

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Upadacitinib (RINVOQ®)

AbbVie Deutschland GmbH & Co. KG

Modul 4 A

Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben.

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	8
Abkürzungsverzeichnis	9
4 Modul 4 – allgemeine Informationen	13
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik.....	21
4.2.1 Fragestellung.....	21
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	23
4.2.3 Informationsbeschaffung.....	25
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	25
4.2.3.2 Bibliografische Literaturrecherche.....	25
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken.....	26
4.2.3.4 Suche auf der Internetseite des G-BA.....	28
4.2.3.5 Selektion relevanter Studien.....	29
4.2.4 Bewertung der Aussagekraft der Nachweise.....	29
4.2.5 Informationssynthese und -analyse.....	30
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	30
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	31
4.2.5.3 Meta-Analysen.....	50
4.2.5.4 Sensitivitätsanalysen.....	52
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	52
4.2.5.6 Indirekte Vergleiche.....	53
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	56
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	56
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	56
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	56
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	58
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken.....	61
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA.....	62
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	62
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	63
4.3.1.2.1 Studiendesign und Studienpopulationen.....	63
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	75
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	77
4.3.1.3.1 Mortalität – RCT.....	77
4.3.1.3.2 Krankheitsaktivität – RCT.....	82
4.3.1.3.2.1 ASAS – RCT.....	82
4.3.1.3.2.2 ASDAS – RCT.....	86

4.3.1.3.2.3	BASDAI – RCT.....	90
4.3.1.3.3	Schmerz.....	92
4.3.1.3.3.1	Patientenberichteter Rückenschmerz – RCT.....	92
4.3.1.3.4	Funktion.....	94
4.3.1.3.4.1	BASFI – RCT.....	94
4.3.1.3.5	Entzündung.....	97
4.3.1.3.5.1	SPARCC MRI Spine – RCT.....	97
4.3.1.3.5.2	MASES – RCT.....	99
4.3.1.3.6	Fatigue (FACIT-F) – RCT.....	101
4.3.1.3.7	Gesundheitsbezogene Lebensqualität.....	104
4.3.1.3.7.1	ASQoL – RCT.....	104
4.3.1.3.8	Verträglichkeit – RCT.....	106
4.3.1.3.9	Subgruppenanalysen – RCT.....	110
4.3.1.4	Liste der eingeschlossenen Studien – RCT.....	112
4.3.2	Weitere Unterlagen.....	112
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien.....	112
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche.....	112
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	113
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen.....	113
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT.....	113
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT.....	116
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT...	116
4.3.2.2	Nicht randomisierte vergleichende Studien.....	116
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien.....	116
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	117
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien.....	118
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	118
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien.....	119
4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien.....	120
4.3.2.3	Weitere Untersuchungen.....	120
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen.....	120
4.3.2.3.2	Charakteristika der weiteren Untersuchungen.....	121
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen.....	121
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen.....	121
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen.....	122
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen.....	122
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	122
4.4.1	Beurteilung der Aussagekraft der Nachweise.....	122
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	124
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht.....	129
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte.....	129
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	129

4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	130
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	130
4.5.4	Verwendung von Surrogatendpunkten.....	130
4.6	Referenzliste.....	131
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		141
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....		145
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		147
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....		148
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		150
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		153

Tabellenverzeichnis

	Seite
Tabelle 4-1: Übersicht über die zu betrachtende Fragestellung bzw. Population.....	14
Tabelle 4-2: Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel.....	15
Tabelle 4-3: Übersicht über die zu betrachtende Fragestellung bzw. Population.....	22
Tabelle 4-4: Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel.....	23
Tabelle 4-5: Nutzendimensionen und patientenrelevante Endpunkte zur Bewertung des Ausmaßes des medizinischen Nutzens von Upadacitinib	34
Tabelle 4-6: Möglichkeiten der systematischen Erfassung von Therapiezielen gemäß S3-Leitlinie.....	36
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	58
Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel.....	62
Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	63
Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	65
Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	67
Tabelle 4-14: Charakterisierung der Studienpopulation: Demografische Charakteristika zu Baseline nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	68
Tabelle 4-15: Charakterisierung der Studienpopulationen: Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel	69
Tabelle 4-16: Charakterisierung der Studienpopulation: Medizinische Vorgeschichte und Begleitbehandlung nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel	71
Tabelle 4-17: Patientenstatus in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	71
Tabelle 4-18: Behandlungs- und Beobachtungsdauer in RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel	72
Tabelle 4-19: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	75

Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	77
Tabelle 4-21: Operationalisierung von Mortalität.....	80
Tabelle 4-22: Bewertung des Verzerrungspotenzials für Mortalität in RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-23: Operationalisierung von ASAS	82
Tabelle 4-24: Bewertung des Verzerrungspotenzials für ASAS in RCT mit dem zu bewertenden Arzneimittel	83
Tabelle 4-25: Ergebnisse für ASAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	83
Tabelle 4-26: Operationalisierung von ASDAS.....	86
Tabelle 4-27: Bewertung des Verzerrungspotenzials für ASDAS in RCT mit dem zu bewertenden Arzneimittel	87
Tabelle 4-28: Ergebnisse für die mittlere Veränderung des ASDAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	88
Tabelle 4-29: Weitere Ergebnisse für ASDAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel	88
Tabelle 4-30: Operationalisierung von BASDAI.....	90
Tabelle 4-31: Bewertung des Verzerrungspotenzials für BASDAI in RCT mit dem zu bewertenden Arzneimittel	90
Tabelle 4-32: Ergebnisse für BASDAI aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel	91
Tabelle 4-33: Operationalisierung von patientenberichtetem Rückenschmerz.....	92
Tabelle 4-34: Bewertung des Verzerrungspotenzials für patientenberichteten Rückenschmerz in RCT mit dem zu bewertenden Arzneimittel.....	93
Tabelle 4-35: Ergebnisse für die mittlere Veränderung des patientenberichteten Rückenschmerzes aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	93
Tabelle 4-36: Operationalisierung von BASFI	94
Tabelle 4-37: Bewertung des Verzerrungspotenzials für BASFI in RCT mit dem zu bewertenden Arzneimittel	95
Tabelle 4-38: Ergebnisse für die mittlere Veränderung des BASFI aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	96
Tabelle 4-39: Operationalisierung von SPARCC MRI Spine.....	97
Tabelle 4-40: Bewertung des Verzerrungspotenzials für SPARCC MRI Spine in RCT mit dem zu bewertenden Arzneimittel.....	97
Tabelle 4-41: Ergebnisse für die mittlere Veränderung des SPARCC MRI Spine aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	98
Tabelle 4-42: Operationalisierung von MASES	99
Tabelle 4-43: Bewertung des Verzerrungspotenzials für MASES in RCT mit dem zu bewertenden Arzneimittel	99

Tabelle 4-44: Ergebnisse für die mittlere Veränderung des MASES aus RCT SELECT- AXIS 2 mit dem zu bewertenden Arzneimittel.....	100
Tabelle 4-45: Operationalisierung von FACIT-F.....	101
Tabelle 4-46: Bewertung des Verzerrungspotenzials für FACIT-F in RCT mit dem zu bewertenden Arzneimittel	102
Tabelle 4-47: Ergebnisse für die mittlere Veränderung des FACIT-F aus RCT SELECT- AXIS 2 mit dem zu bewertenden Arzneimittel.....	102
Tabelle 4-48: Operationalisierung von ASQoL	104
Tabelle 4-49: Bewertung des Verzerrungspotenzials für ASQoL in RCT mit dem zu bewertenden Arzneimittel	104
Tabelle 4-50: Ergebnisse für die mittlere Veränderung des ASQoL aus RCT SELECT- AXIS 2 mit dem zu bewertenden Arzneimittel.....	105
Tabelle 4-51: Operationalisierung von UE	106
Tabelle 4-52: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-53: Ergebnisse für UE aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	107
Tabelle 4-54: Ergebnisse für häufige unerwünschte Ereignisse von speziellem Interesse aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel.....	108
Tabelle 4-55 Matrix der durchgeführten Subgruppenanalysen.....	111
Tabelle 4-56: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für <Studie> und <Effektmodifikator>.....	111
Tabelle 4-57: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	113
Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	114
Tabelle 4-59: Operationalisierung von <Endpunkt xxx>.....	114
Tabelle 4-60: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	115
Tabelle 4-61: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	115
Tabelle 4-62: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	117
Tabelle 4-63: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien	118
Tabelle 4-64: Operationalisierung von <Endpunkt xxx>.....	118
Tabelle 4-65: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	119
Tabelle 4-66: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen	121
Tabelle 4-67: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	121

Tabelle 4-68: Übersicht über die zu betrachtende Fragestellung bzw. Population 126

Tabelle 4-69: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens..... 129

Tabelle 4-70 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung> ... 151

Tabelle 4-71 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie <Studienbezeichnung> 154

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	59
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	60
Abbildung 3: Anteil der Patienten mit einem ASAS40-Ansprechen von Baseline bis Woche 14 (Studie SELECT-AXIS 2, FAS)	85

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AD	Atopische Dermatitis
AMIce	Arzneimittel-Informationssystem (vormals AMIS)
AMIS	Arzneimittel-Informationssystem
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse (Analysis of Covariance)
AS	Ankylosierende Spondylitis
ASAS	Assessment of SpondyloArthritis International Society
ASAS-PR	ASAS-partielle Remission
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASQoL	Fragebogen zur Lebensqualität bei ankylosierender Spondylitis (Ankylosing Spondylitis Quality of Life Questionnaire)
axSpA	Axiale Spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
bDMARD	Biologisches DMARD (Biological DMARD)
BMI	Body Mass Index
BSG	Blutsenkungsgeschwindigkeit
CHMP	Committee for Medicinal Products for Human Use
CMH	Cochran-Mantel-Haenszel
CMQ	Company MedDRA Query
Cochrane	Cochrane Central Register of Controlled Trials
CONSORT	Consolidated Standards of Reporting Trials
COVID	Corona Virus Disease
CRP	C-reaktives Protein
csDMARD	Konventionelles synthetisches DMARD (Conventional Synthetic DMARD)
CSR	Klinischer Studienreport (Clinical Study Report)
CTCAE	Common Terminology Criteria for Adverse Events
CU	Colitis ulcerosa
DIMDI	Deutsches Institut für Medizinische Dokumentation

Abkürzung	Bedeutung
DMARD	Krankheitsmodifizierendes Antirheumatikum (Disease-Modifying Anti-Rheumatic Drug)
DVU	Discovertebrale Einheit (Discovertebral Unit)
EMA	Europäische Arzneimittel-Agentur (European Medicines Agency)
EMBASE	Excerpta Medica Database
EU-CTR	European Union Clinical Trials Registry
EudraCT	European Union Drug Regulating Authorities Clinical Trials Database
EULAR	European League Against Rheumatism
FACIT-F	Functional Assessment of Chronic Illness Therapy – Fatigue
FAS	Full Analysis Set
G-BA	Gemeinsamer Bundesausschuss
HI	Health Index
HLA	Humanes Leukozytenantigen
hsCRP	Hoch sensitives C-reaktives Protein (High sensitivity C-reactive Protein)
ICTRP	International Clinical Trials Registry Platform
IL	Interleukin
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interaktives Response Technologie System (Interactive Response Technology)
ITT	Intention-to-Treat
JAK	Januskinase
k. A.	Keine Angabe
KI	Konfidenzintervall
LS-MW	Least-Squares-Mittelwert
MACE	Major Adverse Cardiovascular Events
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MedDRA	Medical Dictionary for Regulatory Activities
MEDLINE	Medical Literature Analysis and Retrieval System Online
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measures

Abkürzung	Bedeutung
MRI	Magnetic Resonance Imaging
MRT	Magnetresonanztomografie
MTC	Mixed Treatment Comparison
MTX	Methotrexat
MW	Mittelwert
MWD	Mittelwertdifferenz
NCT	National Clinical Trial (US)
NMSC	Nicht melanozytärer Hautkrebs (Non-Melanoma Skin Cancer)
nr-axSpA	Nicht röntgenologische axiale Spondyloarthritis
NRI	Non-Responder-Imputation
NRS	Numerical Rating Scale
NSAR	Nicht steroidales Antirheumatikum
OMERACT	Outcome Measures in Rheumatology
OR	Odds-Ratio
p.o.	Peroral
PsA	Psoriasis-Arthritis
PT	Preferred Term nach MedDRA
PtGADA	Patientenberichtete globale Krankheitsaktivität (Patient's Global Assessment of Disease Activity)
QD	Einmal täglich (Quaque Die)
RA	Rheumatoide Arthritis
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RD	Risikodifferenz
RR	Relatives Risiko
SAP	Statistischer Analyseplan
s.c.	Subkutan
SD	Standardabweichung (Standard Deviation)
SE	Standardfehler (Standard Error)
SF-36	Short Form – 36
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz

Abkürzung	Bedeutung
SMQ	Standardized MedDRA Query
SOC	Systemorganklasse nach MedDRA (System Organ Class in MedDRA)
SPARCC	Spondyloarthritis Research Consortium of Canada
STE	Surrogate Threshold Effects
STIR	Short-Tau-Inversion-Recovery
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes UE
TNF	Tumornekrosefaktor
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
UESI	Unerwünschtes Ereignis von speziellem Interesse
ULN	Obere Normgrenze (Upper Limit of Normal)
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung (des G-BA)
VTE	Venöse thromboembolische Ereignisse
WHO	Weltgesundheitsorganisation (World Health Organization)
ZVT	Zweckmäßige Vergleichstherapie

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Nutzens und Zusatznutzens von Upadacitinib zur Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben.

Der medizinische Nutzen bzw. Zusatznutzen von Upadacitinib wird anhand der folgenden Patientenpopulation untersucht:

Tabelle 4-1: Übersicht über die zu betrachtende Fragestellung bzw. Population

Fragestellung	Population	Intervention	ZVT
Gesamtpopulation			
1	Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben	Upadacitinib	TNF-Inhibitor (Etanercept oder Adalimumab oder Golimumab oder Certolizumab Pegol)
CRP: C-reaktives Protein; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum; TNF: Tumornekrosefaktor; ZVT: Zweckmäßige Vergleichstherapie			

Die zweckmäßige Vergleichstherapie (ZVT) für Upadacitinib wurde in einem Beratungsgespräch am 15. November 2018 (Beratungsanforderung 2018-B-186) durch den Gemeinsamen Bundesausschuss (G-BA) wie folgt festgelegt:

- ein TNF-Inhibitor (Etanercept oder Adalimumab oder Golimumab oder Certolizumab Pegol)

Datenquellen

Die bibliografische Literaturrecherche wurde in den Datenbanken Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica Database (EMBASE) und Cochrane Central Register of Controlled Trials (Cochrane) durchgeführt, wobei für jede einzelne Datenbank eine sequenzielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wurde. Für die Suche wurde die Plattform OVID verwendet.

Die Suche in Studienregistern/Studienergebnisdatenbanken wird in den Portalen [clinicaltrials.gov](https://www.clinicaltrials.gov/) (<https://www.clinicaltrials.gov/>), International Clinical Trials Registry Platform (ICTRP) Search Portal, Suchportal der Weltgesundheitsorganisation (World Health Organization, WHO) (<http://apps.who.int/trialsearch/>) und im European Union Clinical Trials Registry (EU-CTR) (<https://www.clinicaltrialsregister.eu/ctr-search/search>) durchgeführt. Zudem wird im Clinical Data Suchportal der Europäischen Arzneimittel-Agentur (European Medicines Agency, EMA) (<https://clinicaldata.ema.europa.eu>) und im Arzneimittel-Informationssystem (AMIce, vormals AMIS) (www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html) nach Ergebnisberichten zu denjenigen Studien gesucht, die zuvor schon als relevante Studien identifiziert wurden.

Ein-/Ausschlusskriterien für Studien

Diejenigen Studien, die grundsätzlich für die Beantwortung der Fragestellung geeignet sind, werden anhand von Ein- und Ausschlusskriterien identifiziert. Eine zusammenfassende Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-2.

Tabelle 4-2: Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel.

Kriterium	Einschluss		Ausschluss		Begründung
Patientenpopulation	Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende Patientenpopulation	A1	Population des zugelassenen Anwendungsgebiets
Intervention	Behandlung mit Upadacitinib (15 mg) gemäß Fachinformation	E2	<ul style="list-style-type: none"> Behandlung mit Upadacitinib abweichend von der Fachinformation Andere Interventionen 	A2	Zu bewertendes Arzneimittel
Vergleichstherapie	Adalimumab oder Certolizumab Pegol oder Etanercept oder Golimumab	E3	Andere Therapien oder fehlende Übereinstimmung der Vergleichstherapie mit dem deutschen Label	A3	Therapien der ZVT gemäß G-BA-Beratung
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität,	E4	Kein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität,	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der Verfo

Kriterium	Einschluss		Ausschluss		Begründung
	Lebensqualität oder der Verträglichkeit verwendet werden kann		Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann		
Studientyp	RCT	E5	Nicht RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt
Studiendauer	≥ 24 Wochen	E6	< 24 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt. Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z. B. clinicaltrials.gov) können ebenfalls berücksichtigt werden.	E7	<ul style="list-style-type: none"> Berichterstattung liefert keine ausreichenden Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements genügen) Mehrfachpublikation ohne relevante Zusatzinformationen^a 	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich

a: Jeglicher Publikationstyp, in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.

CONSORT: Consolidated Standards of Reporting Trials; CRP: C-reaktives Protein; G-BA: Gemeinsamer Bundesausschuss; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum; RCT: Randomisierte kontrollierte Studie; VerFO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie

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

Zur Beurteilung des Zusatznutzens von Upadacitinib wurde eine Fragestellung formuliert. Für die Bewertung des medizinischen Nutzens und Zusatznutzens von Upadacitinib liegt keine für die Nutzenbewertung relevante Studie, die einen direkten Vergleich und somit eine Aussage zu einem Zusatznutzen gegenüber der ZVT ermöglicht, vor. Dementsprechend wird kein Zusatznutzen beansprucht.

Zur Beschreibung der Wirksamkeit und Verträglichkeit von Upadacitinib in der Indikation aktive nr-axSpA liegt eine placebokontrollierte Studie vor, die hier ergänzend dargestellt wird. Bei der Studie SELECT-AXIS 2 handelt es sich um eine multizentrische, randomisierte, doppelblinde Phase-III-Studie der Evidenzstufe Ib (nach § 11 2. Kapitel 3. Abschnitt der Verfahrensordnung [VerfO] des G-BA).

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die vorliegende Studie SELECT-AXIS 2 (M19-944) ist eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie. Die EMA hat bereits im Zulassungsprozess ein positives Nutzen-Risikoverhältnis für eine Upadacitinib-Behandlung in der Indikation nr-axSpA festgestellt. Die Studienpopulation beinhaltet erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, die unzureichend auf NSAR angesprochen haben. Da keine Daten im Vergleich zur ZVT vorliegen, werden die Ergebnisse der Studie SELECT-AXIS 2 ergänzend dargestellt, um die Wirksamkeit und Sicherheit von Upadacitinib zu beschreiben.

In der Studie SELECT-AXIS 2 wurden patientenrelevante Endpunkte in den Kategorien Morbidität (Krankheitsaktivität, Schmerz, Funktion, Entzündung und Fatigue), gesundheitsbezogene Lebensqualität und Verträglichkeit erhoben.

Die nr-axSpA manifestiert sich in Knochen und Gelenken der unteren Wirbelsäule wie auch der Verbindung zum Becken (Iliosakralgelenk) sowie in benachbarten Weichteilen, wie z. B. Sehnen und Bändern. Patienten mit nr-axSpA leiden unter chronischen Schmerzen und Bewegungseinschränkungen des Achsenskeletts, was zu Einschränkungen der Lebensqualität und des sozialen Lebens führt.

Das primäre Ziel sowohl in der Leitlinie der Assessment of SpondyloArthritis International Society (ASAS)/European League Against Rheumatism (EULAR) als auch in der deutschen S3-Leitlinie bei der Behandlung von Patienten mit nr-axSpA ist die langfristige Verbesserung und Erhaltung der gesundheitsbezogenen Lebensqualität. Die Therapie sollte darauf ausgerichtet sein, eine größtmögliche Schmerzreduktion zu erreichen und gleichzeitig den Erhalt der körperlichen Funktionsfähigkeit, die Reduktion der Steifigkeit, die Verhinderung struktureller Läsionen und den Erhalt der Erwerbsfähigkeit zu gewährleisten. Dabei nimmt die

Reduktion der Krankheitsaktivität, gemessen anhand stringenter Endpunkte, wie z. B. Ankylosing Spondylitis Disease Activity Score (ASDAS) niedrige Krankheitsaktivität, eine zentrale Rolle in der Behandlung von Patienten mit nr-axSpA ein. Die deutsche S3-Leitlinie definiert die Remission als das Erreichen einer partiellen Remission gemäß ASAS-Kriterien oder als das Erreichen eines ASDAS inaktive Erkrankung.

Krankheitsaktivität

Das Erreichen einer niedrigen Krankheitsaktivität bzw. Remission ist ein zentrales Therapieziel. Das ASAS-Ansprechen, der ASDAS und der Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) bilden in der Studie SELECT-AXIS 2 die Krankheitsaktivität ab. Hier konnte durchweg eine stringente Kontrolle der Krankheits-symptomatik durch Upadacitinib im Vergleich zu Placebo beobachtet werden. Eine statistisch signifikante Verbesserung der Krankheitsaktivität konnte auch schon nach zwei Wochen anhand des ASAS40-Ansprechens gezeigt werden. Besonders deutlich zeigten sich die positiven Effekte zur Krankheitsaktivität auch durch den signifikant höheren Anteil der Patienten, die eine Remission (ASAS-partielle Remission [ASAS-PR]; ASDAS inaktive Erkrankung [ASDAS < 1,3]) oder niedrige Krankheitsaktivität (ASDAS $\geq 1,3$ und < 2,1) unter einer Behandlung mit Upadacitinib erreichten.

Schmerzen

Bei Patienten mit nr-axSpA stehen neben Bewegungseinschränkungen des Achsenskeletts insbesondere Schmerzen im Vordergrund, weshalb die Schmerzreduktion eines der wichtigsten Therapieziele ist. Unter einer Upadacitinib-Behandlung kann eine deutliche Verbesserung der patientenberichteten Rückenschmerzen erreicht werden. Diese Schmerzreduktion kann bereits nach zwei Wochen beobachtet werden. Neben den hier dargestellten Daten zum patientenberichteten Rückenschmerz anhand einer patientenberichteten Numerical Rating Scale (NRS) zum Rückenschmerz insgesamt, wurde gezeigt, dass die Einnahme von Upadacitinib über alle Schmerzdomänen hinweg (axialer, peripherer und nächtlicher Schmerz; erhoben mit den Items 2 und 3 des BASDAI bzw. einer patientenberichteten NRS zum nächtlichen Rückenschmerz) zu einer deutlichen und sehr schnellen Verbesserung führt.

Funktion

Neben den Rückenschmerzen ist die zunehmende Steifheit der Wirbelsäule, die zu einer mitunter dramatischen und bewegungseinschränkenden Veränderung der Körperhaltung führen kann, für Patienten mit nr-axSpA von zentraler Bedeutung. In der Studie SELECT-AXIS 2 wurde die Funktion anhand des Bath Ankylosing Spondylitis Functional Index (BASFI) beurteilt. Es konnten signifikante Verbesserungen der Funktionseinschränkungen im Vergleich zu Placebo gezeigt werden.

Entzündung

Entzündungen im Bereich der Wirbelsäule oder auch Entesen können Schmerzen und/oder Schwellungen der peripheren und axialen Gelenke sowie der gelenknahen Strukturen verursachen und erheblich zur Krankheitslast beitragen. Die Ausprägung der Entzündung der Wirbelsäule wurde in der Studie SELECT-AXIS 2 über den Spondyloarthritis Research

Consortium of Canada Magnetic Resonance Imaging (SPARCC MRI) Spine erhoben, das Ausmaß der Enthesitis wurde durch den Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) bestimmt. Auch in diesen bedeutenden Krankheitsmanifestationen zeigte eine Upadacitinib-Therapie eine signifikante Verbesserung der akuten Entzündung gegenüber Placebo.

Fatigue

Fatigue ist für Patienten mit nr-axSpA ein bedeutendes Symptom, das mit Einschränkungen der Funktionalität einhergehen kann und einen erheblichen Einfluss auf Arbeitsfähigkeit und Lebensqualität der Patienten hat. Eine Verbesserung der Fatigue sollte nach dem Empfinden von Patienten ein Therapieziel sein. Das Ausmaß der Fatigue wurde in der Studie SELECT-AXIS 2 mittels des Fragebogens Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F) erhoben. Gegenüber Placebo konnte eine signifikante Verbesserung der Fatigue durch die Upadacitinib-Therapie gezeigt werden.

Gesundheitsbezogene Lebensqualität

Als Behandlungsziel wird von nr-axSpA-Patienten neben der Eliminierung oder Reduktion der Schmerzen die Verbesserung der Funktionsfähigkeit und der Lebensqualität genannt. Die gesundheitsbezogene Lebensqualität wurde in der Zulassungsstudie mittels des Fragebogens zur Lebensqualität bei ankylosierender Spondylitis (Ankylosing Spondylitis Quality of Life Questionnaire, ASQoL) erhoben, es konnte auch hier eine signifikante Verbesserung durch die Upadacitinib-Behandlung gezeigt werden.

Verträglichkeit

Das Auftreten jeglicher unerwünschter Ereignisse in der Studie SELECT-AXIS 2 war vergleichbar unter Upadacitinib und Placebo. In der Studie SELECT-AXIS 2 wurden nur sehr wenige schwerwiegende Infektionen berichtet. Darüber hinaus traten keine Malignitäten, Major Adverse Cardiovascular Events (MACE), venösen thromboembolischen Ereignisse (VTE) und Todesfälle unter der Behandlung mit Upadacitinib auf.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen

Zusatznutzen

Für viele Patienten mit nr-axSpA sind trotz verfügbarer Therapieoptionen in Form von NSAR und biologischer krankheitsmodifizierenden Antirheumatika (Disease-Modifying Anti-Rheumatic Drugs, DMARDs) (Tumornekrosefaktor [TNF]- und Interleukin [IL]-17-Inhibitoren) die Therapieziele einer niedrigen Krankheitsaktivität und besonders einer anhaltenden Remission nach wie vor nicht zu erreichen. Entsprechend besteht weiterhin ein großer Bedarf an neuen Therapieoptionen. Upadacitinib ist neben dem guten Wirksamkeits- und Verträglichkeitsprofil auch aufgrund des neuen Wirkmechanismus eine Bereicherung für die Therapielandschaft.

In der Gesamtschau konnten in der pivotalen Zulassungsstudie SELECT-AXIS 2 in den Wirksamkeitsendpunkten zur Krankheitsaktivität, zu Schmerzen/Rückenschmerzen, zur körperlichen Funktionsfähigkeit und zur Entzündung deutliche Vorteile einer Upadacitinib-

Behandlung gegenüber Placebo bei nr-axSpA-Patienten gezeigt werden. Dies lässt auf eine stringente Kontrolle der Krankheitsaktivität sowie eine schnelle Schmerzreduktion bei den Patienten schließen und trägt zum Erreichen des primären Therapieziels der deutschen S3-Leitlinie, der langfristigen Verbesserung und Erhaltung der gesundheitsbezogenen Lebensqualität der Patienten, bei. Das Auftreten unerwünschter Ereignisse war unter Upadacitinib und Placebo ähnlich. Insgesamt umfasst das Upadacitinib-Studienprogramm klinische Studien der Phase-II/III bei rheumatoider Arthritis (RA), Psoriasis-Arthritis (PsA), ankylosierender Spondylitis (AS), nr-axSpA, Colitis ulcerosa (CU) und atopischer Dermatitis (AD) mit insgesamt über 5.000 Patienten, die bis zu 4,5 Jahre mit Upadacitinib behandelt wurden. Das Gesamtsicherheitsprofil von Upadacitinib ist über die zugelassenen Indikationen konsistent, gut charakterisiert und es wurden keine neuen unerwarteten Nebenwirkungen beobachtet.

Für die oben genannte Population wurde keine Evidenz für einen direkten Vergleich zur ZVT identifiziert. Es liegen zudem keine Studien zu Upadacitinib vor, die für einen adjustierten oder historischen indirekten Vergleich zur ZVT geeignet sind. AbbVie verzichtet daher auf die Durchführung von indirekten Vergleichen. Eine Aussage über einen Zusatznutzen von Upadacitinib gegenüber der ZVT ist somit für die Zielpopulation sowie für die beiden Populationen der bDMARD-naiven und bDMARD-erfahrenen Patienten nicht möglich und ein Zusatznutzen wird entsprechend nicht beansprucht. Aufgrund der fehlenden Evidenz wird auf eine Darstellung der Populationen bDMARD-naive und bDMARD-erfahrene Patienten im Rahmen des Moduls verzichtet.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Nutzens und Zusatznutzens von Upadacitinib zur Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben (1).

Die Bewertung erfolgt gegenüber der vom G-BA festgelegten ZVT anhand patientenrelevanter Endpunkte aus den Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit. Basis der Nutzenbewertung sind randomisierte kontrollierte Studien (Randomized Controlled Trials, RCTs).

Tabelle 4-3: Übersicht über die zu betrachtende Fragestellung bzw. Population

Fragestellung	Population	Intervention	ZVT
Gesamtpopulation			
1	Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben	Upadacitinib	TNF-Inhibitor (Etanercept oder Adalimumab oder Golimumab oder Certolizumab Pegol)
CRP: C-reaktives Protein; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum; TNF: Tumornekrosefaktor; ZVT: Zweckmäßige Vergleichstherapie			

Patientenpopulation

Der medizinische Nutzen bzw. Zusatznutzen von Upadacitinib wird anhand der in Tabelle 4-3 definierten Patientenpopulation untersucht.

Ziel des vorliegenden Modul 4 ist die Bewertung des medizinischen Nutzens und Zusatznutzens einer Behandlung mit Upadacitinib. Upadacitinib (RINVOQ®) wird angewendet zur Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben.

Intervention

Die zu bewertende Intervention ist Upadacitinib. Bei Upadacitinib handelt es sich um einen selektiven und reversiblen Januskinase (JAK)-Inhibitor. Die empfohlene Dosis beträgt eine 15 mg Retardtablette einmal täglich gemäß Fachinformation (1).

Vergleichstherapie

Die ZVT für Upadacitinib wurde in einem Beratungsgespräch am 15. November 2018 (Beratungsanforderung 2018-B-186) durch den G-BA wie folgt festgelegt (2):

- ein TNF-Inhibitor (Etanercept oder Adalimumab oder Golimumab oder Certolizumab Pegol)

Der Festlegung der ZVT durch den G-BA wird gefolgt.

Endpunkte

Die Bewertung des medizinischen Nutzens von Upadacitinib erfolgt anhand der Nutzenkategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit auf Grundlage patientenrelevanter Endpunkte. Eine umfassende Auflistung der

im Rahmen des Dossiers betrachteten patientenrelevanten Endpunkte findet sich in Tabelle 4-5 in Abschnitt 4.2.5.2.

Studientypen

Um die höchste Ergebnissicherheit zu bieten, wird der medizinische Nutzen in diesem Dossier auf Basis von RCTs mit einer Mindestdauer von ≥ 24 Wochen bewertet.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Zur Identifikation relevanter Studien für die Nutzenbewertung wurde eine Suche nach RCTs durchgeführt, in welchen das zu bewertende Arzneimittel direkt mit der ZVT verglichen wurde. Diejenigen Studien, die grundsätzlich für die Beantwortung der Fragestellung gemäß Abschnitt 4.2.1 geeignet waren, wurden anhand von Ein- und Ausschlusskriterien identifiziert. Eine zusammenfassende Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-4.

Tabelle 4-4: Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel.

Kriterium	Einschluss		Ausschluss		Begründung
Patientenpopulation	Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende Patientenpopulation	A1	Population des zugelassenen Anwendungsgebiets

Kriterium	Einschluss		Ausschluss		Begründung
Intervention	Behandlung mit Upadacitinib (15 mg) gemäß Fachinformation	E2	<ul style="list-style-type: none"> Behandlung mit Upadacitinib abweichend von der Fachinformation Andere Interventionen 	A2	Zu bewertendes Arzneimittel
Vergleichstherapie	Adalimumab oder Certolizumab Pegol oder Etanercept oder Golimumab	E3	Andere Therapien oder fehlende Übereinstimmung der Vergleichstherapie mit dem deutschen Label	A3	Therapien der ZVT gemäß G-BA-Beratung
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der Verfo
Studientyp	RCT	E5	Nicht RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnis-sicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt
Studiendauer	≥ 24 Wochen	E6	< 24 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt. Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z. B. clinicaltrials.gov) können ebenfalls berücksichtigt werden.	E7	<ul style="list-style-type: none"> Berichterstattung liefert keine ausreichenden Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, 	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich

Kriterium	Einschluss		Ausschluss		Begründung
			die den Kriterien des CONSORT-Statements genügen) <ul style="list-style-type: none"> Mehrfachpublikation ohne relevante Zusatzinformationen^a 		
<p>a: Jeglicher Publikationstyp, in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; CRP: C-reaktives Protein; G-BA: Gemeinsamer Bundesausschuss; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum; RCT: Randomisierte kontrollierte Studie; VerFO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie</p>					

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE (inklusive „in-process & other non-indexed citations“) und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche wird in den Datenbanken MEDLINE, EMBASE und Cochrane durchgeführt, wobei für jede einzelne Datenbank eine sequenzielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wird. Für die Suche wird die Plattform OVID verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche nach RCTs mit dem zu bewertenden Arzneimittel sind im Abschnitt 4.3.1.1.2 dargestellt.

Die verwendeten Suchstrategien und das Datum der Suche sind in Anhang 4-A beschrieben. Die im Volltext ausgeschlossenen Publikationen sind in Anhang 4-C aufgeführt.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

Eine Suche in öffentlich zugänglichen Studienregistern/ Studienergebnisdatenbanken ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern / Studienergebnisdatenbanken vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern/ Studienergebnisdatenbanken muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien

(Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern sowie Studienergebnisdatenbanken immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern/ Studienergebnisdatenbanken [clinicaltrials.gov](https://www.clinicaltrials.gov) (<https://www.clinicaltrials.gov>), EU Clinical Trials Register (EU-CTR, <https://www.clinicaltrialsregister.eu/ctr-search/search>), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, <http://apps.who.int/trialsearch/>), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (<https://clinicaldata.ema.europa.eu>) sowie dem Arzneimittel-Informationssystem (AMIS, www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html)

durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister/ Studienergebnisdatenbank einzeln und mit einer für das jeweilige Studienregister/ Studienergebnisdatenbank adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Die Suche in Studienregistern/Studienergebnisdatenbanken wurde in den Portalen [clinicaltrials.gov](https://www.clinicaltrials.gov) (<https://www.clinicaltrials.gov/>), ICTRP Search Portal, Suchportal der WHO (<http://apps.who.int/trialsearch/>), im EU-CTR (<https://www.clinicaltrialsregister.eu/ctr-search/search>) durchgeführt. Zudem wurde im Clinical Data Suchportal der EMA (<https://clinicaldata.ema.europa.eu>) und im AMIS (vormals AMIS) (www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html) nach Ergebnisberichten zu denjenigen Studien gesucht, die zuvor schon als relevante Studien identifiziert wurden.

Die Ergebnisse der Studienregistersuche nach RCTs mit dem zu bewertenden Arzneimittel sind in Abschnitt 4.3.1.1.3 dargestellt.

Die verwendeten Suchbegriffe und das Datum der Suchen sind im Anhang 4-B beschrieben. Die ausgeschlossenen Studien sind im Anhang 4-D aufgelistet.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen¹. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

¹ Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staack K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. *BMJ* 2015;350:h796

Auf der Internetseite des G-BA wird eine Suche nach Einträgen zu Studien durchgeführt, die bereits anderweitig identifiziert wurden.

Die Ergebnisse der Suche auf der Internetseite des G-BA sind im Abschnitt 4.3.1.1.4 dargestellt.

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Selektion der Treffer aus der Literaturrecherche, der Suche in Studienregistern/Studien-ergebnisdatenbanken und der Suche auf der Internetseite des G-BA wird von zwei Personen unabhängig voneinander durchgeführt. Die Studien werden anhand der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien auf ihre Relevanz geprüft. Diskrepanzen hinsichtlich der Bewertung werden durch Diskussion gelöst.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips

- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Klassifizierung des Verzerrungspotenzials dient als Hilfe zur Beurteilung der Aussagekraft und Ergebnissicherheit der Studien und dient nicht als Ausschlussgrund für Studien.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurde, entfällt die Beschreibung der Aussagekraft der Nachweise.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-

Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)². Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-³ bzw. STROBE-Statements⁴ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Generell erfolgt die Beschreibung von Design und Methodik der eingeschlossenen Studien unter Berücksichtigung des Consolidated Standards of Reporting Trials (CONSORT)-Statements (Items 2b bis 14).

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, entfällt die Beschreibung des Designs und der Methodik der eingeschlossenen Studien.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n)

² Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

³ Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika

Zur Charakterisierung der in den Studien eingeschlossenen Patientenpopulationen werden folgende demografische Merkmale dargestellt:

- Geschlecht (weiblich, männlich)
- Alter (Mittelwert mit Standardabweichung)
- Body Mass Index (BMI) (Mittelwert mit Standardabweichung)
- Abstammung (weiß, schwarz oder afroamerikanisch, asiatisch, Ureinwohner Amerikas oder Alaskas, multiple)
- Geografische Region (Nordamerika, Süd- und Zentralamerika, Westeuropa, Osteuropa, Asien, Andere)

Als krankheitsbeschreibende Faktoren werden folgende Charakteristika dargestellt:

- Dauer der nr-axSpA-Diagnose (Jahre) (Mittelwert mit Standardabweichung)
- Dauer der nr-axSpA-Symptome (Jahre) (Mittelwert mit Standardabweichung)
- Humanes Leukozytenantigen B27 (HLA-B27) (positiv, negativ)
- ASDAS (C-reaktives Protein, CRP) (Mittelwert mit Standardabweichung)
- BASDAI (Mittelwert mit Standardabweichung)
- ASAS Health Index (HI) (Mittelwert mit Standardabweichung)
- Patientenberichteter Rückenschmerz (Patient's Global Assessment of Total Back Pain) (NRS) (Mittelwert mit Standardabweichung)
- Patientenberichteter Rückenschmerz (BASDAI-Item 2) (NRS) (Mittelwert mit Standardabweichung)
- BASFI (Mittelwert mit Standardabweichung)
- SPARCC MRI Spine (Mittelwert mit Standardabweichung)
- SPARCC MRI (sakroiliakale Gelenke) (Mittelwert mit Standardabweichung)
- MASES (Mittelwert mit Standardabweichung)
- Status der Entzündung anhand Magnetresonanztomografie (MRT) (positiv, negativ)

- Hoch sensitives C-reaktives Protein (High sensitivity C-reactive Protein, hsCRP) (Mittelwert mit Standardabweichung; \leq obere Normgrenze [Upper Limit of Normal, ULN [2,87 mg/l]], $>$ ULN)
- ASQoL (Mittelwert mit Standardabweichung)
- Patientenberichtete globale Krankheitsaktivität (Patient's Global Assessment of Disease Activity, PtGADA) (Mittelwert mit Standardabweichung)

Zu Baseline werden zur medizinischen Vorgeschichte und Begleitbehandlung folgende Charakteristika dargestellt:

- Vorbehandlung mit einem bDMARD (ja, nein)
- Begleitende Behandlung mit NSAR (ja, nein)
- Begleitende Behandlung mit konventionellem synthetischem DMARD (Conventional Synthetic DMARD, csDMARD) (ja, nein)
- Begleitende Behandlung mit oralen Kortikosteroiden (ja, nein)
- Behandlung mit oralen Kortikosteroiden zu Baseline (ja, nein)

Darüber hinaus wird die Patientendisposition mit Angabe der Gründe für einen vorzeitigen Abbruch der Behandlung beschrieben und die Dauer der Behandlung dargestellt.

Patientenrelevante Endpunkte

Zur Bewertung des Ausmaßes des medizinischen Nutzens werden Ergebnisse bezüglich der patientenrelevanten Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit herangezogen. Die dazu operationalisierten patientenrelevanten Endpunkte sind im Folgenden dargestellt (Tabelle 4-5).

Tabelle 4-5: Nutzendimensionen und patientenrelevante Endpunkte zur Bewertung des Ausmaßes des medizinischen Nutzens von Upadacitinib

Nutzendimension	Patientenrelevante Endpunkte
Mortalität	<ul style="list-style-type: none"> Anzahl der Todesfälle (Betrachtung im Rahmen der UE)
Morbidität	
<i>Krankheitsaktivität</i>	<ul style="list-style-type: none"> Anteil an Patienten mit einem ASAS40-Ansprechen und einer ASAS-PR Mittlere Veränderung des ASDAS sowie Anteil an Patienten mit einer ASDAS inaktiven Erkrankung und ASDAS niedrigen Krankheitsaktivität Anteil an Patienten mit einem BASDAI50-Ansprechen
<i>Schmerz</i>	<ul style="list-style-type: none"> Mittlere Veränderung des patientenberichteten Rückenschmerzes
<i>Funktion</i>	<ul style="list-style-type: none"> Mittlere Veränderung des BASFI
<i>Entzündung</i>	<ul style="list-style-type: none"> Mittlere Veränderung des SPARCC MRI Spine Mittlere Veränderung des MASES
<i>Fatigue</i>	<ul style="list-style-type: none"> Mittlere Veränderung des FACIT-F
Gesundheitsbezogene Lebensqualität	<ul style="list-style-type: none"> Mittlere Veränderung des ASQoL
Verträglichkeit	<ul style="list-style-type: none"> Unerwünschte Ereignisse (UE) <ul style="list-style-type: none"> Alle UE Schwerwiegende unerwünschte Ereignisse (SUE) UE, die zum Therapieabbruch führten UE, die zum Tod führten UE von speziellem Interesse (UESI) UE mit $\geq 5\%$ in einem Behandlungsarm
ASAS: Assessment of SpondyloArthritis International Society; ASAS-PR: ASAS-partielle Remission; ASDAS: Ankylosing Spondylitis Disease Activity Score; ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; MRI: Magnetic Resonance Imaging; SPARCC: Spondyloarthritis Research Consortium of Canada; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; UESI: Unerwünschtes Ereignis von speziellem Interesse	

Mortalität

Anzahl der Todesfälle

Bei der nr-axSpA handelt es sich um eine chronische Erkrankung, die in der Regel nicht akut lebensbedrohlich ist (3). Daher ist die Betrachtung der Mortalität als prädefinierter Studienendpunkt in dieser Indikation nicht üblich. Die im vorliegenden Dossier präsentierten Ergebnisse zur Mortalität werden somit im Rahmen der Darstellung der unerwünschten Ereignisse (UE) berichtet.

Operationalisierung

Die im vorliegenden Dossier präsentierten Ergebnisse zur Mortalität beruhen auf der Anzahl aller Todesfälle, die während der Beobachtungsdauer aufgetreten sind. Es erfolgt zusätzlich eine narrative Beschreibung der Todesfälle.

Morbidität

Die nr-axSpA ist eine entzündlich rheumatische Erkrankung und gehört neben der AS zur Gruppe der axialen Spondyloarthritis (axSpA). Das Krankheitsbild der axSpA ist sehr heterogen. Besonders betroffen sind die Knochen und Gelenke der unteren Wirbelsäule wie auch die Verbindung zum Becken (Iliosakralgelenk) sowie benachbarte Weichteile, wie z. B. Sehnen und Bänder (3, 4). Patienten mit nr-axSpA leiden unter chronischen Schmerzen und mit zunehmender Krankheitsdauer können Bewegungseinschränkungen des Achsenskeletts hinzukommen (3). Im Gegensatz zur AS zeigen sich jedoch keine strukturellen Veränderungen der Sakroiliakalgelenke im Röntgenbild, sodass die nr-axSpA als Frühform der AS betrachtet werden kann. Bei etwa 10 – 40 % der nr-axSpA Patienten findet über einen Zeitraum von 2 – 10 Jahren ein Progress zur AS statt (3, 5). In den ersten Jahren kommt es vorwiegend zu Schmerzen an der Wirbelsäule und einer variablen extraskelettalen Beteiligung. Das Leitsymptom der nr-axSpA ist ein tiefsitzender, entzündlicher, teils chronifizierter Rückenschmerz (3, 6). Mit zunehmender Krankheitsdauer und unter Progress zur AS kann es zu Verknöcherungen, insbesondere im Bereich des Achsenskeletts, kommen, was zu weiteren Einschränkungen der Funktionsfähigkeit, der Lebensqualität und des sozialen Lebens führt (3). Das primäre Therapieziel bei der Behandlung von Patienten mit nr-axSpA ist gemäß der Leitlinie der ASAS und der EULAR sowie der deutschen S3-Leitlinie die langfristige Verbesserung der Erkrankung und Erhaltung der gesundheitsbezogenen Lebensqualität. Diese Ziele sollen sowohl durch die Kontrolle der Entzündung und deren Symptome als auch durch die Unterbindung der fortschreitenden strukturellen Veränderungen und den Erhalt der körperlichen Funktionsfähigkeit sowie des sozialen Lebens und der Arbeitsfähigkeit erreicht werden (3, 6).

Den Therapieempfehlungen der Leitlinien liegt der sogenannte Treat-to-Target-Ansatz zugrunde (7). Ein wichtiger Aspekt dieses Konzepts ist, dass die Behandlung gemäß einem vordefinierten Behandlungsziel erfolgt, welches gemeinsam von Patienten und Rheumatologen bestimmt wird (6). Das Konzept wurde 2012 von einem internationalen Team aus Rheumatologen, Dermatologen, Patienten und Gesundheitsexperten für die axiale und periphere Spondyloarthritis entwickelt und 2017 aktualisiert. Es sieht vor, dass das Behandlungsziel, anhand dessen die Therapie überprüft wird, die Remission oder niedrige Krankheitsaktivität sein sollte (7, 8). Auch die aktuelle deutsche S3-Leitlinie sieht bei der Behandlung von Patienten mit nr-axSpA die Reduktion der Krankheitsaktivität und das Erreichen einer klinischen Remission an zentraler Stelle (3). Neben der Remission sind wichtige Ziele die Schmerzreduktion, der Erhalt der körperlichen Funktionsfähigkeit, die Reduktion der Steifigkeit sowie die Verhinderung struktureller Läsionen und der Erhalt der Arbeits- und Erwerbsfähigkeit. In der S3-Leitlinie werden, wie in Tabelle 4-6 berichtet, allgemeine Empfehlungen zu den möglichen Instrumenten für die systematische Erfassung von

Therapiezielen gemacht (3). Die Erfassung der verschiedenen Endpunkte wurde initial für die AS erarbeitet, sie finden im Analogieschluss jedoch auch in der nr-axSpA Anwendung (9).

Tabelle 4-6: Möglichkeiten der systematischen Erfassung von Therapiezielen gemäß S3-Leitlinie

Domäne	Instrument
Allgemeines Patientenurteil	Skala ^a für allgemeine Krankheitsaktivität in der vorhergehenden Woche
Entzündungsparameter	CRP, BSG
Ermüdbarkeit (Fatigue)	BASDAI-Item 1 (Müdigkeit und Erschöpfung)
körperliche Funktionsfähigkeit	BASFI
Periphere Gelenke und Sehnen	Anzahl geschwollener Gelenke (44 Gelenke), validierter Enthesitis-Score (z. B. MASES, San Francisco und Berlin)
Schmerz	Skala ^a für: Nächtliche Schmerzen der Wirbelsäule in der vorhergehenden Woche wegen AS und Schmerzen der Wirbelsäule in der vorhergehenden Woche (s. oben) wegen AS
Steifigkeit	Skala ^a für Dauer Morgensteifigkeit der Wirbelsäule in der vorhergehenden Woche
Wirbelsäulenbeweglichkeit	<u>Thoraxexkursion:</u> Und modifizierter Schober Und Okziput-Wand-Abstand Und zervikale Rotation Und laterale lumbale Flexion oder BASMI
Arbeits- und Erwerbsfähigkeit	Arbeitsunfähigkeitsdauer Subjektive Arbeitsfähigkeit (Work-Ability-Index)
a: Skala ist entweder eine NRS oder VAS AS: Ankylosierende Spondylitis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; BSG: Blutsenkungsgeschwindigkeit; CRP: C-reaktives Protein; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; NRS: Numerical Rating Scale; VAS: Visuelle Analogskala Quelle: (3)	

Zur Bewertung der Morbidität werden im vorliegenden Dossier Messinstrumente zur Erfassung der Krankheitsaktivität (ASAS-Ansprechkriterien, ASDAS, BASDAI), des Schmerzes (patientenberichteter Rückenschmerz), der Funktion bzw. körperlichen Funktionsfähigkeit (BASFI), der Entzündung bzw. Entzündungsparameter (SPARCC MRI Spine, MASES) und der Fatigue (FACIT-F) betrachtet. Diese werden im Folgenden detailliert erläutert.

Krankheitsaktivität

Unbehandelt oder unzureichend behandelt kann die nr-axSpA zu einer erheblichen gesundheitlichen Belastung der Patienten führen, deren Ausmaß hinsichtlich Schmerz und Funktionseinschränkungen sowie Einschränkungen der Lebensqualität und des sozialen Lebens mit der AS vergleichbar ist (10-12). Mit zunehmender Krankheitsdauer droht darüber hinaus ein Progress zur AS, der mit Verknöcherungen, insbesondere im Bereich des Achsenskeletts einhergehen kann (3). Dabei spielt die Krankheitsaktivität eine zentrale Rolle. So konnte in einer Kohortenstudie gezeigt werden, dass mit höherer Krankheitsaktivität die strukturellen Schädigungen der Wirbelsäule zunehmen (13). Durch eine frühe und effektive entzündungshemmende Therapie von nr-axSpA-Patienten, könnten solche dauerhaften Schädigungen verhindert werden (14). Aus diesem Grund ist die Reduktion und Kontrolle der Krankheitsaktivität ein wichtiges Therapieziel für nr-axSpA-Patienten. Neben dem Erreichen einer klinischen Remission steht die Reduktion der Krankheitsaktivität an zentraler Stelle (3). Dementsprechend empfiehlt das Committee for Medicinal Products for Human Use (CHMP), dass in klinischen Studien zur nr-axSpA die Reduktion der Krankheitsaktivität erhoben wird (15). Die Patientenrelevanz der Krankheitsaktivität wurde auch durch den G-BA im Beratungsgespräch bestätigt (2).

Zu den gängigen Messinstrumenten für die Erhebung der Krankheitsaktivität gehören die ASAS-Ansprechkriterien, der ASDAS sowie der BASDAI (3, 15).

ASAS-Ansprechkriterien

Die ASAS-Ansprechkriterien können als Maß für die symptomatische Verbesserung der Erkrankung über die verschiedenen Krankheitsdomänen einer axSpA hinweg herangezogen werden und finden entsprechend Anwendung in der nr-axSpA. Die 2001 definierten vier ASAS-Domänen umfassen das Patientenurteil, Schmerz, Funktion und Entzündung und basieren auf Patientenberichten (16, 17). Sie werden häufig in klinischen Studien verwendet, um die Wirksamkeit einer Behandlung zu beurteilen (15).

Operationalisierung

Die ASAS-Ansprechkriterien setzen sich aus den Domänen PtGADA, patientenberichteter Rückenschmerz, Funktion (BASFI) und der Entzündung (abgeleitet aus den BASDAI-Items 5 und 6 zur Morgensteifigkeit) zusammen. Die Domänen der ASAS-Ansprechkriterien werden jeweils auf einer NRS von 0 bis 10 erhoben und zusammengefasst. Somit ergibt sich eine Skala von 0 bis 40, wobei höhere Werte eine höhere Krankheitsaktivität darstellen. Ein Ansprechen gemäß den ASAS-Kriterien wird in folgenden Kategorien berichtet (17-19):

- ASAS40-Ansprechen: Verbesserung um ≥ 40 % sowie eine absolute Verbesserung um ≥ 2 Einheiten gegenüber Baseline in mindestens 3 der 4 Domänen, wobei es in keiner Domäne zu einer Verschlechterung kommen darf
- ASAS-PR: Wert von maximal 2 in allen Einzeldomänen

Im vorliegenden Dossier werden die Anteile der Patienten mit einem ASAS40-Ansprechen bzw. einer ASAS-PR dargestellt.

Patientenrelevanz

Die ASAS-Kriterien umfassen die vom Patienten wahrgenommene globale Krankheitsaktivität und Funktionsfähigkeit sowie die Symptome Rückenschmerzen und Morgensteifigkeit. Damit decken sie mehrere verschiedene Kerndomänen ab, die in der nr-axSpA eine wichtige Rolle spielen. Es wurde zudem gezeigt, dass ein ASAS40-Ansprechen mit signifikanten Verbesserungen der patientenberichteten Einschätzung ihrer körperlichen Funktion, ihrer Lebensqualität und ihrer Arbeitsproduktivität verbunden ist (20-22). Dementsprechend wird vom CHMP in seiner Leitlinie die Verwendung des ASAS40-Ansprechens als primärer Endpunkt in klinischen Studien empfohlen (15). Auch in den Empfehlungen von ASAS und EULAR werden das ASAS40-Ansprechen und die ASAS-PR als Zielkriterien für die Erfassung der Wirksamkeit von nr-axSpA-Medikamenten genannt (6). Die ASAS-PR wird zudem von der S3-Leitlinie als Endpunkt zur Erfassung der Remission empfohlen (3). Die ASAS-PR wurde in zahlreichen Studien in der Indikation AS und nr-axSpA als Prädiktor für ein klinisches Ansprechen definiert und als primärer oder sekundärer Endpunkt verwendet (23-29). Alle vier Domänen der ASAS-Kriterien werden mittels patientenberichteter Fragebögen beurteilt und bilden somit die subjektive Einschätzung der Patienten selbst ab. Dies spricht für eine hohe Patientenrelevanz des ASAS40-Ansprechens sowie der ASAS-PR.

Validität

Die ASAS-Kriterien wurden speziell für die Erfassung einer symptomatischen Verbesserung bei AS-Patienten entwickelt, sind validiert und im Analogieschluss auf nr-axSpA-Patienten übertragbar. Zunächst wurde ihre Eignung für die Unterscheidung zwischen einer Behandlung mit einem NSAR und Placebo gezeigt. Später wurden die ASAS-Kriterien auch anhand von zwei klinischen Studien mit TNF-Inhibitoren validiert, wobei das ASAS40-Ansprechen erstmals definiert wurde (16, 30).

Ankylosing Spondylitis Disease Activity Score (ASDAS)

Der ASDAS wurde 2008 zur Messung der Krankheitsaktivität mittels eines zusammengesetzten Scores entwickelt, der sowohl vom Patienten berichtete Items als auch objektive Messungen umfasst. Ursprünglich wurden von der ASAS vier verschiedene Versionen des ASDAS entwickelt, unter denen später auf Basis der einfachsten Umsetzbarkeit eine bevorzugte Variante ausgewählt wurde. Der Score ist ein aus fünf Items zusammengesetzter Endpunkt, der einen gemeinsamen Messwert für die Krankheitsaktivität ergibt (31). Der ASDAS wurde für die AS entwickelt, besitzt aber eine ebenso gute Aussagekraft über die Krankheitsaktivität der nr-axSpA (32). Der Endpunkt umfasst neben einem Laborparameter patientenberichtete Einschätzungen zur Krankheitsaktivität, zu Rückenschmerzen, Schwellungen/Schmerzen der peripheren Gelenke und Morgensteifigkeit (31, 33, 34).

Operationalisierung

Der ASDAS setzt sich aus den BASDAI-Items 2 (Rückenschmerz), 3 (Schwellungen/Schmerzen in peripheren Gelenken) und 6 (Morgensteifigkeit), dem PtGADA und dem hsCRP zusammen und wird mit folgender Formel errechnet (19, 33, 34):

- $ASDAS = 0,121 \cdot \text{BASDAI-Item 2} + 0,110 \cdot \text{PtGADA} + 0,073 \cdot \text{BASDAI-Item 3} + 0,058 \cdot \text{BASDAI-Item 6} + 0,579 \cdot \ln(\text{hsCRP} + 1)$

Der ASDAS ist eine bei 0 beginnende und nach oben offene Skala. Dabei stehen höhere Werte für eine höhere Krankheitsaktivität (31). Er lässt sich wie folgt kategorisieren (18, 35, 36):

- $ASDAS < 1,3$: Inaktive Erkrankung
- $1,3 \leq ASDAS < 2,1$: Niedrige Krankheitsaktivität
- $2,1 \leq ASDAS \leq 3,5$: Hohe Krankheitsaktivität
- $ASDAS > 3,5$: Sehr hohe Krankheitsaktivität

Im vorliegenden Dossier wird die mittlere Veränderung des ASDAS zu Baseline dargestellt sowie die Anteile der Patienten, die eine inaktive Erkrankung bzw. eine niedrige Krankheitsaktivität anhand des ASDAS von $< 1,3$ bzw. $< 2,1$ erreichen.

Patientenrelevanz

Der ASDAS wurde zur Messung der Krankheitsaktivität bei AS-Patienten (einschließlich nr-axSpA) entwickelt (32, 33). Er stellt gemäß den Empfehlungen einer internationalen Task Force aus Ärzten und Patienten ein bevorzugtes Messinstrument in der axSpA dar (7). In der deutschen S3-Leitlinie für die axSpA wird der ASDAS zur Erfassung der Krankheitsaktivität empfohlen, speziell wird der $ASDAS < 1,3$ und der $ASDAS < 2,1$ zur Erfassung der Remission bzw. niedrigen Krankheitsaktivität empfohlen (3). Auch in den Empfehlungen von ASAS und EULAR zur Behandlung der axSpA wird der ASDAS als relevantes Instrument zur Messung der Krankheitsaktivität und des Ansprechens auf bDMARDs beschrieben (6). Ebenso empfiehlt das CHMP in seiner axSpA-Leitlinie die Verwendung des ASDAS als eine einfache zusammengesetzte Skala zur Erhebung der Krankheitsaktivität und zur Bewertung der Wirksamkeit von Therapien (15). Seine Patientenrelevanz wird dadurch unterstrichen, dass es sich bei 4 von 5 der Faktoren, die in den Gesamtscore eingehen, um patientenberichtete Parameter handelt (37). Es wurde zudem gezeigt, dass die mittels des ASDAS bestimmte Krankheitsaktivität mit einer Verbesserung in patientenberichteten Endpunkten zu Krankheitsaktivität, Funktion und Lebensqualität sowie mit der radiologischen Progression von Schäden an der Wirbelsäule korreliert (13, 20, 22, 38). Der ASDAS ist ein schnell und einfach anzuwendendes Messinstrument und wird häufig in klinischen Studien in der Indikation axSpA eingesetzt (31).

Validität

Der ASDAS ist ein in den Indikationen AS und nr-axSpA umfassend validiertes Messinstrument, welches sowohl dafür eingesetzt werden kann, zwischen Patienten mit niedriger und hoher Krankheitsaktivität zu unterscheiden, als auch um in klinischen Studien Veränderungen gegenüber Baseline und zwischen verschiedenen Behandlungsarmen zu erfassen (29, 31-34, 39-41).

Der ASDAS und die unter dem Abschnitt Operationalisierung beschriebenen Schwellenwerte sind von der ASAS und von der Outcome Measures in Rheumatology (OMERACT)-Initiative anerkannt (37).

Bath ankylosing Spondylitis Disease Activity Index (BASDAI)

Die Bath-Indizes wurden spezifisch für Patienten mit AS entwickelt, um die Krankheitsaktivität (BASDAI), die körperliche Funktionsfähigkeit (BASFI) und die Beweglichkeit des axialen Skeletts (Bath Ankylosing Spondylitis Metrology Index, BASMI) zu erfassen (31). Mit Ausnahme des BASMI werden sie in klinischen Studien auch zur Beurteilung der Krankheitsaktivität in der nr-axSpA herangezogen (29, 42). Der 1994 entwickelte BASDAI besteht aus sechs Items zur qualitativen und quantitativen Messung der Fatigue, Schmerzen und Morgensteifigkeit anhand der Einschätzung der Patienten über die letzte Woche (43).

Operationalisierung

Der BASDAI setzt sich aus sechs Items zusammen, die jeweils mit einer NRS von 0 bis 10 erfasst werden und folgende Bereiche beinhalten:

- Item 1: Ausprägung der Fatigue/Müdigkeit
- Item 2: Ausprägung der krankheitsbedingten Nacken-, Rücken- oder Hüftschmerzen
- Item 3: Ausprägung der Gelenkschmerzen/-schwellung in anderen Gelenken als den unter Item 2 (periphere Arthritis)
- Item 4: Ausprägung des Druckschmerzes in betroffenen Körperregionen (Enthesitis)
- Item 5: Intensität der Morgensteifigkeit
- Item 6: Dauer der Morgensteifigkeit

Der Gesamtscore für den BASDAI wird anhand der sechs Einzelitems mit folgender Formel berechnet:

- $BASDAI = 0,2 \cdot (\text{Item 1} + \text{Item 2} + \text{Item 3} + \text{Item 4} + 0,5 \cdot \text{Item 5} + 0,5 \cdot \text{Item 6})$

Der Gesamtscore hat eine Spannweite von 0 bis 10, wobei höhere Werte für eine höhere Krankheitsaktivität stehen (18, 19, 43). Im vorliegenden Dossier wird der Anteil der Patienten mit einem BASDAI50-Ansprechen dargestellt, was einer Verbesserung des BASDAI um $\geq 50\%$ gegenüber dem Baseline-Wert entspricht.

Patientenrelevanz

Der BASDAI ist Teil des ASAS-Core-Sets und wird in der deutschen S3-Leitlinie für die axSpA und den Empfehlungen von ASAS und EULAR zur Erfassung der Krankheitsaktivität empfohlen (3, 6, 44). Auch das CHMP empfiehlt in seiner axSpA-Leitlinie die Verwendung des BASDAI als eine einfache zusammengesetzte Skala zur Erhebung der Krankheitsaktivität und zur Bewertung der Wirksamkeit von Therapien (15). Der BASDAI ist ein

patientenberichteter Fragebogen, alle sechs Items werden vom Patienten selbst bewertet. Es wurde gezeigt, dass die mittels des BASDAI ermittelte Krankheitsaktivität einerseits mit der Schmerzintensität, aber auch mit psychiatrischen Symptomen wie z. B. Depressionen und Angstzuständen korreliert (45-48). Weiterhin konnte eine Korrelation des BASDAI mit Messinstrumenten für die Funktionsfähigkeit (BASFI) und die Lebensqualität (Short Form – 36 [SF-36], ASQoL) gezeigt werden (49-54). In früheren Nutzenbewertungsverfahren in der Indikation PsA wurde der BASDAI als patientenrelevanter Endpunkt vom G-BA anerkannt (55, 56). Daher ist davon auszugehen, dass der BASDAI auch in der nr-axSpA für die Erfassung der Krankheitsaktivität als patientenrelevanter Endpunkt geeignet ist.

Validität

Der BASDAI wurde für den Einsatz in der AS entwickelt und darüber hinaus auch für die Beurteilung der Krankheitsaktivität in der nr-axSpA validiert (32, 43, 57-62). Somit ist der BASDAI im gesamten Spektrum der axSpA ein häufig eingesetztes Instrument, das auch von der ASAS empfohlen wird (31, 63). Eine valide Bewertung der einzelnen Items kann sowohl anhand einer visuellen Analogskala (VAS) als auch mithilfe einer NRS erfolgen (64). Eine Verbesserung des BASDAI um mindestens 50 % im Vergleich zum Baseline-Wert stellt eine bedeutsame Verbesserung dar und wird von der EMA als geeignet angesehen, um die klinische Wirksamkeit einer Therapie zu bewerten (15, 31, 65, 66).

Schmerz

Das Leitsymptom der nr-axSpA sind entzündliche, teilweise bereits chronifizierte Rückenschmerzen (3, 6). Diese entzündlichen Rückenschmerzen werden durch Entzündungen in den Sakroiliakgelenken sowie anderen Positionen des Achsenskeletts verursacht (3). Circa 75 % der axSpA-Patienten leiden unter entzündlichen Rückenschmerzen, weshalb gerade dieser Form eine besondere Bedeutung zukommt (3, 67, 68).

Patientenberichteter Rückenschmerz

Dem Rückenschmerz als zentralem Symptom der nr-axSpA kommt eine besondere Bedeutung zu. Daher wird dessen Ausmaß anhand einer patientenberichteten NRS beurteilt und als separater Endpunkt erfasst.

Operationalisierung

Der patientenberichtete Rückenschmerz wird im vorliegenden Dossier anhand einer patientenberichteten NRS erhoben, die den Rückenschmerz insgesamt (Patient's Assessment of Total Back Pain) berücksichtigt. Der Gesamtscore ergibt sich durch die Bewertung von 0 (keine Schmerzen) bis 10 (schwerste vorstellbare Schmerzen) (18, 43). Es wird die mittlere Veränderung des Gesamtscores zu Baseline dargestellt.

Patientenrelevanz

Das Leitsymptom der nr-axSpA sind tiefsitzende entzündliche, teils chronifizierte Rückenschmerzen (3, 6). Viele Patienten leiden sehr unter den Schmerzen und erhoffen sich

von einer Behandlung insbesondere deren Eliminierung oder Reduktion (69). Deshalb hat in der Indikation nr-axSpA gerade diese entzündliche Form des Rückenschmerzes eine besondere Relevanz für die Patienten (3) und muss als patientenrelevanter Endpunkt gewertet werden.

Validität

Die Erfassung des Rückenschmerzes anhand einer patientenberichteten NRS ist etabliert und bildet die Grundlage für eine korrekte Diagnose des entzündlichen Rückenschmerzes bei der axSpA (62, 70). Die NRS zur Erfassung von patientenberichteten Rückenschmerzen ist somit als valides Erhebungsinstrument anzusehen und wurde bereits in vergangenen Nutzenbewertungen regelhaft vom G-BA berücksichtigt (71).

Funktion

Die nr-axSpA kann zu Einschränkungen der körperlichen Funktionsfähigkeit und Mobilität des Achsenskeletts führen. Dementsprechend empfiehlt die Leitlinie die Messung dieser Parameter u. a. anhand des BASFI (Funktionsfähigkeit).

Bath ankylosing spondylitis functional index (BASFI)

Der BASFI wurde 1994 entwickelt, um für die Überwachung der körperlichen Funktionsfähigkeit ein validiertes Messinstrument spezifisch für AS-Patienten zur Verfügung zu haben. Im Analogieschluss wird er auch für die Bewertung von nr-axSpA-Patienten herangezogen (3, 62). Der patientenberichtete Fragebogen erfasst die Funktionsfähigkeit und die Fähigkeit zur Bewältigung von Alltagsaktivitäten (72).

Operationalisierung

Mit dem BASFI wird mittels zehn Fragen die körperliche Funktionsfähigkeit gemessen. Dabei beziehen sich acht Fragen auf die Domänen Bücken, Strecken, Stehen, Drehen, Positionswechsel und Treppensteigen. Die restlichen zwei Fragen beziehen sich auf eine aktive Tagesgestaltung und die Fähigkeit zur körperlichen Anstrengung z. B. beim Sport oder der Gartenarbeit (19, 31). Die Fragen werden jeweils auf einer Skala von 0 (einfach) bis 10 (unmöglich) bewertet und ergeben als Mittelwert der zehn Items den Gesamtscore. Für die Berechnung des BASFI müssen mindestens acht der zehn Fragen beantwortet werden. Bei maximal zwei fehlenden Angaben können diese über den Mittelwert der restlichen Fragen imputiert werden (18). Im vorliegenden Dossier wird die mittlere Veränderung des BASFI zu Baseline dargestellt.

Patientenrelevanz

Neben Schmerzen und Steifheit gehören auch körperliche Funktionsbeeinträchtigungen zu den Hauptbeschwerden von nr-axSpA-Patienten. Der Erhalt der körperlichen Funktionsfähigkeit ist eines der wichtigsten Ziele der Therapie der nr-axSpA und somit ist der BASFI als Maß hierfür patientenrelevant (3). In der deutschen S3-Leitlinie wird zur Messung der körperlichen Funktionsfähigkeit der BASFI empfohlen, welcher daher als Instrument der Wahl zur Erhebung dieses patientenrelevanten Endpunktes angesehen werden kann (3). Der Fragebogen wird vom

Patienten selbst ausgefüllt, womit der Grad der körperlichen Einschränkung der nr-axSpA-Patienten aus ihrer persönlichen Sicht gemessen werden kann. Es wurde gezeigt, dass die mittels des BASFI ermittelte Funktionsfähigkeit in den Indikationen AS und nr-axSpA sowohl mit der Krankheitsaktivität gemäß BASDAI und der Beweglichkeit des axialen Skeletts gemäß BASMI, als auch mit weiteren Parametern und Messinstrumenten für Steifigkeit und Schmerzen korreliert (51, 52, 73, 74). Weiterhin konnte eine Korrelation des BASFI mit der gesundheitsbezogenen Lebensqualität (SF-36) gezeigt werden (49, 50, 75). Auch der G-BA hat im Beratungsgespräch (Beratungsanforderung 2018-B-186) die Patientenrelevanz von Morbiditätseindpunkten, die Bewegungs- und Funktionseinschränkungen erfassen, bestätigt (2).

Validität

Der BASFI ist ein international anerkanntes, valides Erhebungsinstrument der körperlichen Funktionsfähigkeit bei axSpA-Patienten, dessen Verwendung auch von der deutschen S3-Leitlinie empfohlen wird (3). Das Instrument wird seit vielen Jahren in der Routineversorgung und in klinischen Studien der AS und nr-axSpA eingesetzt und wurde mehrfach validiert (3, 31, 62, 72-74, 76-78).

Entzündung

Die nr-axSpA ist von Entzündungen der Gelenke und Enthesen geprägt, die die Symptome begründen. In diesem Dossier wird daher der SPARCC MRI Spine und der MASES dargestellt.

Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging Spine (SPARCC MRI Spine)

Bei Patienten mit nr-axSpA führen die entzündlichen Prozesse in der Wirbelsäule zu Einschränkungen der Wirbelsäulenbeweglichkeit und können im Fall eines Progresses zur AS langfristig Schädigungen der Knochen in Form von Sklerose, Erosion, Ausbildung von Syndesmophyten und Ankylose verursachen. Doch bereits ohne strukturelle Schäden sind nr-axSpA-Patienten durch Bewegungseinschränkungen und Schmerzen stark beeinträchtigt. Der SPARCC MRI Spine wurde 2005 entwickelt, um die frühen Entzündungsprozesse in der Wirbelsäule bei Patienten mit AS schnell und einfach in einem gemeinsamen Score erfassen zu können (79, 80).

Operationalisierung

Der SPARCC MRI Spine berechnet sich aus einer dichotomen Bewertung zu dem Vorhandensein, der Tiefe und der Intensität von Knochenmarködem-Läsionen der Wirbelsäule. Dafür wird mittels MRT die gesamte Wirbelsäule mit der Short-Tau-Inversion-Recovery (STIR)-Bildsequenz nach Entzündungen (Knochenmarködemen) untersucht. Pro Patient und Zeitpunkt werden von jedem Gutachter 23 discovertebrale Einheiten (Discovertebral Units, DVUs) beurteilt und die sechs am schwersten betroffenen DVUs werden von jedem Gutachter ausgewählt und zur Berechnung des SPARCC MRI Spine herangezogen. Für jede dieser sechs DVUs werden drei aufeinanderfolgende sagittale Schnitte in vier Quadranten ausgewertet, um das Ausmaß der Entzündung in allen drei Dimensionen zu ermitteln. Jeder Quadrant wird nach

dem Auftreten eines erhöhten Signals auf STIR entweder mit 0 = „normales Signal“ oder mit 1 = „erhöhtes Signal“ bewertet. Das Auftreten einer Läsion mit erhöhter Signalintensität in einer beliebigen DVU auf jedem der sagittalen Schnitte wird mit einem zusätzlichen Score von 1 bewertet. Schnitte mit einer Läsion, die ein kontinuierlich erhöhtes Signal von einer Tiefe ≥ 1 cm von der Endplatte ausgehend zeigen, sind mit +1 pro Schnitt zu bewerten. Jeder Schnitt kann mit einem Score von maximal 6 bewertet werden, wobei die maximale Punktzahl für alle 6 DVUs 108 beträgt (18). Im vorliegenden Dossier wird die mittlere Veränderung des SPARCC MRI Spine dargestellt.

Patientenrelevanz

Eine charakteristische Eigenschaft der nr-axSpA sind aktive axiale Entzündungen in Form von Knochenmarködemen. Diese entzündlichen Veränderungen der Wirbelsäule können im MRT besser und früher erkannt werden als durch einfache Radiografie. So können im MRT unter Verwendung der STIR-Sequenz auch Knochenmarködeme als Marker für Entzündungen direkt sichtbar gemacht werden, während mittels Radiografie nur die resultierenden Schädigungen der Knochen erfasst werden. Dies erlaubt eine objektive Überwachung der aktiven Erkrankung und der Veränderungen der Krankheitsaktivität innerhalb kürzerer Zeiträume (79-81). Das CHMP empfiehlt in seiner axSpA-Leitlinie die Verwendung des SPARCC MRI Spine zur Messung von entzündlichen Veränderungen (15). Die frühe Erkennung einer Verschlechterung der Krankheitsaktivität ist für den Patienten relevant, da durch Anpassungen der Therapie langfristig Schädigungen und Beweglichkeitseinschränkungen vermieden werden können.

Validität

Der SPARCC MRI Spine wurde speziell für die Untersuchung der Krankheitsaktivität anhand von Entzündungen in der Wirbelsäule entwickelt und in der Indikation AS validiert. Es wurden eine gute Reproduzierbarkeit und eine große Änderungssensibilität gezeigt sowie die Fähigkeit zwischen Behandlungsgruppen zu unterscheiden (79, 80). Insbesondere in der Indikation nr-axSpA, bei der noch keine röntgenologisch nachweisbaren strukturellen Veränderungen vorliegen, ist eine Überwachung der Krankheitsaktivität mithilfe von MRT nützlich. Dementsprechend wird der SPARCC MRI Spine als validiertes Instrument für die Messung der Veränderungen im MRT auch vom CHMP für den Einsatz in klinischen Studien empfohlen (15).

Maastricht Ankylosing Spondylitis Enthesitis Score (MASES)

Der MASES ist ein Maß für die Entzündungen der Sehnenansätze (Enthesitis) bei Patienten mit axSpA, welcher 2003 spezifisch für die Indikation AS entwickelt wurde und mittlerweile auch in der nr-axSpA Anwendung findet (82-85). An einer Enthesitis leiden ca. 30 % der nr-axSpA-Patienten (3, 86). Neben einer Arthritis und einer Daktylitis kann diese Schmerzen und/oder Schwellungen der Gelenke sowie der gelenknahen Strukturen verursachen und erheblich zur Krankheitslast beitragen (87, 88).

Operationalisierung

Der MASES bewertet die Abwesenheit (0) oder das Vorliegen (1) einer Enthesitis an 13 verschiedenen Stellen. Der MASES hat demzufolge eine Spannweite von 0 bis 13, wobei

höhere Werte für eine stärkere Ausbreitung der Enthesitis stehen (18, 19, 83). Im vorliegenden Dossier wird die mittlere Veränderung des MASES im Vergleich zu Baseline dargestellt.

Patientenrelevanz

Bei einer Enthesitis kommt es zu einer Entzündung der Entesen (Sehnenansätze) einhergehend mit Schmerzen und eingeschränkter Funktion. Die deutsche S3-Leitlinie empfiehlt die Verwendung von validierten Instrumenten zur Erfassung der Enthesitis, wie z. B. den MASES (3). Ebenso empfiehlt das CHMP der EMA die Verwendung des MASES (3, 15, 82). Es besteht ein direkter Zusammenhang zwischen dem Vorliegen einer Enthesitis und einer höheren Krankheitsaktivität sowie einer damit einhergehenden niedrigeren Lebensqualität. So wurde beispielsweise gezeigt, dass der MASES mit dem BASDAI und dem BASFI korreliert, welche patientenrelevante Veränderungen in der Krankheitsaktivität bzw. der körperlichen Funktionsfähigkeit messen. Auch mit Symptomen wie Schmerzen und Müdigkeit und mit der Lebensqualität (ASQoL, SF-36) wurden Korrelationen festgestellt (84, 89-92). Die Reduktion der Krankheitsaktivität und die Verminderung von Schmerzen stehen an zentraler Stelle in der Therapie der nr-axSpA; die Ergebnisse des MASES sind hierfür ein direktes Maß, woraus sich die Patientenrelevanz des Erhebungsinstruments ergibt (3). Der G-BA hat die Messung der Enthesitis in früheren Nutzenbewertungen zur PsA als patientenrelevanten Endpunkt eingestuft (55, 56). Auch in der nr-axSpA erachtet der G-BA extraskeletale Manifestationen (z. B. Enthesitis) als patientenrelevant (2).

Validität

Der MASES ist ein international etabliertes, seit vielen Jahren verwendetes, validiertes Instrument zur Beurteilung des Ausmaßes der Enthesitis bei axSpA-Patienten. Das CHMP der EMA und die deutsche S3-Leitlinie empfehlen daher die Verwendung dieses Instruments sowohl in der Indikation AS als auch nr-axSpA (3, 15, 82).

Fatigue

Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)

Fatigue ist ein subjektiver Zustand überwältigender, anhaltender Erschöpfung und verminderter Fähigkeit zu körperlicher und geistiger Arbeit, der durch Ruhe nicht gelindert wird. Auch bei nr-axSpA-Patienten ist es ein häufig auftretendes und wichtiges Symptom (3, 93, 94). Der FACIT-F ist ein Fragebogen, der mithilfe von 13 Fragen das Ausmaß der Fatigue und ihre Auswirkungen auf die täglichen Aktivitäten und Funktionen ermittelt (95-97). Er wurde ursprünglich 1997 entwickelt, um eine präzisere Bewertung der Fatigue im Zusammenhang mit Anämie bei Krebspatienten zu erhalten (98).

Operationalisierung

Die 13 Items des FACIT-F-Fragebogens werden mittels einer fünfstufigen Likert-Skala erfasst. Bewertet werden die Fragen mit einer Punkteskala von 4 („gar nicht“) bis 0 („sehr“). Der Gesamtscore des FACIT-F ergibt sich aus der Summe aller Items, welche mit 13 multipliziert und durch die Anzahl der beantworteten Items geteilt wird. Es müssen mindestens 7 der

13 Items beantwortet sein. Der Gesamtscore hat eine Spannweite von 0 bis 52, wobei ein höherer Wert eine geringere Fatigue bedeutet (18). Im vorliegenden Dossier werden für den FACIT-F Analysen zur mittleren Veränderung zu Baseline vorgelegt.

Patientenrelevanz

Fatigue wird von Patienten mit entzündlichen rheumatischen Erkrankungen nach Schmerz und Steifigkeit als dritthäufigstes Symptom genannt (3). Etwa zwei Drittel der axSpA-Patienten leiden unter einer schweren Fatigue-Symptomatik, wobei kein Unterschied zwischen den Indikationen AS und nr-axSpA besteht (93). Fatigue wird direkt von nr-axSpA-Patienten erfahren und kann mit Einschränkungen der Funktionalität einhergehen sowie einen erheblichen Einfluss auf Arbeitsfähigkeit und Lebensqualität haben (3, 93). Fatigue wird von den Patienten als wichtiges Symptom erachtet und ist somit unmittelbar patientenrelevant (99, 100). Auch der G-BA sieht Fatigue als einen patientenrelevanten krankheitsbedingten Endpunkt für die nr-axSpA an.

Validität

Der FACIT-F ist für axSpA-Patienten ein zuverlässiges, valides Messinstrument zur Bewertung der Fatigue (95, 96). Es konnte gezeigt werden, dass Veränderungen in der Fatigue mit Veränderungen in der Krankheitsaktivität korrelieren (93, 97). Im Beratungsgespräch (Beratungsanforderung 2018-B-186) bestätigte der G-BA, dass es sich beim FACIT-F um ein geeignetes Messinstrument zur Erhebung des patientenrelevanten Endpunktes Fatigue in der Indikation nr-axSpA handelt (2).

Gesundheitsbezogene Lebensqualität

Die Lebensqualität setzt sich aus einzelnen Parametern zusammen, die umschreiben, wie der Patient in seinem Alltag durch die Erkrankung beeinträchtigt wird. Hauptbeschwerden der nr-axSpA-Patienten sind Schmerzen, Steifigkeit, Erschöpfung und eine Beeinträchtigung in der Aktivität und des sozialen sowie beruflichen Lebens (101). In den ersten Jahren der Erkrankung kommt es vorwiegend zu Schmerzen an der Wirbelsäule und einer variablen extraskelletalen Beteiligung. Bereits das Vorhandensein von extraskelletalen Manifestationen hat einen wesentlichen Einfluss auf die Lebensqualität der Patienten, hier scheint insbesondere die psychische Gesundheit im Fokus zu liegen (3). Mit zunehmender Krankheitsdauer kann es unter einem Progress zur AS zu Verknöcherungen, insbesondere im Bereich des Achsenskeletts, kommen, was zu weiteren Einschränkungen der Funktionsfähigkeit, der Lebensqualität und des sozialen Lebens führen kann (3). Ein primäres Therapieziel ist die Erhaltung der gesundheitsbezogenen Lebensqualität (3, 6).

Fragebogen zur Lebensqualität bei ankylosierender Spondylitis (ASQoL)

Der 2003 für die AS entwickelte krankheitsspezifische ASQoL ist ein vom Patienten auszufüllender Fragebogen, der den Einfluss der Erkrankung auf die Lebensqualität erfasst und neben der AS auch in der Indikation nr-axSpA eingesetzt wird (54, 102). Er umfasst 18 Items, die sich auf die Aspekte Schlaf, Stimmungslage, Motivation, Krankheitsbewältigung,

Aktivitäten des täglichen Lebens, Selbstständigkeit, Beziehungen und Sozialleben beziehen (31).

Operationalisierung

Der ASQoL besteht aus 18 Items, die mit 0 oder 1 bewertet werden können. Der Gesamtscore des ASQoL ist ein Summenscore über die bewerteten Items und hat eine Spannweite von 0 bis 18, wobei niedrigere Werte eine bessere Lebensqualität beschreiben. Für eine Berechnung des Gesamtscores müssen mindestens 15 Items des ASQoL bewertet werden, ansonsten kann dem Endpunkt kein Wert zugeordnet werden (18). Im vorliegenden Dossier wird die mittlere Veränderung des ASQoL dargestellt.

Patientenrelevanz

Da die Fragen des ASQoL vom Patienten selbst beantwortet werden, spiegelt der Gesamtscore des ASQoL direkt die Sicht des Patienten bezüglich der eigenen Lebensqualität wider. Die Auswahl der einzelnen Items basierte auf Patienteninterviews, daher handelt es sich um unmittelbar patientenrelevante Aspekte (102). Es wurde zudem gezeigt, dass die mittels des ASQoL ermittelte Lebensqualität sowohl mit Endpunkten zur Krankheitsaktivität (BASDAI), Funktionsfähigkeit (BASFI, BASMI) und Enthesitis (MASES), als auch mit Symptomen wie Schmerzen und Müdigkeit korreliert (53, 54, 89, 102, 103). Gemäß der Leitlinie des CHMP kann der ASQoL als krankheitsspezifische Skala zur Beurteilung der Lebensqualität von nr-axSpA-Patienten eingesetzt werden (15). Auch der G-BA schätzt den ASQoL als geeignet ein, die Auswirkungen der nr-axSpA auf Patienten zu erfassen (2).

Validität

Der ASQoL wurde entwickelt, um die Lebensqualität von AS-Patienten zu erfassen und wurde in dieser Indikation initial validiert (102). Weitere internationale Validierungsstudien bestätigen die Zuverlässigkeit und Validität der Originalversion und zahlreicher Übersetzungen sowie deren Anwendung in der Indikation nr-axSpA (53, 54, 58, 103-107).

Verträglichkeit

Neben der Wirksamkeit eines Arzneimittels ist die Beurteilung der Verträglichkeit des Arzneimittels ein wesentlicher Bestandteil der klinischen Prüfung. Dazu wurden im Rahmen der Studie sämtliche UE erfasst. Die Verringerung von therapieassoziierten Nebenwirkungen ist grundsätzlich als patientenrelevanter Aspekt eines zu bewertenden Arzneimittels anzusehen. Jedoch sind nicht alle untersuchten UE direkt patientenrelevant, da beispielsweise Änderungen von Laborparametern nicht per se für den Patienten spürbar und nicht zwingend symptomatisch sind.

Zur Beurteilung der Verträglichkeit werden die Gesamtraten für folgende UE dargestellt:

- Alle UE
- Schwerwiegende UE (SUE)

- UE, die zum Therapieabbruch führten
- UE, die zum Tod führten
- UE in $\geq 5\%$ der Patienten in einem Behandlungsarm

Die Erfassung der Schwere der UE erfolgt entsprechend den Common Terminology Criteria for Adverse Events (CTCAE) des National Cancer Institute.

Darüber hinaus werden folgende im Prüfplan präspezifizierte, nach Medical Dictionary for Regulatory Activities (MedDRA) Systemorganklasse (System Organ Class, SOC) und Preferred Terms (PT) zusammengefasste UE von speziellem Interesse (UESI) dargestellt:

- Schwerwiegende Infektion
- Opportunistische Infektion (ohne Herpes Zoster und Tuberkulose)
- Mögliche Malignität (alle Typen)
 - Malignität
 - Nicht melanozytärer Hautkrebs (Non-Melanoma Skin Cancer, NMSC)
 - Malignität außer NMSC
 - Lymphom
- Herpes Zoster
- Aktive/latente Tuberkulose
- Gastrointestinale Perforationen
- Anämie
- Neutropenie
- Lymphopenie
- Nierendysfunktion
- Hepatische Störungen
- Bewertete kardiovaskuläre Ereignisse
- Bewertete VTE (einschließlich tiefe Beinvenenthrombose und Lungenembolie)

Datenschnitte

In der Studie SELECT-AXIS 2 lag zum Zeitpunkt der Einreichung des Dossiers ein Datenschnitt vor:

- Datenschnitt zu Woche 14 zum 26. August 2021: Dieser präspezifizierte Datenschnitt wurde durchgeführt, nachdem der letzte Patient die Studienvsiste zu Woche 14 abgeschlossen hat. Für alle weiteren Patienten wurden alle Informationen bis zum

Datum der Visite des letzten Patienten dokumentiert. Alle im Datenschnitt enthaltenen Daten wurden verblindet gegenüber dem Patienten und dem Behandler erhoben.

In diesem Dossier werden wie auch im klinischen Studienreport (Clinical Study Report, CSR) zur Bewertung der Wirksamkeit und Verträglichkeit die Ergebnisse der placebokontrollierten Periode zu Woche 14 dargestellt.

Zwei weitere Datenschnitte waren bis Studienende vorgesehen. Ein Datenschnitt erfolgt im Anschluss an die doppelt-verblindete, placebokontrollierte Periode zu Woche 52. Die Daten und Ergebnisse dieses Datenschnitts lagen zum Zeitpunkt der Dossiererstellung noch nicht vor. Der finale Datenschnitt wird zu Studienende nach weiteren 52 Wochen in der offenen Verlängerung der Studie zu Woche 104 vorgenommen.

Statistische Methoden

Die Basis für die in diesem Dossier vorgelegten Daten zur Wirksamkeit und Verträglichkeit ist der CSR (108).

Die Analyse der binären Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität basiert auf einem Cochran-Mantel-Haenszel (CMH)-Test stratifiziert nach Status der Entzündung anhand MRT und hsCRP-Level an Screening (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN). Zum Vergleich der Behandlungsgruppen wurde die Risikodifferenz (RD) mit dem 95 %-Konfidenzintervall (KI) und p-Wert als Effektmaß berechnet. Aufgrund gleicher Beobachtungszeiten zwischen den Behandlungsarmen ist keine Verzerrung diesbezüglich zu erwarten, sodass das Effektmaß geeignet ist. Fehlende Werte werden mittels Non-Responder Imputation (NRI) unter Einbindung der multiplen Imputation (MI) ersetzt (NRI-MI). Die NRI wurde erweitert, um fehlende Daten im Zusammenhang mit der Corona Virus Disease (COVID)-19-Pandemie entsprechend zu adressieren. Dazu werden fehlende Daten, die aufgrund einer COVID-19-Infektion des Patienten oder aufgrund von logistischen Einschränkungen durch die Pandemie nicht erhoben werden konnten, mittels MI ersetzt. Fehlen die Daten bei einer geplanten Studienvisite jedoch aus anderen Gründen als einer COVID-19-Infektion oder den genannten logistischen Einschränkungen, so werden diese mittels NRI adressiert. Dabei werden Patienten mit einem fehlenden Wert zu einer Studienvisite für den jeweiligen Endpunkt als Non-Responder behandelt. Darüber hinaus werden Patienten, die eine Rescue-Therapie erhalten oder die Therapie vorzeitig abbrechen als Non-Responder angesehen. Im Fall zusammengesetzter, binärer Endpunkte, werden fehlende kontinuierliche Komponenten des zusammengesetzten Endpunkts mittels MI ersetzt und der zusammengesetzte Endpunkt unter Einbindung des entsprechend ersetzten kontinuierlichen Werts abgeleitet. Bei binären Endpunkten, die direkt aus dem kontinuierlichen Wert gebildet werden, werden fehlende Werte ebenfalls zunächst mittels MI ersetzt. Unter Verwendung der Rubins Regel werden die Ergebnisse der 30 imputierten, mittels NRI-MI vervollständigten Datensätze kombiniert, um den aggregierten Behandlungsunterschied, das zugehörige KI und den p-Wert zu erhalten. Dieses Vorgehen für fehlende Werte in binären Endpunkten ist im statistischen Analyseplan (SAP) präspezifiziert.

Für stetige Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität werden Mittelwertdifferenzen (MWD) auf Basis eines Mixed Model for Repeated Measures (MMRM) mit unstrukturierter Kovarianz-Varianz-Matrix oder einer Kovarianzanalyse (ANCOVA) berechnet. Die Berechnung der MWD mit dem MMRM erfolgt mit der Behandlungsgruppe, Visite, der Interaktion aus Behandlung und Visite als feste Effekte sowie den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening-hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) sowie dem zugehörigen Wert zu Baseline als Kovariablen. Das MMRM erfordert keine zusätzliche Ersetzung fehlender Werte und adressiert unter der Missing-at-Random-Annahme auch fehlende Werte aufgrund der COVID-19-Pandemie adäquat. Die Berechnung der MWD auf Basis einer ANCOVA erfolgt mit der Behandlungsgruppe und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening-hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) als feste Effekte und dem zugehörigen Wert zu Baseline als Kovariablen. Für die ANCOVA werden keine fehlenden Werte ersetzt. Dies entspricht der im SAP definierten Analyse für stetige Endpunkte.

Für Endpunkte zur Verträglichkeit werden die Anteile der Patienten mit mindestens einem UE zum Zeitpunkt zu Woche 14 dargestellt.

Die Auswertung der Wirksamkeitsendpunkte erfolgt anhand der Population aller randomisierten, behandelten Patienten entsprechend der zugewiesenen Studienmedikation. Dies entspricht der präspezifizierten Definition des Full Analysis Set (FAS). Die Auswertung der Endpunkte zur Verträglichkeit erfolgt anhand der Population aller behandelten Patienten entsprechend der erhaltenen Studienmedikation. Dies entspricht der präspezifizierten Definition des Safety Analysis Sets.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten nach der Knapp-

Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung⁵ erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁶ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand geeigneter statistische Maße auf Vorliegen von Heterogenität^{7,5} erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0,05$), soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Es wird keine Meta-Analyse durchgeführt, da im entsprechenden Anwendungsgebiet für die zu betrachtende Patientenpopulation lediglich eine Studie zu dem zu bewertenden Arzneimittel vorliegt. Aus diesem Grund wird auf die Beschreibung der Methodik verzichtet.

⁵ Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015: 25-27.

⁶ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

⁷ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

4.2.5.4 Sensitivitätsanalysen

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

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

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

Sensitivitätsanalysen dienen zur Überprüfung der Robustheit der Ergebnisse der Hauptanalyse sowie der Schlussfolgerung gegenüber der Variation methodischer Faktoren.

Da keine direkt oder indirekt vergleichende Evidenz mit dem zu bewertenden Arzneimittel und einer ZVT vorliegt, sind keine Sensitivitätsanalysen im vorliegenden Dossier dargestellt und es wird daher auf die Beschreibung der Methodik verzichtet.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.9).

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Subgruppenanalysen dienen der Identifizierung von möglichen Effektmodifikatoren und sind hilfreich bei der Identifikation von Patientengruppen mit unterschiedlichen Nutzenprofilen.

Da keine direkt oder indirekt vergleichende Evidenz mit dem zu bewertenden Arzneimittel und einer ZVT vorliegt, sind keine Subgruppenanalysen im vorliegenden Dossier dargestellt und es wird daher auf die Beschreibung der Methodik verzichtet.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethoden darstellen⁸. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche⁹ oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung

⁸ Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDs und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.

⁹ Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6): 683-691.

sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)¹⁰ und Rücker (2012)¹¹ vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹².

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen^{13, 14, 15}

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*

¹⁰ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.

¹¹ Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods 2012; 3(4): 312-324.

¹² Schöttker B, Lüthmann D, Boukhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.

¹³ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009; 338: b1147.

¹⁴ Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study BMJ 2011; 343 :d4909

¹⁵ Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods 2013; 4(4): 291-323.

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Eine Suche nach RCTs für einen indirekten Vergleich nach nicht randomisierten vergleichenden Studien oder nach weiteren Untersuchungen wird nicht durchgeführt.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen/abgebrochen/laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
SELECT-AXIS 2 (M19-944)	ja	ja	laufend	<u>Studie 1^a:</u> erste Periode 14 Wochen, zweite Periode bis zu 90 Wochen 1. Datenschnitt nach 14 Wochen 2. Datenschnitt nach 52 Wochen 3. Datenschnitt nach 104 Wochen (Studienende) <u>Studie 2^b:</u> erste Periode 52 Wochen, zweite Periode bis zu 52 Wochen 1. Datenschnitt nach 14 Wochen 2. Datenschnitt nach 52 Wochen 3. Datenschnitt nach 104 Wochen (Studienende)	<u>Studie 1:</u> Upadacitinib 15 mg QD, Placebo ^c <u>Studie 2:</u> Upadacitinib 15 mg QD, Placebo ^d
a: Patienten mit AS b: Patienten mit nr-axSpA c: Patienten im Placebo-Arm wechselten nach 14 Wochen in den Upadacitinib-Arm. d: Patienten im Placebo-Arm wechselten nach 52 Wochen in den Upadacitinib-Arm. AS: Ankylosierende Spondylitis; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie					

Geben Sie an, welchen Stand die Information in Tabelle 4-7 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Aufgelistet sind alle RCTs mit finanzieller Beteiligung des pharmazeutischen Unternehmers für das zu bewertende Arzneimittel, die in der Indikation laufend sind oder durchgeführt wurden (Stand: 04. Juli 2022).

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-7 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
SELECT-AXIS 2 (M19-944)	Kein Vergleich zur ZVT (placebokontrolliert)
RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie	

Die SELECT-AXIS 2 ist keine direkt vergleichende Studie gegen die ZVT. Entsprechend wird diese Studie nicht zur Ableitung eines Zusatznutzens verwendet. Für die SELECT-AXIS 2 liegen Ergebnisse zur Wirksamkeit und Verträglichkeit von einer Behandlung mit Upadacitinib bei nr-axSpA-Patienten vor, die in Abschnitt 4.3.1.3 ergänzend dargestellt werden.

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

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

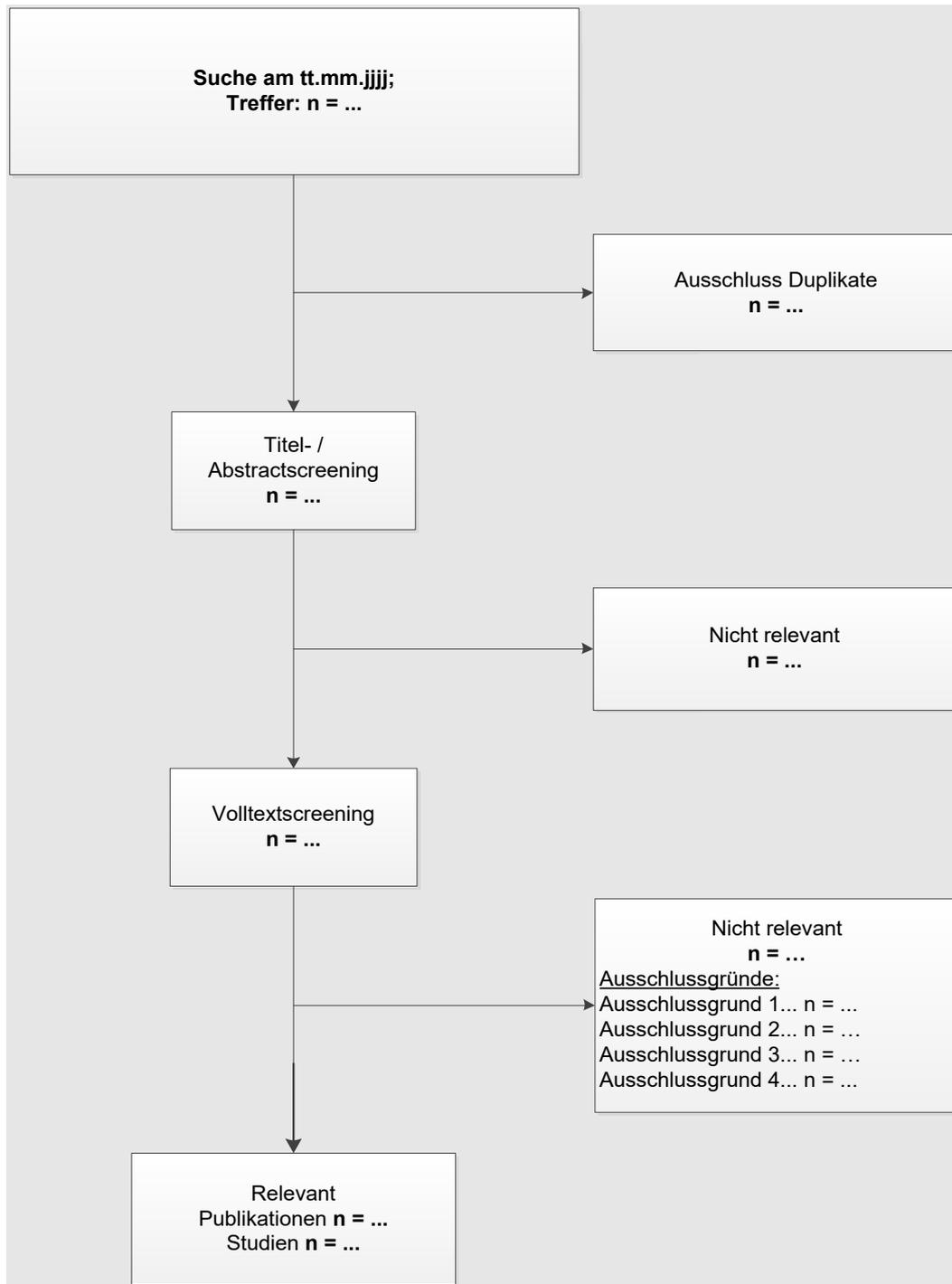


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche zu RCTs mit Upadacitinib fand am 05. Juli 2022 statt und erzielte insgesamt 12 Treffer. Nach Duplikatausschluss (n = 5) verblieben 7 Publikationen, welche unter Berücksichtigung der Ein- und Ausschlusskriterien (Abschnitt 4.2.2) gesichtet

wurden. Im Zuge des Titel- und Abstractscreenings wurden alle 7 Publikationen als nicht relevant eingestuft. Eine Sichtung der Volltexte war daher nicht notwendig (Abbildung 2).

Es wurden im Rahmen der bibliografischen Literaturrecherche keine relevanten RCTs mit dem zu bewertenden Arzneimittel identifiziert, die gemäß Fragestellung (Abschnitt 4.2.1) für die Nutzenbewertung im vorliegenden Anwendungsgebiet nr-axSpA herangezogen werden können.

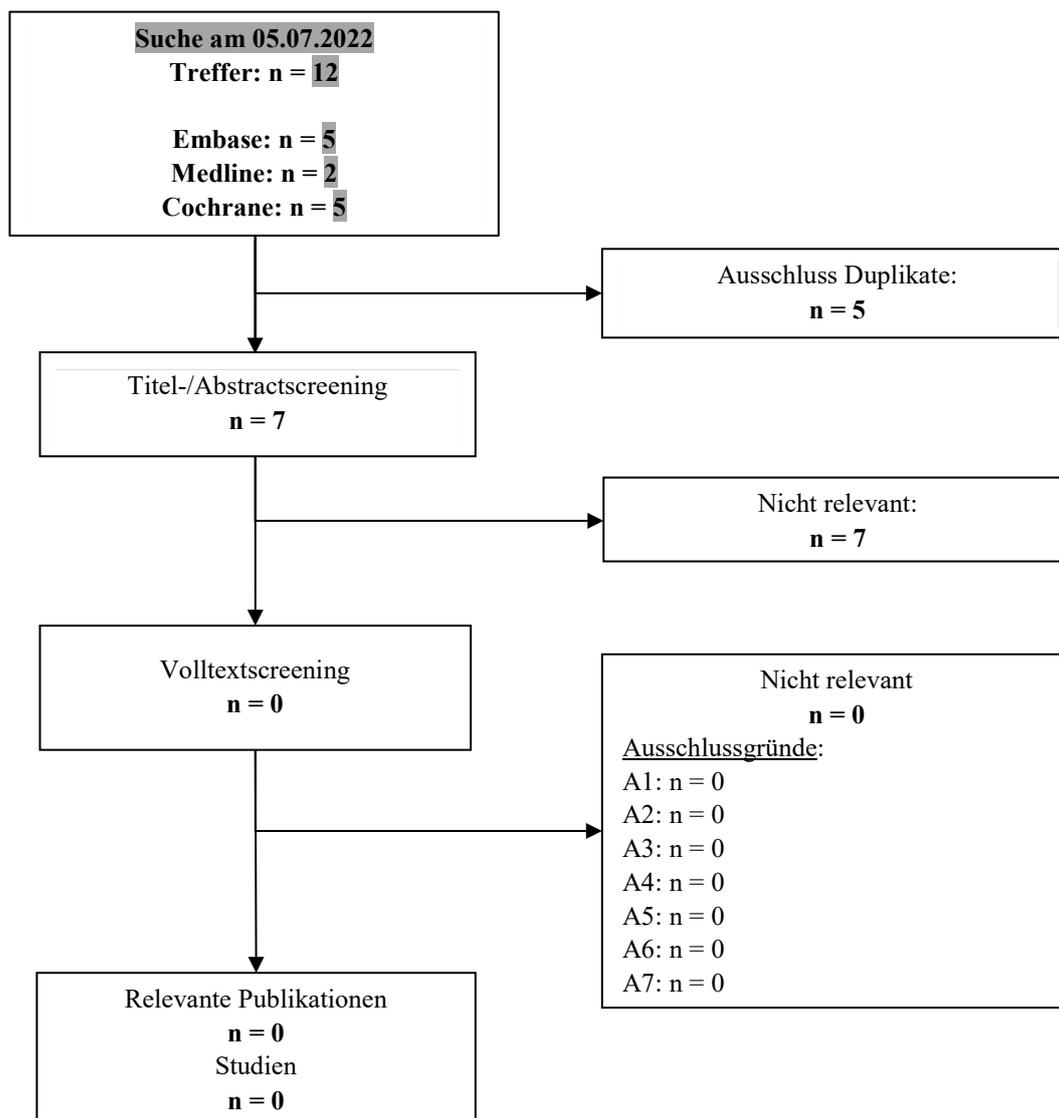


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

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

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnis- datenbank und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
Keine Studie identifiziert.				
<p>a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>EudraCT: European Union Drug Regulating Authorities Clinical Trials Database; NCT: National Clinical Trial (US); RCT: Randomisierte kontrollierte Studie</p>				

Geben Sie an, welchen Stand die Information in Tabelle 4-9 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Studienregistersuche ergab keine relevante RCT mit dem zu bewertenden Arzneimittel, die einen direkten Vergleich mit der ZVT erlaubt.

Stand der Information ist der 05. Juli 2022.

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

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

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	Relevante Quellen ^a	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern/Studienergebnisdatenbanken identifiziert (ja/nein)
Keine Studie identifiziert.				
a: Quellen aus der Suche auf der Internetseite des G-BA G-BA: Gemeinsamer Bundesausschuss; RCT: randomisierte kontrollierte Studie				

Geben Sie an, welchen Stand die Information in Tabelle 4-10 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Es wurden keine relevanten Studien mit Upadacitinib im Anwendungsgebiet identifiziert (Stand: 05. Juli 2022).

4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2, 4.3.1.1.3 und 4.3.1.1.4) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich aller verfügbaren Quellen in Abschnitt 4.3.1.4 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Register-einträge ^c (ja/nein [Zitat])	Publikation und sonstige Quellen ^d (ja/nein [Zitat])
placebokontrolliert						
SELECT-AXIS 2 (M19-944)	ja	ja	nein	ja (108)	ja (109-111)	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Keine Studie indentifiziert.						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.</p> <p>G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie</p>						

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert. Die Ergebnisse der zulassungsrelevanten, placebokontrollierten Phase-III-Studie SELECT-AXIS 2 werden zur Beschreibung der Wirksamkeit und der Verträglichkeit einer Behandlung mit Upadacitinib in Abschnitt 4.3.1.3 dargestellt.

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch

für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/ einfach, verblindet/offen, parallel/cross- over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
SELECT- AXIS 2 (M19-944)	Randomisierte, doppelblinde, parallele, multizentrische, placebo- kontrollierte Phase-III-Studie, bestehend aus zwei unabhängigen Studien	Erwachsene Patienten (≥ 18 Jahre) mit der klinischen Diagnose einer AS entsprechend der modifizierten New York Kriterien für AS, die ein unzureichendes Ansprechen oder eine Unverträglichkeit gegen- über einer vorherigen Therapie mit bDMARDs hatten (Studie 1) ^a oder mit einer klinischen Diagnose einer nr-axSpA entsprechend der 2009 ASAS- Klassifikationskriterien, jedoch ohne das radiologische Kriterium der modifizierten New York-Kriterien für AS (Studie 2)	Upadacitinib 15 mg QD (n = 156) Placebo (n = 158 ^b), Wechsel auf Upadacitinib 15 mg nach Woche 52	Screening: 35 Tage Behandlung: Doppelblinde Behandlungsphase von 52 Wochen, offene Behandlungsphase mit Upadacitinib über weitere 52 Wochen Follow-up: 30 Tage Studiendauer: bis zu 104 Wochen Behandlung (ohne Berücksichtigung von Screening und Follow-up) 1. Datenschnitt: 26. August 2021, 14 Wochen (primärer Datenschnitt) 2. Datenschnitt: 52 Wochen (weiterer Datenschnitt für die Zulassungsbehörden) Finaler Datenschnitt: 104 Wochen Nachbeobachtung bis Woche 152	International in 113 Studien- zentren in 23 Ländern: Argentinien, Australien, Belgien, Brasilien, Bulgarien, China, Deutschland, Frankreich, Israel, Japan, Kanada, Korea, Mexiko, Polen, Russland, Slowakei, Spanien, Taiwan, Tschechische Republik, Türkei, Ukraine, Ungarn, Vereinigte Staaten Zeitraum: 11/2019 – laufend	<u>Primärer Endpunkt:</u> Anteil der Patienten mit ASAS40-Ansprechen zu Woche 14 <u>Patientenrelevante sekundäre Endpunkte:</u> - ASDAS - BASDAI - ASAS - patientenberichteter Rückenschmerz - BASFI - SPARCC MRI - MASES - ASAS HI - PtGADA - ASQoL - Verträglichkeit

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/ einfach, verblindet/offen, parallel/cross- over etc.>	<relevante Charakteristika, z. B. Schweregrad>		<ggf. Run-in, Behandlung, Nachbeobachtung>		
<p>a: Da die beiden Studien innerhalb der Studie SELECT-AXIS 2 unabhängig voneinander sind, die Studie 1 ausschließlich Patienten mit AS und damit nicht das Anwendungsgebiet des vorliegenden Dossiers umfasst, wird Studie 1 im Folgenden nicht weiter dargestellt.</p> <p>b: Ein Patient, der in den Placebo-Arm randomisiert wurde, entschied sich nach der Randomisierung gegen die Studienteilnahme ohne Studienmedikation erhalten zu haben.</p> <p>AS: Ankylosierende Spondylitis; ASDAS; Ankylosing Spondylitis Disease Activity Score; ASAS: Assessment of SpondyloArthritis International Society; ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; bDMARD: Biologisches DMARD; DMARD: Krankheitsmodifizierendes Antirheumatikum; HI: Health Index; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; MRI: Magnetic Resonance Imaging; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; PtGADA: Patientenberichtete globale Krankheitsaktivität; QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie; SPARCC: Spondyloarthritis Research Consortium of Canada</p>						

Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Upadacitinib	Placebo	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
SELECT-AXIS 2	Upadacitinib 15 mg p.o. QD	Placebo zu Upadacitinib 15 mg p.o. QD	<p>Vorbehandlung:</p> <ul style="list-style-type: none"> • Unzureichendes Ansprechen auf mind. 2 Behandlungen mit NSAR über einen Zeitraum von mind. 4 Wochen bei maximal zugelassener oder tolerierter Dosis oder Unverträglichkeit oder Kontraindikation gegenüber NSAR durch den Prüfarzt bestätigt. • Einhalten der jeweils festgelegten Wash-out Periode bei Vorbehandlung mit bDMARDs • Keine Vorbehandlung mit jeglichen JAK-Inhibitoren <p>Patienten, die zu Einschluss der Studie begleitende Therapien erhalten, können diese unter Einhaltung folgender Kriterien fortführen:</p> <ul style="list-style-type: none"> • Erhaltung der laufenden Hintergrundtherapie mit nicht biologischen DMARDs (MTX p.o. oder s.c. \leq 25 mg/Tag, Sulfasalazin \leq 3 g/Tag, Hydroxychloroquin \leq 400 mg/Tag, Chloroquin \leq 400 mg/Tag, Leflunomid \leq 20 mg/Tag oder Apremilast \leq 50mg/Tag) mit stabiler Dosis für mindestens 4 Wochen vor der ersten Gabe der Studienmedikation • Kombination mit bis zu 2 nicht biologischen DMARDs erlaubt (mit Ausnahme der Kombination von MTX und Leflunomid) • Erhaltung der begleitenden Therapie mit stabiler Dosis von NSAR, Tramadol, Kombination von Paracetamol und Codein bzw. Paracetamol und Hydrocodon, nicht opioiden Analgetika, oraler oder inhalativer Kortikosteroide

Studie	Upadacitinib	Placebo	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
In der offenen Verlängerungsperiode ab Woche 52 erhalten alle Patienten Upadacitinib 15 mg p.o. einmal täglich			
bDMARD: Biologisches DMARD; DMARD: Krankheitsmodifizierendes Antirheumatikum; JAK: Januskinase; MTX: Methotrexat; NSAR: Nicht steroidales Antirheumatikum; p.o.: Peroral; QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie; s.c.: Subkutan			

Tabelle 4-14: Charakterisierung der Studienpopulation: Demografische Charakteristika zu Baseline nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Merkmal	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
Geschlecht, n (%)		
weiblich	89 (57,1)	94 (59,9)
männlich	67 (42,9)	63 (40,1)
Alter, Jahre		
MW (SD)	41,6 (12,00)	42,5 (12,44)
BMI (kg/m²)		
MW (SD)	28,2 (6,39)	27,7 (5,23)
Abstammung, n (%)		
weiß	134 (85,9)	127 (80,9)
asiatisch	19 (12,2)	28 (17,8)
andere ^a	3 (1,9)	2 (1,3)
geografische Region, n (%)		
Nordamerika	26 (16,7)	19 (12,1)
Süd- und Zentralamerika	12 (7,7)	13 (8,3)
Westeuropa	24 (15,4)	19 (12,1)
Osteuropa	68 (43,6)	72 (45,9)
Asien ^b	19 (12,2)	27 (17,2)
Andere ^c	7 (4,5)	7 (4,5)
a: Schwarz oder afroamerikanisch, Ureinwohner Amerikas oder Alaskas, multiple b: China, Taiwan, Südkorea und Japan c: Australien und Israel Population: Full Analysis Set BMI: Body Mass Index; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung		

Tabelle 4-15: Charakterisierung der Studienpopulationen: Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Merkmal	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
Dauer der nr-axSpA-Diagnose, Jahre		
MW (SD)	4,5 (5,54)	4,4 (5,83)
Dauer der nr-axSpA-Symptome, Jahre		
n (fehlend)	155 (1)	156 (1)
MW (SD)	9,0 (7,86)	9,2 (8,12)
HLA-B27, n (%)		
positiv	90 (58,8)	93 (59,6)
negativ	63 (41,2)	63 (40,4)
fehlend ^a	3	1
ASDAS (CRP)		
MW (SD)	3,61 (0,674)	3,65 (0,644)
BASDAI (Score 0 - 10)		
MW (SD)	6,82 (1,295)	6,91 (1,215)
ASAS HI		
n (fehlend)	153 (3)	155 (2)
MW (SD)	9,44 (3,575)	9,52 (3,661)
patientenberichteter Rückenschmerz^b (NRS-Score 0 - 10)		
MW (SD)	7,2 (1,55)	7,3 (1,39)
patientenberichteter Rückenschmerz (BASDAI-Item 2) (NRS-Score 0 - 10)		
MW (SD)	7,4 (1,49)	7,4 (1,32)
BASFI (Score 0 – 10)		
MW (SD)	5,89 (2,077)	5,99 (2,139)
SPARCC MRI Spine		
n (fehlend)	139 (17)	147 (10)
MW (SD)	2,72 (6,881)	1,44 (3,724)
SPARCC MRI (sakroiliakale Gelenke)		
n (fehlend)	142 (14)	148 (9)
MW (SD)	4,38 (8,737)	3,49 (7,601)

Merkmal	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
Enthesitis zu Baseline^c, n (%)		
ja	125 (80,1)	125 (79,6)
nein	31 (19,9)	32 (20,4)
MASES (bei Patienten mit Enthesitis zu Baseline^c)		
MW (SD)	4,7 (3,09)	4,7 (3,18)
Status der Entzündung anhand MRT zu Screening, n (%)		
positiv	70 (44,9)	66 (42,0)
negativ	86 (55,1)	91 (58,0)
hsCRP zu Screening, mg/L		
MW (SD)	13,61 (24,794)	10,52 (13,522)
hsCRP zu Screening, n (%)		
≤ ULN	33 (21,2)	31 (19,7)
> ULN	123 (78,8)	126 (80,3)
ASQoL		
n (fehlend)	154 (2)	155 (2)
MW (SD)	11,9 (4,41)	11,9 (4,53)
PtGADA (NRS-Score 0 – 10)		
MW (SD)	6,99 (1,618)	7,30 (1,384)
<p>a: Die Tests wurden abgeschlossen, aber die Genotypen von zwei Patienten in der Upadacitinib-Gruppe konnten nicht zugeordnet werden.</p> <p>b: Patientenberichteter Rückenschmerz berichtet auf einer NRS (0 – 10) auf die Frage „Wie ausgeprägt waren Ihre Rückenschmerzen in den vergangenen 7 Tagen insgesamt?“</p> <p>c: Basierend auf einem MASES > 0 zu Baseline</p> <p>Population: Full Analysis Set</p> <p>ASAS: Assessment of SpondyloArthritis international Society; ASDAS: Ankylosing Spondylitis Disease Activity Score; ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; CRP: C-reaktives Protein; HI: Health Index; HLA: Humanes Leukozytenantigen; hsCRP: Hoch sensitives C-reaktives Protein; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; MRI: Magnetic Resonance Imaging; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; NRS: Numerical Rating Scale; PtGADA: Patientenberichtete globale Krankheitsaktivität; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; SPARCC: Spondyloarthritis Research Consortium of Canada; ULN: Obere Normgrenze</p>		

Tabelle 4-16: Charakterisierung der Studienpopulation: Medizinische Vorgeschichte und Begleitbehandlung nach Behandlungsarm in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Merkmal	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
frühere Behandlung mit bDMARDs, n (%)		
ja	49 (31,4)	54 (34,4)
nein	107 (68,6)	103 (65,6)
begleitende Behandlung mit NSAR, n (%)		
ja	122 (78,2)	115 (73,2)
nein	34 (21,8)	42 (26,8)
begleitende csDMARD-Therapie, n (%)		
ja	41 (26,3)	50 (31,8)
nein	115 (73,7)	107 (68,2)
begleitende Behandlung mit oralen Kortikosteroiden zu Woche 14, n (%)		
ja	18 (11,5)	18 (11,5)
nein	138 (88,5)	139 (88,5)
Behandlung mit oralen Kortikosteroiden zu Baseline, n (%)		
ja	18 (11,5)	17 (10,8)
nein	138 (88,5)	140 (89,2)
Population: Full Analysis Set bDMARD: Biologisches DMARD; csDMARD: Konventionelles synthetisches DMARD; DMARD: Krankheitsmodifizierendes Antirheumatikum; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; NSAR: Nicht steroidales Antirheumatikum; RCT: Randomisierte kontrollierte Studie		

Tabelle 4-17: Patientenstatus in der RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
Patienten mit Abbruch der Studienmedikation bis Woche 14, n (%)	11 (7,1)	7 (4,5)
Primärer Abbruchgrund, n (%)		
UE	4 (2,6)	2 (1,3)
Abbruch auf Wunsch des Patienten	2 (1,3)	1 (0,6)
Lost to follow-up	0	0
fehlende Wirksamkeit	3 (1,9)	3 (1,9)

	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
logistische Einschränkungen aufgrund von COVID-19	1 (0,6)	0
andere	1 (0,6)	1 (0,6)
Population: Full Analysis Set COVID: Corona Virus Disease; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; UE: Unerwünschtes Ereignis		

Tabelle 4-18: Behandlungs- und Beobachtungsdauer in RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Merkmal	SELECT-AXIS 2	
	Upadacitinib N = 156	Placebo N = 157
Behandlungsdauer (Tagen)		
MW (SD)	96,4 (14,11)	96,8 (12,87)
Median	98,0	98,0
Min – Max	5 – 124	1 – 119
Population: Full Analysis Set MW: Mittelwert; N: Anzahl Patienten; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung		

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

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

Studiendesign

Die Studie SELECT-AXIS 2 (M19-944) ist eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie, die die Wirksamkeit und Sicherheit von Upadacitinib im Vergleich zu Placebo bei Patienten mit aktiver AS oder aktiver nr-axSpA untersucht. Die Studie

SELECT-AXIS 2 wurde mit einem übergreifenden Studienprotokoll geplant und umfasst zwei getrennte Studien (Studie 1 und Studie 2). Somit können in einem gemeinsamen Screeningprozess Patienten mit axSpA in die jeweils geeignete Studie eingeschlossen werden. Randomisierung, Datenerhebung sowie weitere Studienaktivitäten laufen anschließend an das Screening unabhängig voneinander. Studie 1 umfasst Patienten mit einer aktiven AS mit unzureichendem Ansprechen auf eine vorherige Behandlung mit bDMARDs und umfasst damit nicht das Anwendungsgebiet im vorliegenden Dossier. Somit wird Studie 1 im Folgenden nicht weiter betrachtet. In die Studie 2 der Studie SELECT-AXIS 2 wurden Patienten mit aktiver nr-axSpA eingeschlossen. Im Folgenden wird mit der Nennung der Studie SELECT-AXIS 2 ausschließlich die relevante Studie 2 beschrieben. In diese wurden erwachsene Patienten (≥ 18 Jahre) mit einer aktiven nr-axSpA eingeschlossen, welche die folgenden Haupteinschlusskriterien erfüllen:

- Klinische Diagnose einer nr-axSpA, die die 2009 ASAS-Klassifikationskriterien für axSpA erfüllen, jedoch nicht die modifizierten New-York-Kriterien für AS erfüllen
- Objektive Anzeichen einer aktiven entzündlichen Erkrankung im Zusammenhang mit einer axSpA mit einem Status der Entzündung der sakroiliakalen Gelenke anhand MRT und/oder einem Screening hsCRP $>$ ULN
- Aktive Erkrankung zu Baseline, definiert als BASDAI ≥ 4 und Patienteneinschätzung der Rückenschmerzen ≥ 4 zu Screening und zu Baseline
- Unzureichendes Ansprechen auf mindestens zwei Behandlungen mit NSAR über insgesamt mindestens 4 Wochen in der maximal möglichen empfohlenen oder tolerierten Dosierung oder eine Therapie mit NSAR ist nicht angezeigt bzw. nicht geeignet

Bei Studieneinschluss konnten Patienten begleitend mit bis zu zwei nicht bDMARDs (z. B. Methotrexat [MTX], Leflunomid, Sulfasalazin oder Hydroxychloroquin in einer stabilen Dosis ≥ 28 Tage vor Baseline) weiterbehandelt werden, mit Ausnahme der Kombination von MTX und Leflunomid. Ebenso waren orale Kortikosteroide, nicht steroidale Entzündungshemmer oder Schmerzmittel (stabile Dosis ≥ 14 Tage vor Baseline) als Begleitmedikation entsprechend dem Studienprotokoll zulässig. Eine vorangegangene Therapie mit einem bDMARD war für den Einschluss in die Studie zulässig, sofern die Behandlung vor Einschluss in die Studie beendet war und die festgelegten Wash-out-Perioden eingehalten wurden. Eine vorangegangene Therapie mit einem JAK-Inhibitor war für den Einschluss in die Studie nicht zulässig.

In die Studie SELECT-AXIS 2 wurden Patienten in 113 Zentren in 23 Ländern in Nordamerika, Ost- und Westeuropa, Asien und Ozeanien rekrutiert und im Verhältnis von 1:1 in zwei Behandlungsarme (Upadacitinib 15 mg: 156, Placebo: 158) randomisiert. Die Randomisierung erfolgte stratifiziert nach hsCRP-Level zu Screening, Status der Entzündung anhand des MRT sowie der vorherigen Therapie mit bDMARDs. Der Anteil an Patienten mit einer vorherigen Therapie mit bDMARDs wurde auf mindestens 20 % und höchstens 35 % festgelegt. Ein Patient, der in den Placebo-Arm randomisiert wurde, entschied sich nach der Randomisierung

und vor jeglicher Gabe von Studienmedikation gegen eine Studienteilnahme. Dieser Patient wird im Folgenden nicht weiter berücksichtigt, sodass das FAS einen Patienten weniger umfasst als ursprünglich randomisiert wurden.

In der ersten Periode nahmen die Patienten entweder Upadacitinib 15 mg oder Placebo einmal täglich für 52 Wochen ein. Im Anschluss an die 52-wöchige, verblindete, placebokontrollierte Studienbehandlung werden alle Patienten in der offenen Verlängerung für weitere 52 Wochen mit Upadacitinib behandelt. Patienten, die in der ersten Periode Placebo erhalten haben, werden in der offenen Verlängerung auf Upadacitinib umgestellt. Nach Woche 104 werden Patienten mit klinischer Remission bis Woche 152 ohne Gabe der Studienmedikation weiter nachbeobachtet und können im Fall des Verlusts der Remission erneut mit Upadacitinib 15 mg behandelt werden. Patienten ohne Remission zu Woche 104 können die Studie mit der Follow-up-Visite beenden oder, sofern im Rahmen der nationalen und lokalen Vorgaben möglich, für einen festgelegten Zeitraum weiter im offenen Design mit Upadacitinib 15 mg behandelt werden. Im vorliegenden Dossier werden die zulassungsrelevanten Ergebnisse des primären Endpunkts sowie weitere patientenrelevante Endpunkte zu Woche 14 dargestellt.

Eine Rescue-Therapie ist für Patienten vorgesehen, die ab der Studienvsiste zu Woche 24 bis Ende der Periode 1 zu Woche 52 an zwei aufeinanderfolgenden Visiten kein ASAS20-Ansprechen zeigen. Die Rescue-Therapie ergibt sich durch Hinzunahme oder Anpassung der Begleitmedikation, die entsprechend im Rahmen des Studienprotokolls möglich ist. Bei Abbruch der Studienmedikation kann auch die Behandlung mit einem bDMARD (TNF- oder IL-17-Inhibitor) begonnen werden. Patienten, die in Periode 1 gemäß der Randomisierung mit Placebo behandelt werden, werden zum Ende der Periode 1 zu Woche 52 auf eine Behandlung mit Upadacitinib umgestellt. Gemäß Studienprotokoll ist vorgesehen, dass Patienten, die ab Woche 64 bis Woche 104 kein ASAS20-Ansprechen an zwei aufeinanderfolgenden Visiten zeigen, die Studienmedikation abbrechen.

Charakterisierung der Studienpopulation

Das durchschnittliche Alter im Upadacitinib-Arm lag bei ca. 42 Jahren, im Placebo-Arm lag es bei ca. 43 Jahren. Im Upadacitinib-Arm waren ca. 57 % der Patienten weiblich, im Placebo-Arm waren es ca. 60 % der Patienten. Der mittlere BMI lag bei Patienten im Upadacitinib-Arm bei ca. 28,2 kg/m², im Placebo-Arm lag dieser bei ca. 27,7 kg/m². Der überwiegende Anteil der Patienten in dieser Studie war weiß (> 80 %). Es wurden Patienten aus Osteuropa (jeweils ca. 44 % bzw. 46 %), Nordamerika (jeweils ca. 17 % bzw. 12 %), Westeuropa (jeweils ca. 15 % bzw. 12 %), Asien (jeweils ca. 12 % bzw. 17 %), Süd- und Zentralamerika (jeweils ca. 8 %) sowie Australien und Israel (jeweils ca. 5 %) eingeschlossen.

Hinsichtlich der krankheitsspezifischen Merkmale waren beide Behandlungsarme konsistent. Die durchschnittliche Erkrankungsdauer lag mit ca. 4,5 Jahren im Upadacitinib-Arm und ca. 4,4 Jahren im Placebo-Arm in einem vergleichbaren Bereich; Symptome der Erkrankung hatten sich jeweils seit ca. 9,0 Jahren im Upadacitinib-Arm bzw. seit ca. 9,2 Jahren im Placebo-Arm manifestiert.

Bezüglich der medizinischen Vorgeschichte und der Begleitbehandlung zu Baseline sind die Behandlungsarme vergleichbar. Ungefähr 31 % bzw. 34 % der Patienten hatten eine vorherige bDMARD-Therapie. Eine begleitende NSAR-Therapie erhielten ca. 78 % der Patienten im Upadacitinib-Arm, im Placebo-Arm waren es ca. 73 % der Patienten. Eine begleitende Behandlung mit einem csDMARD erhielten im Upadacitinib-Arm ca. 26% der Patienten, im Placebo-Arm waren es ca. 32 % der Patienten.

Insgesamt ist von einer guten Vergleichbarkeit im Hinblick auf die Charakterisierung der Studienarme auszugehen.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Studie SELECT-AXIS 2 wurde weltweit durchgeführt. Die Studienpopulation bestand zu knapp 83 % aus weißen Studienteilnehmern. Aufgrund der Strukturgleichheit, v. a. bezüglich der klinischen Parameter zwischen der Studienpopulation und der Zielpopulation im Anwendungsgebiet wird davon ausgegangen, dass die in der Studie SELECT-AXIS 2 beobachteten klinischen Effekte auch in der medizinischen Versorgung unter Alltagsbedingungen auftreten und die Studienergebnisse somit auf den deutschen Versorgungskontext übertragbar sind. Es konnten keine Hinweise identifiziert werden, die einer Übertragbarkeit widersprechen.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-19: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
Nicht zutreffend.							
RCT: Randomisierte kontrollierte Studie							

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Morbidität ^a	Gesundheits- bezogene Lebensqualität	Verträglichkeit
SELECT-AXIS 2	ja	ja	ja	ja
a: Umfasst Krankheitsaktivität, Schmerz, Funktion, Entzündung und Fatigue RCT: Randomisierte kontrollierte Studie				

4.3.1.3.1 Mortalität – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle

Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]). Die Auswertung von Responderanalysen mittels klinischer Relevanzschwellen bei komplexen Skalen soll nach dem folgenden Vorgehen erfolgen:

1. Falls in einer Studie Responderanalysen unter Verwendung einer MID präspezifiziert sind und das Responsekriterium mindestens 15 % der Skalenspannweite des verwendeten Erhebungsinstruments entspricht, sind diese Responderanalysen für die Bewertung darzustellen.

2. Falls präspezifiziert Responsekriterien im Sinne einer MID unterhalb von 15 % der Skalenspannweite liegen, bestehen in diesen Fällen und solchen, in denen gar keine Responsekriterien präspezifiziert wurden, aber stattdessen Analysen kontinuierlicher Daten zur Verfügung stehen, verschiedene Möglichkeiten. Entweder können post hoc spezifizierte Analysen mit einem Responsekriterium von genau 15 % der Skalenspannweite dargestellt werden. Alternativ können Analysen der kontinuierlichen Daten dargestellt werden, für die Relevanzbewertung ist dabei auf ein allgemeines statistisches Maß in Form von standardisierten Mittelwertdifferenzen (SMDs, in Form von Hedges' g) zurückzugreifen. Dabei ist eine Irrelevanzschwelle als Intervall von - 0,2 bis 0,2 zu verwenden: Liegt das zum Effektschätzer korrespondierende Konfidenzintervall vollständig außerhalb dieses Irrelevanzbereichs, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Dies soll gewährleisten, dass der Effekt hinreichend sicher mindestens als klein angesehen werden kann.

3. Liegen sowohl geeignete Responderanalysen (Responsekriterium präspezifiziert mindestens 15 % der Skalenspannweite oder post hoc genau 15 % der Skalenspannweite) als auch Analysen stetiger Daten vor, sind die Responderanalysen darzustellen.

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen: 1. Gesamtrate UE,

2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die

Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

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

Tabelle 4-21: Operationalisierung von Mortalität

Studie	Operationalisierung
SELECT- AXIS 2	Die Erhebung der Mortalität erfolgte im Rahmen der Erhebung der UE und wird daher in Abschnitt 4.3.1.3.8 dargestellt.
UE: Unerwünschtes Ereignis	

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

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

Tabelle 4-22: Bewertung des Verzerrungspotenzials für Mortalität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Nicht zutreffend, Ergebnisse zur Mortalität werden im Rahmen der UE in Abschnitt 4.3.1.3.8 dargestellt.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.2 Krankheitsaktivität – RCT

4.3.1.3.2.1 ASAS – RCT

Tabelle 4-23: Operationalisierung von ASAS

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der ASAS setzt sich zusammen aus den vier Domänen PtGADA, patientenberichteter Rückenschmerz, BASFI sowie der Entzündung (abgeleitet aus dem Mittelwert der BASDAI-Items 5 und 6 zur Morgensteifigkeit) und wird durch den Patienten selbst berichtet.</p> <p>Der ASAS wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 2, Woche 4, Woche 8, Woche 12 und zu Woche 14.</p> <p>Das ASAS40-Ansprechen ist definiert als Verbesserung um $\geq 40\%$ und eine absolute Verbesserung um mindestens 2 Einheiten ab Baseline auf einer Skala von 0 bis 10 in mindestens drei der vier Domänen, wobei es in keiner der Domänen zu einer Verschlechterung kommen darf. Die ASAS-PR ist definiert als Erreichen eines absoluten Scores von ≤ 2 für jede der vier Domänen auf einer Skala von 0 bis 10.</p> <p>Für das vorliegende Dossier werden der Anteil der Patienten mit ASAS40-Ansprechen und der Anteil der Patienten mit ASAS-PR zu Woche 14 untersucht.</p> <p>Der Anteil der Patienten mit einem ASAS40-Ansprechen sowie einer ASAS-PR basiert auf dem FAS. Die statistische Analyse erfolgte mithilfe eines CMH-Tests adjustiert für die Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP.</p> <p>Fehlende Werte sowie Werte nach Behandlungsabbruch wurden mittels NRI-MI imputiert (siehe Abschnitt 4.2.5.2).</p>
<p>ASAS: Assessment of SpondyloArthritis International Society; ASAS-PR: ASAS-partielle Remission; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; CMH: Cochran-Mantel-Haenszel; FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; KI: Konfidenzintervall; MI: Multiple Imputation; MRT: Magnetresonanztomografie; NRI: Non-Responder-Imputation; PtGADA: Patientenberichtete globale Krankheitsaktivität</p>	

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

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

Tabelle 4-24: Bewertung des Verzerrungspotenzials für ASAS in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ASAS: Assessment of SpondyloArthritis International Society; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-25: Ergebnisse für ASAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95 %-KI] p-Wert	OR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Woche 14					
Anteil der Patienten mit ASAS40-Ansprechen	70/156 (44,9)	35/157 (22,5)	k. A.	k. A.	0,22 [0,12; 0,32] p < 0,0001
Anteil der Patienten mit ASAS-PR	29/156 (18,6)	12/157 (7,6)	k. A.	k. A.	0,11 [0,36; 0,18] p = 0,0035

Endpunkt	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95 %-KI] p-Wert	OR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
<p>Imputationsmethode: NRI-MI Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN)</p> <p>Der Behandlungsunterschied, das zugehörige KI und der p-Wert für den Test des Unterschieds zwischen dem Upadacitinib-Arm und dem Placebo-Arm werden auf der Grundlage der MI-Inferenz geschätzt. Die RD und der SE werden für jeden mittels NRI-MI vervollständigten Datensatz mit dem CMH-Test adjustiert für die Stratifizierungsfaktoren geschätzt.</p> <p>Die Anzahl an Patienten mit einem Ansprechen (n) wird berechnet durch die Anzahl an Patienten in der Analyse (N) sowie der aggregierten MI-Responderate (%).</p> <p>Eine RD größer Null stellt einen Vorteil von Upadacitinib dar.</p> <p>ASAS: Assessment of SpondyloArthritis International Society; ASAS-PR: ASAS-partielle Remission; CMH: Cochran-Mantel-Haenszel; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; MI: Multiple Imputation; MRT: Magnetresonanztomografie; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Response; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie; RD: Risikodifferenz; RR: Relatives Risiko; SE: Standardfehler; ULN: Obere Normgrenze</p>					

Unter der Behandlung mit Upadacitinib erreichten etwa doppelt so viele Patienten ein ASAS40-Ansprechen als im Vergleichsarm mit Placebo. Zum herangezogenen Zeitpunkt 14 Wochen nach Therapiebeginn hatten unter der Behandlung mit Upadacitinib statistisch signifikant mehr Patienten (ca. 45 %) ein ASAS40-Ansprechen als im Placebo-Arm (ca. 23 %; RD [95 %-KI]; p-Wert: 0,22 [0,12; 0,32]; $p < 0,0001$). Eine partielle Remission gemäß des ASAS erreichten im Upadacitinib-Arm deutlich mehr Patienten als im Vergleichsarm mit Placebo. Unter der Behandlung mit Upadacitinib hatten ca. 19 % der Patienten und im Placebo-Arm ca. 8 % der Patienten eine ASAS-PR. Der Unterschied ist statistisch signifikant zugunsten von Upadacitinib (RD [95 %-KI]; p-Wert: 0,11 [0,36; 0,18]; $p = 0,0035$). Abbildung 3 macht den Vorteil von Upadacitinib im zeitlichen Verlauf des ASAS40-Ansprechens deutlich. Eine deutliche Verbesserung zeigt sich schon nach zwei Wochen.

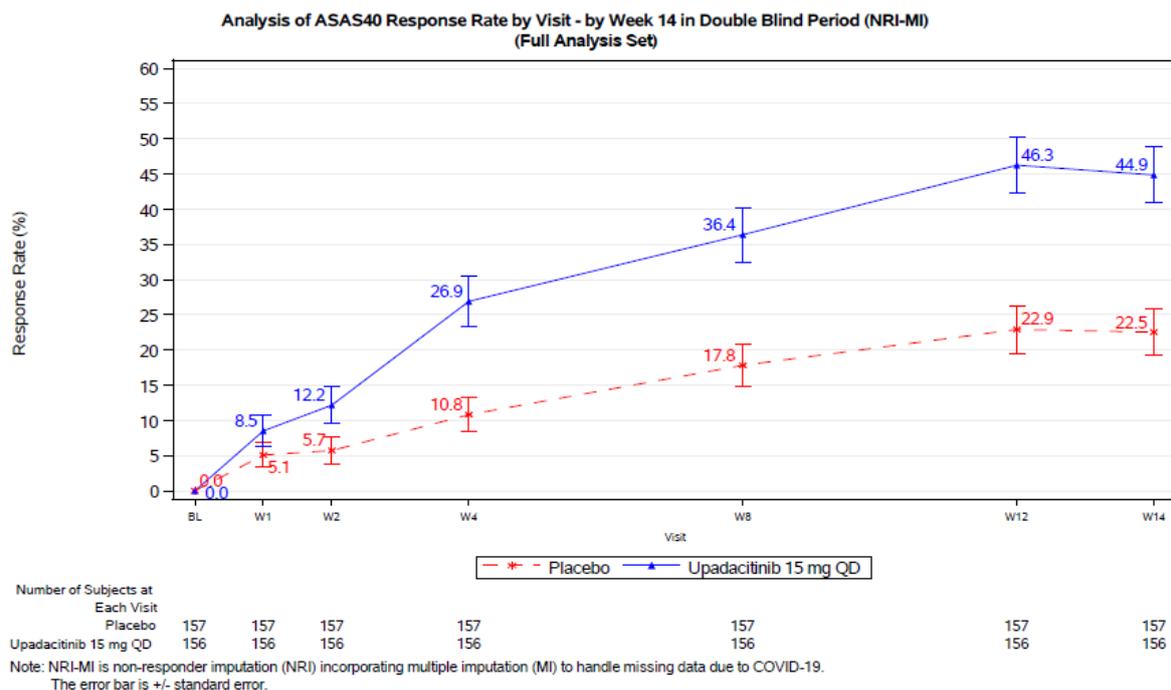


Abbildung 3: Anteil der Patienten mit einem ASAS40-Ansprechen von Baseline bis Woche 14 (Studie SELECT-AXIS 2, FAS)

ASAS: Assessment of SpondyloArthritis International Society; COVID: Corona Virus Disease; FAS: Full Analysis Set; MI: Multiple Imputation; NRI: Non-Responder-Imputation; QD: Einmal täglich

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.2.2 ASDAS – RCT

Tabelle 4-26: Operationalisierung von ASDAS

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der ASDAS setzt sich aus den fünf patientenberichteten Items Rückenschmerzen (Item 2 des BASDAI), peripherer Schmerz und Schwellung (Item 3 des BASDAI), Dauer der Morgensteifigkeit (Item 6 des BASDAI), PtGADA und hsCRP zusammen. Der Gesamtscore des ASDAS ergibt sich anhand einer Formel (Abschnitt 4.2.5.2) und nimmt Werte ≥ 0 an.</p> <p>Der ASDAS wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 2, Woche 4, Woche 8, Woche 12 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des ASDAS von Baseline bis Woche 14 untersucht. Außerdem werden Patienten mit einer inaktiven Erkrankung (Erreichen eines ASDAS $< 1,3$ zu Woche 14) und einer niedrigen Krankheitsaktivität (Erreichen eines ASDAS $< 2,1$ zu Woche 14) dargestellt.</p> <p>Die statistische Analyse für die mittlere Veränderung basiert auf dem FAS und erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie nicht fehlenden eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p> <p>Der Anteil der Patienten mit einer inaktiven Erkrankung sowie einer niedrigen Krankheitsaktivität basiert auf dem FAS. Die statistische Analyse erfolgte mithilfe eines CMH-Tests adjustiert für die Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP.</p> <p>Fehlende Werte sowie Werte nach Behandlungsabbruch wurden mittels NRI-MI imputiert (siehe Abschnitt 4.2.5.2).</p>
<p>ASDAS: Ankylosing Spondylitis Disease Activity Score; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; CMH: Cochran-Mantel-Haenszel; FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; NRI: Non-Responder-Imputation; PtGADA: Patientenberichtete globale Krankheitsaktivität</p>	

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

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

Tabelle 4-27: Bewertung des Verzerrungspotenzials für ASDAS in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ASDAS: Ankylosing Spondylitis Disease Activity Score; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-28: Ergebnisse für die mittlere Veränderung des ASDAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
ASDAS	154	3,61 (k. A.)	2,22 (k. A.)	-1,36 [-1,50, -1,21]	156	3,61 (k. A.)	2,90 (k. A.)	-0,71 [-0,85; -0,56]	-0,65 [-0,85; -0,45] p < 0,0001
<p>Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. ASDAS: Ankylosing Spondylitis Disease Activity Score; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen</p>									

Tabelle 4-29: Weitere Ergebnisse für ASDAS aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95 %-KI] p-Wert	OR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Woche 14					
ASDAS inaktive Erkrankung	22/156 (14,1)	8/157 (5,2)	k. A.	k. A.	0,09 [0,025; 0,15] p = 0,0063
ASDAS niedrige Krankheitsaktivität	66/156 (42,3)	29/157 (18,3)	k. A.	k. A.	0,24 [0,14; 0,33] p < 0,001

Endpunkt	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95 %-KI] p-Wert	OR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
<p>Imputationsmethode: NRI-MI Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN)</p> <p>Der Behandlungsunterschied, das zugehörige KI und der p-Wert für den Test des Unterschieds zwischen dem Upadacitinib-Arm und dem Placebo-Arm werden auf der Grundlage der MI-Inferenz geschätzt. Die RD und der SE werden für jeden mittels NRI-MI vervollständigten Datensatz mit dem CMH-Test adjustiert für die Stratifizierungsfaktoren geschätzt.</p> <p>Die Anzahl an Patienten mit einem Ansprechen (n) wird berechnet durch die Anzahl an Patienten in der Analyse (N) sowie der aggregierten MI-Responderate (%).</p> <p>Eine RD größer Null stellt einen Vorteil von Upadacitinib dar.</p> <p>ASDAS: Ankylosing Spondylitis Disease Activity Score; CMH: Cochran-Mantel-Haenszel; hsCRP: Hochsensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; MI: Multiple Imputation; MRT: Magnetresonanztomografie; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Response; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RD: Risikodifferenz; RCT: Randomisierte kontrollierte Studie; RR: Relatives Risiko; SE: Standardfehler; ULN: Obere Normgrenze</p>					

Zu Woche 14 zeigen sich positive Effekte der Behandlung mit Upadacitinib hinsichtlich der Krankheitsaktivität gemessen anhand des ASDAS. Dabei wird statistische Signifikanz zugunsten von Upadacitinib (MWD [95 %-KI; p-Wert: -0,65 [-0,85; -0,45]; $p < 0,0001$) erreicht. Auch hinsichtlich der Ergebnisse zum Anteil der Patienten mit einer nach dem ASDAS inaktiven Erkrankung sowie niedrigen Krankheitsaktivität zeigen sich deutlich die statistisch signifikanten Vorteile einer Upadacitinib-Behandlung im Vergleich zu Placebo (RD [95 %-KI]; p-Wert: ASDAS inaktive Erkrankung: 0,09 [0,025; 0,15]; $p = 0,0063$; ASDAS niedrige Krankheitsaktivität: 0,24 [0,14; 0,33]; $p < 0,001$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.2.3 BASDAI – RCT

Tabelle 4-30: Operationalisierung von BASDAI

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der BASDAI setzt sich zusammen aus sechs Items, welche jeweils mit einer NRS von 0 bis 10 folgende Bereiche untersuchen: Fatigue, axiale Schmerzen, periphere Arthritis, Enthesitis, sowie Intensität und Dauer der Morgensteifigkeit. Der BASDAI wird durch den Patienten selbst berichtet.</p> <p>Der BASDAI wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 2, Woche 4, Woche 8, Woche 12 und zu Woche 14.</p> <p>Der Gesamtscore des BASDAI ergibt sich anhand einer Formel (Abschnitt 4.2.5.2) und nimmt Werte zwischen 0 und 10 an. Das BASDAI50-Ansprechen ist definiert als Verbesserung des Gesamtscores um $\geq 50\%$ ab Baseline.</p> <p>Für das vorliegende Dossier wird der Anteil der Patienten mit BASDAI50-Ansprechen zu Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse erfolgte mithilfe eines CMH-Tests adjustiert für die Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP.</p> <p>Fehlende Werte sowie Werte nach Behandlungsabbruch wurden mittels NRI-MI imputiert (siehe Abschnitt 4.2.5.2).</p>
<p>BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; CMH: Cochran-Mantel-Haenszel; FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MI: Multiple Imputation; MRT: Magnetresonanztomografie; NRI: Non-Responder-Imputation; NRS: Numerical Rating Scale</p>	

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

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

Tabelle 4-31: Bewertung des Verzerrungspotenzials für BASDAI in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

Stellen Sie die Ergebnisse für den Endpunkt BASDAI50 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-32: Ergebnisse für BASDAI aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95 %-KI] p-Wert	OR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Woche 14					
Anteil der Patienten mit BASDAI50	66/156 (42,3)	35/157 (22,1)	k. A.	k. A.	0,20 [0,10; 0,30] p = 0,0001
<p>Imputationsmethode: NRI-MI Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Der Behandlungsunterschied, das zugehörige KI und der p-Wert für den Test des Unterschieds zwischen dem Upadacitinib-Arm und dem Placebo-Arm werden auf der Grundlage der MI-Inferenz geschätzt. Die RD und der SE werden für jeden mittels NRI-MI vervollständigten Datensatz mit dem CMH-Test adjustiert für die Stratifizierungsfaktoren geschätzt. Die Anzahl an Patienten mit einem Ansprechen (n) wird berechnet durch die Anzahl an Patienten in der Analyse (N) sowie der aggregierten MI-Responderate (%). Eine RD größer Null stellt einen Vorteil von Upadacitinib dar. BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; CMH: Cochran-Mantel-Haenszel; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; MI: Multiple Imputation; MRT: Magnetresonanztomografie; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Response; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie; RD: Risikodifferenz; RR: Relatives Risiko; SE: Standardfehler; ULN: Obere Normgrenze</p>					

Unter der Behandlung mit Upadacitinib erreichten fast doppelt so viele Patienten ein BASDAI50-Ansprechen als im Vergleichsarm mit Placebo. Zu Woche 14 hatten unter der Behandlung mit Upadacitinib statistisch signifikant mehr Patienten (42,3 %) ein BASDAI50-Ansprechen als im Placebo-Arm (22,1 %; RD [95 %-KI]; p-Wert: 0,20 [0,10; 0,30]; p = 0,0001).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum

einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.3 Schmerz

4.3.1.3.3.1 Patientenberichteter Rückenschmerz – RCT

Tabelle 4-33: Operationalisierung von patientenberichtetem Rückenschmerz

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der patientenberichtete Rückenschmerz wird anhand des Rückenschmerzes insgesamt (Patient's Assessment of Total Back Pain) bewertet. Der Gesamtscore ergibt sich durch die Bewertung auf einer NRS zwischen 0 und 10.</p> <p>Der patientenberichtete Rückenschmerz wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 2, Woche 4, Woche 8, Woche 12 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des patientenberichteten Rückenschmerzes von Baseline bis Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie nicht fehlenden eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p>
<p>FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; NRS: Numerical Rating Scale</p>	

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

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

Tabelle 4-34: Bewertung des Verzerrungspotenzials für patientenberichteten Rückenschmerz in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-35: Ergebnisse für die mittlere Veränderung des patientenberichteten Rückenschmerzes aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
Rückenschmerz	154	7,26 (k. A.)	4,29 (k. A.)	-2,91 [-3,27; -2,56]	156	7,32 (k. A.)	5,27 (k. A.)	-2,00 [-2,35; -1,65]	-0,92 [-1,42, -0,41] p = 0,0004
Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen									

Zu Woche 14 zeigen sich positive Effekte zugunsten der Behandlung mit Upadacitinib hinsichtlich des patientenberichteten Rückenschmerzes. Der Unterschied zwischen Upadacitinib und Placebo ist statistisch signifikant zugunsten von Upadacitinib (MWD [95 %-KI]; p-Wert: -0,92 [-1,42, -0,41]; p = 0,0004).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.4 Funktion

4.3.1.3.4.1 BASFI – RCT

Tabelle 4-36: Operationalisierung von BASFI

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der BASFI setzt sich zusammen aus 10 Items zu alltäglichen Aktivitäten, die auf einer Skala von 0 (einfach) bis 10 (unmöglich) hinsichtlich des Schweregrads durch den Patienten bewertet werden. Der Gesamtscore des BASFI ergibt sich aus dem Mittelwert der Antworten auf die 10 Items und nimmt somit Werte zwischen 0 und 10 an.</p> <p>Der BASFI wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 2, Woche 4, Woche 8, Woche 12 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des BASFI von Baseline bis Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie nicht fehlenden eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p>
<p>BASFI: Bath Ankylosing Spondylitis Functional Index; FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie</p>	

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

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

Tabelle 4-37: Bewertung des Verzerrungspotenzials für BASFI in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
BASFI: Bath Ankylosing Spondylitis Functional Index; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-38: Ergebnisse für die mittlere Veränderung des BASFI aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
BASFI	154	5,96 (k. A.)	3,33 (k. A.)	-2,61 [-2,94; -2,29]	156	5,97 (k. A.)	4,47 (k. A.)	-1,47 [-1,79; -1,15]	-1,14 [-1,60; -0,68] p < 0,0001
Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. BASFI: Bath Ankylosing Spondylitis Functional Index; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen									

Zu Woche 14 zeigen sich positive Effekte der Behandlung mit Upadacitinib hinsichtlich der Funktion gemessen anhand des BASFI. Dabei wird statistische Signifikanz zugunsten von Upadacitinib (MWD [95 %-KI]; p-Wert: -1,14 [-1,60; -0,68] p < 0,0001) erreicht.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.5 Entzündung

4.3.1.3.5.1 SPARCC MRI Spine – RCT

Tabelle 4-39: Operationalisierung von SPARCC MRI Spine

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der SPARCC MRI Spine setzt sich zusammen aus einer dichotomen Bewertung zu dem Vorhandensein, der Tiefe und der Intensität von Knochenmarködem-Läsionen der Wirbelsäule. Insgesamt werden dazu 23 DVUs bewertet, die sechs am stärksten betroffenen DVUs durch den Gutachter ausgewählt und zur Berechnung der Punktzahl des SPARCC MRI Spine herangezogen. Der Gesamtscore nimmt Werte bis 108 an.</p> <p>Der SPARCC MRI Spine wurde zu folgenden Messzeitpunkten bestimmt: Screening oder Baseline und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des SPARCC MRI Spine von Baseline bis Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels ANCOVA mit Behandlungsgruppe und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als feste Effekte und dem zugehörigen Baseline-Wert als Kovariable. Es werden die Werte aller Patienten mit Ausgangswert und Wert zu Woche 14 sowie die eingeschlossenen Faktoren und Kovariablen in dem Modell berücksichtigt. Fehlende Werte werden nicht ersetzt.</p>
<p>ANCOVA: Kovarianzanalyse; DVU: Discovertebrale Einheit; FAS: Full Analysis Set; hsCRP: Hochsensitives C-reaktives Protein; MRI: Magnetic Resonance Imaging; MRT: Magnetresonanztomografie; SPARCC: Spondyloarthritis Research Consortium of Canada</p>	

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

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

Tabelle 4-40: Bewertung des Verzerrungspotenzials für SPARCC MRI Spine in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ITT: Intention-to-Treat; MRI: Magnetic Resonance Imaging; RCT: Randomisierte kontrollierte Studie; SPARCC: Spondyloarthritis Research Consortium of Canada						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-41: Ergebnisse für die mittlere Veränderung des SPARCC MRI Spine aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
SPARCC MRI Spine	136	2,64 (k. A.)	1,69 (k. A.)	-0,79 [-1,48; -0,11]	146	1,45 (k. A.)	1,93 (k. A.)	0,34 [-0,32; 1,00]	-1,13 [-2,08; -0,17] p = 0,0206
Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf ANCOVA mit Behandlungsgruppe, den Stratifizierungsfaktoren und dem zugehörigen Baseline-Wert als Kovariable. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. ANCOVA: Kovarianzanalyse; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MRI: Magnetic Resonance Imaging; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert zu Baseline sowie Patienten ohne Wert zu Woche 14 werden aus der Analyse ausgeschlossen, da keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; SPARCC: Spondyloarthritis Research Consortium of Canada; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen									

Zu Woche 14 zeigen sich positive Effekte der Behandlung mit Upadacitinib hinsichtlich der Entzündung gemessen anhand des SPARCC MRI Spine. Dabei wird eine statistische Signifikanz zugunsten von Upadacitinib erreicht (MWD [95 %-KI]; p-Wert: -1,13 [-2,08; -0,17]; p = 0,0206).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.5.2 MASES – RCT

Tabelle 4-42: Operationalisierung von MASES

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der MASES bewertet die Abwesenheit (0) oder das Vorliegen (1) einer Enthesitis an 13 verschiedenen Stellen im Körper. Der Score ergibt sich durch Aufsummieren und nimmt Werte zwischen 0 und 13 an.</p> <p>Der MASES wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 4, Woche 8 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des MASES im Vergleich von Baseline bis Woche 14 basierend auf dem FAS, eingeschränkt auf Patienten mit Enthesitis zu Baseline, untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie nicht fehlenden eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p>
<p>FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie</p>	

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

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

Tabelle 4-43: Bewertung des Verzerrungspotenzials für MASES in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ITT: Intention-to-Treat; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-44: Ergebnisse für die mittlere Veränderung des MASES aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
MASES	124	4,7 (k. A.)	2,4 (k. A.)	-2,3 [-2,7; -1,9]	125	4,8 (k. A.)	3,2 (k. A.)	-1,6 [-2,0; -1,2]	-0,7 [-1,3; -0,1] p = 0,0193
<p>Population: Full Analysis Set, eingeschränkt auf Patienten mit Enthesitis zu Baseline Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen</p>									

Unter der Behandlung mit Upadacitinib zeigen sich positive Effekte der Behandlung gegenüber Placebo hinsichtlich der Entzündung der Sehnenansätze gemessen anhand des MASES. Der Unterschied zwischen den Behandlungen zu Woche 14 ist statistisch signifikant zugunsten von Upadacitinib (MWD [95 %-KI]; p-Wert: -0,7 [-1,3; -0,1]; p = 0,0193).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.6 Fatigue (FACIT-F) – RCT

Tabelle 4-45: Operationalisierung von FACIT-F

Studie	Operationalisierung
SELECT- AXIS 2	<p>Die Erhebung der Fatigue erfolgte mittels des FACIT-F als patientenberichteter Endpunkt. Der FACIT-F ist ein validiertes Messinstrument zur Erfassung der Fatigue. Der Fragebogen zur Fatigue ermittelt mithilfe einer 5-Punkte-Likert-Skala über 13 Fragen einen Gesamtscore, der das Ausmaß der Fatigue beschreibt. Eine genaue Beschreibung dieser Parameter befindet sich in Abschnitt 4.2.5.2.</p> <p>Der FACIT-F wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 2 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des FACIT-F von Baseline bis Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie der eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p>
<p>FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue; FAS: Full Analysis Set; hsCRP: Hoch sensitives C-reaktives Protein; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie</p>	

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

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

Tabelle 4-46: Bewertung des Verzerrungspotenzials für FACIT-F in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-47: Ergebnisse für die mittlere Veränderung des FACIT-F aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
FACIT-F	149	26,6 (k. A.)	35,0 (k. A.)	8,4 [7,0; 9,9]	154	26,0 (k. A.)	31,4 (k. A.)	5,3 [3,9; 6,7]	3,1 [1,1; 5,2] p = 0,0027

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
<p>Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen größer Null stellen einen Vorteil von Upadacitinib dar. FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen</p>									

Zu Woche 14 zeigen sich positive Effekte der Behandlung mit Upadacitinib hinsichtlich der Fatigue gemessen anhand des FACIT-F. Dabei wird eine statistische Signifikanz zugunsten von Upadacitinib erreicht (MWD [95 %-KI]; p-Wert: 3,1[1,1; 5,2] p = 0,0027).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.7 Gesundheitsbezogene Lebensqualität

4.3.1.3.7.1 ASQoL – RCT

Tabelle 4-48: Operationalisierung von ASQoL

Studie	Operationalisierung
SELECT- AXIS 2	<p>Der ASQoL besteht aus 18 Items, die je mit 0 oder 1 vom Patienten selbst bewertet werden. Der Gesamtscore des ASQoL ist ein Summenscore über die bewerteten Items und nimmt Werte von 0 bis 18 an, wobei niedrigere Werte eine bessere Lebensqualität beschreiben. Für eine valide Berechnung des Gesamtscores müssen mindestens 15 Items des ASQoL bewertet werden.</p> <p>Der ASQoL wurde zu folgenden Messzeitpunkten bestimmt: Baseline, Woche 1, Woche 4, Woche 8 und zu Woche 14.</p> <p>Für das vorliegende Dossier wird die mittlere Veränderung des ASQoL im Vergleich zu Baseline bis Woche 14 basierend auf dem FAS untersucht. Die statistische Analyse für die mittlere Veränderung erfolgte mittels MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert und den Stratifizierungsfaktoren Status der Entzündung anhand MRT und Screening hsCRP als Kovariablen. Werte aller Patienten mit Baseline-Wert und mindestens einem weiteren Wert bis Woche 14 sowie nicht fehlenden eingeschlossenen Faktoren und Kovariablen wurden in dem Modell berücksichtigt.</p>
<p>ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; FAS: Full Analysis Set; hsCRP: hoch sensitives C-reaktives Protein; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie</p>	

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

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

Tabelle 4-49: Bewertung des Verzerrungspotenzials für ASQoL in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-50: Ergebnisse für die mittlere Veränderung des ASQoL aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	N	Baseline MW (SD)	Visite MW (SD)	Δ LS-MW [95 %-KI]	Differenz Δ [95 %-KI] p-Wert
Woche 14									
ASQoL	151	11,90 (k. A.)	6,46 (k. A.)	-5,38 [-6,11; -4,65]	154	11,97 (k. A.)	8,86 (k. A.)	-3,15 [-3,87; -2,43]	-2,23 [-3,26; -1,21] p < 0,0001
<p>Population: Full Analysis Set Adjustierung: Status der Entzündung anhand MRT und Screening hsCRP (MRT+/hsCRP > ULN, MRT+/hsCRP ≤ ULN, MRT-/hsCRP > ULN) Mittelwertdifferenzen basieren auf MMRM mit Behandlungsgruppe, Visite und der Interaktion aus Behandlung und Visite als feste Effekte und dem zugehörigen Baseline-Wert sowie den Stratifizierungsfaktoren als Kovariablen. Mittelwertdifferenzen kleiner Null stellen einen Vorteil von Upadacitinib dar. ASQoL: Fragebogen zur Lebensqualität bei ankylosierender Spondylitis; hsCRP: Hoch sensitives C-reaktives Protein; k. A.: Keine Angabe; KI: Konfidenzintervall; LS-MW: Least-Squares-Mittelwert; MMRM: Mixed Model for Repeated Measures; MRT: Magnetresonanztomografie; MW: Mittelwert; N: Anzahl Patienten in der Analyse (Patienten ohne Wert nach Baseline werden aus der Analyse ausgeschlossen, da auf Basis des Baseline-Wertes keine Imputation durchgeführt wird); RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ULN: Obere Normgrenze; Δ: Veränderung im Vergleich zu Baseline bzw. Unterschied zwischen den Behandlungsarmen</p>									

Zu Woche 14 zeigen sich positive Effekte der Behandlung mit Upadacitinib hinsichtlich der gesundheitsbezogenen Lebensqualität gemessen anhand des ASQoL. Dabei wird eine statistische Signifikanz zugunsten von Upadacitinib erreicht (MWD [95 %-KI]; p-Wert: -2,23 [-3,26; -1,21]; p < 0,0001).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum

einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.8 Verträglichkeit – RCT

Tabelle 4-51: Operationalisierung von UE

Studie	Operationalisierung
SELECT- AXIS 2	<p>Zur Beurteilung der Verträglichkeit wurden die Gesamtraten für folgende UE dargestellt:</p> <ul style="list-style-type: none"> • Alle UE • SUE • UE, die zum Therapieabbruch führten • UE, die zum Tod führten • UE von speziellem Interesse • UE in $\geq 5\%$ der Patienten in einem Behandlungsarm <p>Im Prüfplan präspezifizierte UE von speziellem Interesse waren:</p> <ul style="list-style-type: none"> • Schwerwiegende Infektionen (CMQ: Infektion eingeschränkt auf SUE) • Opportunistische Infektion (CMQ: opportunistische Infektion ohne Herpes Zoster und Tuberkulose) • Mögliche Malignität (SMQ [eng]: Malignität) <ul style="list-style-type: none"> ○ Malignität (SMQ [eng]: maligne Tumoren) ○ Nicht melanozytärer Hautkrebs (NMSC) (SMQ [eng]: Hautkrebs ohne CMQ Melanom) ○ Malignität nicht NMSC (maligne Tumoren ohne NMSC) ○ Lymphome (SMQ [weit]: maligne Lymphome) • Herpes Zoster (CMQ: Herpes Zoster) • Aktive/latente Tuberkulose (CMQ: aktive Tuberkulose) • Gastrointestinale Perforation (basierend auf bewerteten Ereignissen) • Anämie (CMQ: nicht hämolytische und nicht aplastische Anämien) • Neutropenie (CMQ: hämatologische Toxizität – Neutropenie) • Lymphopenie (CMQ: hämatologische Toxizität – Lymphopenie) • Nierendysfunktion (SMQ [eng]: akutes Nierenversagen) • Hepatische Störung (SMQ [eng]: Wirkstoff abhängige hepatische Störung) • Bewertete kardiovaskuläre Ereignisse • Bewertete venöse thromboembolische Ereignisse <p>UE wurden von Behandlungsbeginn bis Behandlungsende zuzüglich einer Nachbeobachtungsdauer von 30 Tagen erfasst.</p> <p>Für das vorliegende Dossier wird der Anteil der Patienten mit mindestens einem UE zu Woche 14 dargestellt. Die Analyse der UE basierte auf dem Safety Analysis Set. Dieses umfasste alle Patienten, die mindestens eine Dosis der Studienmedikation erhielten.</p>

Studie	Operationalisierung
CMQ: Company MedDRA Query; MedDRA: Medical Dictionary for Regulatory Activities; NMSC: Nicht melanozytärer Hautkrebs; SMQ: Standardized MedDRA Query; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis	

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

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

Tabelle 4-52: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend.						
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

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

Tabelle 4-53: Ergebnisse für UE aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
UE	75/156 (48,1)	72/157 (45,9)
SUE	4/156 (2,6)	2/157 (1,3)
UE, die zum Therapieabbruch führten	4/156 (2,6)	2/157 (1,3)

Endpunkt	Upadacitinib	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
UE, die zum Tod führten	0/156 (0)	0/157 (0)
UE bei $\geq 5\%$ in einem Behandlungsarm		
COVID-19 ^a	6/156 (3,8)	9/157 (5,7)
Kopfschmerzen ^a	9/156 (5,8)	4/157 (2,5)
a: nach PT Population: Safety Analysis Set COVID: Corona Virus Disease; N: Anzahl Patienten in der Analyse; n: Anzahl von Patienten mit Ereignis; PT: Preferred Term; RCT: Randomisierte kontrollierte Studie; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis		

Die Gesamtrate jeglicher UE war zwischen dem Upadacitinib- und dem Placebo-Arm vergleichbar (ca. 48 % vs. ca. 46 %). Das Auftreten von UE, die zum Therapieabbruch führten, war in etwa vergleichbar in beiden Gruppen (2,6 % der Patienten im Upadacitinib- und 1,3 % der Patienten im Placebo-Arm). Die Anzahl schwerwiegender UE lag mit 2,6 % unter der Behandlung mit Upadacitinib bzw. 1,3 % unter Placebo in einem vergleichbaren Rahmen. In keinem der beiden Behandlungsarme traten Todesfälle auf. Ereignisse, die in mindestens einem Arm bei mehr als 5 % der Patienten auftraten, waren COVID-19 und Kopfschmerzen. Die Raten waren in beiden Behandlungsarmen vergleichbar.

Tabelle 4-54: Ergebnisse für häufige unerwünschte Ereignisse von speziellem Interesse aus RCT SELECT-AXIS 2 mit dem zu bewertenden Arzneimittel

Endpunkt	Upadacitinib	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
Infektion ^a	36/156 (23,1)	36/157 (22,9)
Schwerwiegende Infektion ^b	2/156 (1,3)	1/157 (0,6)
Malignität ^c	0/156 (0)	1/157 (0,6)
- NMSC ^c	0/156 (0)	1/157 (0,6)
Hepatische Störung	4/156 (2,6)	5/157 (3,2)
Anämie	1/156 (0,6)	0/157 (0)
Neutropenie	5/156 (3,2)	0/157 (0)
Herpes Zoster ^d	2/156 (1,3)	1/157 (0,6)

Endpunkt	Upadacitinib	Placebo
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
<p>a: Im CSR wurde zusätzlich das Auftreten aller Infektionen als unerwünschte Ereignisse von speziellem Interesse berichtet und entsprechend hier ergänzend dargestellt.</p> <p>b: Bei den schwerwiegenden Infektionen handelte es sich im Upadacitinib-Arm um eine COVID-19 Lungenentzündung und eine Pyelonephritis, im Placebo-Arm handelte es sich um Haemorrhagisches Fieber mit renalem Syndrom.</p> <p>c: Ein Patient im Placebo-Arm hatte ein Basalzellkarzinom.</p> <p>d: Keine der Herpes Zoster Infektionen wurde der Kategorie schwerwiegend zugeordnet, die beobachteten Herpes Zoster Infektionen waren Ereignisse der Kategorien mild oder moderat.</p> <p>Population: Safety Analysis Set</p> <p>Es werden nur UESI berichtet, die bei mindestens einem Patienten im Verlauf der Studie auftraten.</p> <p>CSR: Klinischer Studienreport; N: Anzahl Patienten in der Analyse; n: Anzahl von Patienten mit Ereignis; NMSC: Nicht melanozytärer Hautkrebs; RCT: Randomisierte kontrollierte Studie; UESI: Unerwünschtes Ereignis von speziellem Interesse</p>		

Insgesamt sind die bis zum Analysezeitpunkt zu Woche 14 erfassten UESI in beiden Behandlungsgruppen überwiegend der Kategorie Infektionen zuzuordnen. Der Anteil an Patienten mit Infektionen im Upadacitinib-Arm war vergleichbar mit dem Anteil im Placebo-Arm. Bei den Infektionen handelte es sich überwiegend um nicht schwere Infektionen. Weiterhin wurden keine VTE, keine MACE oder malignen Ereignisse unter einer Behandlung mit Upadacitinib berichtet. Weitere UESI wurden in geringer Anzahl dokumentiert.

Der Anteil der Patienten mit UE war im Allgemeinen ähnlich in den Upadacitinib- und Placebo-Gruppen. Insgesamt werden die beschriebenen Unterschiede als geringfügig eingestuft. Das Sicherheitsprofil von Upadacitinib entspricht dem bekannten und gut charakterisierten Sicherheitsprofil von Upadacitinib ohne neue unerwartete UE.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorliegt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 dargestellt.

4.3.1.3.9 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.¹⁶

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

¹⁶ unbesetzt

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Da keine direkt oder indirekt vergleichende Evidenz mit dem zu bewertenden Arzneimittel und einer ZVT vorliegt, werden im vorliegenden Dossier keine Subgruppenanalysen dargestellt.

Tabelle 4-55 Matrix der durchgeführten Subgruppenanalysen

Endpunkt Studie	Alter	Geschlecht	<Effekt- modifikator -a>	<Effekt- modifikator -b>	<Effekt- modifikator -c>	<Effekt- modifikator -d>
Nicht zutreffend.						
●: A priori geplante Subgruppenanalyse ○: Posthoc durchgeführte Subgruppenanalyse n. d.: Subgruppenanalyse nicht durchgeführt.						

Stellen Sie anschließend in Tabelle 4-56 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

Tabelle 4-56: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für <Studie> und <Effektmodifikator>

Endpunkt Studie	Alter	Geschlecht	<Effekt- modifikator -a>	<Effekt- modifikator -b>	<Effekt- modifikator -c>	<Effekt- modifikator -d>
Gesamtmortalität						
<Studie 1>	p=0,345	p=0,321	p=0,003	p=0,041	p=0,981	p=0,212
<Studie 2>	p=0,634	p=0,212	p<0,001	k. A.	k. A.	k. A.
<Endpunkt 2>						
...						
k. A.: Keine Angabe						

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere

unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Nicht zutreffend.

4.3.1.4 Liste der eingeschlossenen Studien – RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Es wurden keine für die Nutzenbewertung relevanten Studien identifiziert. Die Ergebnisse zur Wirksamkeit und Verträglichkeit von Upadacitinib bei Patienten mit aktiver nr-axSpA werden anhand der Studie SELECT-AXIS 2 dargestellt.

Studie	Quelle
SELECT-AXIS 2 (M19-944)	Studienbericht inklusive Studienprotokoll und statistischem Analyseplan (108) Studienregistereinträge (109-111)

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer

Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken
- Studien aus der Suche auf der Internetseite des G-BA
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.**

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-57: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in

drei Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs-therapie 1>	<Vergleichs-therapie 2>	<Vergleichs-therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2> <Studie 3>	• •		• •	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-59: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-60: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-61: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- Homogenität der Ergebnisse: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt für den ein indirekter Vergleich vorgenommen wird fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.9.

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken
- Studien aus der Suche auf der G-BA Internetseite
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend.

4.3.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.**

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-62: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-63: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-64: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-65: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.9.

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie***

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-66: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-67: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind,

gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.9.***

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Zur Beurteilung des medizinischen Nutzens von Upadacitinib wurde eine Fragestellung formuliert. Für die Bewertung des medizinischen Nutzens und Zusatznutzens von Upadacitinib

zur Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben, liegt eine Studie vor. Diese Studie entspricht aufgrund der Studiendauer und des fehlenden Vergleichs mit einer ZVT jedoch nicht den G-BA-Anforderungen und wird nicht zur Ableitung eines Zusatznutzens herangezogen. Im Dossier wurde die Wirksamkeit und Verträglichkeit einer Behandlung mit Upadacitinib bei Patienten mit aktiver nr-axSpA, die unzureichend auf mindestens zwei vorangegangene Therapien mit NSAR angesprochen haben, anhand der SELECT-AXIS 2 dargestellt.

Fragestellung 1 (erwachsene Patienten mit aktiver nr-axSpA, die unzureichend auf eine vorangegangene Therapie mit NSAR angesprochen haben)

Evidenzstufe

Die vorliegende Studie SELECT-AXIS 2 (M19-944) ist eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie der Evidenzstufe Ib (nach § 11 2. Kapitel 3. Abschnitt der VerFO des G-BA). Die Studienpopulation entspricht der Fragestellung 1.

Studienqualität

Die Studie SELECT-AXIS 2 wurde bei erwachsenen Patienten mit aktiver nr-axSpA, die unzureichend auf mindestens zwei vorangegangene Therapien mit NSAR angesprochen haben, durchgeführt.

Insgesamt wurden 314 Patienten in der Studie SELECT-AXIS 2 mittels einer computer-generierten Randomisierungssequenz in einem Verhältnis von 1:1 in die beiden Behandlungsarme randomisiert (Upadacitinib: N = 156; Placebo: N = 158). Die Zuteilung erfolgte jeweils durch ein interaktives Response Technologie System (IRT), wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte stratifiziert nach hsCRP-Level an Screening, Status der Entzündung anhand MRT sowie vorheriger Therapie mit bDMARDs. Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte erste Studienperiode (mindestens 52 Wochen bzw. bis alle Patienten die erste Studienperiode abgeschlossen hatten) verblindet. Endpunkte wurden gemäß SAP berichtet und es kann von einer ergebnisunabhängigen Berichterstattung ausgegangen werden.

Die Dauer der Studie beträgt in der ersten Periode 52 Wochen und wird in der zweiten Periode um weitere 52 Wochen bis zum Studienende (mindestens 104 Wochen) fortgesetzt. Die ersten 14 Wochen bis zum primären Endpunkt der Studie sind abgeschlossen, Daten zur ersten Periode der Studie bis Woche 52 liegen noch nicht vor und die zweite Periode ist noch laufend. Patienten und Behandler werden gemäß Studienprotokoll entblindet, nachdem alle Patienten die erste Periode abgeschlossen haben.

Eine Rescue-Therapie ist für Patienten vorgesehen, die ab der Studienvisite zu Woche 24 bis Ende der Periode 1 zu Woche 52 an zwei aufeinanderfolgenden Visiten kein ASAS20-Ansprechen zeigen. Die Rescue-Therapie ergibt sich durch Hinzunahme oder Anpassung der

Begleitmedikation, die entsprechend im Rahmen des Studienprotokolls möglich ist. Bei Abbruch der Studienmedikation kann auch die Behandlung mit einem bDMARD (TNF- oder IL-17-Inhibitor) begonnen werden. Patienten, die in Periode 1 gemäß der Randomisierung mit Placebo behandelt werden, werden zu Woche 52 auf eine Behandlung mit Upadacitinib umgestellt. Gemäß Studienprotokoll ist vorgesehen, dass Patienten, die ab Woche 64 bis Woche 104 kein ASAS20-Ansprechen an zwei aufeinanderfolgenden Visiten zeigen, die Studienmedikation abbrechen.

Validität der herangezogenen Endpunkte

Für die Bewertung des medizinischen Nutzens wurden Daten aus den Kategorien Mortalität, Morbidität, Lebensqualität und Verträglichkeit berücksichtigt. Für die Operationalisierung der entsprechenden Kategorien wurden Daten zu patientenrelevanten Endpunkten verwendet. Die untersuchten Endpunkte wurden mittels validierter und krankheitsspezifischer Instrumente erhoben und gemäß der im Studienprotokoll festgelegten Vorgehensweise ausgewertet (Intention-to-Treat [ITT]-Prinzip).

Im vorliegenden Dossier wird Woche 14 als Analysezeitpunkt dargestellt. Zu diesem Zeitpunkt war die Studie SELECT-AXIS 2 placebokontrolliert. Ein Behandlungswechsel war innerhalb dieses Zeitraums nicht vorgesehen. Für die Beurteilung aller Endpunkte wird aufgrund des fehlenden direkten Vergleichs zur ZVT kein Zusatznutzen beansprucht.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*

- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Hintergrund

Upadacitinib wird angewendet zur Behandlung der aktiven nicht röntgenologischen axialen Spondyloarthritis bei erwachsenen Patienten mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes C-reaktives Protein (CRP) und/oder Nachweis durch Magnetresonanztomografie (MRT), die unzureichend auf nicht steroidale Antirheumatika (NSAR) angesprochen haben.

Die nr-axSpA stellt eine facettenreiche Erkrankung dar und kann das Achsenskelett, die peripheren Gelenke, Entesen und extraartikuläre Domänen betreffen. Im Vordergrund des Krankheitsgeschehens stehen Schmerzen und Bewegungseinschränkungen des Achsenskeletts, was zu Einschränkungen der Lebensqualität und des sozialen Lebens führt. Für viele Patienten mit nr-axSpA sind trotz verfügbarer Therapieoptionen in Form von NSAR und bDMARDs (TNF- und IL-17-Inhibitoren) die Therapieziele einer niedrigen Krankheitsaktivität und besonders einer anhaltenden Remission nach wie vor nicht zu erreichen (23). Somit besteht weiterhin ein großer Bedarf an Therapieoptionen mit neuen Wirkmechanismen bei der nr-axSpA.

Upadacitinib ist der einzige zugelassene JAK-Inhibitor zur Behandlung der nr-axSpA bei erwachsenen Patienten, dessen Wirksamkeit und Sicherheit in der Studie SELECT-AXIS 2 untersucht wurde.

Fragestellung und Population in der Indikation nr-axSpA

Basierend auf der von der Europäischen Kommission zugelassenen Indikation für Upadacitinib wird für die vorliegende frühe Nutzenbewertung folgende Population betrachtet (Tabelle 4-68).

Tabelle 4-68: Übersicht über die zu betrachtende Fragestellung bzw. Population

Fragestellung	Population	Intervention	ZVT
Gesamtpopulation			
1	Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben	Upadacitinib	TNF-Inhibitor (Etanercept oder Adalimumab oder Golimumab oder Certolizumab Pegol)
CRP: C-reaktives Protein; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum; TNF: Tumornekrosefaktor; ZVT: Zweckmäßige Vergleichstherapie			

Für die oben genannte Population wurde keine Evidenz für einen direkten Vergleich zur ZVT identifiziert. Es liegen zudem keine Studien zu Upadacitinib vor, die für einen adjustierten oder historischen indirekten Vergleich zur ZVT geeignet sind. AbbVie verzichtet daher auf die Durchführung von indirekten Vergleichen. Eine Aussage über einen Zusatznutzen von Upadacitinib gegenüber der ZVT ist somit für die Zielpopulation sowie für die beiden Populationen der bDMARD-naiven und bDMARD-erfahrenen Patienten nicht möglich und ein Zusatznutzen wird entsprechend nicht beansprucht. Aufgrund der fehlenden Evidenz wird auf eine Darstellung der Populationen bDMARD-naive und bDMARD-erfahrene Patienten im Rahmen des Moduls verzichtet.

Die Wirksamkeit und Sicherheit von Upadacitinib wurde in der placebokontrollierten Studie SELECT-AXIS 2 untersucht und die Ergebnisse sind im vorliegenden Nutzendossier ergänzend dargestellt.

Medizinischer Nutzen von Upadacitinib – Ergebnisse zur Wirksamkeit und Verträglichkeit

Die vorliegende Studie SELECT-AXIS 2 (M19-944) ist eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie. Die EMA hat bereits im Zulassungsprozess ein positives Nutzen-Risikoverhältnis für eine Upadacitinib-Behandlung in der Indikation nr-axSpA festgestellt. Die Studienpopulation beinhaltet erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, die unzureichend auf NSAR angesprochen haben. Da keine Daten im Vergleich zur ZVT vorliegen, werden die Ergebnisse der Studie SELECT-AXIS 2 ergänzend dargestellt, um die Wirksamkeit und Sicherheit von Upadacitinib zu beschreiben.

In der Studie SELECT-AXIS 2 wurden patientenrelevante Endpunkte in den Kategorien Morbidität (Krankheitsaktivität, Schmerz, Funktion, Entzündung und Fatigue), gesundheitsbezogene Lebensqualität und Verträglichkeit erhoben.

Die nr-axSpA manifestiert sich in Knochen und Gelenken der unteren Wirbelsäule wie auch der Verbindung zum Becken (Iliosakralgelenk) sowie in benachbarten Weichteilen, wie z. B. Sehnen und Bändern (3, 112). Patienten mit nr-axSpA leiden unter chronischen Schmerzen und Bewegungseinschränkungen des Achsenskeletts, was zu Einschränkungen der Lebensqualität und des sozialen Lebens führt (3).

Das primäre Ziel sowohl der ASAS/EULAR- als auch der deutschen S3-Leitlinie bei der Behandlung von Patienten mit nr-axSpA ist die langfristige Verbesserung und Erhaltung der gesundheitsbezogenen Lebensqualität. Die Therapie sollte darauf ausgerichtet sein, eine größtmögliche Schmerzreduktion zu erreichen und gleichzeitig den Erhalt der körperlichen Funktionsfähigkeit, die Reduktion der Steifigkeit, die Verhinderung struktureller Läsionen und den Erhalt der Erwerbsfähigkeit zu gewährleisten (3, 6). Dabei nimmt die Reduktion der Krankheitsaktivität, gemessen anhand stringenter Endpunkte, wie z. B. ASDAS niedrige Krankheitsaktivität, eine zentrale Rolle in der Behandlung von Patienten mit nr-axSpA ein (3). Die deutsche S3-Leitlinie definiert die Remission als das Erreichen einer partiellen Remission gemäß ASAS-Kriterien oder als das Erreichen einer ASDAS inaktive Erkrankung.

Krankheitsaktivität

Wie oben beschrieben ist das Erreichen einer niedrigen Krankheitsaktivität bzw. Remission ein zentrales Therapieziel (6). Das ASAS-Ansprechen, der ASDAS und der BASDAI bilden in der Studie SELECT-AXIS 2 die Krankheitsaktivität ab. Hier konnte durchweg eine stringente Kontrolle der Krankheitssymptomatik durch Upadacitinib im Vergleich zu Placebo beobachtet werden. Eine statistisch signifikante Verbesserung der Krankheitsaktivität konnte auch schon nach zwei Wochen anhand des ASAS40-Ansprechens gezeigt werden (108). Besonders deutlich zeigten sich die positiven Effekte zur Krankheitsaktivität auch durch den signifikant höheren Anteil der Patienten, die eine Remission (ASAS-PR; ASDAS inaktive Erkrankung [ASDAS < 1,3]) oder niedrige Krankheitsaktivität (ASDAS \geq 1,3 und < 2,1) unter einer Behandlung mit Upadacitinib erreichten.

Schmerzen

Bei Patienten mit nr-axSpA stehen neben Bewegungseinschränkungen des Achsenskeletts insbesondere Schmerzen im Vordergrund, weshalb die Schmerzreduktion eines der wichtigsten Therapieziele ist (3). Unter einer Upadacitinib-Behandlung kann eine deutliche Verbesserung der patientenberichteten Rückenschmerzen erreicht werden. Diese Schmerzreduktion kann bereits nach zwei Wochen beobachtet werden (108). Neben den hier dargestellten Daten zum patientenberichteten Rückenschmerz anhand einer patientenberichteten NRS zum Rückenschmerz insgesamt, wurde gezeigt, dass die Einnahme von Upadacitinib über alle Schmerzdomänen (axialer, peripherer und nächtlicher Schmerz; erhoben mit den Items 2 und 3 des BASDAI bzw. einer patientenberichteten NRS zum nächtlichen Rückenschmerz) hinweg zu einer deutlichen und sehr schnellen Verbesserung führt (108).

Funktion

Neben den Rückenschmerzen ist die zunehmende Steifheit der Wirbelsäule, die zu einer mitunter dramatischen und bewegungseinschränkenden Veränderung der Körperhaltung führen

kann, für Patienten mit nr-axSpA von zentraler Bedeutung (3). In der Studie SELECT-AXIS 2 wurde die Funktion anhand des BASFI beurteilt. Es konnten signifikante Verbesserungen der Funktionseinschränkungen im Vergleich zu Placebo gezeigt werden.

Entzündung

Entzündungen im Bereich der Wirbelsäule oder auch Enthesen können Schmerzen und/oder Schwellungen der peripheren und axialen Gelenke sowie der gelenknahen Strukturen verursachen und erheblich zur Krankheitslast beitragen (87, 88). Die Ausprägung der Entzündung der Wirbelsäule wurde in der Studie SELECT-AXIS 2 über den SPARCC MRI Spine erhoben, das Ausmaß der Enthesitis wurde durch den MASES bestimmt. Auch in diesen bedeutenden Krankheitsmanifestationen zeigte eine Upadacitinib-Therapie eine signifikante Verbesserung der akuten Entzündung gegenüber Placebo.

Fatigue

Fatigue ist für Patienten mit nr-axSpA ein bedeutendes Symptom, das mit Einschränkungen der Funktionalität einhergehen kann und einen erheblichen Einfluss auf Arbeitsfähigkeit und Lebensqualität der Patienten hat (3, 93). Eine Verbesserung der Fatigue sollte nach dem Empfinden von Patienten ein Therapieziel sein. Das Ausmaß der Fatigue wurde in der Studie SELECT-AXIS 2 mittels des Fragebogens FACIT-F erhoben. Gegenüber Placebo konnte eine signifikante Verbesserung der Fatigue durch die Upadacitinib-Therapie gezeigt werden.

Gesundheitsbezogene Lebensqualität

Als Behandlungsziel wird von nr-axSpA-Patienten neben der Eliminierung oder Reduktion der Schmerzen die Verbesserung der Funktionsfähigkeit und der Lebensqualität genannt (69). Die gesundheitsbezogene Lebensqualität wurde in der Zulassungsstudie mittels des ASQoL erhoben, es konnte auch hier eine signifikante Verbesserung durch die Upadacitinib-Behandlung gezeigt werden.

Verträglichkeit

Das Auftreten jeglicher unerwünschter Ereignisse in der Studie SELECT-AXIS 2 war vergleichbar unter Upadacitinib und Placebo. In der Studie SELECT-AXIS 2 wurden nur sehr wenige schwerwiegende Infektionen berichtet. Darüber hinaus traten keine Malignitäten, MACE, VTE und Todesfälle unter der Behandlung mit Upadacitinib auf (108).

Schlussfolgerung zur Wirksamkeit und Verträglichkeit von Upadacitinib

Für viele Patienten mit nr-axSpA sind trotz verfügbarer Therapieoptionen in Form von NSAR und bDMARDs (TNF- und IL-17-Inhibitoren) die Therapieziele einer niedrigen Krankheitsaktivität und besonders einer anhaltenden Remission nach wie vor nicht zu erreichen. Entsprechend besteht weiterhin ein großer Bedarf an neuen Therapieoptionen. Upadacitinib ist neben dem guten Wirksamkeits- und Verträglichkeitsprofil auch aufgrund des neuen Wirkmechanismus eine Bereicherung für die Therapielandschaft.

In der Gesamtschau konnten in der pivotalen Zulassungsstudie SELECT-AXIS 2 in den Wirksamkeitsendpunkten zur Krankheitsaktivität, zu Schmerzen/Rückenschmerzen, zur

körperlichen Funktionsfähigkeit und zur Entzündung deutliche Vorteile einer Upadacitinib-Behandlung gegenüber Placebo bei nr-axSpA-Patienten gezeigt werden. Dies lässt auf eine stringente Kontrolle der Krankheitsaktivität sowie eine schnelle Schmerzreduktion bei den Patienten schließen und trägt zum Erreichen des primären Therapieziels der deutschen S3-Leitlinie, der langfristigen Verbesserung und Erhaltung der gesundheitsbezogenen Lebensqualität der Patienten bei. Das Auftreten unerwünschter Ereignisse war unter Upadacitinib und Placebo ähnlich. Insgesamt umfasst das Upadacitinib-Studienprogramm klinische Studien der Phase-II/III bei RA, PsA, AS, nr-axSpA, CU und AD mit insgesamt über 5.000 Patienten, die bis zu 4,5 Jahre mit Upadacitinib behandelt wurden (113-115). Das Gesamtsicherheitsprofil von Upadacitinib ist über die zugelassenen Indikationen konsistent, gut charakterisiert und es wurden keine neuen unerwarteten Nebenwirkungen beobachtet.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-69: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit aktiver nr-axSpA mit objektiven Anzeichen einer Entzündung, angezeigt durch erhöhtes CRP und/oder Nachweis durch MRT, die unzureichend auf NSAR angesprochen haben	kein Zusatznutzen belegt
CRP: C-reaktives Protein; MRT: Magnetresonanztomografie; nr-axSpA: Nicht röntgenologische axiale Spondyloarthritis; NSAR: Nicht steroidales Antirheumatikum	

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

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

Nicht zutreffend.

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

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

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁷, Molenberghs 2010¹⁸). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁹) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und

¹⁷ Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

¹⁸ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

¹⁹ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%- Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006²⁰) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

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

Nicht zutreffend.

4.6 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

References

1. AbbVie Deutschland GmbH & Co. KG. Fachinformation für RINVOQ[®] 15 mg Retardtabletten, RINVOQ[®] 30 mg Retardtabletten, RINVOQ[®] 45 mg Retardtabletten Stand: Juli 2022.
2. Gemeinsamer Bundesausschuss (G-BA). Niederschrift zum Beratungsgespräch gemäß § 8 Absatz 1 AM-NutzenV, Beratungsanforderung 2018-B-186. 2018.

²⁰ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

3. Kiltz U, Braun J, Becker A, Chenot JF, Dreimann M, Hammel L, et al. Langfassung zur S3-Leitlinie Axiale Spondyloarthritis inklusive Morbus Bechterew und Frühformen, Update 2019: Evidenzbasierte Leitlinie der Deutschen Gesellschaft für Rheumatologie (DGRh) und der beteiligten medizinisch-wissenschaftlichen Fachgesellschaften und weiterer Organisationen. *Z Rheumatol.* 2019;78(Suppl 1):3-64.
4. Zhu W, He X, Cheng K, Zhang L, Chen D, Wang X, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. *Bone research.* 2019;7(1):1-16.
5. Protopopov M, Poddubnyy D. Radiographic progression in non-radiographic axial spondyloarthritis. *Expert Review of Clinical Immunology.* 2018;14(6):525-33.
6. van der Heijde D, Ramiro S, Landewé R, Baraliakos X, Van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis.* 2017;76(6):978-91.
7. Smolen JS, Schöls M, Braun J, Dougados M, FitzGerald O, Gladman DD, et al. Treating axial spondyloarthritis and peripheral spondyloarthritis, especially psoriatic arthritis, to target: 2017 update of recommendations by an international task force. *Ann Rheum Dis.* 2018;77(1):3-17.
8. Smolen JS, Braun J, Dougados M, Emery P, Fitzgerald O, Helliwell P, et al. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international task force. *Ann Rheum Dis.* 2014;73(1):6-16.
9. van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. *The Journal of rheumatology.* 1997;24(11):2225-9.
10. Ciurea A, Scherer A, Exer P, Bernhard J, Dudler J, Beyeler B, et al. Tumor necrosis factor α inhibition in radiographic and nonradiographic axial spondyloarthritis: results from a large observational cohort. *Arthritis & Rheumatism.* 2013;65(12):3096-106.
11. Kiltz U, Baraliakos X, Karakostas P, Igelmann M, Kalthoff L, Klink C, et al. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? *Arthritis care & research.* 2012;64(9):1415-22.
12. Wallman JK, Kapetanovic MC, Petersson IF, Geborek P, Kristensen LE. Comparison of non-radiographic axial spondyloarthritis and ankylosing spondylitis patients—baseline characteristics, treatment adherence, and development of clinical variables during three years of anti-TNF therapy in clinical practice. *Arthritis research & therapy.* 2015;17(1):1-11.
13. Ramiro S, van der Heijde D, van Tubergen A, Stolwijk C, Dougados M, van den Bosch F, et al. Higher disease activity leads to more structural damage in the spine in ankylosing spondylitis: 12-year longitudinal data from the OASIS cohort. *Ann Rheum Dis.* 2014;73(8):1455-61.
14. Poddubnyy D, Sieper J. Mechanism of new bone formation in axial spondyloarthritis. *Current Rheumatology Reports.* 2017;19(9):1-9.
15. Committee for Medicinal Products for Human Use (CHMP). Guideline on the Clinical Investigation of Medicinal Products for the Treatment of Axial Spondyloarthritis. 2017. Verfügbar unter: https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-axial-spondyloarthritis-revision-1_en.pdf. [Zugriff am: 10.05.2022]
16. Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. *Arthritis Rheum.* 2001;44(8):1876-86.

17. Landewé R, van Tubergen A. Clinical Tools to Assess and Monitor Spondyloarthritis. *Curr Rheumatol Rep.* 2015;17(7):47.
18. AbbVie Inc. Statistical Analysis Plan M19-944. A Phase 3 Randomized, Placebo-Controlled, Double-Blind Program to Evaluate Efficacy and Safety of Upadacitinib in Adult Subjects with Axial Spondyloarthritis Followed by a Remission-withdrawal Period. Study 2: Non-Radiographic Axial Spondyloarthritis (nr-axSpA). 2021.
19. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis.* 2009;68 Suppl 2:ii1-44.
20. van der Heijde D, Joshi A, Pangan AL, Chen N, Betts K, Mittal M, et al. ASAS40 and ASDAS clinical responses in the ABILITY-1 clinical trial translate to meaningful improvements in physical function, health-related quality of life and work productivity in patients with non-radiographic axial spondyloarthritis. *Rheumatology (Oxford).* 2016;55(1):80-8.
21. Davis JC, Jr., Revicki D, van der Heijde DM, Rentz AM, Wong RL, Kupper H, et al. Health-related quality of life outcomes in patients with active ankylosing spondylitis treated with adalimumab: results from a randomized controlled study. *Arthritis Rheum.* 2007;57(6):1050-7.
22. Dougados M, van der Heijde D, Tsai W-C, Saaibi D, Marshall L, Jones H, et al. Relationship between disease activity status or clinical response and patient-reported outcomes in patients with non-radiographic axial spondyloarthritis: 104-week results from the randomized controlled EMBARK study. *Health and quality of life outcomes.* 2020;18(1):1-7.
23. Baraliakos X, Berenbaum F, Favalli EG, Olivieri I, Ostendorf B, Poddubnyy D, et al. Challenges and Advances in Targeting Remission in Axial Spondyloarthritis. *J Rheumatol.* 2018;45(2):153-7.
24. Rudwaleit M, Claudepierre P, Wordsworth P, Cortina EL, Sieper J, Kron M, et al. Effectiveness, safety, and predictors of good clinical response in 1250 patients treated with adalimumab for active ankylosing spondylitis. *J Rheumatol.* 2009;36(4):801-8.
25. Sieper J, van der Heijde D, Dougados M, Brown LS, Lavie F, Pangan AL. Early response to adalimumab predicts long-term remission through 5 years of treatment in patients with ankylosing spondylitis. *Ann Rheum Dis.* 2012;71(5):700-6.
26. Sieper J, Rudwaleit M, Lenaerts J, Wollenhaupt J, Myasoutova L, Park SH, et al. Partial remission in ankylosing spondylitis and non-radiographic axial spondyloarthritis in treatment with infliximab plus naproxen or naproxen alone: associations between partial remission and baseline disease characteristics. *Rheumatology (Oxford).* 2016;55(11):1946-53.
27. Sieper J, van der Heijde D, Dougados M, Maksymowych WP, Scott BB, Boice JA, et al. A randomized, double-blind, placebo-controlled, sixteen-week study of subcutaneous golimumab in patients with active nonradiographic axial spondyloarthritis. *Arthritis Rheumatol.* 2015;67(10):2702-12.
28. van der Heijde D, Dijkmans B, Geusens P, Sieper J, DeWoody K, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). *Arthritis Rheum.* 2005;52(2):582-91.
29. Braun J, Blanco R, Marzo-Ortega H, Gensler LS, Van den Bosch F, Hall S, et al. Secukinumab in non-radiographic axial spondyloarthritis: subgroup analysis based on key baseline characteristics from a randomized phase III study, PREVENT. *Arthritis research & therapy.* 2021;23(1):1-10.

30. Brandt J, Listing J, Sieper J, Rudwaleit M, van der Heijde D, Braun J. Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. *Ann Rheum Dis.* 2004;63(11):1438-44.
31. Zochling J. Measures of symptoms and disease status in ankylosing spondylitis: Ankylosing Spondylitis Disease Activity Score (ASDAS), Ankylosing Spondylitis Quality of Life Scale (ASQoL), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), Bath Ankylosing Spondylitis Global Score (BAS-G), Bath Ankylosing Spondylitis Metrology Index (BASMI), Dougados Functional Index (DFI), and Health Assessment Questionnaire for the Spondylarthropathies (HAQ-S). *Arthritis Care Res (Hoboken).* 2011;63 Suppl 11:S47-58.
32. Kilic E, Kilic G, Akgul O, Ozgocmen S. Discriminant validity of the Ankylosing Spondylitis Disease Activity Score (ASDAS) in patients with non-radiographic axial spondyloarthritis and ankylosing spondylitis: a cohort study. *Rheumatology international.* 2015;35(6):981-9.
33. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis.* 2009;68(1):18-24.
34. van der Heijde D, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al. ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. *Ann Rheum Dis.* 2009;68(12):1811-8.
35. Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. *Ann Rheum Dis.* 2011;70(1):47-53.
36. Machado PM, Landewé R, Heijde DV. Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. *Ann Rheum Dis.* 2018;77(10):1539-40.
37. Machado PM, Landewé RB, van der Heijde DM. Endorsement of definitions of disease activity states and improvement scores for the Ankylosing Spondylitis Disease Activity Score: results from OMERACT 10. *J Rheumatol.* 2011;38(7):1502-6.
38. Poddubnyy D, Protopopov M, Haibel H, Braun J, Rudwaleit M, Sieper J. High disease activity according to the Ankylosing Spondylitis Disease Activity Score is associated with accelerated radiographic spinal progression in patients with early axial spondyloarthritis: results from the GERman SPondyloarthritis Inception Cohort. *Ann Rheum Dis.* 2016;75(12):2114-8.
39. Machado P, van der Heijde D. How to measure disease activity in axial spondyloarthritis? *Curr Opin Rheumatol.* 2011;23(4):339-45.
40. Pedersen SJ, Sørensen IJ, Hermann KG, Madsen OR, Tvede N, Hansen MS, et al. Responsiveness of the Ankylosing Spondylitis Disease Activity Score (ASDAS) and clinical and MRI measures of disease activity in a 1-year follow-up study of patients with axial spondyloarthritis treated with tumour necrosis factor alpha inhibitors. *Ann Rheum Dis.* 2010;69(6):1065-71.
41. van der Heijde D, Braun J, Dougados M, Sieper J, Pedersen R, Szumski A, et al. Sensitivity and discriminatory ability of the Ankylosing Spondylitis Disease Activity Score in patients treated with etanercept or sulphasalazine in the ASCEND trial. *Rheumatology (Oxford).* 2012;51(10):1894-905.
42. Song I-H, Weiß A, Hermann K-GA, Haibel H, Althoff CE, Poddubnyy D, et al. Similar response rates in patients with ankylosing spondylitis and non-radiographic axial

- spondyloarthritis after 1 year of treatment with etanercept: results from the ESTHER trial. *Annals of the rheumatic diseases*. 2013;72(6):823-5.
43. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol*. 1994;21(12):2286-91.
 44. Zochling J, Sieper J, van der Heijde D, Braun J. Development of a core set of domains for data collection in cohorts of patients with ankylosing spondylitis receiving anti-tumor necrosis factor-alpha therapy. *J Rheumatol*. 2008;35(6):1079-82.
 45. Rahman P, Jovaisas A, Bensen W, Olszynski W, Jaroszynska A, Baer P, et al. 70th Annual Meeting of the Canadian Rheumatology Association. Abstract 79 - Is the BASDAI Score Driven by Pain in Ankylosing Spondylitis Patients Treated with Anti-TNF? *J Rheumatol*. 2015;42:1263-351.
 46. Martindale J, Smith J, Sutton CJ, Grennan D, Goodacre L, Goodacre JA. Disease and psychological status in ankylosing spondylitis. *Rheumatology (Oxford)*. 2006;45(10):1288-93.
 47. Brionez TF, Assassi S, Reveille JD, Green C, Learch T, Diekman L, et al. Psychological correlates of self-reported disease activity in ankylosing spondylitis. *J Rheumatol*. 2010;37(4):829-34.
 48. Durmus D, Sarisoy G, Alayli G, Kesmen H, Çetin E, Bilgici A, et al. Psychiatric symptoms in ankylosing spondylitis: their relationship with disease activity, functional capacity, pain and fatigue. *Compr Psychiatry*. 2015;62:170-7.
 49. Huang JC, Qian BP, Qiu Y, Wang B, Yu Y, Zhu ZZ, et al. Quality of life and correlation with clinical and radiographic variables in patients with ankylosing spondylitis: a retrospective case series study. *BMC Musculoskelet Disord*. 2017;18(1):352.
 50. Ozdemir O. Quality of life in patients with ankylosing spondylitis: relationships with spinal mobility, disease activity and functional status. *Rheumatol Int*. 2011;31(5):605-10.
 51. Karatepe AG, Akkoc Y, Akar S, Kirazli Y, Akkoc N. The Turkish versions of the Bath Ankylosing Spondylitis and Dougados Functional Indices: reliability and validity. *Rheumatol Int*. 2005;25(8):612-8.
 52. Waldner A, Cronstedt H, Stenström CH. The Swedish version of the Bath ankylosing spondylitis disease activity index. Reliability and validity. *Scand J Rheumatol Suppl*. 1999;111:10-6.
 53. Zhao LK, Liao ZT, Li CH, Li TW, Wu J, Lin Q, et al. Evaluation of quality of life using ASQoL questionnaire in patients with ankylosing spondylitis in a Chinese population. *Rheumatol Int*. 2007;27(7):605-11.
 54. Hoepken B, Serrano D, Harris K, Hwang MC, Reveille J. Validation of the Ankylosing Spondylitis Quality of Life assessment tool in patients with non-radiographic axial spondyloarthritis. *Quality of Life Research*. 2021;30(3):945-54.
 55. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Ixekizumab (neues Anwendungsgebiet: Psoriasis-Arthritis) 2018. Verfügbar unter: https://www.g-ba.de/downloads/40-268-5189/2018-08-16_AM-RL-XII_Ixekizumab_nAWG_D-343_TrG-15-10-2018.pdf. [Zugriff am: 10.05.2022]
 56. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie

- (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Tofacitinib (neues Anwendungsgebiet: Psoriasis-Arthritis) 2019. Verfügbar unter: https://www.g-ba.de/downloads/40-268-5588/2019-02-21_AM-RL-XII_Tofacitinib_D-373_TrG.pdf. [Zugriff am: 10.05.2022]
57. Calin A, Nakache JP, Gueguen A, Zeidler H, Mielants H, Dougados M. Defining disease activity in ankylosing spondylitis: is a combination of variables (Bath Ankylosing Spondylitis Disease Activity Index) an appropriate instrument? *Rheumatology (Oxford)*. 1999;38(9):878-82.
 58. Haywood KL, A MG, Jordan K, Dziedzic K, Dawes PT. Disease-specific, patient-assessed measures of health outcome in ankylosing spondylitis: reliability, validity and responsiveness. *Rheumatology (Oxford)*. 2002;41(11):1295-302.
 59. Kwan YH, Tan JJ, Phang JK, Fong W, Lim KK, Koh HL, et al. Validity and reliability of the Ankylosing Spondylitis Disease Activity Score with C-reactive protein (ASDAS-CRP) and Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) in patients with axial spondyloarthritis (axSpA) in Singapore. *Int J Rheum Dis*. 2019;22(12):2206-12.
 60. Brandt J, Westhoff G, Rudwaleit M, Listing J, Zink A, Braun J, et al. Adaption and validation of the bath ankylosing spondylitis disease activity index (BASDAI) for use in Germany. *Zeitschrift für Rheumatologie*. 2003;62(3):264-73.
 61. Bönisch A, Ehlebracht-König I. [The BASDAI-D--an instrument to defining disease status in ankylosing spondylitis and related diseases]. *Z Rheumatol*. 2003;62(3):251-63. Der BASDAI-D--ein Fragebogen zur Erfassung der Krankheitsaktivität bei Spondylitis ankylosans und verwandten Erkrankungen.
 62. van Tubergen A, Black PM, Coteur G. Are patient-reported outcome instruments for ankylosing spondylitis fit for purpose for the axial spondyloarthritis patient? A qualitative and psychometric analysis. *Rheumatology*. 2015;54(10):1842-51.
 63. van der Heijde D, Sieper J, Maksymowych WP, Dougados M, Burgos-Vargas R, Landewé R, et al. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. *Ann Rheum Dis*. 2011;70(6):905-8.
 64. van Tubergen A, Debats I, Ryser L, Londoño J, Burgos-Vargas R, Cardiel MH, et al. Use of a numerical rating scale as an answer modality in ankylosing spondylitis-specific questionnaires. *Arthritis Rheum*. 2002;47(3):242-8.
 65. Magrey M, Ritchlin C. Measuring outcomes in ankylosing spondylitis: pearls and pitfalls. *Curr Opin Rheumatol*. 2019;31(2):109-17.
 66. Machado PM, Raychaudhuri SP. Disease activity measurements and monitoring in psoriatic arthritis and axial spondyloarthritis. *Best Pract Res Clin Rheumatol*. 2014;28(5):711-28.
 67. Braun J, Sieper J. Ankylosing spondylitis. *Lancet*. 2007;369(9570):1379-90.
 68. Rojas-Vargas M, Munoz-Gomariz E, Escudero A, Font P, Zarco P, Almodovar R, et al. First signs and symptoms of spondyloarthritis--data from an inception cohort with a disease course of two years or less (REGISPONSER-Early). *Rheumatology (Oxford)*. 2009;48(4):404-9.
 69. Garrido-Cumbrera M, Poddubnyy D, Gossec L, Galvez-Ruiz D, Bundy C, Mahapatra R, et al. The European Map of Axial Spondyloarthritis: Capturing the Patient Perspective-an Analysis of 2846 Patients Across 13 Countries. *Curr Rheumatol Rep*. 2019;21(5):19.

70. Sieper J, van der Heijde D, Landewe R, Brandt J, Burgos-Vagas R, Collantes-Estevez E, et al. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). *Annals of the rheumatic diseases*. 2009;68(6):784-8.
71. Gemeinsamer Bundesausschuss (G-BA). Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO - Wirkstoff: Volanesorsen. 2019. Verfügbar unter: https://www.g-ba.de/downloads/92-975-3197/2019-08-15_Nutzenbewertung-G-BA_Volanesorsen_D-469.pdf. [Zugriff am: 10.05.2022]
72. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol*. 1994;21(12):2281-5.
73. Ruof J, Sangha O, Stucki G. [Evaluation of a German version of the Bath Ankylosing Spondylitis Functional Index (BASFI) and Dougados Functional Index (D-FI)]. *Z Rheumatol*. 1999;58(4):218-25. Evaluation einer deutschen Version des Bath Ankylosing Spondylitis Functional Index (BASFI) und Dougados Functional Index (D-FI).
74. Roussou E, Sultana S. The Bath Ankylosing Spondylitis Activity and Function Indices (BASDAI and BASFI) and their correlation with main symptoms experienced by patients with spondyloarthritis. *Clin Rheumatol*. 2010;29(8):869-74.
75. Haibel H, Niewerth M, Brandt J, Rudwaleit M, Listing J, Sieper J, et al. [Measurement of quality of life in patients with active ankylosing spondylitis being treated with infliximab-a comparison of SF-36 and SF-12]. *Z Rheumatol*. 2004;63(5):393-401. Messung der Lebensqualität unter Therapie mit Infliximab bei Patienten mit aktiver ankylosierender Spondylitis. Ein Vergleich von SF-36 und SF-12.
76. Jones SD, Calin A, Steiner A. An update on the Bath Ankylosing Spondylitis Disease Activity and Functional Indices (BASDAI, BASFI): excellent Cronbach's alpha scores. *J Rheumatol*. 1996;23(2):407.
77. Madsen OR, Rytter A, Hansen LB, Suetta C, Egsmose C. Reproducibility of the Bath Ankylosing Spondylitis Indices of disease activity (BASDAI), functional status (BASFI) and overall well-being (BAS-G) in anti-tumour necrosis factor-treated spondyloarthropathy patients. *Clin Rheumatol*. 2010;29(8):849-54.
78. Akgul O, Bodur H, Ataman S, Yurdakul FG, Capkin E, Gurer G, et al. Clinical performance of ASAS Health Index in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis: real-world evidence from Multicenter Nationwide Registry. *Rheumatology International*. 2020;40(11):1793-801.
79. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Østergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. *J Rheumatol*. 2007;34(4):862-70.
80. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. *Arthritis Rheum*. 2005;53(4):502-9.
81. Maksymowych WP, Lambert RG, Brown LS, Pangan AL. Defining the Minimally Important Change for the SpondyloArthritis Research Consortium of Canada Spine and Sacroiliac Joint Magnetic Resonance Imaging Indices for Ankylosing Spondylitis. *J Rheumatol*. 2012;39(8):1666-74.

82. Mease PJ, Van den Bosch F, Sieper J, Xia Y, Pangan AL, Song IH. Performance of 3 Enthesitis Indices in Patients with Peripheral Spondyloarthritis During Treatment with Adalimumab. *J Rheumatol.* 2017;44(5):599-608.
83. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewé R, van der Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. *Ann Rheum Dis.* 2003;62(2):127-32.
84. Mease PJ, Liu M, Rebello S, Hua W, McLean RR, Yi E, et al. Characterization of patients with axial spondyloarthritis by enthesitis presence: data from the corona psoriatic arthritis/spondyloarthritis registry. *ACR open rheumatology.* 2020;2(7):449-56.
85. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). *Ann Rheum Dis.* 2013;72(6):815-22.
86. de Winter JJ, van Mens LJ, van der Heijde D, Landewe R, Baeten DL. Prevalence of peripheral and extra-articular disease in ankylosing spondylitis versus non-radiographic axial spondyloarthritis: a meta-analysis. *Arthritis Res Ther.* 2016;18:196.
87. Sieper J, Poddubnyy D. Axial spondyloarthritis. *Lancet.* 2017;390(10089):73-84.
88. Schett G, Lories RJ, D'Agostino MA, Elewaut D, Kirkham B, Soriano ER, et al. Enthesitis: from pathophysiology to treatment. *Nat Rev Rheumatol.* 2017;13(12):731-41.
89. Bodur H, Ataman S, Rezvani A, Buğdaycı DS, Cevik R, Birtane M, et al. Quality of life and related variables in patients with ankylosing spondylitis. *Qual Life Res.* 2011;20(4):543-9.
90. Hamdi W, Chelli-Bouaziz M, Ahmed MS, Ghannouchi MM, Kaffel D, Ladeb MF, et al. Correlations among clinical, radiographic, and sonographic scores for enthesitis in ankylosing spondylitis. *Joint Bone Spine.* 2011;78(3):270-4.
91. Rezvani A, Bodur H, Ataman S, Kaya T, Buğdaycı DS, Demir SE, et al. Correlations among enthesitis, clinical, radiographic and quality of life parameters in patients with ankylosing spondylitis. *Mod Rheumatol.* 2014;24(4):651-6.
92. Palominos PE, de Campos APB, Ribeiro SLE, Xavier RM, Xavier JW, de Oliveira FB, et al. Correlation of enthesitis indices with disease activity and function in axial and peripheral spondyloarthritis: a cross-sectional study comparing MASES, SPARCC and LEI. *Adv Rheumatol.* 2019;59(1):23.
93. Bedaiwi M, Sari I, Thavaneswaran A, Ayearst R, Haroon N, Inman RD. Fatigue in ankylosing spondylitis and nonradiographic axial spondyloarthritis: analysis from a longitudinal observation cohort. *The Journal of rheumatology.* 2015;42(12):2354-60.
94. Bianchi WA, Elias FR, Carneiro S, Bortoluzzo AB, Gonçalves CR, da Silva JAB, et al. Assessment of fatigue in a large series of 1492 Brazilian patients with Spondyloarthritis. *Modern Rheumatology.* 2014;24(6):980-4.
95. Pearson NA, Packham JC, Tutton E, Parsons H, Haywood KL. Assessing fatigue in adults with axial spondyloarthritis: a systematic review of the quality and acceptability of patient-reported outcome measures. *Rheumatology advances in practice.* 2018;2(2):rky017.
96. Revicki DA, Rentz AM, Luo MP, Wong RL. Psychometric characteristics of the short form 36 health survey and functional assessment of chronic illness Therapy-Fatigue subscale for patients with ankylosing spondylitis. *Health and Quality of Life Outcomes.* 2011;9(1):1-11.

97. Chandran V, Bhella S, Schentag C, Gladman DD. Functional assessment of chronic illness therapy-fatigue scale is valid in patients with psoriatic arthritis. *Annals of the rheumatic diseases*. 2007;66(7):936-9.
98. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *Journal of pain and symptom management*. 1997;13(2):63-74.
99. Pilgaard T, Hagelund L, Stallknecht SE, Jensen HH, Esbensen BA. Severity of fatigue in people with rheumatoid arthritis, psoriatic arthritis and spondyloarthritis—Results of a cross-sectional study. *PloS one*. 2019;14(6):e0218831.
100. Strand V, Singh JA. Patient burden of axial spondyloarthritis. *Journal of Clinical Rheumatology*. 2017;23(7):383.
101. Sieper J, Braun J, Dougados M, Baeten D. Axial spondyloarthritis. *Nat Rev Dis Primers*. 2015;1:15013.
102. Doward LC, Spoorenberg A, Cook SA, Whalley D, Helliwell PS, Kay LJ, et al. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. *Ann Rheum Dis*. 2003;62(1):20-6.
103. Pham T, van der Heijde DM, Pouchot J, Guillemin F. Development and validation of the French ASQoL questionnaire. *Clin Exp Rheumatol*. 2010;28(3):379-85.
104. Doward LC, McKenna SP, Meads DM, Twiss J, Revicki D, Wong RL, et al. Translation and validation of non-English versions of the Ankylosing Spondylitis Quality of Life (ASQOL) questionnaire. *Health Qual Life Outcomes*. 2007;5:7.
105. Duruöz MT, Doward L, Turan Y, Cerrahoglu L, Yurtkuran M, Calis M, et al. Translation and validation of the Turkish version of the Ankylosing Spondylitis Quality of Life (ASQOL) questionnaire. *Rheumatol Int*. 2013;33(11):2717-22.
106. Jenks K, Treharne GJ, Garcia J, Stebbings S. The ankylosing spondylitis quality of life questionnaire: validation in a New Zealand cohort. *Int J Rheum Dis*. 2010;13(4):361-6.
107. Hwang MC, Martin M, Harris K, Geerds P, Stark JL, Reveille J. Content validity of the ASQoL for use in a non-radiographic axial spondyloarthritis population: a qualitative study. *Quality of Life Research*. 2020;29(11):3155-66.
108. AbbVie Inc. Clinical Study Report - SELECT-AXIS 2 (M19-944) - A Phase 3 Randomized, Placebo-Controlled, Double-Blind Program to Evaluate Efficacy and Safety of Upadacitinib in Adult Subjects with Axial Spondyloarthritis Followed by a Remission-Withdrawal Period. 2021.
109. ClinicalTrials.gov. NCT04169373 - Titel: A Phase 3 Randomized, Placebo-Controlled, Double-Blind Program to Evaluate Efficacy and Safety of Upadacitinib in Adult Subjects With Axial Spondyloarthritis Followed by a Remission-Withdrawal Period 2022. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/record/NCT04169373>. [Zugriff am: 18.03.2022]
110. EU-CTR. 2019-003229-12 - Titel: A Phase 3 Randomized, Placebo-Controlled, Double-Blind Program to Evaluate Efficacy and Safety of Upadacitinib in Adult Subjects with Axial Spondyloarthritis. 0000. Verfügbar unter: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2019-003229-12/DE>. [Zugriff am: 18.03.2022]
111. ICTRP. NCT04169373 - Titel: A Study to Evaluate Efficacy and Safety of Upadacitinib in Adult Participants With Axial Spondyloarthritis SELECT AXIS 2. 2022. Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT04169373>. [Zugriff am: 18.03.2022]

112. Syrbe U. Neues zur Pathophysiologie der Spondyloarthritis. *Akt Rheumatol.* 2019;44:315-20.
113. Burmester G, Cohen S, Winthrop K, Nash P, Rubbert-Roth A, Deodhar A, et al., editors. Long-Term Safety Profile of Upadacitinib in Patients with Rheumatoid Arthritis, Psoriatic Arthritis, or Ankylosing Spondylitis [abstract]2021. *Arthritis Rheumatol.* 2021.
114. Colombel JF, Panaccione R, Nakase H, Burmester G, Cohen SB, Mease P, et al., editors. P573 The safety profile of upadacitinib maintenance therapy in ulcerative colitis in the Phase 3 U-ACHIEVE study is consistent with that in approved indications. *Journal of Crohn's and Colitis*; 2022.
115. Simpson EL, Papp KA, Blauvelt A, Chu CY, Hong HC, Katoh N, et al. Efficacy and Safety of Upadacitinib in Patients With Moderate to Severe Atopic Dermatitis: Analysis of Follow-up Data From the Measure Up 1 and Measure Up 2 Randomized Clinical Trials. *JAMA Dermatol.* 2022;158(4):404-13.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ²¹] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

²¹ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	Embase Classic+Embase	
Suchoberfläche	Ovid	
Datum der Suche	05.07.2022	
Zeitsegment	1947 bis 01. Juli 2022	
Suchfilter	nach Wong et al. 2006	
#	Suchbegriffe	Ergebnis
1	exp upadacitinib/	1.256
2	(upadacitinib or ABT494 or ABT-494 or ABT 494 or rinvoq).mp.	1.304
3	1310726-60-3.rn.	1.182
4	1 or 2 or 3	1.304
5	exp spondylarthritis/ or exp spondyloarthritis/ or exp axial spondyloarthritis/	10.652
6	(spondylit* or spondyloarthr* or spondylarthr* or axspa or ax-spa).mp.	64.815
7	5 or 6	64.815
8	((without adj radio*) or not-radio* or nonradio* or non-radio* or xray\$1 or x-ray\$1 or nr-axspa or nraxpa or (recent-onset or early-onset)).mp.	885.691
9	7 and 8	5.977
10	((nr adj spondyl*) or (early adj axial) or nraxspa or nr-axspa or (nr adj axial)).mp.	1.320
11	((non-radio* or nonradio* or not-radio*) and (axial or axspa or ax-spa or spondyl*)).mp.	1.609
12	9 or 10 or 11	6.302
13	4 and 12	8
14	(random* or double-blind*).tw. or placebo*.mp.	2.099.212
15	13 and 14	5

Datenbankname	Ovid MEDLINE® and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions®	
Suchoberfläche	Ovid	
Datum der Suche	05.07.2022	
Zeitsegment	1946 bis 01. Juli 2022	
Suchfilter	nach Wong et al. 2006	
#	Suchbegriffe	Ergebnis
1	(upadacitinib or ABT494 or ABT-494 or ABT 494 or rinvoq or 1310726-60-3).mp.	354
2	Exp spondylarthritis/	29.389
3	(spondylit* or spondyloarthr* or spondylarthr* or axspa or ax-spa).mp.	31.733
4	2 or 3	40.577
5	((without adj radio*) or not-radio* or nonradio* or non-radio* or xray\$1 or x-ray\$1 or nr-axspa or nraxpa or (recent-onset or early-onset)).mp.	975.561
6	4 and 5	2.995
7	((nr adj spondyl*) or (early adj axial) or nraxspa or nr-axspa or (nr adj axial)).mp.	426
8	((non-radio* or nonradio* or not-radio*) and (axial or axspa or ax-spa or spondyl*)).mp.	571
9	6 or 7 or 8	3.178
10	1 and 9	2
11	randomi#ed controlled trial.pt. or (randomi#ed or placebo).mp.	1.072.237
12	10 and 11	2

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	05.07.2022	
Zeitsegment	1991 bis Mai 2022	
Suchfilter	Kein Suchfilter verwendet	
#	Suchbegriffe	Ergebnis
1	(upadacitinib or ABT494 or ABT-494 or ABT 494 or rinvoq or 1310726-60-3).mp.	506
2	exp spondylarthritis/	1.490
3	(spondylit* or spondyloarthr* or spondylarthr* or axspa or ax-spa).mp.	3.376
4	2 or 3	3.879
5	((without adj radio* or not-radio* or nonradio* or non-radio* or xray\$1 or x-ray\$1 or nr-axspa or nraxpa or (recent-onset or early-onset)).mp.	27.300
6	4 and 5	548
7	((nr adj spondyl*) or (early adj axial) or nraxspa or nr-axspa or (nr adj axial)).mp.	300
8	((non-radio* or nonradio* or not-radio*) and (axial or axspa or ax-spa or spondyl*)).mp.	345
9	6 or 7 or 8	578
10	1 and 9	5

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	<i>ClinicalTrials.gov</i> (CT.gov)
Internetadresse	https://clinicaltrials.gov/ct2/search/advanced
Datum der Suche	05.07.2022
Suchstrategie	(axial-spondyloarthritis OR axial-spondyloarthropathies OR axial-spondyloarthritis OR axial-spondylarthritis OR early-axial OR nr-axspa OR active-axial OR axial-Spondylitis OR early-ankylosing OR non-radiographic OR nonradiographic) AND (Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq)
Treffer	3

Studienregister	<i>EU Clinical Trials Register</i> (EUCTR)
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	05.07.2022
Suchstrategie	(early-axspa OR early-axial OR nr-axspa OR nonradiographic OR non-radiographic OR axial-Spondyloarthritis OR axial-Spondylarthritis OR active-axial OR early-ankylosing OR nraxspa OR axial-spondyloarthritis OR axial-Spondylitis) AND (Upadacitinib OR “ABT 494” OR Rinvoq OR 1310726-60-3)
Treffer	2

Studienregister	<i>WHO International Clinical Trials Registry Platform Search Portal (ICTRP)</i>
Internetadresse	http://apps.who.int/trialsearch
Datum der Suche	05.07.2022
Suchstrategie	nr-axspa OR non-radiographic OR nonradiographic OR early-axial OR early-axspa OR early-ankylosing OR active-axial OR axial-spondyloarthropathy OR axial-spondyloarthritis OR axial-spondylarthritis OR axial-Spondylitis [Condition] Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq OR 1310726-60-3 [Intervention]
Treffer	3

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der /den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Es wurden im Rahmen der bibliografischen Literaturrecherche keine Volltexte gesichtet. Der Anhang entfällt daher.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträge auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Studien-ID	Titel	Ausschlussgrund
Clinicaltrials.gov			
(1)	NCT04169373	A Study to Evaluate Efficacy and Safety of Upadacitinib in Adult Participants With Axial Spondyloarthritis. ClinicalTrials.gov. 2021. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04169373	A3 Vergleichstherapie
(2)	NCT04846244	A Study of the Change in Early and Sustained Pain Control in Radiographic Axial Spondylarthritis in Adult Participants Receiving Upadacitinib. ClinicalTrials.gov. 2022. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04846244	A1 Population
(3)	NCT05094128	A Study to Assess Disease Activity in Adult Participants With Ankylosing Spondylitis Who Receive Upadacitinib in a Real-world Setting. ClinicalTrials.gov. 2022. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT05094128	A1 Population
EU-CTR			
(4)	2009-010643-14	A Multicenter Study of the Efficacy and Safety of the Human Anti-TNF Monoclonal Antibody Adalimumab in Subjects with Axial Spondyloarthritis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-010643-14	A2 Intervention
(5)	2019-003229-12	A Phase 3 Randomized, Placebo-Controlled, Double-Blind Program to Evaluate Efficacy and Safety of Upadacitinib in Adult Subjects with Axial Spondyloarthritis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-003229-12	A3 Vergleichstherapie
ICTRP			
(6)	EUCTR2019-003229-12-SK	Evaluation of upadacitinib in adult subjects with axial spondyloarthritis. ICTRP. 2021. [Zugriffsdatum: 24.05.2022]. Verfügbar unter:	A3 Vergleichstherapie

Nr.	Studien-ID	Titel	Ausschlussgrund
		http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUC-TR2019-003229-12-SK	
(7)	NCT04169373	A Study to Evaluate Efficacy and Safety of Upadacitinib in Adult Participants With Axial Spondyloarthritis. ICTRP. 2021. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04169373	A3 Vergleichstherapie
(8)	NCT05094128	A Study to Assess Disease Activity in Adult Participants With Ankylosing Spondylitis Who Receive Upadacitinib in a Real-world Setting. ICTRP. 2022. [Zugriffsdatum: 24.05.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT05094128	A1 Population

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-70 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-70 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-70 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung>

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genauere Ziele, Fragestellung und Hypothesen	
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	
4b	Studienorganisation und Ort der Studiendurchführung	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	

Item ^a	Charakteristikum	Studieninformation
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

Nicht zutreffend.

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Da keine direkt vergleichenden Studien mit dem zu bewertenden Arzneimittel gegenüber einer der benannten ZVT identifiziert wurden, wird auf die Einschätzung des Verzerrungspotenzials verzichtet.

Tabelle 4-71 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie <Studienbezeichnung>

Studie: _____

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: _____

1. Verblindung der Endpunkterheber ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:_____
_____**2. Adäquate Umsetzung des ITT-Prinzips** ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:_____
_____**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:_____
_____**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:_____
_____**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig hoch

Begründung für die Einstufung:

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

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

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

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:
