19	- ⁹⁾	21	- ⁹⁾	18	- ⁹⁾

¹⁾ Skalenspannweite von 0 „keine Schmerzen“ bis 10 für „schlimmste vorstellbare Schmerzen“.

²⁾ Prozentualer Anteil bezogen auf alle Personen, die geantwortet haben; Konfidenzintervalle wurden mit der Clopper-Pearson-Methode für binomiale Verteilungen berechnet.

³⁾ Zu Baseline wiesen 11 Studienteilnehmende (33,3 %) einen Wert < 2 auf und konnten sich demzufolge nicht um 2 Punkte verbessern.

⁴⁾ Zu Baseline wiesen 4 Studienteilnehmende (12,1 %) einen Wert > 8 auf und konnten sich demzufolge nicht um 2 Punkte verschlechtern.

⁵⁾ Zu Baseline wiesen 10 Studienteilnehmende (38,4 %) einen Wert < 2 auf und konnten sich demzufolge nicht um 2 Punkte verbessern.

⁶⁾ Zu Baseline wiesen 3 Studienteilnehmende (11,5 %) einen Wert > 8 auf und konnten sich demzufolge nicht um 2 Punkte verschlechtern.

⁷⁾ Zu Baseline wiesen 13 Studienteilnehmende (46,4 %) einen Wert < 2 auf und konnten sich demzufolge nicht um 2 Punkte verbessern.

⁸⁾ Zu Baseline wies 1 Studienteilnehmende/r (3,6 %) einen Wert > 8 auf und konnten sich demzufolge nicht um 2 Punkte verschlechtern.

⁹⁾ Keine Ergebnisdarstellung, da Rücklaufquote unter 70 %.

¹⁰⁾ Zu Baseline wiesen 9 Studienteilnehmende (31,0 %) einen Wert < 2 auf und konnten sich demzufolge nicht um 2 Punkte verbessern.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; MW: Mittelwert; NRS-11: Numerical Rating Scale; PN: Plexiformes Neurofibrom; SD: Standardabweichung.

Tabelle 14: Responderanalysen zur Schmerzlinderung für „Schmerz des Ziel-PN“; FAS der Studie SPRINT, Phase II, Stratum 1 im Alter von 8 bis 18 Jahren (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen zur Schmerzlinderung	Schmerzlinderung gemäß Definition 1¹⁾ N = 34	Schmerzlinderung gemäß Definition 2²⁾ N = 34
Vor Zyklus 3, n/N (% [95%-KI] ³⁾)	9/15 (60,0 [32,3; 83,7])	10/17 (58,8 [32,9; 81,6])
Vor Zyklus 5, n/N (% [95%-KI] ³⁾)	12/15 (80,0 [51,9; 95,7])	13/17 (76,5 [50,1; 93,2])
Vor Zyklus 9, n/N (% [95%-KI] ³⁾)	12/15 (80,0 [51,9; 95,7])	13/17 (76,5 [50,1; 93,2])
Vor Zyklus 13, n/N (% [95%-KI] ³⁾)	11/14 (78,6 [49,2; 95,3])	11/16 (68,8 [41,3; 89,0])
Innerhalb des gesamten Beobachtungszeitraums		
Personen mit Schmerzlinderung ⁴⁾ , n (%)	14 (41,2)	15 (44,1)
Mediane Dauer bis zur Schmerzlinderung (Monate) [95%-KI] ⁵⁾	1,87 [1,74; 3,55]	3,71 [1,87; n. b.]

¹⁾ Schmerzlinderung ist definiert nur für Personen mit einem Score ≥ 2 Punkte zu Baseline: Schmerzlinderung entspricht einer Abnahme von ≥ 2 Punkten ohne Erhöhung des Schmerzmittel-Scores.

²⁾ Schmerzlinderung ist definiert für Personen mit einem Score ≥ 2 Punkte zu Baseline ODER für Personen mit einem Score < 2 Punkte und einem Schmerzmittel-Score ≥ 2 Punkte zu Baseline: Schmerzlinderung entspricht einer Abnahme von ≥ 2 Punkten ohne Erhöhung des Schmerzmittel-Scores ODER für Personen mit einem Score < 2 Punkte und einem Schmerzmittel-Score ≥ 2 Punkte zu Baseline entspricht eine Schmerzlinderung einem stabilen oder verminderten NRS-11-Score bei gleichzeitiger Abnahme des Schmerzmittel-Scores um mindestens eine Stufe.

³⁾ KI wurden mit der Clopper-Pearson-Methode für binomiale Verteilungen berechnet.

⁴⁾ Prozentualer Anteil bezogen auf das FAS.

⁵⁾ Berechnet mittels Kaplan-Meier Methode; prozentuale Anteil bezogen auf alle Personen.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; n. b.: nicht berechenbar; NRS-11: Numerical Rating Scale; PN: Plexiformes Neurofibrom;

Motorfunktion

Grooved Pegboard Test

Laut Studienprotokoll sollte die Erhebung des Grooved Pegboard Tests bei Kindern und Jugendlichen ab einem Alter von 5 Jahren mit einem PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks erfolgen. Die Ergebnisdarstellung im Studienbericht erfolgt für die Population mit Motor-PN-bezogener Morbidität. Die Anzahl an Personen mit PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks wird im Studienbericht mit 25 angegeben. Die Teilnahmequoten werden bezogen auf diese Population berechnet. Ab der Studienvsiste vor Zyklus 25 lag die Teilnahmequote unter 70 %, sodass die Ergebnisse nicht für die Nutzenbewertung herangezogen werden. Es werden sowohl die Rohwerte für die für den Test benötigte Zeit als auch die Z-Scores dargestellt (Tabelle 15). Die Auswertung auf Basis der Rohwerte ist dadurch limitiert, dass Kinder bis einschließlich 8 Jahren nur 2/5 des Tests durchführten und die Werte zwischen den Altersgruppen damit schlecht vergleichbar sind. Dieses Problem wird durch die Bildung und Auswertung altersabhängiger Z-Scores umgangen. Allerdings fehlen Angaben zur verwendeten Normstichprobe, sodass die Validität und Repräsentativität der Z-Score Werte nicht beurteilt werden kann.

Tabelle 15: Veränderungen im Grooved Pegboard Test; FAS ab 5 Jahren mit einem PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Veränderungen im Grooved Pegboard Test	n	Zeit [s] Median (min; max)	Z-Score für Zeit¹⁾ Median (min; max)	Fallengelassene Pegs Median (min, max)
Alle Teilnehmenden – dominante Hand (N = 25)				
Baseline	25	86,1 (41,0; 164,7)	1,1 (-1,0; 12,4)	1 (0; 6)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	24	1,9 (-96,7; 63,0)	0,06 (-10,7; 4,3)	0 (-4; 11)
Vor Zyklus 9	23	-10,1 (-117,0; 350,4)	-0,54 (-13,0; 33,7)	0 (-4; 11)
Vor Zyklus 13	22	-2,46 (-61,3; 202,4)	-0,35 (-12,2; 19,5)	0 (-3; 6)
Alle Teilnehmenden – nicht-dominante Hand (N = 25)				
Baseline	24	100,1 (46,0; 505,9)	2,2 (-0,9; 39,2)	1 (0; 18)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	23	-7,0 (-217,6; 152,4)	-0,35 (-18,6; 4,5)	-1 (-16; 1)
Vor Zyklus 9	23	-13,2 (-178,4; 92,2)	-1,19 (-9,4; 5,7)	0 (-17; 4)
Vor Zyklus 13	21	-11,7 (-121,9; 134,8)	-0,62 (-11,0; 10,1)	0 (-16; 5)
Teilnehmende mit unilateraler PN – beeinträchtigte Hand (N = 17)				
Baseline	16	88,8 (41,0; 505,9)	1,46 (-0,9; 39,2)	1,5 (0; 18)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	16	-6,3 (-205,9; 152,4)	-0,38 (-18,6; 4,5)	-0,5 (-16; 7)
Vor Zyklus 9	15	-1,0 (-102,0; 63,1)	-0,54 (-3,7; 5,7)	-1 (-17; 3)
Vor Zyklus 13	15	-2,7 (-121,9; 70,4)	-0,50 (-11,0; 0,6)	-1 (-16; 2)
Teilnehmende mit unilateraler PN – nicht-beeinträchtigte Hand (N = 17)				
Baseline	17	92,5 (41,0; 300,0)	1,78 (-1,0; 12,4)	1 (0; 8)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	17	-0,01 (-217,6; 63,0)	-0,10 (-10,9; 4,3)	0 (-7; 3)
Vor Zyklus 9	17	-22,8 (-178,4; 350,4)	-1,23 (-13,0; 33,7)	0 (-7; 11)
Vor Zyklus 13	16	-3,6 (-111,0; 202,4)	-0,47 (-12,2; 19,5)	0 (-3; 6)
Teilnehmende mit bilateraler PN – dominante Hand (N = 8)				
Baseline	8	88,9 (51,0; 137,0)	1,93 (-0,1; 6,8)	0 (0; 1)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	7	-14,0 (-36,0; 3,0)	-0,57 (-3,5; 0,3)	0 (-1; 1)
Vor Zyklus 9	7	-12,0 (-33,8; 84,0)	-0,55 (-3,3; 3,4)	0 (-1; 1)
Vor Zyklus 13	6	-18,6 (-43,1; 84,0)	-1,28 (-4,1; 11,3)	0 (-1; 0)
Teilnehmende mit bilateraler PN – nicht-dominante Hand (N = 8)				
Baseline	8	94,2 (46,0; 209,8)	2,14 (-0,2; 6,2)	0 (0; 4)
<i>Veränderung zu Baseline</i>				
Vor Zyklus 5	7	-2,0 (-55,7; 26,0)	-0,18 (-2,2; 1,2)	0 (-2; 1)
Vor Zyklus 9	7	4,9 (-98,9; 92,2)	0,44 (-3,2; 4,2)	0 (-1; 3)
Vor Zyklus 13	6	5,67 (-12,9; 134,8)	0,92 (-1,2; 10,1)	1 (-1; 5)

¹⁾ Alters- (und geschlechts-)standardisierter Z-Score.

Abkürzungen: FAS: Full Analysis Set; PN: Plexiformes Neurofibrom;

PROMIS

Die Darstellung der Ergebnisse zur körperlichen Funktionsfähigkeit erhoben mittels PROMIS ist auf den Zeitraum bis zur Studienvsiste vor Zyklus 13 beschränkt, da danach die Rücklaufquote bei unter 70 %, bezogen auf das FAS mit Motor-PN bezogener Morbidität, liegt. Die körperliche Funktionsfähigkeit wird für Kinder ≥ 8 Jahre direkt über die Befragung der Kinder erhoben und ist damit über die Patientenversion der PROMIS-Scores „Mobilität“ und „Obere Extremitäten“ abgebildet (Tabelle 16). Für die Eltern-/Betreuerversion des PROMIS liegen keine separaten Auswertungen für Kinder < 8 Jahre vor. Die Auswertung der Eltern-/Betreuerversion innerhalb des FAS ab 5 Jahren wird nicht in der Nutzenbewertung dargestellt.

Tabelle 16: Veränderungen der PROMIS-Skalen „Mobilität“ und „Obere Extremitäten“; FAS im Alter von 8 bis 18 Jahren mit Motor-PN-bezogener Morbidität der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Veränderungen PROMIS	PROMIS patientenberichtet „Mobilität“ ⁽¹⁾ N = 24		PROMIS patientenberichtet „Obere Extremitäten“ ⁽¹⁾ N =24	
	n	LSM [95%-KI], p-Wert	n	LSM [95%-KI], p-Wert
Baseline, MW (SD)	23	46,57 (6,54)	22	45,95 (12,91)
<i>Veränderung zu Baseline²⁾</i>				
Vor Zyklus 3	21	0,69 [-2,44; 3,82], 0,65	21	0,34 [-2,64; 3,32], 0,81
Vor Zyklus 5	22	1,83 [-0,69; 4,35], 0,15	21	-0,09 [-2,10; 1,92], 0,93
Vor Zyklus 9	22	1,01 [-2,06; 4,08], 0,50	21	-1,40 [-4,82; 2,03], 0,40
Vor Zyklus 13	20	1,75 [-0,70; 4,19], 0,15	19	1,76 [-0,88; 4,39], 0,18

¹⁾ Dargestellt sind T-Scores basierend auf der US-Allgemeinbevölkerung mit einem Mittelwert von 50 und einer Standardabweichung von 10. Höhere Werte entsprechen einer besseren körperlichen Funktion.

²⁾ Berechnet mittels MMRM mit Messwert für die jeweiligen Zeitpunkte, Baseline-Wert, Alter, Anzahl an Morbiditäten zu Baseline und Baseline x Zeitpunkt-Interaktion.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; LSM: Least Square Mean; MMRM: Mixed Effect Model Repeat Measures; MW: Mittelwert; PN: Plexiformes Neurofibrom; PROMIS: Patient-Reported Outcomes Measurement Information System; SD: Standardabweichung.

Sehschärfe

Insgesamt 10 Personen wiesen ein orbitales PN auf. Für 5 dieser Personen war keine Verbesserung der Sehschärfe möglich, da das durch das PN betroffene Auge entfernt worden war ($n = 2$), das Sehen auf Licht- und Lagewahrnehmung begrenzt war ($n = 1$) oder die Betroffenen keine Lichtwahrnehmung ($n = 2$) mehr aufwiesen. Dies wurde bei der Berechnung der Rücklaufquoten berücksichtigt. Zu den Studienvisiten 5 und 9 lagen Ergebnisse bezüglich Verbesserung und Verschlechterung nur für 3 Teilnehmende vor. Der Rücklauf wird als zu niedrig erachtet und die Ergebnisse nicht dargestellt.

Tabelle 17: Responderanalysen für Verbesserungen und Verschlechterungen der Sehschärfe; FAS mit PN der Augenhöhle der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen Sehschärfe	Von PN-betroffenes Auge HOTV N = 10		Von PN-nicht-betroffenes Auge HOTV N = 10	
	n	n (% ¹⁾)	n	n (% ¹⁾)
Baseline (logMAR) MW (SD)	5	0,54 (0,38)	7	0,01 (0,11)
Vor Zyklus 13 Verbesserung um $\geq 0,2$ logMAR Verschlechterung um $\geq 0,2$ logMAR	4	0 (0) 2 (50)	6	0 (0) 1 (16,7)
Vor Zyklus 25 Verbesserung um $\geq 0,2$ logMAR Verschlechterung um $\geq 0,2$ logMAR	4	0 (0) 1 (25)	6	0 (0) 1 (16,7)

¹⁾ Prozentualer Anteil bezogen auf Patientinnen und Patienten ohne fehlende Daten zur jeweiligen Studienvisite.

Abkürzungen: FAS: Full Analysis Set; logMAR: Logarithm of the Minimum Angle of Resolution; MW: Mittelwert; PN: Plexiformes Neurofibrom; SD: Standardabweichung.

Symptomcheckliste

Die Ergebnisse zur Veränderung der mit der Symptomcheckliste erfassten Symptome ist in Tabelle 18 zusammengefasst. Die Rücklaufquote beläuft sich ab dem Erhebungszeitpunkt vor Zyklus 25 auf unter 70 % ($N = 34$, Rücklaufquote 68 %), weshalb keine Darstellung der Ergebnisse ab diesem Zeitpunkt erfolgt.

Für die Symptome „Schlafprobleme“, „Kopfschmerzen“, „Würgen“, „Schnarchen“, „Häufiges Aufwachen in der Nacht“, „Husten“, „Kurzatmigkeit bei Belastung“, „Schwäche“, „Muskelschmerzen“ und „Kribbeln“ wird häufiger (≥ 10 Prozentpunkte Unterschied zum Erhebungszeitpunkt vor Zyklus 13) von einer Verbesserung der Symptome als einer Verschlechterung der Symptome berichtet. Umgekehrt wird für die Symptome „Anschwellen der Füße/Hände“ und „Diarrhö“ zum Zeitpunkt vor Zyklus 13 häufiger (≥ 10 Prozentpunkte Unterschied) über eine Verschlechterung der Symptome berichtet als über eine Verbesserung. Zu früheren Erhebungszeitpunkten wurden für die Symptome „Übelkeit“, „Erhöhter Appetit“, „Unterleibsschmerzen“, „Erbrechen“ und „Schwindel“ ebenfalls häufiger Verschlechterungen als Verbesserungen berichtet.

Tabelle 18: Responderanalysen der Symptomatik; FAS der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
n	50	49	48	47	44
<i>Müdigkeit/Fatigue</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 17 (34,0) 2: 21 (42,0) 3: 5 (10,0) 4: 4 (8,0) 5: 3 (6,0)	15 (30,6) 18 (36,7)	18 (37,5) 14 (29,2)	17 (36,2) 15 (31,9)	17 (38,6) 13 (29,5)
<i>Schlafprobleme</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 22 (44,0) 2: 12 (24,0) 3: 7 (14,0) 4: 7 (14,0) 5: 2 (4,0)	20 (40,8) 9 (18,4)	18 (37,5) 4 (8,3)	19 (40,4) 8 (17,0)	17 (38,6) 6 (13,6)
<i>Verringerter Appetit</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 34 (68,0) 2: 6 (12,0) 3: 3 (6,0) 4: 5 (10,0) 5: 2 (4,0)	13 (26,5) 8 (16,3)	13 (27,1) 6 (12,5)	11 (23,4) 7 (14,9)	10 (22,7) 6 (13,6)
<i>Erhöhter Appetit</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 36 (72,0) 2: 9 (18,0) 3: 3 (6,0) 4: 1 (2,0) 5: 1 (2,0)	8 (16,3) 16 (32,7)	6 (12,5) 21 (43,8)	8 (17,0) 7 (14,9)	10 (22,7) 10 (22,7)
<i>Kopfschmerzen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 28 (56,0) 2: 13 (26,0) 3: 7 (14,0) 4: 2 (4,0) 5: 0	14 (28,6) 8 (16,3)	15 (31,3) 9 (18,8)	16 (34,0) 10 (21,3)	16 (36,4) 6 (13,6)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Sehstörungen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 43 (86,0) 2: 5 (10,0) 3: 2 (4,0) 4: 0 5: 0	5 (10,2) 3 (6,1)	6 (12,5) 6 (12,5)	5 (10,6) 2 (4,3)	5 (11,4) 1 (2,3)
<i>Vermindertes Hörvermögen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 47 (94,0) 2: 1 (2,0) 3: 2 (4,0) 4: 0 5: 0	3 (6,1) 2 (4,1)	3 (6,3) 4 (8,3)	3 (6,4) 1 (2,1)	3 (6,8) 4 (9,1)
<i>Wunde Stellen im Mund</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 45 (90,0) 2: 4 (8,0) 3: 1 (2,0) 4: 0 5: 0	4 (8,2) 5 (10,2)	5 (10,4) 4 (8,3)	4 (8,5) 7 (14,9)	5 (11,4) 6 (13,6)
<i>Schwierigkeiten beim Schlucken</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 3 (6,0) 3: 1 (2,0) 4: 0 5: 2 (4,0)	5 (10,2) 0	5 (10,4) 0	5 (10,6) 0	5 (11,4) 1 (2,3)
<i>Würgen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 43 (86,0) 2: 5 (10,0) 3: 2 (4,0) 4: 0 5: 0	6 (12,2) 2 (4,1)	7 (14,6) 1 (2,1)	6 (12,8) 0	6 (13,6) 0
<i>Schnarchen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 31 (62,0) 2: 7 (14,0) 3: 4 (8,0) 4: 3 (6,0) 5: 5 (10,0)	14 (28,6) 6 (12,2)	12 (25,0) 5 (10,4)	16 (34,0) 4 (8,5)	15 (34,1) 3 (6,8)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Häufiges Aufwachen in der Nacht</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 26 (52,0) 2: 11 (22,0) 3: 7 (14,0) 4: 4 (8,0) 5: 2 (4,0)	16 (32,7) 10 (20,4)	16 (33,3) 6 (12,5)	17 (36,2) 7 (14,9)	17 (38,6) 7 (15,9)
<i>Husten</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 26 (52,0) 2: 17 (34,0) 3: 5 (10,0) 4: 1 (2,0) 5: 1 (2,0)	18 (36,7) 14 (28,6)	20 (41,7) 8 (16,7)	11 (23,4) 9 (19,1)	17 (38,6) 8 (18,2)
<i>Giemen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 2 (4,0) 3: 2 (4,0) 4: 1 (2,0) 5: 1 (2,0)	6 (12,2) 4 (8,2)	6 (12,5) 4 (8,3)	6 (12,8) 2 (4,3)	5 (11,4) 1 (2,3)
<i>Schwierigkeiten beim Atmen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 5 (10,0) 3: 1 (2,0) 4: 0 5: 0	5 (10,2) 2 (4,1)	3 (6,3) 2 (4,2)	6 (12,8) 1 (2,1)	3 (6,8) 1 (2,3)
<i>Schmerzen in der Brust</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 48 (96,0) 2: 1 (2,0) 3: 1 (2,0) 4: 0 5: 0	2 (4,1) 5 (10,2)	2 (4,2) 1 (2,1)	1 (2,1) 3 (6,4)	1 (2,3) 3 (6,8)
<i>Herzklopfen/-flattern</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 46 (92,0) 2: 3 (6,0) 3: 1 (2,0) 4: 0 5: 0	3 (6,1) 1 (2,0)	3 (6,3) 1 (2,1)	3 (6,4) 1 (2,1)	3 (6,8) 2 (4,5)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Kurzatmigkeit bei Belastung</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 36 (72,0) 2: 8 (16,0) 3: 5 (10,0) 4: 1 (2,0) 5: 0	12 (24,5) 3 (6,1)	10 (20,8) 4 (8,3)	10 (21,3) 5 (10,6)	9 (20,5) 3 (6,8)
<i>Kurzatmigkeit in Ruhe</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 47 (94,0) 2: 3 (6,0) 3: 0 4: 0 5: 0	3 (6,1) 0	3 (6,3) 0	3 (6,4) 2 (4,3)	2 (4,5) 0
<i>Anschwellen der FüÙe/Hände</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 49 (98,0) 2: 0 3: 0 4: 1 (2,0) 5: 0	1 (2,0) 4 (8,2)	1 (2,1) 5 (10,4)	1 (2,1) 4 (8,5)	0 6 (13,6)
<i>Unterleibsschmerzen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 33 (66,0) 2: 11 (22,0) 3: 5 (10,0) 4: 0 5: 1 (2,0)	7 (14,3) 14 (28,6)	6 (12,5) 7 (14,6)	6 (12,8) 10 (21,3)	10 (22,7) 10 (22,7)
<i>Sodbrennen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 46 (92,0) 2: 2 (4,0) 3: 2 (4,0) 4: 0 5: 0	3 (6,1) 3 (6,1)	3 (6,3) 2 (4,2)	2 (4,3) 1 (2,1)	2 (4,5) 2 (4,5)
<i>Übelkeit</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 38 (76,0) 2: 8 (16,0) 3: 4 (8,0) 4: 0 5: 0	7 (14,3) 17 (34,7)	5 (10,4) 14 (29,2)	6 (12,8) 13 (27,7)	8 (18,2) 10 (22,7)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Erbrechen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 6 (12,0) 3: 0 4: 0 5: 0	4 (8,2) 12 (24,5)	4 (8,3) 11 (22,9)	2 (4,3) 6 (12,8)	3 (6,8) 5 (11,4)
<i>Diarrhö</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 39 (78,0) 2: 6 (12,0) 3: 4 (8,0) 4: 0 5: 1 (2,0)	6 (12,2) 12 (24,5)	5 (10,4) 8 (16,7)	6 (12,8) 13 (27,7)	5 (11,4) 10 (22,7)
<i>Verstopfungen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 38 (76,0) 2: 7 (14,0) 3: 5 (10,0) 4: 0 5: 0	7 (14,3) 7 (14,3)	7 (14,6) 5 (10,4)	6 (12,8) 3 (6,4)	5 (11,4) 6 (13,6)
<i>Stuhlinkontinenz</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 45 (90,0) 2: 4 (8,0) 3: 1 (2,0) 4: 0 5: 0	4 (8,2) 5 (10,2)	2 (4,2) 1 (2,1)	3 (6,4) 4 (8,5)	4 (9,1) 3 (6,8)
<i>Schmerzen beim Urinieren</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 46 (92,0) 2: 3 (6,0) 3: 1 (2,0) 4: 0 5: 0	4 (8,2) 1 (2,0)	3 (6,3) 2 (4,2)	3 (6,4) 3 (6,4)	4 (9,1) 3 (6,8)
<i>Erhöhte Harnfrequenz / Erhöhter Harndrang</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 40 (80,0) 2: 5 (10,0) 3: 5 (10,0) 4: 0 5: 0	8 (16,3) 5 (10,2)	7 (14,6) 5 (10,4)	6 (12,8) 3 (6,4)	6 (13,6) 2 (4,5)

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Schwierigkeiten beim Beginn des Urinierens</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 4 (8,0) 3: 1 (2,0) 4: 1 (2,0) 5: 0	4 (8,2) 2 (4,1)	4 (8,3) 2 (4,2)	2 (4,3) 3 (6,4)	3 (6,8) 2 (4,5)
<i>Harninkontinenz</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 42 (84,0) 2: 2 (4,0) 3: 3 (6,0) 4: 2 (4,0) 5: 1 (2,0)	7 (14,6) 1 (2,1)	7 (14,9) 1 (2,1)	7 (14,9) 4 (8,5)	6 (13,6) 2 (4,5)
<i>Schwäche</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 34 (68,0) 2: 4 (8,0) 3: 6 (12,0) 4: 3 (6,0) 5: 3 (6,0)	14 (28,6) 2 (4,1)	16 (33,3) 2 (4,2)	14 (29,8) 5 (10,6)	15 (34,1) 2 (4,5)
<i>Muskelschmerzen</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 28 (56,0) 2: 9 (18,0) 3: 7 (14,0) 4: 4 (8,0) 5: 2 (4,0)	19 (38,8) 5 (10,2)	20 (41,7) 3 (6,3)	18 (38,3) 3 (6,4)	16 (36,4) 4 (9,1)
<i>Schwindel</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 44 (88,0) 2: 5 (10,0) 3: 1 (2,0) 4: 0 5: 0	2 (4,1) 7 (14,3)	3 (6,3) 4 (8,3)	2 (4,3) 4 (8,5)	4 (9,1) 5 (11,4)
<i>Taubheitsgefühl</i> Verbesserung ³⁾ Verschlechterung ⁴⁾	1: 45 (90,0) 2: 3 (6,0) 3: 1 (2,0) 4: 1 (2,0) 5: 0	5 (10,2) 0	3 (6,3) 1 (2,1)	5 (10,6) 1 (2,1)	3 (6,8) 0

Studie SPRINT, Phase II, Stratum 1 Responderanalysen der Symptomatik	Baseline ¹⁾ N = 50	Vor Zyklus 3 N = 50	Vor Zyklus 5 N = 50	Vor Zyklus 9 N = 50	Vor Zyklus 13 N = 50
	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)	n (% ²⁾)
<i>Kribbeln</i>	1: 41 (82,0)				
Verbesserung ³⁾	2: 5 (10,0)	8 (16,3)	6 (12,5)	6 (12,8)	6 (13,6)
Verschlechterung ⁴⁾	3: 2 (4,0)	2 (4,1)	2 (4,2)	2 (4,3)	0
	4: 1 (2,0)				
	5: 1 (2,0)				

¹⁾ Antwortkategorien: 1 – überhaupt nicht; 2 – ein wenig; 3– etwas; 4 – ziemlich viel; 5 – viel

²⁾ Prozentualer Anteil bezogen auf alle Personen, die geantwortet haben.

³⁾ Verbesserung um mindestens eine Antwortkategorie.

⁴⁾ Verschlechterung um mindestens eine Antwortkategorie.

Abkürzungen: FAS: Full Analysis Set.

Globalbeurteilung der klinischen Veränderung

Die Darstellung der patientenberichteten Ergebnisse der klinischen Veränderung erhoben mittels GIC ist auf den Zeitraum bis zur Studienvsiste vor Zyklus 13 beschränkt, da danach die Rücklaufquote, bezogen auf das FAS im Alter von 8 bis 18 Jahren, bei unter 70 % lag. Die klinische Veränderung wird für Kinder ≥ 8 Jahre direkt über die Befragung der Kinder erhoben und ist damit über die Patientenversion des GIC abgebildet (Tabelle 19). Für die Eltern-/Betreuerversion des GIC liegen keine separaten Auswertungen für Kinder < 8 Jahre vor. Die Auswertung der Eltern-/Betreuerversion innerhalb des FAS ab 5 Jahren wird nicht in der Nutzenbewertung dargestellt.

Zur Studienvsiste vor Zyklus 13 berichten 52 bzw. 59 % der Kinder und Jugendlichen, welche den GIC selbst ausfüllten (Alter 8–18 Jahre), ihr Tumorschmerz bzw. ihre Tumor-assoziierte Morbidität sei „viel besser“ oder „sehr viel besser“ im Vergleich zu vor Beginn der Behandlung. Für den Gesamtschmerz berichteten 41 % einen „viel besseren“ oder „sehr viel besseren“ Zustand zur Studienvsiste vor Zyklus 13. Eine Zusammenfassung über die Antwortkategorien für Verschlechterungen im GIC liegt nicht vor. Die Darstellung aller Antwortkategorien (Tabelle 4-168 im Dossier, Modul 4 [4]) zeigt, dass Verschlechterungen insgesamt nur selten berichtet werden (maximale Häufigkeit ≤ 10 %).

Tabelle 19: Responderanalyse zum GIC; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalyse zum GIC	Patientenberichteter GIC (N = 34)							
	Vor Zyklus 3		Vor Zyklus 5		Vor Zyklus 9		Vor Zyklus 13	
	n	n (% ¹⁾)	n	n (% ¹⁾)	n	n (% ¹⁾)	n	n (% ¹⁾)
Tumorschmerz Verbesserung ²⁾ Verschlechterung ³⁾	26	12 (46,2) k. A.	30	14 (46,7) k. A.	30	16 (53,3) k. A.	29	15 (51,7) k. A.
Gesamtschmerz Verbesserung ²⁾ Verschlechterung ³⁾	30	9 (30,0) k. A.	30	12 (40,0) k. A.	30	10 (33,3) k. A.	29	12 (41,4) k. A.
Tumor-assoziierte Morbidität Verbesserung ²⁾ Verschlechterung ³⁾	23	- ⁴⁾ - ⁴⁾	29	14 (48,3) k. A.	30	13 (43,3) k. A.	29	17 (58,6) k. A.

¹⁾ Prozentualer Anteil bezogen auf alle Personen, die geantwortet haben.

²⁾ Antwortkategorien „sehr viel besser“ oder „viel besser“.

³⁾ Für Verschlechterungen wurde keine über die Antwortkategorien aggregierte Analyse vorgelegt.

⁴⁾ Keine Darstellung, da Rücklaufquote unter 70 %.

Abkürzungen: FAS: Full Analysis Set; GIC: Global Impression of Change; k. A.: keine Angabe.

3.4 Lebensqualität

PedsQL

Die Darstellung der patientenberichteten Ergebnisse zur gesundheitsbezogenen Lebensqualität erhoben mittels PedsQL ist auf den Zeitraum bis zur Studiervisite vor Zyklus 13 beschränkt, da danach die Rücklaufquote, bezogen auf das FAS im Alter von 8 bis 18 Jahren, bei unter 70 % lag. Die gesundheitsbezogene Lebensqualität wird für Kinder ≥ 8 Jahre direkt über die Befragung der Kinder erhoben und ist damit über die Patientenversion des PedsQL abgebildet (Tabelle 20). Für die Eltern-/Betreuerversion des PedsQL liegen keine separaten Auswertungen für Kinder < 8 Jahre vor. Die Auswertung der Eltern-/Betreuerversion innerhalb des gesamten FAS wird nicht in der Nutzenbewertung dargestellt.

Die Baseline-Werte für den PedsQL sind im Median zu Baseline bereits relativ hoch (Tabelle 20), was auf eine hohe gesundheitsbezogene Lebensqualität vieler Teilnehmenden zu Baseline schließen lässt. Der Median für den Gesamtwert liegt bei 82 Punkten, weshalb aufgrund der Skalenspannweite von 0 bis 100 eine Verbesserung um 15 Punkte für einen Teil der Studienpopulation nicht möglich ist. Daher wird zusätzlich zu den Responderanalysen auch die kontinuierliche Analyse mittels MMRM in der Nutzenbewertung dargestellt (Tabelle 21); MMRM-Analysen liegen nur für den Gesamtwert vor. Anzumerken ist weiterhin, dass der Anteil an Teilnehmenden, welche für die PedsQL-Subskala „Schulische Funktion“ mehr als 50 % der Items beantwortet hat, unter 70 % liegt, weshalb die Ergebnisse in der Nutzenbewertung nicht dargestellt werden. Der Median der Subskala „Schulische Funktion“ lag zu Baseline bei 75 Punkten und damit niedriger als der Median der drei anderen Subskalen („Körperliche Funktion“: 84 Punkte; „Emotionale Funktion“: 85 Punkte; „Soziale Funktion“: 80 Punkte). Es muss daher davon ausgegangen werden, dass die Einschränkungen der schulischen Funktion für die Patientenpopulation am stärksten sind. Daher kann die fehlende Beantwortung der Items zur schulischen Funktion im Studienverlauf die Ergebnisse des Gesamtwerts zugunsten von Selumetinib verzerren.

Tabelle 20: Responderanalyse für den PedsQL; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Responderanalyse PedsQL	PedsQL – patientenberichtet (N = 34)									
	Baseline		Vor Zyklus 3		Vor Zyklus 5		Vor Zyklus 9		Vor Zyklus 13	
	n	MW (SD) Median (min, max)	n	n (% [95%-KI] ¹⁾	n	n (% [95%-KI] ¹⁾	n	n (% [95%-KI] ¹⁾	n	n (% [95%-KI] ¹⁾
Gesamtwert Verbesserung ²⁾ Verschlechterung ³⁾	33	73,9 (20,7) 81,8 (13,0; 96,7)	31	6 (19,4 [7,5; 37,5]) k. A.	31	6 (19,4 [7,5; 37,5]) k. A.	31	9 (29,0 [14,2; 48,0]) k. A.	29	7 (24,1 [10,3; 43,5]) k. A.
Körperliche Funktion Verbesserung ²⁾ Verschlechterung ³⁾	33	75,5 (23,0) 84,4 (15,6; 100)	31	8 (25,8 [11,9; 44,6]) k. A.	31	8 (25,8 [11,9; 44,6]) k. A.	31	4 (12,9 [3,6; 29,8]) k. A.	29	10 (34,5 [17,9; 54,3]) k. A.
Emotionale Funktion Verbesserung ²⁾ Verschlechterung ³⁾	33	75,9 (22,8) 85,0 (5,0; 100)	31	11 (35,5 [19,2; 54,6]) k. A.	31	10 (32,3 [16,7; 51,4]) k. A.	31	10 (32,3 [16,7; 51,4]) k. A.	29	11 (37,9 [20,7; 57,7]) k. A.
Soziale Funktion Verbesserung ²⁾ Verschlechterung ³⁾	33	75,9 (23,4) 80,0 (0; 100)	30	10 (33,3 [17,3; 52,8]) k. A.	31	9 (29,0 [14,2; 48,0]) k. A.	31	8 (25,8 [11,9; 44,6]) k. A.	29	9 (31,0 [15,3; 50,8]) k. A.
Schulische Funktion Verbesserung ²⁾ Verschlechterung ³⁾	28	66,3 (27,5) 75,0 (10,0; 100)	20	- ⁴⁾ - ⁴⁾	17	- ⁴⁾ - ⁴⁾	23	- ⁴⁾ - ⁴⁾	23	- ⁴⁾ - ⁴⁾

¹⁾ Prozentualer Anteil bezogen auf alle Personen, die geantwortet haben; KI wurden mit der Clopper-Pearson-Methode für binomiale Verteilungen berechnet.

²⁾ Verbesserung um mindestens 15 Punkte.

³⁾ Verschlechterung um mindestens 15 Punkte.

⁴⁾ Keine Darstellung, da Rücklaufquote unter 70 %.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; MW: Mittelwert; PedsQL: Pediatric Quality of Life Inventory; SD: Standardabweichung.

Tabelle 21: Veränderungen des PedsQL-Gesamtwerts; FAS im Alter von 8 bis 18 Jahren der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Veränderung PedsQL-Gesamtwert	PedsQL ¹⁾ [Punkte] N = 34	
	n	LSM [95%-KI], p-Wert ²⁾
<i>Baseline</i>		
MW (SD)	33	73,9 (20,7)
Median (min, max)		81,8 (13,0; 96,7)
<i>Veränderung zu Baseline</i>		
Vor Zyklus 3	31	6,55 [2,81; 10,28]; 0,001
Vor Zyklus 5	31	4,68 [0,34; 9,02], 0,04
Vor Zyklus 9	31	5,29 [0,74; 9,84], 0,02
Vor Zyklus 13	29	6,68 [1,34; 12,03], 0,02

¹⁾ Höhere Werte repräsentieren eine höhere Lebensqualität.

²⁾ LSM, KI und p-Wert adjustiert für Studienvsiste, Baseline-Wert, Alter, Anzahl an Morbiditäten zu Baseline und Baseline x Studienvsiste-Interaktion mittels MMRM.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; LSM: Least Square Mean; MMRM: Mixed Effect Model Repeat Measures; MW: Mittelwert; PedsQL: Pediatric Quality of Life Inventory; SD: Standardabweichung.

3.5 Sicherheit

Ergebnisse zur Sicherheit innerhalb der Studie SPRINT, Phase I und Phase II, bis zum Datenschnitt vom 29.06.2018 sind in Tabelle 22 zusammengefasst. Angaben zur Dauer der Beobachtungszeit für die Sicherheit liegen nicht vor. Die mediane Behandlungsdauer mit Selumetinib wird mit 1.597 Tagen (\pm 52,5 Monaten oder 4,4 Jahren) für Phase I und 801,5 Tagen (\pm 26,4 Monaten) bzw. 1.027,5 Tagen (\pm 33,8 Monaten) für die Datenschnitte vom 29.06.2018 bzw. 29.03.2019 für Phase II angegeben. Die Nachbeobachtung der Sicherheit sollte bis 30 Tage nach Beendigung der Behandlung erfolgen.

Innerhalb der Phase II (N = 50) wurden bei 62 % der Teilnehmenden UE Grad 3 oder höher beobachtet, bei 24 % SUE und bei 12 % UE, die zum Abbruch der Studienmedikation führten (Tabelle 22). Die Gesamtanzahl an von UE (inklusive UE Grad 3 oder höher), UE, die zum Abbruch der Studienmedikation führten, sowie SUE betroffenen Teilnehmenden innerhalb der Phase II blieb bis zum Datenschnitt vom 29.03.2019 konstant.

Tabelle 22: Unerwünschte Ereignisse; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT Unerwünschte Ereignisse Personen mit mindestens einem ...	Phase I N = 6 n (%)	Phase II N = 50 n (%)
UE	6 (100)	49 (98)
UE CTCAE-Grad \geq 3	5 (83)	31 (62)
SUE	2 (33)	12 (24)
UE, das zum Abbruch der Studienmedikation führte ¹⁾	0	6 (12)

¹⁾ Die Studienteilnehmenden erhielten die Studienmedikation bis zum Auftreten einer Krankheitsprogression, inakzeptabler UE, der Rücknahme der Einverständniserklärung oder der Entscheidung des ärztlichen Prüfpersonals, je nachdem, was früher auftrat. Diese möglichen Therapieabbruchgründe, die vor einem potentiellen Abbruch aufgrund von UE auftreten können, stellen somit ein konkurrierendes Ereignis dar, weshalb die Ergebnissicherheit und Interpretierbarkeit der Ergebnisse eingeschränkt ist.

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

Unerwünschte Ereignisse

Die Häufigkeit von UE jeglichen Schweregrads nach Systemorganklasse und Preferred Term bis zum Datenschnitt vom 29.06.2018 ist in Tabelle 23 zusammengefasst. Teilnehmende waren häufig von UE folgender Systemorganklassen betroffen: „Erkrankungen der Haut und des subkutanen Gewebes“ (98 %), „Untersuchungen“ (98 %), „Erkrankungen des Gastrointestinaltrakts“ (96 %), „Stoffwechsel- und Ernährungsstörungen“ (96 %), „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (84 %), „Erkrankungen der Atemwege, des Brustraums und des Mediastinums“ (82 %), „Infektionen und Infestationen“ (82 %), „Erkrankungen der Muskeln und des Bindegewebes“ (70 %), „Erkrankungen des Nervensystems“ (70 %), „Erkrankungen der Nieren und Harnwege“ (50%) sowie „Erkrankungen des Bluts und des Lymphsystems“ (46 %).

Bis zum Datenschnitt vom 29.03.2019 waren nur wenige Teilnehmende zusätzlich von UE betroffen. Diese betrafen die Systemorganklassen „Erkrankungen des Bluts und Lymphsystems“ (+ 2 Personen), „Psychiatrische Erkrankungen“ (+ 3 Personen), „Erkrankungen des Ohrs und des Labyrinths“ (+ 3 Personen), „Vaskuläre Erkrankungen“ (+ 1 Person), „Erkrankungen des Gastrointestinaltrakts“ (+ 1 Person), „Erkrankungen der Muskeln und des Bindegewebes“ (+ 1 Person), „Erkrankungen der Nieren und der Harnwege“ (+ 2 Personen) und „Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen“ (+ 1 Person).

Tabelle 23: UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT UE mit Inzidenz ≥ 10 %¹⁾ MedDRA-Systemorganklasse <i>Preferred Term</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Erkrankungen der Haut und des subkutanen Gewebes	6 (100)	49 (98)
Trockene Haut	4 (67)	30 (60)
Dermatitis akneiform	3 (50)	25 (50)
Pruritus	3 (50)	23 (46)
Ausschlag makulo-papulös	4 (67)	18 (36)
Ekzem	3 (50)	13 (26)
Alopezie	4 (67)	12 (24)
Änderung der Haarfarbe	3 (50)	11 (22)
Untersuchungen	6 (100)	49 (98)
Kreatinphosphokinase im Blut erhöht	6 (100)	38 (76)
Aspartat-Aminotransferase erhöht	4 (67)	23 (46)
Alanin-Aminotransferase erhöht	3 (50)	17 (34)
Neutrophilenzahl vermindert	2 (33)	16 (32)
Lipase erhöht	0	13 (26)
Kreatinin im Blut erhöht	2 (33)	14 (28)
Lymphozytenzahl erhöht	2 (33)	10 (20)
Ejektionsfraktion erniedrigt	1 (17)	11 (22)

Studie SPRINT UE mit Inzidenz ≥ 10 %¹⁾ MedDRA-Systemorganklasse Preferred Term	Phase I N = 6 n (%)	Phase II N = 50 n (%)
Alkalische Phosphatase im Blut erhöht	3 (50)	10 (20)
Lymphozytenzahl vermindert	4 (67)	10 (20)
Amylase erhöht	0	7 (14)
Hämoglobin erhöht	4 (67)	10 (20)
Leukozytenzahl vermindert	2 (33)	10 (20)
Thrombozytenzahl vermindert	0	6 (12)
Gewicht erhöht	1 (17)	5 (10)
Erkrankungen des Gastrointestinaltrakts	6 (100)	48 (96)
Erbrechen	4 (67)	41 (82)
Diarrhö	5 (83)	35 (70)
Nausea	4 (67)	33 (66)
Stomatitis	2 (33)	25 (50)
Abdominalschmerz	2 (33)	22 (44)
Schmerzen Oberbauch	3 (50)	20 (40)
Obstipation	4 (67)	17 (34)
Bauch aufgetrieben	1 (17)	5 (10)
Stoffwechsel- und Ernährungsstörungen	6 (100)	48 (96)
Hypoalbuminämie	3 (50)	25 (50)
Hypoglykämie	3 (50)	14 (28)
Hypokaliämie	1 (17)	13 (26)
Hypokalziämie	2 (33)	12 (24)
Verminderter Appetit	0	11 (22)
Hyperglykämie	4 (67)	11 (22)
Hyperkaliämie	4 (67)	11 (22)
Hypernatriämie	2 (33)	9 (18)
Hyponatriämie	0	8 (16)
Hypophosphatämie	2 (33)	8 (16)
Dehydratation	1 (17)	7 (14)
Hypomagnesiämie	1 (17)	5 (10)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	6 (100)	42 (84)
Pyrexie	5 (83)	28 (56)
Fatigue	4 (67)	28 (56)
Grippeähnliche Erkrankung	1 (17)	15 (30)

Studie SPRINT UE mit Inzidenz ≥ 10 %¹⁾ MedDRA-Systemorganklasse Preferred Term	Phase I N = 6 n (%)	Phase II N = 50 n (%)
Schmerz	0	11 (22)
Ödem peripher	1 (17)	6 (12)
Erkrankungen der Atemwege, des Brustraums und des Mediastinums	6 (100)	41 (82)
Schmerzen im Oropharynx	1 (17)	24 (48)
Husten	6 (100)	20 (40)
Nasenverstopfung	4 (67)	17 (34)
Allergische Rhinitis	0	15 (30)
Epistaxis	2 (33)	14 (28)
Rhinorrhoe	1 (17)	6 (12)
Infektionen und Infestationen	6 (100)	41 (82)
Paronchie	3 (50)	23 (46)
Otitis media	3 (50)	14 (28)
Infektion der oberen Atemwege	3 (50)	8 (16)
Pharyngitis	0	6 (12)
Hautinfektion	1 (17)	8 (16)
Pharyngitis durch Streptokokken	3 (33)	7 (14)
Sinusitis	3 (50)	6 (12)
Zellulitis	1 (17)	5 (10)
Konjunktivitis	1 (17)	5 (10)
Erkrankungen der Muskeln und des Bindegewebes	5 (83)	35 (70)
Schmerz in einer Extremität	1 (17)	18 (36)
Rückenschmerzen	3 (50)	9 (18)
Nackenschmerzen	1 (17)	8 (16)
Arthralgie	0	5 (10)
Erkrankungen des Nervensystems	5 (83)	35 (70)
Kopfschmerz	5 (83)	24 (48)
Schwindel	0	14 (28)
Erkrankungen der Nieren und Harnwege	3 (50)	25 (50)
Hämaturie	1 (17)	11 (22)
Proteinurie	0	11 (22)
Harninkontinenz	1 (17)	7 (14)
Erkrankungen des Bluts und des Lymphsystems	2 (33)	23 (46)
Anämie	2 (33)	21 (42)

Studie SPRINT UE mit Inzidenz ≥ 10 %¹⁾ MedDRA-Systemorganklasse <i>Preferred Term</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen	6 (100)	22 (44)
Sturz	2 (33)	13 (26)
Psychiatrische Erkrankungen	2 (33)	20 (40)
Schlaflosigkeit	2 (33)	10 (20)
Angst	0	5 (10)
Augenerkrankungen	2 (33)	17 (34)
Tränensekretion verstärkt	0	8 (16)
Vaskuläre Erkrankungen	3 (50)	14 (28)
Hypertonie	2 (33)	9 (18)
Herzerkrankungen	2 (33)	12 (24)
Sinustachykardie	2 (33)	10 (20)
Erkrankungen des Ohrs und des Labyrinths	1 (17)	9 (18)
Ohrenschmerzen	1 (17)	5 (10)
Erkrankungen des Immunsystems	2 (33)	6 (12)
Überempfindlichkeit	2 (33)	6 (12)

¹⁾ Innerhalb Phase II.

Abkürzungen: FAS: Full Analysis Set; MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

Unerwünschte Ereignisse CTCAE-Grad 3 oder höher

UE des Schweregrad 3 oder höher nach Systemorganklasse und Preferred Term mit Auftreten bis zum Datenschnitt vom 29.06.2018 sind in Tabelle 24 aufgelistet. UE CTCAE-Grad 3 oder höher betrafen die Systemorganklassen „Erkrankungen des Gastrointestinaltrakts“ (24 %), „Untersuchungen“ (20 %), „Infektionen und Infestationen“ (18 %), „Erkrankungen der Haut und des subkutanen Gewebes“ (10 %), „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (8 %), „Erkrankungen der Atemwege, des Brustraums und des Mediastinums“ (8 %), „Erkrankungen des Nervensystems“ (6 %) und „Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen“ (6 %). Bis zum Datenschnitt vom 29.03.2019 traten wenige Ereignisse in weiteren Teilnehmenden auf. Diese betrafen die Systemorganklassen „Erkrankungen des Gastrointestinaltrakts“ (+ 2 Personen), „Erkrankungen der Haut und des subkutanen Gewebes“ (+ 1 Person) und „Untersuchungen“ (+ 1 Person).

Tabelle 24: UE CTCAE-Grad 3 oder höher mit Inzidenz $\geq 5\%$ nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT UE CTCAE-Grad 3 oder höher mit Inzidenz $\geq 5\%$¹⁾ MedDRA-Systemorganklasse <i>Preferred Term</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Erkrankungen des Gastrointestinaltrakts	2 (33)	12 (24)
Diarrhö	0	8 (16)
Erbrechen	2 (33)	3 (6)
Untersuchungen	1 (17)	10 (20)
Gewicht erhöht	0	3 (6)
Kreatinphosphokinase im Blut erhöht	1 (17)	3 (6)
Infektionen und Infestationen	1 (17)	9 (18)
Paronchie	1 (17)	3 (6)
Erkrankungen der Haut und des subkutanen Gewebes	1 (17)	5 (10)
Erkrankungen der Atemwege, des Brustraums und des Mediastinums	0	4 (8)
Hypoxie	0	4 (8)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	1 (17)	4 (8)
Pyrexie	1 (17)	4 (8)
Erkrankungen des Nervensystems	0	3 (6)
Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen	0	3 (6)

¹⁾ Innerhalb Phase II.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; FAS: Full Analysis Set; MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

Schwerwiegende Unerwünschte Ereignisse

Angaben zur Häufigkeit aufgetretener SUE nach Systemorganklasse und Preferred Term sind in Tabelle 25 dargestellt. Mit einer Inzidenz von über 5 % traten SUE in der Systemorganklasse „Infektionen und Infestationen“ (12 %) und „Erkrankungen des Gastrointestinaltrakts“ (6 %) auf. Die Häufigkeit an von UE dieser Systemorganklasse betroffenen Teilnehmenden blieb bis zum Datenschnitt vom 29.03.2019 konstant.

Tabelle 25: SUE mit Inzidenz \geq 5 % nach MedDRA-Systemorganklasse und Preferred Term; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT SUE mit Inzidenz \geq 5 %¹⁾ <i>MedDRA-Systemorganklasse</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Infektionen und Infestationen	0	6 (12)
Erkrankungen des Gastrointestinaltrakts	1 (17)	3 (6)

¹⁾ Innerhalb Phase II.

Abkürzungen: FAS: Full Analysis Set; MedDRA: Medical Dictionary for Regulatory Activities; SUE: Schwerwiegendes unerwünschtes Ereignis.

Unerwünschte Ereignisse von besonderem Interesse

UE von besonderem Interesse bis zum Datenschnitt vom 29.06.2018 sind in Tabelle 26 zusammengefasst. Bis zum Datenschnitt vom 29.03.2019 waren einige weitere Teilnehmende von UE von besonderem Interesse betroffen: „Effekte der Erythropenie“ (+ 2 Personen), „Effekte der Leukopenie“ (+ 1 Person), „Muskelbezogene Effekte“ (+ 1 Person), „Ausschlag, akneiform“ (+ 1 Person), „Ausschlag, nicht-akneiform“ (+ 1 Person).

Tabelle 26: UE von besonderem Interesse und zugeordnete Preferred Terms mit Inzidenz \geq 5 %¹⁾; FAS bzw. Sicherheitspopulation der Studie SPRINT, Phase I und Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT UE von besonderem Interesse <i>Preferred Term</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Muskelbezogene Effekte	6 (100)	43 (86)
Kreatinphosphokinase im Blut erhöht	6 (100)	38 (76)
Kreatinin erhöht	2 (33)	14 (28)
Hypokalziämie	2 (33)	12 (24)
Muskuloskelettale Schmerzen	0	3 (6)
Ausschlag, nicht akneiform	5 (83)	35 (70)
Pruritus	3 (50)	23 (46)
Ausschlag	0	3 (6)
Ausschlag makulo-papulös	4 (67)	18 (36)
Ausschlag, akneiform	3 (50)	25 (50)
Dermatitis akneiform	3 (50)	25 (50)

Studie SPRINT UE von besonderem Interesse <i>Preferred Term</i>	Phase I N = 6 <i>n (%)</i>	Phase II N = 50 <i>n (%)</i>
Effekte der Entzündung der oralen Mundschleimhaut	2 (33)	25 (50)
Stomatitis	2 (33)	25 (50)
Erkrankungen des Nagels	4 (67)	23 (46)
Paronchie	3 (50)	23 (46)
Effekte der Leukopenie	5 (83)	22 (44)
Lymphozytenzahl vermindert	4 (67)	10 (20)
Neutrophilenzahl vermindert	2 (33)	16 (32)
Leukozytenzahl vermindert	2 (33)	10 (20)
Effekte der Erythropenie	2 (33)	21 (42)
Anämie	2 (33)	21 (42)
Effekte der Herzinsuffizienz	2 (33)	18 (36)
Ejektionsfraktion erniedrigt	1 (17)	11 (22)
Ödem peripher	1 (17)	6 (12)
Netzhaut betreffende Effekte	1 (17)	8 (16)
Verschwommenes Sehen	1 (17)	4 (8)
Effekte der Thrombozytopenie	0	6 (12)
Thrombozytenzahl vermindert	0	6 (12)
Physeale Dysplasie	k. A.	0

¹⁾ Innerhalb Phase II.

Abkürzungen: FAS: Full Analysis Set; k. A.: keine Angabe; MedDRA: Medical Dictionary for Regulatory Activities; UE: Unerwünschtes Ereignis.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Selumetinib

Selumetinib (Koselugo®) ist zugelassen für Kinder ab 3 Jahren und Jugendliche zur Behandlung von symptomatischen, inoperablen plexiformen Neurofibromen (PN) bei Neurofibromatose Typ 1 (NF1). Von der EMA wurde eine bedingte Zulassung erteilt mit der Auflage Ergebnisse der pivotalen Studie SPRINT mit einer längeren Nachbeobachtungszeit (Datenschnitte: Phase I: 27.02.2021; Phase II: 31.03.2021) bis zum 31.03.2022 sowie Ergebnisse einer nicht-interventionellen Post-Authorisation Safety Study (PASS) bis zum 31.12.2027 bei der Zulassungsbehörde vorzulegen.

Die für die Nutzenbewertung herangezogene pivotal Studie SPRINT spiegelt die Zulassungspopulation größtenteils wider. In Phase II (Stratum 1) der Studie SPRINT wurden Personen mit NF1 und inoperablen PN, die eine signifikante Morbidität verursachten, eingeschlossen. Zudem wurden, wie in klinischen Studien üblich, Anforderungen bezüglich des Performance-Status ($\geq 70\%$ für Karnofsky-Index bzw. Lansky-Index) und einer adäquaten hämatologischen sowie Leber-, Nieren- und Herzfunktion gestellt. Personen mit dem Nachweis eines Optikusglioms, eines malignen Glioms, eines MPNST oder einer anderen Krebserkrankung, die eine Behandlung mit Chemo- oder Strahlentherapie nötig machten, waren ausgeschlossen, genauso wie Personen mit spezifischen Augenkrankungen. Der Einschluss in die Studie SPRINT erfolgte für Kinder ab 3 Jahren, welche in der Lage waren, die Studienmedikation in Form einer Kapsel oral einzunehmen. Insgesamt wurden nur 2 Kinder, die jünger als 5 Jahre waren, in die Zulassungsstudie eingeschlossen. Jüngere Kinder sind in der Studienpopulation daher unterrepräsentiert. Insgesamt scheint die Zulassungspopulation etwas breiter als die Studienpopulation zu sein.

Gemäß Fachinformation [5] beträgt die zugelassene Dosierung 25 mg/m² Körperoberfläche, zweimal täglich oral eingenommen (alle 12 Stunden). Die Behandlung mit Selumetinib soll so lange fortgesetzt werden, wie ein klinischer Nutzen beobachtet wird bzw. bis zum Auftreten einer Progression der PN oder einer inakzeptablen Toxizität [5]. Die Anwendung von Selumetinib erfolgte FI-konform in Phase II der Studie SPRINT und in einem Studienarm der Phase I. Die Ergebnisse der Phase I für die FI-konforme Dosierung werden in der Nutzenbewertung ebenfalls berichtet. Für die Phase I galten größtenteils ähnliche Ein- und Ausschlusskriterien, ein wesentlicher Unterschied zur Phase II ist, dass nicht ausschließlich symptomatische Patientinnen und Patienten eingeschlossen wurden. Da in Phase I die Untersuchung der Sicherheit im Vordergrund stand, können die erhobenen Daten zur Sicherheit für die Nutzenbewertung herangezogen werden.

4.2 Design und Methodik der Studie

Bei Phase I und Phase II der Studie SPRINT handelt es sich um offene, einarmige, multizentrische Studien mit Selumetinib bei pädiatrischen Patientinnen und Patienten mit NF1 und inoperablen PN. Die Studie wird vom NCI Cancer Therapy Evaluation Program gesponsert und wurde an 4 Studienzentren in den USA durchgeführt. In beiden Phasen erfolgte die Behandlung in Zyklen von 28 Tagen ohne Einnahmepausen zwischen den Zyklen. Teilnehmende (mit Krankheitsprogression innerhalb von 1,5 Jahren vor Studieneintritt) erhalten das Studienmedikament so lange sie keine schweren Nebenwirkungen haben und sich die Krankheit nicht verschlimmert. Bei Teilnehmenden ohne Krankheitsprogression in den 1,5 Jahren vor Studieneintritt und ohne Ansprechen (Reduktion des Tumolvolumen um $\geq 20\%$) sollte die Behandlung nach 2 Jahre beendet werden. Die Patientinnen und Patienten sollten für insgesamt 7 Jahre nach Beginn der Behandlung mit Selumetinib oder 5 Jahre

nach Beendigung der Behandlung mit Selumetinib nachverfolgt werden, je nachdem was eine längere Nachbeobachtung bietet.

Für Phase I der Studie SPRINT wird der Datenschnitt der primären Analyse vom 29.06.2018 in der Nutzenbewertung dargestellt. Für Phase II wird ebenfalls der Datenschnitt der primären Analyse vom 29.06.2018 dargestellt. Für die Endpunktkategorie „Sicherheit“, inklusive Todesfälle, liegen für Phase II neuere Daten eines durch die FDA getriggerten Datenschnitts vom 29.03.2019 vor, welche für diese Endpunkte zur Nutzenbewertung ergänzend herangezogen werden. Ein weiterer neuerer Datenschnitt wurde von der EMA im Zusammenhang mit der bedingten Zulassung von Selumetinib beauftragt (Phase I: 27.02.2021; Phase II: 31.03.2021). Der pU gibt an, dass ihm für diese Datenschnitte zum Zeitpunkt der Dossiereinreichung noch keine Daten bzw. Studienberichte vorliegen. Zum Zeitpunkt, der in der Nutzenbewertung dargestellten Datenschnitte waren beide Studienphasen noch nicht abgeschlossen.

In den Studienarm mit der FI-konformen Dosierung der Phase I wurden insgesamt 6 Personen eingeschlossen, welche alle bis zum Datenschnitt vom 29.06.2018 behandelt wurden. Die mediane Behandlungsdauer in Phase I betrug 1.597 Tage (\approx 52,5 Monaten oder 4,4 Jahren).

In Phase II wurden insgesamt 50 Personen eingeschlossen. Innerhalb der Phase II brachen bis zum Datenschnitt vom 29.03.2019 20 Personen (40 %) die Behandlung mit Selumetinib ab; 12 % aufgrund von UE, 10 % aufgrund einer Krankheitsprogression, 8 % aufgrund der Entscheidung des ärztlichen Prüfungspersonals, 4 % weil sie die Behandlungsperiode abgeschlossen hatten, 4 % wegen der Rücknahme der Einverständniserklärung und 2 % wegen Protokollverletzungen. Für 2 Teilnehmende (4 %) wird über eine Wiederaufnahme der Behandlung, nach Abbruch der Einnahme der Studienmedikation, berichtet. Die mediane Behandlungsdauer in Phase II betrug 1.027,5 Tage (\approx 33,8 Monaten) für den Datenschnitt vom 29.03.2019. Angaben zur Beobachtungsdauer der Phase II konnten in den Studienunterlagen nicht identifiziert werden. Für alle Studienteilnehmenden sowohl in Phase I als auch in Phase II wurden Dosisunterbrechungen registriert. Ca. 80 % der Studienpopulation ($n = 5$ in Phase I, $n = 41$ in Phase II) wiesen ein UE auf, welches zu einer Dosisunterbrechung führte. Angaben zu Anzahl an Teilnehmenden mit einer Dosisreduktion bzw. Dosiserhöhung konnten in den Studienunterlagen nicht identifiziert werden. In Phase I wiesen 3 Teilnehmende (50 %) ein UE auf, das zu einer Dosisreduktion führte und in Phase II 13 Teilnehmende (26 %). Bis zum Datenschnitt vom 29.03.2019 brachen 6 Teilnehmende (12 %) ihre Studienteilnahme in der Phase II ab, davon 3 aufgrund des Entzugs der Einverständniserklärung, 2 aus anderen Gründen und 1 Person wird als Lost to Follow-up verzeichnet. Folglich befanden sich zum Zeitpunkt des Datenschnitts 30 Personen in der Behandlungsphase und 14 Personen, die ihre Behandlung abgebrochen oder beendet hatten, befanden sich in der Nachbeobachtungsphase der Phase II der Studie SPRINT.

Das mediane Alter zu Baseline lag bei 6,8 Jahren in Phase I und 10,2 Jahren in Phase II. Die Studienpopulation war überwiegend europäischer Abstammung. Zwischen Diagnose der NF1-Erkrankung und Beginn der Behandlung mit Selumetinib lagen im Median ca. 8 Jahre. Zwischen Diagnose des PN und Beginn der Behandlung mit Selumetinib lagen im Median ca. 6 Jahre.

Innerhalb der Studie SPRINT wurde das klinisch relevanteste, inoperable PN, welches mittels volumetrischer 3D-MRT-Messung erfasst werden konnte, als Ziel-PN definiert. Daneben konnten maximal 2 weitere, inoperable PN als zusätzliche Nicht-Zielläsionen für weitere 3D-MRT-Untersuchungen ausgewählt werden. Die dem Ziel-PN zugeschriebenen Morbiditäten in Phase II waren „Entstellung“ (88 %), „Motordysfunktion“ (66 %), „Blockierung der Atemwege“ (32 %), „Störung der Darm- und Blasenfunktion“ (20 %), „Einschränkung des Sehens“ (20 %) und „Andere Dysfunktion“ (24 %). Für

Phase I liegen keine Angaben dazu vor. 52 % der Studienteilnehmenden der Phase II der Studie SPRINT gaben an Schmerzen aufgrund der Ziel-PN zu haben. Das Ziel-PN verursachte dabei 1 bis 4 Morbiditäten (Median = 3). Angaben zur Anzahl weiterer PN außer dem Ziel-PN konnten in den Studienunterlagen nicht identifiziert werden. Für Phase II wird lediglich im Studienbericht angegeben, dass keine Nicht-Ziel-PN definiert wurden.

Indirekte Vergleiche mit externen Kontrollstudien

Um den Zusatznutzen über die einarmige, pivotale Studie nachzuweisen legt der pU einen indirekten Vergleich mit externen Kontrollstudien [19,37,38,52,54] vor. Für 4 Studien [19,37,38,52] basiert der Vergleich auf publizierten Studien, welche aber zum Teil große Unterschiede hinsichtlich der eingeschlossenen Studienpopulation aufweisen und für die mangels ausführlicher Informationen zu Baseline-Charakteristika keine hinreichende Vergleichbarkeit mit der Studienpopulation der Studie SPRINT nachgewiesen werden kann. Für keine der 4 Studien wurden ein Studienprotokoll, ein SAP sowie eine detaillierte Ergebnisdarstellung eingereicht, sodass relevante Angaben zum Nachweis der Vergleichbarkeit der Studienpopulationen und Endpunkte fehlen. Daneben führt der pU einen indirekten Vergleich mit einer prospektiven Beobachtungsstudie (NCI-08-C-0079) zum natürlichen Verlauf der NF1-Erkrankung und mit dem Placebo-Arm der RCT 01-C-0222 durch. Die Studie zum natürlichen Verlauf dient als Rahmenprotokoll für das klinische Studienprogramm des NCI POB, welches auch die Studie SPRINT koordiniert hat. Die eingeschlossene Studienpopulation wird in einer der Publikationen zur Studie als wenig repräsentativ für die allgemeine NF1-Population beschrieben, sondern umfasst eher Patientinnen und Patienten, die aufgrund der Schwere ihrer Erkrankung um eine Behandlung ihrer PN ersuchen [2]. Für diese Studie wurde ein Studienprotokoll sowie ein Bericht über eine mittels Propensity-Score-adjustierte Analyse für den indirekten Vergleich vorgelegt. Die Analyse wird auch im Studienbericht der Phase II der SPRINT-Studie beschrieben. Dennoch fehlen Angaben bspw. zur Operationalisierung der erhobenen Endpunkte. Zudem gibt es Unterschiede in den Ein- und Ausschlusskriterien (z. B. keine Notwendigkeit einer PN-bezogenen Morbidität; keine Kriterien bezüglich Leber-, Lungen-, Nieren- und Herzfunktion und Infektionen; kein Ausschluss von Personen mit Gliom oder MPNST, welche eine Behandlung mit Chemo- oder Strahlentherapie erforderten), sodass die Patientenpopulation insgesamt breiter ist als die der Studie SPRINT. Auch die Studienpopulation des Placebo-Arms der Studie 01-C-0222 unterscheidet sich wesentlich von der der SPRINT-Studie, da in die Studie 01-C-0222 bei ansonsten sehr ähnlichen Einschlusskriterien nur Personen mit einer progressiven Erkrankung ($\geq 20\%$ Zunahme des Volumens der PN, oder $\geq 13\%$ Zunahme in zweidimensionalen Messungen, oder $\geq 6\%$ Zunahme in eindimensionaler Messung über die 2 letzten volumetrischen MRT-Scans oder innerhalb eines Jahres vor Einschluss in die Studie) eingeschlossen wurden. In der Studie SPRINT lag der Anteil an Personen mit progressiver Erkrankung ($\geq 20\%$ Zunahme des PN-Volumens innerhalb von 12 bis 15 Monaten vor dem Einschluss) bei 42 %. Weitere Unterschiede zwischen den Studienpopulationen zu Baseline zeigen sich bezüglich demographischer Faktoren wie Alter (Teilnehmende in der Studie SPRINT, Phase II waren ca. 2 Jahre älter als in den Kontrollstudien) und bezüglich Krankheitscharakteristika (Tumorzusatz ca. 1,5 x größer in der Studie SPRINT als in externen Kontrollstudien). Da für den indirekten Vergleich vom pU nur die zwei nicht patientenrelevanten Endpunkte „Tumorzusatz“ und „Progressionsfreies Überleben“ herangezogen wurden (siehe Abschnitt 2.3.2), werden der indirekte Vergleich und die dafür eingereichten externen Kontrollstudien nicht für die Nutzenbewertung herangezogen.

Eine Bewertung der indirekten Vergleiche mit der Beobachtungsstudie zum natürlichen Verlauf und mit dem Placebo-Arm der RCT 01-C-0222 erfolgte im Rahmen des Zulassungsprozesses bei der

FDA und der EMA. Die FDA erachtete die Ergebnisse der externen Kontrollstudien zwar als unterstützend, aber für vergleichende Analysen als ungeeignet. Dies wird mit der Heterogenität der Patientenpopulation in Bezug auf gemessene und nicht gemessene Faktoren (z. B. Krankheitscharakteristika, Patientencharakteristika, Behandlungscharakteristika usw.) sowie der fehlenden Bestätigung des Ansprechens durch ein unabhängiges Komitee in den externen Kontrollstudien begründet [45]. Weiterhin werden potentielle Unterschiede zwischen den Studien angemerkt, welche die Ein- und Ausschlusskriterien für die Patientinnen und Patienten, die Häufigkeit der Endpunkterhebung und die Endpunktdefinition einschließen. Eine Adjustierung für potentielles Confounding oder Bias sei aufgrund des Fehlens von Kovariateninformationen für die externen Kontrollstudien nicht möglich, sodass die Ergebnisse der Vergleiche nicht interpretierbar sind [45]. Die EMA sieht insgesamt ähnliche Limitationen und kommt zu dem Schluss, dass auch wenn ein positiver Effekt für Selumetinib auf das PFS beobachtet werden kann das genaue Ausmaß dieses Effekts unbekannt bleibt. Aussagekräftige, vergleichende Analysen der SPRINT-Studie mit den externen Kontrollstudien werden aufgrund der Unterschiede in den Studienpopulationen und möglichem Bias von der EMA als schwer durchführbar eingeschätzt.

4.3 Mortalität

Weder in Phase I noch in Phase II wurden Todesfälle verzeichnet. Eine Interpretation und Bewertung dieses Ergebnisses ist aufgrund der fehlenden Kontrollgruppe nicht möglich. Ein Effekt von Selumetinib auf die Mortalität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.4 Morbidität

Die Erhebung von Morbiditätsendpunkten erfolgte ausschließlich in Phase II der Studie SPRINT. Für die Nutzenbewertung werden die Morbiditätsendpunkte „Intensität des schlimmsten Schmerzes“, „Grooved Pegboard Test“, „PROMIS – Körperliche Funktionsfähigkeit Mobilität / Obere Extremitäten“, „Sehschärfe“, „Symptomcheckliste“ und „Global Impression of Change“ als patientenrelevante Endpunkte herangezogen. Insgesamt ist anzumerken, dass viele der Morbiditätsendpunkte nur in Teilpopulationen mit spezifischen Funktionseinschränkungen, z. B. mit Motordysfunktion, erhoben wurden. Dies ist aufgrund der vielfältigen und von Person zu Person sehr unterschiedlichen Symptome der Erkrankung größtenteils nachvollziehbar. Jedoch kann je nach Lokalisation des Tumors nicht immer ausgeschlossen werden, dass bei einem Tumorwachstum auch weitere Dysfunktionen auftreten können. Kritisch anzumerken ist, dass der Effekt der Behandlung mit Selumetinib auf ein wesentliches Symptom der Erkrankung, die Entstellung, unerfasst bleibt. Entstellungen wurden zu Baseline von 88 % der Studienpopulation der Phase II berichtet. Dieses Symptom wurde jedoch nur über Photographien dokumentiert, für die eine objektive Auswertung als nicht machbar eingestuft wurde. Eine direkte, patientenberichtete Erfassung erfolgte nicht. Die Auswirkung von Entstellungen könnte sich gegebenenfalls in der Lebensqualität widerspiegeln. Ebenfalls weitgehend unberücksichtigt bleibt innerhalb der Nutzenbewertung der Effekt von Selumetinib auf die Atemfunktion, da die Erfassung der Atemfunktion lediglich über nicht unmittelbar patientenrelevante Endpunkte erfolgte. Einige Symptome der Atemfunktion wie „Giemen“, „Schwierigkeiten beim Atmen“, „Kurzatmigkeit bei Belastung“ und „Kurzatmigkeit in Ruhe“ wurden jedoch über die Symptomcheckliste erfasst. Ähnlich ist die Situation für die Darm- und Blasenfunktion. Hier erfolgte die Erfassung über einen nicht ausreichend validierten Fragebogen, der nicht für die Nutzenbewertung herangezogen werden kann. Daneben erfolgte eine Erfassung der Symptome über die Symptomcheckliste mit Items für „Verstopfungen“, „Stuhlinkontinenz“, „Schmerzen beim Urinieren“, „Erhöhte Harnfrequenz / Erhöhter

Harndrang“, „Schwierigkeiten beim Beginn des Urinierens“ und „Harninkontinenz“. Im Bereich der Motorfunktion werden der Grooved Pegboard Test sowie die PROMIS-Skalen „Mobilität“ und „Obere Extremitäten“, welche der PROMIS-Domäne „Körperliche Funktionsfähigkeit“ zugeordnet sind, zur Nutzenbewertung herangezogen. Für den Grooved Pegboard Test werden leichte bis moderate Verbesserungen festgestellt. Allerdings kann aufgrund des einarmigen Studiendesigns nicht unterschieden werden, ob diese Verbesserungen durch die Behandlung mit Selumetinib oder bspw. durch Übungseffekte bedingt sind. Für Teilnehmende mit unilateralem PN sind die Verbesserungen für die beeinträchtigte Hand ähnlich stark ausgeprägt wie für die nicht beeinträchtigte Hand, was eher für den Einfluss von Übungseffekten sprechen würde (Z-Score zur Studienvs. vor Zyklus 13 (Median (min; max)): beeinträchtigte Hand: -0,50 (-11,0; 0,6); nicht-beeinträchtigte Hand: -0,47 (-12,2; 19,5)). Für die beiden PROMIS-Skalen „Mobilität“ und „Obere Extremitäten“ werden zur Studienvs. vor Zyklus 13 leichte Verbesserungen im FAS im Alter von 8 bis 18 Jahren beobachtet (Veränderung zu Baseline (LSM [95%-KI], p-Wert): „Mobilität“: 1,75 [-0,70; 4,19], 0,15; „Obere Extremitäten“: 1,76 [-0,88; 4,39], 0,18). Für das FAS mit einem Alter unter 8 Jahren und Motor-PN-bezogener Morbidität liegen keine separaten Auswertungen der Eltern-/Betreuerversion des PROMIS-Fragebogens vor, sondern nur Auswertungen der Eltern-/Betreuerversion für die gesamte Studienpopulation ab 5 Jahren, sodass diese Population in der Nutzenbewertung unberücksichtigt bleibt.

Auch in Bezug auf die Intensität des schlimmsten Schmerzes wurden Teilnehmende unter 8 Jahren nicht berücksichtigt, da sie von der patientenberichteten Schmerzerhebung ausgeschlossen waren. Von Teilnehmenden im Alter von 8 bis 18 Jahren wurde zur Studienvs. vor Zyklus 13 von 58,6 % (95%-KI: [38,9; 76,5]) eine Verbesserung um mindestens 2 Punkte auf der 11-stufigen NRS für den schlimmsten Schmerz eines von den Teilnehmenden bestimmten PN berichtet. Kein Kind oder Jugendlicher zeigte eine Verschlechterung um 2 Punkte oder mehr zur Studienvs. vor Zyklus 13. Bei der Interpretation der Responderanalysen ist zu beachten, dass zu Baseline 33,3 % der Teilnehmenden einen Wert < 2 auf der NRS für Schmerzintensität von 0 bis 10 aufwiesen (0 „keine Schmerzen“ bis 10 „schlimmste vorstellbare Schmerzen“) und somit die Responderdefinition für eine Verbesserung nicht erreichen konnten. Daneben wiesen 12,1 % der Studienteilnehmenden einen Wert > 8 auf und konnten sich demzufolge nicht um 2 Punkte verschlechtern. Die Ergebnisse deuten damit auf eine Verbesserung der Intensität des schlimmsten Schmerzes für die überwiegende Anzahl an Kindern und Jugendlichen im Alter von 8 bis 18 Jahren mit Schmerzen zu Baseline hin. Die Analysen zu Schmerzlinderungen, welche auch den Konsum vom Schmerzmitteln berücksichtigen, bestätigen die Ergebnisse. Die Ergebnissicherheit und Interpretierbarkeit der Ergebnisse ist aufgrund des offenen Studiendesigns ohne Kontrollgruppe jedoch eingeschränkt. Aufgrund des Bezugszeitraums von 7 Tagen und der Erhebung des Endpunkts im Abstand von ca. 3 Monaten ermöglicht die Operationalisierung keine Aussagen darüber, inwieweit eine dauerhafte Schmerzlinderung erreicht wird.

In die Analyse zur Sehschärfe konnten insgesamt nur 5 Teilnehmende für das vom PN-betroffene Auge einbezogen werden. Es wurden keine Verbesserungen und eine Verschlechterung um $\geq 0,2$ logMAR für das vom PN-betroffene Auge beobachtet. Die Ergebnissicherheit und Interpretierbarkeit der Ergebnisse ist aufgrund des offenen Studiendesigns ohne Kontrollgruppe und des kleinen Stichprobenumfangs jedoch sehr eingeschränkt.

Angaben zur mittels Symptomcheckliste erfassten Symptomatik liegen für die gesamte Studienpopulation vor, da die Symptomcheckliste von Eltern oder Betreuer ausgefüllt werden sollte, wenn die Teilnehmenden selbst dazu nicht in der Lage waren. Die Ergebnisse der Symptomcheckliste lassen sich folgendermaßen zusammenfassen: Für die Symptome „Schlafprobleme“, „Häufiges Aufwachen in der Nacht“, „Kopfschmerzen“, „Würgen“, „Schnarchen“, „Husten“, „Kurzatmigkeits bei Belastung“, „Schwäche“, „Muskelschmerzen“ und „Kribbeln“ wird häufiger (≥ 10 % Unterschied zur

Studienvisite vor Zyklus 13) von einer Verbesserung der Symptome als einer Verschlechterung berichtet. Umgekehrt wird für die Symptome „Anschwellen der Füße/Hände“ und „Diarrhö“ zum Zeitpunkt vor Zyklus 13 häufiger ($\geq 10\%$ Unterschied) über eine Verschlechterung der Symptome berichtet als über eine Verbesserung. Zu früheren Erhebungszeitpunkten wurde für die Symptome „Erhöhter Appetit“, „Unterleibsschmerzen“, „Übelkeit“, „Erbrechen“ und „Schwindel“ ebenfalls häufiger von Verschlechterungen als von Verbesserungen berichtet. Symptome, für die häufiger eine Verschlechterung berichtet wird, könnten eher die Verträglichkeit von Selumetinib abbilden als die Morbidität. Die Ergebnissicherheit und Interpretierbarkeit der Ergebnisse ist aufgrund des offenen Studiendesigns ohne Kontrollgruppe jedoch eingeschränkt.

Zur Studienvisite vor Zyklus 13 berichten 52 bzw. 59 % der Kinder und Jugendlichen, welche den GIC selbst ausfüllten (Alter 8–18 Jahre), ihr Tumorschmerz bzw. ihre Tumor-assoziierte Morbidität sei „viel besser“ oder „sehr viel besser“ im Vergleich zu vor Beginn der Behandlung. Für den Gesamtschmerz berichteten 41 % einen „viel besseren“ oder „sehr viel besseren“ Zustand zur Studienvisite vor Zyklus 13. Eine Zusammenfassung über die Antwortkategorien für Verschlechterungen im GIC liegt nicht vor. Die Darstellung aller Antwortkategorien (Tabelle 4-168 im Dossier) zeigt, dass Verschlechterungen insgesamt nur selten (maximale Häufigkeit $\leq 10\%$) berichtet werden. Bei der Interpretation der Ergebnisse ist zu beachten, dass die Erfassung von Veränderungen mittels GIC in der Forschung kritisch diskutiert wird [23,26,39,40], da die Befragten bei der rückwirkenden Einschätzung einer Veränderung in der Lage sein müssen, sowohl ihren gegenwärtigen Zustand als auch ihren Ausgangszustand zu bewerten und einen Unterschied im Sinne einer Verbesserung, Verschlechterung oder keiner Veränderung zu quantifizieren. Größtenteils werde bei dieser Abschätzung allerdings der gegenwärtige Gesundheitszustand erfasst. Dieser Effekt sei umso größer, je länger das Erinnerungsintervall ist (Recall Bias). Die Interpretation der Ergebnisse für den GIC fällt in diesem Zusammenhang schwer, insbesondere weil es sich bei den Befragten um Kinder und Jugendliche handelte, denen diese vergleichende Einschätzung evtl. schwerfällt, und der Zeitraum bis zur letzten berücksichtigten Studienvisite vor Zyklus 13 mit ca. einem Jahr sehr lang ist. Insgesamt ist aber davon auszugehen, dass der Therapieeffekt von den Teilnehmenden eher positiv wahrgenommen wurde. Für das FAS mit einem Alter unter 8 Jahren liegen keine separaten Auswertungen der Eltern-/Betreuerversion des GIC vor, sondern nur Auswertungen der Eltern-/Betreuerversion für die gesamte Studienpopulation ab 5 Jahren vor, sodass diese Population in den Nutzenbewertung unberücksichtigt bleibt.

Bei allen Auswertungen zur Morbidität ist zu beachten, dass keine Erhebung der Morbiditätseindpunkte nach Abbruch der Einnahme der Studienmedikation erfolgte. Die gewonnenen Daten spiegeln dadurch nicht die Morbidität der in die Studie eingeschlossenen gesamten Studienpopulation wider. Weiterhin beziehen die Auswertungen immer nur Teilnehmende mit Angaben zum jeweiligen Erhebungszeitpunkt und zu Baseline ein. In Abhängigkeit von der Rücklauf- bzw. Teilnahmequote bzw. des Gesundheitszustands der Non-Responder können die Angaben für die relativen Häufigkeiten oder Veränderungen über die Zeit daher verzerrt sein.

Zusammenfassende Einschätzung zur Morbidität

Innerhalb der Phase II der Studie SPRINT zeigen sich im zeitlichen Verlauf Verbesserungen für den schlimmsten Schmerz und der mittels GIC berichteten Globalbeurteilung der klinischen Veränderung für Teilnehmende im Alter von 8 bis 18 Jahren. Für jüngere Studienteilnehmende liegen keine (separaten) Auswertungen vor. Eine Interpretation und Bewertung der Veränderungen in der Morbidität ist aufgrund der fehlenden Kontrollgruppe nicht möglich. Der Effekt von Selumetinib auf die Morbidität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.5 Lebensqualität

Die Lebensqualität wurde in der Studie SPRINT mithilfe des PedsQL für Kinder und Jugendliche im Alter von 8 bis 18 Jahren durch die Patientenversion des PedsQL und für die gesamte Studienpopulation mit der Eltern-/Betreuerversion des PedsQL erhoben. Die Erhebung erfolgte ausschließlich in Phase II. Die patientenberichteten Baseline-Werte für den PedsQL sind im Median bereits relativ hoch (Median 82 für Gesamtwert der Skalenspannweite von 0 bis 100), was auf eine hohe allgemeine gesundheitsbezogene Lebensqualität vieler Teilnehmenden zu Baseline schließen lässt. Da es sich beim PedsQL jedoch um einen generischen Fragebogen zur Lebensqualität handelt, erscheint es möglich, dass indikationsspezifische Aspekte der Lebensqualität, wie bspw. Stigmatisierung aufgrund von Entstellungen, durch den PedsQL nicht ausreichend erfasst werden. Ein indikationsspezifischer Fragebogen wurde nicht eingesetzt.

Die MMRM-Analysen des selbstberichteten PedsQL-Gesamtwerts für Teilnehmende im Alter von 8 bis 18 Jahren weisen auf eine Verbesserung der Lebensqualität im Studienverlauf hin (Veränderung zu Baseline (LSM [95%-KI], p-Wert): 6,68 Punkte [1,34; 12,03], 0,02).

Die Responderanalysen zeigen für 24,1 % der Studienteilnehmenden (95%-KI: [10,3; 43,5]) im Alter von 8 bis 18 Jahren eine Verbesserung um 15 Punkte im Gesamtwert. Der Anteil an Respondern für eine Verbesserung ist für die Subskalen „Körperliche Funktion“ (34,5 % [17,9; 54,3]), „Emotionale Funktion“ (37,9 % [20,7; 57,7]) und „Soziale Funktion“ (31,0 % [15,3; 50,8]) sogar noch etwas höher. Die Responderanalysen der 4. Subskala „Schulische Funktion“ kann leider nicht zur Nutzenbewertung herangezogen werden, da die Rücklaufquote, welche bereits zu Baseline niedriger war als die der anderen Subskalen, im Studienverlauf auf unter 70 % fällt. Der Median der Subskala zur schulischen Funktion lag zu Baseline bei 75 Punkten und war damit niedriger als der Median der 3 anderen Subskalen („Körperliche Funktion“: 84 Punkte; „Emotionale Funktion“: 85 Punkte; „Soziale Funktion“: 80 Punkte). Es muss daher davon ausgegangen werden, dass die Einschränkungen der schulischen Funktion für die Patientenpopulation am stärksten sind. Daher kann die fehlende Beantwortung der Items zur schulischen Funktion im Studienverlauf die Ergebnisse des Gesamtwerts zugunsten von Selumetinib verzerren. Responderanalysen für Verschlechterungen um 15 Punkte konnten in den Studienunterlagen nicht identifiziert werden. Ebenfalls wurden keine separaten Auswertungen der Eltern-/Betreuerversion des PedsQL für Kinder < 8 Jahre vorgelegt, um deren Lebensqualität abzubilden.

Auch bei den Auswertungen zur Lebensqualität ist zu beachten, dass keine Erhebung nach Abbruch der Einnahme der Studienmedikation erfolgte. Die gewonnenen Daten spiegeln dadurch nicht die Lebensqualität der in die Studie eingeschlossenen gesamten Studienpopulation wider. Weiterhin beziehen die Auswertungen immer nur Teilnehmende mit Angaben zum jeweiligen Erhebungszeitpunkt und zu Baseline ein. In Abhängigkeit von der Rücklauf- bzw. Teilnahmequote bzw. des Gesundheitszustands der Non-Responder können die Angaben für die relativen Häufigkeiten oder Veränderungen über die Zeit daher verzerrt sein.

Zusammenfassende Einschätzung zur Lebensqualität

Die Ergebnisse der Phase II der Studie SPRINT weisen auf eine Verbesserung der allgemeinen, gesundheitsbezogenen Lebensqualität im Studienverlauf hin. Allerdings sind Verzerrungen durch Nonresponse der Subskala zur schulischen Funktion zugunsten von Selumetinib möglich. Eine Interpretation und Bewertung der Veränderungen in der Lebensqualität ist aufgrund der fehlenden Kontrollgruppe nicht möglich. Der Effekt von Selumetinib auf die Lebensqualität kann auf Grundlage der vorgelegten Daten nicht beurteilt werden.

4.6 Sicherheit

Angaben zur Dauer der Beobachtungszeit für die Sicherheit liegen nicht vor. Die mediane Behandlungsdauer mit Selumetinib wird mit 1.597 Tagen (\pm 52,5 Monaten oder 4,4 Jahren) für Phase I und 801,5 Tagen (\pm 26,4 Monaten) bzw. 1.027,5 Tagen (\pm 33,8 Monaten) für die Datenschnitte vom 29.06.2018 bzw. 29.03.2019 für Phase II angegeben. Die Nachbeobachtung der Sicherheit sollte bis 30 Tage nach Beendigung der Behandlung erfolgen.

Innerhalb der Phase II wurden bei 62 % der Teilnehmenden UE Grad 3 oder höher, bei 24 % SUE und bei 12 % UE, die zum Abbruch der Studienmedikation führten, beobachtet. Teilnehmende waren häufig von UE folgender Systemorganklassen betroffen: „Erkrankungen der Haut und des subkutanen Gewebes“ (98 %), „Untersuchungen“ (98 %), „Erkrankungen des Gastrointestinaltrakts“ (96 %), „Stoffwechsel- und Ernährungsstörungen“ (96 %), „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (84%), „Erkrankungen der Atemwege, des Brustraums und des Mediastinums“ (82 %), „Infektionen und Infestationen“ (82 %), „Erkrankungen der Muskeln und des Bindegewebes“ (70 %), „Erkrankungen des Nervensystems“ (70 %), „Erkrankungen der Nieren und Harnwege“ (50%) sowie „Erkrankungen des Bluts und des Lymphsystems“ (50%). UE CTCAE-Grad 3 oder höher betrafen die Systemorganklassen „Erkrankungen des Gastrointestinaltrakts“ (24 %), „Untersuchungen“ (20 %), „Infektionen und Infestationen“ (18 %), „Erkrankungen der Haut und des subkutanen Gewebes“ (10%), „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (8%), „Erkrankungen der Atemwege, des Brustraums und des Mediastinums“ (8%), „Erkrankungen des Nervensystems“ (6%) und „Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen“ (6%). Mit einer Inzidenz von über 5 % traten SUE in den Systemorganklassen „Infektionen und Infestationen“ (12 %) und „Erkrankungen des Gastrointestinaltrakts“ (6 %) auf.

Die Häufigkeit des Auftretens von UE/SUE in Phase I ist insgesamt vergleichbar mit Phase II, aufgrund der kleinen Stichprobengröße ergeben sich gelegentlich kleinere oder größere Abweichungen der Inzidenzen. Eine (meta-analytische) Zusammenfassung der Inzidenzen von UE in Phase I (FI-konforme Dosierung) und Phase II wurde nicht vorgelegt.

Für die Interpretation der Ergebnisse zu den UE ist zu beachten, dass möglicherweise Ereignisse der Grunderkrankung, z. B. Pruritus als UE erfasst wurden.

Als wesentliche UE, die das Sicherheitsprofil von Selumetinib charakterisieren werden „Erbrechen“, „Ausschlag“, „Kreatinphosphokinase im Blut erhöht“, „Diarrhö“, „Nausea“, „Asthenische Ereignisse“, „Trockene Haut“, „Fieber“, „Akneiformer Hautausschlag“, „Hypoalbuminämie“, „Aspartat-Aminotransferase erhöht“ und „Paronychie“ von der EMA benannt.

Zusammenfassende Einschätzung zur Sicherheit

Die Sicherheit der Therapie mit Selumetinib kann aufgrund der fehlenden Kontrollgruppe nicht abschließend bewertet werden. Schwere UE treten bei ungefähr zwei Dritteln der Teilnehmenden auf (62 %), SUE bei ca. einem Viertel (24 %). Ein kleiner Teil der Patientinnen und Patienten (ca. 12 %) bricht die Therapie aufgrund von UE ab.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen. Die Einleitung und Überwachung der Therapie mit Selumetinib soll durch in der Therapie von Patientinnen und Patienten mit NF1-bedingten Tumoren erfahrene Fachärztinnen und -ärzte für Innere Medizin und Hämatologie und Onkologie erfolgen.

6 Zusammenfassung der Nutzenbewertung

Selumetinib ist zugelassen für Kinder ab 3 Jahren und Jugendliche zur Behandlung von symptomatischen, inoperablen plexiformen Neurofibromen (PN) bei Neurofibromatose Typ 1 (NF1). Die Nutzenbewertung von Selumetinib basiert auf der zulassungsbegründenden Studie SPRINT, einer offenen, einarmigen, multizentrischen Studie zur Sicherheit und Wirksamkeit von Selumetinib bei Kindern und Jugendlichen mit symptomatischen, inoperablen PN aufgrund von NF1.

Die vom pU vorgelegten Daten zu indirekten Vergleichen mit der Studie SPRINT wurden aufgrund fehlender Vergleichbarkeit mit der Studienpopulation und fehlender Patientenrelevanz der untersuchten Endpunkte als für die Nutzenbewertung ungeeignet eingeschätzt und nicht herangezogen.

Die Ergebnisse der patientenrelevanten Endpunkte der Studie SPRINT, Phase II werden in der folgenden Tabelle zusammengefasst. 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 sowie 3.2 bis 3.5 der Nutzenbewertung.

Tabelle 27: Zusammenfassende Darstellung der Ergebnisse; FAS der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt: 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Zusammenfassende Darstellung der Ergebnisse		Selumetinib N = 50	
Endpunkt	n	Ergebnis	
Mortalität			
Todesfälle	50	n (%): 0	
Morbidität			
<i>Schlimmster Schmerz Patienten-bestimmtes PN (NRS-11)</i>	34 ¹⁾		
Verbesserung um ≥ 2 Punkte zur Visite vor Zyklus 13	29	n (%): 17 (58,6)	
Verschlechterung um ≥ 2 Punkte zur Visite vor Zyklus 13	29	n: 0	
<i>Grooved Pegboard Test [z-Score]²⁾</i>	25 ³⁾		
<i>Dominante Hand</i>			
Baseline	25	Median (min; max): 1,1 (-1,0; 12,4)	
Veränderung von Baseline zur Visite vor Zyklus 13	22	Median (min; max): -0,35 (-12,2; 19,5)	
<i>Nicht dominante Hand</i>			
Baseline	24	Median (min; max): 2,2 (-0,9; 39,2)	
Veränderung von Baseline zur Visite vor Zyklus 13	21	Median (min; max): -0,62 (-11,0; 10,1)	
<i>PROMIS⁴⁾</i>	33 ⁵⁾		
<i>Mobilität</i>			
Baseline	24 ⁵⁾	MW (SD): 46,57 (6,54)	
Veränderung von Baseline zur Visite vor Zyklus 13	20	LSM [95%-KI]: 1,75 [-0,70; 4,19]; p = 0,15	
<i>Obere Extremitäten</i>			
Baseline	24 ⁵⁾	MW (SD): 45,95 (12,91)	
Veränderung von Baseline zur Visite vor Zyklus 13	19	LSM [95%-KI]: 1,76 [-0,88; 4,39]; p = 0,18	
<i>Sehschärfe⁶⁾</i>	10 ⁷⁾		
Baseline (logMAR)	5	MW (SD): 0,54 (0,38)	
Verbesserung um $\geq 0,2$ logMAR zur Visite vor Zyklus 25	4	n: 0	
Verschlechterung um $\geq 0,2$ logMAR zur Visite vor Zyklus 25	4	n (%): 1 (25)	

Studie SPRINT, Phase II, Stratum 1 Zusammenfassende Darstellung der Ergebnisse		Selumetinib N = 50	
Endpunkt	n	Ergebnis	
<i>Symptomcheckliste</i>	50		
Müdigkeit/Fatigue Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 17 (38,6) n (%): 13 (29,5)	
Schlafprobleme Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 17 (38,6) n (%): 6 (13,6)	
Verringerter Appetit Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 10 (22,7) n (%): 6 (13,6)	
Erhöhter Appetit Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 10 (22,7) n (%): 10 (22,7)	
Kopfschmerzen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 16 (36,4) n (%): 6 (13,6)	
Sehstörungen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 1 (2,3)	
Vermindertes Hörvermögen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n (%): 4 (9,1)	
Wunde Stellen im Mund Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 6 (13,6)	
Schwierigkeiten beim Schlucken Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 1 (2,3)	
Würgen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 6 (13,6) n: 0	
Schnarchen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 15 (34,1) n (%): 3 (6,8)	
Häufiges Aufwachen in der Nacht Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 17 (38,6) n (%): 7 (15,9)	
Husten Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 17 (38,6) n (%): 8 (18,2)	
Giemen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 1 (2,3)	
Schwierigkeiten beim Atmen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n (%): 1 (2,3)	

Studie SPRINT, Phase II, Stratum 1 Zusammenfassende Darstellung der Ergebnisse		Selumetinib N = 50	
Endpoint	n	Ergebnis	
Schmerzen in der Brust Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 1 (2,3) n (%): 3 (6,8)	
Herzklopfen/-flattern Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n (%): 2 (4,5)	
Kurzatmigkeit bei Belastung Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 9 (20,5) n (%): 3 (6,8)	
Kurzatmigkeit in Ruhe Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 2 (4,5) n: 0	
Anschwellen der FüÙe/Hände Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n: 0 n (%): 6 (13,6)	
Unterleibsschmerzen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 10 (22,7) n (%): 10 (22,7)	
Sodbrennen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 2 (4,5) n (%): 2 (4,5)	
Ùbelkeit Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 8 (18,2) n (%): 10 (22,7)	
Erbrechen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n (%): 5 (11,4)	
Diarrhö Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 10 (22,7)	
Verstopfungen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 5 (11,4) n (%): 6 (13,6)	
Stuhlinkontinenz Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 4 (9,1) n (%): 3 (6,8)	
Schmerzen beim Urinieren Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 4 (9,1) n (%): 3 (6,8)	
Erhöhte Harnfrequenz / Erhöhter Harndrang Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 6 (13,6) n (%): 2 (4,5)	
Schwierigkeiten beim Beginn des Urinierens Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n (%): 2 (4,5)	

Studie SPRINT, Phase II, Stratum 1 Zusammenfassende Darstellung der Ergebnisse		Selumetinib N = 50	
Endpunkt	n	Ergebnis	
Harninkontinenz Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 6 (13,6) n (%): 2 (4,5)	
Schwäche Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 15 (34,1) n (%): 2 (4,5)	
Muskelschmerzen Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 16 (36,4) n (%): 4 (9,1)	
Schwindel Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 4 (9,1) n (%): 5 (11,4)	
Taubheitsgefühl Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 3 (6,8) n: 0	
Kribbeln Verbesserung ⁸⁾ zur Visite vor Zyklus 13 Verschlechterung ⁹⁾ zur Visite vor Zyklus 13	44	n (%): 6 (13,6) n: 0	
<i>Global Impression of Change</i>	34 ¹⁰⁾		
Tumorschmerz Verbesserung ¹¹⁾ zur Visite vor Zyklus 13 Verschlechterung ¹²⁾ zur Visite vor Zyklus 13	29	n (%): 15 (51,7) n (%): k. A.	
Gesamtschmerz Verbesserung ¹¹⁾ zur Visite vor Zyklus 13 Verschlechterung ¹²⁾ zur Visite vor Zyklus 13	29	n (%): 12 (41,4) n (%): k. A.	
Tumor-assoziierte Morbidität Verbesserung ¹¹⁾ zur Visite vor Zyklus 13 Verschlechterung ¹²⁾ zur Visite vor Zyklus 13	29	n (%): 17 (58,6) n (%): k. A.	
Lebensqualität			
<i>PedsQL</i>	34 ¹⁰⁾		
Gesamtwert Verbesserung ¹³⁾ zur Visite vor Zyklus 13 Verschlechterung ¹⁴⁾ zur Visite vor Zyklus 13	29	n (%): 7 (24,1) n (%): k. A.	
Körperliche Funktion Verbesserung ¹³⁾ zur Visite vor Zyklus 13 Verschlechterung ¹⁴⁾ zur Visite vor Zyklus 13	29	n (%): 10 (34,5) n (%): k. A.	
Emotionale Funktion Verbesserung ¹³⁾ zur Visite vor Zyklus 13 Verschlechterung ¹⁴⁾ zur Visite vor Zyklus 13	29	n (%): 11 (37,9) n (%): k. A.	
Soziale Funktion Verbesserung ¹³⁾ zur Visite vor Zyklus 13 Verschlechterung ¹⁴⁾ zur Visite vor Zyklus 13	29	n (%): 9 (31,0) n (%): k. A.	
Schulische Funktion	23	_15)	

Studie SPRINT, Phase II, Stratum 1 Zusammenfassende Darstellung der Ergebnisse		Selumetinib N = 50
Endpoint	n	Ergebnis
Sicherheit		
UE CTCAE-Grad ≥ 3	50	n (%): 31 (62)
SUE	50	n (%): 12 (24)
UE, das zum Abbruch der Studienmedikation führte ¹⁶⁾	50	n (%): 6 (12)
<i>UE von besonderem Interesse</i>		
Muskelbezogene Effekte	50	n (%): 43 (86)
Ausschlag, nicht akneiform	50	n (%): 35 (70)
Ausschlag, akneiform	50	n (%): 25 (50)
Effekte der Entzündung der oralen Mundschleimhaut	50	n (%): 25 (50)
Leukopenie	50	n (%): 22 (44)
Erythropenie	50	n (%): 21 (42)
Herzinsuffizienz	50	n (%): 18 (36)
Die Netzhaut betreffende Effekte	50	n (%): 8 (16)
Thrombozytopenie	50	n (%): 6 (12)
Physeale Dysplasie	50	n: 0

¹⁾ FAS im Alter von 8 bis 18 Jahren zu Baseline.

²⁾ Aufgrund fehlender Angaben zur Normstichprobe kann die Validität und Repräsentativität der Normstichprobe für die Berechnung der Z-Scores nicht beurteilt werden.

³⁾ FAS ab 5 Jahren mit einem PN in den oberen Extremitäten oder mit bekannter Kompression des zervikalen oder oberen thorakalen Rückenmarks

⁴⁾ Dargestellt sind T-Scores basierend auf der US-Allgemeinbevölkerung mit einem Mittelwert von 50 und einer Standardabweichung von 10. Höhere Werte entsprechen einer besseren körperlichen Funktion

⁵⁾ FAS mit Motor-PN-bezogenen Morbidität. Patientenberichtete Ergebnisse liegen nur für Teilnehmende im Alter von 8 bis 18 Jahren vor (N = 24). Für die Eltern-/Betreuerversion des PROMIS liegen keine separaten Auswertungen für Kinder < 8 Jahre vor.

⁶⁾ Sehschärfe des von PN betroffenen Auges ermittelt mittels HOTV-Sehtafel.

⁷⁾ FAS mit PN der Augenhöhle. Für 5 Personen war keine Verbesserung der Sehschärfe möglich, da das durch das PN betroffene Auge entfernt worden war (n = 2), das Sehen auf Licht- und Lagewahrnehmung begrenzt war (n = 1) oder die Betroffenen keine Lichtwahrnehmung (n = 2) mehr aufwiesen. Diese Personen wurden in der Auswertung nicht berücksichtigt.

⁸⁾ Verbesserung um mindestens eine Antwortkategorie.

⁹⁾ Verschlechterung um mindestens eine Antwortkategorie.

¹⁰⁾ Ergebnisse liegen nur für Teilnehmende im Alter von 8 bis 18 Jahren vor (N = 34). Für die Eltern-/Betreuerversion liegen keine separaten Auswertungen für Kinder < 8 Jahre vor.

¹¹⁾ Antwortkategorien „sehr viel besser“ oder „viel besser“.

¹²⁾ Für „Verschlechterungen“ wurde keine über die Antwortkategorien aggregierte Analyse vorgelegt.

¹³⁾ Verbesserung um mindestens 15 Punkte.

¹⁴⁾ Verschlechterung um mindestens 15 Punkte.

¹⁵⁾ Keine Darstellung da Rücklaufquote unter 70 %.

¹⁶⁾ Die Studienteilnehmenden erhielten die Studienmedikation bis zum Auftreten einer Krankheitsprogression, inakzeptabler UE, der Rücknahme der Einverständniserklärung oder der Entscheidung des ärztlichen Prüfpersonals, je nachdem, was früher auftrat. Diese möglichen Therapieabbruchgründe, die vor einem potentiellen Abbruch aufgrund von UE auftreten können, stellen somit ein konkurrierendes Ereignis dar, weshalb die Ergebnissicherheit und Interpretierbarkeit der Ergebnisse eingeschränkt ist.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; FAS: Full Analysis Set; GIC: Global Impression of Change; k. A.: keine Angabe; KI: Konfidenzintervall; logMAR: Logarithm of the Minimum Angle of Resolution; LSM: Least Square Mean; MW: Mittelwert; PedsQL: Pediatric Quality of Life Inventory; NRS-11: Numerical Rating Scale; PN: Plexiformes Neurofibrom; PROMIS: Patient-Reported Outcomes Measurement Information System; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **Afshar K, Mirbagheri A, Scott H, MacNeily AE.** Development of a symptom score for dysfunctional elimination syndrome. *J Urol* 2009;182(4 Suppl):1939-1943.
2. **Akshintala S, Baldwin A, Liewehr DJ, Goodwin A, Blakeley JO, Gross AM, et al.** Longitudinal evaluation of peripheral nerve sheath tumors in neurofibromatosis type 1: growth analysis of plexiform neurofibromas and distinct nodular lesions. *Neuro Oncol* 2020;22(9):1368-1378.
3. **American Thoracic Society.** ATS Statement: Guidelines for the Six-Minute Walk Test. *American Journal of Respiratory and Critical Care Medicine* 2002;166(1):111-117.
4. **AstraZeneca.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A: Selumetinib, Behandlung von symptomatischen, inoperablen plexiformen Neurofibromen bei Kindern ab 3 Jahren und Jugendlichen mit Neurofibromatose Typ 1; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 12.08.2021.
5. **AstraZeneca.** Koselugo 10 mg / 25 mg Hartkapseln [online]. 06.2021. Berlin. [Zugriff: 19.08.2021]. URL: <http://www.fachinfo.de>.
6. **AstraZeneca.** A phase I/II study of the mitogen activated protein kinase Kinase (MEK) 1 inhibitor selumetinib (AZD6244; HYD Sulfate) in children with neurofibromatosis type 1 (NF1) and inoperable plexiform neurofibromas (PN) (SPRINT phase I); clinical study report [unveröffentlicht]. 08.08.2019.
7. **AstraZeneca.** A phase I/II study of the mitogen activated protein kinase Kinase (MEK) 1 inhibitor selumetinib (AZD6244; HYD Sulfate) in children with neurofibromatosis type 1 (NF1) and inoperable plexiform neurofibromas (PN) (SPRINT phase II stratum 1); clinical study report [unveröffentlicht]. 28.08.2019.
8. **AstraZeneca.** Selumetinib for the treatment of paediatric patients aged 3 years and above, with neurofibromatosis type 1 (NF1) and symptomatic, inoperable plexiform neurofibromas (PN); 90 days safety update [unveröffentlicht]. 03.12.2019.
9. **Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J.** Validation and properties of the verbal numeric scale in children with acute pain. *PAIN* 2010;149(2):216-221.
10. **Baldo F, Grasso AG, Cortellazzo Wiel L, Maestro A, Trojniak MP, Murru FM, et al.** Selumetinib in the treatment of symptomatic intractable plexiform neurofibromas in neurofibromatosis type 1: a prospective case series with emphasis on side effects. *Paediatr Drugs* 2020;22(4):417-423.
11. **Ciesla N, Dinglas V, Fan E, Kho M, Kuramoto J, Needham D.** Manual muscle testing: a method of measuring extremity muscle strength applied to critically ill patients. *J Vis Exp* 2011(50):2632.
12. **DeWalt DA, Gross HE, Gipson DS, Selewski DT, DeWitt EM, Dampier CD, et al.** PROMIS® pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. *Quality of Life Research* 2015;24(9):2195-2208.

13. **DeWitt EM, Stucky BD, Thissen D, Irwin DE, Langer M, Varni JW, et al.**
Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory.
Journal of clinical epidemiology 2011;64(7):794-804.
14. **Dombi E, Baldwin A, Marcus LJ, Fisher MJ, Weiss B, Kim A, et al.**
Activity of selumetinib in neurofibromatosis type 1-related plexiform neurofibromas.
N Engl J Med 2016;375(26):2550-2560.
15. **Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al.**
Core outcome measures for chronic pain clinical trials: IMMPACT recommendations.
Pain 2005;113(1-2):9-19.
16. **Escolar DM, Henricson EK, Mayhew J, Florence J, Leshner R, Patel KM, et al.**
Clinical evaluator reliability for quantitative and manual muscle testing measures of strength in children. Muscle & Nerve 2001;24(6):787-793.
17. **Espírito Santo V, Passos J, Nzwalo H, Carvalho I, Santos F, Martins C, et al.**
Selumetinib for plexiform neurofibromas in neurofibromatosis type 1: a single-institution experience. J Neurooncol 2020;147(2):459-463.
18. **Fisher MJ, Avery RA, Allen JC, Ardern-Holmes SL, Bilaniuk LT, Ferner RE, et al.**
Functional outcome measures for NF1-associated optic pathway glioma clinical trials.
Neurology 2013;81(21 supplement 1):S15-S24.
19. **Fisher MJ, Basu S, Dombi E, Yu JQ, Widemann BC, Pollock AN, et al.**
The role of [18F]-fluorodeoxyglucose positron emission tomography in predicting plexiform neurofibroma progression. J Neurooncol 2008;87(2):165-171.
20. **Gajdosik RL, Bohannon RW.** Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. Phys Ther 1987;67(12):1867-1872.
21. **Geiger R, Strasak A, Tremel B, Gasser K, Kleinsasser A, Fischer V, et al.** Six-minute walk test in children and adolescents. J Pediatr 2007;150(4):395-399, 399.e391-392.
22. **Getz LM, Dobson V, Luna B, Mash C.** Interobserver reliability of the Teller Acuity Card procedure in pediatric patients.
Investigative Ophthalmology & Visual Science 1996;37(1):180-187.
23. **Guyatt GH, Norman GR, Juniper EF, Griffith LE.** A critical look at transition ratings.
J Clin Epidemiol 2002;55(9):900-908.
24. **Holmes JM, Beck RW, Repka MX, Leske DA, Kraker RT, Blair RC, et al.**
The amblyopia treatment study visual acuity testing protocol.
Archives of Ophthalmology 2001;119(9):1345-1353.
25. **Irwin DE, Gross HE, Stucky BD, Thissen D, DeWitt EM, Lai JS, et al.**
Development of six PROMIS pediatrics proxy-report item banks.
Health and Quality of Life Outcomes 2012;10(1):22.
26. **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.

27. **Lafayette Instrument.** Grooved Pegboard Test user instructions - model 32025 [online]. 2002. [Zugriff: 19.08.2021]. URL: <https://www.advys.be/docs/GroovedPegboardTestManual.pdf>.
28. **Leone JF, Gole GA, Mitchell P, Kifley A, Pai ASI, Rose KA.** Visual acuity testability and comparability in Australian preschool children: The Sydney Paediatric Eye Disease Study. *Eye* 2012;26(7):925-932.
29. **Martin S, Nelson Schmitt S, Wolters PL, Abel B, Toledo-Tamula MA, Baldwin A, et al.** Development and validation of the English Pain Interference Index and Pain Interference Index-Parent report. *Pain Med* 2015;16(2):367-373.
30. **Martin S, Wolters PL, Toledo-Tamula MA, Schmitt SN, Baldwin A, Starosta A, et al.** Acceptance and commitment therapy in youth with neurofibromatosis type 1 (NF1) and chronic pain and their parents: A pilot study of feasibility and preliminary efficacy. *Am J Med Genet A* 2016;170(6):1462-1470.
31. **Mash C, Dobson V, Carpenter N.** Interobserver agreement for measurement of grating acuity and interocular acuity differences with the Teller acuity card procedure. *Vision Research* 1995;35(2):303-312.
32. **Mathiowetz V, Wiemer DM, Federman SM.** Grip and pinch strength: norms for 6- to 19-year-olds. *American Journal of Occupational Therapy* 1986;40(10):705-711.
33. **Medical Research Council.** Aids to the examination of the peripheral nervous system,. London: Her Majesty's Stationery Office; 1976.
34. **Miró J, Castarlenas E, Huguet A.** Evidence for the use of a numerical rating scale to assess the intensity of pediatric pain. *European Journal of Pain* 2009;13(10):1089-1095.
35. **Mylius CF, Paap D, Takken T.** Reference value for the 6-minute walk test in children and adolescents: a systematic review. *Expert Review of Respiratory Medicine* 2016;10(12):1335-1352.
36. **National Cancer Institute (NCI).** Natural history study and longitudinal assessment of children, adolescents, and adults with neurofibromatosis type 1; study protocol 08-C-0079 I [unveröffentlicht]. 16.01.2017.
37. **Nguyen R, Dombi E, Widemann BC, Solomon J, Fuensterer C, Kluwe L, et al.** Growth dynamics of plexiform neurofibromas: a retrospective cohort study of 201 patients with neurofibromatosis 1. *Orphanet J Rare Dis* 2012;7(75):1-6.
38. **Nguyen R, Ibrahim C, Friedrich RE, Westphal M, Schuhmann M, Mautner VF.** Growth behavior of plexiform neurofibromas after surgery. *Genet Med* 2013;15(9):691-697.
39. **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.
40. **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.

41. **Plotkin SR, Davis SD, Robertson KA, Akshintala S, Allen J, Fisher MJ, et al.** Sleep and pulmonary outcomes for clinical trials of airway plexiform neurofibromas in NF1. *Neurology* 2016;87(7 Suppl 1):S13-20.
42. **Siu K, Al-Harbi S, Clark H, Thabane L, Cheng J, Tarnopolsky M, et al.** Feasibility and Reliability of Muscle Strength Testing in Critically Ill Children. *Journal of pediatric intensive care* 2015;4(4):218-224.
43. **Teller DY, McDonald MA, Preston K, Sebris SL, Dobson V.** Assessment of visual acuity in infants and children; the acuity card procedure. *Developmental Medicine & Child Neurology* 1986;28(6):779-789.
44. **Thissen D, Liu Y, Magnus B, Quinn H, Gipson DS, Dampier C, et al.** Estimating minimally important difference (MID) in PROMIS pediatric measures using the scale-judgment method. *Qual Life Res* 2016;25(1):13-23.
45. **U.S. Food and Drug Administration (FDA).** Koselugo (selumetinib): application number 213756Orig1s000: multidiscipline review [online]. 03.04.2020. Silver Spring (USA): FDA. [Zugriff: 19.08.2021]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/213756Orig1s000MultidisciplineR.pdf.
46. **Varni JW, Burwinkle TM, Seid M, Skarr D.** The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambul Pediatr* 2003;3(6):329-341.
47. **Varni JW, Magnus B, Stucky BD, Liu Y, Quinn H, Thissen D, et al.** Psychometric properties of the PROMIS (R) pediatric scales: precision, stability, and comparison of different scoring and administration options. *Qual Life Res* 2014;23(4):1233-1243.
48. **Varni JW, Seid M, Kurtin PS.** PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001;39(8):800-812.
49. **Varni JW, Thissen D, Stucky BD, Liu Y, Gorder H, Irwin DE, et al.** PROMIS(R) Parent Proxy Report Scales: an item response theory analysis of the parent proxy report item banks. *Qual Life Res* 2012;21(7):1223-1240.
50. **Varni JW, Thissen D, Stucky BD, Liu Y, Magnus B, Quinn H, et al.** PROMIS(R) Parent Proxy Report Scales for children ages 5-7 years: an item response theory analysis of differential item functioning across age groups. *Qual Life Res* 2014;23(1):349-361.
51. **Wang Y-C, Magasi SR, Bohannon RW, Reuben DB, McCreath HE, Bubela DJ, et al.** Assessing dexterity function: a comparison of two alternatives for the NIH Toolbox. *Journal of hand therapy : official journal of the American Society of Hand Therapists* 2011;24(4):313-321.
52. **Well L, Döbel K, Kluwe L, Bannas P, Farschtschi S, Adam G, et al.** Genotype-phenotype correlation in neurofibromatosis type-1: NF1 whole gene deletions lead to high tumor-burden and increased tumor-growth. *PLoS Genet* 2021;17(5):e1009517.
53. **Wicksell RK, Melin L, Lekander M, Olsson GL.** Evaluating the effectiveness of exposure and acceptance strategies to improve functioning and quality of life in longstanding pediatric pain--a randomized controlled trial. *Pain* 2009;141(3):248-257.

54. **Widemann BC, Dombi E, Gillespie A, Wolters PL, Belasco J, Goldman S, et al.**
Phase 2 randomized, flexible crossover, double-blinded, placebo-controlled trial of the farnesyltransferase inhibitor tipifarnib in children and young adults with neurofibromatosis type 1 and progressive plexiform neurofibromas. *Neuro Oncol* 2014;16(5):707-718.
55. **Wolters PL, Martin S, Merker VL, Tonsgard JH, Solomon SE, Baldwin A, et al.**
Patient-reported outcomes of pain and physical functioning in neurofibromatosis clinical trials. *Neurology* 2016;87(7 Suppl 1):S4-S12.

Anhang

Gefähigkeit (6MWT)

Der prozentuale Anteil an Personen mit nicht-fehlenden Werten für den 6MWT lag zu den Erhebungszeitpunkten nach Zyklus 13 unter 70 %, weshalb keine Darstellung der Daten mehr erfolgt.

Tabelle 28: Veränderungen der 6-Minuten-Gehstrecke; FAS ab 5 Jahren mit PN der unteren Extremitäten, Kompression des Rückenmarks und PN der oberen Atemwege der Studie SPRINT, Phase II, Stratum 1 (Datenschnitt 29.06.2018)

Studie SPRINT, Phase II, Stratum 1 Veränderungen der 6-Minuten-Gehstrecke	n	6-Minuten-Gehstrecke (in Metern) N = 34
<i>Baseline</i>	30	
MW (SD)		500,1 (164,1)
Median (min, max)		455 (135; 850)
<i>Veränderung zu Baseline</i>		<i>LSM [95%-KI], p-Wert¹⁾</i>
Vor Zyklus 5	27	2,3 [-30,9; 35,6], 0,89
Vor Zyklus 9	28	17,4 [-8,1;42,9], 0,17
Vor Zyklus 13	25	-2,9 [-42,8; 37,0], 0,88

¹⁾ LSM, KI und p-Wert adjustiert für Studienvisite, Baseline-Wert, Alter, Anzahl an Morbiditäten zu Baseline und Baseline x Studienvisite-Interaktion mittels MMRM.

Abkürzungen: FAS: Full Analysis Set; KI: Konfidenzintervall; LSM: Least Square Mean; MMRM: Mixed Effect Model Repeat Measures; MW: Mittelwert; PN: Plexiformes Neurofibrom; SD: Standardabweichung.