

Xanthinoxidasehemmer

Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie
der Gicht

A horizontal bar composed of 18 colored segments in various shades of blue and grey. A dark blue segment in the middle contains the text 'EVIDENZBERICHT' in white, uppercase letters.

EVIDENZBERICHT

Projekt: V22-06C

Version: 1.0

Stand: 16.08.2023

IQWiG-Berichte – Nr. 1618

DOI: 10.60584/V22-06C

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Xanthinoxidasehemmer – Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht

Auftraggeber

Bundesministerium für Gesundheit

Datum des Auftrags

20.12.2022

Interne Projektnummer

V22-06C

DOI-URL

<https://dx.doi.org/10.60584/V22-06C>

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Xanthinoxidasehemmer; Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht [online]. 2023 [Zugriff: TT.MM.JJJJ]. URL: <https://dx.doi.org/10.60584/V22-06C>.

Schlagwörter

Allopurinol, Febuxostat, Gicht, Systematische Übersicht

Keywords

Allopurinol, Febuxostat, Gout, Systematic Review

Dieser Bericht wurde ohne die Beteiligung externer Sachverständiger erstellt.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Anna Margraf
- David Endres
- Tatjana Hermanns
- Claudia Mischke
- Wiebke Sieben
- Sibylle Sturtz
- Andrea Tasar

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vi
Abbildungsverzeichnis	viii
Abkürzungsverzeichnis.....	ix
1 Hintergrund.....	1
2 Fragestellung.....	2
3 Projektverlauf.....	3
3.1 Zeitlicher Verlauf des Projekts	3
3.2 Spezifizierungen und Änderungen im Projektverlauf.....	3
4 Methodik gemäß Projektskizze	5
4.1 Tabellarische Darstellung der Kriterien für den Studieneinschluss	5
4.1.1 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf	6
4.1.2 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	6
4.2 Informationsbeschaffung.....	6
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten	6
4.2.2 Fokussierte Informationsbeschaffung von Studien	7
4.2.3 Anwendung von Limitierungen auf Datenbankebene	8
4.2.4 Selektion relevanter Studien	8
4.3 Informationsdarstellung und Synthese.....	9
4.3.1 Darstellung der Einzelstudien.....	9
4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial).....	10
4.3.3 Metaanalysen	12
4.3.4 Subgruppenmerkmale und andere Effektmodifikatoren.....	12
4.3.5 Bewertung der Qualität der Evidenz	13
4.3.5.1 Abwertung der Qualität der Evidenz	14
4.3.5.2 Aufwertung der Qualität der Evidenz	15
5 Ergebnisse	17
5.1 Ergebnisse der Informationsbeschaffung	17
5.2 Charakteristika der für die Evidenzdarstellung berücksichtigten Studien	19
5.3 Übersicht der Endpunkte	31
5.4 Ergebnisse zu Endpunkten	33
5.4.1 Ergebnisse zum kritischen Endpunkt Auftreten akuter Gichtanfälle.....	33

5.4.2	Ergebnisse zum kritischen Endpunkt Serumharnsäurekonzentration.....	37
5.4.3	Ergebnisse zum kritischen Endpunkt unerwünschte Ereignisse	44
5.4.4	Ergebnisse zum wichtigen Endpunkt Auftreten von Tophi.....	51
5.4.5	Ergebnisse zum wichtigen Endpunkt Tophusgröße	53
6	Zusammenfassung und Anmerkungen zu den Ergebnissen	55
6.1	Zusammenfassung der Ergebnisse	55
6.2	Anmerkungen zu den Ergebnissen	55
7	Literatur	57
8	Studienlisten	61
8.1	Liste der identifizierten systematischen Übersichten.....	61
8.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	61
Anhang A	Kriterien des Verzerrungspotenzials.....	66
Anhang B	Suchstrategien	71
B.1	Bibliografische Datenbanken	71
B.2	Studienregister	73
B.3	Weitere Informationsquellen und Suchtechniken	74

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	5
Tabelle 2: Studienpool der eingeschlossenen Studien	19
Tabelle 3: Charakteristika der für die Evidenzdarstellung berücksichtigten Studien	20
Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien.....	21
Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien.....	24
Tabelle 6: Charakterisierung der Studienpopulationen.....	29
Tabelle 7: Matrix der einbezogenen Endpunkte.....	32
Tabelle 8: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Auftreten akuter Gichtanfälle (binäre Analyse).....	33
Tabelle 9: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Auftreten akuter Gichtanfälle (binäre Analyse).....	34
Tabelle 10: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (binäre Analyse).....	37
Tabelle 11: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (binäre Analyse).....	38
Tabelle 12: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (stetige Analyse)	41
Tabelle 13: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (stetige Analyse)	42
Tabelle 14: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt unerwünschte Ereignisse (binäre Analyse)	44
Tabelle 15: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt unerwünschte Ereignisse (binäre Analyse)	47
Tabelle 16: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den wichtigen Endpunkt Auftreten von Tophi (stetige Analyse)	51
Tabelle 17: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den wichtigen Endpunkt Auftreten von Tophi (stetige Analyse)	52
Tabelle 18: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den wichtigen Endpunkt Tophusgröße (stetige Analyse)	53
Tabelle 19: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den wichtigen Endpunkt Tophusgröße (stetige Analyse)	54
Tabelle 20: Kriterien des Verzerrungspotenzials zum Endpunkt Auftreten akuter Gichtanfälle.....	66
Tabelle 21: Kriterien des Verzerrungspotenzials zum Endpunkt Serumharnsäurekonzentration	67
Tabelle 22: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse – Studienabbruch wegen UE	68

Tabelle 23: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse
– SUE 69

Tabelle 24: Kriterien des Verzerrungspotenzials zum Endpunkt Auftreten von Tophi 69

Tabelle 25: Kriterien des Verzerrungspotenzials zum Endpunkt Tophusgröße..... 70

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung	18
Abbildung 2: Metaanalyse, Febuxostat versus Placebo, Auftreten akuter Gichtanfälle, bis 80 mg/ 80 mg und 120 mg Febuxostat – Behandlungszeitraum bis 3 Monate, Effektmaß: OR.....	36
Abbildung 3: Metaanalyse, Febuxostat versus Placebo, Endpunkt Serumharnsäurekonzentration (Erreichen des Grenzwerts), bis 80 mg/ 80 mg und 120 mg Febuxostat – 6 Monate, Effektmaß: OR	40
Abbildung 4: Metaanalyse, Febuxostat versus Placebo, Endpunkt Serumharnsäurekonzentration (Prozentuale Reduktion), 80 mg und 120 mg Febuxostat – 6 Monate, Effektmaß: MWD.....	43
Abbildung 5: Metaanalyse, Allopurinol versus Placebo, Endpunkt unerwünschte Ereignisse (Abbruch aufgrund von UE) – Studienende, Effektmaß: OR	45
Abbildung 6: Metaanalyse, Allopurinol vs. Placebo, Endpunkt unerwünschte Ereignisse (SUE) – Studienende, Effektmaß: OR.....	46
Abbildung 7: Metaanalyse, Febuxostat vs. Placebo, Endpunkt unerwünschte Ereignisse (Abbruch aufgrund von UE), bis 120 mg Febuxostat – Studienende, Effektmaß: OR.....	49
Abbildung 8: Metaanalyse, Febuxostat vs. Placebo, Endpunkt unerwünschte Ereignisse (SUE), bis 120 mg Febuxostat – Studienende, Effektmaß: OR	50

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
BMG	Bundesministerium für Gesundheit
DGRh	Deutsche Gesellschaft für Rheumatologie e. V.
DVG	Digitale-Versorgung-Gesetz
G-BA	Gemeinsamer Bundesausschuss
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KI	Konfidenzintervall
MWD	Mittelwertdifferenz
OR	Odds Ratio
PICO	Population, Intervention, Comparison, Outcome
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	absolute Risikodifferenz
SMD	Standardized Mean Difference (standardisierte Mittelwertdifferenz)
SÜ	systematische Übersicht
SUE	schwerwiegende unerwünschte Ereignisse
UE	unerwünschte Ereignisse

1 Hintergrund

Auf Basis des Digitale-Versorgung-Gesetzes (DVG) kann die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF) dem Bundesministerium für Gesundheit (BMG) Leitlinien vorschlagen, bei denen das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) gemäß SGB V (§§ 139a Abs. 3 Nr. 3, 139b Abs. 6) deren Entwicklung oder Aktualisierung mit Evidenzrecherchen unterstützen soll [1].

Der vorliegende Auftrag umfasst die Beantwortung von Fragestellungen zur Weiterentwicklung der interdisziplinären S2e-Leitlinie „Diagnostik und Therapie der Gicht“ (Registernummer der AWMF: 060-005) [2] zu einer S3-Leitlinie.

2 Fragestellung

Ziel des Evidenzberichts ist die Darstellung der Evidenz bezüglich der Effekte der Xanthinoxidasehemmer Allopurinol und Febuxostat jeweils im Vergleich zu Placebo bei erwachsenen Patientinnen und Patienten mit Gicht.

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Das IQWiG wurde am 20.12.2022 vom BMG beauftragt, die Deutsche Gesellschaft für Rheumatologie e. V. (DGRh) bei der Weiterentwicklung einer interdisziplinären S2e-Leitlinie „Diagnostik und Therapie der Gicht“ (Registernummer der AWMF: 060-005) [2] zu unterstützen.

Insgesamt wurden 6 Population-Intervention-Comparison-Outcome(PICO)-Fragen von der Leitliniengruppe in Abstimmung mit Patientenvertreterinnen und -vertretern und mit Beratung durch die AWMF formuliert, zu denen das IQWiG jeweils auf Basis einer mit der Leitlinienkoordination abgestimmten Projektskizze einen Evidenzbericht erstellte. Die Projektbearbeitung im IQWiG begann mit dem Kick-off-Treffen, an dem die Leitlinienkoordination und Ansprechpersonen des IQWiG teilnahmen. Das Kick-off-Treffen hat am 21.02.2023 stattgefunden. Die Evidenzberichte wurden an die Leitlinienkoordination sowie an die zuständige Ansprechperson für die Leitlinie bei der AWMF sowie an das BMG übermittelt.

Nach Abschluss aller Evidenzberichte für diesen Auftrag wurden diese zusammen an die Gremien des IQWiG und das BMG übermittelt sowie 4 Wochen später auf der Website des IQWiG veröffentlicht.

3.2 Spezifizierungen und Änderungen im Projektverlauf

Es ergaben sich folgende Spezifizierungen oder Änderungen im Evidenzbericht:

- Abschnitt 4.1: Der kritische Endpunkt „Normalisierung der Harnsäurekonzentration (< 360 µmol/l [6 mg/dl]) über mindestens 6 Monate“ wurde umbenannt in „Serumharnsäurekonzentration (vorzugsweise Serumharnsäurekonzentration < 360 µmol/l [< 6 mg/dl] über mindestens 6 Monate und / oder Reduktion der Serumharnsäurekonzentration über mindestens 6 Monate)“. Der Endpunkt „unerwünschte Ereignisse“ wurde als kritisch für die Entscheidungsfindung definiert. Die Änderungen erfolgten auf Wunsch der Leitliniengruppe.
- Abschnitt 4.1: Zur Prüfintervention wurde folgende Spezifizierung vorgenommen: Grundlage für die Evidenzdarstellung von Arzneimitteln ist i. d. R. die in Deutschland bestehende Zulassung. Es wird daher geprüft, ob die Anwendung der in den Studien eingesetzten Prüf- und Vergleichsinterventionen im Rahmen des für Deutschland gültigen Zulassungsstatus erfolgte. Voraussetzung für die Evidenzdarstellung von Arzneimitteln ist zudem die Marktverfügbarkeit in Deutschland.
- Abschnitt 4.2.2: Als weitere Informationsquelle wurden die Webseiten des IQWiG und Gemeinsamen Bundesausschusses (G-BA) durchsucht.

- Abschnitt 4.2.3: Die Selektion von systematischen Übersichten wurde auf das Publikationsdatum ab 01/2020 eingeschränkt.
- Abschnitt 4.3.2: Die Bewertung der Kriterien des Verzerrungspotenzials erfolgte mit „ja“ statt „adäquat“ bzw. mit „nein“ statt „nicht adäquat“. 2 Kriterien des Verzerrungspotenzials wurden umbenannt: von „Erzeugung der Randomisierungssequenz“ zu „Adäquate Erzeugung der Randomisierungssequenz“ und von „Umsetzung des Intention-to-treat(ITT)-Prinzips“ zu „Intention-to-treat(ITT)-Prinzip adäquat umgesetzt“.

4 Methodik gemäß Projektskizze

Die folgenden Abschnitte geben den Wortlaut der Berichtsmethodik aus der Projektskizze wieder. Über diese Methodik hinausgehende Spezifizierungen oder Änderungen der Methoden im Projektverlauf werden in Abschnitt 3.2 erläutert. Im folgenden Text wird an den entsprechenden Stellen auf diesen Abschnitt verwiesen.

Die PICO-Fragestellung wird durch die Darstellung von Evidenzprofilen beantwortet. Die Erstellung der Evidenzprofile erfolgt auf Grundlage der methodischen Vorgaben von Grading of Recommendations Assessment, Development and Evaluation (GRADE) [3] und ist konform zu den Allgemeinen Methoden [4]. Das PICO-Schema wurde von der Leitliniengruppe zur interdisziplinären S2e-Leitlinie „Diagnostik und Therapie der Gicht“ (Registernummer der AWMF: 060-005) [2] festgelegt.

4.1 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in den Studienpool eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Population: erwachsene Patientinnen und Patienten mit Gicht
E2	Prüfintervention: Behandlung mit den Xanthinoxidasehemmern Allopurinol oder Febuxostat
E3	Vergleichsintervention: Behandlung mit Placebo
E4	kritische Endpunkte: Auftreten akuter Gichtanfälle, Normalisierung der Harnsäurekonzentration (< 360 µmol/l [6 mg/dl]) über mindestens 6 Monate wichtige Endpunkte: Auftreten von Tophi, Tophusgröße, Funktionsfähigkeit, unerwünschte Ereignisse Bei mehr als 7 Endpunkten sind kritische Endpunkte den wichtigen Endpunkten vorzuziehen.
E5	Studientyp: RCTs
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar ^a
<p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [5] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [6] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie</p>	

(Zu diesem Vorgehen gab es eine Änderung/ Spezifizierung im Projektverlauf, siehe Abschnitt 3.2.)

4.1.1 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf

Sofern sich im Projektverlauf Änderungen im Zulassungsstatus der darzustellenden Interventionen ergeben, werden die Kriterien für den Studieneinschluss gegebenenfalls an die neuen Zulassungsbedingungen angepasst. Die jeweils vorgenommenen Änderungen werden im Evidenzbericht explizit vermerkt.

4.1.2 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei ca. 80 % der eingeschlossenen Patientinnen und Patienten diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen und Patienten vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien E1, E2 und E3 bei weniger als ca. 80 % erfüllt sind, können ggf. für die Evidenzdarstellung unberücksichtigt bleiben.

4.2 Informationsbeschaffung

4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung der Projektskizze erfolgte eine Recherche nach systematischen Übersichten (SÜs) in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), der International Health Technology Assessment (HTA) Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

Die Suche fand am 21.02.2023 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang B.

Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen den beiden aufgelöst.

Es wird geprüft, ob mindestens 1 hochwertige und aktuelle SÜ infrage kommt, deren Informationsbeschaffung als Grundlage für die Evidenzdarstellung verwendet werden kann (im Folgenden: Basis-SÜ). Dafür erfolgt eine Bewertung der Qualität der Informationsbeschaffung dieser SÜ(s). Kann mindestens 1 diesbezüglich hochwertige und aktuelle Basis-SÜ identifiziert werden, werden die zugrunde liegenden Studien bzw. Dokumente von 1 Person auf deren Relevanz für die vorliegende Evidenzdarstellung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der eingeschlossenen Studien oder die Datenextraktion werden nicht übernommen.

Die finale Entscheidung, ob und wenn ja welche SÜ(s) als Basis-SÜ(s) für die Evidenzdarstellung herangezogen werden, erfolgt nach Fertigstellung der Projektskizze anhand der darin

festgelegten Kriterien. In jedem Fall werden die Referenzlisten der identifizierten SÜ(s) hinsichtlich relevanter Primärstudien gesichtet (siehe Abschnitt 4.2.2).

4.2.2 Fokussierte Informationsbeschaffung von Studien

Für den Fall, dass mindestens 1 SÜ als Basis-SÜ für die Informationsbeschaffung des Evidenzberichts verwendet werden kann (siehe Abschnitt 4.2.1), wird diese für die Informationsbeschaffung von Studien für den von der Übersicht abgedeckten Zeitraum herangezogen. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien bzw. Dokumenten für den nicht von der Übersicht abgedeckten Zeitraum. Gegebenenfalls wird auf eine ergänzende fokussierte Informationsbeschaffung von Studien ganz verzichtet.

Für den Fall, dass keine Basis-SÜ identifiziert werden kann, findet eine systematische Recherche für den gesamten relevanten Zeitraum statt.

Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Cochrane Central Register of Controlled Trials
- Studienregister (es erfolgt eine Einschränkung auf Studienergebnisse)
 - U.S. National Institutes of Health. ClinicalTrials.gov
 - World Health Organization. International Clinical Trials Registry Platform Search Portal
 - European Medicines Agency. EU Clinical Trials Register

Weitere Informationsquellen und Suchtechniken

- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter SÜs
 - Identifizieren von Studienregistereinträgen zu eingeschlossenen Studien
- Autorenanfrage(n)

(Zu diesem Vorgehen gab es eine Änderung / Spezifizierung im Projektverlauf, siehe Abschnitt 3.2.)

4.2.3 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von SÜs

Die Suchen wurden auf das Publikationsdatum ab 01/2012 eingeschränkt. Die MEDLINE-Suchstrategie enthält Limitierungen auf deutsch- und englischsprachige Publikationen [4] sowie auf Humanstudien.

(Zu diesem Vorgehen gab es eine Änderung / Spezifizierung im Projektverlauf, siehe Abschnitt 3.2.)

Fokussierte Informationsbeschaffung von Studien

Hinsichtlich des Publikationszeitraums gibt es keine Einschränkung. Sollte die Informationsbeschaffung auf Grundlage einer Basis-SÜ erfolgen, wird eine entsprechende zeitliche Einschränkung in Betracht gezogen (siehe Abschnitt 4.2.2). Mit der MEDLINE-Suchstrategie werden folgende Publikationstypen ausgeschlossen: Kommentare und Editorials, da diese i. d. R. keine Studien enthalten [7]. Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen [4] sowie auf Humanstudien (MEDLINE). In der Cochrane Central Register of Controlled Trials Suche werden Einträge aus Studienregistern ausgeschlossen.

4.2.4 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Recherche

Duplikate werden mit Hilfe des Literaturverwaltungsprogrammes EndNote entfernt. Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 1) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Die Rechercheergebnisse aus den folgenden Informationsquellen werden von 1 Person auf Studien gesichtet:

- Studienregister
- Referenzlisten identifizierter SÜs

Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

4.3 Informationsdarstellung und Synthese

4.3.1 Darstellung der Einzelstudien

Alle für den Evidenzbericht notwendigen Informationen werden aus den Unterlagen zu den für die Evidenzdarstellung berücksichtigten Studien in standardisierte Tabellen extrahiert. Ergeben sich im Abgleich der Informationen aus unterschiedlichen Dokumenten zu einer Studie (aber auch aus multiplen Angaben zu einem Aspekt innerhalb eines Dokumentes selbst) Diskrepanzen, die auf die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wird dies an den entsprechenden Stellen des Berichts dargestellt.

Die Extraktion erfolgt durch 1 Person und wird von einer 2. Person auf Grundlage der Studien kontrolliert.

Die Ergebnisse zu den in den Studien berichteten kritischen und wichtigen Endpunkten werden im Bericht vergleichend dargestellt. Anhand inhaltlicher und methodischer Kriterien kann eine Auswahl der dargestellten Operationalisierungen zu den Endpunkten, für die verwertbare Daten vorliegen, erfolgen. Hierbei erfolgt die Auswahl unter Berücksichtigung der Datenlage für den gesamten Evidenzbericht. Operationalisierungen aus qualitativ guten Studien und / oder Studien mit größeren Stichproben werden bevorzugt dargestellt. Werden in der überwiegenden Zahl der Studien Ergebnisse zu einer spezifischen Operationalisierung dargestellt, kann die Darstellung im Bericht darauf begrenzt werden. Werden in den Studien Auswertungen mittels verschiedener Modelle berichtet, die nicht metaanalytisch zusammengefasst werden können, kann die Darstellung auf eine Auswertungsart beschränkt werden. Es wird geprüft, ob die Ergebnisse, die nicht dargestellt werden, zu denen der ausgewählten Operationalisierung(en) bzw. Auswertung(en) passen.

Das Vorgehen zur Bewertung des Einflusses der Studienlimitationen auf die berichtsrelevanten Ergebnisse wird in Abschnitt 4.3.2 endpunktspezifisch pro Studie beschrieben. Nach dieser Bewertung werden die Informationen zu den ausgewählten Endpunkten in Evidenzprofilen zusammengeführt und analysiert [8]. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in den Abschnitten 4.3.3 bis 4.3.4 beschriebenen Verfahren eingesetzt.

Für binäre Daten wird als relative Effektschätzung primär das Odds Ratio (OR) herangezogen. Dabei wird im Fall von 0 Ereignissen in einem Studienarm bei der Berechnung von Effekt und Konfidenzintervall (KI) der Korrekturfaktor 0,5 in beiden Studienarmen verwendet. Für absolute Effekte wird vorrangig die absolute Risikodifferenz (RD) verwendet. Das Basisrisiko

wird in der Regel durch den Median des Risikos der Vergleichsgruppe(n) der eingeschlossenen Einzelstudie(n) bestimmt. Auf Grundlage dieses angenommenen Basisrisikos werden mithilfe der relativen Gesamtschätzung der Metaanalyse das absolute Risiko in der Interventionsgruppe und die Risikodifferenz errechnet. Zur Abschätzung der quantitativen Ergebnisunsicherheit werden ergänzend die oberen und unteren KI-Grenzen der RD basierend auf den KI-Grenzen der relativen Gesamtschätzung der Metaanalyse berechnet. Wird keine Metaanalyse durchgeführt, erfolgt eine Berechnung der RD pro Studie durch die Risiken in den beiden Behandlungsgruppen; als zugehöriges KI wird standardmäßig dasjenige nach der Wilson-Score-Methode [9] angegeben. Sollte das Ergebnis des CSZ-Tests [10] qualitativ nicht zum Ergebnis des KI nach der Wilson-Score-Methode, sondern zum Wald-KI passen, wird dieses angegeben.

Für stetige Daten wird primär die Mittelwertdifferenz (MWD) herangezogen. Falls notwendig (z. B. wenn verschiedene Skalen gepoolt werden sollen oder dies zur Bewertung der Qualität der Evidenz benötigt wird), wird außerdem eine standardisierte Mittelwertdifferenz (SMD, z. B. Hedges' g) angegeben.

Ergebnisse können ggf. im Evidenzbericht unberücksichtigt bleiben, wenn ein großer Anteil der in die Auswertung eigentlich einzuschließenden Personen nicht in der Auswertung berücksichtigt worden ist. Für die Entscheidung hierüber wird sich an einem Anteil von ca. 70 % orientiert, der in der Auswertung mindestens berücksichtigt sein sollte.

Die Ergebnisse können ggf. auch dann unberücksichtigt bleiben, wenn der Unterschied der Anteile nicht berücksichtigter Personen zwischen den Gruppen sehr groß ist. Als Orientierung für die Entscheidung dient ein Unterschied von mehr als 15 Prozentpunkten.

4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial)

Der Einfluss möglicher Studienlimitationen auf die Ergebnisse wird zunächst endpunktspezifisch pro Studie insbesondere anhand der unten aufgeführten Kriterien bewertet. Dazu erfolgt jeweils eine Bewertung mit „adäquat“, „unklar“ oder „nicht adäquat“. Die Voraussetzungen für eine Bewertung mit „adäquat“ sowie wesentliche Aspekte, welche zu einer Bewertung mit „nicht adäquat“ führen können, werden im Folgenden erläutert. Eine Bewertung mit „unklar“ erfolgt grundsätzlich dann, wenn keine bzw. keine ausreichenden Angaben zur adäquaten Bewertung zur Verfügung stehen.

Erzeugung der Randomisierungssequenz

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass aus der Beschreibung der Studie hervorgeht, dass die Gruppenzuteilung rein zufällig erfolgte und die Erzeugung der Zuteilungssequenz beschrieben und geeignet ist. Die Antwortmöglichkeit „nicht adäquat“

besteht im Falle eines Evidenzberichts auf Basis von RCTs nicht, da die Studie in diesem Fall als nicht randomisierte Studie aufgefasst, ausgeschlossen und nicht weiter bewertet wird.

Verdeckung der Gruppenzuteilung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass aus der Beschreibung der Studie hervorgeht, dass die Zuteilung der Patientinnen und Patienten in die verschiedenen Studiengruppen den Personen, die die Gruppenzuteilung veranlassen oder über den Einschluss der Patientinnen und Patienten entscheiden, nicht bekannt ist. Die Antwortmöglichkeit „nicht adäquat“ besteht im Falle eines Evidenzberichts auf Basis von RCTs nicht, da die Studie in diesem Fall als nicht randomisierte Studie aufgefasst, ausgeschlossen und nicht weiter bewertet wird.

Verblindung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass alle zu verblindenden Personengruppen (Patientinnen und Patienten, behandelnde Personen, endpunkterhebende Personen, Datenanalytistinnen und Datenanalysten) jeweils explizit benannt und die Methode, die eine Verblindung gewährleistet, beschrieben und geeignet ist.

Umsetzung des Intention-to-treat(ITT)-Prinzips

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass randomisierte Patientinnen und Patienten entsprechend ihrer Gruppenzuteilung ausgewertet und der Anteil der fehlenden Werte gering ist oder sie in geeigneter Weise berücksichtigt wurden. Regelmäßig erfolgt eine Bewertung mit „adäquat“, wenn der Anteil der nicht in der Auswertung berücksichtigten Patientinnen und Patienten insgesamt kleiner als 10 % oder der Unterschied der Anteile zwischen den Gruppen kleiner als 5 Prozentpunkte ist. Wurden fehlende Werte in geeigneter Weise berücksichtigt, hängt die Bewertung des Ergebnisses von vielen Aspekten ab: u. a. der Häufigkeit fehlender Werte, der Art des Auswertungsverfahrens oder der Gründe für das Auftreten fehlender Werte.

Ergebnisunabhängige Berichterstattung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass der Endpunkt gemäß einer prospektiven Planung (in der Regel im Studienregistereintrag) operationalisiert, erfasst und ausgewertet wird oder relevante Abweichungen von den Studienautorinnen und Studienautoren plausibel erklärt werden.

Im Evidenzprofil werden abschließend die Studienlimitationen studienübergreifend und endpunktspezifisch durch die Bewertungen „nicht schwerwiegend“, „schwerwiegend“ und „sehr schwerwiegend“ dargestellt.

(Zu diesem Vorgehen gab es eine Änderung / Spezifizierung im Projektverlauf, siehe Abschnitt 3.2.)

4.3.3 Metaanalysen

Die geschätzten Effekte und KIs aus den Studien werden mittels Forest Plots dargestellt. Die Heterogenität zwischen den Studien wird mithilfe des statistischen Tests auf Vorliegen von Heterogenität [11] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten gemäß der Methode nach Knapp und Hartung. Ist das Konfidenzintervall nach Knapp-Hartung schmäler als das Konfidenzintervall nach DerSimonian-Laird, wird die Knapp-Hartung-Schätzung mit Ad-hoc-Varianzkorrektur weiter betrachtet. Ansonsten wird die Schätzung ohne Ad-hoc-Varianzkorrektur betrachtet. Die Schätzung des Heterogenitätsparameters erfolgt nach Paule-Mandel [12]. Als Ergebnis wird der gemeinsame Effekt mit Konfidenzintervall im Forest Plot dargestellt. Da die Heterogenität im Fall sehr weniger (2 bis 4) Studien nicht verlässlich geschätzt werden kann, wird im Fall von 2 Studien aus pragmatischen Gründen ein Modell mit festem Effekt verwendet, sofern keine deutlichen Gründe dagegen sprechen. Bei 3 und 4 Studien erfolgt vorzugsweise eine Metaanalyse nach der Methode von Knapp und Hartung, sofern das KI informativ ist und kein Widerspruch zu einer qualitativen Zusammenfassung besteht. Andernfalls wird ein Modell mit festem Effekt verwendet, wenn keine deutlichen Gründe dagegensprechen. Ist ein Modell mit festem Effekt nicht vertretbar, bleibt es bei der Schätzung nach Knapp und Hartung.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird untersucht, welche Faktoren eine vorhandene Heterogenität möglicherweise verursachen. Dazu zählen methodische und klinische Faktoren, sogenannte Effektmodifikatoren. Falls vorhandene Heterogenität durch solche Faktoren zumindest zum Teil erklärt werden kann, so wird der Studienpool nach diesen Faktoren aufgespaltet und die weiteren Berechnungen erfolgen in den getrennten Studienpools. Kann die Heterogenität nicht erklärt werden, so wird, sofern möglich, ebenfalls ein gemeinsamer (gepoolter) Effekt berechnet. Dabei ist zu beachten, dass ein gepoolter Effekt bei bedeutsamer Heterogenität möglicherweise nicht sinnvoll interpretierbar ist. Von der Berechnung eines gepoolten Effekts wird abgesehen, falls sich die KIs der eingehenden Studien nur gering oder gar nicht überlappen und gleichzeitig die Effektschätzungen der Studien in unterschiedliche Richtungen weisen.

4.3.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In den Evidenzbericht werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus

Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-regressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter.

Sollten sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

4.3.5 Bewertung der Qualität der Evidenz

Alle für den Evidenzbericht relevanten Ergebnisse werden hinsichtlich einer Beeinflussung durch Faktoren, die zu einer Abwertung oder ggf. Aufwertung der Qualität der Evidenz führen können, überprüft. Für jeden für die Entscheidung kritischen und wichtigen Endpunkt wird eine studienübergreifende Aussage zur Qualität der Evidenz bezüglich des jeweiligen Vertrauens in die Effektschätzung getroffen. Hierzu erfolgt eine Einteilung der Qualität der Evidenz entsprechend der 4 Stufen der GRADE-Guideline in hoch, moderat, niedrig und sehr niedrig [13,14]:

- Das Vertrauen in die Effektschätzung ist hoch: Eine hohe Qualität der Evidenz bedeutet, dass der wahre Effekt sehr sicher nahe bei der Effektschätzung liegt.
- Das Vertrauen in die Effektschätzung ist mäßig: Eine moderate Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich nahe bei der Effektschätzung liegt, aber die Möglichkeit besteht, dass er relevant verschieden ist.
- Das Vertrauen in die Effektschätzung ist begrenzt: Eine niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt durchaus relevant verschieden zur Effektschätzung sein kann.
- Das Vertrauen in die Effektschätzung ist sehr gering: Eine sehr niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich relevant verschieden von der Effektschätzung ist.

Die Bewertung erfolgt durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Ergebnissen aus RCTs wird im 1. Bewertungsschritt eine hohe, Ergebnissen aus Beobachtungsstudien eine niedrige Qualität der Evidenz attestiert [13]. Von dieser rein

designbedingten Einschätzung ausgehend wird bei Studienlimitationen, inkonsistenten Effekten, Indirektheit (eingeschränkter Übertragbarkeit) der Effekte, einem anzunehmenden Publikationsbias oder aufgrund fehlender Genauigkeit der Effekte die Qualität der Evidenz abgewertet sowie bei großen Effekten, einer Dosis-Wirkungs-Beziehung oder, wenn die Berücksichtigung aller potenziellen Confounder zu einer Effekterhöhung / Verstärkung des beobachteten Ergebnisses führen würde, die Qualität der Evidenz ggf. aufgewertet. Unterstützend für diese Bewertungsschritte können Sensitivitätsanalysen durchgeführt werden. Alle Ab- und Aufwertungen der Qualität der Evidenz werden durch Fußnoten gekennzeichnet.

4.3.5.1 Abwertung der Qualität der Evidenz

A: Studienlimitationen

Die Qualität der Evidenz wird aufgrund von schwerwiegenden oder sehr schwerwiegenden Studienlimitationen in der Regel um 1 oder 2 Stufen abgewertet, wenn starke Limitierungen in einem oder mehreren der in Abschnitt 4.3.2 genannten Kriterien das Vertrauen in die Effektschätzung beeinträchtigen. Die endpunktbezogene studienübergreifende Bewertung der Studienlimitationen erfolgt unter Berücksichtigung des Einflusses der einzelnen Studien auf die Effektschätzung für jeden Endpunkt [15].

B: Inkonsistente (heterogene) Effekte

Die Qualität der Evidenz wird bei Inkonsistenz (unerklärter Heterogenität) zwischen Studienergebnissen in der Regel um 1 oder 2 Stufen abgewertet. Die Einschätzung einer möglichen Heterogenität erfolgt anhand der in Abschnitt 4.3.3 beschriebenen Kriterien.

Falls für binäre Daten Metaanalysen für 2 Effektmaße durchgeführt wurden und eine der beiden einen statistisch signifikanten Gruppenunterschied anzeigt, wird die Inkonsistenz anhand dieser Metaanalyse bewertet. Ist der Gruppenunterschied für beide Effektmaße nicht statistisch signifikant, wird die Bewertung in der Regel anhand des relativen Effektmaßes vorgenommen.

Bei stetigen Daten wird die Bewertung in der Regel anhand der Metaanalyse der Mittelwertdifferenz vorgenommen. Falls die fehlende Genauigkeit anhand der standardisierten Mittelwertdifferenz bewertet wurde, wird die Inkonsistenz auch anhand dieser bewertet.

Eine Abwertung um 1 Stufe erfolgt in der Regel, wenn der Heterogenitätstest ein signifikantes Ergebnis liefert. Um 2 Stufen kann z. B. abgewertet werden, wenn aufgrund der Heterogenität der Studienergebnisse keine gepoolte Effektschätzung berechnet wird.

C: Indirektheit

Die Qualität der Evidenz wird bei starker oder sehr starker Indirektheit (eingeschränkter Übertragbarkeit) in der Regel um 1 oder 2 Stufen abgewertet. Indirektheit kann auf Abweichungen zwischen dem PICO und den Einschlusskriterien der Studien basieren [16].

D: Publikationsbias

Die Qualität der Evidenz wird in der Regel um 1 Stufe abgewertet, wenn ein Publikationsbias wahrscheinlich oder sehr wahrscheinlich ist.

Kriterien für die Annahme eines Publikationsbias umfassen eine Evidenzgrundlage, welche vorrangig auf kleinen Studien und wenigen Ereignissen beruht, welche einen beobachteten Zusammenhang zwischen Größe der Effektschätzung und Studiengröße (oder Genauigkeit der Effekte) zeigt, welche Studienabbrüche auf der Grundlage von hohen Effektschätzungen aufweist oder welche Auffälligkeiten in der gesichteten Datenlage erkennen lässt [17].

E: Fehlende Genauigkeit der Effekte

Die Qualität der Evidenz wird wegen geringer oder sehr geringer Genauigkeit der Effekte in der Regel um 1 oder 2 Stufen abgewertet. Maßgeblich hierfür ist neben der Lage und Breite des 95 %-KI auch die Wahl des Metaanalysemodells (zufällige Effekte oder fester Effekt) (siehe Abschnitt 4.3.3). Außerdem können sehr kleine Fallzahlen zu einer Abwertung wegen fehlender Genauigkeit führen [18]. Bei ausreichender Patientenzahl wird regelhaft nicht abgewertet, wenn das KI eines Effekts (relativ oder absolut im Fall binärer Daten bzw. MWD oder SMD im Fall stetiger Daten) einen signifikanten Unterschied anzeigt.

Andernfalls erfolgt regelhaft eine Abwertung wegen fehlender Präzision, falls das KI (des relativen Effekts bei binären Daten bzw. Hedges' g bei stetigen Daten) einen mittelgroßen Effekt und gleichzeitig auch den Nulleffekt überdeckt. Bei binären Daten wird unter einem mittelgroßen Effekt eine Halbierung oder eine Verdopplung (0,5 oder 2) des relativen Effekts verstanden. Bei stetigen Daten wird von einem mittelgroßen Effekt ausgegangen, wenn das KI für eine standardisierte Mittelwertdifferenz die Grenze 0,5 oder $-0,5$ umfasst. In diesen Fällen, in denen ein mittelgroßer Effekt und gleichzeitig auch der Nulleffekt überdeckt wird, sollte eine Abwertung wegen fehlender Genauigkeit um 1 Stufe in Betracht gezogen werden. Eine Abwertung um 2 Stufen wird in Fällen in Betracht gezogen, in denen die zuvor genannten KI-Grenzen beide, d. h. z. B. im Fall binärer Daten sowohl 0,5 als auch 2 im KI enthalten sind.

4.3.5.2 Aufwertung der Qualität der Evidenz

Eine Aufwertung kommt bei methodisch guten, nicht randomisierten vergleichenden Studien infrage, sofern insbesondere folgende Aspekte nicht dagegen sprechen: relevante Verzerrungsrisiken, plausible Confounder, die den Effekt verringern könnten, statistisch nicht signifikante Effekte oder unpräzise Effekte (mit nicht hinreichend schmalen KIs) [19].

A: Große Effekte

Die Aufwertung der Qualität der Evidenz ist in der Regel um 1 Stufe bei mittelgroßen Effekten möglich, wenn beispielsweise das relative Risiko bei binären Daten zwischen 2 und 5 bzw. zwischen 0,2 und 0,5 liegt. Gleiches gilt für stetige Daten, wenn die SMD zwischen 0,5 und 0,8 bzw. $-0,8$ und $-0,5$ liegt. Eine Aufwertung um 1 weitere Stufe kann bei großen Effekten erfolgen, wenn beispielsweise das relative Risiko bei binären Daten über 5 bzw. unter 0,2 oder bei stetigen Daten die SMD über 0,8 bzw. unter $-0,8$ liegt.

B: Dosis-Wirkungs-Beziehung

Die Qualität der Evidenz wird bei einer Dosis-Wirkungs-Beziehung regelhaft aufgewertet [19].

C: Confounding

Die Qualität der Evidenz wird bei einer möglichen Verstärkung des beobachteten Effektes unter Berücksichtigung von potenziellen Confoundern oder Verzerrungsquellen regelhaft aufgewertet [19].

5 Ergebnisse

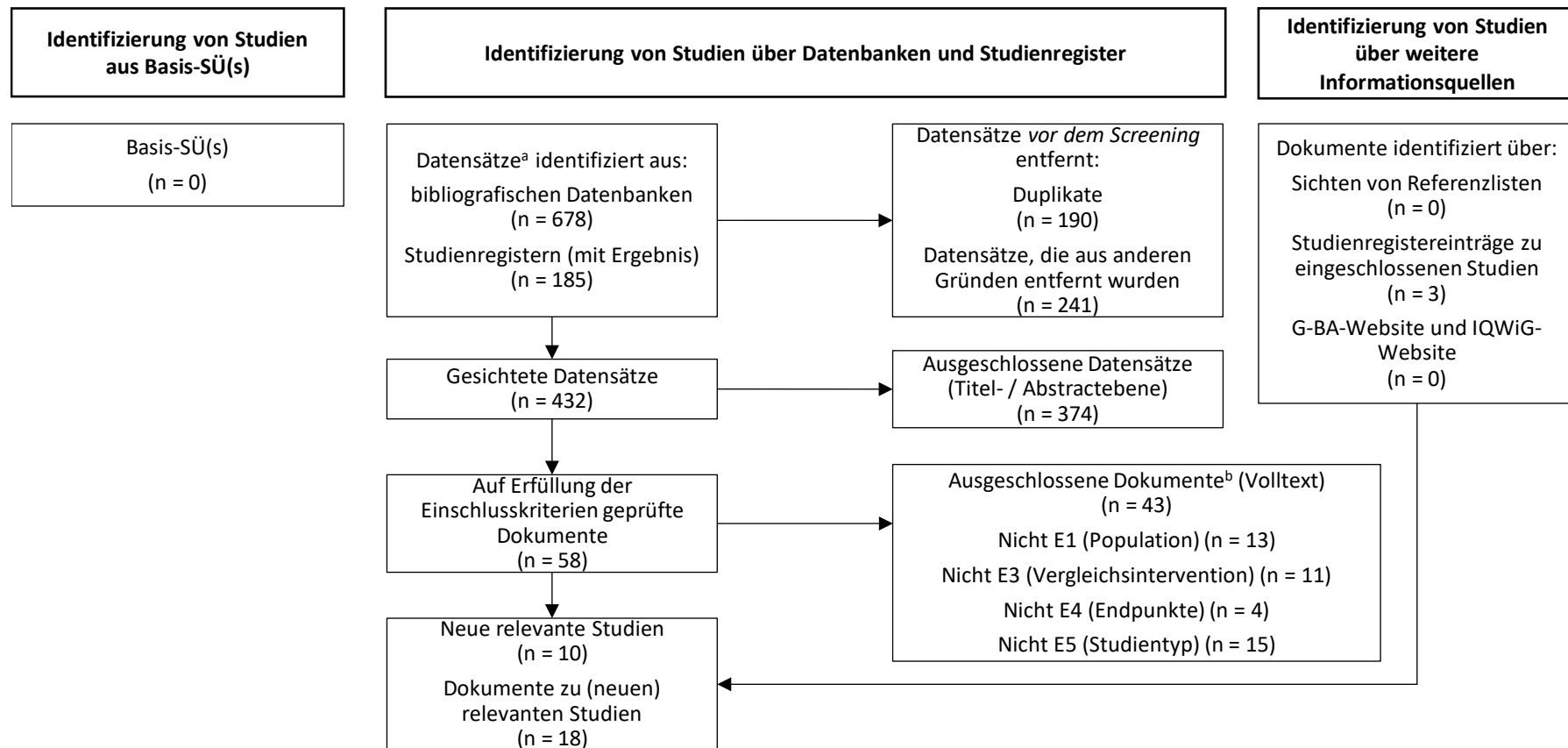
5.1 Ergebnisse der Informationsbeschaffung

Abbildung 1 zeigt das Ergebnis der Informationsbeschaffung inklusive Studienselektion gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken (letzte Suche am 13.03.2023) und die Suche in Studienregistern (letzte Suche am 27.03.2023) sind in Anhang B dargestellt.

Die Referenzen der als Volltexte geprüften, jedoch ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt 8.2.

Autorenanfragen wurden nicht gestellt.

Von den 7 identifizierten SÜs (siehe Abschnitt 8.1) wurde keine als Basis-SÜ herangezogen.



a. Datensatz: Titel oder Abstract (oder beides) einer Studie, der in einer Datenbank (z. B. MEDLINE) oder auf einer Website indiziert ist.

b. nicht gelistete Ein- und Ausschlusskriterien (n = 0)

Abbildung 1: Ergebnis der Informationsbeschaffung

Resultierender Studienpool

Durch die verschiedenen Rechenschritte wurden insgesamt 10 relevante Studien identifiziert (siehe Tabelle 2).

Tabelle 2: Studienpool der eingeschlossenen Studien

Vergleich Studie	Verfügbare Dokumente	
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern
Allopurinol vs. Placebo		
NCT01988402	nein	ja [20] / ja [20]
Poiley 2016	ja [21]	ja [22] / nein
Taylor 2012 ^a	ja [23]	ja [24] / nein
Allopurinol vs. Febuxostat vs. Placebo		
Schumacher 2008 (APEX)	ja [25]	ja [26] / ja [26]
Febuxostat vs. Placebo		
Becker 2005	ja [27]	ja [28] / ja [28]
Dalbeth 2017	ja [29]	ja [30] / ja [30]
Gunawardhana 2018	ja [31]	ja [32] / ja [32]
Huang 2020	ja [33]	nein / nein
Saag 2016	ja [34]	ja [35] / ja [35]
Saag 2019	ja [36]	ja [37] / ja [37]
a. Diese Studie wird in der weiteren Evidenzdarstellung nicht berücksichtigt, da sie in der Gesamtschau der Studienlage keinen relevanten Beitrag zum Evidenzkörper liefert (kleine Studiengröße [N = 57], sehr kurze Dauer [10 Tage zum relevanten Vergleich], nur 1 Ereignis in beiden Studienarmen für 1 relevanten Endpunkt).		

1 der 10 eingeschlossenen Studien erfüllte zwar die formalen Einschlusskriterien, jedoch wurden die Daten für den Evidenzbericht nicht herangezogen. Daher wurde die Studie in der weiteren Evidenzdarstellung nicht berücksichtigt (siehe Fußnote Tabelle 2).

5.2 Charakteristika der für die Evidenzdarstellung berücksichtigten Studien

In Tabelle 3 werden die Charakteristika der für die Evidenzdarstellung berücksichtigten Studien dargestellt. Die in den jeweiligen Studien untersuchten Interventionen werden in Tabelle 4 charakterisiert. Die wesentlichen Ein- und Ausschlusskriterien für Patientinnen und Patienten in den berücksichtigten Studien sowie die Charakteristika der jeweiligen Studienpopulationen sind in Tabelle 5 und Tabelle 6 aufgeführt.

Tabelle 3: Charakteristika der für die Evidenzdarstellung berücksichtigten Studien (mehreseitige Tabelle)

Vergleich Studie	Studien-design	Patientenzahl N	Ort und Zeitraum der Durchführung	Studiendauer
Allopurinol vs. Placebo				
NCT01988402	RCT	IAllopurinol 100 mg / 200 mg: 16 CPlacebo: 19	USA, 2007 bis 2013	28 Tage
Poiley 2016 ^a	RCT	IAllopurinol 300 mg + Colchicin: 55 ^b IAllopurinol 300 mg: 55 CPlacebo: 28	USA, Kanada, Georgien, 2014 bis 2015	12 Wochen
Allopurinol vs. Febuxostat vs. Placebo				
Schumacher 2008	RCT	IFebuxostat 80 mg: 267 IFebuxostat 120 mg: 269 IFebuxostat 240 mg: 134 IAllopurinol 100 mg / 300 mg: 268 CPlacebo: 134	USA, 2003 bis 2004	28 Wochen
Febuxostat vs. Placebo				
Becker 2005	RCT	IFebuxostat 40 mg: 37 IFebuxostat 80 mg: 40 IFebuxostat 120 mg: 38 CPlacebo: 38	USA, 2001	28 Tage
Dalbeth 2017	RCT	IFebuxostat 40 mg / 80 mg: 157 CPlacebo: 157	USA, 2010 bis 2013	24 Monate
Gunawardhana 2018	RCT	IFebuxostat IR 40 mg: 37 ^c IFebuxostat XR 40 mg: 39 ^c IFebuxostat IR 80 mg: 37 ^c IFebuxostat XR 80 mg: 38 ^c CPlacebo: 38	USA, 2014 bis 2015	3 Monate
Huang 2020	RCT	IFebuxostat 80 mg: 78 CPlacebo: 78	China, 2014 bis 2015	24 Wochen
Saag 2016	RCT	IFebuxostat 60 mg: 32 IFebuxostat 40 mg / 80 mg: 32 CPlacebo: 32	USA, 2010 bis 2012	12 Monate
Saag 2019	RCT	IFebuxostat IR 40 mg: 357 ^c IFebuxostat XR 40 mg: 355 ^c IFebuxostat IR 80 mg: 357 ^c IFebuxostat XR 80 mg: 357 ^c CPlacebo: 357	USA, 2015 bis 2016	3 Monate
<p>a. In dieser Studie waren 2 weitere Arme ohne relevante Vergleiche für die Fragestellung des Evidenzberichts enthalten. Daher werden diese nicht dargestellt.</p> <p>b. Da im Placeboarm kein Colchicin als Begleitmedikation verabreicht wurde, ist dieser Studienarm nicht mit Placebo vergleichbar und wird daher nicht weiter berücksichtigt.</p> <p>c. Febuxostat IR entspricht der zugelassenen Formulierung mit sofortiger Freigabe des Wirkstoffes. Diese wurde in dieser Studie mit einer Formulierung mit verzögerter Freigabe (Febuxostat XR) und Placebo verglichen. Beide Formulierungen werden berücksichtigt.</p>				

Tabelle 3: Charakteristika der für die Evidenzdarstellung berücksichtigten Studien (mehreseitige Tabelle)

Vergleich Studie	Studien-design	Patientenzahl N	Ort und Zeitraum der Durchführung	Studiendauer
C: Vergleichsgruppe; I: Interventionsgruppe; IR: immediate release (sofortige Freigabe); k. A.: keine Angabe; N: Anzahl randomisierter Patientinnen und Patienten; RCT: randomisierte kontrollierte Studie; XR: extended release (verzögerte Freigabe)				

Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien (mehreseitige Tabelle)

Vergleich Studie	Intervention	Vergleich	Begleittherapie
Allopurinol vs. Placebo			
NCT01988402	Allopurinol Dosierung: ▪ Tag 1-14: 1x 100 mg ▪ Tag 15-28: 1x 200 mg Dauer der Anwendung: ▪ 28 Tage	Placebo	▪ k. A.
Poiley 2016	Allopurinol Dosierung: ▪ 1x 300 mg/Tag Dauer der Anwendung: 12 Wochen	Placebo	▪ keine
Allopurinol vs. Febuxostat vs. Placebo			
Schumacher 2008	Dosierung: ▪ Interventionsgruppe 1: Febuxostat 80 mg/Tag ▪ Interventionsgruppe 2: Febuxostat 120 mg/Tag ▪ Interventionsgruppe 3: Febuxostat 240 mg/Tag ▪ Interventionsgruppe 4: Allopurinol 100 mg bzw. 300 mg ^a Dauer der Anwendung: 28 Wochen	Placebo	▪ vor der Randomisierung 2 Wochen Auswaschphase für Patientinnen und Patienten, die vor der Studie harnsäuresenkende Medikamente eingenommen haben ▪ Colchicin 0,6 mg/Tag oder Naproxen 2-mal 250 mg/Tag zur Gichtanfall-Prophylaxe ▫ während der 2-wöchigen Auswaschphase und ▫ während der ersten 8 Wochen der Intervention für alle Patientinnen und Patienten ▪ akute Gichtanfälle wurden nach Ermessen der Prüferärztin bzw. des Prüferarztes behandelt

Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien
(mehrsseitige Tabelle)

Vergleich Studie	Intervention	Vergleich	Begleittherapie
Febuxostat vs. Placebo			
Becker 2005	Febuxostat Dosierung: ▪ Interventionsgruppe 1: 1x 40 mg/Tag ▪ Interventionsgruppe 2: 1x 80 mg/Tag ▪ Interventionsgruppe 3: 1x 120 mg/Tag Dauer der Anwendung: 28 Tage	Placebo	<ul style="list-style-type: none"> ▪ 2x pro Tag 0,6 mg Colchicin zur Gichtanfall-Prophylaxe <ul style="list-style-type: none"> ▫ während der 2-wöchigen Auswaschphase vor Randomisierung für Patientinnen und Patienten, die vor der Studie harnsäuresenkende Medikamente eingenommen haben und ▫ während der ersten 2 Wochen der Intervention für alle Patientinnen und Patienten ▪ akute Gichtanfälle wurden nach Ermessen der Prüferärztin bzw. des Prüferarztes behandelt
Dalbeth 2017	Febuxostat Dosierung: ▪ 1x 40 mg bzw. 80 mg/Tag ^b Dauer der Anwendung: 24 Monate	Placebo	<ul style="list-style-type: none"> ▪ Gichtanfall-Prophylaxe für alle Patientinnen und Patienten während der ersten 6 Monate: <ul style="list-style-type: none"> ▫ 1x pro Tag 0,6 mg Colchicin oder 2x pro Tag 250 mg Naproxen plus 1x pro Tag 15 mg Lansoprazol oder ▫ eine andere Gichtanfall-Prophylaxe nach Ermessen des Prüferarztes bei vorliegenden Komorbiditäten
Gunawardhana 2018	Febuxostat Dosierung: ▪ Interventionsgruppe 1: 1x 40 mg IR/Tag ▪ Interventionsgruppe 2: 1x 40 mg XR/Tag ▪ Interventionsgruppe 3: 1x 80 mg IR/Tag ▪ Interventionsgruppe 4: 1x 80 mg XR/Tag Dauer der Anwendung: 3 Monate	Placebo	<ul style="list-style-type: none"> ▪ vor der Randomisierung 21 Tage Auswaschphase für Patientinnen und Patienten, die vor der Studie harnsäuresenkende Medikamente eingenommen haben und Einnahme von 0,6 mg Colchicin 1x alle 2 Tage zur Gichtanfall-Prophylaxe ▪ Gichtanfall-Prophylaxe während der gesamten Intervention für alle Patientinnen und Patienten 1x alle 2 Tage 0,6 mg Colchicin oder 2x pro Tag 250 mg Naproxen oder 1 anderes NSAR oder Prednison
Huang 2020	Febuxostat Dosierung: ▪ 1x 80 mg/Tag ^c Dauer der Anwendung: 24 Wochen	Placebo	k. A.

Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien (mehreseitige Tabelle)

Vergleich Studie	Intervention	Vergleich	Begleittherapie
Saag 2016	Febuxostat Dosierung: ▪ Interventionsgruppe 1: 2x 30 mg/Tag ▪ Interventionsgruppe 2: 1x 40 mg bzw. 80 mg/Tag ^b Dauer der Anwendung: 12 Monate	Placebo	<ul style="list-style-type: none"> ▪ vor der Randomisierung 21 Tage Auswaschphase für Patientinnen und Patienten, die vor der Studie harnsäuresenkende Medikamente eingenommen haben ▪ 1x alle 2 Tage 0,6 mg Colchicin oder ≤ 10 mg Prednison (kurzzeitig höhere Dosierungen erlaubt) zur Gichtanfall-Prophylaxe <ul style="list-style-type: none"> ▫ während der Auswaschphase und ▫ während der ersten 6 Monate der Intervention für alle Patientinnen und Patienten
Saag 2019	Febuxostat Dosierung: ▪ Interventionsgruppe 1: 1x 40 mg IR/Tag ▪ Interventionsgruppe 2: 1x 40 mg XR/Tag ▪ Interventionsgruppe 3: 1x 80 mg IR/Tag ▪ Interventionsgruppe 4: 1x 80 mg XR/Tag Dauer der Anwendung: 3 Monate	Placebo	<ul style="list-style-type: none"> ▪ vor der Randomisierung 21 Tage Auswaschphase für Patientinnen und Patienten, die vor der Studie harnsäuresenkende Medikamente eingenommen haben mit Einnahme von 0,6 mg Colchicin 1x alle 2 Tage zur Gichtanfall-Prophylaxe ▪ Gichtanfall-Prophylaxe während der gesamten Intervention für alle Patientinnen und Patienten <ul style="list-style-type: none"> ▫ 1x alle 2 Tage 0,6 mg Colchicin für Patientinnen und Patienten mit einer eGFR von 15-59 ml/min ▫ 1x pro Tag 0,6 mg Colchicin für Patientinnen und Patienten mit einer eGFR von ≥ 60 ml/min oder ▫ 2x pro Tag 250 mg Naproxen plus 1x pro Tag 15 mg Lansoprazol bei Patientinnen und Patienten mit einer eGFR von ≥ 50 ml/min oder eine andere Gichtanfall-Prophylaxe nach Ermessen des Prüfarztes
<p>a. Patientinnen und Patienten mit normaler Nierenfunktion (Serumkreatinin ≤ 1,5 mg/dl) erhielten 300 mg Allopurinol, Patientinnen und Patienten mit eingeschränkter Nierenfunktion (Serumkreatinin > 1,5 bis ≤ 2,0 mg/dl) erhielten 100 mg Allopurinol.</p> <p>b. Alle Patientinnen und Patienten erhielten ab Tag 1 40 mg Febuxostat. Patientinnen und Patienten mit einer Serumharnsäurekonzentration von ≥ 6 mg/dl an Tag 14 erhielten ab Monat 1 80 mg Febuxostat.</p> <p>c. Darreichungsform: Aufgelöst in 200 ml Wasser</p> <p>eGFR: estimated GFR (geschätzte glomeruläre Filtrationsrate); IR: immediate release (sofortige Freigabe); k. A.: keine Angabe; NSAR: nichtsteroidales Antirheumatikum; XR: extended release (verzögerte Freigabe)</p>			

Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Allopurinol vs. Placebo		
NCT01988402	<ul style="list-style-type: none"> ▪ Gicht mit nachgewiesenen Uratkristallen ▪ Patienteneinschluss innerhalb der ersten 72 Stunden nach Beginn eines akuten Gichtanfalls ▪ Zusätzlich eines der folgenden Kriterien: <ul style="list-style-type: none"> ▫ mindestens 2 Gichtattacken in den letzten 12 Monaten ▫ Tophi ▫ Nierensteine ▪ Harnsäure im 24-Stunden-Sammelurin über 1000 mg 	<ul style="list-style-type: none"> ▪ eGFR < 50 ml/min ▪ Einnahme von Allopurinol in den letzten 6 Monaten ▪ Krebstherapie ▪ gleichzeitige Einnahme von Azathioprin oder Cyclophosphamid ▪ Alanin-Aminotransferase, Aspartat-Aminotransferase oder alkalische Phosphatase > 1,25-fache Obergrenze des Normalwerts ▪ Schmerzen vor der Gichterkrankung im betroffenen Gelenk von mehr als 3 auf einer Likert-Skala von 1-10 ▪ neurologische Defizite um das betroffene Gelenk
Poiley 2016	<ul style="list-style-type: none"> ▪ Alter: 18 bis 75 Jahre ▪ Gicht, nach den Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ ≥ 3 Gichtanfälle in den letzten 12 Monaten ▪ SHS 7,5 bis 12 mg/dl ▪ keine harnsäuresenkende Therapie oder Einnahme von Colchicin in den letzten 2 Wochen 	<ul style="list-style-type: none"> ▪ geschätzte Kreatinin-Clearance < 60 ml/min/1,73 m² ▪ fraktionierte Urat Ausscheidung > 10 % ▪ vorangegangene Nierensteine ▪ Leberfunktionstests und Kreatinkinasewerte ≤ 3-fache des oberen Normalbereichs ▪ sekundäre Hyperurikämie ▪ Xanthinurie ▪ kein eingestellter Blutdruck und keine normalen EKG-Ergebnisse ▪ BMI > 42 kg/m² ▪ andere Erkrankung, die die Studiendurchführung beeinträchtigen könnte
Allopurinol vs. Febuxostat vs. Placebo		
Schumacher 2008	<ul style="list-style-type: none"> ▪ Alter: 18 bis 85 Jahre ▪ Hyperurikämie (SHS ≥ 8,0 mg/dl) und Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ normale Nierenfunktion (Serumkreatinin ≤ 1,5 mg/dl) oder eingeschränkte Nierenfunktion (Serumkreatinin > 1,5 bis ≤ 2,0 mg/dl) 2 Tage vor Studienbeginn 	<ul style="list-style-type: none"> ▪ Unverträglichkeit von Allopurinol, Naproxen oder Colchicin ▪ vorangegangene Nierensteine ▪ Konsum von ≥ 14 alkoholischen Getränken pro Woche ▪ hepatische Dysfunktion > 1,5-fache Obergrenze des Normalwerts von Alanin-Aminotransferase und Aspartat-Aminotransferase ▪ andere schwere Erkrankungen

Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Febuxostat vs. Placebo		
Becker 2005	<ul style="list-style-type: none"> ▪ Erwachsene ▪ Hyperurikämie (SHS \geq 8,0 mg/dl) und Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR 	<ul style="list-style-type: none"> ▪ Serumkreatinin $>$ 1,5 mg/dl ▪ gleichzeitige Therapie mit harnsäuresenkenden Mitteln, Azathioprin, 6-Mercaptopurin oder Arzneimitteln, die Acetylsalicylsäure ($>$ 325 mg) oder andere Salicylate enthalten ▪ BMI $>$ 50 kg/m² ▪ Vorerkrankungen: Xanthinurie, Lebererkrankung, Hepatische Dysfunktion ▪ Änderung der Thiazid- oder Steroidtherapie (innerhalb von 1 Monat nach Beginn der Studie) oder der Hormonersatztherapie/ Gabe von oralen Kontrazeptiva (innerhalb von 3 Monaten nach Beginn der Studie) ▪ vorangegangener Alkoholmissbrauch oder Konsum von \geq 14 alkoholischen Getränken pro Woche
Dalbeth 2017	<ul style="list-style-type: none"> ▪ Alter: Männer \geq 18 Jahre; Frauen \geq 45 Jahre, die \geq 2 Jahre postmenopausal sind oder Frauen \geq 55 Jahre, die eine Hormonersatztherapie erhalten ▪ SHS \geq 7 mg/dl ▪ Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ frühe Gicht (definiert als 1 oder 2 Gichtanfälle in den letzten 12 Monaten) 	<ul style="list-style-type: none"> ▪ vorangegangene harnsäuresenkende Therapie ▪ sekundäre Hyperurikämie ▪ rheumatoide Arthritis ▪ aktive peptische Ulkuskrankheit ▪ vorangegangene Xanthinurie ▪ Myokardinfarkt oder Schlaganfall innerhalb von 90 Tagen vor dem Einschluss ▪ Krebs, der seit \geq 5 Jahren nicht in Remission ist ▪ eGFR $<$ 60 ml/min/1,73 m² ▪ Serumkreatinin $>$ 2,0 mg/dl ▪ Alanin-Aminotransferase und / oder Aspartat-Aminotransferase $>$ 2,09-fache Obergrenze des Normalwerts ▪ Drogen- oder Alkoholmissbrauch in den letzten 5 Jahren

Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Gunawardhana 2018	<ul style="list-style-type: none"> ▪ Alter: ≥ 18 Jahre ▪ Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ SHS ≥ 8 mg/dl 4 Tage vor der Randomisierung oder am Tag des Wiederholungstests ▪ Nierenerkrankung mit moderater Funktionseinschränkung mit eGFR von ≥ 30 und < 60 ml/min/1,73 m² <ul style="list-style-type: none"> ▫ 21 Tage vor der Randomisierung für Patientinnen und Patienten, die harnsäuresenkende Medikamente einnehmen, ▫ bzw. 4 Tage vor der Randomisierung oder an einem Wiederholungstest für Patientinnen und Patienten, die keine harnsäuresenkenden Medikamente einnehmen ▪ mindestens 1 Gichtanfall in den letzten 12 Monaten 	<ul style="list-style-type: none"> ▪ Einnahme eines der Studienmedikamente innerhalb der letzten 30 Tage vor Studienbeginn ▪ sekundäre Hyperurikämie ▪ vorangegangene Xanthinurie ▪ vorangegangener Myokardinfarkt, Schlaganfall, Hospitalisierung aufgrund von instabiler Angina pectoris, kardiale oder zerebrale Revaskularisation, Hospitalisierung aufgrund einer transitorischen ischämischen Attacke ▪ Krebs in den letzten 5 Jahren (ausgenommen Basalzellkarzinome der Haut) ▪ vorangegangener Drogen- oder Alkoholmissbrauch ▪ rheumatoide Arthritis ▪ aktive peptische Ulkuskrankheit ▪ andere schwere Erkrankung, die die Studientherapie, Sicherheit oder Compliance beeinträchtigen würde
Huang 2020	<ul style="list-style-type: none"> ▪ Alter: 18 bis 70 Jahre ▪ Gicht, die durch einen Arzt bzw. eine Ärztin diagnostiziert wurde ▪ SHS ≥ 8 mg/dl 	<ul style="list-style-type: none"> ▪ vorangegangene Autoimmunerkrankung ▪ Nierenerkrankung ▪ Krebs ▪ Blutkrankheit

Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Saag 2016	<ul style="list-style-type: none"> ▪ Männer im Alter ≥ 18 Jahre oder postmenopausale Frauen ≥ 45 Jahre ▪ Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ Keine Tophi vorhanden ▪ SHS $> 7,0$ mg/dl ▪ eGFR ≥ 15 bis 50 ml/min/1,73 m² 	<ul style="list-style-type: none"> ▪ sekundäre Hyperurikämie ▪ Xanthinurie ▪ Tophöse Gicht ▪ Einnahme von Acetylsalicylsäure > 325 mg/Tag bis 35 Tage vor der Randomisierung ▪ Unverträglichkeit von Allopurinol, Febuxostat oder Colchicin ▪ akutes Koronarsyndrom in den letzten 90 Tagen ▪ Nierenerkrankung im Endstadium (eGFR < 15 ml/min/1,73 m²) oder hohe Wahrscheinlichkeit, während der Studie dialysepflichtig zu werden ▪ Alanin-Aminotransferase und / oder Aspartat-Aminotransferase > 2-fache Obergrenze des Normalwerts ▪ vorangegangener Drogen- oder Alkoholmissbrauch ▪ Konsum von ≥ 14 alkoholischen Getränken pro Woche ▪ andere schwere Erkrankung, die aus Sicht der Untersuchenden die Studientherapie, Sicherheit oder Compliance beeinträchtigen würde

Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Saag 2019	<ul style="list-style-type: none"> ▪ Alter: ≥ 18 Jahre ▪ Gicht, nach den vorläufigen Kriterien für die Klassifizierung der akuten Arthritis der primären Gicht des ACR ▪ SHS ≥ 8 mg/dl 4 Tage vor der Randomisierung oder an einem Wiederholungstest ▪ mindestens 1 Gichtanfall in den letzten 12 Monaten ▪ normale Nierenfunktion ▪ mindestens 30 % der Studienpopulation eine Nierenerkrankung mit moderater bis schwerer Funktionseinschränkung (eGFR ≥ 15 bis 59 ml/min/1,73 m²), von denen mindestens 85 Personen eine schwere Nierenerkrankung haben (eGFR ≥ 15 bis 29 ml/min/1,73 m²) 	<ul style="list-style-type: none"> ▪ Einnahme eines der Studienmedikamente innerhalb der letzten 30 Tage vor Studienbeginn ▪ sekundäre Hyperurikämie ▪ vorangegangene Xanthinurie ▪ vorangegangene kardiovaskuläre Erkrankung inklusive Myokardinfarkt, Schlaganfall, Hospitalisierung aufgrund von instabiler Angina, kardiale oder zerebrale Revaskularisation, Hospitalisierung aufgrund einer transitorischen ischämischen Attacke (außer bei Patientinnen und Patienten mit schwerer Nierenerkrankung) ▪ bei Patientinnen und Patienten mit schwerer Nierenerkrankung: Myokardinfarkt oder Schlaganfall in den letzten 90 Tagen ▪ Krebs in den letzten 5 Jahren (ausgenommen Basalzellkarzinome der Haut) ▪ vorangegangener Drogen- oder Alkoholmissbrauch ▪ Konsum von ≥ 14 alkoholischen Getränken pro Woche ▪ rheumatoide Arthritis ▪ aktive peptische Ulkuskrankheit ▪ andere schwere Erkrankungen, die die Studientherapie, Sicherheit oder Compliance beeinträchtigen würde
<p>ACR: American College of Rheumatology; BMI: Body-Mass-Index; eGFR: estimated GFR (geschätzte glomeruläre Filtrationsrate); k. A.: keine Angabe; SHS: Serumharnsäure</p>		

Tabelle 6: Charakterisierung der Studienpopulationen (mehrsseitige Tabelle)

Vergleich Studie Gruppe	N	Alter [Jahre] MW (SD)	Geschlecht [w/m] %	Baseline Serum- harnsäure MW (SD) [mg/dl]	Vorliegen von Tophi n (%)	Anzahl bisheriger Gichtanfälle: n (%)	Dauer der Gicht [Jahre] MW (SD)
Allopurinol vs. Placebo							
NCT01988402							
Allopurinol 100 mg/200 mg	16	61 (40-84) ^a	13/88	k. A.	6 (38)	4 (0-12) ^a	5,5 (0-20) ^a
Placebo	19	53 (31-68) ^a	0/100	k. A.	6 (32)	7 (1-35) ^a	4,9 (0-17) ^a
Poiley 2016							
Allopurinol 300 mg	54	53 (9)	4/96	9,0 (1,4)	10 (19)	4,6 (2) ^b	k. A.
Placebo	28	50 (10)	7/93	9,1 (1,4)	6 (21)	4,7 (3) ^b	k. A.
Allopurinol vs. Febuxostat vs. Placebo							
Schumacher 2008							
Allopurinol 100 mg/300 mg ^c	268	52 (12)	7/93	k. A.	76 (28)	k. A.	11 (9)
Febuxostat 80 mg	267	51 (12)	6/94	k. A.	62 (24)	k. A.	11 (9)
Febuxostat 120 mg	269	51 (12)	5/95	k. A.	79 (30)	k. A.	12(9)
Febuxostat 240 mg	134	54 (13)	6/94	k. A.	36 (27)	k. A.	11 (9)
Placebo	134	52 (12)	8/92	k. A.	44 (33)	k. A.	10 (8)
Febuxostat vs. Placebo							
Becker 2005							
Febuxostat 40 mg	37	52 (14)	11/89	9,24 (0,94)	6 (16)	k. A.	k. A.
Febuxostat 80 mg	40	55 (13)	5/95	9,92 (1,30)	10 (25)	k. A.	k. A.
Febuxostat 120 mg	38	56 (11)	13/87	9,58 (1,11)	11 (29)	k. A.	k. A.
Placebo	38	52 (13)	16/84	9,87 (1,33)	9 (24)	k. A.	k. A.
Dalbeth 2017							
Febuxostat 40 mg/80 mg ^d	157	51 (12)	9/91	8,7 (1,40)	14 (9)	1: 125 (80) 2: 32 (20)	k. A.
Placebo	157	50 (12)	8/92	8,8 (1,33)	24 (15)	1: 113 (72) 2: 44 (28)	k. A.

Tabelle 6: Charakterisierung der Studienpopulationen (mehrsseitige Tabelle)

Vergleich Studie Gruppe	N	Alter [Jahre] MW (SD)	Geschlecht [w/m] %	Baseline Serum- harnsäure MW (SD) [mg/dl]	Vorliegen von Tophi n (%)	Anzahl bisheriger Gichtanfälle: n (%)	Dauer der Gicht [Jahre] MW (SD)
Gunawardhana 2018							
Febuxostat IR 40 mg	37	61 (10)	32/68	9,8 (1,4)	k. A.	1-3: 22 (60) 4-6: 6 (16) > 6: 9 (24)	k. A.
Febuxostat XR 40 mg	39	64 (11)	33/67	9,6 (1,3)	k. A.	1-3: 27 (70) 4-6: 9 (23) > 6: 3 (8)	k. A.
Febuxostat IR 80 mg	37	64 (10)	24/76	9,6 (1,1)	k. A.	1-3: 23 (62) 4-6: 9 (24) > 6: 5 (14)	k. A.
Febuxostat XR 80 mg	38	61 (12)	24/76	9,8 (1,4)	k. A.	1-3: 22 (58) 4-6: 9 (24) > 6: 7(18)	k. A.
Placebo	38	65 (13)	32/68	9,7 (1,2)	k. A.	1-3: 24 (63) 4-6: 10 (26) > 6: 4 (11)	k. A.
Huang 2020							
Febuxostat 80 mg	78	43 (12)	k. A.	k. A.	k. A.	k. A.	k. A.
Placebo	78	43 (10)	k. A.	k. A.	k. A.	k. A.	k. A.
Saag 2016							
Febuxostat 60 mg	32	67 (11)	22/78	10,4 (1,43)	k. A.	k. A.	k. A.
Febuxostat 40 mg/80 mg ^e	32	64 (8)	19/81	10,4 (1,70)	k. A.	k. A.	k. A.
Placebo	32	66 (12)	19/81	10,8 (1,96)	k. A.	k. A.	k. A.

Tabelle 6: Charakterisierung der Studienpopulationen (mehreseitige Tabelle)

Vergleich Studie Gruppe	N	Alter [Jahre] MW (SD)	Geschlecht [w/m] %	Baseline Serum- harnsäure MW (SD) [mg/dl]	Vorliegen von Tophi n (%)	Anzahl bisheriger Gichtanfälle: n (%)	Dauer der Gicht [Jahre] MW (SD)
Saag 2019							
Febuxostat IR 40 mg	357	56 (11)	13/87	9,6 (1,2)	k. A.	1-3: 200 (56) 4-6: 97 (27) > 6: 61 (17)	k. A.
Febuxostat XR 40 mg	355	55 (13)	12/88	9,5 (1,2)	k. A.	1-3: 213 (60) 4-6: 92 (26) > 6: 50 (14)	k. A.
Febuxostat IR 80 mg	357	55 (11)	12/88	9,6 (1,3)	k. A.	1-3: 203 (57) 4-6: 93 (26) > 6: 60 (17)	k. A.
Febuxostat XR 80 mg	357	55 (12)	10/91	9,7 (1,3)	k. A.	1-3: 214 (60) 4-6: 85 (24) > 6: 58 (16)	k. A.
Placebo	357	54 (12)	12/89	9,7 (1,4)	k. A.	1-3: 196 (55) 4-6: 102 (29) > 6: 58 (16)	k. A.
<p>a. MW (Minimum-Maximum) b. Gichtanfälle in den letzten 12 Monaten MW (SD) c. Patientinnen und Patienten, die 100 mg Allopurinol erhielten n = 10; Patientinnen und Patienten, die 300 mg Allopurinol erhielten n = 258 d. Patientinnen und Patienten, die nach 1 Monat von 40 mg auf 80 mg umgestellt wurden n = 60 e. Patientinnen und Patienten, die nach 1 Monat von 40 mg auf 80 mg umgestellt wurden n = 20</p> <p>IR: immediate release (sofortige Freigabe); k. A.: keine Angabe; m: männlich; MW: Mittelwert; N: Anzahl randomisierter (bzw. eingeschlossener) Patientinnen und Patienten; n: Anzahl der Patientinnen und Patienten mit Merkmal; SD: Standardabweichung; w: weiblich; XR: extended release (verzögerte Freigabe)</p>							

5.3 Übersicht der Endpunkte

Aus 9 Studien wurden Daten zu kritischen und wichtigen Endpunkten extrahiert. Die Tabelle 7 zeigt die Übersicht der dargestellten Daten zu den von der Leitlinienkoordination definierten Endpunkten aus den berücksichtigten Studien. Zu dem Endpunkt Funktionsfähigkeit wurde in keiner Studie Daten berichtet.

Tabelle 7: Matrix der einbezogenen Endpunkte

Vergleich Studie	Kritische Endpunkte			Wichtige Endpunkte		
	Auftreten akuter Gichtanfälle	Serumharnsäurekonzentration	Unerwünschte Ereignisse	Auftreten von Tophi	Tophusgröße	Funktionsfähigkeit
Allopurinol vs. Placebo						
NCT01988402	–	–	●	–	–	–
Poiley 2016	–	–	●	–	–	–
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	●	●	●	●	●	–
Febuxostat vs. Placebo						
Becker 2005	●	–	●	–	–	–
Dalbeth 2017	●	●	●	–	–	–
Gunawardhana 2018	●	–	●	–	–	–
Huang 2020	–	●	●	–	–	–
Saag 2016	–	●	●	–	–	–
Saag 2019	●	–	●	–	–	–
●: Daten werden im Evidenzbericht dargestellt –: Daten wurden für den Evidenzbericht nicht herangezogen oder in der Studie nicht berichtet.						

5.4 Ergebnisse zu Endpunkten

5.4.1 Ergebnisse zum kritischen Endpunkt Auftreten akuter Gichtanfälle

Tabelle 8: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Auftreten akuter Gichtanfälle (binäre Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Anteil an Patientinnen und Patienten, die eine Behandlung eines selbst berichteten Gichtanfalls benötigen – Behandlungsdauer 8 Wochen Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [25]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	61 / 268	27 / 134	20	1,17 [0,70; 1,94]	3 [-6; 11]	moderat
a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppe der eingeschlossenen Studie (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken) b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen. d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt. C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz												

Tabelle 9: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Auftreten akuter Gichtanfälle (binäre Analyse) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Anteil an Patientinnen und Patienten, die eine Behandlung eines Gichtanfalls benötigen, 40 mg und 40 mg / 80 mg Febuxostat – Behandlungsdauer bis 3 Monate												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 4 [27,29,3 1,36]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	220 / 982	113 / 590	17	1,34 [0,58; 3,09]	5 [-6; 22]	niedrig
Anteil an Patientinnen und Patienten, die eine Behandlung eines Gichtanfalls benötigen, 80 mg und 120 mg Febuxostat^f – Behandlungsdauer bis 3 Monate												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 4 [25,27,3 1,36]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	430 / 1398	119 / 567	20	1,90 [0,84; 4,28]	12 [-3; 32]	niedrig
Anteil an Patientinnen und Patienten, die eine Behandlung eines Gichtanfalls benötigen, 40 mg / 80 mg Febuxostat – Behandlungsdauer 24 Monate												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [29]	schwer-wiegend ^g	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	46 / 157	65 / 157	41	0,59 [0,37; 0,94]	-12 [-22; -2]	moderat

Tabelle 9: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Auftreten akuter Gichtanfälle (binäre Analyse) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
<p>a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppe(n) der eingeschlossenen Studie(n) (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken)</p> <p>b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren in 2 Studien unklar. Das ITT-Prinzip wurde in 1 Studie nicht adäquat umgesetzt.</p> <p>e. Das 95 %-KI des OR überdeckt 1 und 2. Somit können weder Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>f. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt einen statistisch signifikanten Effekt zuungunsten der Intervention (OR [95 %-KI]: 4,21 [2,45; 7,23]).</p> <p>g. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Febuxostat vs. Placebo
 Auftreten akuter Gichtanfälle

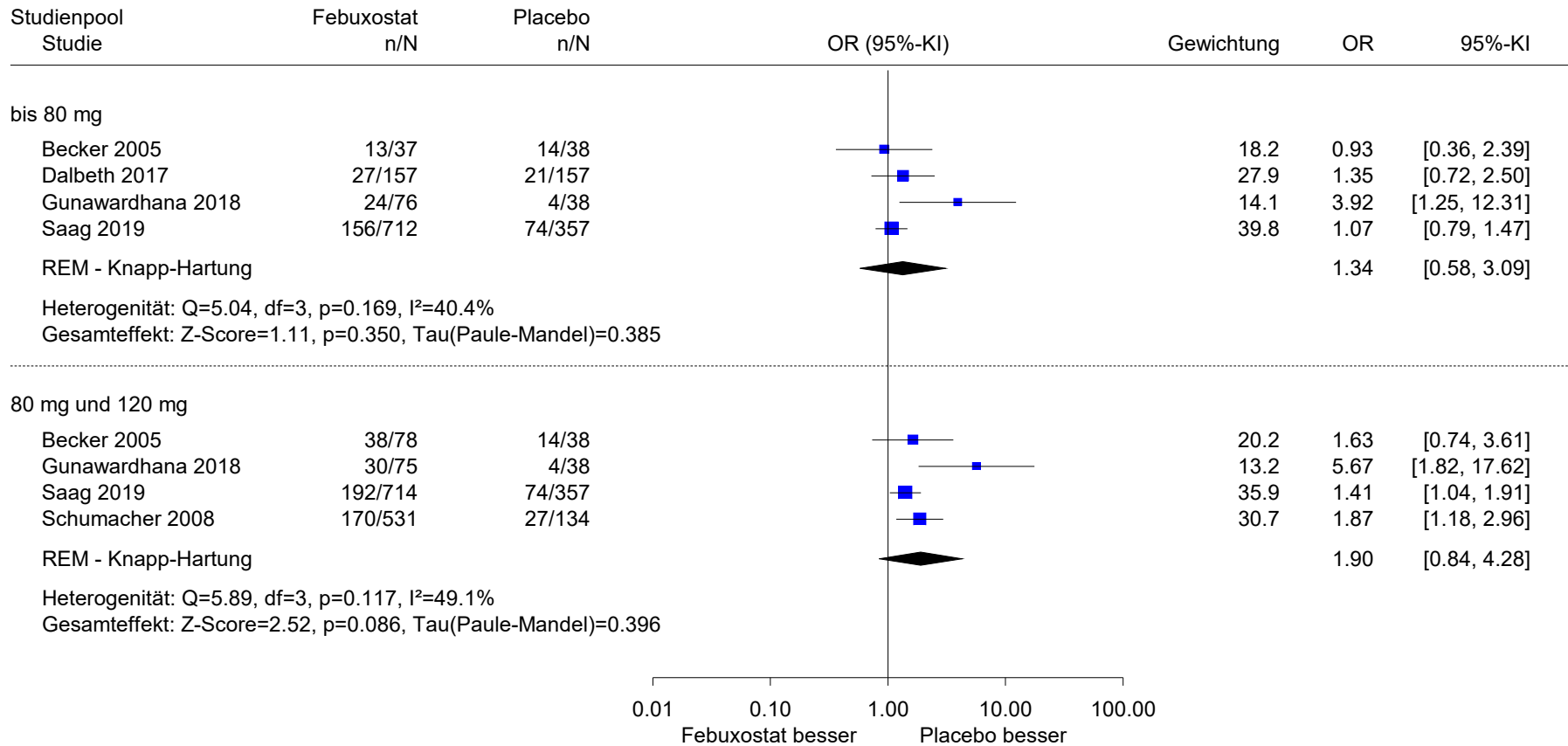


Abbildung 2: Metaanalyse, Febuxostat versus Placebo, Auftreten akuter Gichtanfälle, bis 80 mg/ 80 mg und 120 mg Febuxostat –
 Behandlungszeitraum bis 3 Monate, Effektmaß: OR

5.4.2 Ergebnisse zum kritischen Endpunkt Serumharnsäurekonzentration

Tabelle 10: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (binäre Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) nach einer Behandlungsdauer von 6 Monaten erreicht haben												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [25]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	85 / 208	1 / 99	1	67,72 [9,27; 495,06]	40 [32; 47]	moderat
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) über einen Behandlungszeitraum von 3 Monaten erreicht haben												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [25]	schwer-wiegend ^e	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	60 / 268	0 / 134	0	78,05 [4,79; 1272,94]	22 [17; 28]	moderat
<p>a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppe der eingeschlossenen Studie (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken)</p> <p>b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>e. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Tabelle 11: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (binäre Analyse) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) nach einer Behandlungsdauer von 6 Monaten erreicht haben, 60 mg und 40 mg / 80 mg Febuxostat												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 2 [29,34]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	121 / 188	3 / 189	1	96,98 [32,17; 292,41]	47 [23; 73]	moderat
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) nach einer Behandlungsdauer von 6 Monaten erreicht haben, 80 mg und 120 mg Febuxostat^e												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 2 [25,33]	schwer-wiegend ^f	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	310 / 410	1 / 167	1	268,20 [52,78; 1362,87]	57 [21; 87]	moderat
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) im Behandlungszeitraum 3 bis 6 Monaten erreicht haben, 80 mg und 120 mg Febuxostat^g												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [25]	schwer-wiegend ^h	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	301 / 531	0 / 134	0	351,86 [21,78; 5684,37]	57 [52; 61]	moderat

Tabelle 11: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (binäre Analyse) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) nach einer Behandlungsdauer von 24 Monaten erreicht haben, 40 mg / 80 mg Febuxostat Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 1 [29]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	99 / 157	9 / 157	6	28,07 [13,30; 59,24]	57 [48; 65]	moderat
a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppe(n) der eingeschlossenen Studie(n) (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken) b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen. d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt. e. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt in die gleiche Richtung (OR [95 %-KI]: 1528,80 [174,99; 13356,67]). f. Die adäquate Erzeugung der Randomisierungssequenz war in 1 Studie unklar. Die Verdeckung der Gruppenzuteilung war unklar. Das ITT-Prinzip war nicht adäquat umgesetzt. g. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt in die gleiche Richtung (OR [95 %-KI]: 585,47 [35,58; 9633,69]). h. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. i. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt. C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz												

Febuxostat vs. Placebo
Erreichen des Grenzwertes der Serumharnsäurekonzentration

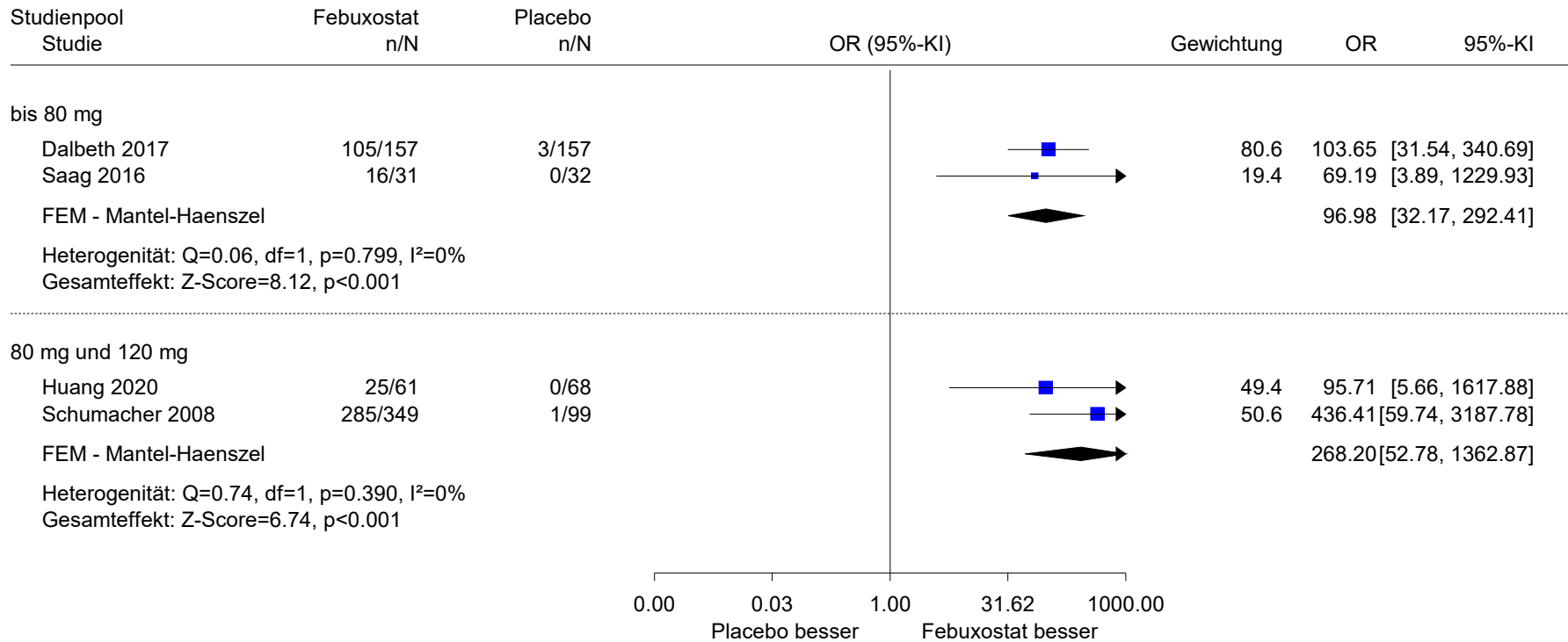


Abbildung 3: Metaanalyse, Febuxostat versus Placebo, Endpunkt Serumharnsäurekonzentration (Erreichen des Grenzwertes), bis 80 mg/ 80 mg und 120 mg Febuxostat – 6 Monate, Effektmaß: OR

Tabelle 12: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Prozentuale Reduktion der Serumharnsäurekonzentration nach einer Behandlungsdauer von 6 Monaten										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [25]	schwer-wiegend ^c	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	208; 99	-30,80 [-34,19; -27,41]	-2,18 [-2,48; -1,88]	moderat
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

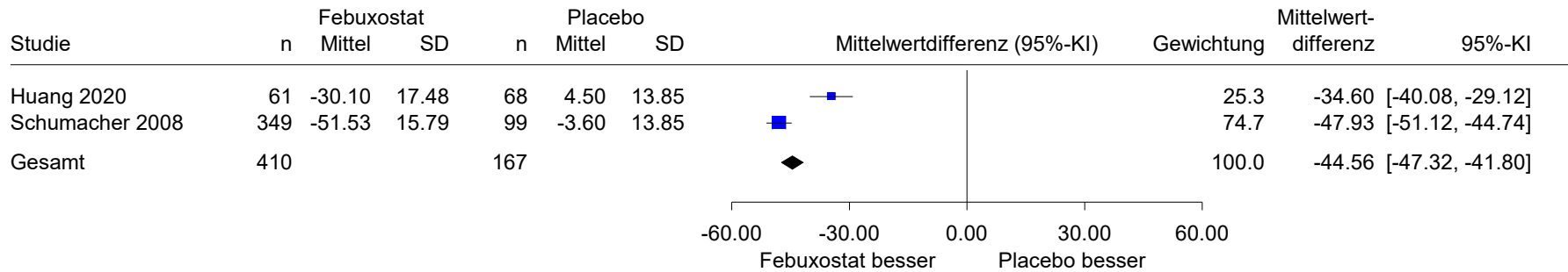
Tabelle 13: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt Serumharnsäurekonzentration (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Absolute Reduktion der Serumharnsäurekonzentration nach einer Behandlungsdauer von 6 Monaten, 60 mg und 40 mg / 80 mg Febuxostat										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [34]	schwer-wiegend ^c	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	63; 32	-4,76 [-5,52; -4,00]	-2,68 [-3,25; -2,10]	moderat
Prozentuale Reduktion der Serumharnsäurekonzentration nach einer Behandlungsdauer von 6 Monaten, 80 mg und 120 mg Febuxostat^d										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 2 [25,33]	schwer-wiegend ^e	schwer-wiegend ^f	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	410; 167	-44,56 [-47,32; -41,80]	-2,82 [-3,07; -2,57]	moderat
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>d. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt in die gleiche Richtung (Mittelwertdifferenz [95 %-KI]: -64,20 [-68,89; -59,51]).</p> <p>e. Die adäquate Erzeugung der Randomisierungssequenz war bei 1 Studie unklar. Die Verdeckung der Gruppenzuteilung war unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>f. Bedeutsame Heterogenität: $p < 0,001$; $I^2 = 94,1 \%$, die jedoch nicht in die Gesamtbewertung der Qualität der Evidenz einfließt, da alle Studien einen Vorteil für die Vergleichsgruppe berichten und sich die Unsicherheit somit auf die Größe der Effektschätzung und nicht auf das Bestehen dieses Vorteils bezieht.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

Febuxostat vs. Placebo

Prozentuale Reduktion der Serumharnsäurekonzentration

Modell mit festem Effekt - Inverse Varianz



Heterogenität: $Q=16.96$, $df=1$, $p<0.001$, $I^2=94.1\%$

Gesamteffekt: $Z\text{-Score}=-31.66$, $p<0.001$

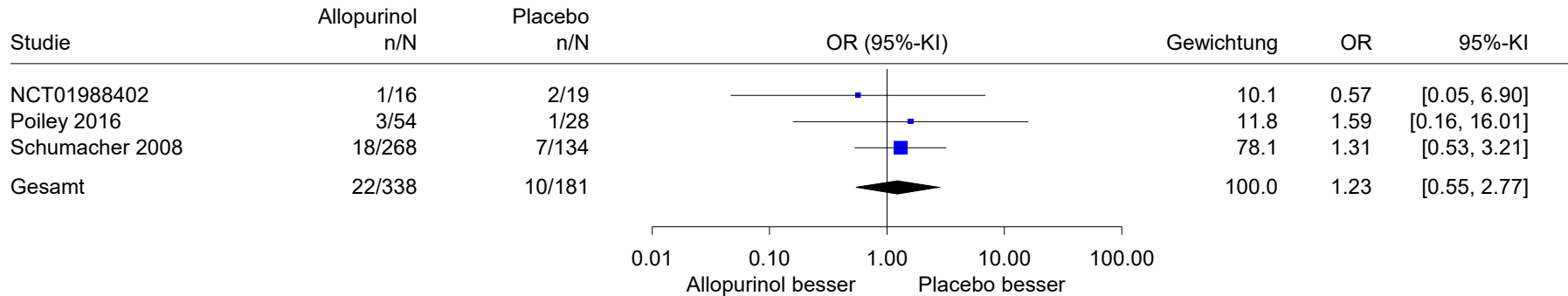
Abbildung 4: Metaanalyse, Febuxostat versus Placebo, Endpunkt Serumharnsäurekonzentration (Prozentuale Reduktion), 80 mg und 120 mg Febuxostat – 6 Monate, Effektmaß: MWD

5.4.3 Ergebnisse zum kritischen Endpunkt unerwünschte Ereignisse

Tabelle 14: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den kritischen Endpunkt unerwünschte Ereignisse (binäre Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Studien-/ Therapieabbruch aufgrund von UE – Studienende												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 3 [20,21,25]	schwerwiegend ^d	nicht schwerwiegend	nicht schwerwiegend	unentdeckt	schwerwiegend ^e	keine	22 / 338	10 / 181	5	1,23 [0,55; 2,77]	1 [-2; 8]	niedrig
SUE – Studienende												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 3 [20,21,25]	schwerwiegend ^d	nicht schwerwiegend	nicht schwerwiegend	unentdeckt	sehr schwerwiegend ^f	keine	11 / 338	4 / 181	1	1,54 [0,23; 10,32]	1 [-1; 12]	sehr niedrig
<p>a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppen der eingeschlossenen Studien (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken)</p> <p>b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war in 1 Studie nicht adäquat umgesetzt.</p> <p>e. Das 95 %-KI des OR überdeckt 1 und 2. Somit können weder Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>f. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Allopurinol vs. Placebo
 Abbruch wegen UE
 Modell mit zufälligen Effekten - Knapp und Hartung



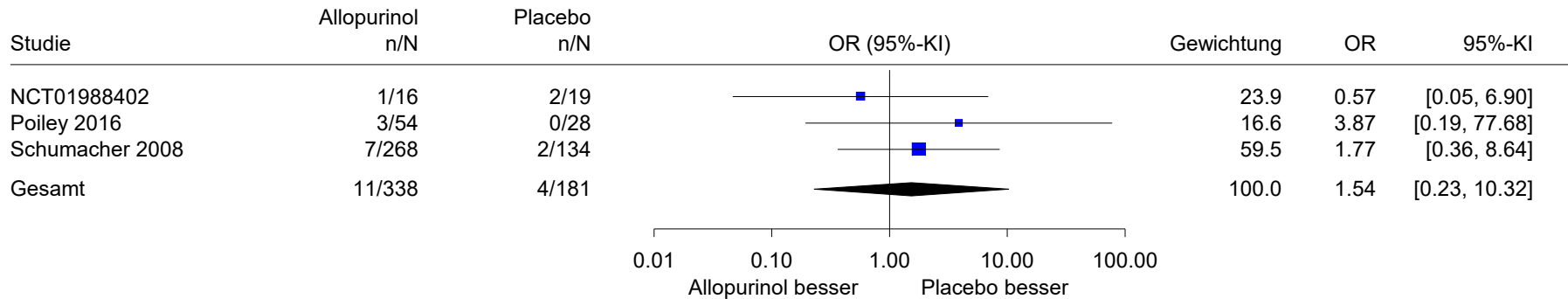
Heterogenität: $Q=0.43$, $df=2$, $p=0.805$, $I^2=0\%$
 Gesamteffekt: $Z\text{-Score}=1.09$, $p=0.389$, $\text{Tau(Paule-Mandel)}=0$

Abbildung 5: Metaanalyse, Allopurinol versus Placebo, Endpunkt unerwünschte Ereignisse (Abbruch aufgrund von UE) – Studienende, Effektmaß: OR

Allopurinol vs. Placebo

SUE

Modell mit zufälligen Effekten - Knapp und Hartung



Heterogenität: $Q=1.01$, $df=2$, $p=0.602$, $I^2=0\%$

Gesamteffekt: $Z\text{-Score}=0.97$, $p=0.435$, $\text{Tau(Paule-Mandel)}=0$

Abbildung 6: Metaanalyse, Allopurinol vs. Placebo, Endpunkt unerwünschte Ereignisse (SUE) – Studienende, Effektmaß: OR

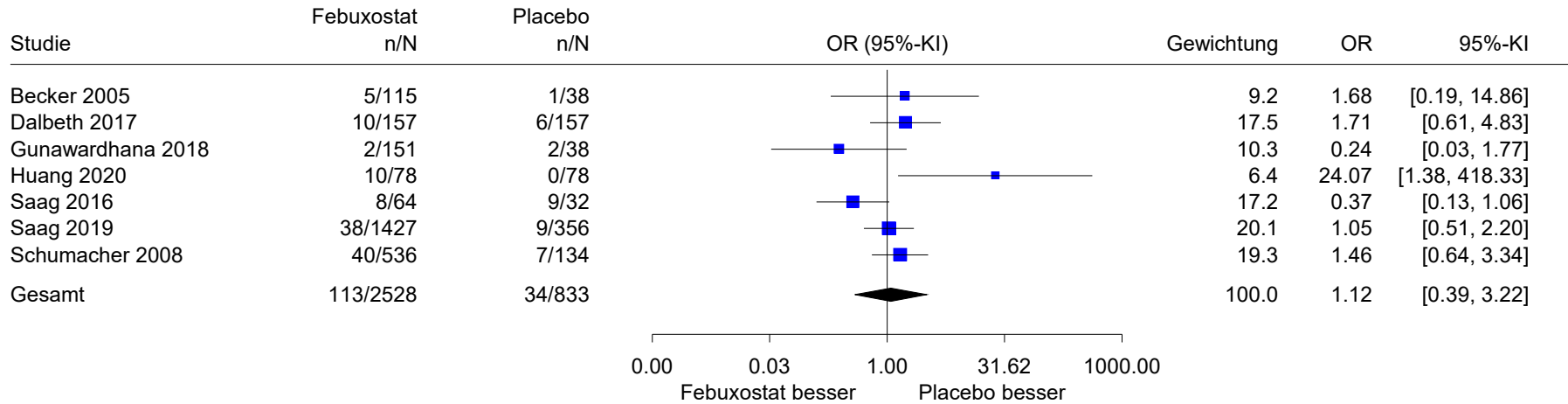
Tabelle 15: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt unerwünschte Ereignisse (binäre Analyse)(mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Studien-/ Therapieabbruch aufgrund von UE, bis 120 mg Febuxostat – Studienende^d												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 7 [25,27,29,31,33,34,36]	schwer-wiegend ^e	schwer-wiegend ^f	nicht schwer-wiegend	unentdeckt	sehr schwer-wiegend ^g	keine	113 / 2528	34 / 833	4	1,12 [0,39; 3,22]	0 [-2; 8]	sehr niedrig
SUE, bis 120 mg Febuxostat – Studienende^h												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.												
RCT; 6 [25,27,29,31,34,36]	schwer-wiegend ⁱ	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	89 / 2450	28 / 755	2	1,22 [0,83; 1,81]	0 [-0,3; 1]	moderat

Tabelle 15: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den kritischen Endpunkt unerwünschte Ereignisse (binäre Analyse)(mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
<p>a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppen der eingeschlossenen Studien (der Begriff „Risiko“ bezieht sich hier sowohl auf Endpunkte, deren Ereignisse sich für die Patientinnen und Patienten negativ als auch auf solche, deren Ereignisse sich positiv auswirken)</p> <p>b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>c. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>d. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt keinen statistisch signifikanten Effekt (OR [95 %-KI]: 1,95 [0,75; 5,05]).</p> <p>e. Die adäquate Erzeugung der Randomisierungssequenz war in 4 Studien unklar. Die Verdeckung der Gruppenzuteilung war in 5 Studien unklar. Das ITT-Prinzip war in 4 Studien nicht adäquat umgesetzt.</p> <p>f. Bedeutsame Heterogenität: $p = 0,0497$; $I^2 = 52,4 \%$</p> <p>g. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>h. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt keinen statistisch signifikanten Effekt (OR [95 %-KI]: 2,56 [0,49; 13,42]).</p> <p>i. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren in 4 Studien unklar. Das ITT-Prinzip war in 5 Studien nicht adäquat umgesetzt.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Febuxostat vs. Placebo
 Abbruch wegen UE
 Modell mit zufälligen Effekten - Knapp und Hartung



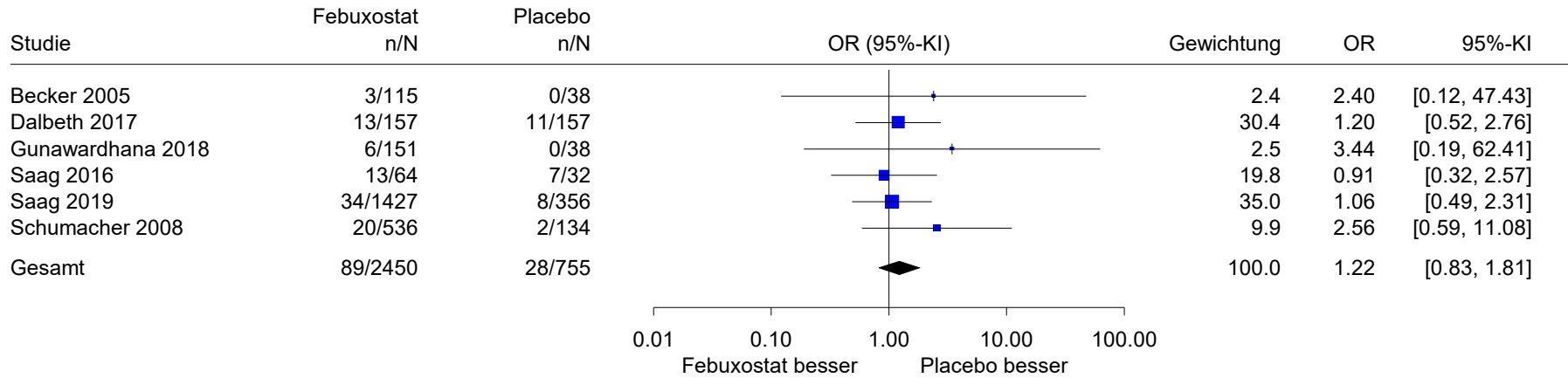
Heterogenität: $Q=12.61$, $df=6$, $p=0.050$, $I^2=52.4\%$
 Gesamteffekt: $Z\text{-Score}=0.26$, $p=0.805$, $\text{Tau(Paule-Mandel)}=0.889$

Abbildung 7: Metaanalyse, Febuxostat vs. Placebo, Endpunkt unerwünschte Ereignisse (Abbruch aufgrund von UE), bis 120 mg Febuxostat – Studienende, Effektmaß: OR

Febuxostat vs. Placebo

SUE

Modell mit zufälligen Effekten - Knapp und Hartung



Heterogenität: $Q=2.14$, $df=5$, $p=0.830$, $I^2=0\%$

Gesamteffekt: $Z\text{-Score}=1.33$, $p=0.242$, $\text{Tau(Paule-Mandel)}=0$

Abbildung 8: Metaanalyse, Febuxostat vs. Placebo, Endpunkt unerwünschte Ereignisse (SUE), bis 120 mg Febuxostat – Studienende, Effektmaß: OR

5.4.4 Ergebnisse zum wichtigen Endpunkt Auftreten von Tophi

Tabelle 16: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den wichtigen Endpunkt Auftreten von Tophi (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Reduktion der Anzahl der Tophi bei Personen mit tastbarem Tophus zu Baseline – Behandlungsdauer 28 Wochen										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [25]	schwer-wiegend ^c	nicht zutreffend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^d	keine	47; 22	-0,35 [-0,79; 0,09]	-0,41 [-0,92; 0,10]	niedrig
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>d. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder mittelgroße Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

Tabelle 17: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den wichtigen Endpunkt Auftreten von Tophi (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Reduktion der Anzahl der Tophi bei Personen mit tastbarem Tophus zu Baseline, 80 mg und 120 mg Febuxostat^c – Behandlungsdauer 28 Wochen										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [25]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	66; 22	-0,48 [-1,06; 0,09]	-0,41 [-0,90; 0,08]	niedrig
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt in die gleiche Richtung (Mittelwertdifferenz [95 %-KI]: -0,36 [-0,99; 0,26]).</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder mittelgroße Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

5.4.5 Ergebnisse zum wichtigen Endpunkt Tophusgröße

Tabelle 18: Evidenzprofil für den Vergleich Allopurinol vs. Placebo für den wichtigen Endpunkt Tophusgröße (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Prozentuale Reduktion der Tophusgröße bei Personen mit tastbarem Tophus zu Baseline – Behandlungsdauer 28 Wochen										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [25]	schwer-wiegend ^c	nicht zutreffend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^d	keine	46; 21	3,93 [-31,08; 38,94]	0,06 [-0,46; 0,57]	niedrig
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>d. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zugunsten der Prüfintervention noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

Tabelle 19: Evidenzprofil für den Vergleich Febuxostat vs. Placebo für den wichtigen Endpunkt Tophusgröße (stetige Analyse)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design; N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Prozentuale Reduktion der Tophusgröße bei Personen mit tastbarem Tophus zu Baseline, 80 mg und 120 mg Febuxostat^c – Behandlungsdauer 28 Wochen										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Interventionsgruppe.										
RCT; 1 [25]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	61; 21	-6,18 [-36,34; 23,97]	-0,10 [-0,60; 0,39]	niedrig
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Kriterien des Verzerrungspotenzials pro Studie pro Endpunkt sind Anhang A zu entnehmen.</p> <p>c. Das Ergebnis für die Dosierung 240 mg zeigt in die gleiche Richtung (Mittelwertdifferenz [95 %-KI]: -6,06 [-33,03; 20,92]).</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung waren unklar. Das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder mittelgroße Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie</p>										

6 Zusammenfassung und Anmerkungen zu den Ergebnissen

6.1 Zusammenfassung der Ergebnisse

Für die untersuchte Fragestellung zum Vergleich Allopurinol bzw. Febuxostat versus Placebo bei Patientinnen und Patienten mit Gicht wurden insgesamt 10 Studien identifiziert. Von diesen 10 Studien lieferten 9 verwertbare Daten zu 3 kritischen Endpunkten und 2 wichtigen Endpunkten.

Für den kritischen Endpunkt Serumharnsäurekonzentration ($< 360 \mu\text{mol/l}$ [6 mg/dl] über mindestens 6 Monate und/oder Reduktion der Serumharnsäurekonzentration über mindestens 6 Monate) zeigten sich in beiden Vergleichen statistisch signifikante große Effekte mit einem Vorteil für die Interventionsgruppe. Bei Febuxostat zeigten sich diese Effekte für alle untersuchten Dosierungen und Behandlungsdauern. Für den kritischen Endpunkt Auftreten akuter Gichtanfälle zeigte sich im Vergleich Febuxostat versus Placebo nach 2 Jahren ein statistisch signifikanter Effekt zugunsten der Interventionsgruppe. Das Ergebnis für die Dosierung Febuxostat 240 mg zeigt einen statistisch signifikanten Effekt zuungunsten der Intervention bei einer Behandlungsdauer von 8 Wochen. Für den Vergleich Allopurinol versus Placebo zeigte sich kein statistisch signifikanter Effekt bei einer Behandlungsdauer von 8 Wochen. Sowohl für den kritischen Endpunkt unerwünschte Ereignisse (Abbruch aufgrund von UE und SUE), als auch für die wichtigen Endpunkte Auftreten von Tophi sowie Tophusgröße zeigten sich in beiden Vergleichen keine statistisch signifikanten Effekte. Das Vertrauen in die Effektschätzungen ist mäßig bis sehr gering (moderate bis sehr niedrige Qualität der Evidenz), was meist auf schwerwiegenden Studienlimitationen sowie breiten KI der Effektschätzungen und somit einer geringen Genauigkeit der Ergebnisse beruhte. Zum wichtigen Endpunkt Funktionsfähigkeit wurden keine Studien identifiziert.

6.2 Anmerkungen zu den Ergebnissen

Die eingeschlossenen Studien unterschieden sich hinsichtlich verschiedener Aspekte, wie den Dosierungen, Darreichungsformen und der Behandlungsdauer mit den Xanthinoxidasehemmern sowie bezüglich des Zeitpunktes des Beginns der Therapie (während oder nach einem Gichtanfall) und den eingesetzten Begleittherapien. In Deutschland sind für Febuxostat Dosierungen von 80 mg pro Tag und 120 mg pro Tag zugelassen [38]. Für die USA besteht eine Zulassung für 40 mg und 80 mg Febuxostat [39]. In 6 Studien wurden Dosierungen von 40 mg, 60 mg oder 240 mg Febuxostat sowie mögliche Dosisanpassungen von 40 mg auf 80 mg Febuxostat verwendet (siehe Tabelle 4). Diese entsprachen nicht der deutschen Zulassung. Die Ergebnisse für diese Dosierungen wurden für die Endpunkte Auftreten akuter Gichtanfälle, Serumharnsäurekonzentration, Auftreten von Tophi und Tophusgröße getrennt zu den in Deutschland zugelassenen Dosierungen dargestellt. Für den Endpunkt unerwünschte Ereignisse wurden die Ergebnisse der Dosierungen bis 120 mg Febuxostat gemeinsam dargestellt und die der Dosierung 240 mg Febuxostat separat berichtet. Aufgrund der deutlich

höheren als der in Deutschland zugelassenen Dosierung von 240 mg Febuxostat wurde diese jeweils lediglich ergänzend dargestellt. Darüber hinaus kamen Darreichungsformen mit verzögerter Freisetzung von Febuxostat zur Anwendung, die nicht der deutschen Zulassung entsprachen. Die Unterschiede in der Darreichungsform fanden keine Berücksichtigung in der Ergebnisdarstellung, es zeigten sich in den Studien aber auch keine Unterschiede diesbezüglich.

Die für den kritischen Endpunkt Serumharnsäurekonzentration festgelegte Operationalisierung „Serumharnsäurekonzentration < 360 µmol/l (6 mg/dl) über mindestens 6 Monate“ wurde in keiner der eingeschlossenen Studien berichtet. In der Studie Schumacher 2008 wurde der Endpunkt „Erreichen des Grenzwertes über einen Behandlungszeitraum von 3 Monaten“ berichtet und im Evidenzbericht separat dargestellt. Die Ergebnisse zur Operationalisierung „Erreichen des Grenzwertes nach 6 Monaten“ wurden aus Studien ausgewählt, die Daten zu mindestens 1 weitere Messung der Serumharnsäure zwischen Baseline und 6 Monaten berichteten. Es wurden nur Ergebnisse aus Studien berichtet, die bei den Zwischenmessungen bereits einen ansteigenden Anteil von Patientinnen und Patienten aufwiesen, die den Grenzwert erreicht hatten. Somit kann eine Zufallsmessung nach 6 Monaten ausgeschlossen und Rückschlüsse auf die Serumharnsäurekonzentration in diesem Zeitraum gezogen werden.

Die Darstellung der Ergebnisse zu unerwünschten Ereignissen wurde begrenzt auf die Darstellung der Studien- / Therapieabbrüche wegen UE und SUE. Für diese Operationalisierungen zeigten sich keine statistisch signifikanten Effekte zuungunsten der Interventionsgruppen. Daher wurde auf die Darstellung spezifischer UE verzichtet.

Colchicin ist in Deutschland nur dann für die Gichtanfall-Prophylaxe zugelassen, wenn nicht steroidale Antirheumatika (NSAR) kontraindiziert sind oder nicht vertragen werden [40]. In den Studien, in denen eine Gichtanfall-Prophylaxe durchgeführt wurde, wurde Colchicin als Arzneimittel der ersten Wahl oder gleichberechtigt mit anderen Arzneimitteln eingesetzt. Daher ist unklar, inwieweit die Ergebnisse der eingeschlossenen Studien auf den deutschen Versorgungskontext übertragbar sind.

7 Literatur

1. Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG). Bundesgesetzblatt Teil 1 2019; (49): 2562-2584.
2. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Diagnostik und Therapie der Gicht; angemeldetes Leitlinienvorhaben [online]. 2021 [Zugriff: 21.02.2023]. URL: <https://register.awmf.org/de/leitlinien/detail/060-005>.
3. Schünemann H, Brożek J, Guyatt G et al. GRADE Handbook [online]. 2013 [Zugriff: 19.02.2021]. URL: <https://gdt.gradeapro.org/app/handbook/handbook.html>.
4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 6.1 [online]. 2022 [Zugriff: 27.01.2022]. URL: <https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf>.
5. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 19.02.2021]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.
6. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. <https://dx.doi.org/10.1136/bmj.c869>.
7. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. <https://dx.doi.org/10.1002/jrsm.1425>.
8. Guyatt GH, Oxman AD, Santesso N et al. GRADE guidelines: 12. Preparing summary of findings tables—binary outcomes. J Clin Epidemiol 2013; 66(2): 158-172. <https://dx.doi.org/10.1016/j.jclinepi.2012.01.012>.
9. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 1998; 17(8): 873-890. [https://dx.doi.org/10.1002/\(sici\)1097-0258\(19980430\)17:8<873::aid-sim779>3.0.co;2-i](https://dx.doi.org/10.1002/(sici)1097-0258(19980430)17:8<873::aid-sim779>3.0.co;2-i).
10. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. Comput Stat Data Anal 1994; 17(5): 555-574. [https://dx.doi.org/10.1016/0167-9473\(94\)90148-1](https://dx.doi.org/10.1016/0167-9473(94)90148-1).
11. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.
12. Veroniki AA, Jackson D, Viechtbauer W 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; (10 Suppl 1): 25-27.

13. Balslem H, Helfand M, Schunemann HJ et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64(4): 401-406.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.015>.
14. Meerpohl JJ, Langer G, Perleth M et al. GRADE-Leitlinien: 3. Bewertung der Qualität der Evidenz (Vertrauen in die Effektschätzer). *Z Evid Fortbild Qual Gesundhwes* 2012; 106(6): 449-456. <https://dx.doi.org/10.1016/j.zefq.2012.06.013>.
15. Guyatt GH, Oxman AD, Vist G et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol* 2011; 64(4): 407-415.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.017>.
16. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011; 64(12): 1303-1310.
<https://dx.doi.org/10.1016/j.jclinepi.2011.04.014>.
17. Guyatt GH, Oxman AD, Montori V et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol* 2011; 64(12): 1277-1282.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.011>.
18. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines 6. Rating the quality of evidenced—imprecision. *J Clin Epidemiol* 2011; 64(12): 1283-1293.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.012>.
19. Guyatt GH, Oxman AD, Sultan S et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12): 1311-1316.
<https://dx.doi.org/10.1016/j.jclinepi.2011.06.004>.
20. 59th Medical Wing. Does Allopurinol Prolong a Treated, Acute Gout Flare? [online]. 2018 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT01988402>.
21. Poiley J, Steinberg AS, Choi YJ et al. A Randomized, Double-Blind, Active- and Placebo-Controlled Efficacy and Safety Study of Arhalofenate for Reducing Flare in Patients With Gout. *Arthritis & Rheumatology* 2016; 68(8): 2027-2034.
<https://dx.doi.org/10.1002/art.39684>.
22. CymaBay Therapeutics. Evaluate the Efficacy and Safety of Arhalofenate for Preventing Flares and Reducing Serum Uric Acid in Gout Patients [online]. 2018 [Zugriff: 23.06.2023]. URL: <https://classic.clinicaltrials.gov/ct2/show/NCT02063997>.
23. Taylor TH, Mecchella JN, Larson RJ et al. Initiation of allopurinol at first medical contact for acute attacks of gout: a randomized clinical trial. *Am J Med* 2012; 125(11): 1126-1134.e7.
<https://dx.doi.org/10.1016/j.amjmed.2012.05.025>.
24. White River Junction Veterans Affairs Medical Center. Initiation of Allopurinol at First Medical Contact for Acute Attacks of Gout [online]. 2011 [Zugriff: 23.06.2023]. URL: <https://classic.clinicaltrials.gov/ct2/show/NCT01310673>.

25. Schumacher HR Jr, Becker MA, Wortmann RL et al. Effects of febuxostat versus allopurinol and placebo in reducing serum urate in subjects with hyperuricemia and gout: a 28-week, phase III, randomized, double-blind, parallel-group trial. *Arthritis Rheum* 2008; 59(11): 1540-1548. <https://dx.doi.org/10.1002/art.24209>.
26. Takeda. Phase 3, Febuxostat, Allopurinol and Placebo-Controlled Study in Gout Subjects. (APEX) [online]. 2012 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT00174915>.
27. Becker MA, Schumacher HR, Jr., Wortmann RL et al. Febuxostat, a novel nonpurine selective inhibitor of xanthine oxidase: a twenty-eight-day, multicenter, phase II, randomized, double-blind, placebo-controlled, dose-response clinical trial examining safety and efficacy in patients with gout. *Arthritis Rheum* 2005; 52(3): 916-923. <https://dx.doi.org/10.1002/art.20935>.
28. Takeda. Dose-Response, Safety and Efficacy of Febuxostat in Subjects With Gout [online]. 2011 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT00174967>.
29. Dalbeth N, Saag KG, Palmer WE et al. Effects of Febuxostat in Early Gout: A Randomized, Double-Blind, Placebo-Controlled Study. *Arthritis & Rheumatology* 2017; 69(12): 2386-2395. <https://dx.doi.org/10.1002/art.40233>.
30. Takeda. Effect of Febuxostat on Joint Damage in Hyperuricemic Subjects With Early Gout [online]. 2014 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT01078389>.
31. Gunawardhana L, Becker MA, Whelton A et al. Efficacy and safety of febuxostat extended release and immediate release in patients with gout and moderate renal impairment: phase II placebo-controlled study. *Arthritis Res Ther* 2018; 20(1): 99. <https://dx.doi.org/10.1186/s13075-018-1593-0>.
32. Takeda. Efficacy and Safety of Extended Release and Immediate Release Febuxostat in Participants With Gout and Moderate Renal Impairment [online]. 2016 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT02128490>.
33. Huang YY, Ye Z, Gu SW et al. The efficacy and tolerability of febuxostat treatment in a cohort of Chinese Han population with history of gout. *J Int Med Res* 2020; 48(5). <https://dx.doi.org/10.1177/0300060520902950>.
34. Saag KG, Whelton A, Becker MA et al. Impact of Febuxostat on Renal Function in Gout Patients With Moderate-to-Severe Renal Impairment. *Arthritis & Rheumatology* 2016; 68(8): 2035-2043. <https://dx.doi.org/10.1002/art.39654>.
35. Takeda. Effect of Febuxostat on Renal Function in Patients With Gout and Moderate to Severe Renal Impairment [online]. 2013 [Zugriff: 12.04.2023]. URL: <https://ClinicalTrials.gov/show/NCT01082640>.

36. Saag KG, Becker MA, Whelton A et al. Efficacy and Safety of Febuxostat Extended and Immediate Release in Patients With Gout and Renal Impairment: A Phase III Placebo-Controlled Study. *Arthritis & Rheumatology* 2019; 71(1): 143-153.

<https://dx.doi.org/10.1002/art.40685>.

37. Takeda. Efficacy and Safety of Extended Release and Immediate Release Febuxostat in Participants With Gout [online]. 2016 [Zugriff: 12.04.2023]. URL:

<https://ClinicalTrials.gov/show/NCT02139046>.

38. European Medicines Agency. Febuxostat Mylan [online]. 2023 [Zugriff: 25.07.2023]. URL:

<https://www.ema.europa.eu/en/medicines/human/EPAR/febuxostat-mylan>.

39. Administration FaD. Abbreviated New Drug Application (ANDA): 205374 [online]. 2023 [Zugriff: 14.08.2023]. URL:

<https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=205374>.

40. Tiofarma. Colchicin Tiofarma 0,5 mg Tabletten, Colchicin Tiofarma 1 mg Tabletten; Fachinformation [online]. 2019 [Zugriff: 08.08.2023]. URL:

<https://portal.dimdi.de/abda/servlet/Gate?language=GERM&accessid=abdafachkreise>.

41. Wong SS, 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.

42. Lefebvre C, Glanville J, Briscoe S et al. *Cochrane Handbook for Systematic Reviews of Interventions; Version 6; Technical Supplement to Chapter 4: Searching for and selecting studies* [online]. 2019 [Zugriff: 19.02.2021]. URL:

<https://training.cochrane.org/handbook/version-6/chapter-4-tech-suppl>.

8 Studienlisten

8.1 Liste der identifizierten systematischen Übersichten

1. Al-Abdouh A, Khan SU, Barbarawi M et al. Effects of Febuxostat on Mortality and Cardiovascular Outcomes: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Mayo Clinic Proceedings Innovations, Quality & Outcomes* 2020; 4(4): 434-442. <https://dx.doi.org/10.1016/j.mayocpigo.2020.04.012>.
2. Deng H, Zhang BL, Tong JD et al. Febuxostat Use and Risks of Cardiovascular Disease Events, Cardiac Death, and All-cause Mortality: Metaanalysis of Randomized Controlled Trials. *J Rheumatol* 2021; 48(7): 1082-1089. <https://dx.doi.org/10.3899/jrheum.200307>.
3. Fan M, Liu J, Zhao B et al. Comparison of efficacy and safety of urate-lowering therapies for hyperuricemic patients with gout: a meta-analysis of randomized, controlled trials. *Clin Rheumatol* 2021; 40(2): 683-692. <https://dx.doi.org/10.1007/s10067-020-05272-4>.
4. Lin YJ, Lin SY, Lin CH et al. Evaluation of urate-lowering therapy in hyperuricemia patients: a systematic review and Bayesian network meta-analysis of randomized controlled trials. *Clin Rheumatol* 2020; 39(5): 1633-1648. <https://dx.doi.org/10.1007/s10067-019-04893-8>.
5. Sun SS, Zhang DH, Shi Y et al. Efficacy and safety of urate-lowering treatments in patients with hyperuricemia: A comprehensive network meta-analysis of randomized controlled trials. *J Clin Pharm Ther* 2020; 45(4): 729-742. <https://dx.doi.org/10.1111/jcpt.13156>.
6. Zhang S, Xie Q, Xie S et al. The association between urate-lowering therapies and treatment-related adverse events, liver damage, and major adverse cardiovascular events (MACE): A network meta-analysis of randomized trials. *Pharmacotherapy* 2021; 41(9): 781-791. <https://dx.doi.org/10.1002/phar.2609>.
7. Zhao L, Cao L, Zhao TY et al. Cardiovascular events in hyperuricemia population and a cardiovascular benefit-risk assessment of urate-lowering therapies: a systematic review and meta-analysis. *Chin Med J (Engl)* 2020; 133(8): 982-993. <https://dx.doi.org/10.1097/cm9.0000000000000682>.

8.2 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

Nicht E1

1. Goldfarb DS, MacDonald PA, Gunawardhana L et al. Randomized controlled trial of febuxostat versus allopurinol or placebo in individuals with higher urinary uric acid excretion and calcium stones. *Clinical Journal of The American Society of Nephrology: CJASN* 2013; 8(11): 1960-1967. <https://dx.doi.org/10.2215/cjn.01760213>.
2. Goldfarb DS, MacDonald PA, Hunt B et al. Febuxostat in gout: serum urate response in uric acid overproducers and underexcretors. *J Rheumatol* 2011; 38(7): 1385-1389. <https://dx.doi.org/10.3899/jrheum.101156>.

3. Kamatani N, Fujimori S, Hada T et al. An allopurinol-controlled, randomized, double-dummy, double-blind, parallel between-group, comparative study of febuxostat (TMX-67), a non-purine-selective inhibitor of xanthine oxidase, in patients with hyperuricemia including those with gout in Japan: phase 3 clinical study. JCR: Journal of Clinical Rheumatology 2011; 17(4 Suppl 2): S13-18. <https://dx.doi.org/10.1097/RHU.0b013e31821d36cc>.
4. Kamatani N, Fujimori S, Hada T et al. An allopurinol-controlled, multicenter, randomized, open-label, parallel between-group, comparative study of febuxostat (TMX-67), a non-purine-selective inhibitor of xanthine oxidase, in patients with hyperuricemia including those with gout in Japan: phase 2 exploratory clinical study. JCR: Journal of Clinical Rheumatology 2011; 17(4 Suppl 2): S44-49. <https://dx.doi.org/10.1097/RHU.0b013e31821d352f>.
5. Kamatani N, Fujimori S, Hada T et al. Placebo-controlled double-blind dose-response study of the non-purine-selective xanthine oxidase inhibitor febuxostat (TMX-67) in patients with hyperuricemia (including gout patients) in japan: late phase 2 clinical study. JCR: Journal of Clinical Rheumatology 2011; 17(4 Suppl 2): S35-43. <https://dx.doi.org/10.1097/RHU.0b013e31821d351d>.
6. Kamatani N, Fujimori S, Hada T et al. Placebo-controlled, double-blind study of the non-purine-selective xanthine oxidase inhibitor Febuxostat (TMX-67) in patients with hyperuricemia including those with gout in Japan: phase 3 clinical study. JCR: Journal of Clinical Rheumatology 2011; 17(4 Suppl 2): S19-26. <https://dx.doi.org/10.1097/RHU.0b013e31821d36de>.
7. Kumar B, Agarwal PK. Comparative evaluation of efficacy and safety profile of febuxostat with allopurinol in patients with hyperuricemia and gout. International journal of pharma medicine and biological sciences 2013; 2(4): 52-56.
8. Zhang F, Liu Z, Jiang L et al. A Randomized, Double-Blind, Non-Inferiority Study of Febuxostat Versus Allopurinol in Hyperuricemic Chinese Subjects With or Without Gout. Rheumatology & Therapy 2019; 6(4): 543-557. <https://dx.doi.org/10.1007/s40744-019-00173-8>.

Nicht E3

1. Becker MA, MacDonald PA, Hunt B et al. Treating hyperuricemia of gout: safety and efficacy of febuxostat and allopurinol in older versus younger subjects. Nucleosides, Nucleotides & Nucleic Acids 2011; 30(12): 1011-1017. <https://dx.doi.org/10.1080/15257770.2011.603715>.
2. Becker MA, MacDonald PA, Hunt BJ et al. Diabetes and gout: efficacy and safety of febuxostat and allopurinol. Diabetes Obes Metab 2013; 15(11): 1049-1055. <https://dx.doi.org/10.1111/dom.12135>.

3. Becker MA, Schumacher HR, Espinoza LR et al. The urate-lowering efficacy and safety of febuxostat in the treatment of the hyperuricemia of gout: the CONFIRMS trial. *Arthritis Res Ther* 2010; 12(2): R63. <https://dx.doi.org/10.1186/ar2978>.
4. Becker MA, Schumacher HR Jr, Wortmann RL et al. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *N Engl J Med* 2005; 353(23): 2450-2461. <https://dx.doi.org/10.1056/NEJMoa050373>.
5. Desideri G, Rajzer M, Gerritsen M et al. Effects of intensive urate lowering therapy with febuxostat in comparison with allopurinol on pulse wave velocity in patients with gout and increased cardiovascular risk: the FORWARD study. *European Heart Journal Cardiovascular Pharmacotherapy* 2022; 8(3): 236-242. <https://dx.doi.org/10.1093/ehjcvp/pvaa144>.
6. Gibson T, Simmonds HA, Potter C et al. A controlled study of the effect of long term allopurinol treatment on renal function in gout. *Adv Exp Med Biol* 1980; 122A: 257-262. https://dx.doi.org/10.1007/978-1-4615-9140-5_42.
7. Hao G, Duan W, Sun J et al. Effects of febuxostat on serum cytokines IL-1, IL-4, IL-6, IL-8, TNF- α and COX-2. *Exp Ther Med* 2019; 17(1): 812-816. <https://dx.doi.org/10.3892/etm.2018.6972>.
8. Huang X, Du H, Gu J et al. An allopurinol-controlled, multicenter, randomized, double-blind, parallel between-group, comparative study of febuxostat in Chinese patients with gout and hyperuricemia. *Int J Rheum Dis* 2014; 17(6): 679-686. <https://dx.doi.org/10.1111/1756-185x.12266>.
9. Jackson RL, Hunt B, MacDonald PA. The efficacy and safety of febuxostat for urate lowering in gout patients ≥ 65 years of age. *BMC Geriatr* 2012; 12: 11. <https://dx.doi.org/10.1186/1471-2318-12-11>.
10. Mackenzie IS, Ford I, Nuki G et al. Long-term cardiovascular safety of febuxostat compared with allopurinol in patients with gout (FAST): a multicentre, prospective, randomised, open-label, non-inferiority trial. *Lancet* 2020; 396(10264): 1745-1757. [https://dx.doi.org/10.1016/s0140-6736\(20\)32234-0](https://dx.doi.org/10.1016/s0140-6736(20)32234-0).
11. O'Dell JR, Brophy MT, Pillinger MH et al. Comparative Effectiveness of Allopurinol and Febuxostat in Gout Management. *NEJM Evidence* 2022; 1(3). <https://dx.doi.org/10.1056/evidoa2100028>.
12. Wang S. The efficacy of febuxostat and allopurinol in the treatment of gout with hyperuricemia. *Pak J Pharm Sci* 2018; 31(4(Special)): 1623-1627.
13. Wells AF, MacDonald PA, Chefo S et al. African American patients with gout: efficacy and safety of febuxostat vs allopurinol. *BMC Musculoskelet Disord* 2012; 13: 15. <https://dx.doi.org/10.1186/1471-2474-13-15>.

14. White WB, Saag KG, Becker MA et al. Cardiovascular Safety of Febuxostat or Allopurinol in Patients with Gout. *N Engl J Med* 2018; 378(13): 1200-1210.

<https://dx.doi.org/10.1056/NEJMoa1710895>.

15. Xu S, Liu X, Ming J et al. A phase 3, multicenter, randomized, allopurinol-controlled study assessing the safety and efficacy of oral febuxostat in Chinese gout patients with hyperuricemia. *Int J Rheum Dis* 2015; 18(6): 669-678. <https://dx.doi.org/10.1111/1756-185x.12648>.

16. Yu KH, Lai JH, Hsu PN et al. Safety and efficacy of oral febuxostat for treatment of HLA-B*5801-negative gout: a randomized, open-label, multicentre, allopurinol-controlled study. *Scand J Rheumatol* 2016; 45(4): 304-311.

<https://dx.doi.org/10.3109/03009742.2015.1099729>.

Nicht E4

1. Hill EM, Sky K, Sit M et al. Does starting allopurinol prolong acute treated gout? A randomized clinical trial. *JCR: Journal of Clinical Rheumatology* 2015; 21(3): 120-125.

<https://dx.doi.org/10.1097/rhu.0000000000000235>.

2. Jia E, Zhang Y, Ma W et al. Initiation of febuxostat for acute gout flare does not prolong the current episode: a randomized clinical trial. *Rheumatology* 2021; 60(9): 4199-4204.

<https://dx.doi.org/10.1093/rheumatology/keaa908>.

3. Kim HA, Seo YI, Song YW. Four-week effects of allopurinol and febuxostat treatments on blood pressure and serum creatinine level in gouty men. *J Korean Med Sci* 2014; 29(8): 1077-1081. <https://dx.doi.org/10.3346/jkms.2014.29.8.1077>.

4. Saag KG, Becker MA, White WB et al. Evaluation of the Relationship Between Serum Urate Levels, Clinical Manifestations of Gout, and Death From Cardiovascular Causes in Patients Receiving Febuxostat or Allopurinol in an Outcomes Trial. *Arthritis & Rheumatology* 2022; 74(9): 1593-1601. <https://dx.doi.org/10.1002/art.42160>.

Nicht E5

1. Becker MA, Schumacher HR, MacDonald PA et al. Clinical efficacy and safety of successful longterm urate lowering with febuxostat or allopurinol in subjects with gout. *J Rheumatol* 2009; 36(6): 1273-1282. <https://dx.doi.org/10.3899/jrheum.080814>.

2. Chohan S, Becker MA, MacDonald PA et al. Women with gout: efficacy and safety of urate-lowering with febuxostat and allopurinol. *Arthritis Care Res (Hoboken)* 2012; 64(2): 256-261. <https://dx.doi.org/10.1002/acr.20680>.

3. Delbarre F, Amor B, Auscher C et al. Treatment of gout with allopurinol. A study of 106 cases. *Ann Rheum Dis* 1966; 25(6 Suppl): 627-633.

https://dx.doi.org/10.1136/ard.25.Suppl_6.627.

4. Gunther R, Knapp E. Effect of allopurinol (Zyloric) on uric acid, creatinine, fasting blood sugar and plasma lipids in patients with gout. *Wien Klin Wochenschr* 1970; 82(5): 78-82.
5. Houpt JB. The effect of allopurinol (HPP) in the treatment of gout. *Arthritis Rheum* 1965; 8(5): 899-904. <https://dx.doi.org/10.1002/art.1780080455>.
6. Nikolov NP, Seymour S, Neuner R et al. Underscoring the Importance of Allopurinol in Treating Gout: Results of a Food and Drug Administration-Mandated Safety Trial. *Arthritis & Rheumatology* 2020; 72(6): 877-878. <https://dx.doi.org/10.1002/art.41266>.
7. Pohar S, Murphy G. Febuxostat for prevention of gout attacks. *Issues Emerg Health Technol* 2006; (87): 1-4.
8. Rundles RW, Metz EN, Silberman HR. Allopurinol in the treatment of gout. *Ann Intern Med* 1966; 64(2): 229-258. <https://dx.doi.org/10.7326/0003-4819-64-2-229>.
9. Scott JT, Hall AP, Grahame R. Allopurinol in treatment of gout. *Br Med J* 1966; 2(5509): 321-327. <https://dx.doi.org/10.1136/bmj.2.5509.321>.
10. Singh JA. Can racial disparities in optimal gout treatment be reduced? Evidence from a randomized trial. *BMC Med* 2012; 10(15). <https://dx.doi.org/10.1186/1741-7015-10-15>.
11. Smyth CJ. Studies with allopurinol (HPP) in patients with tophaceous gout. *Arthritis Rheum* 1965; 8(5): 907-910. <https://dx.doi.org/10.1002/art.1780080457>.
12. Taylor TH, Mecchella JN, Larson RJ et al. Allopurinol during acute gout attacks did not differ from delayed allopurinol for pain or recurrence. *Ann Intern Med* 2013; 158(8): Jc6. <https://dx.doi.org/10.7326/0003-4819-158-8-201304160-02006>.
13. Timilsina S, Brittan K, O'Dell JR et al. Design and Rationale for the Veterans Affairs "Cooperative Study Program 594 Comparative Effectiveness in Gout: Allopurinol vs. Febuxostat" Trial. *Contemp Clin Trials* 2018; 68: 102-108. <https://dx.doi.org/10.1016/j.cct.2018.03.015>.
14. Yu TF. The effect of allopurinol in primary and secondary gout. *Arthritis Rheum* 1965; 8(5): 905-906. <https://dx.doi.org/10.1002/art.1780080456>.
15. Zollner N, Schattenkirchner M. [Allopurinol in the treatment of gout and uric acid nephrolithiasis]. *Dtsch Med Wochenschr* 1967; 92(14): 654-660. <https://dx.doi.org/10.1055/s-0028-1103816>.

Anhang A Kriterien des Verzerrungspotenzials

Tabelle 20: Kriterien des Verzerrungspotenzials zum Endpunkt Auftreten akuter Gichtanfälle

Vergleich Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein	unklar	ja
Febuxostat vs. Placebo						
Becker 2005	unklar	unklar	unklar	ja	unklar	ja
Dalbeth 2017	unklar	unklar	unklar	nein	ja	ja
Gunawardhana 2018	ja	ja	ja	ja	ja	ja
Saag 2019	ja	ja	unklar	ja	unklar	ja
ITT: Intention to treat						

Tabelle 21: Kriterien des Verzerrungspotenzials zum Endpunkt Serumharnsäurekonzentration

Vergleich Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein ^a / ja ^b	unklar	ja
Febuxostat vs. Placebo						
Dalbeth 2017	unklar	unklar	unklar	nein	ja	ja
Huang 2020	ja	unklar	unklar	nein	unklar	ja
Saag 2016	unklar	unklar	ja	nein	unklar	ja
a. für die Operationalisierungen „Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) nach 6 Monaten erreicht haben“ und Reduktion der Serumharnsäurekonzentration nach 6 Monaten“ b. für die Operationalisierung „Anteil der Patientinnen und Patienten, die den Grenzwert von < 360 µmol/l (6 mg/dl) über die Monate 3 bis 6 erreicht haben“ ITT: Intention to treat						

Tabelle 22: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse – Studienabbruch wegen UE

Vergleich Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Placebo						
NCT01988402	unklar	unklar	unklar	ja	unklar	ja
Poiley 2016	unklar	unklar	unklar	ja	ja	ja
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein	unklar	ja
Febuxostat vs. Placebo						
Becker 2005	unklar	unklar	unklar	ja	unklar	ja
Dalbeth 2017	unklar	unklar	unklar	nein	ja	ja
Gunawardhana 2018	ja	ja	ja	ja	ja	ja
Huang 2020	ja	unklar	unklar	ja	unklar	ja
Saag 2016	unklar	unklar	ja	nein	ja	ja
Saag 2019	ja	ja	unklar	nein	unklar	ja
ITT: Intention to treat						

Tabelle 23: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse – SUE

Vergleich Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Placebo						
NCT01988402	unklar	unklar	unklar	ja	unklar	ja
Poiley 2016	unklar	unklar	unklar	ja	ja	ja
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein	unklar	ja
Febuxostat vs. Placebo						
Becker 2005	unklar	unklar	unklar	ja	unklar	ja
Dalbeth 2017	unklar	unklar	unklar	nein	ja	ja
Gunawardhana 2018	ja	ja	ja	nein	ja	ja
Saag 2016	unklar	unklar	ja	nein	ja	ja
Saag 2019	ja	ja	unklar	nein	unklar	ja
ITT: Intention to treat						

Tabelle 24: Kriterien des Verzerrungspotenzials zum Endpunkt Auftreten von Tophi

Vergleich Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein	unklar	ja
ITT: Intention to treat						

Tabelle 25: Kriterien des Verzerrungspotenzials zum Endpunkt Tophusgröße

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Allopurinol vs. Febuxostat vs. Placebo						
Schumacher 2008	unklar	unklar	unklar	nein	unklar	ja
ITT: Intention to treat						

Anhang B Suchstrategien

B.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to February 20, 2023

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [41] – High specificity strategy

#	Searches
1	exp Gout/
2	gout*.ti,ab.
3	toph*.ti,ab.
4	podagra.ti,ab.
5	pseudogout.ti,ab.
6	or/1-5
7	Cochrane database of systematic reviews.jn.
8	(search or MEDLINE or systematic review).tw.
9	meta analysis.pt.
10	or/7-9
11	10 not (exp animals/ not humans.sh.)
12	and/6,11
13	12 and (english or german or multilingual or undetermined).lg.
14	..l/ 13 yr=2012-Current

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
#1	"gout"[mhe]
#2	gout*
#3	toph*
#4	#1 OR #2 OR #3
#5	(*) FROM 2012 TO 2023
#6	#4 AND #5

Suche nach Primärstudien

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to March 10, 2023

Es wurde folgender Filter übernommen:

- RCT: Lefebvre [42] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision)

#	Searches
1	exp Gout/
2	gout*.ti,ab.
3	toph*.ti,ab.
4	podagra.ti,ab.
5	pseudogout.ti,ab.
6	or/1-5
7	allopurinol*.mp.
8	febuxostat*.mp.
9	or/7-8
10	randomized controlled trial.pt.
11	controlled clinical trial.pt.
12	(randomized or placebo or randomly).ab.
13	clinical trials as topic.sh.
14	trial.ti.
15	or/10-14
16	exp animals/ not humans.sh.
17	15 not 16
18	and/6,9,17
19	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/
20	hi.fs. or case report.mp.
21	or/19-20
22	18 not 21
23	22 and (english or german or multilingual or undetermined).lg.

2. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials: Issue 2 of 12, February 2023

#	Searches
#1	[mh "Gout"]
#2	gout*:ti,ab
#3	toph*:ti,ab
#4	podagra:ti,ab
#5	pseudogout:ti,ab
#6	#1 or #2 or #3 or #4 or #5
#7	allopurinol*
#8	febuxostat*
#9	#7 or #8
#10	#6 and #9
#11	#10 not (*clinicaltrial*gov* or *trialssearch*who* or *clinicaltrialsregister*eu* or *anzctr*org*au* or *trialregister*nl* or *irct*ir* or *isrctn* or *controlled*trials*com* or *drks*de*):so
#12	#11 not ((language next (afr or ara or aze or bos or bul or car or cat or chi or cze or dan or dut or es or est or fin or fre or gre or heb or hrv or hun or ice or ira or ita or jpn or ko or kor or lit or nor or peo or per or pol or por or pt or rom or rum or rus or slo or slv or spa or srp or swe or tha or tur or ukr or urd or uzb)) not (language near/2 (en or eng or english or ger or german or mul or unknown)))
#13	#12 in Trials

B.2 Studienregister

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: <http://www.clinicaltrials.gov>
- Eingabeoberfläche: Expert Search

Suchstrategie
AREA[ResultsFirstSubmitDate] NOT MISSING AND AREA[ConditionSearch] (gout)

2. EU Clinical Trials Register

Anbieter: European Medicines Agency

- URL: <https://www.clinicaltrialsregister.eu/ctr-search/search>
- Eingabeoberfläche: Basic Search

Suchstrategie
gout AND (febuxostat* OR "TEI 6720" OR TEI-6720 OR TEI6720 OR allopurin*)

3. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <https://trialsearch.who.int>
- Eingabeoberfläche: Standard Search

Suchstrategie
gout [with Results only]

B.3 Weitere Informationsquellen und Suchtechniken

G-BA-Website und IQWiG-Website

G-BA

URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/>

Suchbegriffe
Gicht

IQWiG

URL: <https://www.iqwig.de/projekte/projekte-und-ergebnisse/>

Suchbegriffe
Gicht