

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: Cannabidiol

Datum der Veröffentlichung: 16. August 2021

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abkürzungsverzeichnis	4
Hintergrund	5
1 Fragestellung	6
2 Auswahl und Methodik der relevanten Studien.....	7
2.1 Studienbasis für die Nutzenbewertung	7
2.2 Aufbau und Design der eingeschlossenen Studie	8
2.3 Endpunkte.....	15
2.3.1 Mortalität.....	16
2.3.2 Morbidität.....	16
2.3.3 Lebensqualität	34
2.3.4 Sicherheit	39
2.3.5 Übersicht der Erhebungszeitpunkte.....	43
2.4 Statistische Methoden	44
2.5 Verzerrungspotential auf Studien- und Endpunktebene	48
3 Ergebnisse der eingeschlossenen Studie.....	50
3.1 Studiencharakteristika und Studienmedikation	50
3.2 Mortalität	55
3.3 Morbidität	56
3.4 Lebensqualität	63
3.5 Sicherheit.....	64
4 Diskussion der Methodik und Ergebnisse.....	67
4.1 Zulassungsstatus und Zulassungspopulation von Cannabidiol	67
4.2 Design und Methodik der Studie	68
4.3 Mortalität	71
4.4 Morbidität	71
4.5 Lebensqualität	75
4.6 Sicherheit.....	76
5 Anforderung an eine qualitätsgesicherte Anwendung	78
6 Zusammenfassung der Nutzenbewertung	79
Referenzen	81

Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis	7
Tabelle 2:	Charakterisierung der Studie GWEP1521 (verblindete Studienperiode).....	8
Tabelle 3:	Relevante Änderungen des Studienprotokolls der Studie GWEP1521 nach Beginn der Rekrutierung (06.04.2016).....	12
Tabelle 4:	Charakterisierung der Intervention der Studie GWEP1521.....	13
Tabelle 5:	Zusammenfassung der Endpunktbewertung der Studie GWEP1521	15
Tabelle 6:	Problemskalen und Syndromskalen der jeweiligen Versionen der Achenbach-Verhaltens-Checklisten	28
Tabelle 7:	Adaptive Level nach v-Skalenwert und Standardwert für Vineland-II	33
Tabelle 8:	Erhebungszeitpunkte der berücksichtigten Endpunkte in Studie GWEP1521.....	43
Tabelle 9:	Verzerrungspotential der Studie GWEP1521.....	48
Tabelle 10:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie GWEP1521.....	48
Tabelle 11:	Allgemeine Angaben der Studie GWEP1521.....	50
Tabelle 12:	Charakterisierung der Studienpopulation GWEP1521	52
Tabelle 13:	Angaben zur Exposition mit der Studienmedikation in Studie GWEP1521.....	54
Tabelle 14:	Begleitmedikationen ≥ 10 % innerhalb eines Behandlungsarms in Studie GWEP1521	55
Tabelle 15:	Ergebnisse zum Endpunkt „Häufigkeit epileptischer Anfälle“ in Studie GWEP1521; Behandlungszeitraum; ITT-Population.....	56
Tabelle 16:	Ergebnisse zum Endpunkt „Häufigkeit TSC-assoziierten Anfälle“ in Studie GWEP1521; Behandlungszeitraum (Sensitivitätsanalyse); ITT-Population	57
Tabelle 17:	Ergebnisse für den Endpunkt „Häufigkeit TSC-assoziierte Anfälle“ in Studie GWEP1521; Responderanalysen (Reduktion ≥ 25 , ≥ 50 , ≥ 75 , 100 %; Anstieg > 0 %); Behandlungszeitraum; ITT-Population	58
Tabelle 18:	Ergebnisse für den Endpunkt „Häufigkeit Anfälle gesamt“ in Studie GWEP1521; Responderanalysen (Reduktion ≥ 25 , ≥ 50 , ≥ 75 , 100 %; Anstieg > 0 %); Behandlungszeitraum; ITT-Population.....	58
Tabelle 19:	Ergebnisse für den Endpunkt „Status epilepticus“ in Studie GWEP1521; Behandlungszeitraum; ITT-Population.....	59
Tabelle 20:	Ergebnisse für den Endpunkt „Globaler Eindruck der Veränderung“ in Studie GWEP1521; Responderanalysen; Behandlungszeitraum; ITT-Population.....	60
Tabelle 21:	Ergebnisse für den Endpunkt „Globaler Eindruck der Veränderung“ in Studie GWEP1521; ordinal logistische Regression; Behandlungszeitraum; ITT-Population	61
Tabelle 22:	Ergebnisse zum Endpunkt „Achenbach-Verhaltens-Checkliste“ in Studie GWEP1521; Behandlungszeitraum; ITT-Population.....	62
Tabelle 23:	Ergebnisse zum Endpunkt „Unerwünschte Ereignisse“ während der verblindeten Behandlungsphase in Studie GWEP1521; Sicherheitspopulation.....	64
Tabelle 24:	Ergebnisse zum Endpunkt „UE mit Inzidenz ≥ 10 % nach MedDRA-System- organklassen und Preferred Terms“ in Studie GWEP1521; Sicherheitspopulation ...	65
Tabelle 25:	Zusammenfassende Darstellung der Ergebnisse der Studie GWEP1521.....	79

Abkürzungsverzeichnis

ABCL	Adult Behavior Checklist
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse
CBCL	Child Behavior Checklist
CGIC	Caregiver Global Impression of Change
CGICSD	Caregiver Global Impression of Change in Seizure Duration
CMH	Cochran-Mantel-Haenszel
C-SSRS	Columbia-Suicide Severity Rating Scale
EPAR	European Public Assessment Report
ESC	Epilepsy Study Consortium
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
IVRS	Interactive Voice Response System
KI	Konfidenzintervall
MedDRA	Medical Dictionary for Regulatory Activities
OLE	Offene Extensionsstudie
PGIC	Patient Global Impression of Change
pU	pharmazeutischer Unternehmer
QOLCE	Quality of Life in Childhood Epilepsy
QOLIE-31-P	Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SCQ	Social Communication Questionnaire
SGB	Sozialgesetzbuch
SGIC	Subject Global Impression of Change
SGICSD	Subject Global Impression of Change in Seizure Duration
SUE	Schwerwiegendes unerwünschtes Ereignis
TEAE	Treatment Emergent Adverse Event
TSC	Tuberöse Sklerose (Tuberous Sclerosis Complex)
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA
Vineland-II	Vineland Adaptive Behavior Scales, Second Edition

Hintergrund

Cannabidiol 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 Cannabidiol 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 Cannabidiol in seiner Sitzung am 10. August 2021 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 14. Mai 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 16. August 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

Cannabidiol (Epidyolex®) 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 Patientinnen und Patienten in folgender Indikation eingeschlossen:

Patienten ab 2 Jahren für die adjuvante Behandlung von Krampfanfällen im Zusammenhang mit Tuberöser Sklerose (TSC).

Die empfohlene Anfangsdosis in der Indikation TSC liegt bei 2,5 mg/kg zweimal täglich (5 mg/kg/Tag) für eine Woche. Danach soll die Dosis auf eine therapeutische Erhaltungs-Dosierung von 5 mg/kg zweimal täglich (10 mg/kg/Tag) erhöht werden. Basierend auf dem individuellen Ansprechen und der Verträglichkeit kann die Dosierung bis zu einer maximalen Dosierung von 12,5 mg/kg zweimal täglich (25 mg/kg/Tag) gesteigert werden.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als rele- vant für die Ableitung des Zusatz- nutzens erachtet	Studie relevant für die Nutzen- bewertung	Ausschlussgrund
Studien zum Wirkstoff				
GWEP1521 ¹⁾	Ja	Ja	Ja	-
GWEP1521 Open-Label Extension	(Ja) ²⁾³⁾	Nein	Nein	• Laufende, offene, einarmige Extensionsstudie. • Maximal mögliche Dosisescalation von 50 mg/kg/Tag Cannabidiol außerhalb der FI-konformen Anwendung.

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ In die OLE konnten alle Testpersonen wechseln, die die doppelt-blinde Studienphase abgeschlossen haben. Dabei wurden alle Testpersonen zunächst auf eine Dosis von 25 mg/kg/Tag titriert, die anschließend angepasst werden konnte, um die optimale Dosis für jede Testperson zu finden mit einer maximalen Dosisescalation auf 50 mg/kg/Tag. Das primäre Zielkriterium der OLE war die Erfassung der Sicherheit. Von 224 in Studie GWEP1521 randomisierten Testpersonen wechselten 199 (88,8 %) in die OLE. Ein Studienbericht oder Studiendaten für die OLE wurden nicht eingereicht. Im EPAR ist ein Interim-Datenschnitt vom 26.02.2019 aufgeführt [15]. Gemäß diesen Daten wurde die Mehrheit der Testpersonen (n = 156; 78 %) mit einer modalen Dosis (Dosis die jeweils am häufigsten je Testperson verabreicht wurde) von ≤ 25 mg/kg/Tag Cannabidiol behandelt. Mangels Datengrundlage und aufgrund eines überwiegenden Dosispektrums außerhalb der FI-konformen Anwendung wird die OLE nicht für die vorliegende Nutzenbewertung berücksichtigt. Das Studienende ist für Februar 2022 vorgesehen.

³⁾ Für Studie GWEP1521 und OLE liegen ein gemeinsames Studienprotokoll sowie ein SAP vor. Ergebnisse für die OLE wurden nicht eingereicht.

Abkürzungen: EPAR: European Public Assessment Report; FI: Fachinformation; OLE: Offene Extensionsstudie; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan.

Für die vorliegende Nutzenbewertung wird der Studienarm mit einer Dosierung von 50 mg/kg/Tag Cannabidiol nicht berücksichtigt, da diese Dosis außerhalb der fachinformationskonformen Anwendung liegt, die maximal 25 mg/kg/Tag Cannabidiol vorsieht.

Zur Nutzenbewertung für Cannabidiol wurden folgende Studien und Daten herangezogen:

- Herstellerdossier zu Cannabidiol [21].
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [15].
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie GWEP1521 [18,19,20].
- Publikation zum Wirkstoff: Thiele et al. Add-On Cannabidiol Treatment for Drug-Resistant Seizures in Tuberous Sclerosis Complex: A Placebo-Controlled Randomized Clinical Trial [37].

2.2 Aufbau und Design der eingeschlossenen Studie

Die Unterlagen zur Nutzenbewertung für Cannabidiol basieren auf der Zulassungsstudie GWEP1521 (GWPCARE6). Studie und Intervention werden in den Tabellen 2–4 charakterisiert. Dabei werden ausschließlich die Angaben der verblindeten, abgeschlossenen Studienperiode dargestellt; Angaben zur offenen Extensionsstudie (OLE; Studienende geplant für Februar 2022) wurden nicht extrahiert.

Tabelle 2: Charakterisierung der Studie GWEP1521 (verblindete Studienperiode)

Charakteristikum	Beschreibung
Design	Die Studie GWEP1521 ist eine doppelblinde, randomisierte, placebokontrollierte, multi-zentrische Studie zur Untersuchung der Wirksamkeit und Sicherheit von Cannabidiol bei Kindern und Erwachsenen im Alter von 1 bis 65 Jahren mit TSC, deren Krampfanfälle nicht ausreichen kontrolliert sind. Die Randomisierung der Studienpopulation erfolgte im Verhältnis 2:2:1:1 auf die Studienarme Cannabidiol 25 mg/kg/Tag, Cannabidiol 50 mg/kg/Tag und Placebo¹⁾ (25 mg/kg/Tag und 50 mg/kg/Tag) stratifiziert nach Altersgruppen (1–6, 7–11, 12–17 und ≥ 18 Jahre). Die Behandlungsdauer betrug 16 Wochen. Das primäre Zielkriterium der Studie war die Reduktion der Häufigkeit von Krampfanfällen. Die Studie lässt sich in folgende Phasen unterteilen: • Baseline-Periode (4 Wochen) • Behandlungsphase (16 Wochen) mit ◦ Titrationsphase (4 Wochen) ◦ Erhaltungsphase (12 Wochen) • 10-tägiges Ausschleichen der Studienmedikation mit 4-wöchiger Sicherheits-Follow-up-Periode <u>oder</u> Übergang in eine offene Extensionsstudie (OLE). In die OLE konnten alle Testpersonen wechseln, die die doppelt blinde Studienphase abgeschlossen haben. Dabei wurden alle Testpersonen zunächst auf eine Dosis von 25 mg/kg/Tag titriert, die anschließend angepasst werden konnte, um die optimale Dosis für jede Testperson zu finden mit einer maximalen Dosis Eskalation auf 50 mg/kg/Tag.
Population	Wesentliche Einschlusskriterien • Männliche und weibliche Testpersonen im Alter von 1 bis 65 Jahren. • Die Testperson und die betreuende Person waren bereit und in der Lage (nach Meinung des ärztlichen Prüfpersonals), alle Anforderungen der Studie (einschließlich eines genauen Tagebuchs und IVRS-Abschlusses) zu erfüllen. • Gut dokumentierte klinische Vorgeschichte der Epilepsie, die durch derzeitige Antiepileptika nicht vollständig kontrolliert wird. • Klinische Diagnose der TSC nach den von der Internationalen Konsensuskonferenz zu TSC 2012 vereinbarten Kriterien. • Einnahme von einem oder mehreren Antiepileptika in einer Dosis, die vor dem Screening mindestens 4 Wochen lang stabil war. • Alle Medikamente oder Maßnahmen zur Behandlung von Epilepsie (einschließlich ketogener Diät und jeglicher Neurostimulationsgeräte für Epilepsie) müssen vor dem Screening einen Monat lang stabil gewesen sein, und die Testperson muss bereit sein, während der gesamten Studie ein gleichbleibendes Schema beizubehalten. • Testpersonen waren bereit, alle Faktoren, von denen erwartet wurde, dass sie die Anfälle beeinflussen, stabil zu halten (wie z. B. die Menge des Alkoholkonsums und des Rauchens). • Am Ende des Baseline-Zeitraums mussten die Testpersonen zusätzlich folgende Kriterien erfüllen: ◦ Mindestens 8 Anfälle in den ersten 28 Tagen des Baseline-Zeitraums, wobei mindestens ein Anfall pro Woche in 3 der 4 Wochen auftrat (Anfälle beinhalteten: fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der

Charakteris- tikum	Beschreibung
	Wahrnehmung; fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung; fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen und generalisierten Krampfanfällen [tonisch-klonisch, tonisch, klonisch oder atonisch] entwickelten, die zählbar waren). - o Mindestens 90 % der Anrufe beim IVRS (1 Anruf pro Tag vorgesehen) während der ersten 28 Tage des Baseline-Zeitraums (mindestens 25 erledigte Anrufe). Wesentliche Ausschlusskriterien • Pseudo-Anfälle in der Vergangenheit. • Weitere klinisch signifikante instabile Erkrankungen neben der Epilepsie. • Testperson hatte andere Erkrankungen als Epilepsie in den 4 Wochen vor dem Screening oder der Randomisierung, die nach Einschätzung des ärztlichen Prüfpersonals die Anfallshäufigkeit beeinflussen könnten. • Vollnarkose in den 4 Wochen vor dem Screening oder der Randomisierung. • Operation aufgrund der Epilepsie in den 6 Monaten vor dem Screening. • Die Testperson wurde während der verblindeten Phase der Studie für eine Epilepsie-Operation oder einen Eingriff unter Vollnarkose in Betracht gezogen. • Einnahme von Felbamat weniger als ein Jahr vor dem Screening. • Einnahme eines oralen mTOR-Inhibitor. • Klinisch signifikante, abnorme Laborwerte nach Einschätzung des Prüfpersonals. • Bekannte oder vermutete Überempfindlichkeit gegen Cannabinoide oder einen der Hilfsstoffe des Präparats, wie z. B. Sesamöl. • Jegliche Vorgeschichte suizidalen Verhaltens oder Suizidgedanken vom Typ 4 oder 5 auf der C-SSRS im letzten Monat oder beim Screening. • Testperson hat in den drei Monaten vor dem Screening Cannabis (gelegentlicher Drogenkonsum oder medizinisches Cannabis) oder Medikamente auf Cannabinoid-basis konsumiert oder in der Vergangenheit konsumiert und war nicht bereit, für die Dauer der Studie darauf zu verzichten. • Tumorwachstum, das nach Ansicht des ärztlichen Prüfpersonals den primären Endpunkt beeinträchtigen könnte. • Klinisch signifikante Anomalien im EKG nach Ansicht des ärztlichen Prüfpersonals, die beim Screening oder bei der Randomisierung gemessen wurden, oder andere kardiovaskuläre Erkrankungen, die die Fähigkeit zum Ablesen der EKG beeinträchtigen könnten. • Signifikante Beeinträchtigung der Leberfunktion bei der Screening-Visite (Tag -35) oder bei der Randomisierungsvisite (Tag 1), definiert als eine der folgenden: - o Serum-ALT oder -AST > 5 × ULN. - o Gesamtbilirubin* ≥ 2 × ULN oder INR > 1,5 (*TBL ≥ 2 × ULN-Ausschluss galt nicht für Personen, bei denen die Gilbert-Krankheit diagnostiziert wurde). - o Serum-ALT oder -AST ≥ 3 × ULN mit Vorhandensein von Müdigkeit, Übelkeit, Erbrechen, Schmerzen oder Empfindlichkeit im rechten oberen Quadranten, Fieber, Hautausschlag und/oder Eosinophilie (> 5 %). Dieses Kriterium konnte erst nach Vorliegen der Laborergebnisse bestätigt werden. • Erhalt eines Prüfpräparats weniger als 12 Wochen vor der Screening-Visite. • Testpersonen mit einer anderen signifikanten Krankheit oder Störung, die nach Ansicht des Prüfpersonals entweder die Testperson aufgrund der Teilnahme an der klinischen Studie gefährdet, das Ergebnis der klinischen Studie beeinflusst oder die Fähigkeit der Testperson zur Teilnahme an der klinischen Studie beeinträchtigen könnte. • Alle nach einer körperlichen Untersuchung der Patientin / des Patienten festgestellten Anomalien, die nach Einschätzung des ärztlichen Prüfpersonals die Sicherheit der Patientin / des Patienten bei einer Teilnahme an der Studie gefährden würden. • Bekannte oder vermutete Vorgeschichte von Alkohol- oder Drogenmissbrauch.

Charakteristikum	Beschreibung
Intervention und Zahl der Patienten	Screening: N = 255 Randomisierung: N = 224 • N = 75 Intervention: 25 mg/kg/Tag Cannabidiol* • N = 73 Intervention: 50 mg/kg/Tag Cannabidiol • N = 76 Kontrolle: Placebo¹⁾ * Die maximale FI-konforme Dosierung beträgt 25 mg/kg/Tag; somit bewertungsrelevanter Studienarm für die vorliegende Nutzenbewertung.
Ort und Zeitraum der Durchführung	Studienzentren 46 internationale Studienzentren, davon 25 in den USA, 6 in Polen, 5 in Spanien, 4 in Australien, 4 in Großbritannien und 2 in den Niederlanden. Studienzeitraum Datum Studienbeginn: 06.04.2016 Datum Studienabschluss: 15.02.2019 Entblindung der Datenbank: 27.04.2019 Datum Studienbericht: 23.12.2019
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	Primärer Endpunkt Veränderung der Häufigkeit TSC-assoziiierter Anfälle²⁾ während der Behandlung im Vergleich zu Baseline. Wichtige sekundäre Endpunkte • Anzahl der Responder, definiert als Reduktion der TSC-assoziierten Anfallshäufigkeit im Vergleich zu Baseline um $\geq 50\%$. • Globaler Patienten-/Betreuer-Eindruck der Veränderung (CGIC / SGIC). • Veränderungen der Gesamtzahl der Anfälle. Sekundäre Endpunkte • Anzahl der Responder, definiert als Reduktion der TSC-assoziierten Anfallshäufigkeit im Vergleich zu Baseline um $\geq 25\%$, $\geq 50\%$, $\geq 75\%$ oder 100% (gesamt und alle 4 Wochen). • Anzahl Patientinnen und Patienten mit einer Verschlechterung der TSC-assoziierten Anfallshäufigkeit im Vergleich zu Baseline um $> 25\%$, keine Veränderung ≤ -25 bis $\geq 0\%$, keine Veränderung > 0 bis $< +25\%$, Verbesserung um ≥ 25 bis $< 50\%$, Verbesserung um ≥ 50 bis $< 75\%$ oder Verbesserung um $\geq 75\%$³⁾. • Veränderung in der Anzahl der Tage ohne TSC-assoziierte Anfälle. • Veränderung in der Anzahl anderer Anfälle (Absencen, myoklonische-, fokale sensorische- und infantile/epileptische Spasmen). • Globaler arztbewerteter Eindruck der Veränderung (PGIC). • Wachstum und Entwicklung bei Personen unter 18 Jahren (IGF1-Level und Tanner-Staging). • Lebensqualität bestimmt mittels QOLCE (für Personen von 2 bis 18 Jahren) oder QOLIE-31-P (für Personen ab 19 Jahren). • Anzahl der Hospitalisierungen aufgrund von Epilepsie. • Sicherheit (einschließlich C-SSRS, Missbrauchspotential). Explorative Endpunkte • Zusammengesetzter fokaler Anfallswert (Häufigkeit x Schwere). • Veränderung in der Häufigkeit von Anfalls-Subtypen. • Veränderung in der Einnahme von Notfallmedikation. • Veränderung in der Anzahl der Status-epilepticus-Episoden (konvulsiv und nicht-konvulsiv). • Veränderung der Anfallsdauer (bestimmt mittels CGICSD / SGICSD). • Bewertung des adaptiven Verhaltens (mittels Vineland-II). • Wechsler-Test (Vorschule, Primärstufe, Kinder, Erwachsene). • Veränderung in der Achenbach-Verhaltens-Checkliste.

Charakteristikum	Beschreibung
	• Kennzeichen von Autismus: Wert des Fragebogens zur sozialen Kommunikation (SCQ). • Plasmakonzentrationen von begleitend eingenommenen Antiepileptika vor und nach der Behandlung mit Cannabidiol.
Subgruppenanalysen	Folgende Subgruppenanalysen waren gemäß SAP für den primären Endpunkt sowie den wichtigen sekundären Endpunkt „≥ 50 % Reduktion der TSC-assoziierten Anfallsfrequenz“ geplant: • Altersgruppe (1–6, 7–11, 12–17, ≥ 18 Jahre) • Geschlecht (männlich, weiblich) • Region (USA, Rest der Welt) • Einnahme von Clobazam (ja, nein) • Einnahme von Valproinsäure (ja, nein) • Einnahme von Levetiracetam (ja, nein) • Einnahme von Vigabartin (ja, nein) • Durchschnittliche Häufigkeit von Sturzanfällen pro 28 Tage zu Baseline (≤ Beobachtungstertil 1, > Beobachtungstertil 1 und ≤ Beobachtungstertil 2, > Beobachtungstertil 2) • Anzahl begleitender antiepileptischer Medikamente (< 3, ≥ 3) • Anzahl früherer antiepileptischer Medikamente (< 5, ≥ 5) • Anzahl begleitender und früherer antiepileptischer Medikamente (< 8, ≥ 8) Die Effekte weiterer antiepileptischer Medikamente werden ebenfalls getestet, falls dieses antiepileptische Medikament von > 25 % der Testpersonen eingenommen wird.

¹⁾ Randomisierung erfolgte in die Studienarme: Placebo 25 mg/kg/Tag (n = 38), Placebo 50 mg/kg/Tag (n = 38). Die Placebo-Gruppen wurden für die Wirksamkeits- und Sicherheitsanalysen gepoolt.

²⁾ TSC-assoziierte Anfälle beinhalten: Fokalmotorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der Wahrnehmung; fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung; fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen und generalisierten Krampfanfällen [tonisch-klonisch, tonisch, klonisch oder atonisch] entwickelten, die zählbar waren).

³⁾ Im Studienprotokoll ist die Responderklassifikation "keine Verbesserung" in einer Kategorie definiert als -25% bis 25% Reduktion der TSC-assoziierten Anfälle. In SAP und Studienbericht ist "keine Verbesserung" dagegen weiter unterteilt in 2 Gruppen: ≤ -25 bis ≥ 0% und > 0 bis < +25 %.

Abkürzungen: ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CGIC: Caregiver Global Impression of Change; CGICSD: Caregiver Global Impression of Change in Seizure Duration; C-SSRS: Columbia-Suicide Severity Rating Scale; FI: Fachinformation; IGF1: Insulin-like Growth Factor1; INR: International Normalized Ratio; IVRS: Interactive Voice Response System; PGIC: Patient Global Impression of Change; QOLCE: Quality of Life in Childhood Epilepsy; QOLIE-31-P: Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version; SAP: Statistischer Analyseplan; SCQ: Social Communication Questionnaire; SGIC: Subject Global Impression of Change; SGICSD: Subject Global Impression of Change in Seizure Duration; TBL: Total Bilirubin Level; TSC: Tuberöse Sklerose; ULN: Upper Limit of Normal; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

Im Studienprotokoll der Studie GWEP1521 werden 7 Änderungen beschrieben, die die verblindete Studienphase betreffen. 6 dieser Protokolländerungen traten nach Einschluss der ersten Person in die Studie (06.04.2016) in Kraft. Wie viele Patientinnen und Patienten zu diesem Zeitpunkt bereits in die Studie eingeschlossen waren ist nicht bekannt. Die relevanten Änderungen nach Studienbeginn finden sich in der nachfolgenden Tabelle 3.

Tabelle 3: Relevante Änderungen des Studienprotokolls der Studie GWEP1521 nach Beginn der Rekrutierung (06.04.2016)

Protokollversion	Wesentliche Protokolländerungen
Amendment 2 (25.08.2016)	• Reduktion der Häufigkeit der Erhebung folgender Instrumente: QOLCE / QOLIE-31P, PGIC, CGIC / SGIC, Wechsler-Test, CBCL / ABCL, SCQ und Vineland-II. • Einfügen einer detaillierten Beschreibung für den PGIC. • Änderung der statistischen Methoden zur Analyse der primären und sekundären Endpunkte von einer Kovarianzanalyse zu einem Wilcoxon-Rangsummentest inklusive der Annahme, dass für diesen Test mehr Testpersonen erforderlich sind. Die Zielstichprobengröße wurde auf 210 erhöht. • Der Aufzählungspunkt "erfüllte die Eignungskriterien nicht" wurde von "muss" zu "kann aus der Studie entfernt werden" verändert. • Die primären und sekundären Endpunkte wurden präziser definiert (z. B. Reduktion der TSC-assoziierten Anfälle um $\geq 50\%$ als wichtiger sekundärer Endpunkt). • Änderungen infolge zusätzlicher Anforderungen der MHRA: - o Geänderte Formulierungen in den Ausschlusskriterien, die die Überwachung von Leberenzymen betreffen. - o Klarstellung, dass bei Personen mit Morbus Gilbert ein erhöhter Bilirubinwert nur dann als Ausschlussgrund gilt, wenn er mit erhöhten ALT- oder AST-Werten einhergeht. - o Im Vereinigten Königreich darf die Aufnahme von Personen im Alter zwischen 12 und 23 Monaten erst dann beginnen, wenn 15 Patientinnen und Patienten im Alter von über 23 Monaten mindestens 4 Wochen lang behandelt wurden und keine neuen Sicherheitsprobleme auftraten. • Klarstellung, dass der Konsum von Cannabis, ausgenommen das Prüfpräparat, in der Studie untersagt ist. • Klarstellung, dass die Testperson und/oder die betreuende Person ein Training zur Identifikation von Anfällen erhält.
Amendment 3 (05.12.2016)	Anpassung der Kriterien für ein Ausscheiden aus der Studie in Bezug auf die Überwachung von Leberenzymen („und“- statt „oder“-Kriterium), um das Protokoll in Übereinstimmung mit den FDA-Leitlinien zu leberenzymbezogenen Ausschlusskriterien zu bringen.
Amendment 4 (27.06.2017)	• Sekundäre Endpunkte wurden unterteilt in: Wichtige sekundäre Endpunkte, sonstige sekundäre Endpunkte und explorative Endpunkte. • Die statistische Analyse wurde in Folge der Neukategorisierung der sekundären Endpunkte geändert. Die hierarchische Teststrategie der wichtigen sekundären Endpunkte wurde klar definiert. • Die Verabreichung von Cannabidiol über eine gastrostomische/nasogastrische Ernährungssonde wurde nach Rücksprache mit dem medizinischen Berater als Option hinzugefügt.
Amendment 5 (07.08.2018)	• Hinzufügen des Einschlusskriteriums "Einnahme von einem oder mehreren Antiepileptika in einer Dosis, die vor dem Screening mindestens 4 Wochen lang stabil war", um sicherzustellen, dass in Frage kommende Personen ein oder mehrere Antiepileptika in einer Dosis einnehmen müssen. • Spezifizierung des Einschlusskriteriums „Gut dokumentierte klinische Vorgeschichte der Epilepsie“ um den Passus „die durch derzeitige Antiepileptika nicht vollständig kontrolliert wird“. • Klarstellung, dass Dosisreduzierungen der Begleitmedikation mit Antiepileptika aus klinischen Gründen nach Diskussion mit dem medizinischen Berater des Sponsors zulässig sind. • Spezifizierung, dass nur Personen, die die verblindete Phase der Studie erfolgreich abgeschlossen haben, an der OLE teilnehmen können.

Protokollversion	Wesentliche Protokolländerungen
	- Klarstellung, dass bei Personen mit tiefgreifender kognitiver Beeinträchtigung im Alter von 6 Jahren oder älter, bei denen der Abschluss der C-SSRS nicht angemessen ist, die Suizidalität durch das klinische Urteil des ärztlichen Prüfpersonals nach Befragung der Patientin / des Patienten beurteilt wird. - Klarstellung, dass die verblindete Phase dieser Studie vor Abschluss der OLE geschlossen und entblindet wird und dass der SAP für die verblindete Studienperiode vor der Entblindung finalisiert wird.
Amendment 6 (06.09.2018)	- Aktualisierung der statistischen Analyse des primären Endpunkts von einem Wilcoxon-Rangsummentest auf eine negativ-binomiale Regressionsanalyse. Diese Änderung wurde auch für die Analysen der weiteren Anfallsendpunkte vollzogen. - Entfernen der Bezeichnung „prozentuale Änderung“ aus der Formulierung des primären Endpunkts und für die prozentuale Änderung bei anderen Anfallsarten unter sekundären Endpunkten im gesamten Protokoll. - Hinzufügen der ersetzten primäre Analyse mit dem Wilcoxon-Rangsummentest als weitere Sensitivitätsanalyse. - Gegebenenfalls wurden andere Sensitivitätsanalysen von Wilcoxon-Rangsummentests auf negativ-binomiale Regressionsanalysen geändert.
Amendment 7 (23.04.2019)	Veränderung der Reihenfolge der hierarchischen Teststrategie, sodass alle Endpunkte mit TSC-assoziierten Anfällen zuerst getestet werden.

Abkürzungen: ABCL: Adult Behavior Checklist; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CGIC: Caregiver Global Impression of Change; CBCL: Child Behavior Checklist; C-SSRS: Columbia-Suicide Severity Rating Scale; FDA: U. S. Food and Drug Administration; MHRA: Medicines and Healthcare products Regulatory Agency; OLE: Offene Extensionsstudie; PGIC: Patient Global Impression of Change; QOLCE: Quality of Life in Childhood Epilepsy; QOLIE-31-P: Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version; SAP: Statistischer Analyseplan; SCQ: Social Communication Questionnaire; SGIC: Subject Global Impression of Change; TSC: Tuberöse Sklerose; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

Tabelle 4: Charakterisierung der Intervention der Studie GWEP1521

Intervention	Kontrolle
Cannabidiol (Epidyolex®)¹⁾ <u>Dosierung und Art der Anwendung³⁾</u> 25 mg/kg/Tag bzw. 50 mg/kg/Tag (Erhaltungsperiode). - Zweimal täglich (morgens und abends) als orale Lösung (100 mg/ml) für 16 Wochen. - Titrationperiode maximal 4 Wochen, mit einer Anfangsdosis von zweimal täglich 2,5 mg/kg (gesamt 5 mg/kg/Tag); schrittweise Erhöhung um 5 mg/kg/Tag alle 2 Tage auf insgesamt 25 mg/kg/Tag. Die Titration von 25 mg/kg/Tag auf 50 mg/kg/Tag erfolgt in kleineren Schritten von 2,5 mg/kg/Tag alle 2 Tage; danach Erhaltungsperiode bis Studienvisite 9 (Tag 113 ± 3).	Placebo²⁾ <u>Dosierung und Art der Anwendung</u> Wie Intervention, ohne den aktiven Wirkstoff.
Personen, die im Anschluss an die Erhaltungsperiode nicht unmittelbar in die OLE eintraten, senkten ihre Dosis allmählich um 10 % pro Tag ab. Dieses Vorgehen war auch für Personen vorgesehen, die die Studie frühzeitig beendeten, außer eine weitere Einnahme war aufgrund eines UE nicht möglich. Personen, die im Anschluss an die Erhaltungsperiode direkt in die OLE eintraten, durchliefen zunächst eine zweiwöchige verblindete Übergangsphase. Dabei wurde das Prüfpräparat schrittweise auf 25 mg/kg/Tag auftitriert, während das Prüfpräparat der verblindeten Studienphase gleichzeitig schrittweise abgesetzt wurde, sodass alle Testpersonen am Ende der Übergangsphase 25 mg/kg/Tag erhielten.	

Intervention	Kontrolle
Begleitmedikation • Cannabidiol wird zusammen mit anderen AED verabreicht. • Die Dosierungen der begleitenden AED mussten vor dem Screening mindestens 4 Wochen lang stabil gehalten werden und während des gesamten verblindeten Studienzeitraumes stabil bleiben. • Sofern sich die Plasmakonzentrationen der begleitenden AED infolge der Applikation der Intervention verändern, oder falls Nebenwirkungen aufgrund der veränderten Plasmakonzentrationen auftreten, kann die Dosierung der begleitenden AED in Abhängigkeit der medizinischen Notwendigkeit und nach Rücksprache mit dem medizinischen Prüfpersonal des pU modifiziert werden. • Während der randomisierten Studienphase ist die Initiierung neuer AED untersagt. • Nicht-pharmakologische Therapien (z. B. ketogene Diät, VNS) mussten ebenfalls mindestens 4 Wochen vor dem Screening und während des gesamten Studienzeitraums stabil sein. Weitere erlaubte Begleitmedikation Verwendung von Notfallmedikamenten bei Bedarf. Nicht erlaubte Begleitmedikation Die Einnahme folgender Medikationen ist für die Dauer der Studie ab Abgabe der Einwilligungserklärung untersagt: • Jegliche neu initiierten Medikamente oder Maßnahmen gegen Epilepsie (einschließlich ketogener Diät und VNS) oder Änderungen der Dosierung. • Gelegentlicher Cannabis-Konsum, medizinisches Cannabis oder synthetische Cannabinoid-basierte Medikamente (einschließlich Sativex®). • Einnahme eines anderen Prüfpräparats in einer klinischen Studie. • Dauerhafte Einnahme systemischer Steroide (außer inhalierte Medikation zur Asthmabehandlung) oder anderer täglicher Medikationen, von denen bekannt ist, dass sie Epilepsie verschlimmern. Eine Ausnahme gilt für prophylaktische Medikamente, z. B. für idiopathisches nephrotisches Syndrom oder Asthma. Besonderes Augenmerk soll auf Arzneimittel oder Metaboliten gelegt werden, die Substrate von Cytochrom P450 2C19 sind, wie N-desmethyloclobazam oder solche die allein oder vorwiegend von UDP-glukuronosyltransferase1A9 und 2B7 metabolisiert werden.	

¹⁾ Die Prüfsubstanz in Form einer ölhaltigen Lösung setzt sich zusammen aus 100 mg/ml Cannabidiol, 79 mg/ml Ethanol, 0,5 mg/ml Sucralose, 0,2 mg/ml Erdbeeraroma, Sesamöl zum Auffüllen der Lösung bis 1 ml.

²⁾ Die Kontrollsubstanz (Placebo) setzt sich zusammen aus 79 mg/ml Ethanol, 0,5 mg/ml Sucralose, 0,2 mg/ml Erdbeeraroma, Sesamöl zum Auffüllen der Lösung bis 1 ml.

³⁾ In der Fachinformation zu Epidyolex wird darauf verwiesen, dass die Einnahme konsequent entweder mit oder ohne Nahrungsaufnahme erfolgen sollte, da Nahrung zu einer Erhöhung des Cannabidiolspiegels führen kann. In Studie GWEP1521 waren keine Einschränkungen der Nahrungsaufnahme im Zusammenhang mit der Prüfsubstanz vorgegeben [22].

Abkürzungen: AED: Antiepileptisches Medikament; OLE: Offene Extensionsstudie; pU: pharmazeutischer Unternehmer; UE: Unerwünschtes Ereignis; VNS: Vagusnerv-Stimulation.

2.3 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie GWEP1521

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle ¹⁾	Mortalität	ja	ja
Epileptische Anfälle • TSC-assoziierte Anfälle²⁾ • Anfälle gesamt • Weitere Anfälle • Zusammengesetzter fokaler Anfallswert • Subtypen von Anfällen	Morbidität	ja	ja
		ja	ja
		ja	ja
		ja	nein
		ja	nein
Status epilepticus		ja	ja
Globaler Patienten-/Betreuer-Eindruck der Veränderung (mittels CGIC / SGIC)		ja	ja
Globaler arztbewerteter Eindruck der Veränderung (mittels PGIC)		ja	nein
Veränderung der Anfallsdauer (mittels CGICSD / SGICSD)		ja	nein
Epilepsiebedingte Hospitalisierung		ja	ergänzend
Verwendung von Notfallmedikation		ja	nein
Soziale Kommunikation (mittels SCQ) ⁴⁾		ja	nein
Achenbach-Verhaltens-Checkliste ²⁾		ja	ja
Bewertung des adaptiven Verhaltens (mittels Vineland-II) ²⁾		ja	ja
Lebensqualität bei kindlicher Epilepsie (mittels QOLCE)	Lebensqualität	ja	ja
Lebensqualität bei Epilepsie (mittels QOLIE-31-P)		ja	nein
Unerwünschte Ereignisse	Sicherheit	ja	ja
Suizidalität (mittels C-SSRS)		nein	ja

¹⁾ Todesfälle wurden im Rahmen der Sicherheitserhebung erfasst.

²⁾ Veränderung in der Häufigkeit von TSC-assoziierten Anfällen im Vergleich zur Baseline-Periode primärer Endpunkt der Studie.

³⁾ Vom pU der Kategorie „Lebensqualität“ zugeordnet.

Abkürzungen: CGIC: Caregiver Global Impression of Change; CGICSD: Caregiver Global Impression of Change in Seizure Duration; C-SSRS: Columbia-Suicide Severity Rating Scale; PGIC: Patient Global Impression of Change; pU: pharmazeutischer Unternehmer; QOLCE: Quality of Life in Childhood Epilepsy; QOLIE-31-P: Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version; SCQ: Social Communication Questionnaire; SGIC: Subject Global Impression of Change; SGICSD: Subject Global Impression of Change in Seizure Duration; TSC: Tuberöse Sklerose; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

2.3.1 Mortalität

Todesfälle

Der Endpunkt „Todesfälle“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Todesfälle wurde in den Studien nicht als separater Endpunkt erfasst. Todesfälle wurden im Rahmen der Sicherheitsbewertung anhand der unerwünschten Ereignisse mit Todesfolge erhoben (siehe 2.3.4).

- Erhebungszeitpunkte: Im Rahmen der Sicherheitsbewertung über den gesamten Studienzeitraum bis zur Sicherheits-Follow-up-Visite.
- Auswertung: Auf Basis der Sicherheitspopulation.

Bewertung

Die Operationalisierung ist nachvollziehbar.

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“.

Validität

Todesfälle waren kein eigenständiger Endpunkt in der Studie GWEP1521. Todesfälle waren jedoch Bestandteil der Sicherheitserhebung. Die Erhebung des Endpunkts wird als valide erachtet. Für die Nutzenbewertung werden die Ergebnisse der Behandlungsperiode herangezogen.

2.3.2 Morbidität

Epileptische Anfälle

Der Endpunkt „Epileptische Anfälle“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Folgende epileptische Anfallstypen wurden in der Studie GWEP1521 unterschieden und erfasst:

- Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der Wahrnehmung
- Fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung
- Fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen entwickelten
- Generalisierte Krampfanfälle:
 - o tonisch-klonische Anfälle
 - o tonische Anfälle
 - o klonische Anfälle
 - o atonische Anfälle
- Weitere Anfälle
 - o Absencen
 - o myoklonische Anfälle

- o fokal-sensorische Anfälle
- o infantile/epileptische Spasmen
- Status-epilepticus-Episoden

Patientinnen und Patienten oder die jeweilige betreuende Person sollten Anfallsart und -anzahl sowie die Schwere der fokalen Anfälle täglich von Baseline (Studienvisite 2) bis Behandlungsende mittels Interactive Voice Response System (IVRS) erheben. Aus Konsistenzgründen sollte die Einschätzung immer durch dieselbe Person erfolgen.

Die betreuenden Personen bzw. die Patientinnen und Patienten erhielten zu Studienvisite 2 (Tag -28) ein Training zur Identifikation der bei der Testperson zu erwartenden Anfallstypen (individuelles Anfallsprofil) und zum richtigen Umgang mit dem IVRS. Falls eine neue Anfallsart während der Studie auftrat, sollte das Studienzentrum die betreuende Person erneut trainieren. Zudem erhielten auch alle Testzentren ein Training zum Umgang mit dem IVRS.

Für eine zentrumsübergreifende Konsistenz wurde ein unabhängiges externes „Epilepsy Study Consortium“ (ESC) zur Bestätigung, Beratung und Klassifikation von Anfällen eingesetzt. Nach der Screening-Visite sollte das Prüfpersonal die Dokumentation der Krankheitsgeschichte der Testpersonen zur Verifizierung und Bestätigung der Anfallstypen an das ESC übersenden. Im Fall neu auftretender Anfallstypen während der Studie erfolgte ebenfalls eine Prüfung durch das ESC.

Der pU bietet verschiedene Auswertungen der epileptischen Anfälle an. Als Baselinephase wurden die 28 Tage vor Randomisierung bestimmt. Als Baseline sind alle Anfallsdaten als Durchschnitt pro 28 Tage definiert, die vom Tag der Studienvisite 2 bis zu Tag 1 der verblindeten Studienperiode erhoben wurden.

Häufigkeit TSC-assoziiierter Anfälle

TSC-assoziierte Anfälle waren definiert als fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen und generalisierten Krampfanfällen [tonisch-klonisch, tonisch, klonisch oder atonisch] entwickelten, die zählbar waren. Die Anfallshäufigkeit war definiert als Durchschnitt pro 28 Tage.

Folgende Operationalisierungen liegen für die TSC-assoziierten Anfälle vor:

- Veränderung der Häufigkeit TSC-assoziiierter Anfälle (Durchschnitt pro 28 Tage) während des Behandlungszeitraums im Vergleich zu Baseline (primärer Endpunkt).
- Responderanalysen
 - o Anzahl Responder, definiert als solche mit einer Reduktion um ≥ 25 , ≥ 50 , ≥ 75 oder 100 % in der Häufigkeit TSC-assoziiierter Anfälle gegenüber Baseline (insgesamt und 4-wöchentlich), die die Studie nicht abgebrochen haben.
 - o Responder, definiert als solche mit einer Verschlechterung TSC-assoziiierter Anfälle im Vergleich zu Baseline um > 25 %, keine Veränderung ≤ -25 bis ≥ 0 %, keine Veränderung > 0 bis $< +25$ %, Verbesserung um ≥ 25 bis < 50 %, Verbesserung um ≥ 50 bis < 75 % oder Verbesserung um ≥ 75 %.
- Zeit bis zur Baseline-Frequenz TSC-assoziiierter Anfälle definiert als die Anzahl der berichteten Tage im IVRS von Tag 1, bis die kumulierte Anzahl erlittener TSC-assoziiierter Anfälle größer oder gleich der Anzahl der Anfälle während der Baseline-Periode ist.
- Die Veränderung der Anzahl der anfallsfreien Tage im Vergleich zu Baseline wurde als Durchschnitt pro 28 Tage errechnet aus der Anzahl der Tage ohne TSC-assoziierte Anfälle je Periode und der Anzahl der im IVRS berichteten Tage für diese Periode.

Häufigkeit Anfälle gesamt

Berücksichtigt und zusammengefasst wurden alle Anfälle der in der Studie erhobenen Anfallstypen. Die Operationalisierung erfolgt analog zum primären Endpunkt für die Veränderung der Häufigkeit im Vergleich zu Baseline und die Responderanalysen.

Häufigkeit weiterer Anfälle

Weitere Anfälle umfassen Absencen, myoklonische Anfälle, fokal-sensorische Anfälle und infantile/epileptische Spasmen. Testpersonen ohne Anfälle eines bestimmten Subtyps während der Baseline-Periode werden für die Auswertung des jeweiligen Subtyps nicht berücksichtigt. Zusätzlich wurden außerdem die Responder analog zu den TSC-assoziierten Anfällen ausgewertet.

Zusammengesetzter fokaler Anfallswert

Für die Berechnung des fokalen Anfallswerts wird die Anzahl der fokal-motorischen Anfälle ohne Beeinträchtigung des Bewusstseins mit 1 multipliziert, die Anzahl der fokalen Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung mit 2 und die Anzahl der fokalen Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen entwickelten, mit 3. Die Operationalisierung erfolgt analog zum primären Endpunkt.

Häufigkeit von Anfalls-Subtypen

Die Häufigkeit jedes erhobenen Anfalls-Subtyps wird dargestellt. Analysiert werden lediglich fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen entwickelten, tonisch-klonische Anfälle sowie tonische Anfälle. Testpersonen ohne den jeweiligen Anfalls-Subtyp zu Baseline werden nicht für die Auswertung dieses Subtyps berücksichtigt. Die Operationalisierung erfolgt analog der Analyse für den primären Endpunkt. Zusätzlich wurden außerdem die Responder und die Anfallsfreiheit analog zu den TSC-assoziierten Anfällen ausgewertet.

- Erhebungszeitpunkte: Täglich von Baseline (Studienvisite 2) bis zum Übergang in die OLE oder bis Studienvisite 11 (Personen, die nicht an der OLE teilnahmen).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist weitgehend nachvollziehbar. Anfälle sollten entweder durch die betreuende Person oder die Patientinnen und Patienten selbst berichtet werden. Unklar bleibt, ob die Patientinnen und Patienten geistig und körperlich fähig waren selbst eine Bewertung ihrer Anfälle vorzunehmen.

Patientenrelevanz

Der Endpunkt wird als patientenrelevant bewertet.

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 Erkrankung TSC ist gekennzeichnet durch eine Vielzahl von Anfällen und Anfallstypen, die je nach Verlauf und Krankheitsschwere erheblich variieren könne. Epileptische Anfälle im Rahmen der TSC zählen zu den behandlungsresistenten Formen. Dabei sind alle Anfallsarten häufig stark

refraktär gegenüber konventionellen antiepileptischen Therapien [23]. Eine komplette Anfallsfreiheit wird zumeist nicht erreicht, weshalb für die Patientinnen und Patienten eine Reduktion der Anfallsschwere und -häufigkeit ein relevantes Behandlungsziel ist.

Validität

Die Erhebung der Häufigkeit von Anfällen erfolgte täglich standardisiert mittels IVRS.

Die Anfälle wurden folgenden Typen zugeordnet: Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung, fokale Anfälle, die sich zu bilateralen generalisierten konvulsiven Anfällen und generalisierten Krampfanfällen entwickelten, generalisierte Krampfanfälle mit tonisch-klonischen, tonischen, klonischen und atonischen Anfällen sowie weitere Anfälle, die Absenzen, myoklonische Anfälle, fokal-sensorische Anfälle und infantile/epileptische Spasmen beinhalten. Darüber hinaus wurden Status-epilepticus-Episoden erfasst.

Aus den Studienunterlagen geht nicht eindeutig hervor, wie umfangreich Anfälle im IVRS dokumentiert wurden, oder ob lediglich der Anfallstyp notiert wurde. In einem Training zum Umgang mit dem IVRS und den Tagebüchern wurden die Testpersonen oder die jeweils betreuenden Personen bei der Baseline-Visite dahingehend geschult, auf Basis der Krankheitsgeschichte die erwarteten Anfallstypen der jeweiligen Testperson zu erkennen. Im Fall von Fehlklassifikationen bzw. Abweichungen von dem Anfallsprofil sollte zur Verifizierung der Anfallstypen ein unabhängiges ESC eingesetzt werden. Welche Anfallsdaten dem ESC hierfür zu Verfügung standen ist unklar.

Die Erhebung der Anfallshäufigkeit wird insgesamt als adäquat erachtet. Es ist zu erwarten, dass die betreuenden Personen erfahren in der Detektion und Klassifizierung von epileptischen Anfällen sind, sodass die Erhebung des Endpunkts als ausreichend valide eingeschätzt wird.

Zur Validierung der angegebenen Responseschwellen legt der pU keine Unterlagen vor. Eine Reduktion der Anfallshäufigkeit ist aufgrund der schweren Symptomatik bei Patientinnen und Patienten mit TSC ein primäres Therapieziel und daher von hoher klinischer Relevanz [16].

Aus der Vielzahl an Auswertungen werden letztendlich nur folgende Operationalisierungen als relevant für die vorliegende Bewertung eingestuft:

- TSC-assozierte Anfälle
 - o Veränderung der Häufigkeit TSC-assoziierter Anfälle (Durchschnitt pro 28 Tage) während des Behandlungszeitraums im Vergleich zu Baseline.
 - o Responderanalysen
 - Anzahl Responder, definiert als Reduktion ≥ 25 , ≥ 50 , ≥ 75 oder 100 % in der Häufigkeit TSC-assoziierter Anfälle gegenüber Baseline.
 - Anzahl Responder (post hoc), definiert als Anstieg > 0 % von TSC-assozierten Anfällen im Vergleich zu Baseline.
- Anfälle gesamt
 - o Veränderung der Häufigkeit der Gesamtzahl der Anfälle (Durchschnitt pro 28 Tage) während des Behandlungszeitraums im Vergleich zu Baseline.
 - o Responderanalysen
 - Anzahl Responder, definiert als Reduktion ≥ 25 , ≥ 50 , ≥ 75 oder 100 % in der Häufigkeit der Gesamtzahl der Anfälle gegenüber Baseline.
 - Anzahl Responder (post hoc), definiert als Anstieg > 0 % von TSC-assozierten Anfällen im Vergleich zu Baseline.
- Veränderung der Häufigkeit der weiteren Anfälle (Durchschnitt pro 28 Tage) während des Behandlungszeitraums im Vergleich zu Baseline.

Die weiteren Auswertungen zum Endpunkt „Epileptische Anfälle“, wie die Subtypen von Anfällen und der zusammengesetzte fokale Anfallswert, werden als nicht bewertungsrelevant eingestuft, da sie bereits durch die genannten Operationalisierungen erfasst sind.

Status epilepticus

Der Endpunkt „Status epilepticus“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Status epilepticus ist definiert als jegliche Art von Anfall (konvulsiv und nicht-konvulsiv), der 30 Minuten oder länger andauert.

Die Erhebung von Anfällen erfolgt täglich mittels IVRS (siehe Endpunkt „Epileptische Anfälle“). Aus Konsistenzgründen sollte die Erfassung des Endpunkts immer durch dieselbe Person erfolgen.

- Erhebungszeitpunkte: Täglich von Baseline (Studienvisite 2) bis zum Übergang in die OLE oder bis Studienvisite 11 (Personen, die nicht an der OLE teilnahmen).
- Auswertung: Auf Basis der ITT-Population durchgeführt.

Bewertung

Die Operationalisierung ist nachvollziehbar. Anfälle sollten entweder durch die betreuende Person oder die Testpersonen selbst berichtet werden. Unklar bleibt, ob die Patientinnen und Patienten geistig und körperlich fähig waren eine eigenständige Bewertung ihrer Anfälle vorzunehmen.

Patientenrelevanz

Der Endpunkt „Status epilepticus“ wird als patientenrelevant bewertet.

Bei einem Status epilepticus handelt es sich per Definition um eine langandauernde epileptische Aktivität, je nach Anfallstyp auch mit Bewusstseinsstörungen verbunden. Episoden von Status epilepticus sind sehr belastende Ereignisse, deren Komplikationen lebensbedrohlich sein können [13,14,33].

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

Die Erfassung von Anfällen erfolgte standardisiert mittels IVRS. Aus den Studienunterlagen geht nicht eindeutig hervor, welche Informationen bei Auftreten eines Anfalls dokumentiert werden sollten. Im Studienbericht werden lediglich Anzahl und Art des Anfalls spezifiziert. Da die Patientinnen und Patienten oder die betreuenden Personen jedoch eine Einführung zur Nutzung des IVRS erhielten, wird die Erhebung des Endpunkts insgesamt als adäquat erachtet.

Globaler Eindruck der Veränderung (mittels CGIC / SGIC)

Der Endpunkt „Globaler Eindruck der Veränderung“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Mittels „Caregiver Global Impression of Change“ (CGIC) bzw. „Subject Global Impression of Change“ (SGIC) sollte der allgemeine Funktionsstatus der Patientinnen und Patienten bewertet werden. Hierfür sollten die Studienteilnehmenden bzw. die betreuenden Personen jeweils vor Einleitung der Studienmedikation während Studienvisite 3 (Tag 1) eine kurze schriftliche, formlose Beschreibung über den Gesamtzustand als Gedankenstütze für spätere Studienvisiten verfassen. In Studienvisite 10 erfolgte eine Abfrage zur Veränderung des Gesamtzustands der behandelten Person mittels einer 7-Punkte-Skala im Vergleich zu Studienbeginn durch

- die betreuende Person (CGIC), entspricht einer Fremdeinschätzung, und/oder
- die behandelte Person selbst (SGIC), entspricht einer Selbsteinschätzung.

Antwortoptionen der Skala waren: „sehr stark verbessert“ (1), „stark verbessert“ (2), „geringfügig verbessert“ (3), „keine Veränderung“ (4), „geringfügig verschlechtert“ (5), „stark verschlechtert“ (6), „sehr stark verschlechtert“ (7). Aus Konsistenzgründen sollte die Erfassung des Endpunkts immer durch dieselbe betreuende Person erfolgen.

Zur primären Auswertung der globalen Veränderung werden beide Instrumente berücksichtigt, wobei der CGIC priorisiert wird, falls sowohl CGIC als auch SGIC zum jeweiligen Erhebungszeitpunkt beantwortet wurden. Eine zusätzliche Sensitivitätsanalyse berücksichtigte ausschließlich den CGIC.

Zusätzlich erfolgte in Modul 4 des Nutzendossiers eine Auswertung für den Anteil der Personen mit einer Verbesserung des Allgemeinzustands („leicht verbessert“, „stark verbessert“ oder „sehr stark verbessert“).

- Erhebungszeitpunkte: Schriftliche Beschreibung des Gesamtzustands als Gedankenstütze zu Studienvisite 3 (Tag 1), Erhebung während Studienvisite 10.
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Aus den Unterlagen geht nicht eindeutig hervor, ob zu der geplanten Studienvisite regelhaft eine Beantwortung der Skala durch sowohl betreuende Person als auch Patientin / Patient vorgesehen war oder ob eine Selektion für bestimmte Patientenpopulationen (z. B. nach kognitiver Einschränkung oder dem Alter) erfolgte. Zudem ist fraglich, warum der betreuerbewertete CGIC für die Auswertung priorisiert wurde, falls sowohl CGIC als auch SGIC beantwortet vorlagen.

Eine Auswertung der Personen mit einer Verbesserung des Allgemeinzustands („leicht verbessert“, „stark verbessert“ oder „sehr stark verbessert“) wurde vom pU post hoc vorgenommen. Responderanalysen, die eine Verschlechterung berücksichtigen, liegen dagegen nicht vor.

Patientenrelevanz

Die Einschätzung der Patientin / des Patienten zur Veränderung ihres/seines Gesundheitszustands wird als patientenrelevant bewertet.

In der vorliegenden Therapiesituation ist zusätzlich eine Bewertung des aktuellen Gesundheitszustands der Testperson durch die betreuende Person vorgesehen. Grundsätzlich wird die Selbsteinschätzung der betroffenen Person zu ihrem Krankheitszustand im Vergleich zu einer Fremdeinschätzung für die Nutzenbewertung bevorzugt. Im vorliegenden Anwendungsgebiet gestaltet sich die Erhebung in Form einer Selbsteinschätzung zumindest für einen Teil der Studienpopulation schwierig, sodass eine Einschätzung durch die betreuende Person im Rahmen der Nutzenbewertung als adäquat erachtet wird.

Validität

Mit dem CGIC bzw. SGIC wird der Gesamtzustand einer Patientin / eines Patienten erfasst und bewertet, inwieweit sich seit Beginn der Therapie eine Verbesserung oder Verschlechterung des Gesamtzustandes eingestellt hat. Dazu wurde an den geplanten Erhebungszeitpunkten von den betreuenden Personen oder den Patientinnen / Patienten selbst die Veränderung des Gesamtzustands im Vergleich zu Baseline erfragt. Die Bewertung erfolgte auf einer 7-teiligen Skala von „sehr stark verbessert“ (1) bis „sehr stark verschlechtert“ (7). Aus Konsistenzgründen sollte die Erhebung des CGIC immer durch dieselbe Person erfolgen. Die Erfassung des Endpunkts erscheint entsprechend standardisiert und valide.

Informationen zur Validität der Skalen konnten weder aus den Unterlagen des pU noch mittels orientierender Literaturrecherche identifiziert werden, sodass eine abschließende Bewertung nicht getroffen werden kann. Problematisch ist zudem, dass zumindest für einen Teil der Studienpopulation keine Selbsteinschätzung durch die Patientin / den Patienten, sondern eine Fremdeinschätzung durch die betreuende Person vorliegt. Im Kontext der Schwere der TSC-Erkrankung mit kognitiven Beeinträchtigungen und einer verzögerten Entwicklung, in Verbindung mit dem teilweise jungen Alter der Personen, erscheint eine Fremdeinschätzung durch die betreuende Person an dieser Stelle angemessen. Da zudem von einem engen Patienten-Betreuer-Verhältnis in dem hier vorliegenden Anwendungsgebiet der TSC ausgegangen werden kann, ist eine valide Fremdeinschätzung der Veränderung des Gesamtzustands plausibel. Dabei wird jegliche Veränderung im Sinne einer Verbesserung des Gesundheitszustands als relevant erachtet.

Für die Auswertung ist eine Kombination der Ergebnisse der beiden Instrumente vorgesehen, wobei der CGIC priorisiert wurde, falls sowohl CGIC als auch SGIC für eine Person beantwortet vorlagen. Strukturelle Unterschiede in der Bewertung der Veränderung des Gesundheitszustands zwischen einer Selbsteinschätzung durch die Patientinnen / Patienten und einer Fremdeinschätzung durch die betreuende Person können nicht ausgeschlossen werden. Aus diesem Grund wird die getrennte Auswertung von CGIC bzw. SGIC für die Nutzenbewertung primär als relevant erachtet.

Globaler arztbewerteter Eindruck der Veränderung (mittels PGIC)

Der Endpunkt „Globaler arztbewerteter Eindruck der Veränderung“ wird in der Nutzenbewertung aufgrund fehlender Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Für die globale arztbewertete Veränderung des Gesamteindrucks (PGIC) sollte der allgemeine Gesundheitsstatus der Patientinnen und Patienten durch das ärztliche Studienpersonal bewertet werden. Hierfür sollte das Prüfpersonal vor Einleitung der Studienmedikation während Studienvisite 3 (Tag 1) eine kurze schriftliche, formlose Beschreibung über den Gesamtzustand als Gedankenstütze für spätere Studienvisiten verfassen. In der Studienvisite 10, zum Ende der Behandlung, erfolgte eine

Abfrage zur Veränderung des Gesamtzustands der behandelten Person mittels einer 7-Punkte-Skala (1 = „sehr stark verbessert“ bis 7 = „sehr stark verschlechtert“) im Vergleich zum Behandlungsbeginn. Aus Konsistenzgründen wurde die Erfassung des Endpunkts durch dieselbe behandelnde Person bevorzugt.

Für den Endpunkt wurden post hoc zudem Responderanalysen für den Anteil der Patientinnen und Patienten mit einer Verbesserung des Allgemeinzustands („leicht verbessert“, „stark verbessert“ oder „sehr stark verbessert“) durchgeführt.

- Erhebungszeitpunkte: Schriftliche Beschreibung des Gesamtzustands als Gedankenstütze zu Studienvisite 3 (Tag 1), Erhebung während Studienvisite 10.
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar. Responderanalysen für eine Verschlechterung der globalen arztbewerteten Veränderung des Gesamteindrucks liegen nicht vor.

Patientenrelevanz

Der globale Gesundheitszustand der Patientinnen und Patienten wird prinzipiell als patientenrelevant erachtet.

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

Mit dem PGIC wird der Gesamtzustand durch das ärztliche Studienpersonal erfasst und bewertet, inwieweit sich seit Beginn der Therapie eine Verbesserung oder Verschlechterung des Gesundheitszustands eingestellt hat. Gemäß Studienprotokoll sollte „bevorzugt“ die gleiche Prüfperson die Erhebung des PGIC bei einer Testperson an den beiden relevanten Studienvisiten vornehmen. Unklar ist, in wie vielen Fällen die Bewertungen von unterschiedlichen Ärztinnen und Ärzten vorgenommen wurde.

Allgemein erscheint fraglich, inwiefern ärztliches Prüfpersonal im Rahmen einer Fremdbewertung Veränderungen im globalen Gesundheitszustand der Testpersonen valide erfassen kann. Eine schriftliche Dokumentation soll zu Studienvisite 3 zwar als Gedankenstütze fungieren, dennoch begegnet das ärztliche Prüfpersonal den Patientinnen und Patienten lediglich in größerem zeitlichen Abstand bei den Studienvisiten. Da zudem von einem engen Patienten-Betreuer-Verhältnis in dem hier vorliegenden Anwendungsgebiet der TSC ausgegangen werden kann, wird die Fremdbewertung des globalen Eindrucks der Veränderung mittels CGIC als valider eingeschätzt, verglichen mit der Fremdbewertung durch das Prüfpersonal.

Veränderung der Anfallsdauer (mittels CGICSD / SGICSD)

Der Endpunkt „Veränderung der Anfallsdauer“ wird in der Nutzenbewertung aufgrund fehlender Validität nicht dargestellt.

Operationalisierung

Beschreibung

Die Dauer der Anfälle wurde mittels Caregiver- und Subject Global Impression of Change in Seizure Duration (CGICSD / SGICSD) erhoben. Hierfür sollten die Studienteilnehmenden bzw. die betreuenden Personen jeweils vor Einleitung der Studienmedikation während Studienvisite 3 (Tag 1) eine kurze schriftliche Einschätzung zur durchschnittlichen Dauer je Anfallstyp als Gedankenstütze für spätere Studienvisiten verfassen. In Studienvisite 10 erfolgte eine Abfrage zur Veränderung der durchschnittlichen Dauer je Anfallstyp der behandelten Person mittels einer 3-Punkte-Skala im Vergleich zu Studienbeginn durch

- die betreuende Person (CGICSD), entspricht einer Fremdeinschätzung, und/oder
- die behandelte Person selbst (SGICSD), entspricht einer Selbsteinschätzung.

Antwortoptionen der Skala waren (1) Verringerung der durchschnittlichen Dauer, (2) keine Veränderung, (3) Erhöhung der durchschnittlichen Dauer. Niedrigere Werte auf der Skala entsprechen einer kürzeren Anfallsdauer. Aus Konsistenzgründen sollte die Erhebung des CGICSD immer durch dieselbe Betreuungsperson erfolgen. Falls sowohl Betreuer- als auch Patienten-Version vorlagen, wurde die Betreuer-Version für die Auswertung berücksichtigt. In die Analyse eines jeden Anfallstyps wurden gemäß Studienbericht nur Personen eingeschlossen, die mindestens einen Anfall des jeweiligen Typs in der Studie erlitten hatten.

- Erhebungszeitpunkte: Schriftliche Beschreibung der durchschnittlichen Anfallsdauer als Gedankenstütze an Studienvisite 3 (Tag 1), Erhebung während Studienvisite 10 (Behandlungsende).
- Auswertung: Für jeden Anfallstyp wurden gemäß Studienbericht nur die Testpersonen der ITT-Population in die Analyse eingeschlossen, die mindestens einen Anfall des jeweiligen Typs zu einem beliebigen Zeitpunkt in der Studie berichtet haben.

Bewertung

Die Operationalisierung des Endpunkts ist weitestgehend nachvollziehbar. Aus den Unterlagen geht nicht eindeutig hervor, ob zu der geplanten Studienvisite regelhaft eine Beantwortung der Skala durch sowohl betreuende Person als auch Patientin / Patient vorgesehen war, oder ob eine Selektion für bestimmte Patientenpopulationen (z. B. nach kognitiver Einschränkung oder dem Alter) erfolgte. Zudem ist fraglich, warum der betreuerbewertete CGICSD für die Auswertung priorisiert wurde, falls sowohl CGICSD als auch SGICSD vorlagen. Weiterhin bestehen unterschiedliche Angaben darüber, welche Personen in die Analyse eingehen. Nach Angabe im Studienbericht werden alle Testpersonen berücksichtigt, die mindestens einen Anfall des jeweiligen Typs zu einem beliebigen Zeitpunkt in der Studie berichtet haben. In Modul 4 ist beschrieben, dass alle Testpersonen in die Analyse eingingen, die während des Behandlungszeitraums mindestens einen Anfall des jeweiligen Typs hatten.

Patientenrelevanz

Der Endpunkt „Veränderung der Anfallsdauer“ wird als patientenrelevant erachtet.

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

Mit dem CGICSD und dem SGICSD wurde der Gesamteindruck zur Veränderung der durchschnittlichen Dauer der Anfälle ermittelt. Dazu wurde am Ende der Behandlung von den betreuenden Personen oder den Patientinnen / Patienten selbst die Veränderung in der Dauer für jeden Anfalls-Subtyp im Vergleich zu Baseline erfragt.

Die kumulierte, subjektive Bewertung der Dauer eines Anfallstyps lässt keine direkte Schlussfolgerung auf die Häufigkeit und Schwere von epileptischen Anfällen zu. Darüber hinaus ist unklar, wie sensitiv relevante Veränderungen der Anfallsdauer auf einer 3-stufigen Skala durch die subjektive Bewertung der betreuenden Person im Falle des CGICSD erfasst werden können. Darüber hinaus ist schwer abschätzbar, ob die von einem Anfall betroffenen Personen die Dauer eines Anfalls selbst bestimmen können. Dies gilt insbesondere für Anfälle mit einer Beeinträchtigung des Bewusstseins. Zuletzt gehen in die Auswertung für die Bewertung der Veränderung der Anfallsdauer des jeweiligen Anfallstyps nur Personen in die Analyse ein, die diesen Anfallstyp während der Studie auch erlitten hatten. Die Anzahl der Personen in der jeweiligen Analyse liegt teilweise erheblich unter der der ITT-Population. Zudem ist der Umgang mit Personen unklar, die während der Baseline-Periode einen Anfall eines bestimmten Anfallstyps erlitten hatten, nicht jedoch während der Behandlungsphase oder umgekehrt.

Die Validität des Endpunkts wird kritisch gesehen; auf eine Darstellung der Ergebnisse des CGICSD / SGICSD wird verzichtet. Schwere und lang andauernde epileptische Anfälle sind über den Endpunkt „Status epilepticus“ abgedeckt.

Verwendung von Notfallmedikation

Der Endpunkt „Verwendung von Notfallmedikation“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

Operationalisierung

Beschreibung

In Studie GWEP1521 wurde erhoben an wie vielen Tagen seit der letzten Studienvisite Notfallmedikamente verwendet wurden. Aus der Anzahl der berichteten Tage und der Anzahl der Tage, an denen Notfallmedikamente eingenommen wurde, wird daraufhin die durchschnittliche Häufigkeit der Notfallmedikation pro 28 Tage für den Behandlungs- und den Baseline-Zeitraum berechnet.

Ausgewertet wurde die durchschnittliche Häufigkeit der Notfallmedikation pro 28 Tage im Behandlungszeitraum gegenüber dem Baseline-Zeitraum.

- Erhebungszeitpunkte: Bei jeder Studienvisite und telefonisch im Sicherheits-Follow-up.
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

Die Einnahme einer Notfallmedikation stellt per se kein patientenrelevantes Ereignis dar. Ereignisse, in deren Folge Notfallmedikationen eingenommen werden, werden in der Studie GWEP1521 bereits direkt in den anfallsbezogenen Endpunkten erhoben.

Validität

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

Epilepsiebedingte Hospitalisierungen

Der Endpunkt „Epilepsiebedingte Hospitalisierungen“ wird in der Nutzenbewertung ergänzend berücksichtigt.

Operationalisierung

Beschreibung

Epilepsiebedingte Hospitalisierungen umfassen die Anzahl der stationären Krankenhausaufenthalte, die nach Ansicht des Prüfpersonals epilepsiebedingt waren. Die epilepsiebedingten Krankenhausaufenthalte wurden bei jeder Studienvsiste erfasst und im CRF sowie als schwerwiegende unerwünschte Ereignisse dokumentiert.

- Erhebungszeitpunkte: Zu jeder Studienvsiste ab Studienvsiste 2 (Tag -28) über den gesamten Studienzeitraum (einschließlich Sicherheits-Follow-up).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist nicht vollständig nachvollziehbar. Hospitalisierungen sollten sowohl im CRF als auch im Rahmen der Erfassung der SUE dokumentiert werden. Welche Dokumentation letztendlich als Datengrundlage für die Auswertung herangezogen wurde ist unklar. Weiterhin geht aus der Dokumentation im CRF nicht eindeutig hervor, wie ein Krankenhausaufenthalt definiert ist. Das heißt, ob eine Übernachtung im Krankenhaus erforderlich ist, oder auch ambulante oder teilstationäre Aufnahmen ausreichen.

Patientenrelevanz

Der Endpunkt „Epilepsiebedingte Hospitalisierung“ wird als patientenrelevant bewertet. Berücksichtigt werden nicht alle in der Studie aufgetretenen Hospitalisierungen, sondern nur solche mit Epilepsiebezug. Dabei bleibt unklar, nach welchen Kriterien die Bewertung erfolgte bzw. ob es definierte Kriterien für einen Epilepsiebezug gab. Prinzipiell sind alle in der Studie aufgetretenen Hospitalisierungen relevant.

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

Hospitalisierungen wurden bei jeder Studienvsiste ab Tag 1 berichtet und vom Prüfpersonal als epilepsiebedingt beurteilt. Kriterien, in welchen Fällen es sich um eine epilepsiebedingte Hospitalisierung handelt, konnten nicht identifiziert werden. Da es sich um eine multizentrische Studie handelt bleibt unklar, ob regionale Unterschiede bestehen, die zu einer Verzerrung der Ergebnisse führen können. Zudem finden sich keine Informationen zur Übertragbarkeit auf den deutschen Versorgungskontext; keines der Studienzentren befindet sich in Deutschland. Die Validität des Endpunkts kann abschließend nicht beurteilt werden.

Soziale Kommunikation (mittels SCQ)

Der Endpunkt „Soziale Kommunikation“ wird in der Nutzenbewertung aufgrund fehlender Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Der Fragebogen „Social Communication Questionnaire“ (SCQ) sollte von den betreuenden Personen aller Patientinnen und Patienten über 4 Jahren mit einem geistigen Alter von mindestens 2 Jahren ausgefüllt werden. Das Instrument erfasst das Verhalten über den Zeitraum der vergangenen 3 Monate mittels 40 Items, die jeweils mit „Ja“ oder „Nein“ beantwortet werden. Den Antworten werden Werte von 0 oder 1 zugeschrieben, sodass niedrigere Werte einen besseren Zustand bedeuten.

Die Antwort des 1. Items („Ist sie/er in der Lage mit kurzen Phrasen oder Sätzen zu kommunizieren“) wird nicht gewertet, sondern zur Bestimmung der verbleibenden Items genutzt, die für „Abnormale Sprache“ in den Gesamtwert eingehen: Falls die Antwort „Ja“ ist, dann werden die 6 Items zu „Abnormale Sprache“ berücksichtigt und der Gesamtwert reicht von 0 bis 39 Punkten; falls die Antwort für Item 1 „Nein“ ist, werden diese 6 Items nicht berücksichtigt woraus sich ein Gesamtwert von 0 bis 33 Punkten ergibt.

Aus den Angaben werden neben dem Gesamtwert folgende Domänenwerte gebildet:

- Wechselseitige soziale Interaktion (Items 9, 10, 19, 26–33, 36 37, 39, 40)
- Kommunikation (Items 2–6, 20–25, 34, 35)
- Stereotype Verhaltensmuster (Items 7, 8, 11–16)

Ausgewertet wird die Veränderung des Gesamtwerts und der Domänenwerte im Vergleich zu Baseline.

- Erhebungszeitpunkte: Zu Behandlungsbeginn (Studienvisite 3) und zu Behandlungsende (Studienvisite 10).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist nur teilweise nachvollziehbar. Unklar ist, wie das geistige Alter der Testpersonen ermittelt wird und inwiefern Testpersonen, für die der SCQ nicht ausgefüllt wurde, in die Analysen eingingen.

Patientenrelevanz

Kognitive Beeinträchtigungen und Störungen des autistischen Spektrums sind prinzipiell patienten-relevant und treten in der Indikation TCS häufig auf [24,34].

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

Beim SCQ handelt es sich um etabliertes Instrument zur Erfassung von abnormen sozialen Interaktions- und Kommunikationsmustern sowie stereotypen Verhaltensweisen im Vorfeld einer eingehenderen klinischen Diagnostik, insbesondere bei Personen mit Verdacht auf Autismus oder anderen Störungen des autistischen Spektrums ab einem Alter von 4 Jahren bzw. einem Entwicklungsalter von mindestens 2 Jahren. Es folgt den diagnostischen Leitlinien von ICD-10 und DSM-IV-TR.

Die Fähigkeit des Instruments zwischen Patientengruppen mit Störungen des autistischen Spektrums und der allgemeinen Bevölkerung zu diskriminieren ist gut untersucht [6]; auch in der Indikation TSC wurde das Instrument untersucht [17]. Unklar ist, inwiefern sich der Fragebogen für die Evaluation von bestehenden kognitiven Beeinträchtigungen in klinischen Studien eignet. Entsprechende Untersuchungen konnten nicht identifiziert werden. Aus diesem Grund und den Unklarheiten bei der Operationalisierung werden die Ergebnisse des SCQ nicht in der Nutzenbewertung berücksichtigt.

Achenbach-Verhaltens-Checkliste

Der Endpunkt „Achenbach-Verhaltens-Checkliste“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Bei den Achenbach-Verhaltens-Checklisten handelt es sich um fremdbewertete Fragebögen zur Erfassung von Verhaltenssymptomen und emotionalen Symptomen bei Kindern und Erwachsenen. Es existieren 2 Versionen, die „Child Behavior Checklist“ (CBCL), die abhängig vom Alter eingesetzt wird, und die „Adult Behavior Checklist“ (ABCL) für Erwachsene:

- „CBCL (1,5–5 Jahre)“ für Kinder im Alter von 18 Monaten bis ≤ 5 Jahren mit 100 Items; Bezugszeitraum sind die letzten 2 Monate.
- „CBCL (6–18 Jahre)“ für Kinder im Alter von 6 bis ≤ 17 Jahren mit 113 Items, Bezugszeitraum sind die letzten 6 Monate.
- „ABCL (18–59 Jahre)“ für Erwachsene im Alter von 18 bis ≤ 59 Jahren mit 126 Items.

Die Antwortoptionen für den Hauptteil befinden sich auf einer Likert-Skala von 0 (= nicht wahr), 1 (= etwas oder manchmal wahr) bis 2 (= sehr wahr oder oft wahr). Höhere Werte bedeuten mehr berichtete Probleme. Ähnliche Items werden wie folgt gruppiert und zu Syndromskalen-Werten zusammengefasst, die weiter zu Problemskalen aggregiert werden.

Tabelle 6: Problemskalen und Syndromskalen der jeweiligen Versionen der Achenbach-Verhaltens-Checklisten

Version der Checkliste	Problemskala	Syndromskalen
CBCL (1,5–5 Jahre)	Internalisierung	Emotionale Reaktivität
		Ängstlich/Depressiv
		Körperliche Beschwerden
		Sozialer Rückzug
	Externalisierung	Aufmerksamkeitsprobleme
		Aggressives Verhalten
	Weitere	Andere Probleme
	Gesamt	Schlafprobleme

Version der Checkliste	Problemskala	Syndromskalen
CBCL (6–18 Jahre)	Internalisierung	Ängstlich/Depressiv
		Rückzug/Depressiv
		Körperliche Beschwerden
	Externalisierung	Regelverletzendes Verhalten
		Aggressives Verhalten
	Weitere	Soziale Probleme
		Denk- und repetitive Probleme
		Aufmerksamkeitsprobleme
		Andere Probleme
	Gesamt	
ABCL (18–59 Jahre)	Internalisierung	Ängstlich/depressiv
		Rückzug
		Körperliche Beschwerden
	Externalisierung	Aggressives Verhalten
		Regelverletzendes Verhalten
		Aufdringlich
	Weitere	Denk- und repetitive Probleme
		Aufmerksamkeitsprobleme
		Andere Probleme
	Gesamt	

Abkürzungen: ABCL: Adult Behavior Checklist; CBCL: Child Behavior Checklist.

Zusätzlich können 3 Dimensionen zum Substanzgebrauch (Tabak, Alkohol, Medikamente/Drogen) betrachtet werden. Für „CBCL (6–18 Jahre)“ kann, neben den in Tabelle 7 aufgeführten Skalen, noch eine weitere Skala für die Gesamtkompetenz mit den Subskalen „Aktivitäten“, „Soziale Kompetenzen“ und „Schule“ gebildet werden:

- **Aktivität:** Ist zusammengesetzt aus weiteren 6 Items, die nicht in den 100 Items der „CBCL (6–18 Jahre)“ enthalten sind und nicht in die Problemskalen eingehen. Falls mehr als 1 Item nicht beantwortet worden ist, wird kein Gesamtwert für die Aktivität gebildet. Es bestehen klare Regeln für die Imputation eines fehlenden Items. Bewertet werden neben der Anzahl der (1) Sportarten, (2) Hobbys und (3) Arbeiten bzw. Hausarbeiten, die das Kind ausführt, wie diese jeweils bezüglich aufgewendeter Zeit und Fähigkeiten des Kindes in Relation mit anderen Kindern der Altersgruppe zu werten sind.
- **Soziale Kompetenzen:** Ist zusammengesetzt aus weiteren 6 Items, die nicht in den 100 Items des CBCL/6-18 enthalten sind und nicht in die Problemskalen eingehen. Falls mehr als 1 Item nicht beantwortet worden ist, wird kein Gesamtwert für die Aktivität gebildet. Es bestehen klare Regeln für die Imputation eines fehlenden Items. Bewertet werden Anzahl der (1) Mitgliedschaften in Organisationen, Teams, Gruppen inklusive (2) Vergleich mit der Altersgruppe des Kindes, (3) Anzahl der Freunde, (4) Wöchentliche Aktivitäten mit Freunden, (5) Umgang mit Sozialkontakten im Vergleich zu anderen Kindern dieser Altersgruppe und (6) Fähigkeit allein zur Spielen und Arbeiten.
- **Schule:** Ist zusammengesetzt aus weiteren 4 Items, die nicht in den 100 Items des CBCL/6-18 enthalten sind und nicht in die Problemskalen eingehen. Falls mehr als ein 1 Item nicht beantwortet worden ist, wird kein Gesamtwert für die Aktivität gebildet. Bewertet wird (1)

Leistung in schulischen Fächern, (2) Vorhandensein schulischer Sonderbehandlungen, (3) Wiederholung von Jahrgangsstufen sowie (4) leistungsassoziierte oder anderweitige Probleme in der Schule.

Auch für „ABCL (18–59 Jahre)“ können noch 2 Skalen für Freunde und kritische Items gebildet werden:

- Die Skala der kritischen Items besteht aus 19 Items, die bereits in den in Tabelle 7 genannten Syndromskalen enthalten sind.
- Neben den 123 Items werden darüber hinaus noch 4 Items zu Freundschaften abgefragt über die Anzahl der Freunde, die Häufigkeit des Kontakts, die Umgänglichkeit und die Häufigkeit der Familien- oder Freundschaftsbesuche. Antworten werden auf einer ordinalen Skala mit 4 Ausprägungen abgegeben. Den Antworten werden Punktwerte von 0 bis 3 zugewiesen und für den Gesamtwert der Freunde-Skala summiert. Falls ein Item nicht beantwortet wurde, wird kein Gesamtwert gebildet.

Darüber hinaus können die Symptome auch zu 6 DSM-orientierten Skalen („Depressive Symptome“, „Angstsymptome“, „Körperliche Symptome“, „Symptome vermeidender Persönlichkeit“, „Unaufmerksamkeit-Hyperaktivitätssymptome“, „Symptome antisozialer Persönlichkeit“) zusammengefasst werden, was jedoch in der vorliegenden Studie nicht erfolgt ist.

Im Fall der „CBCL (1,5–5 Jahre)“ wurde kein Score für die Sprachentwicklungs-Abfrage gebildet. Für die „ABCL (18–59 Jahre)“ erfolgte kein Scoring für die (Ehe-)Partner-Funktionsskala oder die Skala zum Konsum von Tabak, Alkohol und Drogen Rohwerte der 3 Checklisten für die Syndrom- und Problemskalen werden für jede Erhebung errechnet und die Veränderung im Vergleich zu Baseline berichtet.

Fehlende Angaben

CBCL: Falls Angaben für mehr als 8 Items (ausgenommen Item 100 in der „CBCL (1,5–5 Jahre)“ und ausgenommen Items 56h und 114 in der „CBCL (6–18 Jahre)“ fehlen, werden keine Syndrom- und Problemskalen gebildet. Bei der „ABCL (18–59 Jahre)“ werden die Items 2, 4, 15, 49, 73, 80, 88, 98, 106, 109, 110 und 123 nicht auf die fehlenden Items angerechnet und den Items 56a und 56g wird bei fehlenden Angaben der Wert 0 zugeschrieben.

- Erhebungszeitpunkte: Zu Behandlungsbeginn (Studienvisite 3) und zu Behandlungsende (Studienvisite 10).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar.

Patientenrelevanz

Verhaltensprobleme und emotionale Funktionen werden als patientenrelevante Endpunkte bewertet. Die Erkrankung TSC geht mit einer Vielzahl von Verhaltensproblemen und kognitiven Auffälligkeiten einher [10,11]. Diese sind für die Patientinnen und Patienten selbst und ihr Umfeld von hoher Relevanz.

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

Die Achenbach-Verhaltens-Checklisten sind verschiedene, nach Altersgruppen (1–5, 6–18, 18–59 Jahre) aufgeteilte Instrumente, um internalisierendes Verhalten (wie Depression und Angst), externalisierendes Verhalten (wie Aggression), Stress, obsessiv-zwanghaftes Verhalten oder träges kognitives Tempo der Patientinnen und Patienten durch ihre betreuende Person zu bewerten. Es handelt sich um ein generisches Instrument zur Erfassung psychischer Störungen. Für alle Checklisten liegen umfangreiche Manuals vor, in denen Informationen zu Entwicklung, Auswertung und Validierung gegeben werden [2,3,4]. Angaben zu Reliabilität und Validität der Instrumente sind in diesen Manuals dokumentiert. Die Problem- und Syndromskalen wurden empirisch über Faktoranalysen generiert [1]. Die Achenbach-Verhaltens-Checklisten sind bereits vielfach in unterschiedlichen Indikationen eingesetzt worden, auch im Krankheitsbild der Epilepsie. Auch im Indikationsgebiet TSC wird der Einsatz von Verhalts-Checklisten zur Erfassung von Verhaltensproblemen und kognitiven Auffälligkeiten empfohlen [10].

Ausreichende Informationen zu Änderungssensitivität der Instrumente konnten dagegen nicht identifiziert werden. Wie sensitiv Veränderungen aufgrund einer Intervention erfasst werden können, auch vor dem Hintergrund der relativ kurzen Behandlungsphase, ist fraglich. Untersuchungen zu einer klinischen Relevanzschwelle liegen für das Instrument nicht vor. Eine etablierte klinische Relevanzschwelle wird weder in der Literatur noch seitens des pU beschrieben.

Bewertung des adaptiven Verhaltens (mittels Vineland-II)

Der Endpunkt „Adaptives Verhalten“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die Bewertung des adaptiven Verhaltens erfolgte mittels „Vineland Adaptive Behavior Scales, Second Edition“ (Vineland-II). Ziel der Vineland-II ist es die persönlichen und sozialen Fähigkeiten von Kindern oder Erwachsenen in alltäglichen Situationen zu erfassen; im Besonderen bei Personen, die geistig zurückgeblieben sind oder die Schwierigkeiten haben andere Testsituationen zu bestehen. Die Beurteilung des adaptiven Verhaltens erfolgt jeweils durch die betreuende Person. Es existieren verschiedene Versionen des Instruments. Die vorliegende Betreuerversion umfasst 5 Domänen, die wiederum 2 bis 3 Subdomänen (insgesamt 11 Subdomänen) enthalten. Alle Items sind zudem zur Beantwortung an spezifische Altersspannen geknüpft, weshalb nur bestimmte Items für eine Person beantwortet werden müssen:

Adaptive Verhaltensdomänen

- Kommunikation (99 Items)
 - o Rezeptiv (≥ 0 Jahre)
 - o Expressiv (≥ 0 Jahre)
 - o Geschriebene Sprache (≥ 3 Jahre)
- Fähigkeiten des alltäglichen Lebens (109 Items)
 - o Persönlich (≥ 0 Jahre)
 - o Häuslich (≥ 1 Jahre)
 - o Gemeinschaft (≥ 1 Jahre)
- Soziale Fähigkeiten (99 Items)
 - o Zwischenmenschliche Beziehungen (≥ 0 Jahre)
 - o Spielen und Freizeit (≥ 0 Jahre)
 - o Bewältigungsstrategien (≥ 1 Jahre)

- Motorische Fähigkeiten (76 Items)
 - o Grobmotorik (≥ 0 bis < 7 Jahre)
 - o Feinmotorik (≥ 0 bis < 7 Jahre)

Maladaptives Verhalten

- Index für maladaptives Verhalten (36 Items)
 - o Internalisierung – Teil A (≥ 3 Jahre)
 - o Externalisierung – Teil B (≥ 3 Jahre)
 - o Weitere – Teil C (≥ 3 Jahre)
- Maladaptive verhaltenskritische Elemente (Teil D, 14 Items, ≥ 3 Jahre)

Die Antwortoptionen entsprechen einer Häufigkeitsabfrage des jeweiligen Items und reichen von 2 (= üblicherweise), 1 (= manchmal), 0 (= niemals) bis DK (= weiß nicht). Bei einigen Items existiert als weitere Antwortoption N/O (= nicht möglich), sofern bislang keine entsprechende Situation aufgetreten ist. Beispielsweise die Arbeitsleistung, sofern die betroffene Person nicht arbeitstätig ist. In der Domäne „Maladaptives Verhalten“ beschränken sich die Antwortoptionen auf 2 (= oft), 1 (= manchmal) und 0 (=nie). In der Subdomäne „Maladaptive verhaltenskritische Elemente“ soll zusätzlich das Ausmaß der Kategorie mit schwerwiegend (S) oder moderat (M) bewertet werden. Abgefragt wird dabei der Zustand zum aktuellen Zeitpunkt.

Für jede Subdomäne wird ein Rohwert auf Basis der darin enthaltenen Items berechnet. Diese Skalenwerte der Subdomänen werden zu den Domänenwerten zusammengefasst. Die Domänen „Kommunikation“, „Selbstständigkeit“, „Soziale Fähigkeiten“ und „Motorische Fähigkeiten“ ergeben zusammen den Summenwert für das adaptive Verhalten. Für den Index für maladaptives Verhalten wird der Summenwert (Rohwert) der Werte der 3 Subdomänen (Teil A–C) gebildet.

Anhand dieser Rohwerte und dem Alter der Patientinnen/Patienten wird folgendes bestimmt:

- v-Skalenwert: Ist eine Art Standardwert-Skala (standardisiert nach Alter), die die individuelle, relative Funktionsfähigkeit beschreibt; Punktzahl 1–24.
- 90%-Konfidenzintervall (KI) für die v-Skalenwerte.
- Anpassungsfähigkeit (Adaptives Level): Ein Mittel, um die Leistung eines Individuums mit Begriffen zu beschreiben, die vielseitig sind (gering, mittelmäßig gering, angemessen, mittelmäßig hoch, hoch); für das maladaptive Verhalten: durchschnittlich, erhöht oder klinisch signifikant.
- Altersäquivalent: Gibt das durchschnittliche Alter für den errechneten Wert an (nicht für Index des maladaptiven Verhaltens und maladaptive verhaltenskritische Elemente).

Für jede adaptive Verhaltensdomäne wird die Summe der v-Skalenwerte in Verbindung mit dem Alter der Patientinnen und Patienten herangezogen, um unter anderem folgende Werte zu bestimmen:

- Standardwerte für die Domäne (standardisiert nach Alter); Punktwerte 20–160.
- 90%-KI für den Standardwert der Domäne.
- Adaptives Level (gering, mittelmäßig gering, adäquat, mittelmäßig hoch, hoch).

Über das Aufsummieren der Standardwerte der adaptiven Verhaltensdomänen wird ein Gesamtwert für das adaptive Verhalten bestimmt, wobei die Domäne „Motorische Fähigkeiten“ bei Personen ≥ 7 Jahren nicht berücksichtigt wird.

Zur Berechnung des Index für maladaptives Verhalten müssen alle Items innerhalb eines Abschnitts beantwortet werden, andernfalls wird der Wert nicht berechnet. Sofern ≥ 2 Items einer Subdomäne

entweder mit „weiß nicht“ oder nicht beantwortet wurden, werden sowohl Subdomänenwert als auch Domänen- und Summenwert für das adaptive Verhalten als fehlend gewertet.

Die adaptiven Levels, die den v-Skalenwerten und den Standardwerten entsprechen, sind in Tabelle 7 dargestellt. Jedes adaptive Level für das adaptive Verhalten wird mit einem Wert von 1 (= gering) bis 5 (= hoch) kodiert. Das adaptive Level für den Index für maladaptives Verhalten wird mit einem Wert von 1 (= klinisch signifikant) bis 3 (= durchschnittlich) kodiert.

Tabelle 7: Adaptive Level nach v-Skalenwert und Standardwert für Vineland-II

Adaptives Level	v-Skalenwert für Subdomänen und Index für maladaptives Verhalten	Standardwert für Domäne und zusammengefasstes adaptives Verhalten
Domänen adaptives Verhalten		
gering	1–9	20–70
mittelmäßig gering	10–12	71–85
adäquat	13–17	86–114
mittelmäßig hoch	18–20	115–129
hoch	21–24	130–160
Maladaptives Verhalten		
klinisch signifikant	21–24	
erhöht	18–20	
durchschnittlich	1–17	

- Erhebungszeitpunkte: Zu Behandlungsbeginn (Studienvsiste 2) und zu Behandlungsende (Studienvsiste 10).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Der pU ordnet den Endpunkt der Kategorie „Lebensqualität“ zu. Dieser Einschätzung wird nicht gefolgt. Da die Vineland-II ein Instrument zur Erfassung des adaptiven Verhaltens ist und Entwicklungsstörungen zur Krankheitssymptomatik der TSC gehören, wird der Endpunkt der Kategorie „Morbidity“ zugeordnet.

Patientenrelevanz

Neben den Anfällen sind häufig geistige Einschränkung unterschiedlicher Schweregrade, Verhaltensstörungen und psychische Erkrankungen Kennzeichen der TSC. Die Erkrankung TSC geht mit einer Vielzahl von Verhaltensproblemen und kognitiven Auffälligkeiten [10,11] einher.

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

Die Vineland-II ist ein standardisiertes, sehr umfangreiches und aufwändiges Instrument zur Bewertung der Gesamtleistung der adaptiven Funktionalität. Der Fragebogen wurde für Personen im Alter von 0 bis 90 Jahren entwickelt. In der vorliegenden Studie wurde die Betreuerversion des Fragebogens verwendet. Das Instrument umfasst die Domänen „Kommunikation“, „Fähigkeiten

des alltäglichen Lebens“, „Soziale Fähigkeiten“ und „Motorische Fähigkeiten“ sowie eine zusätzliche, optionale Domäne zum maladaptiven Verhalten [36]. Die 4 adaptiven Verhaltensdomänen bilden den Gesamtwert für das adaptive Verhalten. Für das Instrument liegen auch Normwerte spezieller Patientenpopulationen vor, wie z. B. Menschen mit Autismus.

Der pU führt keine Quellen zur Validierung des Endpunkts oder zum Einsatz des Instruments in der vorliegenden Indikation an. Validierungsstudien an Patientinnen und Patienten mit TSC konnten auch in einer orientierenden Recherche nicht identifiziert werden. In mehreren Studien wurden die psychometrischen Gütekriterien jedoch bei unterschiedlichen Patientenpopulationen in jeglichen Altersgruppen untersucht. Es zeigten sich u. a. eine gute Reliabilität (Test-Retest- und Interrater-Reliabilität), interne Konsistenz, Inhalts-, Kriteriums- und Konstruktvalidität. Die Vineland-II wurde zudem bereits bei Patientinnen und Patienten mit Epilepsie im Kindesalter eingesetzt [5,9,26,27,35].

Aufgrund der generischen Art des Fragebogens und der vielfältigen Validierung der Vineland-II in unterschiedlichen Patientenpopulationen, darunter auch Personen mit Entwicklungsstörungen, wird die Validität des Instruments als hinreichend angesehen. Wie sensitiv mit dem Instrument Veränderungen aufgrund einer Intervention erfasst werden können ist fraglich. Untersuchungen zu einer klinischen Relevanzschwelle liegen für das Instrument nicht vor. Eine etablierte klinische Relevanzschwelle wird weder in der Literatur noch seitens des pU beschrieben.

Unklarheit besteht weiterhin darüber, wie sich die v-Skalenwerte aus den Rohwerten der einzelnen Skalen ergeben. Darüber hinaus gibt es keine Angaben über die Referenzpopulation, die den Standardwerten zugrunde liegt.

Die Erhebung des Endpunkts erfolgt durch eine Fremdeinschätzung des altersabhängigen, adäquaten, adaptiven Verhaltens. Dies erscheint in der vorliegenden Indikation akzeptabel; zum einen aufgrund der Krankheitssymptomatik, zum anderen, weil davon auszugehen ist, dass die betreuende Person den Studienteilnehmenden täglich betreut und dessen Funktionalität entsprechend gut einschätzen kann. Die Ergebnisse des Endpunkts werden für die Nutzenbewertung von Cannabidiol herangezogen.

2.3.3 Lebensqualität

Lebensqualität bei kindlicher Epilepsie (mittels QOLCE)

Der Endpunkt „Lebensqualität bei kindlicher Epilepsie“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Der Fragebogen „Quality of Life in Childhood Epilepsy“ (QOLCE) ist ein krankheitsspezifisches Instrument zur Einschätzung der Auswirkungen einer Epilepsie auf die Alltagsfunktionen bei Kindern mit Epilepsie im Alter von 4 bis 18 Jahren in verschiedenen Lebensbereichen. Der Fragebogen besteht aus insgesamt 77 Items und ist von den Eltern oder der betreuenden Person auszufüllen. Zusammengefasst ergeben die Items 16 Subskalen und 7 Domänen:

- Körperliche Aktivitäten
 - o Körperlichen Einschränkungen (10 Items)
 - o Energie/Fatigue (2 Items)
- Wohlbefinden
 - o Depression (4 Items)
 - o Angst (6 Items)

- o Kontrolle/Hilflosigkeit (4 Items)
- o Selbstvertrauen (5 Items)
- Kognition
 - o Aufmerksamkeit (5 Items)
 - o Erinnerungsvermögen (6 Items)
 - o Sprache (8 Items)
 - o Sonstige kognitive Fähigkeiten (3 Items)
- Soziale Aktivitäten
 - o Interaktion (3 Items)
 - o Soziale Aktivitäten (3 Items)*
 - o Stigma (1 Item)
- Verhalten (15 Items)
- Allgemeine Gesundheit (generische Subskala mit 1 Item)*
- Lebensqualität (1 Item)*

Die Antwortoptionen befinden sich mehrheitlich auf einer qualitativen 6-Punkte-Skala von 1 (= sehr häufig / die ganze Zeit / exzellent) bis 5 (= niemals / zu keiner Zeit / schlecht) sowie 6 (= nicht zutreffend) und variieren je nach Sektion des Fragebogens (*Ausnahme: In der Subskala „Soziale Aktivitäten“ befindet sich ein Item mit einer qualitativen 5-Punkte-Skala, wobei im Vergleich zu den weiteren Items die Antwortoption „nicht zutreffend“ entfällt. Dasselbe gilt auch für die Items zur Lebensqualität und zum allgemeinen Gesundheitszustand). Bezugszeitraum sind die vergangenen 4 Wochen. Zur Auswertung erfolgte eine lineare Transformation der Items auf eine 100-Punkte-Skala (1 = 0; 2 = 25; 3 = 50; 4 = 75; 5 = 100), sodass höhere Werte in den Einzelitems oder Skalen für eine bessere Funktionalität stehen. Mit „nicht zutreffend“ bewertete Items werden als fehlend gewertet. Aus allen Items einer Subskala werden Summenwerte über die Mittelwerte gebildet. Der Gesamtwert der Lebensqualität resultiert aus den berechneten Mittelwerten der Subskalen und liegt ebenfalls zwischen 0 und 100 Punkten.

Werte für die Subskalen wurden nur gebildet, wenn mindestens die Hälfte aller Items der jeweiligen Subskala beantwortet wurden. Antworten der Kategorie „nicht zutreffend“ werden wie fehlende Werte gewertet. Ein Gesamtwert wird nur gebildet, sofern weniger als die Hälfte aller Subskalen als fehlend gelten. Die Daten sollten gemäß Handbuch ausgewertet werden.

In Studie GWEP1521 sollte der Fragebogen ausschließlich bei Patientinnen und Patienten im Alter von 2 bis 18 Jahren eingesetzt werden. Die primäre Analyse erfolgte jedoch für alle ausgefüllten Fragebögen, unabhängig vom Alter der Patientinnen und Patienten. Zusätzlich werden Sensitivitätsanalysen für die Altersgruppe der 2- bis 18-Jährigen dargestellt.

- Erhebungszeitpunkte: Zu Behandlungsbeginn (Studienvsiste 3) und zu Behandlungsende (Studienvsiste 10).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

Die Operationalisierung des QOLCE ist nachvollziehbar. Die primär geplante Analyse erfolgte unter Einbezug aller Patientinnen und Patienten, unabhängig von ihrem Alter, für die ein beantworteter Fragebogen vorlag. Da der Fragebogen maßgeblich für Personen im Alter von 4 bis 18 Jahren entwickelt und validiert wurde, ist fraglich, inwiefern auch Kinder unter 4 Jahren und Erwachsene über 18 Jahre mit diesem Instrument adäquat bewertet werden können.

Patientenrelevanz

Der Endpunkt Lebensqualität mittels QOLCE wird als patientenrelevant erachtet.

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

Der QOLCE ist ein Instrument zur Messung der gesundheitsbezogenen Lebensqualität bei Kindern und Jugendlichen zwischen 4 und 18 Jahren mit Epilepsie. Der Fragebogen besteht aus 77 Items in 5 Domänen und 16 Subskalen. Die Entwicklung des Fragebogens beinhaltete sowohl Literaturrecherchen als auch Befragungen von betroffenen Kindern bzw. deren Eltern und medizinischem Personal. In einer Querschnittsstudie mit 63 Eltern von Kindern mit refraktärer Epilepsie wurden anschließend die psychometrischen Gütekriterien des QOLCE untersucht [30]. Infolge von Item-Korrelationsanalysen der jeweiligen Subskalen reduzierte sich die Item-Anzahl von ursprünglich 91 auf 77. Reliabilität konnte ebenfalls über gute Werte bei der internen Konsistenz gemessen am Cronbach's alpha (Subskalen und Gesamtwert 0,72–0,93) und hohe Korrelationen bei der Test-Retest-Reliabilität in einer weiteren Untersuchung gezeigt werden [8]. Parallel zum QOLCE wurden in der Validierungsstudie 2 generische Lebensqualitätsfragebögen bei Kindern eingesetzt, der „Child Health Questionnaire“ (CHQ) und die CBCL. Die Korrelationen ausgewählter Subskalen der generischen Referenzinstrumente mit Subskalen des QOLCE in der konvergenten und diskriminanten Konstruktvalidität bewegten sich in den zu erwartenden Größenordnungen. Zudem zeigte sich ein signifikanter Zusammenhang zwischen der berichteten Anfallsschwere und dem Gesamtwert des QOLCE sowie fast allen Subskalen. Insgesamt scheinen die Items des Fragebogens Aspekte der Lebensqualität von Kindern mit Epilepsie gut abzubilden. In einer weiteren Studie an therapierefraktären Epilepsiepatientinnen und -patienten mit geistiger Beeinträchtigung gemessen am IQ zeigten sich im Sinne der konvergenten Validität moderat signifikante Korrelationen bei den kognitiven und sozialen Subskalen sowie dem Gesamtwert [29].

Eine weitere Untersuchung fokussierte auf die gesundheitsbezogene Lebensqualität bei Patientinnen und Patienten mit unterschiedlichen Epilepsie-Syndromen, jedoch ohne Berücksichtigung der TSC. In allen Subskalen des QOLCE sowie im Gesamtwert zeigten sich signifikant niedrigere Werte bei den symptomatischen Syndromen im Vergleich zu den idiopathischen Syndromen [32]. Die Fähigkeit des Instruments Veränderungen zu detektieren wurde lediglich (longitudinal) an einer Kohorte von Kindern mit Epilepsie untersucht, die eine Epilepsie-Chirurgie erhielten. Kindern, die nach Intervention anfallsfrei waren, wurden signifikant höhere und damit verbesserte Werte für fast alle Subskalen sowie den Gesamtwert attestiert, verglichen mit Kindern, die weiterhin Anfälle erlitten [31]. Neben der dargelegten australischen Validierung des QOLCE liegt auch eine Validierungsstudie für eine US-amerikanische Patientenpopulation vor [32]. Nach der Übersetzung des ursprünglichen Fragebogenentwurfs mit 91 Items und einer methodisch vergleichbaren, separaten Validierungsstudie resultierte mit dem USQOLCE eine amerikanische Version des Fragebogens mit 79 Items und Veränderungen in den Subskalen „Angst“, „Interaktionen“ und „Verhalten“. Die Unterschiede zwischen den Fragebögen werden von der Autorengruppe auf interkulturelle Unterschiede zurückgeführt.

Insgesamt zeigte der QOLCE in den mittels orientierender Literaturrecherche identifizierten Studien ausreichende bis gute Daten für die Reliabilität und Validität bei Kindern mit Epilepsie. Eine weitere durchgeführte Validierung an Patientinnen und Patienten im Indikationsgebiet TSC im Alter

von 2 bis 11 Jahren, anhand post hoc ausgewerteter Daten einer Phase-III-Studie, konnte die Ergebnisse weitgehend bestätigen [12]. Wie viele Kinder unter 4 Jahren in der Population enthalten waren, ist nicht beschreiben. Die Eignung des Fragebogens zur validen Bewertung der Lebensqualität von Personen unter 4 Jahren ist weiterhin unklar. Daher ist auch kritisch zu sehen, dass der QOLCE auch bei Patientinnen und Patienten ab einem Jahr eingesetzt wurde, obwohl das Instrument erst für Kinder ab 4 Jahren entwickelt und validiert ist. Daten dazu, wie viele Testpersonen in der Studie GWEP1521 jünger als 4 Jahre waren, liegen nicht vor. In der Altersgruppe der 1- bis 6-Jährigen waren 29 % der Testpersonen des Interventionsarms und 28% des Kontrollarms enthalten. Aufgrund des engen Patienten-Betreuer-Verhältnisses, von dem im hier vorliegenden Anwendungsgebiet ausgegangen werden kann, erscheint eine Fremdbeurteilung der Lebensqualität plausibel und adäquat. Keine Informationen liegen ebenfalls in Bezug auf die klinische Relevanz (MID) von Veränderungen vor.

Zur Auswertung des Fragebogens werden Skalenwerte über den Mittelwert der jeweiligen Skalenitems sowie ein Gesamtwert über den ungewichteten Mittelwert aller Skalen berichtet. Für die Nutzenbewertung werden sowohl die Ergebnisse für die einzelnen Subskalen als auch der Gesamtwert berücksichtigt [32].

Lebensqualität bei Epilepsie (mittels QOLIE-31-P)

Der Endpunkt „Lebensqualität bei Epilepsie“ wird in der Nutzenbewertung aufgrund fehlender Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Der „Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version“ (QOLIE-31-P) ist ein krankheitsspezifischer Fragebogen zur Erfassung der gesundheitsbezogenen Lebensqualität bei erwachsenen Patientinnen und Patienten ab 18 Jahren mit Epilepsie. Er besteht aus insgesamt 38 Items und ist patientenberichtet. Die Items lassen sich in folgende 7 Subskalen sowie eine Abfrage zum Gesundheitsstatus unterteilen:

- Energie/Müdigkeit (5 Items)
- Stimmung (6 Items)
- Tägliche Aktivitäten (6 Items)
- Kognition (7 Items)
- Medikationseffekte (4 Items)
- Angst vor Anfällen (6 Items)
- Allgemeine Lebensqualität (3 Items)
- Gesundheitsstatus (1 Item)

Ein Item aus jeder Subskala, abgesehen von „Gesundheitsstatus“, fordert die Befragten auf, die Belastung der jeweiligen Subskala in Bezug auf ihre Gesamtsituation zu bewerten.

In Abhängigkeit vom jeweiligen Item befanden sich die Antwortoptionen auf einer 4-, 5- oder 6-Punkte-Likert-Skala. Zur Auswertung erfolgte eine lineare Transformation der Items auf eine Skala von 0 bis 100 Punkten gemäß Scoring Manual, wobei höhere Werte einer besseren Lebensqualität entsprechen. Für alle Items einer Subskala wurde ein Mittelwert gebildet und dieser mit dem Wert des Items „Belastung“ gewichtet. Über alle gewichteten Skalenwerte wurde zusätzlich ein Gesamtwert gebildet.

Werte für die Skalen des QOLIE-31-P sollten nur gebildet werden, wenn mindestens die Hälfte aller Items der jeweiligen Subskala beantwortet wurde. Die Daten sollten gemäß Handbuch evaluiert werden.

Die Antwortoption für die Abfrage zum Gesundheitszustand entspricht einer visuellen Analogskala mit 100 Punkten.

Sollte eine Person nicht in der Lage sein den Fragebogen selbstständig auszufüllen, bestand die Möglichkeit, dass die betreuende Person unterstützt.

In Studie GWEP1521 sollte der Fragebogen ausschließlich bei Patientinnen und Patienten ab 19 Jahren eingesetzt werden. Die primäre Analyse erfolgte jedoch mit allen Personen, die den Fragebogen tatsächlich ausgefüllt haben, unabhängig vom Alter. Zusätzlich werden Sensitivitätsanalysen für die Altersgruppe ab 19 Jahren dargestellt.

- Erhebungszeitpunkte: Zu Behandlungsbeginn (Studienvisite 3) und zu Behandlungsende (Studienvisite 10).
- Auswertung: Auf Basis der ITT-Population.

Bewertung

In Modul 4 des Nutzendossiers und den Studienunterlagen finden sich teils widersprüchliche Angaben zur Operationalisierung des QOLIE-31-P. Im Studienprotokoll ist angegeben, dass der Fragebogen aus 38 Items zu Gesundheitszustand und täglichen Aktivitäten bei Personen mit Epilepsie besteht. Gemäß Angaben in Modul 4 umfasst der Fragebogen nur 31 Items, was der Anzahl des QOLIE-31 entsprechen würde. Es kann nicht nachvollzogen werden, welche Items letztendlich in die in Modul 4 dargestellten Analysen des QOLIE-31-P eingingen. Gemäß SAP sollte der Fragebogen nach dem Scoring Manual der QOLIE-Gruppe für den QOLIE-31-P ausgewertet werden. Das Item zum Gesundheitsstatus auf der visuellen Analogskala scheint allgemein keine Berücksichtigung gefunden zu haben.

Die betreuende Person konnte die Patientin / den Patienten bei der Beantwortung der Fragen unterstützen, wobei unklar bleibt, wie die Unterstützung definiert war und ob die betreuende Person selbst die Fragen beantworten durfte.

Patientenrelevanz

Der Endpunkt „Lebensqualität bei Epilepsie“ (mittels QOLIE-31-P) wird als patientenrelevant erachtet.

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

Der QOLIE-31-P ist eine Abwandlung des krankheitsspezifischen Epilepsie-Fragebogens QOLIE 31. Im Vergleich zur ursprünglichen Version wurde jeder der Subskalen, abgesehen von der Subskala „Gesundheitsstatus“, ein Item hinzugefügt, welches die Befragten auffordert die Belastung der jeweiligen Subskala bezüglich der Gesamtsituation zu bewerten. Dieses Item diente der Berechnung eines gewichteten Skalenwerts und auch eines gewichteten Gesamtwerts für den QOLIE-31-P.

Für den QOLIE-31-P konnten keine Informationen in Bezug auf die Validität identifiziert werden, weder für Patientinnen und Patienten mit Epilepsie noch mit TSC. Eine vom pU zitierte Validierungsstudie betrachtet zwar den QOLIE-31-P in der Indikation TSC, allerdings gingen in die Auswertungen lediglich die ungewichteten Subskalen und der ungewichtete Gesamtwert ein [12]. Für die gewichteten Auswertungen liegen weiterhin keine Daten vor. Aufgrund der deutlichen Veränderungen im gewichteten QOLIE-31-P im Vergleich zum QOLIE-31 erscheint eine Übertragung der Evidenz aus Validierungsstudien zum QOLIE-31 nicht sinnvoll [25].

2.3.4 Sicherheit

Unerwünschte Ereignisse

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

Operationalisierung

Beschreibung

Unerwünschte Ereignisse (UE) waren definiert als alle neu auftretenden unerwünschten oder unbeabsichtigten Ereignisse oder Symptome (inklusive abnormaler Laborwerte), Diagnosen oder Verschlechterung bestehender Beschwerden, die zwischen der Screening-Visite (Studienvisite 1) und der Sicherheits-Follow-up-Visite auftraten (Studienvisite 12), mit oder ohne Bezug zur Studienmedikation. Patientenindividuell zu erwartende Anfallstypen sollten nicht als UE erfasst werden. Jegliche Verschlechterung inklusive Veränderung im Anfallsmuster oder in der Anfallschwere wurde als UE dokumentiert. Elektive Hospitalisierungen für bereits bestehende Beschwerden oder elektive Prozeduren wurden ebenfalls nicht als UE definiert. Ausgenommen davon waren UE, die während der Hospitalisierung zu einer Verlängerung des Krankenhausaufenthalts führten; diese wurden als schwerwiegende unerwünschte Ereignisse (SUE) definiert. Bei einem tödlichen Ereignis sollte die zugrundeliegende Ursache als SUE berichtet werden. Alle berichteten UE wurden mit dem Medical Dictionary for Regulatory Activities (MedDRA) (Version 19.1) kodiert.

Ein TEAE (Treatment Emergent Adverse Event) war definiert als UE, das erstmalig im verblindeten Studienzeitraum zwischen dem Tag der Verabreichung der 1. Dosis der Studienmedikation und dem Tag der 1. Dosis der Studienmedikation in der Titrationsperiode der OLE auftrat.

Als SUE wurden für beide Studien UE definiert, die

- zum Tode führten,
- lebensgefährdend waren,
- zur Hospitalisierung bzw. Verlängerung einer Hospitalisierung führten,
- zu einer anhaltenden oder schwerwiegenden Behinderung/Beeinträchtigung führten,
- zu einer kongenitalen Anomalie / angeborenen Fehlbildung führten,
- als medizinisch bedeutsam durch das ärztliche Prüfpersonal angesehen wurden.

Wichtige medizinische Ereignisse müssen nicht unmittelbar lebensbedrohlich oder tödlich sein oder zu einer Hospitalisierung führen, können aber dazu führen, dass Maßnahmen unternommen werden, um die oben genannten Ereignisse zu vermeiden. Diese sollten auch regelhaft als schwerwiegend eingestuft werden. Beispiele können sein: Intensive Behandlung in der Notaufnahme oder zu Hause auf Grund eines allergischen Bronchospasmus; Blutdyskrasie oder Anfälle, die nicht zu einer Hospitalisierung führten; Entwicklung einer Drogenabhängigkeit bzw. Drogenmissbrauch. Der Sponsor stuft alle Status-epilepticus-Ereignisse als medizinisch signifikantes SUE ein.

Zusätzlich wurden potentielle Fälle von durch Medikamente induzierte Leberschäden untersucht. Sofern a priori definierte Kriterien (erhöhte Laborparameter und ggf. Symptome) zutrafen, waren die Bedingungen für ein Ausscheiden aus der Studie erfüllt und sie wurden als medizinisch bedeutende Fälle angesehen.

UE von besonderem Interesse (gemäß SAP)

- Drogenentzug
- Drogenmissbrauch und Abhängigkeit

Schweregrad von UE

Für die Beschreibung des Schweregrads eines UE sollten die Terme „mild“, „moderat“ oder „schwer“ verwendet und stets der höchste Schweregrad eines UE dokumentiert werden. Die Beurteilung des Schweregrads sollte nach klinischem Urteilsvermögen vorgenommen werden.

UE, die zum Therapieabbruch führen: Es liegt im Ermessen des Prüfpersonals zu entscheiden, ob ein UE ausreichend schwer ist, um den Therapieabbruch der Testperson zu rechtfertigen. Zudem kann eine Testperson freiwillig die Therapie abbrechen, sofern sie ein aus ihrer Sicht intolerables UE erlebt.

- Erhebungszeitpunkte: Ab Screening (Studienvisite 1) über den gesamten Studienzeitraum bis zur Sicherheits-Follow-up-Visite (Studienvisite 12 oder OLE). Als TEAE werden alle UE ab Behandlungsbeginn über die gesamte verblindete Studienperiode bis zum Erhalt der ersten OLE-Dosis gewertet.
- Auswertung: Auf Basis der Sicherheitspopulation.

Bewertung

Die Operationalisierung ist überwiegend nachvollziehbar. TEAE waren definiert als UE, die erstmalig zwischen dem Tag der Verabreichung der 1. Dosis der Studienmedikation und dem Tag der 1. Dosis der Studienmedikation in der Titrationsperiode der OLE auftraten. Nicht beschrieben ist, bis zu welchem Zeitpunkt bzw. welcher Studienvisite UE bei Personen, die die Studie abbrachen oder nicht in die OLE wechselten, als TEAE gewertet wurden.

Bei der Schweregradeinteilung der UE wurde in beiden Studien keine Definition der Terme vorgenommen, sondern diese lag im Ermessen des ärztlichen Prüfpersonals. Es bleibt unklar, inwieweit eine einheitliche Schweregradeinteilung vor dem Hintergrund einer multizentrischen Studie vorgenommen werden konnte.

Drogenentzug sowie Drogenmissbrauch und Abhängigkeit waren im SAP als UE von besonderem Interesse definiert. Darüber hinaus waren im Studienbericht mit „Selbstverletzende Gedanken“, „Euphorie“ und „Schwangerschaft“ weitere, nicht präspezifizierte UE von besonderem Interesse definiert. Gründe für dieses Vorgehen konnten nicht identifiziert werden.

In der Studie wurden als UE Verschlechterung von Anfällen und Änderung im Muster von Anfällen definiert. Diese wurden durch das Data Safety Monitoring Committee im Rahmen der Untersuchung der Sicherheitsdaten evaluiert. Unklar bleibt, ob hierfür Systemorganklassen und Preferred Terms bzw. alternative Erhebungsformen verwendet wurden.

Patientenrelevanz

UE werden als patientenrelevant erachtet.

Entsprechend § 2 Satz 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

Eine Schweregradeinteilung war in beiden Studien vorgesehen. Es wurde keine Definition für die Terme „mild“, „moderat“ und „schwer“ a priori festgelegt. Die Beurteilung sollte ausschließlich nach klinischem Ermessen erfolgen. Dies wird kritisch gesehen, da eine subjektive Beurteilung der UE eine einheitliche Bewertung erschwert. Aus diesem Grund wird von der Darstellung der UE in Abhängigkeit der Schweregradeinteilung abgesehen.

Suizidalität (mittels C-SSRS)

Der Endpunkt „Suizidalität“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Die „Columbia-Suicide Severity Rating Scale“ (C-SSRS) ist ein standardisiertes klinisches Interview zur systematischen Bewertung und strengen Überwachung suizidaler UE während der gesamten Studie sowie anderer Situationen. Erfasst wird das Ausmaß an suizidalen Gedanken und welche Maßnahmen ergriffen wurden (einschließlich Vorbereitung) sich das Leben zu nehmen (Selbstmordverhalten). Die C-SSRS wurde vom ärztlichen Prüfpersonal ausgefüllt oder einer entsprechend qualifizierten Vertretung (= Jede Person, die innerhalb der vergangenen 2 Jahre ein C-SSRS-Training absolviert hat oder die C-SSRS kontinuierlich in der Studie seit der Teilnahme am Training eingesetzt hat). Die Erhebung sollte, wenn möglich, während der gesamten Studie von derselben Person durchgeführt werden. Die Befragung erfolgte nur für Personen ab 6 Jahren; zudem mussten die Befragten nach Ansicht des ärztlichen Prüfpersonals in der Lage sein die Fragen zu verstehen und zu beantworten. Falls dies nicht der Fall war, konnte die Suizidalität durch das klinische Prüfpersonal auf Basis einer Befragung der Testperson eingeschätzt werden.

Das Instrument existiert in verschiedenen Versionen und umfasst je nach Version 13 bis 18 Items. In den Studien kam eine C-SSRS bei Patientinnen und Patienten im Alter von ≥ 19 Jahren zum Einsatz und eine „C-SSRS Children's“ bei allen 6- bis 18-Jährigen, für die jedoch nochmals verschiedene Versionen existieren [7]; aus den Unterlagen geht die eingesetzte Version nicht eindeutig hervor. Ferner liegen keine Angaben dazu vor, wie sich die C-SSRS für Kinder von der Version für Erwachsene unterscheidet.

Der Bezugszeitraum für die Baseline-Befragung umfasst das bisherige Leben, bei allen weiteren Studienvisiten bezieht sich die Befragung auf die Zeit zwischen der letzten und der aktuellen Studienvisite.

Der Fragebogen besteht aus den folgenden Kategorien:

Suizidgedanken

- Ich wünschte, ich wäre tot.
- Unspezifische aktive Suizidgedanken.
- Aktive Suizidgedanken mit allen Möglichkeiten (kein Plan) ohne Handlungsabsicht.
- Aktive Suizidgedanken mit einer gewissen Handlungsabsicht, ohne konkreten Plan.
- Aktive Suizidgedanken mit konkretem Plan und Absicht.

Intensität der Gedanken: Angabe des schlimmsten Gedankens und wie häufig diese Gedanken auftraten.

Suizidales Verhalten (Ja-/Nein-Antworten; teilweise Abfrage zur Anzahl der Versuche)

- Aktueller Versuch.
- Unterbrochener Versuch.
- Gescheiterter Versuch.
- Vorbereitendes Handeln oder Verhalten.
- Suizidales Verhalten; in den Studienvisiten nach Baseline wird ein weiteres Item abgefragt.
- Vollendeter Suizid.

Die Inzidenz der Personen mit jeglichen Suizidgedanken und jeglichem suizidalem Verhalten während der Behandlungsperiode je Behandlungsarm wird deskriptiv dargestellt.

Wenn die C-SSRS bei Screening oder Baseline nicht beantwortet wurde, wird die Patientin / der Patient nicht für die Auswertungen „Auftreten oder Verschlechterung von Suizidgedanken oder -verhalten“ berücksichtigt.

- Erhebungszeitpunkte: Während Studienvisiten 1 (Screening), 3, 4, 5, 6, 7, 9, 10, 11.
- Auswertung: Auf Basis der Sicherheitspopulation.

Bewertung

Die Operationalisierung ist weitgehend nachvollziehbar.

Die C-SSRS ist ein standardisiertes Instrument zur Bewertung von Selbstmordverhalten und -vorstellungen; sie wird u. a. zur strengen Überwachung der Suizidalität in klinischen Studien eingesetzt.

Bewertet werden die Kategorien jegliche Suizidalität sowie Suizidgedanken und suizidales Verhalten; vollständige Suizidalität, Auftreten von Suizidgedanken, Verschlechterung der Suizidgedanken und Auftreten von suizidalem Verhalten. Darüber hinaus erfolgte eine Auswertung für die Kategorien „Suizidgedanken“ und „Suizidales Verhalten“ pro abgefragtem Item. Angaben zur Bildung von Summenwerten finden sich weder im Nutzendossier des pU noch in den Studienunterlagen.

Patientenrelevanz

Suizidalität und selbstverletzendes Verhalten sind patientenrelevante Endpunkte.

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

Validität und Reliabilität der C-SSRS wurden in 3 multizentrischen Studien untersucht [28]: Jugendliche / Junge Erwachsene nach Suizidversuch (N = 124); Jugendliche / Junge Erwachsene mit Depressionen (N = 312); Erwachsenen Personen, die wegen psychischer Probleme eine Notaufnahme aufsuchten (N = 237). Die C-SSRS zeigt dabei gute divergente und konvergente Validität, interne Reliabilität und Veränderungssensitivität.

Untersuchungen an Kindern wurden nicht identifiziert, sodass eine Übertragbarkeit für einen Teil der vorliegenden Population unsicher ist. Es erscheint fraglich, wie sachgerecht bzw. aussagekräftig eine Einschätzung von Selbstmordverhalten und -gedanken bei kognitiven Beeinträchtigungen erfolgen kann. Da der Endpunkt nur für Personen ab 6 Jahren und für diejenigen, die in der Lage waren die Fragen zu verstehen und zu beantworten, erhoben wurde, ist eine Selektion der eingeschlossenen Population durch das ärztliche Prüfpersonal vorgenommen worden. Die Endpunkterhebung wird insgesamt kritisch gesehen.

2.3.5 Übersicht der Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 8.

Tabelle 8: Erhebungszeitpunkte der berücksichtigten Endpunkte in Studie GWEP1521

Studienvisite (Tag)	1¹⁾ (-35)	2²⁾ (-28)	3³⁾ (1)	4 (15)	5 (29)	6 (43)	7 (57)	8 (71)	9 (85)	10⁴⁾ (113)	11⁵⁾ (123)	12⁶⁾ (151)	Safety Calls⁷⁾
Endpunkt													
Todesfälle ⁸⁾	X	X	X	X	X	X	X	X	X	X	X	X	X
Epileptische Anfälle ⁹⁾		täglich											
Status epilepticus ⁹⁾		täglich											
CGIC / SGIC			X ¹⁰⁾							X			
Epilepsiebedingte Hospitalisierung		X	X	X	X	X	X	X	X	X	X	X	X
Vineland-II			X							X			
Achenbach-Verhaltens-Checkliste			X							X			
QOLCE			X							X			
Unerwünschte Ereignisse	X	X	X	X	X	X	X	X	X	X	X	X	X
C-SSRS	X		X	X	X	X	X		X	X	X		

¹⁾ Studienvisite 1 entspricht der Screening-Visite.

²⁾ Studienvisite 2 entspricht der Baseline-Visite.

³⁾ Studienvisite 3 entspricht der Randomisierung und dem Behandlungsbeginn.

⁴⁾ Studienvisite 10 entspricht der Visite am Ende der Behandlung an Tag 113 oder früher bei einem Studienabbruch der Testperson. Patientinnen / Patienten konnten in eine OLE übertreten. Personen, die die Behandlung beendeten, schlichen die Studienmedikation über 10 Tage aus.

⁵⁾ Studienvisite 11: Für Testpersonen, die nicht an der OLE teilgenommen haben oder die die Therapie abgebrochen haben und die Studienmedikation ausschleichen.

⁶⁾ Sicherheits-Follow-up-Visite für alle Testpersonen, die nicht an der OLE teilgenommen haben oder die die Therapie abgebrochen haben.

⁷⁾ Sicherheitsbezogene Telefonanrufe erfolgen alle 2 Tage während der anfänglichen Titrationsperiode, eine Woche nach der Titration und für Testpersonen, die nicht in die OLE übergehen, wöchentlich von Studienvisite 10 bis Studienvisite 12.

⁸⁾ Im Rahmen der Sicherheitserhebung erfasst.

⁹⁾ Epileptische Anfälle und dementsprechend auch Status epilepticus wurden täglich mittels IVRS erhoben bis zum Ende des Ausschleichens der Studienmedikation (Studienvisite 9).

¹⁰⁾ Entspricht nicht einer eigentlichen Erhebung des CGIC / SGIC, sondern einer schriftlichen Erfassung des Gesundheitszustands als Gedankenstütze für spätere Erhebungen.

Abkürzungen: CGIC: Caregiver Global Impression of Change; C-SSRS: Columbia-Suicide Severity Rating Scale; IVRS: Interactive Voice Response System; OLE: Offene Extensionsstudie; QOLCE: Quality of Life in Childhood Epilepsy; SGIC: Subject Global Impression of Change; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

2.4 Statistische Methoden

Analysepopulationen

Intention-to-Treat (ITT): Beinhaltet alle Personen, die auf die Studienarme randomisiert wurden und stellt die primäre Analysepopulation für alle Wirksamkeitsanalysen dar. Patientendaten wurden nach Behandlungsgruppe ausgewertet, in die sie randomisiert wurden.

Sicherheitspopulation: Umfasst alle auf die Studienarme randomisierten Personen, die mindestens eine Dosis der Studienbehandlung erhalten haben. Patientendaten wurden nach erhaltener Behandlung ausgewertet.

Statistische Analysen

Alle statistischen Auswertungen entsprechen 2-seitigen Signifikanztests auf dem 5%-Signifikanzniveau. Die Nullhypothese wird bei einem Unterschied zwischen den Behandlungsgruppen auf einem Alpha-Level von 0,05 verworfen.

Bei Endpunkten, die mittels parametrischer Verfahren analysiert werden, wird, falls angemessen, die Annahme der Normalverteilung und Homogenität der Varianz mittels Residualplots und dem Shapiro-Wilk-Test auf Normalverteilung überprüft. Falls die Annahme der Normalverteilung verletzt ist, werden nicht-parametrische Tests durchgeführt und die parametrischen Tests als Sensitivitätsanalysen dargestellt.

Primärer Endpunkt (Häufigkeit TSC-assoziierter Anfälle)

Primäre Nullhypothese: Nach einer Behandlungsdauer von 16 Wochen zeigt sich kein Unterschied in der Veränderung der Häufigkeit TSC-assoziierter Anfälle im Vergleich zu Baseline zwischen den Studienarmen 25 mg/kg/Tag Cannabidiol und Placebo.

Die primäre Analyse erfolgt mittels negativer Binomialregression für die Summe der TSC-assozierten Anfälle während der Baseline- und der Behandlungsperiode als Responsevariablen. Im Rahmen eines allgemein linearen, negativ-binomialen Modells wurde ein Mixed-Effekt-Modell mit wiederholten Messungen (MMRM) für alle während der Baseline- und der Behandlungsperiode beobachteten TSC-assozierten Anfälle durchgeführt. Das Modell beinhaltet die Stratifizierung nach Altersgruppe (2–5, 6–12, 13–18 Jahre), Zeitpunkt, Behandlungsarm und Interaktion zwischen Behandlungsarm und Zeitpunkt als fixe Effekte und Patientin / Patient als Zufallseffekt. Log-transformierte Anzahl von Tagen mit berichteten Anfällen wurde als Offset-Variable berücksichtigt. Die Zeitvariable entspricht einem Indikator für die Baseline- und die Behandlungsperiode.

Die TSC-assozierte Anfallshäufigkeit (28-Tage-Durchschnitt) sowie die prozentuale Veränderung der Anfallshäufigkeit werden gegenüber Baseline dargestellt. Die prozentuale Veränderung gegenüber dem Ausgangswert konvulsiver Anfallshäufigkeit berechnet sich folgendermaßen:

$$\left(\frac{\text{Häufigkeit während Behandlungsperiode}}{\text{Häufigkeit während Baselineperiode}} - 1 \right) \times 100$$

Die Anfallshäufigkeit für jeden Zeitraum basierte dabei auf einem 28-Tage-Durchschnitt:

$$\frac{\text{Anzahl Anfälle in der Periode}}{\text{Anzahl berichteter Tage für IVRS in der Periode}} \times 28$$

Für den primären Endpunkt waren darüber hinaus eine Vielzahl von Sensitivitätsanalysen geplant, u. a.:

- Für das Per-Protokoll-Analyseset.
- Wilcoxon-Rangsummentest für prozentuale Veränderung der Häufigkeit von TSC-assoziierten Anfällen während der Erhaltungsperiode gegenüber Baseline unter Verwendung von Hodges-Lehmann.
- Rang-basierte Kovarianzanalyse (ANCOVA) für die prozentuale Veränderung von Baseline bei TSC-assoziiierter Anfallshäufigkeit während der Behandlungsperiode.
- ANCOVA mit log-transformierter TSC-assoziiierter Anfallshäufigkeit während der Behandlungsperiode.
- ANCOVA mit Baseline und Stratifizierungsfaktor Altersgruppe als Kovariate und Behandlungsgruppe als fixen Faktor.
- Wiederholung der primären Analyse für die Erhaltungsperiode (nur Testpersonen mit mind. 7 Anfallstagen).
- Wiederholung der primären Analyse für die Titrationsperiode und alle 4 Wochen während der Erhaltungsperiode.
- Wiederholung der primären Analyse, wobei fehlende Werte im Behandlungszeitraum mit dem schlechtesten Wert (höchste Anzahl an Sturzanfällen) folgender Imputationsmethoden ersetzt werden: „Last Observation Carried Forward“ (LOCF), „Next Observation Carried Backward“ (NOCB) oder der durchschnittlichen Anfallszahl aufgezeichneter Tage.
- Wiederholung der primären Analyse mit multiplen Imputationen, um Daten unter der Annahme „Missing not at Random“ (MNAR) zu imputieren.

Sekundäre Endpunkte

Epileptische Anfälle

Veränderungen in der Gesamtzahl der Anfälle, der Anfälle gesamt, die Anzahl weiterer Anfälle und des fokalen Anfallswerts wurden analog zum primären Endpunkt ausgewertet.

Die Analyse der Responder erfolgte mittels Cochran-Mantel-Haenszel (CMH)-Test stratifiziert nach Altersgruppe. Für die Responderauswertungen aller Anfälle präsentiert der pU mit Modul 4 zudem post hoc ein Relatives Risiko (RR). Das zugrundeliegende Modell ist nicht beschrieben.

Status-epilepticus-Episoden

Gemäß SAP war lediglich eine deskriptive Darstellung der Personen mit Status-epilepticus-Episoden für die Baseline- und die Behandlungsperiode vorgesehen. In Modul 4 wurden post hoc die Testpersonen mit Status-epilepticus-Episoden im Behandlungszeitraum mittels RR und CMH-Test ausgewertet. Es liegen keine Angaben zur Verwendung von Stratifikationsfaktoren sowie dem zugrundeliegenden Modell vor.

CGIC / SGIC

Die Auswertung erfolgte gemäß Studienprotokoll am Ende der Behandlungsperiode über eine ordinal logistische Regression mittels „Proportional Odds Model“. Falls die Annahme der „Proportional Odds“ nicht erfüllt sind, werden die Werte zusätzlich mittels „Cochran-Armitage Trend Test“ analysiert. Für das Modul 4 wurden post hoc zusätzlich Analysen mittels ANCOVA, mit Behandlung als festem Effekt, sowie Responderanalysen mittels CMH-Test durchgeführt. Für die Responderauswertungen präsentiert der pU zudem post hoc ein RR. Das dem RR zugrundeliegende Modell ist nicht beschrieben. Aufgrund der ordinal skalierten Daten und dem Vorliegen einer ordinal logistischen Regression wird die Analyse mittels ANCOVA nicht berücksichtigt.

Epilepsiebedingte Hospitalisierungen

Gemäß SAP war lediglich eine deskriptive Darstellung der Personen mit epilepsiebedingter Hospitalisierung vorgesehen. In Modul 4 wurden post hoc die Testpersonen mit epilepsiebedingten Hospitalisierungen im Behandlungszeitraum mittels RR und CMH-Test stratifiziert nach Altersgruppen ausgewertet. Das dem RR zugrundeliegende Modell ist nicht beschrieben.

Vineland-II, Achenbach-Verhaltens-Checkliste, QOLCE

Gemäß SAP sollte die Veränderung zu Baseline mittels ANCOVA analysiert werden. Als Kovariaten dienten Baseline-Wert und stratifizierte Altersgruppe, Behandlungsgruppe war der fixe Faktor. In Modul 4 wurde zusätzlich die standardisierte Mittelwertdifferenz nach Hedges' g mit 95%-KI angegeben.

Für Modul 4 beschreibt der pU eine erneute Analyse der Endpunkte „Vineland-II“, „Achenbach-Verhaltens-Checklisten“ und „QOLCE“, in der die Testpersonen des Hochdosis-Studienarms nicht berücksichtigt wurden, wobei auch eine Analyse der Veränderung der einzelnen Syndromskalen und der weiteren Subskalen erfolgte.

Unerwünschte Ereignisse

Für Studie GWEP1521 wurden post hoc in Modul 4 Odds Ratios (OR), RR und Risikodifferenz (RD) berichtet – jeweils mit 95%-KI und p-Wert, berechnet mittels CMH-Test stratifiziert nach Altersgruppen.

Umgang mit fehlenden Daten

Primärer Endpunkt

Für Personen, die die Studie vorzeitig abbrechen, wird die primäre Analyse auf Basis der bis zu dem Zeitpunkt des Studienabbruchs verfügbaren Daten durchgeführt.

Unerwünschte Ereignisse

Bei fehlenden oder unvollständigen Daten zum Auftreten eines UE wird das frühestmögliche Datum des Auftretens angenommen. Sofern keine Angaben für die Schwere eines UE vorliegen, wird ein „Worst-case“-Ansatz verfolgt und „schwer“ imputiert.

Subgruppenanalysen

Subgruppenanalysen wurden in der ITT-Population für den primären Endpunkt und den wichtigen sekundären Endpunkt „Reduktion der TSC-assoziierten Anfallshäufigkeit um mindestens 50 %“ für alle Subgruppen geplant, die in Tabelle 2 gelistet sind.

Für den primären Endpunkt wird eine negativ-binomiale Regressionsanalyse durchgeführt. Das Modell wird jeweils um Kovariaten für jeden zu untersuchenden Effekt ergänzt, sowohl individuell als auch für die Interaktion mit Zeitpunkt, Behandlungsarm und Interaktion zwischen Behandlungsarm und Zeitpunkt. Die jeweiligen Effekte werden in getrennten Modellen getestet.

Für den wichtigen sekundären Endpunkt werden Testpersonen mit einer Reduktion der Anfallsfrequenz $\geq 50\%$ mittels logistischer Regression modelliert, mit stratifizierter Altersgruppe und Behandlungsarm als Kovariaten. Das Modell enthält zusätzlich Kovariaten für jedes Level des untersuchten Effekts, individuell und mit Interaktion mit dem Behandlungsarm. Die jeweiligen Effekte werden in getrennten Modellen getestet.

Versionen des SAP

Die finale Version V2 des SAP der Studie GWEP1521 ist auf den 24. April 2019 datiert; die ursprüngliche Version des SAP vom 17. April 2019 ist lediglich eine Woche älter. Folgende relevante Änderung vor Entblindung sind im SAP V2 beschrieben: Aktualisierung der Hierarchie der Analysen für Konsistenz mit dem Protokoll Version 8.

Addendum zum SAP vom 28. November 2019

Post hoc nach Entblindung der Studie trat ein Addendum zum SAP in Kraft, in dem folgende zusätzliche Analysen beschrieben sind, um die klinische Wirksamkeit zu unterstützen:

- Darstellung weiterer kumulativer Verteilungsplots für die prozentuale Reduktion der TSC-assoziierten Anfallsfrequenz getrennt für die Subgruppe Clobazam-Einnahme (Ja/Nein).
- Überprüfung der Robustheit der Ergebnisse aufgrund vieler Therapieabbrüche in den Interventionsgruppen durch Post-hoc-Definition folgender Endpunkte:
 - o Zeit bis zum 10ten oder 20sten Anfall nach Behandlungsbeginn.
 - o Durchführung der primären Analyse, wobei die Anfallshäufigkeit von Testpersonen, die die Behandlung abbrachen, mit der Anfallshäufigkeit dieser Testpersonen während der Baseline-Periode imputiert wurde.
 - o Zur Erfassung des Einsetzens der Wirksamkeit wird der primäre Endpunkt sowie der wichtige sekundäre Endpunkt „Reduktion der Anfallshäufigkeit um $\geq 50\%$ “ an jedem Tag nach Behandlungsbeginn kumuliert ausgewertet.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 9: Verzerrungspotential der Studie GWEP1521

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
GWEP1521	Ja ¹⁾	Ja	Ja	Ja	Nein	Ja ²⁾	Niedrig

¹⁾ Das Randomisierungsschema wurde zentral durch einen unabhängigen Statistiker unter Verwendung von zufällig wechselnden Blöcken generiert.

²⁾ 10 Personen (13 %) der Interventionsgruppe brachen die Therapie ab verglichen mit 1 Person (1 %) der Kontrollgruppe.

Das Verzerrungspotential auf Studienebene wird als niedrig eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 10 dargestellt.

Tabelle 10: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie GWEP1521

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
Todesfälle	Ja	Ja	Nein	Nein	Niedrig
Morbidität					
Epileptische Anfälle					
TSC-assoziierte Anfälle	Ja	Ja	Nein ¹⁾	Ja ²⁾	Hoch
Anfälle gesamt	Ja	Ja	Nein ¹⁾	Ja ²⁾	Hoch
Weitere Anfälle	Ja	Ja	Nein ¹⁾	Ja ²⁾³⁾	Hoch
Status epilepticus	Ja	Ja	Nein	Ja ²⁾⁴⁾	Hoch
Globaler Eindruck der Veränderung (mittels CGIC / SGIC)	Ja	Ja ⁵⁾	Ja ⁶⁾	Ja ⁷⁾	Hoch
Bewertung des adaptiven Verhaltens (mittels Vineland-II)	Ja	Nein ⁸⁾	Nein	Nein	Hoch
Achenbach-Verhaltens-Checkliste	Ja	Ja	Nein	Nein	Niedrig

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Lebensqualität					
Lebensqualität bei kindlicher Epilepsie (mittels QOLCE)	Ja	Nein ⁹⁾	Nein	Ja ¹⁰⁾	Hoch
Sicherheit					
Unerwünschte Ereignisse	Ja	Ja	Nein	Nein	Niedrig
Suizidalität (mittels C-SSRS)	Ja	Nein ¹¹⁾	Nein	Nein	Unklar

¹⁾ Die Methodik der statistischen Analyse für die Anfallsdaten wurde erst nach Studienende, jedoch vor Entblindung, final festgelegt. Das ersetzte statistische Modell wurde als Sensitivitätsanalyse für den primären Endpunkt ausgewertet.

²⁾ Für Personen, die die Studie vorzeitig abbrachen, wurde die primäre Analyse auf Basis der bis zum Studienabbruch verfügbaren Daten durchgeführt. Therapieabbrüche waren zwischen den Studienarmen ungleich verteilt und traten bei 10 Personen (13%) des Cannabidiol- und 1 Person (1 %) des Kontrollarms auf. Eine Veränderung, insbesondere Steigerung der Anfallshäufigkeit nach Ende der Behandlung, kann nicht ausgeschlossen werden, weshalb von einer Verzerrung ausgegangen wird.

³⁾ Lediglich 12 Personen (16 %) der Cannabidiol- und 15 (20%) der Kontrollgruppe berichteten während der Baseline-Periode mindestens einen „weiteren Anfall“.

⁴⁾ Aufgrund der geringen Anzahl an Ereignissen wurde die im Studienprotokoll geplante Analyse nicht durchgeführt; der Endpunkt wurde im Studienbericht rein deskriptiv ausgewertet. In Modul 4 des Nutzendossiers wurde die Auswertung der Testpersonen mit mindestens einem Status epilepticus während der Behandlungsperiode dargestellt.

⁵⁾ In die Responderanalyse zum globalen Eindruck der Veränderung (kombinierte Auswertung von CGIC / SGIC) gehen ausschließlich die Testpersonen mit einem Rücklauf zum Behandlungsende ein; dies sind auf Seiten der Interventionsgruppe 93 % der ITT-Population und 100 % der Kontrollgruppe. Eine Analyse, welche die fehlenden Werte aus der Interventionsgruppe als Non-Responder wertet, liegt nicht vor. Die Auswirkungen dieses Aspekts auf das Verzerrungspotential werden jedoch als gering eingeschätzt, da es keine Hinweise auf eine deutliche Veränderung des Effekts gibt.

⁶⁾ Die Responderanalyse entspricht einem post hoc definierten Endpunkt. Es wurden ausschließlich Ergebnisse für eine Verbesserung dargestellt, nicht für Verschlechterungen. Eine separate Responderauswertung des CGIC liegt nicht vor.

⁷⁾ Responderanalysen wurden lediglich für eine kombinierte Auswertung des CGIC / SGIC vorgelegt. Fremd- und Selbstbewertung können sich strukturell unterscheiden. Für die Nutzenbewertung wird eine separate Auswertung für den CGIC als relevant erachtet. Bei der gemeinsamen Auswertung von CGIC / SGIC basierte die Erhebung lediglich für jeweils 4 Personen der Interventions- und der Kontrollgruppe auf dem SGIC.

⁸⁾ In die Berechnung der Domänenwerte sowie den Gesamtwert für das adaptive Verhalten gehen weniger als 70 % der ITT-Population ein.

⁹⁾ In der Studie GWEP1521 sollte der Fragebogen gemäß Protokoll ausschließlich bei Personen im Alter von 2 bis 18 Jahren eingesetzt werden. Entgegen dieser Vorgabe wurde der QOLCE auch für einigen Personen im Alter von ≥ 19 Jahren an Stelle des QOLIE-31-P beantwortet. Wie die Selektion dieser Testpersonen erfolgte und wie viele Testpersonen den QOLCE statt des QOLIE-31-P ausfüllten, ist unklar. In Modul 4 des Nutzendossiers präsentiert der pU Post-hoc-Auswertungen für alle Personen ≤ 18 Jahre. Die Rückläufe zwischen den einzelnen Subskalen sind zum Teil stark unterschiedlich ausgeprägt und liegen, bezogen auf die Studienpopulation 18 Jahre und jünger, lediglich bei 64 % unter Cannabidiol und 77 % unter Placebo.

¹⁰⁾ Es bestehen Diskrepanzen bezüglich der Testpersonen, die in die Analyse der jeweiligen Subskalen / den Gesamtwert eingehen zwischen den Daten aus dem Studienbericht und den post hoc analysierten Daten für Modul 4. Im CSR war eine Sensitivitätsanalyse des QOLCE mit allen Testpersonen im Alter von 2 bis 18 Jahren geplant. Für Modul 4 werden alle Testpersonen ≤ 18 Jahre in der Auswertung berücksichtigt. Es ist unklar, ob die Diskrepanzen darauf zurückzuführen sind.

¹¹⁾ Der Endpunkt ist nur für Personen ab 6 Jahren und für diejenigen, die geistig fähig waren die Fragen zu verstehen und zu beantworten, erhoben worden. Diese Einschätzung erfolgte durch das Prüfpersonal, sodass eine Selektion der eingeschlossenen Population durch das ärztliche Prüfpersonal vorgenommen wurde. Aus den Studienunterlagen ist nicht ersichtlich, wie viele Personen für den Endpunkt berücksichtigt wurden.

Abkürzungen: CGIC: Caregiver Global Impression of Change; CSR: Clinical Study Report; C-SSRS: Columbia-Suicide Severity Rating Scale; ITT: Intention-to-Treat; pU: pharmazeutischer Unternehmer; QOLCE: Quality of Life in Childhood Epilepsy; QOLIE-31-P: Quality Of Life In Epilepsy Inventory-31 – Patient Weighted Version; SGIC: Subject Global Impression of Change.; TSC: Tubuläre Sklerose; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Insgesamt 255 Patientinnen und Patienten wurden für den Einschluss in die Studie GWEP1521 gescreent, 31 wurden ausgeschlossen und 224 letztendlich auf die 4 Studienarme randomisiert. Für die vorliegende Nutzenbewertung werden der Studienarm mit einer Dosierung von Cannabidiol 25 mg/kg/Tag und die Placebo-Arme berücksichtigt; für die Auswertung wurden die beiden Placebo-Arme zu einer Population zusammengefasst.

In Tabelle 11 sind die allgemeinen Angaben der Studie GWEP1521 beschrieben.; ITT- und Sicherheitspopulation sind identisch. In der Interventionsgruppe brachen mit 13 % deutlich mehr Testpersonen die Einnahme der Studienmedikation ab im Vergleich zu Placebo (1 %), wobei dies am häufigsten auf UE zurückzuführen war. Die mediane Behandlungsdauer in den Behandlungsarmen war gleich. Von den Patientinnen und Patienten mit abgeschlossener Behandlungsphase entschied sich lediglich 1 Testperson aus jedem Behandlungsarm gegen eine Teilnahme an der OLE. Gemäß Modul 4 trat bei 19 Personen (25 %) des Interventions- und 13 (17 %) des Kontrollarms eine relevante Protokollverletzung auf. Die Angaben zwischen den Auswertungen, den detaillierten Fallbeschreibungen im Studienbericht und den Angaben in Modul 4 divergieren leicht. Ungefähr die Hälfte aller Protokollverletzungen sind gemäß Studienbericht laborbedingt. Darunter fallen durchgeführte einzelne Labortests, die nicht erforderlich waren, und erforderliche Labortests, die nicht durchgeführt wurden. Protokollverletzungen von Ein- und Ausschlusskriterien betreffen ebenfalls lediglich Labortests. Eine Erhöhung von Begleitmedikationen während der Studie erfolgte bei 4 Personen der Interventions- und 1 Person der Kontrollgruppe. Die Einwilligung betreffende Protokollverletzungen beziehen sich auf nicht rechtzeitig ausgefüllte Einwilligungserklärungen. Im Fall der Untersuchungen handelt es sich um eine mangelhafte Nachverfolgung von UE nach Studienabbruch und fehlende C-SSRS-Erhebungen während der Studie bei einem Zentrum, von denen jeweils 1 Person aus dem Interventions- und Placebo-Arm betroffen waren. Bei 3 Personen des Cannabidiol-Arms ereigneten sich Protokollverletzungen bezüglich der Dosierung der Studienmedikation: eine fehlerhafte Titrierung, eine fehlerhafte Dosisreduktion und eine fehlerhafte Einnahme der Studienmedikation. Bei den weiteren Protokollverletzungen handelt sich um eine zu späte Randomisierung und eine falsch festgestellte Annahme, dass die Studie beendet ist.

Tabelle 11: Allgemeine Angaben der Studie GWEP1521

Studie GWEP1521 Allgemeine Angaben	Cannabidiol	Placebo¹⁾
ITT-Population, n (%)	75 (100)	76 (100)
Sicherheitspopulation, n (%)	75 (100)	76 (100)
Abbruch der Einnahme der Studienmedikation während der Behandlungsphase, n (%)	10 (13,3)	1 (1,3)
Aufgrund von UE	8 (10,7)	0
Entzug der Einverständniserklärung	1 (1,3)	1 (1,3)
Andere	1 (1,3)	0

Studie GWEP1521 Allgemeine Angaben	Cannabidiol	Placebo¹⁾
Relevante Protokollverletzungen, n (%)	19 (25,3) ²⁾	13 (17,1) ²⁾
Einwilligung	2 (2,7)	2 (2,6)
Einnahme ausgeschlossener Medikationen	1 (1,3)	0
Labortests	9 (12,0)	8 (10,5)
Untersuchungen	2 (2,7)	2 (2,6)
Studienmedikation	2 (2,7)	1 (1,3)
Datum der Studienvisite	2 (2,7)	0
Andere	2 (2,7)	0
Mediane Behandlungsdauer in Tagen (min; max)	113 (9; 121)	113 (106; 123)
Mediane Beobachtungsdauer in Tagen (min; max)	k. A. ³⁾	k. A. ⁴⁾
Übergang in die Titrierungsperiode, n (%)	1 (1,3)	1 (1,3)
Übergang in OLE, n (%)	64 (85,3)	75 (98,7)

¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

²⁾ Anzahl der Personen mit mindestens einer relevanten Protokollverletzung.

³⁾ Keine Angaben im Studienbericht. In Modul 4 liegen Angaben hierfür vor, allerdings entsprechen diese der medianen Behandlungsdauer.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; OLE: Offene Extensionsstudie; UE: Unerwünschtes Ereignis;

In Tabelle 12 sind die wesentlichen Charakteristika der Studienpopulation dargestellt. Die Stratifizierung erfolgte nach Altersgruppen, was ein ausgeglichenes mittleres Alter von 14 Jahren in beiden Behandlungsarmen zur Folge hatte. Fast drei Viertel der Patientenpopulation waren Kinder und Kleinkinder. In etwa die Hälfte der Testpersonen nahm über Studienzentren in den USA teil; die verbleibenden Testpersonen wurden fast alle in Europa rekrutiert. Im Mittel nahmen Testpersonen sowohl im Behandlungs- als auch im Placebo-Arm zu Studienbeginn fast 3 Antiepileptika ein und hatten knapp 5 weitere Antiepileptika in ihrer Behandlungshistorie.

Der bei den meisten Testpersonen während der Baseline-Periode auftretende Anfallstyp waren fokale Anfälle mit Beeinträchtigung des Bewusstseins oder der Wahrnehmung (Typ 2), gefolgt von fokal-motorischen Anfällen ohne Beeinträchtigung des Bewusstseins (Typ 1). Bei lediglich 16 % der Personen aus der Cannabidiol- und 20 % der Placebo-Gruppe traten neben TSC-assoziierten Anfällen noch weitere Anfallsarten während der Baseline-Periode auf, wie Absenzen oder infantile oder partiell epileptische Spasmen. Im Median unterschied sich die Anzahl der TSC-assoziierten Anfälle und der Gesamtzahl der Anfälle mit jeweils etwas mehr als 50 Anfällen kaum zwischen den Behandlungsarmen.

Bei den Begleiterkrankungen zu Studienbeginn traten „Entwicklungsverzögerungen“ (19 vs. 30 %) etwas häufiger unter Placebo auf, dafür „Angst“ etwas häufiger unter Cannabidiol. Weiterhin waren „Angiofibrome“ und „Störungen auf dem Autistischen Spektrum“ bei ca. einem Viertel der Patientenpopulation prävalent. Im Wesentlichen waren die Charakteristika der Studienpopulation zwischen den Studienarmen ausgeglichen. Keine Informationen liegen für die Krankheitsdauer vor.

Tabelle 12: Charakterisierung der Studienpopulation GWEP1521

Studie GWEP1521 Charakterisierung der Studienpopulation	Cannabidiol	Placebo¹⁾
<i>Alter (Jahre)</i> MW (SD) Median (min; max)	14,1 (10,8) 11,6 (1,1; 56,8)	13,8 (10,6) 10,9 (1,2; 55,8)
<i>Altersgruppe (Jahre)²⁾, n (%)</i> 1–6 Jahre 7–11 Jahre 12–17 Jahre 18–65 Jahre	21 (28,0) 18 (24,0) 16 (21,3) 20 (26,7)	22 (28,9) 18 (23,7) 16 (21,1) 20 (26,3)
<i>Geschlecht, n (%)</i> weiblich männlich	32 (42,7) 43 (57,3)	31 (40,8) 45 (59,2)
<i>Abstammung (genetisch), n (%)</i> kaukasisch/weiß schwarz/afroamerikanisch indigene Völker Amerikas/Alaskas asiatisch andere	68 (90,7) 2 (2,7) 1 (1,3) 1 (1,3) 3 (4,0)	67 (88,2) 0 0 3 (3,9) 6 (7,9)
<i>Land, n (%)</i> USA Polen Australien Spanien Niederlande UK	36 (48,0) 23 (30,7) 5 (6,7) 6 (8,0) 3 (4,0) 2 (2,7)	42 (55,3) 16 (21,1) 10 (13,2) 2 (2,6) 3 (3,9) 3 (3,9)
<i>Körpergewicht zu Baseline (kg)</i> MW (SD) Median (min; max)	42,8 (24,0) 40,4 (12,1; 143,7)	44,6 (27,9) 36,9 (10,4; 115,7)
<i>Körpergröße zu Baseline (cm)</i> MW (SD) Median (min; max)	142,1 (27,1) 146 (81; 196)	n = 75 ³⁾ 138,7 (29,5) 141 (80; 182)
<i>Anzahl eingenommener Antiepileptika zu Studienbeginn</i> MW (SD) Median (min; max)	2,7 (1,0) 3 (0 ⁴⁾ ; 4)	2,6 (0,9) 3 (1; 5)
<i>Anzahl eingenommener Antiepileptika zu Studienbeginn, n (%)</i> 0 1 2 3 4 5	1 (1,3) ⁴⁾ 9 (12,0) 20 (26,7) 30 (40,0) 15 (20) 0	0 8 (10,5) 27 (35,5) 32 (42,1) 7 (9,2) 2 (2,6)
<i>Anzahl vormals eingenommener Antiepileptika</i> MW (SD) Median (min; max)	4,8 (3,2) 4 (0; 13)	4,6 (3,3) 4 (0; 15)
<i>Krankheitsdauer in Jahren</i> Zeit ab Diagnose bis Randomisierung MW (SD) Median (min; max)	k. A. k. A.	k. A. k. A.

Studie GWEP1521 Charakterisierung der Studienpopulation	Cannabidiol	Placebo¹⁾
TSC-assoziierte Anfälle während Baseline-Periode, n (%)	75 (100)	76 (100)
Atonisch	10 (13,3)	13 (17,1)
Klonisch	3 (4,0)	2 (2,6)
Tonisch	27 (36,0)	15 (19,7)
Tonisch-klonisch	22 (29,3)	14 (18,4)
Typ-1 fokal-motorisch	29 (38,7)	33 (43,4)
Typ-2 fokal	46 (61,3)	50 (65,8)
Typ-3 fokal	17 (22,7)	24 (31,6)
Weitere Anfälle während Baseline-Periode, n (%)	12 (16,0)	15 (19,7)
Absencen	5 (6,7)	7 (9,2)
Infantile oder epileptische Spasmen	5 (6,7)	3 (3,9)
Myoklonisch	3 (4,0)	4 (5,3)
Partiell (fokal-)sensorisch	2 (2,7)	3 (3,9)
<i>Anzahl TSC-assoziiierter Anfälle während Baseline-Periode pro 28 Tage</i>		
MW (SD)	78,0 (83,4)	89,2 (101,8)
Median (min; max)	56,0 (7,7; 427,7)	54,0 (8,0; 558,0)
<i>Gesamtzahl der Anfälle während Baseline-Periode pro 28 Tage</i>		
MW (SD)	80,9 (84,7)	118,32 (211,9)
Median (min; max)	56,0 (7,7; 427,7)	56,5 (8,0; 1660,1)
<i>Begleiterkrankungen zu Studienbeginn nach MedDRA-PT, n (%)⁵⁾</i>		
Hamartom	8 (10,7)	12 (15,8)
Subependymale noduläre Heterotopie	5 (6,7)	9 (11,8)
Obstipation	12 (16,0)	5 (6,6)
Entwicklungsverzögerung	14 (18,7)	23 (30,3)
Jahreszeitbedingte Allergie	8 (10,7)	6 (7,9)
Arzneimittelüberempfindlichkeit	8 (10,7)	4 (5,3)
Angiofibrom	21 (28,0)	16 (21,1)
Rhabdomyom	16 (21,3)	20 (26,3)
Angiomyolipom der Niere	15 (20,0)	15 (19,7)
melanozytischer Naevus	5 (6,7)	8 (10,5)
intellektuelle Behinderung	9 (12,0)	7 (9,2)
Störungen auf dem autistischen Spektrum	17 (22,7)	21 (27,6)
Schlaflosigkeit	9 (12,0)	6 (7,9)
Angst	12 (16,0)	3 (3,9)
Nierenzyste	10 (13,3)	17 (22,4)
Hauthypopigmentierung	14 (18,7)	14 (18,4)
Hypertonie	8 (10,7)	3 (3,9)

¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

²⁾ Stratifikationsfaktor.

³⁾ Nach Angaben aus Modul 4 wurde die Körpergröße einer Person des Kontrollarms beim Screening in Zoll gemessen, aber in cm angegeben. Die Daten wurden ausgeschlossen.

⁴⁾ Eine Testperson wurde in die Studie aufgenommen, obwohl sie keine laufenden begleitenden Antiepileptika einnahm. Die Testperson wurde nach Protokollversion 5 aufgenommen, die in Abschnitt 6.1 nicht ausdrücklich vorsah, dass die Testpersonen ein oder mehrere andere Antiepileptika einnehmen mussten. Das Protokoll wurde in der Folge durch entsprechende Änderungen aktualisiert.

⁵⁾ Begleiterkrankungen, die bei > 10 % der Testpersonen in einem Behandlungsarm auftraten.

Abkürzungen: k. A.: keine Angabe; MedDRA: Medical Dictionary for Regulatory Activities; MW: Mittelwert; PT: Preferred Term; SD: Standardabweichung; TSC: Tuberöse Sklerose.

Exposition mit der Studienmedikation

Daten zur Exposition mit der Studienmedikation sind in Tabelle 13 dargestellt. Die mediane Dauer der Exposition war zwischen den Behandlungsarmen mit 113 Tagen gleich. Aufgrund der vermehrten Behandlungsabbrüche in der Interventionsgruppe lag die mittlere Expositionsdauer mit 102 Tagen etwas unter der der Placebo-Gruppe mit 113 Tagen. Insgesamt zeigte sich mit 94 % im Interventions- und 100 % im Kontrollarm insbesondere unter Placebo eine hohe Behandlungs-Compliance bezogen auf die Anzahl der Tage, an denen die Testpersonen sich in der Behandlungsphase befanden. Angaben zu Dosisanpassungen der Studien- und Begleitmedikationen konnten nicht identifiziert werden.

Tabelle 13: Angaben zur Exposition mit der Studienmedikation in Studie GWEP1521

Studie GWEP1521 Exposition mit der Studienmedikation	Cannabidiol	Placebo¹⁾
<i>Dauer der Exposition (Tage)</i> MW (SD) Median (min; max)	102,6 (28,7) 113 (9; 121)	113,4 (2,7) 113 (106; 123)
<i>Tage ohne Einnahme der Studienmedikation</i> MW (SD) Median (min; max)	0,1 (0,7) 0 (0; 5)	0,0 (0,2) 0 (0; 1)
<i>Compliance der Behandlung (%)¹⁾</i> MW (SD) Median (min; max)	96,4 (11,2) 100 (40; 100)	99,7 (0,7) 100 (96; 100)
<i>Dosisanpassungen (Studienmedikation und Begleitmedikation), n (%)</i> Unterbrechung Erhöhung Reduktion	k. A.	k. A.

¹⁾ Zeit von Tag 1 bis zum Ende der Behandlung / Abbruch abzüglich aller Tage ohne Einnahme der Studienmedikation.

Abkürzungen: k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung.

Begleitmedikation

Wie aus Tabelle 14 hervorgeht, wurden antiepileptische Begleittherapien bei den Testpersonen in der Studie GWEP1521 eher selten durchgeführt. Als antiepileptische Begleitmedikationen wurden Clobazam, Valproinsäure, Vigabatrin, Levetiraceta und Lamotrigin am häufigsten eingesetzt. Vigabatrin wurde dabei etwas häufiger in der Interventionsgruppe (37 vs. 22 %) eingesetzt und Clobazam etwas häufiger in der Kontrollgruppe (23 vs. 33 %). Rund ein Drittel der Personen aus der Kontroll- und Behandlungsgruppe gab die begleitende Einnahme von Notfallmedikationen an.

Tabelle 14: Begleitmedikationen ≥ 10 % innerhalb eines Behandlungsarms in Studie GWEP1521

Studie GWEP1521 Begleitmedikationen ≥ 10 %	Cannabidiol n (%)	Placebo ¹⁾ n (%)
Antiepileptische Begleittherapien ≥ 10 % in einem Behandlungsarm		
Vagusnerv-Stimulation	10 (13,3)	8 (10,5) ²⁾
Antiepileptische Begleitmedikationen ≥ 10 % in einem Behandlungsarm		
Benzodiazepine und Derivate	25 (33,3)	32 (42,1)
Clobazam	17 (22,7)	25 (32,9)
Carboxamide und Derivate	30 (40,0)	24 (31,6)
Oxcarbazepin	13 (17,3)	10 (13,2)
Carbamazepin	11 (14,7)	6 (7,9)
Rufinamid	7 (9,3)	8 (10,5)
Fettsäuren-Derivate	42 (56,0)	46 (60,5)
Valproinsäure ³⁾	29 (38,7)	35 (46,1)
Vigabatrin	28 (37,3)	17 (22,4)
Andere Antiepileptika	56 (74,7)	58 (76,3)
Levetiracetam	19 (25,3)	24 (31,6)
Lamotrigin	17 (22,7)	18 (23,7)
Lacosamid	16 (21,3)	12 (15,8)
Topiramat	7 (9,3)	12 (15,8)
Zonisamid	9 (12,0)	7 (9,2)
Andere Begleitmedikationen ≥ 10 % in einem Behandlungsarm		
Jegliche andere Begleitmedikation	60 (80,0)	67 (88,2)
Ibuprofen	15 (20,0)	19 (25,0)
Paracetamol	18 (24,0)	9 (11,8)
Colecalciferol (Vitamin D3)	12 (16,0)	8 (10,5)
Sirolimus	5 (6,7)	9 (11,8)
Melatonin	11 (14,7)	7 (9,2)
Vitamine	8 (10,7)	5 (6,6)
Odansetron	8 (10,7)	0
Notfallmedikationen ≥ 10 % in einem Behandlungsarm		
Jegliche Notfallmedikation	25 (33,3)	26 (34,2)
Benzodiazepine und Derivate	25 (33,3)	26 (34,2)
Diazepam	20 (26,7)	16 (21,1)

¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

²⁾ Angaben für 1 Person fehlend.

³⁾ Unter dem Begriff Valproinsäure wurden mehrere Preferred Terms zusammengefasst.

3.2 Mortalität

Todesfälle wurden in der Studie GWEP1521 im Rahmen der Sicherheitserhebung erfasst. Im Verlauf der verblindeten Studienphase traten keine Todesfälle auf.

3.3 Morbidität

Epileptische Anfälle

Patientinnen und Patienten oder die jeweilige betreuende Person sollten Anfallsart und -anzahl von Baseline an (Studienvsiste 2) täglich mittels IVRS erheben. Im Mittel betrug die Anzahl fehlender Werte im IVRS 5 bzw. 6 Tage in der Interventions- bzw. Kontrollgruppe. Knapp die Hälfte der Testpersonen hatte keinen oder maximal einen fehlenden Tag im IVRS. Insgesamt weist das Anfallstagebuch sowohl in der Interventions- als auch der Kontrollgruppe eine hohe Vollständigkeit auf. Zu berücksichtigen ist jedoch, dass offensichtlich nur der Zeitraum bis zu einem Therapieabbruch gewertet wurde.

In Tabelle 15 sind die Ergebnisse für die Reduktion der Häufigkeit TSC-assoziiierter Anfälle, Anfälle gesamt und weiterer Anfälle jeweils im Vergleich zu Baseline dargestellt. In der Analyse mittels negativ-binomialen Modell zeigte sich ein statistisch signifikanter Vorteil unter Cannabidiol im Vergleich zu Placebo für die Häufigkeit der TSC-assoziierten Anfälle und der Anfälle gesamt. Kein statistisch signifikanter Unterschied zeigte sich dagegen für die Häufigkeit der weiteren Anfälle.

Tabelle 15: Ergebnisse zum Endpunkt „Häufigkeit epileptischer Anfälle“ in Studie GWEP1521; Behandlungszeitraum¹⁾²⁾; ITT-Population

Studie GWEP1521 Häufigkeit epileptischer Anfälle	Cannabidiol ³⁾ N = 75		Placebo ³⁾⁴⁾ N = 76		Cannabidiol vs. Placebo
	Baseline [95%-KI] ⁵⁾	Verhältnis Behandlung/ Baseline [95%-KI]	Baseline [95%-KI] ⁵⁾	Verhältnis Behandlung/ Baseline [95%-KI]	Verhältnis Cannabidiol/ Placebo ³⁾ [95%-KI] p-Wert
Häufigkeit TSC-assoziiierter Anfälle ⁶⁾⁷⁾	51,6 40,8; 65,3	0,51 [0,45; 0,59]	54,8 43,4; 69,3	0,73 [0,64; 0,84]	0,70 [0,58; 0,85] 0,0003
Häufigkeit Anfälle gesamt ⁸⁾	53,9 [42,4; 68,6]	0,52 [0,45; 0,60]	60,7 [47,8; 77,1]	0,73 [0,64; 0,84]	0,71 [0,58; 0,86] 0,0007
Häufigkeit weitere Anfälle ⁹⁾	11,3 [3,4; 37,0]	0,61 [0,30; 1,22]	11,3 [3,4; 37,0]	0,39 [0,21; 0,72]	1,56 [0,61; 3,97] 0,3470

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Post-hoc-Auswertungen des pU für das Nutzendossier. Der Effektschätzer und der p-Wert aus dem Studienbericht weichen leicht von diesen Werten ab. Der pU liefert keine Begründung für die Abweichung in diesem Endpunkt. An anderer Stelle in Modul 4 werden die Abweichungen darauf zurückgeführt, dass Testpersonen aus dem nicht zugelassenen Hochdosis-Behandlungsarm nicht in die Analyse eingeschlossen wurden.

³⁾ Effektschätzer entspricht negativ-binomialen Modell; Stratifizierungsfaktoren: Altersgruppe, Zeitpunkt (Baseline und Behandlung), Behandlungsgruppe und Interaktion zwischen Zeitpunkt und Behandlung als fixe Effekte und Patientin/Patient als Zufallseffekt. Log-transformierte Anzahl Tage mit Anfallsberichten als Offset.

⁴⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

⁵⁾ Baseline-Wert näherungsweise aus dem Studienbericht dargestellt, da in Modul 4 des Nutzendossiers keine Baseline-Werte für das Modell berichtet sind.

⁶⁾ Primäre Wirksamkeitsanalyse.

⁷⁾ TSC-assoziierte Anfälle umfassen: Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins; fokale Anfälle mit Beeinträchtigung des Bewusstseins; fokale Anfälle, die in bilaterale generalisierte konvulsive Anfälle übergehen; tonisch-klonische, tonische, klonische oder atonische Anfälle.

⁸⁾ „Anfälle gesamt“ umfasst alle in der Studie erhobenen Anfallstypen.

⁹⁾ „Weitere Anfälle“ umfasst Absenzen, myoklonische Anfälle, partiell sensorische Anfälle und infantile oder epileptische Spasmen.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; TSC: Tuberöse Sklerose.

Sensitivitätsanalysen

Gemäß SAP war für den primären Endpunkt „Reduktion der Häufigkeit TSC-assoziiierter Anfälle“ eine Reihe von Sensitivitätsanalysen geplant. Aufgrund der Aktualisierung der statistischen Analyse des primären Endpunkts von einem Wilcoxon-Rangsummentest auf eine negativ-binomiale Regressionsanalyse in einem Protokoll-Amendment kurz vor Studienende, werden die Ergebnisse der vorherigen primären Wirksamkeitsanalyse in Tabelle 16 dargestellt. In dieser Auswertung zeigt sich ein statistisch signifikantes Ergebnis zugunsten von Cannabidiol. Diese Ergebnisse bestätigten sich auch in den weiteren geplanten Sensitivitätsanalysen im Studienbericht, für die jedoch in Modul 4 keine Daten vorgelegt wurden.

Tabelle 16: Ergebnisse zum Endpunkt „Häufigkeit TSC-assoziiierter Anfälle“ in Studie GWEP1521; Behandlungszeitraum¹⁾ (Sensitivitätsanalyse); ITT-Population

Studie GWEP1521 Häufigkeit TSC- assoziierter Anfälle	Cannabidiol N = 75		Placebo²⁾ N = 76		Cannabidiol vs. Placebo
	Baseline Median [IQR]	% Ver- änderung Median [IQR]	Baseline Median [IQR]	% Ver- änderung Median [IQR]	Differenz der Mediane³⁾ [95%-KI] p-Wert⁴⁾
Häufigkeit TSC- assoziierter Anfälle ⁵⁾	78,0 (33,0; 114,0)	-43,4 [-67,8; -13,6]	74,3 (40,5; 114,0)	-20,1 [-47,1; 3,1]	-18,77 [-31,32; -6,28] 0,0038

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

³⁾ Differenz der Mediane nach Hodges-Lehmann.

⁴⁾ Wilcoxon-Rangsummentest.

⁵⁾ TSC-assoziierte Anfälle umfassen: Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins; fokale Anfälle mit Beeinträchtigung des Bewusstseins; fokale Anfälle, die in bilaterale generalisierte konvulsive Anfälle übergehen; tonisch-klonische, tonische, klonische oder atonische Anfälle.

Abkürzungen: IQR: Interquartilsabstand (Interquartile Range); ITT: Intention-to-Treat; KI: Konfidenzintervall; TSC: Tuberöse Sklerose.

Responderanalysen

Wie in Tabelle 17 dargestellt, zeigte sich für den Endpunkt „Reduktion der TSC-assoziierten Anfälle um mindestens 50 %“ im Vergleich zu Baseline mit 27 Personen (36 %) in der Interventions- und 17 (22 %) in der Kontrollgruppe zwar ein numerischer Vorteil für Cannabidiol, der jedoch nicht statistisch signifikant ausfiel; dies gilt auch für eine Reduktion um ≥ 25 %. Für das Responder-Kriterium ≥ 75 % zeigte sich dagegen ein statistisch signifikanter Vorteil von Cannabidiol im Vergleich zu Placebo. Lediglich 1 Person im Interventionsarm erzielte eine vollständige Anfallsfreiheit (Reduktion 100 %). Bei 17 (23 %) bzw. 21 (28 %) Personen zeigte sich entweder keine Veränderung oder eine Steigerung in der Anfallshäufigkeit. Eine statistische Auswertung hierfür liegt nicht vor.

Tabelle 17: Ergebnisse für den Endpunkt „Häufigkeit TSC-assoziierte Anfälle“ in Studie GWEP1521; Responderanalysen (Reduktion ≥ 25 , ≥ 50 , ≥ 75 , 100 %; Anstieg > 0 %); Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521 Häufigkeit TSC- assoziierter Anfälle ²⁾	Cannabidiol 25 mg/kg/Tag N = 75	Placebo ³⁾ N = 76	Cannabidiol 25 mg/kg/Tag vs. Placebo
	Personen mit Ereignis ⁴⁾ n (%)	Personen mit Ereignis ⁴⁾ n (%)	RR [95%-KI] ⁵⁾ ; p-Wert ⁶⁾
Reduktion ≥ 25 %	43 (57,3)	33 (43,4)	1,29 [0,93; 1,77]; 0,091
Reduktion ≥ 50 %	27 (36,0)	17 (22,4)	1,61 [0,96; 2,69]; 0,0692
Reduktion ≥ 75 %	12 (16,0)	0	25,33 [1,53; 420,24] ⁷⁾ ; 0,0003
Reduktion 100°%	1 (1,3)	0	n. b.; 0,3173
Anstieg ≥ 0 °%	17 (22,7)8)	21 (27,6)8)	k. A.

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ TSC-assoziierte Anfälle umfassen: Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins; fokale Anfälle mit Beeinträchtigung des Bewusstseins; fokale Anfälle, die in bilaterale generalisierte konvulsive Anfälle übergehen; tonisch-klonische, tonische, klonische oder atonische Anfälle.

³⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

⁴⁾ Personen, die während der Behandlungsperiode aus der Studie ausgeschieden sind, wurden als Non-Responder gewertet.

⁵⁾ Post-hoc-Auswertungen für das Nutzendossier. Keine Angaben zu verwendetem Modell und verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.

⁶⁾ CMH-Test stratifiziert nach Altersgruppen.

⁷⁾ Modell ohne Stratifikationsfaktoren. In einer der Behandlungsgruppen treten keine Ereignisse auf. Aufgrund der vorliegenden Ereigniszahlen wird nur von geringfügigen Änderungen des Effektschätzers im Fall einer Stratifizierung ausgegangen.

⁸⁾ Eigene Berechnung über den Anteil der Testpersonen mit einem Anstieg ≥ 25 % und einer Veränderung von ≥ 0 bis > 25 °% in der Häufigkeit TSC-assoziiierter Anfälle im Vergleich zu Baseline.

Abkürzungen: CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; n. b.: nicht berechenbar; RR: Relatives Risiko; TSC: Tuberöse Sklerose.

Die Reduktionen der Häufigkeit der Anfälle gesamt sind in Tabelle 18 dargestellt. Es besteht lediglich für die Reduktion um 75 % im Vergleich zu Baseline ein statistisch signifikanter Unterschied. Bei 18 (24 %) bzw. 23 (30 %) Personen zeigte sich entweder keine Veränderung oder eine Steigerung in der Anfallshäufigkeit. Eine statistische Auswertung hierfür liegt nicht vor.

Tabelle 18: Ergebnisse für den Endpunkt „Häufigkeit Anfälle gesamt“ in Studie GWEP1521; Responderanalysen (Reduktion ≥ 25 , ≥ 50 , ≥ 75 , 100 %; Anstieg > 0 %); Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521 Häufigkeit Anfälle gesamt ²⁾	Cannabidiol N = 75	Placebo ³⁾ N = 76	Cannabidiol vs. Placebo
	Personen mit Ereignis ⁴⁾ n (%)	Personen mit Ereignis ⁴⁾ n (%)	RR [95%-KI] ⁵⁾ ; p-Wert ⁶⁾
Reduktion ≥ 25 %	42 (56,0)	34 (44,7)	1,22 [0,89; 1,68]; 0,1733
Reduktion ≥ 50 %	26 (34,7)	16 (21,1)	1,64 [0,96; 2,79]; 0,0655
Reduktion ≥ 75 %	12 (16,0)	1 (1,3)	12,07 [1,61; 90,44]; 0,0015
Reduktion 100°%	1 (1,3)	0	n. b.; 0,3173
Anstieg ≥ 0 °%	18 (24,0)7)	23 (30,3)7)	k. A.

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Anfälle gesamt umfasst alle in der Studie erhobenen Anfallstypen.

³⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

⁴⁾ Personen, die während der Behandlungsperiode aus der Studie ausgeschieden sind, wurden als Non-Responder gewertet.

⁵⁾ Post-hoc-Auswertungen für das Nutzendossier. Keine Angaben zu verwendetem Modell und verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.

⁶⁾ CMH-Test stratifiziert nach Altersgruppen.

⁷⁾ Eigene Berechnung über den Anteil der Testpersonen mit einem Anstieg $\geq 25\%$ und einer Veränderung von ≥ 0 bis $> 25\%$ in der Häufigkeit der Anfälle gesamt im Vergleich zu Baseline.

Abkürzungen: CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; n. b.: nicht berechenbar; RR: Relatives Risiko.

Subgruppenanalyse

In Modul 4 wurden keine Subgruppenanalysen dargestellt. Gemäß SAP waren Subgruppenanalysen für den primären Endpunkt „Veränderung der Häufigkeit TSC-assoziiierter Anfälle“ sowie den sekundären Endpunkt „Reduktion der TSC-assoziierten Anfälle um $\geq 50\%$ “ im Vergleich zur Baseline-Periode“ geplant. Eine statistisch signifikante Effektmodifikation liegt lediglich für das Merkmal „Region“ ($p=0,0205$) im Endpunkt „Reduktion der TSC-assoziierten Anfälle um $\geq 50\%$ “ vor, analysiert mittels logistischer Regression. RR liegen als Effektschätzer nicht vor. Es zeigten sich auch keine relevanten Interaktionen bezüglich der Altersgruppen.

Status epilepticus

Unter der Behandlung mit Cannabidiol 25 mg/kg/Tag trat bei 5 Personen mindestens eine Status-epilepticus-Episode auf, im Kontrollarm bei 7 Personen. In der post hoc durchgeführten Analyse zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.

Tabelle 19: Ergebnisse für den Endpunkt „Status epilepticus“ in Studie GWEP1521; Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521	Cannabidiol N = 75		Placebo ²⁾ N = 76		Cannabidiol vs. Placebo
	Personen mit Ereignis		Personen mit Ereignis		RR [95%-KI] ⁴⁾ ; p-Wert ⁵⁾
	Baseline <i>n</i> (%)	Behandlungs- zeitraum <i>n</i> (%)	Baseline <i>n</i> (%)	Behandlungs- zeitraum <i>n</i> (%)	
Status epilepticus ³⁾	0	5 (6,7)	3 (7,0)	7 (9,2)	0,78 [0,25; 2,44]; 0,5636

¹⁾ Titrations- und Erhaltungsphase.

²⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

³⁾ Status epilepticus ist definiert als jegliche Art von Anfall (konvulsiv und nicht-konvulsiv) der 30 Minuten oder länger andauert.

⁴⁾ Post-hoc-Auswertungen für das Nutzendossier. Keine Angaben zu verwendetem Modell und verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.

⁵⁾ P-Wert basierend auf CMH-Test. Keine Angaben zur Verwendung von Stratifikationsfaktoren in der post hoc für das Nutzendossier durchgeführten Analyse durch den pU.

Abkürzungen: CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Globaler Eindruck der Veränderung (mittels CGIC / SGIC)

Der globale Eindruck der Veränderung wurde sowohl mit dem CGIC als auch dem SGIC erhoben. Zu Baseline sollten die betreuenden Personen bzw. die Testpersonen kurze formlose Aufzeichnungen als Gedankenstütze formulieren. Die Erhebung des Endpunkts erfolgte daraufhin einmalig bei Studienvisite 10 am Ende der Behandlung. Lediglich 6 Personen (8 %) des Behandlungsarms Cannabidiol 25 mg/kg/Tag und 4 (5 %) der Placebo-Gruppe beantworteten die patientenberichtete Version des Instruments (SGIC). Für die betreuerberichtete Version des Instruments liegen Rückläufe von 66 Personen (88 %) des Behandlungsarms Cannabidiol 25 mg/kg/Tag und von 72 (95 %) der Placebo-Gruppe vor. In die kombinierte Auswertung – wobei die betreuerberichtete Version priorisiert wurde, falls bei einer Testperson Rückläufe für den CGIC und den SGIC vorlagen – gingen 70 Personen (93 %) des Behandlungsarms Cannabidiol 25 mg/kg/Tag und alle 76 (100 %) der

Placebo-Gruppe ein. Aufgrund der möglichen strukturellen Unterschiede von selbst- und fremdbewerteten Erhebungen wird eine zusammenfassende Auswertung kritisch gesehen. Bei der gemeinsamen Auswertung von CGIC und SGIC basierte die Erhebung lediglich bei jeweils 4 Personen der Interventions- und der Kontrollgruppe auf dem SGIC. Diese sind gemäß Tabelle 21 wie folgt verteilt: Cannabidiol: stark verbessert $n = 1$, gering verbessert $n = 2$ und stark verschlechtert $n = 1$; Placebo: sehr stark verbessert $n = 1$, gering verbessert $n = 2$ und nicht verändert $n = 1$. Jeweils 3 Personen erfüllten somit das Responder-Kriterium für eine Verbesserung, sodass die Ergebnisse der kombinierten CGIC-/SGIC-Analysen aufgrund der Gleichverteilung ergänzend berichtet werden.

Post hoc wertet der pU den Anteil der Testpersonen mit einer Verbesserung des globalen Eindrucks der Veränderung mittels Responderanalysen für die kombinierte Version des Instruments aus. Wie aus Tabelle 20 hervorgeht, zeigte sich bei 45 Personen (60 %) der Interventions- und 27 (36 %) der Kontrollgruppe eine Verbesserung im CGIC zum Ende der Behandlung. Separate Responderanalysen für den CGIC sowie Responderanalysen für eine Verschlechterung des globalen Eindrucks der Veränderung liegen nicht vor.

Tabelle 20: Ergebnisse für den Endpunkt „Globaler Eindruck der Veränderung“ in Studie GWEP1521; Responderanalysen; Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521 Globaler Eindruck der Veränderung	Cannabidiol N = 75		Placebo ²⁾ N = 76		Cannabidiol vs. Placebo RR [95%-KI] ⁵⁾ ; p-Wert ⁶⁾
	N ³⁾	Personen mit Ereignis n (%) ⁴⁾	N ³⁾	Personen mit Ereignis n (%) ⁴⁾	
Globaler Eindruck der Veränderung ⁷⁾ – Verbesserung					
CGIC	66	45 (60,0) ⁸⁾	72	27 (35,5) ⁸⁾	k. A.
CGIC / SGIC ⁹⁾	70	48 (68,6) ¹⁰⁾	76	30 (39,5)	1,74 [1,26; 2,39]; 0,0004
Globaler Eindruck der Veränderung ¹¹⁾ – Verschlechterung					
CGIC	66	7 (9,3) ⁸⁾	72	4 (5,3) ⁸⁾	k. A.
CGIC / SGIC	70	8 (10,7) ⁸⁾	76	4 (5,3) ⁸⁾	k. A.

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

³⁾ Anzahl Personen mit Rücklauf.

⁴⁾ Bezogen auf die ITT-Population.

⁵⁾ Keine Angaben zu verwendetem Modell und verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung in Modul 4 des Nutzen-dossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.

⁶⁾ P-Wert basierend auf CMH-Test stratifiziert nach Altersgruppen.

⁷⁾ Eine Verbesserung im globalen Eindruck der Veränderung ist definiert als (1) sehr stark verbessert, (2) stark verbessert, (3) gering verbessert.

⁸⁾ Eigene Berechnung durch Addition der entsprechenden Ergebnisse aus Tabelle 21.

⁹⁾ Ergänzende Darstellung, da keine Responderanalysen für den CGIC vorliegen. Die Auswertung basiert auf allen Testpersonen, für die ein Rücklauf entweder im CGIC oder SGIC zum Ende der Behandlung vorlag. Auswertungen, in der Personen ohne Rücklauf als Non-Responder gewertet werden, liegen nicht vor.

¹⁰⁾ Prozentzahl bezogen auf alle Testpersonen mit Rücklauf.

¹¹⁾ Eine Verschlechterung im globalen Eindruck der Veränderung ist definiert als (4) gering verschlechtert, (5) stark verschlechtert, (6) sehr stark verschlechtert.

Abkürzungen: CGIC: Caregiver Global Impression of Change; CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; RR: Relatives Risiko; SGIC: Subject Global Impression of Change.

Im SAP war a priori eine Auswertung des globalen Eindrucks der Veränderung mittels ordinal logistischer Regression geplant. Tabelle 21 stellt die Ergebnisse dieser Auswertung sowie den Anteil der Testpersonen in der jeweiligen Merkmalsausprägung inklusive fehlender Werte dar. Sowohl für die

separate Auswertung des CGIC als auch die kombinierte Auswertung von CGIC und SGIC war die Proportional-Odds-Annahme verletzt (p-Wert von 0,0164 und 0,0138), sodass gemäß SAP zusätzlich ein „Cochran-Armitage Trend Test“ durchgeführt wurde. Beide Auswertungen zeigten ebenfalls einen statistisch signifikanten Effekt zugunsten von Cannabidiol, der etwas geringer ausfällt.

Tabelle 21: Ergebnisse für den Endpunkt „Globaler Eindruck der Veränderung“ in Studie GWEP1521; ordinal logistische Regression; Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521 Globaler Eindruck der Veränderung	Cannabidiol N = 75	Placebo²⁾ N = 76	Cannabidiol vs. Placebo
	Studienvisite 10 <i>n (%)</i>	Studienvisite 10 <i>n (%)</i>	OR³⁾ [95%-KI]; p-Wert
CGIC			
sehr stark verbessert	10 (13,3)	2 (2,6)	n. v.; 0,0043 ⁵⁾
stark verbessert	12 (16,0)	10 (13,2)	
gering verbessert	23 (30,7)	15 (19,7)	
nicht verändert	14 (18,7)	41 (53,9)	
gering verschlechtert	4 (5,3)	3 (3,9)	
stark verschlechtert	2 (2,7)	1 (1,3)	
sehr stark verschlechtert	1 (1,3)	0 (0,0)	
fehlend ⁴⁾	9 (12,0)	4 (5,3)	
CGIC / SGIC			
sehr stark verbessert	10 (13,3)	3 (3,9)	n. v.; 0,0040 ⁶⁾⁷⁾
stark verbessert	13 (17,3)	10 (13,2)	
gering verbessert	25 (33,3)	17 (22,4)	
nicht verändert	14 (18,7)	42 (55,3)	
gering verschlechtert	4 (5,3)	3 (3,9)	
stark verschlechtert	3 (4,0)	1 (1,3)	
sehr stark verschlechtert	1 (1,3)	0	
fehlend ⁴⁾	5 (6,7)	0	

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

³⁾ Ordinal logistische Regression auf Basis eines kumulativen Logit-Modells mit Behandlungsgruppe und Altersgruppe als Faktoren. Es liegt jeweils lediglich ein OR für die ordinal logistische Regression vor. Informationen bezüglich der Einteilung der Kategorien und deren Anzahl bzw. die sich daraus ergebende Bestimmung einer Referenzkategorie konnten nicht identifiziert werden. Aufgrund dieser Unklarheiten bezüglich der Interpretation des OR werden diese als nicht verwertbar eingestuft.

⁴⁾ Fehlende Werte gehen nicht in die Analyse ein.

⁵⁾ Angaben basieren auf den Auswertungen im Studienbericht. Proportional-Odds-Annahme verletzt (p = 0,0164); p-Wert für den „Cochran-Armitage Trend Test“ als Sensitivitätsanalyse: 0,0116.

⁶⁾ Post-hoc-Auswertungen des pU für das Nutzendossier. Der p-Wert aus dem Studienbericht weicht leicht ab. Der pU liefert keine Begründung für die Abweichung in diesem Endpunkt. An anderer Stelle in Modul 4 werden die Abweichungen darauf zurückgeführt, dass Testpersonen aus dem nicht zugelassenen Hochdosis-Behandlungsarm nicht in die Analyse eingeschlossen wurden.

⁷⁾ Proportional-Odds-Annahme verletzt (p = 0,0138); p-Wert für den „Cochran-Armitage Trend Test“ als Sensitivitätsanalyse: 0,0265.

Abkürzungen: CGIC: Caregiver Global Impression of Change; ITT: Intention-to-Treat; KI: Konfidenzintervall; n. v.: nicht verwertbar; OR: Odds Ratio; pU: pharmazeutischer Unternehmer; SGIC: Subject Global Impression of Change.

Epilepsiebedingte Hospitalisierungen

Unter der Behandlung mit Cannabidiol 25 mg/kg/Tag wurden während der Behandlungsperiode für 8 Personen (10 %) epilepsiebedingte Hospitalisierungen berichtet, unter Placebo für 1 (1,3 %). Statistische Auswertungen für diese Daten konnten nicht identifiziert werden. Das vom pU post hoc berechnete und in Modul 4 angegebene RR bezieht sich auf Daten, die von den Angaben aus dem Studienbericht abweichen und nicht nachvollzogen werden können. Die vorliegenden Unterschiede werden seitens des pU nicht begründet.

Achenbach-Verhaltens-Checkliste

Zur Erfassung von Verhaltens- und emotionalen Symptomen wurden 3 unterschiedliche Verhaltens-Checklisten in Abhängigkeit vom Alter der Testpersonen eingesetzt. Der Anteil der ITT-Population im Alter 1,5–5 Jahre für die Analyse der „CBCL (1,5–5 Jahre)“ umfasste 15 Personen in der Cannabidiol- und 17 in der Placebo-Gruppe; 40 bzw. 39 Personen waren zwischen einschließlich 6 und 17 Jahren und in der ITT-Population für „CBCL (6–18 Jahre)“ enthalten. 20 Personen in jedem Behandlungsarm waren älter als 18 Jahre und fielen in die ITT-Population für „ABCL (18–59 Jahre)“.

Die Rücklaufquoten bezogen auf das Altersspektrum, in dem die jeweilige Checkliste eingesetzt wurde, lagen für alle Subskalen und Gesamtwerte über 70 %. In keiner der Subskalen oder Gesamtwerte zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Tabelle 22: Ergebnisse zum Endpunkt „Achenbach-Verhaltens-Checkliste“ in Studie GWEP1521; Behandlungszeitraum¹⁾; ITT-Population

Studie GWEP1521 Checkliste Subskala/ Gesamtwert	Cannabidiol ²⁾			Placebo ²⁾³⁾			Cannabidiol vs. Placebo
	n (%) ⁴⁾	Baseline MW (SD)	Veränderung LS-MW (SE) ⁵⁾	n (%) ⁴⁾	Baseline MW (SD)	Veränderung LS-MW (SE) ⁵⁾	LS-MWD [95%-KI] ²⁾ ; p-Wert ⁵⁾
CBCL (1,5–5 Jahre)							
Gesamt	14 (93)	62,3 (25,2)	-9,48 (4,15)	15 (88)	49,5 (28,1)	-5,49 (4,01)	-3,99 [-16,03; 8,05]; 0,5020
CBCL (6–18 Jahre)							
Gesamt	30 (75)	46,1 (27,1)	-4,02 (3,76)	32 (82)	46,1 (20,9)	-3,76 (3,86)	-0,27 [-8,04; 7,51]; 0,9457
Aktivitäten	28 (70)	5,1 (4,0)	0,45 (0,44)	32 (82)	4,1 (3,4)	1,06 (0,43)	-0,61 [-1,53; 0,31]; 0,1908
Soziale Kompetenzen	29 (73)	2,8 (2,2)	0,01 (0,33)	31 (79)	3,0 (2,4)	0,07 (0,34)	-0,06 [-0,75; 0,63]; 0,8616
Schule ⁶⁾	15 (38)	-	-	15 (38)	-	-	-
Gesamt- kompetenz ⁶⁾	14 (35)	-	-	13 (33)	-	-	-
ABCL (18–59 Jahre)							
Gesamt	18 (90)	44,4 (31,0)	-2,15 (5,01)	15 (75)	43,0 (21,0)	4,24 (5,49)	-6,39 [-21,56; 8,79]; 0,3968
Kritische Elemente	17 (85)	4,9 (4,1)	-1,07 (0,54)	15 (75)	6,0 (3,1)	-0,05 (0,59)	-1,02 [-2,66; 0,63]; 0,2151
Gesamt- auffälligkeit	18 (90)	5,7 (3,4)	0,39 (0,49)	15 (75)	4,9 (4,2)	0,26 (0,53)	0,13 [-1,35; 1,61]; 0,8571

¹⁾ Umfasst Titrations- und Erhaltungszeitraum.

²⁾ Post-hoc-Auswertungen des pU für das Nutzendossier. Der Effektschätzer und der p-Wert aus dem Studienbericht weichen leicht von diesen Werten ab. In Modul 4 werden die Abweichungen darauf zurückgeführt, dass Testpersonen aus dem nicht zugelassenen Hochdosis-Behandlungsarm nicht in die Analyse eingeschlossen wurden.

³⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

⁴⁾ Rücklaufquoten jeweils bezogen auf die Anzahl der Testpersonen der ITT-Population, die in das Altersspektrum der jeweiligen Checkliste fallen: „CBCL (1,5–5 Jahre)“ mit 15 Personen in der Cannabidiol- und 17 in der Placebo-Gruppe; „CBCL (6–18 Jahre)“ mit 40 bzw. 39 Personen und ABCL mit 20 Personen in jedem Behandlungsarm. Angaben sind Modul 4 entnommen.

⁵⁾ ANCOVA mit Baseline-Wert und Altersgruppe als Kovariaten und Behandlung als festem Effekt.

⁶⁾ Rücklaufquote geringer als 70 % für diese Skala, die Daten werden nicht dargestellt.

Abkürzungen: ABCL: Adult Behavior Checklist; ANCOVA: Kovarianzanalyse; CBCL: Child Behavior Checklist; ITT: Intention-to-Treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert; LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung; SE: Standardfehler.

Bewertung des adaptiven Verhaltens (mittels Vineland-II)

Die Ergebnisse für den Endpunkt „Bewertung des adaptiven Verhaltens“ werden nicht berichtet. Die Rücklaufquoten der Vineland-II-Skala liegen lediglich für 4 der 11 Subdomänen zum adaptiven Verhalten sowie für keine Subdomäne zum maladaptiven Verhalten über 70 % in beiden Behandlungsarmen. In keiner der 5 Domänen lag die Rücklaufquote in beiden Behandlungsarmen bei 70 % oder höher. Für den Gesamtwert des adaptiven Verhaltens beträgt die Rücklaufquote nur 57 % in der Behandlungsgruppe Cannabidiol 25 mg/kg/Tag und 62 % in der Placebo-Gruppe.

3.4 Lebensqualität

Gemäß Protokoll sollte der QOLCE in der Studie GWEP1521 ausschließlich bei Studienteilnehmenden im Alter von 2 bis 18 Jahren eingesetzt werden; Personen ab 19 Jahren erhielten den QOLIE-31-P. Die Ergebnisse des QOLIE-31-P werden aufgrund unklarer Validität des Instruments sowie geringer Rücklaufquoten in Bezug auf die Testpersonen über 18 Jahre (50 % im Placebo- und maximal 70 % im Behandlungsarm), nicht dargestellt.

Im Laufe der Studie GWEP1521 wurde der QOLCE teilweise auch für Studienteilnehmende ≥ 18 Jahre durch die betreuende Person beantwortet und ausgewertet. Es ist unklar, ob eine Selektion dieser Patientinnen und Patienten erfolgte und falls ja, nach welchen Kriterien. Daher sind für die Nutzenbewertung lediglich die Ergebnisse der Personen im Alter von 2 bis einschließlich 18 Jahren relevant, für die der Einsatz des Instruments a priori geplant war. Wie viele Personen in der jeweiligen Behandlungsgruppe jünger als 4 Jahre waren und damit außerhalb des beanspruchten und validierten Bereichs des QOLCE liegen, kann aus den vorliegenden Daten nicht abgeschätzt werden.

In Modul 4 stellt der pU Post-hoc-Analysen für alle Testpersonen bis einschließlich 18 Jahren dar. Diese Population umfasst 55 Personen der Cannabidiol- und 56 der Kontrollgruppe. Unklar ist, an welchem Zeitpunkt die Altersbestimmung erfolgte (Studien- oder Behandlungsbeginn). Die Rücklaufquoten in Bezug auf die Population der Studie im Alter von 18 Jahren und jünger lag für den Gesamtwert des QOLCE bei 64 % im Cannabidiol- und 77 % im Placebo-Arm. Innerhalb der Subskalen zeigten sich stark unterschiedliche Rücklaufquoten mit 40–96 % im Cannabidiol- und 36–96 % im Placebo-Arm. Anhand der unterschiedlichen Rücklaufquoten stellt sich die Frage, ob das gesamte Instrument für alle Testpersonen eine vergleichbare Relevanz besitzt.

Des Weiteren liegen Diskrepanzen zwischen den Daten des Modul 4 und des Studienberichts vor. Der pU begründet diese Abweichungen damit, dass für die post hoc durchgeführten Analysen nur die beiden Studienarme Cannabidiol 25 mg/kg/Tag und Placebo eingeschlossen wurden; der nicht zulassungskonforme Studienarm Cannabidiol 50 mg/kg/Tag wurde nicht berücksichtigt. Allerdings liegen auch Abweichungen bei den jeweils in die Auswertung eingehenden Personen vor. Diese liegen bei den post hoc durchgeführten Analysen für die Subskalen und den Gesamtwert mehrheitlich über der Anzahl der Personen, die in entsprechende Analysen im Studienbericht eingingen. Der pU liefert keine Rationale für sein Vorgehen.

Aus den genannten Gründen werden die vorliegenden Daten zur Lebensqualität gemessen anhand des QOLCE als nicht aussagekräftig eingeschätzt und nicht dargestellt.

3.5 Sicherheit

Für die Auswertung der Sicherheit wurden alle behandlungsbezogenen UE (TEAE) berücksichtigt, die während der verblindeten Studienphase auftraten, also ab Behandlungsbeginn bis einschließlich dem Erhalt der 1. Dosis in der OLE. Die mediane Behandlungsdauer war in beiden Behandlungsarmen mit 113 Tagen gleich (siehe Tabelle 11). Aufgrund vermehrter Therapieabbrüche unter Cannabidiol 25 mg/kg/Tag liegt die mittlere Behandlungsdauer im Interventionsarm mit 102 Tagen unter der des Placebo-Arms mit 113 Tagen. Für die Beobachtungsdauer konnten keine Informationen identifiziert werden. Die Sicherheitspopulation ist deckungsgleich mit der ITT-Population. Bei fast allen Testpersonen trat mindestens 1 UE auf. SUE und UE, die zum Therapieabbruch führten, wurden statistisch signifikant häufiger in der Interventionsgruppe beobachtet.

Tabelle 23: Ergebnisse zum Endpunkt „Unerwünschte Ereignisse“ während der verblindeten Behandlungsphase in Studie GWEP1521; Sicherheitspopulation

Studie GWEP1521 Unerwünschte Ereignisse	Cannabidiol N = 75 n (%)	Placebo ¹⁾ N = 76 n (%)	Cannabidiol vs. Placebo RR [95%-KI] ²⁾ ; p-Wert ³⁾
Personen mit mindestens einem ...			
UE	70 (93,3)	72 (94,7)	-
SUE	16 (21,3)	2 (2,6)	8,22 [1,96; 34,45]; 0,0004
UE, das zum Therapieabbruch führte	8 (10,7)	2 (2,6)	4,04 [0,90; 18,20]; 0,0484

¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

²⁾ Keine Angaben zum verwendeten Modell und den verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.

³⁾ CMH-Test stratifiziert nach Altersgruppen.

Abkürzungen: CMH: Cochran-Mantel-Haenszel; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Unerwünschte Ereignisse

UE mit einer Inzidenz von $\geq 10\%$ in einem der Studienarme sind in Tabelle 24 aufgeführt. UE der Systemorganklassen „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und „Untersuchungen“ traten statistisch signifikant häufiger unter Cannabidiol auf. Der Preferred Term „Fieber“ trat statistisch signifikant häufiger unter Cannabidiol auf und die Preferred Terms „Gamma-Glutamyltransferase erhöht“, „Alanin-Aminotransferase erhöht“ und „Aspartat-Aminotransferase erhöhte“ sogar ausschließlich unter Cannabidiol.

Tabelle 24: Ergebnisse zum Endpunkt „UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklassen und Preferred Terms“ in Studie GWEP1521; Sicherheitspopulation

Studie GWEP1521 UE mit Inzidenz ≥ 10 % MedDRA-Systemorganklasse Preferred Term	Cannabidiol N = 75 <i>n</i> (%)	Placebo¹⁾ N = 76 <i>n</i> (%)	Cannabidiol vs. Placebo RR [95%-KI]²⁾; p-Wert³⁾
Erkrankungen des Blutes und des Lymphsystems	8 (10,7)	4 (5,3)	2,06 [0,65; 6,51]; 0,2178
Erkrankungen des Gastrointestinaltrakts	37 (49,3)	27 (35,5)	1,39 [0,96; 2,01]; 0,0799
Erbrechen	13 (17,3)	7 (9,2)	1,86 [0,79; 4,37]; 0,144
Obstipation	8 (10,7)	6 (7,9)	1,38 [0,51; 3,74]; 0,557
Diarrhoe	23 (30,7)	19 (25,0)	1,22 [0,74; 2,03]; 0,4295
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	21 (28,0)	11 (14,5)	2,02 [1,06; 3,85]; 0,0404
Fieber	14 (18,7)	6 (6,7)	2,51 [1,05; 5,99]; 0,0428
Infektionen und parasitäre Erkrankungen	41 (54,7)	37 (48,7)	1,11 [0,83; 1,48]; 0,4344
Infektion der oberen Atemwege	7 (9,3)	10 (13,2)	0,71 [0,29; 1,75]; 0,469
Nasopharyngitis	11 (14,7)	12 (15,8)	0,93 [0,45; 1,95]; 0,8367
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	5 (6,7)	10 (13,2)	0,49 [0,18; 1,35]; 0,1891
Untersuchungen	30 (40,0)	11 (14,5)	2,77 [1,50; 5,12]; 0,0005
Gamma-Glutamyltransferase erhöht	12 (16,0)	0	25,33 [1,53; 420,24] ⁴⁾ ; 0,0003
Alanin-Aminotransferase erhöht	9 (12,0)	0	19,25 [1,14; 324,93] ⁴⁾ ; 0,0021
Aspartat-Aminotransferase erhöht	8 (10,7)	0	17,22 [1,01; 293,18] ⁴⁾ ; 0,0039
Stoffwechsel- und Ernährungsstörungen	20 (26,7)	13 (17,1)	1,54 [0,84; 2,84]; 0,1597
Appetit vermindert	15 (20,0)	9 (11,8)	1,71 [0,81; 3,63]; 0,1751
Erkrankungen des Nervensystems	27 (36,0)	32 (42,1)	0,86 [0,58; 1,28]; 0,4450
Somnolenz	10 (13,3)	7 (9,2)	1,40 [0,56; 3,50]; 0,4269
Psychiatrische Erkrankungen	19 (25,3)	16 (21,1)	1,21 [0,68; 2,16]; 0,5442
Erkrankungen der Atemwege, des Brustraums und Mediastinums	15 (20,0)	11 (14,5)	1,41 [0,70; 2,84]; 0,3603
Husten	8 (10,7)	5 (6,6)	1,63 [0,56; 4,74]; 0,3709
Erkrankungen der Haut und des Unterhautgewebes	11 (14,7)	6 (7,9)	1,82 [0,71; 4,67] 0,1872

- ¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.
²⁾ Keine Angaben zum verwendeten Modell und den verwendeten Stratifizierungsfaktoren. Aufgrund der Darstellung im Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen.
³⁾ CMH-Test stratifiziert nach Altersgruppen.
⁴⁾ Modell ohne Stratifikationsfaktoren. In einer der Behandlungsgruppen treten keine Ereignisse auf. Aufgrund der vorliegenden Ereigniszahlen wird nur von geringfügigen Änderungen des Effektschätzers im Fall einer Stratifizierung ausgegangen.
- Abkürzungen: CMH: Cochran-Mantel-Haenszel; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

Schwerwiegende unerwünschte Ereignisse

Bei den SUE nach Systemorganklasse und Preferred Term trat lediglich die Systemorganklasse „Infektionen und parasitäre Erkrankungen“ in einem Behandlungsarm bei mindestens 5 % der Personen während der verblindeten Behandlungsphase der Studie GWEP1521 auf. 4 Personen (5,3 %) im Behandlungsarm erlitten ein SUE dieser Systemorganklasse im Vergleich zu 1 Person (1,3 %) in der Kontrollgruppe. Der Effekt war statistisch nicht signifikant: RR: 4,05 (95%-KI: [0,47; 35,11]; $p = 0,1685$).

Unerwünschte Ereignisse, die zum Therapieabbruch führten

Bei den UE, die zum Therapieabbruch führten trat lediglich die Systemorganklasse „Erkrankungen der Haut und des Unterhautgewebes“ in einem Behandlungsarm bei mindestens 5 % der Personen während der verblindeten Behandlungsphase der Studie GWEP1521 auf. 4 Personen (5,3 %) im Behandlungsarm erlitten ein UE dieser Systemorganklasse, das zum Therapieabbruch führte, im Vergleich zu keiner Person in der Kontrollgruppe. Der Effekt war statistisch signifikant: RR: 9,12 (95%-KI: [0,50; 166,47]; $p = 0,0429$).

Unerwünschte Ereignisse von besonderem Interesse

Im Studienbericht werden „Selbstverletzende Gedanken“, „Euphorie“ und „Schwangerschaft“ als UE von besonderem Interesse dargestellt. Lediglich der Preferred Term „Euphorie“ trat bei einer Person im Behandlungsarm Cannabidiol 25 mg/kg/Tag auf.

Suizidalität (mittels C-SSRS)

Die C-SSRS zur Erhebung von suizidalem Verhalten wurde bei Personen ab 6 Jahren eingesetzt, die nach Ansicht des ärztlichen Prüfpersonals in der Lage waren die Fragen zu verstehen und zu beantworten. Bei Personen im Alter von 6 bis einschließlich 18 Jahren kam die kindgerechte Version des Fragebogens zum Einsatz. Für wie viele Personen Daten an den jeweiligen Erhebungszeitpunkten vorlagen ist unklar. In der Behandlungsgruppe Cannabidiol 25 mg/kg/Tag waren 44 Personen (72 %) 7 Jahre oder älter, in der Interventionsgruppe traf dies ebenfalls auf 44 Personen (71 %) zu.

Zur Screening-Visite gab eine Person aus der Kontrollgruppe einen Suizidgedanken („Ich wünschte ich wäre tot“) an. Zu keinem weiteren Erhebungszeitpunkt wurden von einer Person suizidales Verhalten oder suizidale Absichten angegeben.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Cannabidiol

Cannabidiol (Epidyolex®) wird gemäß Zulassung als adjuvante Behandlung von Patientinnen und Patienten ab 2 Jahren mit Krampfanfällen im Zusammenhang mit Tuberöser Sklerose (TSC) angewendet [22]. Weiterhin besitzt Cannabidiol in Kombination mit Clobazam eine Zulassung zur adjuvanten Behandlung von Patientinnen und Patienten ab 2 Jahren mit Krampfanfällen im Zusammenhang mit dem Lennox-Gastaut-Syndrom oder dem Dravet-Syndrom. Für Cannabidiol liegt eine Zulassung 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 vor. Gemäß Fachinformation wurde die Erstzulassung am 23. September 2019 erteilt. Datum der Zulassungserweiterung um das Anwendungsgebiet TSC war der 16. April 2021. Im vorliegenden Dokument wird eine Bewertung von Cannabidiol in der Indikation TSC vorgenommen. Die empfohlene Anfangsdosis von Cannabidiol beträgt 5 mg/kg/Tag für eine Woche. Nach einer Woche soll die Dosis auf zweimal täglich 5 mg/kg (10 mg/kg/Tag) erhöht werden und das klinische Ansprechen sowie die Verträglichkeit bewertet werden [22]. Basierend auf dem klinischen Ansprechen und der Verträglichkeit kann die Dosis in wöchentlichen Schritten von täglich 5 mg/kg/Tag bis zu einer empfohlenen Höchstdosis von 25 mg/kg/Tag weiter erhöht werden. Jede Dosissteigerung über 10 mg/kg/Tag bis zur empfohlenen Höchstdosis von 25 mg/kg/Tag sollte unter Abwägung des individuellen Nutzens und Risikos erfolgen.

Das klinische Studienprogramm von Cannabidiol in der Indikation TSC besteht ausschließlich aus einer Phase-III-Studie. Die Studie GWEP1521 liegt der Zulassung zugrunde und untersuchte Wirksamkeit und Sicherheit von Cannabidiol bei Kindern und Erwachsenen im Alter von 1 bis 65 Jahren mit TSC, deren Krampfanfälle nicht ausreichend kontrolliert sind. Cannabidiol wurde entweder in einer Dosierung von 25 mg/kg/Tag oder 50 mg/kg/Tag eingesetzt und mit Placebo verglichen. Der Behandlungsarm mit der Dosierung 50 mg/kg/Tag liegt außerhalb der zugelassenen Dosierung von Cannabidiol und wird nicht für die vorliegende Nutzenbewertung berücksichtigt. Für die Nutzenbewertung herangezogen wird der Behandlungsarm mit der Dosierung 25 mg/kg/Tag. Die Dosierung von konstant 25 mg/kg/Tag liegt am oberen Ende der fachinformationskonformen Dosierungsspanne und somit deutlich oberhalb der empfohlenen Erhaltungsdosis von 10 mg/kg/Tag. In Studie GEWP1521 erfolgte zudem keine patientenindividuelle Dosistitration nach dem jeweiligen klinischen Ansprechen und der Verträglichkeit. Die finale Erhaltungsdosis im Studienarm Cannabidiol 25 mg/kg/Tag war bereits an Tag 9 der Titrationsperiode erreicht. In der Fachinformation ist für diese Auftitration ein wesentlich längerer Zeitraum mit wöchentlichen Steigerungsschritten um 5 mg/kg/Tag vorgesehen. Mögliche Auswirkungen auf Wirksamkeit und Sicherheit können nicht abgeschätzt werden. Weiterhin erfolgte die Einnahme der Studienmedikation ohne Rücksicht auf Mahlzeiten. Abweichend davon beschreibt die Fachinformation, dass die Einnahme konsequent entweder mit oder ohne Nahrungsaufnahme eingenommen werden soll. Wie variabel die Einnahme mit Hinblick auf die zeitliche Nähe zu Mahlzeiten erfolgte und ob dies einen Einfluss auf die Wirksamkeit hat ist nicht dokumentiert. Studienteilnehmende mussten für den Studieneinschluss die klinische Diagnose der TSC sowie eine gut dokumentierte klinische Vorgeschichte der Epilepsie aufweisen, die durch die derzeit eingenommenen Antiepileptika nicht vollständig kontrolliert war. Zusätzlich war die Einnahme von einem oder mehreren Antiepileptika in einer Dosis erforderlich, die vor dem Screening mindestens 4 Wochen lang stabil war. Am Ende des Baseline-Zeitraums mussten die Testpersonen mindestens 8 TSC-assoziierte Anfälle in den ersten 28 Tagen erlitten haben. Fast alle Studienzentren befindet sich in den

USA und Europa. Von einer Übertragbarkeit der Zulassungspopulation auf den deutschen Versorgungskontext ist auszugehen.

Die Eingrenzung auf 8 TSC-assoziierte Anfälle innerhalb von 28 Tagen findet sich nicht im zugelassenen Anwendungsgebiet von Cannabidiol wieder. Angaben dazu, welcher Anteil der an TSC erkrankten Personen mit Krampfanfällen unter einer vergleichbaren Symptomatik trotz Einnahme antiepileptischer Therapie leidet, konnten nicht identifiziert werden. Daher ist unklar, inwiefern es sich bei der Studienpopulation um stärker betroffene Patientinnen und Patienten gegenüber der Zulassungspopulation handelt. Weiterhin besteht die Zulassung von Cannabidiol für Personen ab 2 Jahren, wohingegen in die Zulassungsstudie auch Kinder im Alter von 1 Jahr eingeschlossen waren und Erwachsene bis maximal 65 Jahren berücksichtigt wurden. Daten dazu, wie viele Kinder im Alter von 1 Jahr in die Studie eingeschlossen wurden, liegen nicht vor. Die Altersspanne von 18 bis 65 Jahren umfasste insgesamt nur etwas mehr als 25 % der Testpersonen.

4.2 Design und Methodik der Studie

Bei der Studie GWEP1521 handelt es sich um eine doppelblinde, placebokontrollierte, multizentrische, randomisierte Phase-III-Studie mit 4 Behandlungsarmen. Untersucht wurden Wirksamkeit und Sicherheit von Cannabidiol 25 mg/kg/Tag und 50 mg/kg/Tag als adjuvante antiepileptische Behandlung im Vergleich zu Placebo. Insgesamt 224 Patientinnen und Patienten wurden im Verhältnis 2:2:1:1 auf die Studienarme Cannabidiol 25 mg/kg/Tag, Cannabidiol 50 mg/kg/Tag, Placebo 25 mg/kg/Tag bzw. 50 mg/kg/Tag randomisiert. Primäres Zielkriterium der Studie war die Reduktion der Häufigkeit TSC-assoziiierter Krampfanfälle während der Behandlungsperiode im Vergleich zur Baseline-Periode. Für die Auswertungen wurden beide Placebo-Arme zusammengefasst. Die Randomisierung erfolgte stratifiziert nach Altersgruppe (1–6, 7–11, 12–17, ≥ 18 Jahre). An das Screening und die Baseline-Visite schloss sich eine 28-tägige Baseline-Periode an, in der die Studienteilnehmenden mindestens 8 TSC-assoziierte Anfälle erleiden mussten. Nach Randomisierung folgte eine 4-wöchige Titrationsphase, in der die Studienteilnehmenden auf die gewünschte Dosis des Prüf- oder Kontrollpräparats auftitriert wurden, mit der sie in der 12-wöchigen Erhaltungsperiode behandelt wurden. Am Ende der 16-wöchigen Studienmedikation war es für die Studienteilnehmenden möglich die Behandlung in einer offenen Extensionsstudie (OLE) fortzusetzen. Für alle Personen, die nicht direkt in die OLE übergingen, erfolgte eine schrittweise Reduktion (Ausschleichen) der Dosis über 10 Tage (10 % pro Tag) und eine 4-wöchige Nachbeobachtung. Im Fall einer Teilnahme an der OLE wurden alle Testpersonen in einer 2-wöchigen, verblindeten Titrationsperiode zunächst auf eine Dosis von Cannabidiol 25 mg/kg/Tag titriert, die anschließend bis zu einer maximalen Dosis auf 50 mg/kg/Tag eskaliert werden konnte, um die optimale Dosis für jede Testperson zu finden.

Die doppelt-verblindete Studienphase beinhaltete eine Behandlungsphase von insgesamt 16 Wochen. Eine Behandlungsdauer von 16 Wochen bei einer langfristigen, meist lebenslangen Erkrankung wie der TSC wird als kritisch angesehen, insbesondere vor dem Hintergrund der Auswirkungen auf die Nachhaltigkeit eines Therapieeffekts und die Langzeitsicherheit. Der OLE liegt zwar eine längere Beobachtungszeit zugrunde, diese Studienphase ist jedoch nicht vergleichend und setzt Cannabidiol bis zu einer nicht zulassungskonformen Dosierung von 50 mg/kg/Tag ein. Die Studie ist laufend; eine Zwischenanalyse für die zulassungskonform behandelten Personen liegt nicht vor.

Wie bereits beschrieben ist für die Nutzenbewertung nur der Vergleich der Behandlungsarme Cannabidiol 25 mg/kg/Tag und Placebo in der verblindeten Studienphase relevant. ITT- und Sicherheitspopulation der Studie GWEP1521 sind identisch und bestehen aus 75 Personen des Behandlungsarms Cannabidiol 25 mg/kg/Tag und 76 im gepoolten Placebo-Arm. Unter Cannabidiol brachen mit 10 Personen (13 %) deutlich mehr Personen die Einnahme der Studienmedikation im Vergleich zu Placebo ($n=1$ (1 %)) ab. Hauptgrund war das Auftreten von UE (8 Personen, 11 %) im Interventionsarm. Abgesehen von 1 Person je Behandlungsarm entschieden sich alle behandelten Personen, die die Behandlungsphase abgeschlossen haben, für eine Teilnahme an der OLE.

Gemessen am Median ist die Behandlungsdauer mit 113 Tagen zwischen den Behandlungsarmen vergleichbar. Die mittlere Expositionsdauer lag mit 102 Tagen im Interventionsarm etwas unter der der Placebo-Gruppe mit 113 Tagen. Die mediane Beobachtungsdauer in Modul 4 des Nutzen dossiers ist mit den gleichen Werten (Median / min; max) wie die Beobachtungsdauer angegeben, wobei der Bezug unklar ist. Diese Daten konnten anhand des Studienberichts nicht verifiziert werden. Bezogen auf alle Tage, die sich die Testpersonen in der Behandlungsphase befanden, zeigt sich mit 96 % unter Cannabidiol 25 mg/kg/Tag und 100 % unter Placebo eine hohe Behandlungcompliance mit der Studienmedikation.

Die Charakteristika der randomisierten Studienpopulation zu Studienbeginn bzw. zu Baseline waren insgesamt balanciert zwischen den Behandlungsarmen. Es zeigt sich auch eine Gleichverteilung zu Baseline innerhalb der Altersgruppen (1–6, 7–11, 12–17, ≥ 18 Jahre). Mit einem Anteil von fast 75 % der Patientenpopulation im Alter von 17 Jahren oder jünger und einem medianen Alter von 11,6 bzw. 10,9 Jahren war ein junges Patientenkontinuum in der Studie eingeschlossen. Angaben zur Krankheitsdauer zu Studienbeginn liegen nicht vor. Geringfügige Abweichungen von > 10 Prozentpunkten zwischen den Behandlungsarmen zeigten sich bei einigen Begleiterkrankungen wie „Entwicklungsverzögerungen“ und „Angst“ und dem Anteil der Personen mit tonischen und tonisch-klonischen Anfällen während der Baseline-Periode, die vermehrt in der Interventionsgruppe auftraten. Bei den antiepileptischen Begleitmedikationen wurde Vigabatrin häufiger in der Interventionsgruppe (37 vs. 22 %) eingesetzt und Clobazam etwas häufiger in der Kontrollgruppe (23 vs. 33 %). Durch die heterogene Verteilung lassen sich jedoch keine Konsequenzen für die Nutzenbewertung ableiten. Im Median nahmen die Testpersonen 3 antiepileptische Begleitmedikationen zu Studienbeginn ein. Notfallmedikationen wurden jeweils von ca. einem Drittel der Studienteilnehmenden beansprucht. Neben den antiepileptischen Begleitmedikationen mussten auch die nicht-pharmakologischen Therapien (ketogene Diät, Vagusnerv-Stimulation) mindestens 4 Wochen vor dem Screening und während des gesamten Studienzeitraums stabil gehalten werden. Inwiefern Therapieoptimierungen im Studienverlauf notwendig waren, ist nicht beschrieben. Protokollverletzungen bezüglich Therapieanpassungen von antiepileptischen Begleitmedikationen sind lediglich für 4 Personen des Interventions- und 1 Person des Kontrollarms angegeben. Es sind Wechselwirkungen von Cannabidiol mit weiteren antiepileptischen Arzneimitteln möglich [37].

Für Studie GWEP1521 sind insgesamt 7 Protokolländerungen beschreiben, davon 6 nach Einschluss der ersten Person. Dabei wurde die statistische Analyse des primären Endpunkts und der primären Auswertungen der anderen Anfallsendpunkte (Anfälle gesamt, weitere Anfälle) mehrfach geändert, von einer ANCOVA zu einem Wilcoxon-Rangsummentest und final auf eine negativ-binomiale Regressionsanalyse. Da alle Amendments vor Entblindung der Datenbank (27. April 2019) implementiert wurden und die genannten Analysen als Sensitivitätsanalysen für den primären Endpunkt vorliegen, wird der Einfluss dieser Änderung auf die Nutzenbewertung als gering eingestuft.

Erst mit Amendment 4 (7. August 2018) wurde die „Einnahme von einem oder mehreren Antiepileptika in einer Dosis, die vor dem Screening mindestens 4 Wochen lang stabil war“ als Einschlusskriterium definiert. Es ist unklar, wie viele Personen zu diesem Zeitpunkt bereits in die Studie eingeschlossen wurden. Aus den Charakteristika der Patientenpopulation (siehe Tabelle 12) geht jedoch hervor, dass abgesehen von einer Person aus dem Behandlungsarm Cannabidiol 25 mg/kg/Tag alle Studienteilnehmenden mindestens ein weiteres Antiepileptikum einnahmen; diese Person wurde zudem nach der Änderung der Protokollversion in die Studie eingeschlossen.

Für das Nutzendossier führt der pU Post-hoc-Auswertungen aller in Modul 4 berichteten Endpunkte durch. Die angewendeten Modelle sind zum Teil nicht vollständig nachvollziehbar beschrieben. Abgesehen von den binär ausgewerteten Endpunkten und der Sensitivitätsanalyse des primären Endpunkts über den Wilcoxon-Rangsummentest, weichen Effektschätzer und p-Wert aller weiteren Endpunkte aus dem Studienbericht leicht von den Post-hoc-Auswertungen für das Nutzendossier ab. Der pU legt lediglich für die Endpunkte „QOLCE“, „QOLIE-31-P“ und „Vineland-II“ eine Begründung für die Abweichung vor. Danach resultieren die Abweichungen zu den im Studienbericht beschriebenen Ergebnissen daraus, dass in den vorliegenden Analysen nur die beiden Studienarme Cannabidiol 25 mg/kg/Tag und Placebo eingeschlossen wurden, nicht aber die Patientinnen und Patienten des Studienarms mit der nicht zugelassenen Dosierung Cannabidiol 50 mg/kg/Tag. Eine Klärung seitens des pU, inwiefern diese Begründung auch für die geringfügigen Abweichungen der Ergebnisse der anderen Endpunkte zutrifft, ist wünschenswert. Für das Nutzendossier stellt der pU bei den dichotomen Auswertungen (Responderanalysen, Personen mit Status epilepticus und Auswertungen der UE) neben dem stratifizierten CMH-Test post hoc RR ohne Beschreibung des zugrundeliegenden Modells und der eingehenden Stratifikationsfaktoren dar. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen. Eine Klarstellung ist dennoch wünschenswert.

Sowohl die erste Version des SAP (17.04.2019) als auch die finale Version (24.04.2019) wurden erst nach Studienabschluss (15.02.2019) erstellt, jedoch vor Entblindung der Datenbank (27.04.2019). In einem weiteren Addendum zum SAP (28.11.2019) nach Entblindung der Studie sind weitere Analysen beschrieben, insbesondere um die Robustheit der Ergebnisse aufgrund vieler Therapieabbrüche in den Interventionsgruppen zu prüfen. Post hoc durchgeführte Analysen für das Nutzendossier analog zu den weiteren Endpunkten liegen jedoch nicht vor.

In den Studienunterlagen wurden Subgruppenanalysen für den primären Endpunkt und den sekundären Endpunkt „Reduktion der Häufigkeit TSC-assoziiierter Anfälle um $\leq 50\%$ im Vergleich zu Baseline“ geplant und durchgeführt. Untersucht wurden unter anderem Alter, Geschlecht, Region und Clobazam als Begleitmedikation. Für die weiteren Endpunkte existieren keine Subgruppenanalysen. Allgemein liegen keine Auswertungen getrennt für Kinder und Erwachsene vor. Inwiefern tatsächlich keine Unterschiede in der Wirksamkeit und Sicherheit von Cannabidiol zwischen Kindern und Erwachsenen bestehen, kann somit nicht abschließend beurteilt werden.

Das Verzerrungspotential auf Studienebene wird für Studie GWEP1521 insgesamt mit gering bewertet. Aufgrund der Abweichungen von den Dosierungsempfehlungen der Fachinformation ist die Aussagesicherheit der Ergebnisse dennoch eingeschränkt.

4.3 Mortalität

Todesfälle wurde in der Studie GWEP1521 im Rahmen der Sicherheitserhebung erfasst. Es traten keine Todesfälle im Studienverlauf auf.

Zusammenfassende Einschätzung zur Wirksamkeit

Das Verzerrungspotential des Endpunkts wird als gering eingeschätzt. Dennoch erscheint eine abschließende Bewertung der Mortalität aufgrund der kurzen Behandlungsdauer in der Studie GWEP1521 nicht möglich.

4.4 Morbidität

Epileptische Anfälle

Die Erhebung epileptischer Anfälle erfolgte täglich über ein elektronisches IVRS, entweder durch die Testperson selbst oder ihre betreuende Person. Zu Baseline erhielten die betreuenden Personen bzw. die Patientinnen und Patienten ein Training zur Identifikation der bei der Testperson zu erwartenden Anfallstypen und zum richtigen Umgang mit dem IVRS. Zur Verifizierung der Anfallstypen wurde zudem ein unabhängiges ESC eingesetzt. Im Wesentlichen aufgrund des engen Patienten-Betreuer-Verhältnisses wird davon ausgegangen, dass betreuende Personen Krampfanfälle valide erfassen können. Andererseits ist fraglich, wie valide eine Erfassung der Anfälle bzw. der Anfallsdauer durch die Testperson selbst ist. Es konnten keine Informationen darüber identifiziert werden, wie viele Testpersonen die Anfälle selbst bewerteten und berichteten bzw. welche Informationen im Fall eines Anfalls im IVRS dokumentiert wurden. Insgesamt wird jedoch von einer validen Erfassung epileptischer Anfälle ausgegangen.

Die Vollständigkeit der Eintragungen im IVRS war hoch, bezieht sich jedoch offensichtlich auf die Tage, an denen sich die Testperson in der Behandlungsphase befand. Fehlende Werte aufgrund von Therapieabbrüchen sind nicht abzuschätzen. Dabei brachen mit 13 % deutlich mehr Personen in der Interventionsgruppe die Einnahme der Studienmedikation ab, im Vergleich zu Placebo (1 %). Für Personen, die die Studie vorzeitig abbrachen, wurde die primäre Analyse auf Basis der bis zum Studienabbruch verfügbaren Daten durchgeführt.

TSC-assozierte Anfälle

Primärer Endpunkt der Studie GWEP1521 war die „Veränderung der Häufigkeit TSC-assoziierter Anfälle während der Behandlungsperiode im Vergleich zu Baseline“. Unter Cannabidiol 25 mg/kg/Tag verringerte sich die Häufigkeit TSC-assoziierter Anfälle im Median um 43 %, verglichen mit 20 % in der Kontrollgruppe. In der Analyse mittels negativ-binomiale Modell zeigte sich ein statistisch signifikanter Vorteil unter Cannabidiol im Vergleich zu Placebo (0,70; 95%-KI: [0,58; 0,85]; $p = 0,0003$). Für den primären Endpunkt war eine Reihe von Sensitivitätsanalysen geplant, die den statistisch signifikanten Vorteil von Cannabidiol gegenüber Placebo bestätigen. Zudem zeigten sich keine statistisch signifikanten Interaktionen bei den Subgruppenanalysen.

In den Responderanalysen liegt ebenfalls ein statistisch signifikanter Effekt zugunsten von Cannabidiol vor, jedoch lediglich für eine Reduktion der Anfallshäufigkeit um ≥ 75 % (RR: 25,33 (95%-KI: [1,53; 420,24]; $p = 0,0003$)). Für Reduktionen um ≥ 25 und ≥ 50 % stellten sich dagegen keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen ein, es bestehen jedoch numerische Vorteile. Nur 1 Person im Interventionsarm erreichte eine vollständige Freiheit

von TSC-assoziierten Anfällen. Statistische Auswertungen für einen Anstieg oder keine Veränderung in der Anfallshäufigkeit ($\geq 0\%$), die bei 17 Personen (23 %) des Interventions- und 21 (28 %) des Kontrollarms auftrat, liegen nicht vor. Im Endpunkt „Reduktion der TSC-assoziierten Anfälle um $\geq 50\%$ “ zeigte sich lediglich für das Merkmal „Region“ eine statistisch signifikante Interaktion ($p = 0,0205$), die mangels Bestätigung im primären Endpunkt und mangels Subgruppenanalysen für weitere Endpunkte jedoch von untergeordneter Relevanz ist.

Das Verzerrungspotential für den Endpunkt wird aufgrund der deutlich höheren Anzahl an Therapieabbrüchen im Interventionsarm mit hoch bewertet. Da die primäre Analyse auf Basis der bis zu dem Zeitpunkt des Studienabbruchs verfügbaren Anfallsdaten durchgeführt wird und nach dem Ende der Behandlung mit Cannabidiol eine Veränderung, insbesondere auch eine Steigerung in der Anfallshäufigkeit, plausibel ist, kann eine Verzerrung zugunsten von Cannabidiol nicht ausgeschlossen werden.

Anfälle gesamt

Bezogen auf die Gesamtzahl der in der Studie GWEP1521 aufgetretenen Anfälle zeigten sich vergleichbare Ergebnisse zu den TSC-assoziierten Anfällen. Die Veränderung der Häufigkeit der Anfälle gesamt während der Behandlungsperiode im Vergleich zur Baseline-Periode mittels negativ-binomialen Modell war statistisch signifikant zugunsten von Cannabidiol (0,71 (95%-KI: [0,58; 0,86]; $p = 0,0013$)). Auch die Responderanalyse zur Reduktion der Anfälle gesamt um $\geq 75\%$ zeigte einen statistisch signifikanten Effekt zugunsten von Cannabidiol (12,07 (95%-KI: [1,61; 90,44]; $p = 0,0007$)). Für die weiteren Responder-Kriterien sowie die Anfallsfreiheit zeigten sich jedoch keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen. Statistische Auswertungen für einen Anstieg oder keine Veränderung in der Anfallshäufigkeit liegen auch für die Anfälle gesamt nicht vor.

Das Verzerrungspotential für den Endpunkt wird aufgrund der deutlich höheren Anzahl an Therapieabbrüchen im Interventionsarm mit hoch bewertet. Da die primäre Analyse auf Basis der bis zu dem Zeitpunkt des Studienabbruchs verfügbaren Anfallsdaten durchgeführt wird und nach dem Ende der Behandlung mit Cannabidiol eine Veränderung, insbesondere auch eine Steigerung in der Anfallshäufigkeit, plausibel ist, kann eine Verzerrung zugunsten von Cannabidiol nicht ausgeschlossen werden.

Weitere Anfälle

Lediglich 12 Personen (16 %) der Interventions- und 15 (20 %) der Kontrollgruppe erlitten während der Baseline-Periode mindestens einen weiteren Anfallstyp. Bei der Veränderung der Häufigkeit von „Weitere Anfälle“ zeigte sich kein statistisch signifikanter Effekt zwischen den Behandlungsarmen. Aufgrund der geringen Prävalenz von „Weitere Anfälle“ während der Baseline-Periode (vgl. Tabelle 12) wird auf eine Darstellung der Responderanalysen verzichtet.

Das Verzerrungspotential für den Endpunkt wird in Folge der deutlich höheren Anzahl an Therapieabbrüchen im Interventionsarm mit hoch bewertet. Die Aussagekraft des Endpunkts ist, auch aufgrund der geringen Prävalenz von „Weitere Anfälle“, in der Studienpopulation zusätzlich eingeschränkt.

Status epilepticus

Ausgewertet wurde ausschließlich der Anteil der Personen mit mindestens einer Status-epilepticus-Episode während der Behandlungsperiode. Falls eine Person mehrere Status-epilepticus-Episoden erlitten hatte, findet dies keine Berücksichtigung. Die Auswertung wurde post hoc mittels CMH-Test

durchgeführt. Es liegen keine Angaben vor, ob eine Stratifizierung der Analyse nach Altersgruppen erfolgte. Unter der Behandlung mit Cannabidiol 25 mg/kg/Tag trat bei 5 Personen mindestens eine Status-epilepticus-Episode auf, im Kontrollarm bei 7 Personen. Der Effekt ist statistisch nicht signifikant. Das Verzerrungspotential für den Endpunkt wird aufgrund der deutlich höheren Anzahl an Therapieabbrüchen im Interventionsarm mit hoch bewertet. Da die primäre Analyse auf Basis der bis zu dem Zeitpunkt des Studienabbruchs verfügbaren Anfallsdaten durchgeführt wird und nach dem Ende der Behandlung mit Cannabidiol eine Veränderung, insbesondere auch eine Steigerung in der Anfallshäufigkeit und -dauer, plausibel ist, kann eine Verzerrung zugunsten von Cannabidiol nicht ausgeschlossen werden.

Globaler Eindruck der Veränderung (mittels CGIC / SGIC)

Der globale Eindruck der Veränderung sollte in der Studie GWEP1521 mittels CGIC – durch die betreuende Person – und/oder mittels SGIC – durch die Testperson selbst – bewertet werden. Im SAP wurde bereits von sehr wenigen Rückläufen im SGIC ausgegangen, weshalb keine eigenständigen Analysen für den SGIC geplant waren.

Beim CGIC handelt es sich um eine Fremdeinschätzung; dies erscheint jedoch im Kontext der Schwere der TSC, mit vielfach kognitiven Beeinträchtigungen und einer verzögerten Entwicklung, in Verbindung mit dem teilweise sehr jungen Alter der Studienteilnehmenden bei einem Teil der Studienpopulation adäquat. Allerdings ist nicht beschrieben, ob es für das Studienprozedere und die Studiendurchführung selbst Vorgaben gab, wann eine Fremdeinschätzung angemessener als eine Selbsteinschätzung anzusehen ist und wie viele Testpersonen davon betroffen waren; oder, ob dies allein von den Testpersonen bzw. behandelnden Personen zu entscheiden war. Primär war eine gemeinsame Auswertung von CGIC und SGIC vorgesehen, wobei der CGIC priorisiert wurde, falls sowohl CGIC als auch SGIC für eine Testperson vorlagen. Prinzipiell wird eine gemeinsame Auswertung von CGIC und SGIC kritisch gesehen, da sich Eigen- und Fremdbewertung strukturell unterscheiden können; getrennte Auswertungen sind wünschenswert. Im Vergleich zu separaten Auswertungen des CGIC gehen bei der gemeinsamen Auswertung je Behandlungsgruppe lediglich 4 Werte des SGIC zusätzlich ein, wovon jeweils 3 das Responder-Kriterium für eine Verbesserung erfüllten. Von einer starken Beeinflussung der Ergebnisse ist nicht auszugehen.

Präspezifiziert in den Studienunterlagen war eine Analyse mittels ordinal logistischer Regression. Im Nutzendossier präsentiert der pU zusätzlich dazu Responderanalysen für eine Verbesserung im globalen Eindruck der Veränderung. Für die Nutzenbewertung werden primär die Responderanalysen herangezogen. Responderanalysen, die ausschließlich den CGIC berücksichtigen, liegen nicht vor. Die vom pU dargestellten Responderanalysen zum CGIC / SGIC beziehen sich ausschließlich auf Testpersonen mit einem Rücklauf. Zum Behandlungsende erreichten mehr Personen unter Cannabidiol 25 mg/kg/Tag (48 von 70) als unter Placebo (30 von 76) eine sehr starke, starke oder geringfügige Verbesserung (RR: 1,74 (95%-KI: [1,26; 2,39]; $p = 0,0004$)) im CGIC / SGIC. Fehlende Werte wurden aus der Analyse ausgeschlossen, was auf 5 Personen der Interventionsgruppe zutrifft. Auswertungen, in denen die fehlenden Werte als Non-Responder eingehen, liegen nicht vor. Responderanalysen zu einer möglichen Verschlechterung des globalen Eindrucks der Veränderung liegen ebenfalls nicht vor; aus den Daten geht jedoch hervor, dass sich 7 Personen (9,3 %) der Interventions- und 4 (5,3 %) der Kontrollgruppe im CGIC im Laufe der Studie verschlechterten. Bei der Interpretation der Responderanalysen ist zu berücksichtigen, dass sich mehr als die Hälfte der Kontrollgruppe weder verschlechtert, noch verbessert hatte (vgl. Tabelle 21).

Aufgrund der Einschränkungen bei den Responderanalysen und der A-priori-Definition werden ergänzend die Auswertungen des globalen Eindrucks der Veränderung mittels ordinal logistischer Regression berücksichtigt. Dabei zeigte sich sowohl für die separate Auswertung des CGIC ($p = 0,0043$) als auch die kombinierte Auswertung von CGIC und SGIC ($p = 0,0040$) ein vergleichbarer, statistisch signifikanter Effekt zugunsten von Cannabidiol. Einschränkend ist zu erwähnen, dass bei beiden Auswertungen die Proportional-Odds-Annahme verletzt war. Der in diesem Fall gemäß SAP zusätzlich durchzuführende „Cochran-Armitage Trend Test“ zeigte jedoch ebenfalls einen statistisch signifikanten Effekt zugunsten von Cannabidiol, der jedoch etwas geringer ausfällt.

Aufgrund der teilweise post hoc durchgeführten, selektiven Analysen wird das Verzerrungspotential des Endpunkts insgesamt mit hoch bewertet.

Bewertung des adaptiven Verhaltens (mittels Vineland-II)

Die Bewertung des adaptiven Verhaltens erfolgte mittels Vineland-II, einem komplexen Instrument zur Erfassung persönlicher und sozialer Fähigkeiten von Kindern oder Erwachsenen in alltäglichen Situationen; insbesondere bei Personen, die geistig zurückgeblieben sind oder mit Schwierigkeiten andere Testsituationen zu bestehen. Entgegen der Einstufung durch den pU als Instrument zur Erfassung der Lebensqualität wird die Vineland-II als Instrument zur Erfassung der Morbidität eingestuft.

Insbesondere aufgrund der geringen Rücklaufquoten für den Gesamtwert des adaptiven Verhaltens von lediglich 57 % in der Behandlungsgruppe Cannabidiol 25 mg/kg/Tag und 62 % in der Placebo-Gruppe, aber auch aufgrund der mangelhaften Rücklaufquoten in den Subdomänen, wird von einer Darstellung der Ergebnisse der Vineland-II abgesehen.

Achenbach-Verhaltens-Checkliste

Verhaltens- und emotionale Symptome sollten in der Studie GWEP1521 über drei unterschiedliche Verhaltens-Checklisten erfasst werden, die in Abhängigkeit vom Alter der Testpersonen eingesetzt wurden; „CBCL (1,5–5 Jahre)“ und „CBCL (6–18 Jahre)“ für Kinder und „ABCL (18–59 Jahre)“ für Erwachsene. Aufgrund der stratifizierten Randomisierung nach Altersgruppen erscheint der altersabhängige Einsatz verschiedener Erhebungsinstrumente prinzipiell adäquat. Da jedoch keine gemeinsame Auswertung der Checklisten über die gesamte Studienpopulation möglich ist, resultieren daraus 3 Subpopulationen, in denen die jeweilige Checkliste eingesetzt wird. Neben einem potentiellen Verlust an statistischer Power sind die Ergebnisse nur für den Teil der Patientenpopulation aussagekräftig, die sich im Altersspektrum der jeweiligen Checkliste befinden. Entgegen der Einstufung durch den pU als Instrumente zur Erfassung der Lebensqualität werden die Achenbach-Verhaltens-Checklisten als Instrumente zur Erfassung der Morbidität eingestuft.

Die Rücklaufquoten bezogen auf das Altersspektrum, in dem die jeweilige Checkliste eingesetzt wurde, lagen für alle Subskalen und Gesamtwerte über 70 %. Bei einem geringen Verzerrungspotential zeigten sich jedoch keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Fraglich ist allgemein, ob relevante Veränderungen bei Verhaltenssymptomen innerhalb der relativ kurzen Behandlungsphase von 16 Wochen überhaupt erfassbar sind oder ob hierfür längere Beobachtungsphase erforderlich sind.

Zusammenfassende Einschätzung zur Morbidität

Für den Endpunkt „Epileptische Anfälle“ zeigten sich bei hohem Verzerrungspotential homogen statistisch signifikante Vorteile für Cannabidiol gegenüber der Kontrollgruppe für die Reduktion der Häufigkeit TSC-assoziiierter Anfälle und eine Anfallsreduktion um mindestens 75 %. Auch für die weiter gefassten „Anfälle gesamt“ zeigten sich Vorteile von Cannabidiol bei der Anfallshäufigkeit und einer Reduktion um mindestens 75 % im Vergleich zu Baseline.

Bei einem hohen Verzerrungspotential für den Endpunkt „Globaler Eindruck der Veränderung“ liegen sowohl für die Responderanalysen als auch für die Auswertung mittels ordinal logistischer Regression statistisch signifikante Vorteile von Cannabidiol im Vergleich zur Placebo-Gruppe vor. Für alle weiteren Endpunkte zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen, oder die Ergebnisse konnten aufgrund von Limitationen nicht für die Nutzenbewertung herangezogen werden.

Aufgrund der kurzen randomisierten Studienphase ist die Ergebnissicherheit allgemein limitiert.

4.5 Lebensqualität

Da der QOLIE-31-P in seiner in Studie GWEP1521 eingesetzten Form mit gewichteten Skalen und einem gewichteten Mittelwert als nicht validiert eingestuft wurde, erfolgt die Bewertung der Lebensqualität nur für einen Teil der Patientenpopulation der Studie mittels QOLCE. Dieser sollte gemäß Studienprotokoll ausschließlich bei Personen im Alter von 2 bis 18 Jahren eingesetzt werden. In Studie GWEP1521 liegen, entgegen der geplanten Erhebungen im Protokoll, Rückläufe bei einigen Patientinnen und Patienten außerhalb des vorgesehenen Altersspektrums vor. Es ist unklar, ob eine Selektion dieser Patientinnen und Patienten erfolgte und falls ja, nach welchen Kriterien. Daher sind für die Nutzenbewertung lediglich die Ergebnisse der Testpersonen im Alter von 2 bis einschließlich 18 Jahren relevant, für die der Einsatz des Instruments a priori geplant war. Diese Altersgruppe wurde gemäß Studienunterlagen als Sensitivitätsanalyse berücksichtigt.

In Modul 4 präsentiert der pU post hoc dagegen Analysen für alle Testpersonen bis einschließlich 18 Jahren. Auf Basis der Daten aus Modul 4 liegen die Rücklaufquoten in Bezug auf die Population der Studie im Alter von 18 Jahren und jünger für den Gesamtwert des QOLCE bei 64 % unter Cannabidiol und 77 % unter Placebo. Innerhalb der Subskalen zeigten sich mit 40–96 % unter Cannabidiol und 36–96 % im Placebo-Arm stark unterschiedliche Rücklaufquoten. Anhand der unterschiedlichen Rücklaufquoten stellt sich die Frage, ob einzelne Subskalen bzw. das gesamte Instrument für alle Testpersonen eine vergleichbare Relevanz besitzen. Hinzu kommen Unterschiede von mehr als 10 % in den Rücklaufquoten auf Ebene des Gesamtwerts zwischen den Behandlungsarmen. Auch die Diskrepanzen zwischen den Auswertungen aus Modul 4 und den a priori geplanten Auswertungen im Studienbericht können nicht abschließend erklärt werden (siehe Kapitel 3.4); bzw. ist fraglich, ob diese Diskrepanzen durch den Einbezug von Personen im Alter von 1 Jahr erklärt werden können. Bei dieser Patientenpopulation war der Einsatz des QOLCE jedoch nicht vorgesehen. Eine Aufklärung dieser Diskrepanzen wäre wünschenswert.

Insgesamt können die Auswertungen nicht für die Nutzenbewertung herangezogen werden.

Zusammenfassende Einschätzung zur Lebensqualität

Es liegen keine verwertbaren Daten zur Lebensqualität vor.

4.6 Sicherheit

In der Studie GWEP1521 entspricht die ITT-Population der Sicherheitspopulation. Dementsprechend haben alle Testpersonen mindestens eine Dosis der Studienmedikation erhalten, auf die sie randomisiert wurden. Die mediane Behandlungsdauer der Behandlungsarme war mit 113 Tagen gleich. Aufgrund der vermehrten Behandlungsabbrüche in der Interventionsgruppe lag die mittlere Expositionsdauer mit 102 Tagen etwas unter der der Placebo-Gruppe mit 113 Tagen. Daten für die Beobachtungsdauer bezüglich UE konnten nicht identifiziert werden.

Unerwünschte Ereignisse

Das Auftreten von UE wurde in Studie GWEP1521 ab der Screening-Visite (Studienvisite 1) erhoben. Für die Auswertung werden, mit den TEAE, alle UE berücksichtigt, die während der verblindeten Studienphase auftraten; einschließlich dem Tag der Einnahme der 1. Dosis in der OLE. Inwiefern Ereignisse von Therapieabbrüchern aus der 10-tägigen Ausschleichphase und dem 4-wöchigen Sicherheits-Follow-up nach der Studienvisite zu Behandlungsende als behandlungsbezogene UE gewertet wurden, ist nicht beschrieben. Eine statistische Auswertung der UE war in den Studienunterlagen nicht geplant. Für das Nutzendossier erfolgte post hoc eine Analyse der Unterschiede zwischen den Behandlungsgruppen mittels CMH-Test, stratifiziert nach Altersgruppen. Zusätzlich dazu stellt der pU RR ohne Beschreibung des zugrundeliegenden Modells oder der eingehenden Stratifikationsfaktoren dar. Aufgrund der Darstellung in Modul 4 des Nutzendossiers wird von einer Stratifizierung nach Altersgruppen ausgegangen. Eine Klarstellung ist dennoch wünschenswert.

Mit 93 % im Behandlungsarm Cannabidiol 25 mg/kg/Tag und 95 % im Placebo-Arm, erlitten fast alle Personen mindestens ein UE; UE nach Schweregrad wurden aufgrund einer nicht standardisierten Erhebung nicht berücksichtigt. SUE traten statistisch signifikant häufiger (RR: 8,22 (95%-KI: [1,96; 34,45]; $p=0,0004$)) in der Interventionsgruppe ($n=16$; 21 %) vergleichen mit der Kontrollgruppe ($n=2$; 3 %) auf. Auch bei den UE, die zum Therapieabbruch führten, zeigte sich ein statistisch signifikanter Effekt (RR: 4,04 (95%-KI: [0,90; 18,20]; $p=0,0484$)) zuungunsten der Interventionsgruppe, wobei 8 Personen (11 %) der Behandlungsgruppe Cannabidiol 25 mg/kg/Tag die Therapie vorzeitig abbrachen, verglichen mit 2 Personen (3 %) der Placebo-Gruppe.

In der Auswertung der UE mit Inzidenz ≥ 10 % nach MedDRA-Systemorganklasse und Preferred Term zeigte sich auf Ebene der Systemorganklasse lediglich für „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (RR: 2,02 (95%-KI: [1,06; 3,85]; $p=0,0404$)) und „Untersuchungen“ (2,77 (95%-KI: [1,50; 5,12]; $p=0,0005$)) ein statistisch signifikanter Effekt zwischen den Vergleichsarmen zuungunsten von Cannabidiol 25 mg/kg/Tag. Weitere statistisch signifikante Effekte zuungunsten der Intervention traten bei den Preferred Terms „Fieber“ (RR: 2,51 (95%-KI: [1,05; 5,99]; $p=0,0428$)), „Gamma-Glutamyltransferase erhöht“ (RR: 25,33 (95%-KI: [1,53; 420,24]; $p=0,0003$)), „Alanin-Aminotransferase erhöht“ (RR: 19,25 (95%-KI: [1,14; 324,93]; $p=0,0021$)) und „Aspartat-Aminotransferase erhöht“ (RR: 17,22 (95%-KI: [1,01; 293,18]; $p=0,0039$)) auf. UE der Preferred Terms der erhöhten Leberwerte traten ausschließlich in der Interventionsgruppe auf. Bei den SUE nach Systemorganklassen liegt kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen vor. Für UE nach Systemorganklasse und Preferred Term, die zum Therapieabbruch führten, zeigte sich lediglich bei der Systemorganklasse „Erkrankungen der Haut und des Unterhautgewebes“ ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen ($p=0,0429$), der zuungunsten von Cannabidiol ausfiel. Aufgrund der im Studiendesign vorgesehenen fixen Dosierung der Prüfmedikation am oberen Ende des zugelassenen Dosisspektrums in

der Studie GWEP1521, kann eine mögliche Überschätzung der UE im Cannabidiol-Arm gegenüber einer fachinformationskonformen Dosistitration nach individuellem Risiko und Nutzen nicht ausgeschlossen werden. Mangels Daten von Testpersonen in einer niedrigeren Dosierung kann dies jedoch nicht quantifiziert werden.

Im EPAR wird ein klarer dosisabhängiger Anstieg im Auftreten von behandlungsbezogenen UE, SUE und UE, die zum Therapieabbruch führten beschrieben, unabhängig davon, ob die UE vom Studienpersonal als behandlungsbezogen gewertet wurden [15]. Dieser dosisbezogene Anstieg zeigte sich auch bei der Erhöhung von Alanin- und Aspartat-Aminotransferase bei Patientinnen und Patienten mit TSC, wobei in einigen Fällen die Konsenskriterien medikamenteninduzierter Leberschädigungen erfüllt waren. Dazu wurde in die Fachinformation auch ein Warnhinweis vor hepatozellulären Schädigungen aufgenommen [22]. Weiterhin warnt die Fachinformation allgemein vor Suizidgedanken und suizidalem Verhalten, für die sich ein leicht erhöhtes Risiko in Studien anderer Antiepileptika zeigte. Gleichzeitig wurde ein Gewichtsverlust bzw. eine verminderte Gewichtszunahme beobachtet, die scheinbar dosisbezogen ist. Da Personen mit klinisch signifikanten kardiovaskulären Beeinträchtigungen nicht in die Studie aufgenommen wurden, können keine Aussagen zu dieser Population getroffen werden [22].

Suizidalität (mittels C-SSRS)

Die Abfrage der C-SSRS erfolgt lediglich bei Personen, die 6 Jahre oder älter sind und nach Ansicht des ärztlichen Prüfpersonals in der Lage waren die Fragen zu verstehen und zu beantworten. Bei Personen im Alter von 6 bis einschließlich 18 Jahren kam die kindgerechte Version des Fragebogens zum Einsatz. Aus den Studienunterlagen geht nicht hervor, bei wie vielen Studienteilnehmenden das Instrument überhaupt eingesetzt wurde. Aus diesem Grund wurden die Daten der C-SSRS nicht für die Nutzenbewertung herangezogen.

Zusammenfassende Einschätzung zur Sicherheit

Statistisch signifikante Nachteile im Endpunkt „Unerwünschte Ereignisse“ treten ausschließlich zuungunsten von Cannabidiol auf. Diese betreffen das Auftreten von SUE und UE, die zum Therapieabbruch führten, sowie einzelne Systemorganklassen und Preferred Terms, darunter insbesondere solche, die sich auf erhöhte Leberwerte beziehen. Das Verzerrungspotential für UE wird als gering eingestuft. Für „Suizidalität“ lagen keine verwertbaren Daten vor. In der Gesamtschau resultiert daraus ein Nachteil von Cannabidiol bei den UE. Zu berücksichtigen ist, dass Cannabidiol in der Studie GWEP1521 in der Behandlungsphase in der maximal zugelassenen Erhaltungsdosis von 25 mg/kg/Tag eingesetzt wurde. Eine Titration nach klinischem Ansprechen sowie Verträglichkeit war nicht möglich. Aufgrund des dosisabhängigen Auftretens von UE ereignen sich im Fall einer niedrigeren Erhaltungsdosis womöglich weniger UE.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen. Die Behandlung mit Cannabidiol soll nur von ärztlichem Personal eingeleitet und überwacht werden, das Erfahrung in der Therapie von Patientinnen und Patienten mit Epilepsie hat.

6 Zusammenfassung der Nutzenbewertung

Cannabidiol (Epidyolex®) zugelassen bei Personen ab 2 Jahren zur adjuvanten Behandlung von Krampfanfällen im Zusammenhang mit Tuberöser Sklerose (TSC), dem Lennox-Gastaut-Syndrom oder dem Dravet-Syndrom. Die Nutzenbewertung von Cannabidiol in der Indikation TSC basiert auf der zulassungsbegründenden Studie GWEP1521 (GWPCARE6). Es handelt sich um eine randomisierte, doppelblinde, placebokontrollierte, multizentrische Phase-III-Studie.

Die Ergebnisse der Studie werden in der folgenden Tabelle 25 unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Cannabidiol ein statistisch signifikant positiver Effekt bei hohem oder unklarem (↑) oder niedrigem (↑↑) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem (↓) oder niedrigem (↓↓) Verzerrungspotential bzw. kein Unterschied (↔) gezeigt werden konnte.

Tabelle 25: Zusammenfassende Darstellung der Ergebnisse der Studie GWEP1521

Endpunkt	Cannabidiol (N = 75) vs. Placebo (N = 76 ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Mortalität		
Todesfälle ²⁾	Keine Ereignisse aufgetreten.	n. b.
Morbidität		
Häufigkeit TSC-assoziiierter Anfälle ¹⁶⁾	Verhältnis Cannabidiol/Placebo ³⁾ : 0,70 [0,58; 0,85]; 0,0003	↑
<i>Häufigkeit TSC-assoziiierter Anfälle¹⁶⁾ – Responder</i>		
Reduktion konvulsiver Anfälle ≥ 25 %	RR: 1,29 [0,93; 1,77]; 0,091 ⁵⁾	↔
Reduktion konvulsiver Anfälle ≥ 50 %	RR: 1,61 [0,96; 2,69]; 0,0692 ⁵⁾	↔
Reduktion konvulsiver Anfälle ≥ 75 %	RR: 25,33 [1,53; 420,24] ⁴⁾ ; 0,0003 ⁵⁾	↑
Reduktion konvulsiver Anfälle ≥ 100 %	n. b.; 0,3173 ⁶⁾	↔
Anstieg konvulsiver Anfälle ≥ 0 %	k. A. ⁷⁾	-
Häufigkeit der Anfälle gesamt ¹⁷⁾	Verhältnis Cannabidiol/Placebo ³⁾ : 0,71 [0,58; 0,86]; 0,0007	↑
<i>Häufigkeit Anfälle gesamt – Responder</i>		
Reduktion konvulsiver Anfälle ≥ 25 %	RR: 1,22 [0,89; 1,68]; 0,1733 ⁵⁾	↔
Reduktion konvulsiver Anfälle ≥ 50 %	RR: 1,64 [0,96; 2,79]; 0,0655 ⁵⁾	↔
Reduktion konvulsiver Anfälle ≥ 75 %	RR: 12,07 [1,61; 90,44] ⁴⁾ ; 0,0015 ⁵⁾	↑
Reduktion konvulsiver Anfälle ≥ 100 %	n. b.; 0,3173 ⁶⁾	↔
Anstieg konvulsiver Anfälle ≥ 0 %	k. A. ⁷⁾	-
Häufigkeit weiterer Anfälle ¹⁸⁾	Verhältnis Cannabidiol/Placebo ³⁾ : 1,56 [0,61; 3,97]; 0,3470 ⁵⁾	↔
Status epilepticus ⁸⁾	RR: 0,78 [0,25; 2,44]; 0,5636 ⁵⁾	↔
<i>Globaler Eindruck der Veränderung – Responder für eine Verbesserung⁹⁾</i>		
CGIC	k. A.	-
CGIC / SGIC	RR: 1,74 [1,26; 2,39]; 0,0004 ³⁾	↑
<i>Globaler Eindruck der Veränderung – Responder für eine Verschlechterung¹⁰⁾</i>		
CGIC	k. A.	-
CGIC / SGIC	k. A.	-

Endpunkt	Cannabidiol (N = 75) vs. Placebo (N = 76 ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Bewertung des adaptiven Verhaltens (mittels Vineland-II) ¹¹⁾	₁₁₎	-
Achenbach-Verhaltens-Checkliste CBCL (1,5–5 Jahre): Gesamtwert CBCL (6–18 Jahre): Gesamtwert CBCL (6–18 Jahre): Soziale Kompetenz CBCL (6–18 Jahre): Aktivitäten CBCL (6–18 Jahre): Schule CBCL (6–18 Jahre): Gesamtkompetenz ABCL (18–59 Jahre): Gesamtwert ABCL (18–59 Jahre): Kritische Elemente ABCL (18–59 Jahre): Gesamtaufälligkeit	LS-MWD: -3,99 [-16,03; 8,05]; 0,5020 LS-MWD: -0,27 [-8,04; 7,51]; 0,9457 LS-MWD: -0,06 [-0,75; 0,63]; 0,8616 LS-MWD: -0,61 [-1,53; 0,31]; 0,1908 ₁₂₎ ₁₂₎ LS-MWD: -6,39; [-21,56; 8,79]; 0,3968 LS-MWD: -1,02 [-2,66; 0,63]; 0,2151 LS-MWD: 0,13 [-1,35; 1,61]; 0,8571	↔
Lebensqualität		
Lebensqualität bei kindlicher Epilepsie (mittels QOLCE)	₁₃₎	-
Sicherheit		
UE	-	-
SUE	RR: 8,22 [1,96; 34,45]; 0,0004 ⁵⁾	↓↓
Therapieabbruch aufgrund von UE	RR: 4,04 [0,90; 18,20]; 0,0484 ⁵⁾	↓↓
Suizidalität (mittels C-SSRS) ¹⁴⁾	k. A.	-

¹⁾ Die beiden Studienarme Placebo 25 mg/kg/Tag und Placebo 50 mg/kg/Tag wurden gepoolt.

²⁾ Todesfälle sind während des Studienzeitraums nicht aufgetreten.

³⁾ Negativ-binomiales Modell; Stratifizierungsfaktoren: Altersgruppe, Zeitpunkt (Baseline und Behandlung), Behandlungsgruppe und Interaktion zwischen Zeitpunkt und Behandlung als fixe Effekte und Patientin/Patient als Zufallseffekt. Log-transformierte Anzahl Tage mit Anfallsberichten als Offset.

⁴⁾ Modell ohne Stratifikationsfaktoren. In einer der Behandlungsgruppen treten keine Ereignisse auf. Aufgrund der vorliegenden Ereigniszahlen wird nur von geringfügigen Änderungen des Effektschätzers im Fall einer Stratifizierung ausgegangen.

⁵⁾ P-Wert basierend auf CMH-Test stratifiziert nach Altersgruppen.

⁶⁾ TSC-assoziierte Anfallsfreiheit und Freiheit von Anfällen gesamt wurde jeweils lediglich von 1 Person im Interventionsarm im Studienzeitraum erreicht.

⁷⁾ Es wurden keine statistischen Auswertungen zum Endpunkt vorgelegt.

⁸⁾ Definiert als jegliche Art von Anfall (konvulsiv und nicht-konvulsiv), der 30 Minuten oder länger andauert.

⁹⁾ Verbesserung definiert als die Punktwerte 1 (sehr stark verbessert), 2 (stark verbessert) und 3 (leicht verbessert).

¹⁰⁾ Verschlechterung definiert als die Punktwerte (4) gering verschlechtert, (5) stark verschlechtert, (6) sehr stark verschlechtert.

¹¹⁾ Die Ergebnisse zur Vineland-II wurden nicht berücksichtigt, da die Rücklaufquoten lediglich in einzelnen Subdomänen über 70 % liegen.

¹²⁾ Aufgrund niedriger Rücklaufquoten wurden die Auswertungen nicht berücksichtigt.

¹³⁾ Die Ergebnisse zum QOLCE wurden nicht berücksichtigt, da die Rücklaufquoten in Bezug auf die zulassungskonforme ITT-Population lediglich in wenigen Subskalen über 70 % liegen.

¹⁴⁾ Unklar, bei wie vielen Studienteilnehmenden das Instrument eingesetzt wurde.

¹⁵⁾ TSC-assoziierte Anfälle umfassen: Fokal-motorische Anfälle ohne Beeinträchtigung des Bewusstseins; fokale Anfälle mit Beeinträchtigung des Bewusstseins; fokale Anfälle, die in bilaterale generalisierte konvulsive Anfälle übergehen; tonisch-klonische, tonische, klonische oder atonische Anfälle.

¹⁶⁾ „Anfälle gesamt“ umfasst alle in der Studie erhobenen Anfallstypen.

¹⁷⁾ „Weitere Anfälle“ umfasst Absencen, myoklonische, partiell sensorische Anfälle und infantile oder epileptischer Spasmen.

Abkürzungen: ABCL: Adult Behavior Checklist; CGIC: Caregiver Global Impression of Change; CBCL: Child Behavior Checklist; CMH: Cochran-Mantel-Haenszel; C-SSRS: Columbia-Suicide Severity Rating Scale; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz; n. b.: nicht berechenbar; QOLCE: Quality of Life in Childhood Epilepsy; RR: Relatives Risiko; SGIC: Subject Global Impression of Change; SUE: Schwerwiegendes unerwünschtes Ereignis; TSC: Tuberöse Sklerose; UE: Unerwünschtes Ereignis; Vineland-II: Vineland Adaptive Behavior Scales, Second Edition.

Referenzen

1. **Achenbach TM.** Child Behavior Checklist. In: Kreutzer JS, Deluca J, Caplan B (Eds). Encyclopedia of Clinical Neuropsychology. New York: Springer; 2011. S. 546-552.
2. **Achenbach TM, Rescorla LA.** Adult Behaviour Checklist (ABCL/ASR). Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2003.
3. **Achenbach TM, Rescorla LA.** Manual for ASEBA School-Age Forms and Profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2001.
4. **Achenbach TM, Rescorla LA.** Manual for the ASEBA Preschool Forms and Profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2000.
5. **Berg AT, Caplan R, Baca CB, Vickrey BG.** Adaptive behavior and later school achievement in children with early-onset epilepsy. Dev Med Child Neurol 2013;55(7):661-667.
6. **Chandler S, Charman T, Baird G, Simonoff E, Loucas T, Meldrum D, et al.** Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. J Am Acad Child Adolesc Psychiatry 2007;46(10):1324-1332.
7. **Columbia Lighthouse Project.** The Columbia Protocol for Research [online]. New York (USA): Columbia University; 2019. [Zugriff: 16.06.2021]. URL: <https://cssrs.columbia.edu/the-columbia-scale-c-ssrs/cssrs-for-research/>.
8. **Connolly AM, Sabaz M, Lawson JA, Bye AM, Cairns DR.** Quality of life in childhood epilepsy: validating the QOLCE. J Paediatr Child Health 2005;41(3):157-158.
9. **De Bildt A, Kraijer D, Sytema S, Minderaa R.** The psychometric properties of the Vineland Adaptive Behavior Scales in children and adolescents with mental retardation. J Autism Dev Disord 2005;35(1):53-62.
10. **De Vries P, Humphrey A, McCartney D, Prather P, Bolton P, Hunt A.** Consensus clinical guidelines for the assessment of cognitive and behavioural problems in Tuberous Sclerosis. Eur Child Adolesc Psychiatry 2005;14(4):183-190.
11. **De Vries PJ, Belousova E, Benedik MP, Carter T, Cottin V, Curatolo P, et al.** TSC-associated neuropsychiatric disorders (TAND): findings from the TOSCA natural history study. Orphanet J Rare Dis 2018;13(1):157.
12. **De Vries PJ, Franz DN, Curatolo P, Nabbout R, Neary M, Herbst F, et al.** Measuring health-related quality of life in tuberous sclerosis complex - psychometric evaluation of three instruments in individuals with refractory epilepsy. Front Pharmacol 2018;9:964.
13. **DeLorenzo RJ, Garnett LK, Towne AR, Waterhouse EJ, Boggs JG, Morton L, et al.** Comparison of status epilepticus with prolonged seizure episodes lasting from 10 to 29 minutes. Epilepsia 1999;40(2):164-169.
14. **Donner EJ, Camfield P, Brooks L, Buchhalter J, Camfield C, Loddenkemper T, et al.** Understanding death in children with epilepsy. Pediatr Neurol 2017;70:7-15.
15. **European Medicines Agency (EMA).** Epidyolex (Cannabidiol): European public assessment report (Variation assessment report) EMEA/H/C/004675/II/0005 [online]. Amsterdam (NED): EMA; 2021. [Zugriff: 16.06.2021]. URL: https://www.ema.europa.eu/en/documents/variation-report/epidyolex-h-c-4675-ii-0005-epar-assessment-report-variation_en.pdf.

16. **European Medicines Agency (EMA).** Guideline on clinical investigation of medicinal products in the treatment of epileptic disorders, CHMP/EWP/566/98 Rev.2/Corr [online]. Amsterdam (NED): EMA; 2010. [Zugriff: 16.06.2021]. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-epileptic-disorders-revision-2_en.pdf.
17. **Granader YE, Bender HA, Zemon V, Rath S, Nass R, Macallister WS.** The clinical utility of the Social Responsiveness Scale and Social Communication Questionnaire in tuberous sclerosis complex. *Epilepsy Behav* 2010;18(3):262-266.
18. **GW Pharma, GW Research.** A double-blind, randomized, placebo-controlled study to investigate the efficacy and safety of cannabidiol (GWP42003-P, CBD) as add-on therapy in patients with tuberous sclerosis complex who experience inadequately-controlled seizures (GWEP1521): Clinical study report: blinded phase [unveröffentlicht]. 2019.
19. **GW Pharma, GW Research.** A double-blind, randomized, placebo-controlled study to investigate the efficacy and safety of cannabidiol (GWP42003-P, CBD) as add-on therapy in patients with tuberous sclerosis complex who experience inadequately-controlled seizures (GWEP1521): Statistical analysis plan [unveröffentlicht]. 2019.
20. **GW Pharma, GW Research.** A double-blind, randomized, placebo-controlled study to investigate the efficacy and safety of cannabidiol (GWP42003-P, CBD) as add-on therapy in patients with tuberous sclerosis complex who experience inadequately-controlled seizures (GWEP1521): study protocol [unveröffentlicht]. 2015.
21. **GW Pharma (International).** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Cannabidiol (Epidyolex): Bei Patienten ab 2 Jahren für die adjuvante Behandlung von Krampfanfällen im Zusammenhang mit Tuberöser Sklerose (TSC); Modul 4 C: Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 14.05.2021.
22. **GW Pharma (International).** Epidyolex 100 mg/ml Lösung zum Einnehmen: Fachinformation [online]. Berlin (GER): 2021. [Zugriff: 21.07.2021]. URL: <http://www.fachinfo.de>.
23. **Hamer HM, Pfafflin M, Baier H, Bosebeck F, Franz M, Holtkamp M, et al.** Characteristics and healthcare situation of adult patients with tuberous sclerosis complex in German epilepsy centers. *Epilepsy Behav* 2018;82:64-67.
24. **Joinson C, O'Callaghan FJ, Osborne JP, Martyn C, Harris T, Bolton PF.** Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex. *Psychol Med* 2003;33(2):335-344.
25. **May TW, Pfafflin M, Cramer JA.** Psychometric properties of the German translation of the QOLIE-31. *Epilepsy Behav* 2001;2(2):106-114.
26. **Middleton HA, Keene RG, Brown GW.** Convergent and discriminant validities of the Scales of Independent Behavior and the revised Vineland Adaptive Behavior Scales. *Am J Ment Retard* 1990;94(6):669-673.
27. **Perry A, Factor DC.** Psychometric validity and clinical usefulness of the Vineland Adaptive Behavior Scales and the AAMD Adaptive Behavior Scale for an autistic sample. *J Autism Dev Disord* 1989;19(1):41-55.

28. **Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al.** The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011; 168(12):1266-1277.
29. **Sabaz M, Cairns DR, Lawson JA, Bleasel AF, Bye AM.** The health-related quality of life of children with refractory epilepsy: a comparison of those with and without intellectual disability. *Epilepsia* 2001;42(5):621-628.
30. **Sabaz M, Cairns DR, Lawson JA, Nheu N, Bleasel AF, Bye AM.** Validation of a new quality of life measure for children with epilepsy. *Epilepsia* 2000;41(6):765-774.
31. **Sabaz M, Lawson JA, Cairns DR, Duchowny MS, Resnick TJ, Dean PM, et al.** The impact of epilepsy surgery on quality of life in children. *Neurology* 2006;66(4):557-561.
32. **Sabaz M, Lawson JA, Cairns DR, Duchowny MS, Resnick TJ, Dean PM, et al.** Validation of the quality of life in childhood epilepsy questionnaire in American epilepsy patients. *Epilepsy Behav* 2003;4(6):680-691.
33. **Shepherd CW, Gomez MR, Lie JT, Crowson CS.** Causes of death in patients with tuberous sclerosis. *Mayo Clin Proc* 1991;66(8):792-796.
34. **Smalley SL, Tanguay PE, Smith M, Gutierrez G.** Autism and tuberous sclerosis. *J Autism Dev Disord* 1992;22(3):339-355.
35. **Sparrow SS, Cicchetti DV, Balla DA.** Vineland Adaptive Behavior Scales, Second Edition (Vineland-II). London: Pearson; 2005.
36. **Sparrow SS, Cicchetti DV, Balla DA.** Vineland adaptive behavior scales: Second edition (Vineland II), survey interview form/caregiver rating form. Livonia: Pearson; 2005.
37. **Thiele EA, Bebin EM, Bhathal H, Jansen FE, Kotulska K, Lawson JA, et al.** Add-on cannabidiol treatment for drug-resistant seizures in tuberous sclerosis complex: a placebo-controlled randomized clinical trial. *JAMA Neurology* 2021;78(3):285-292.