

Allopurinol versus Febuxostat

Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie
der Gicht

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The word 'EVIDENZBERICHT' is written in white, uppercase letters on a dark blue segment that spans across the middle of the bar.

EVIDENZBERICHT

Projekt: V22-06D

Version: 1.0

Stand: 14.08.2023

IQWiG-Berichte – Nr. 1617

DOI: 10.60584/V22-06D

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Allopurinol versus Febuxostat – Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht

Auftraggeber

Bundesministerium für Gesundheit

Datum des Auftrags

20.12.2022

Interne Projektnummer

V22-06D

DOI-URL

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

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. Allopurinol versus Febuxostat; Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht [online]. 2023 [Zugriff: TT.MM.JJJJ]. URL: <https://dx.doi.org/10.60584/V22-06D>.

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

- Ulrike Paschen
- Tatjana Hermanns
- Anna Margraf
- Claudia Mischke
- Wiebke Sieben
- Andrea Tasar

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vi
Abbildungsverzeichnis	vii
Abkürzungsverzeichnis.....	viii
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	4
4.1 Tabellarische Darstellung der Kriterien für den Studieneinschluss	4
4.1.1 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf	5
4.1.2 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	5
4.2 Informationsbeschaffung.....	5
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten	5
4.2.2 Fokussierte Informationsbeschaffung von Studien	6
4.2.3 Anwendung von Limitierungen auf Datenbankebene	7
4.2.4 Selektion relevanter Studien	7
4.3 Informationsdarstellung und Synthese.....	8
4.3.1 Darstellung der Einzelstudien.....	8
4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial).....	9
4.3.3 Metaanalysen	11
4.3.4 Subgruppenmerkmale und andere Effektmodifikatoren.....	11
4.3.5 Bewertung der Qualität der Evidenz	12
4.3.5.1 Abwertung der Qualität der Evidenz	13
4.3.5.2 Aufwertung der Qualität der Evidenz	14
5 Ergebnisse	16
5.1 Ergebnisse der Informationsbeschaffung	16
5.2 Charakteristika der für die Evidenzdarstellung berücksichtigten Studien	18
5.3 Übersicht der Endpunkte	26
5.4 Ergebnisse zu Endpunkten	27
5.4.1 Ergebnisse zum kritischen Endpunkt Gesamtmortalität.....	27

5.4.1.1	Metaanalysen	29
5.4.2	Ergebnisse zum kritischen Endpunkt MACE.....	30
5.4.2.1	Metaanalysen	32
5.4.3	Ergebnisse zum kritischen Endpunkt weitere unerwünschte Ereignisse.....	33
5.4.3.1	Metaanalysen	36
6	Zusammenfassung und Anmerkungen zu den Ergebnissen	37
6.1	Zusammenfassung der Ergebnisse	37
6.2	Anmerkungen zu den Ergebnissen	38
7	Literatur	40
8	Studienlisten	45
8.1	Liste der identifizierten systematischen Übersichten.....	45
8.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	46
Anhang A	Kriterien des Verzerrungspotenzials.....	51
Anhang B	Übersicht über die Ergebnisse der MACE-Einzelkomponenten.....	53
Anhang C	Suchstrategien	54
C.1	Bibliografische Datenbanken	54
C.2	Studienregister	56
C.3	Weitere Informationsquellen und Suchtechniken	57

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	4
Tabelle 2: Studienpool der eingeschlossenen Studien	18
Tabelle 3: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien	18
Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien.....	19
Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien.....	22
Tabelle 6: Charakterisierung der Studienpopulationen.....	24
Tabelle 7: Matrix der einbezogenen Endpunkte.....	26
Tabelle 8: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für die Gesamtmortalität (Zeit-bis-Ereignis-Analyse)	27
Tabelle 9: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für die Gesamtmortalität (binäre Analyse)	28
Tabelle 10: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für MACE (Zeit- bis-Ereignis-Analyse).....	30
Tabelle 11: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für MACE (binäre Analyse)	31
Tabelle 12: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für weitere unerwünschte Ereignisse – SUEs (binäre Analyse).....	33
Tabelle 13: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für weitere unerwünschte Ereignisse – Gichtanfälle (binäre Analyse)	34
Tabelle 14: Kriterien des Verzerrungspotenzials zum Endpunkt Gesamtmortalität	51
Tabelle 15: Kriterien des Verzerrungspotenzials zum Endpunkt MACE	51
Tabelle 16: Kriterien des Verzerrungspotenzials zum Endpunkt weitere unerwünschte Ereignisse	52
Tabelle 17: Ergebnisse – MACE-Einzelkomponenten	53

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung	17
Abbildung 2: Metaanalyse, Allopurinol versus Febuxostat, Gesamtmortalität, Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES]); Effektmaß: HR	29
Abbildung 3: Metaanalyse, Allopurinol versus Febuxostat, Gesamtmortalität, Nachbeobachtung 26 bis 72 Wochen; Effektmaß: OR	29
Abbildung 4: Metaanalyse, Allopurinol versus Febuxostat, MACE, Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES]); Effektmaß: HR.....	32
Abbildung 5: Metaanalyse, Allopurinol versus Febuxostat, MACE, Nachbeobachtung 26 bis 72 Wochen; Effektmaß: OR.....	32
Abbildung 6: Metaanalyse, Allopurinol versus Febuxostat, SUE, Nachbeobachtung maximal 6 Jahre; Effektmaß: OR	36
Abbildung 7: Metaanalyse, Allopurinol versus Febuxostat, SUE, Nachbeobachtung 28 bis 72 Wochen; Effektmaß: OR	36

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BMG	Bundesministerium für Gesundheit
DGRh	Deutsche Gesellschaft für Rheumatologie e. V.
DVG	Digitale-Versorgung-Gesetz
EMA	Europäische Arzneimittel-Agentur
FDA	Food and Drug Administration
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
MACE	Major adverse cardiovascular Events
MWD	Mittelwertdifferenz
NYHA	New York Heart Association
OR	Odds Ratio
PICO	Population, Intervention, Comparison, Outcome
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	absolute Risikodifferenz
SMD	Standardized Mean Difference (standardisierte Mittelwertdifferenz)
SUE	schwerwiegendes unerwünschtes Ereignis
SÜ	systematische Übersicht

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 einer Behandlung mit Allopurinol im Vergleich zu einer Behandlung mit Febuxostat bei erwachsenen Patientinnen und Patienten mit Gicht und

- einer kardiovaskulären Komorbidität (z. B. koronare Herzerkrankung) oder
- mindestens 1 weiteren Risikofaktor für eine kardiovaskuläre Komorbidität (z. B. Hypertonie, Diabetes mellitus, Dyslipidämie, Niereninsuffizienz) oder
- einer kardiovaskulären Komorbidität und mindestens 1 weiteren Risikofaktor für eine kardiovaskuläre Komorbidität.

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: 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 Websites 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 und <ul style="list-style-type: none"> ▪ einer kardiovaskulären Komorbidität (z. B. koronare Herzerkrankung) oder ▪ mindestens 1 weiteren Risikofaktor für eine kardiovaskuläre Komorbidität (z. B. Hypertonie, Diabetes mellitus, Dyslipidämie, Niereninsuffizienz) oder ▪ einer kardiovaskulären Komorbidität und mindestens 1 weiteren Risikofaktor für eine kardiovaskuläre Komorbidität
E2	Prüfintervention: Behandlung mit Allopurinol
E3	Vergleichsintervention: Behandlung mit Febuxostat
E4	kritische Endpunkte: schwere kardiovaskuläre Ereignisse (MACE) ^a , Gesamt mortalität, weitere unerwünschte Ereignisse
E5	Studientyp: RCTs
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar ^b
<p>a. Es werden Kombinationen aus Ereignissen mit einem vergleichbaren Schweregrad bevorzugt. Zudem müssen die Ergebnisse der Komponenten des kombinierten Endpunkts berichtet sein.</p> <p>b. 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; MACE: Major adverse cardiovascular Events; 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 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 C.

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 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 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

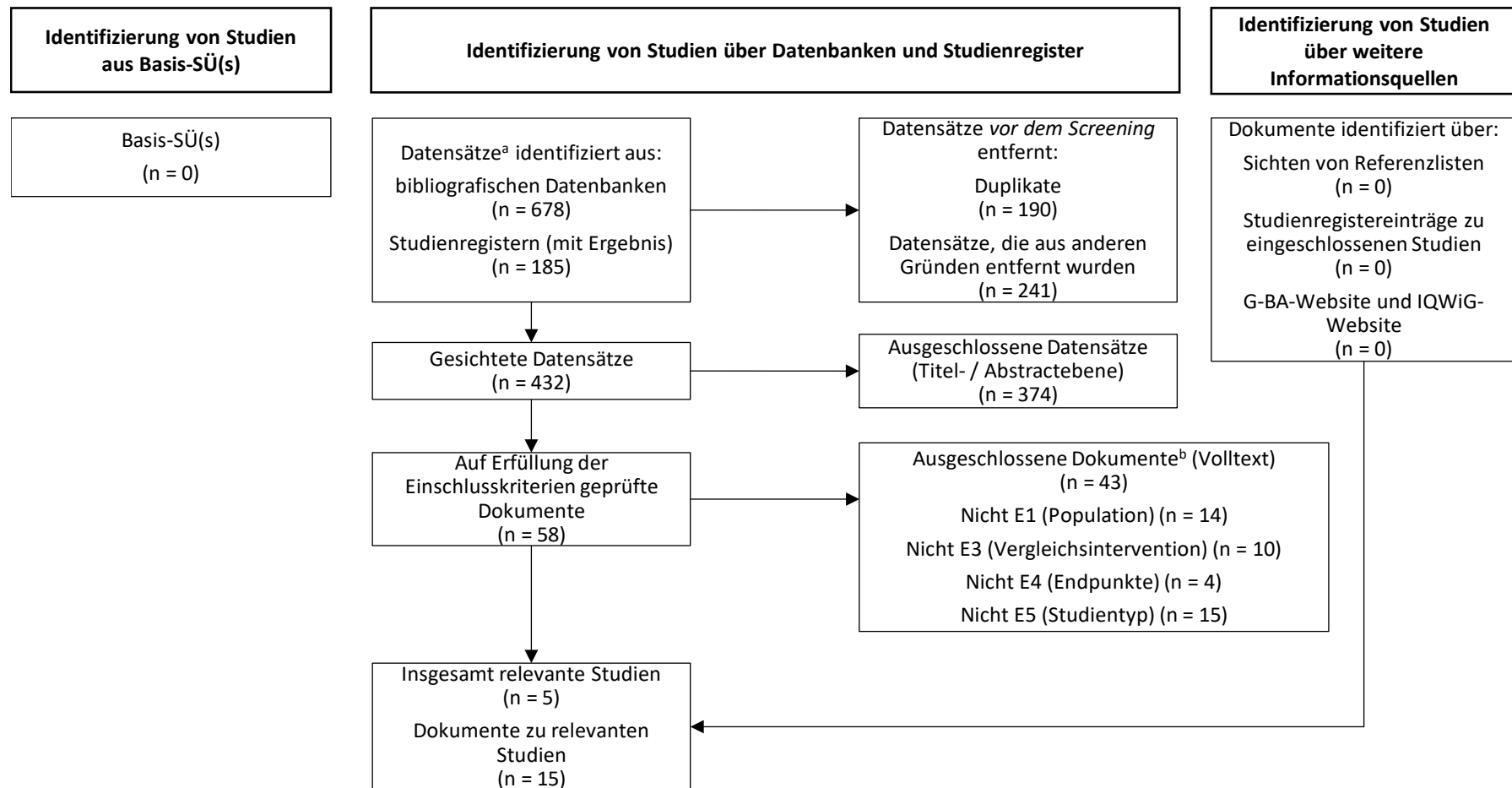
5.1 Ergebnisse der Informationsbeschaffung

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 C 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 9 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 5 relevante Studien identifiziert (siehe Tabelle 2).

Tabelle 2: Studienpool der eingeschlossenen Studien

Studie	Verfügbare Dokumente	
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern
Becker 2010	ja [20-24]	ja [25] / ja [25]
Desideri 2022	ja [26]	ja [27,28] / ja [27]
Mackenzie 2020 (FAST)	ja [29]	ja [30] / ja [30]
O'Dell 2022	ja [31]	ja [32] / ja [32]
White 2018 (CARES)	ja [33]	ja [34] / ja [34]

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 in die Evidenzdarstellung eingeschlossenen Studien

Studie	Studien-design	Patientenzahl N	Ort und Zeitraum der Durchführung	Dauer der Nachbeobachtung
Becker 2010	RCT	I: 756 C: 1513 ^a	USA; 2007 bis 2008	6 Monate
Desideri 2022	RCT	I: 98 C: 99	Italien, Niederlande, Deutschland, Serbien, Polen; 2015 bis 2017	36 Wochen
Mackenzie 2020 (FAST)	RCT	I: 3065 C: 3063	Vereinigtes Königreich, Dänemark, Schweden; 2011 bis 2019	median 1467 Tage
O'Dell 2022	RCT	I: 468 C: 472	USA; 2017 bis 2021	72 Wochen
White 2018 (CARES)	RCT	I: 3097 C: 3101	USA, Kanada, Mexiko; 2010 bis 2017	median 32 Monate
a. 1 Febuxostat-Gruppe mit 40 mg und 1 Febuxostat-Gruppe mit 80 mg C: Vergleichsgruppe; I: Interventionsgruppe; N: Anzahl randomisierter Patientinnen und Patienten; RCT: randomisierte kontrollierte Studie				

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

Studie	Intervention	Vergleich	Begleittherapie
Becker 2010	Allopurinol Dosierung: ▪ eCLcr \geq 60 bis < 90 ml/min: 300 mg/Tag ▪ eCLcr \geq 30 bis < 60 ml/min: 200 mg/Tag Dauer der Anwendung: ▪ 6 Monate	Febuxostat Dosierung: ▪ 40 mg/Tag bzw. 80 mg/Tag ^a Dauer der Anwendung: ▪ 6 Monate	Gichtanfallprophylaxe: ▪ Colchicin (0,6 mg/Tag) oder ▪ Naproxen 2-mal 250 mg/Tag ^b + Lansoprazol 15 mg/Tag Dauer der Anwendung: ▪ 6 Monate
Desideri 2022	Allopurinol Dosierung: ▪ 100 mg bis 600 mg/Tag ^c Dauer der Anwendung: ▪ 36 Wochen	Febuxostat Dosierung: ▪ 80 mg bis 120 mg/Tag ^c Dauer der Anwendung: ▪ 36 Wochen	Gichtanfallprophylaxe: ▪ Colchicin 0,5 bis 1 mg/Tag ▪ Bei Colchicin-Intoleranz: Naproxen-Natrium 2-mal 550 mg/Tag + Omeprazol 20 bis 40 mg/Tag Dauer der Anwendung: ▪ k. A. ^d
Mackenzie 2020 (FAST)	Allopurinol Dosierung: ▪ 100 mg bis 900 mg/Tag ^{e, f} Dauer der Anwendung: ▪ durchschnittlich 3 Jahre	Febuxostat Dosierung: ▪ 80 mg bis 120 mg/Tag ^e Dauer der Anwendung: ▪ durchschnittlich 3 Jahre	Gichtanfallprophylaxe: ▪ Colchicin 1- bis 2-mal 0,5 mg/Tag oder ▪ Naproxen 2-mal 250 mg/Tag oder 500 mg/Tag + Omeprazol 20 mg/Tag oder Ranitidin 2-mal 300 mg/Tag oder ▪ Meloxicam 7,5 mg/Tag oder 15 mg/Tag + Omeprazol 20 mg/Tag oder Ranitidin 2-mal 300 mg/Tag oder ▪ Diclofenac 2- bis 3-mal 50 mg/Tag + Omeprazol 20 mg/Tag oder Ranitidin 2-mal 300 mg/Tag Dauer der Anwendung: ▪ 6 Monate

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

Studie	Intervention	Vergleich	Begleittherapie
O'Dell 2022	Allopurinol Dosierung: ▪ 100 mg bis 800 mg/Tag ^g Dauer der Anwendung: ▪ 72 Wochen	Febuxostat Dosierung: ▪ 40 mg bis 120 mg/Tag ^h ▪ Dosistitration bis Harnsäurewert < 6 mg/dl oder 120 mg/Tag ⁱ Dauer der Anwendung: ▪ 72 Wochen	Gichtanfallprophylaxe: eGFR ≥ 60 ml/min: ▪ Colchicin 1- bis 2-mal 0,6 mg/Tag oder ▪ Naproxen 2-mal 250 mg/Tag oder 500 mg/Tag ^b oder ▪ nicht steroidale Antirheumatika ^j oder ▪ Glukokortikoide ^j eGFR < 60 ml/min ^b : ▪ Colchicin 1-mal 0,5 mg/jeden 2. Tag bis 1-mal 0,5 mg/Tag oder ▪ Prednison 10 mg/Tag Dauer der Anwendung: ▪ 24 bis 48 Wochen
White 2018 (CARES)	Allopurinol Dosierung: ▪ eCLcr ≥60 ml/min: 300 mg bis 600 mg/Tag ^e ▪ eCLcr ≥ 30 bis < 60 ml/min: 200 mg bis 400 mg/Tag ^e Dauer der Anwendung: ▪ ungefähr 5 Jahre	Febuxostat Dosierung: ▪ 40 mg bis 80 mg/Tag ^{e, k} Dauer der Anwendung: ▪ ungefähr 5 Jahre	Gichtanfallprophylaxe: ▪ Colchicin 0,6 mg/Tag ▪ bei Colchicin-Unverträglichkeit: Naproxen 2-mal 250 mg/Tag ^b + Lansoprazol 15 mg/Tag oder ▪ bei Colchicin- und Naproxen-Unverträglichkeit: nicht steroidale Antirheumatika oder Prednison Dauer der Anwendung: ▪ 6 Monate

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

Studie	Intervention	Vergleich	Begleittherapie
	<p>a. 50 % der Patientinnen und Patienten wurden mit 40 mg Febuxostat pro Tag behandelt (1 Febuxostat-Gruppe mit 40 mg/Tag und 1 Febuxostat-Gruppe mit 80 mg/Tag)</p> <p>b. Patientinnen und Patienten mit einer eCLcr < 50 ml/min (Becker 2010 und White 2018) bzw. eGFR < 60 ml/min (O'Dell 2022) erhielten kein Naproxen.</p> <p>c. Dosistitration bei Harnsäurewert > 6 mg/dl zwischen Woche 2 und Woche 10</p> <p>d. in der Anfangsphase der Studie</p> <p>e. Dosistitration bei Harnsäurewert > 6 mg/dl</p> <p>f. Die Patientinnen und Patienten begannen mit der Dosis, mit der sie zum Zeitpunkt der Rekrutierung behandelt wurden. 0,1 % der Patientinnen und Patienten erhielten zum Zeitpunkt der Rekrutierung 900 mg Allopurinol pro Tag. Im Studienverlauf erhielten 3,9 % der Patientinnen und Patienten 500 mg bis 900 mg Allopurinol pro Tag.</p> <p>g. Die Patientinnen und Patienten, die bereits vor der Studie mit Allopurinol behandelt wurden, erhielten die Dosierung wie vor Studienbeginn. Bei allen anderen erfolgte eine Dosistitration bei einem Harnsäurewert > 6 mg/dl bis zur 33. Woche.</p> <p>h. Die mediane Febuxostat-Dosis betrug 40 mg/Tag. Somit erhielten mindestens die Hälfte der Patientinnen und Patienten 40 mg Febuxostat pro Tag.</p> <p>i. Dosistitration bei Harnsäurewert > 6 mg/dl bis zur 33. Woche; nach Warnung der FDA Reduktion der maximalen Dosis auf 80 mg/Tag [35]</p> <p>j. keine Angaben zur Dosierung</p> <p>k. 61,0 % der Patientinnen und Patienten wurden mit 40 mg Febuxostat pro Tag behandelt.</p> <p>eCLcr: geschätzte Kreatinin-Clearance; eGFR: geschätzte glomeruläre Filtrationsrate; FDA: Food and Drug Administration; k. A.: keine Angabe</p>		

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

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Becker 2010	<ul style="list-style-type: none"> ▪ Alter 18 bis 85 Jahre ▪ Gicht ▪ Serumharnsäure $\geq 8,0$ mg/dl 	<ul style="list-style-type: none"> ▪ Komorbidität <ul style="list-style-type: none"> ▫ schwere Niereninsuffizienz eCLcr < 30 ml/min
Desideri 2022	<ul style="list-style-type: none"> ▪ Alter ≥ 18 Jahre ▪ Gicht, ohne Gichtanfälle innerhalb der letzten 4 Wochen ▪ keine frühere Behandlung mit harnsäuresenkenden Medikamenten^a ▪ Serumharnsäure > 8 mg/dl ▪ kardiovaskulärer Risiko-Score nach Joint Task Force of the European Society of Cardiology und anderen Fachgesellschaften zwischen 5 und 15 % [36] (≤ 7 % bei Vorliegen von Diabetes mellitus Typ 2) 	<ul style="list-style-type: none"> ▪ Komorbidität <ul style="list-style-type: none"> ▫ Myokardinfarkt oder Schlaganfall ▫ Herzinsuffizienz, NYHA-Stadium III oder IV ▫ schwere periphere arterielle Verschlusskrankheit ▫ schwere Niereninsuffizienz (eCLcr < 30 ml/min) ▫ Leberinsuffizienz ▫ Lebererkrankung oder -funktionsstörung ▫ Diabetes mellitus Typ 1 ▫ Krebserkrankung innerhalb der letzten 5 Jahre (außer Basalzellkarzinome) ▫ entzündliche Arthritis ▪ weitere medikamentöse Behandlung <ul style="list-style-type: none"> ▫ Azathioprin, Mercaptopurin, Theophyllin, Meclofenamat, Sulfinpyrazon, Trimethoprim sulfamethoxazol, Cyclophosphamid, Benzbromaron, Pyrazinamid, Captopril und Enalapril (bei Allopurinol), Tegafur, Pegloticase, Tacrolimus
Mackenzie 2020 (FAST)	<ul style="list-style-type: none"> ▪ Alter: ≥ 60 Jahre ▪ Gicht, die einer harnsäuresenkenden Behandlung bedarf und in den 6 Monaten vor Studienbeginn mit Allopurinol behandelt wurde ▪ Komorbidität <ul style="list-style-type: none"> ▫ Vorliegen von mindestens 1 zusätzlichen kardiovaskulären Risikofaktor^b 	<ul style="list-style-type: none"> ▪ Komorbidität <ul style="list-style-type: none"> ▫ Myokardinfarkt oder Schlaganfall innerhalb der letzten 6 Monate ▫ Herzinsuffizienz, NYHA-Stadium III oder IV ▫ schwere Niereninsuffizienz (eGFR < 30 ml/min) ▫ moderate oder schwere Leberfunktionsstörung ▫ Krebserkrankung (außer nicht melanotischer Hautkrebs und andere indolente Krebstypen, die keine Behandlung erfordern) innerhalb der letzten 5 Jahre ▪ gichtbezogen <ul style="list-style-type: none"> ▫ akuter Gichtanfall zum Zeitpunkt der Rekrutierung oder innerhalb von 14 Tagen des Abklingens eines akuten Gichtanfalls

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

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
O' Dell 2022	<ul style="list-style-type: none"> ▪ Alter \geq 18 Jahre (Männer) bzw. \geq 50 Jahre (Frauen) ▪ Gicht ▪ Serumharnsäure \geq 6,8 mg/dl 	<ul style="list-style-type: none"> ▪ Komorbidität <ul style="list-style-type: none"> ▫ schwere Niereninsuffizienz (eGFR $<$ 30 ml/min) ▫ Leberfunktionsstörung ▫ Krebserkrankung (außer nicht melanotischer Hautkrebs) innerhalb der letzten 5 Jahre ▪ gichtbezogen <ul style="list-style-type: none"> ▫ früheres Nichterreichen der Zielerumharnsäure trotz Behandlung mit Allopurinol $>$ 300 mg/Tag ▫ Serumharnsäure $>$ 15 mg/dl ▪ Weitere medikamentöse Behandlung <ul style="list-style-type: none"> ▫ frühere Behandlung mit Febuxostat ▫ Behandlung mit Azathioprin, Mercaptopurin, Didanosin, Cyclophosphamid, Probenecid, Lesinurad oder Pegloticase
White 2018 (CARES)	<ul style="list-style-type: none"> ▪ Alter $>$ 50 Jahre (Männer) bzw. $>$ 55 Jahre (Frauen) ▪ Gicht ▪ Komorbidität <ul style="list-style-type: none"> ▫ schwere kardiovaskuläre oder zerebrovaskuläre Erkrankung^c ▪ Serumharnsäure \geq 7,0 mg/dl zum Screening-Zeitpunkt oder Serumharnsäure \geq 6,0 mg/dl und unzureichend behandelte Gicht (1 oder mehr Gichtanfälle in den letzten 12 Monaten und / oder Vorliegen eines Tophus) 	<ul style="list-style-type: none"> ▪ Komorbidität <ul style="list-style-type: none"> ▫ Myokardinfarkt oder Schlaganfall innerhalb der letzten 60 Tage ▫ schwere Niereninsuffizienz (eClcr $<$ 30 ml/min) ▫ Leberfunktionsstörung ▫ Krebserkrankung (außer Basalzellkarzinome) innerhalb der letzten 5 Jahre ▪ medikamentöse Behandlung <ul style="list-style-type: none"> ▫ Behandlung mit harnsäuresenkenden Medikamenten innerhalb der letzten 7 Tage vor Randomisierung ▫ Behandlung mit Febuxostat innerhalb der letzten 30 Tage vor dem Screening-Zeitpunkt
<p>a. oder frühere Behandlung mit harnsäuresenkenden Medikamenten, aber nicht innerhalb des Monats vor Studienbeginn und nur solange die Behandlung mit harnsäuresenkenden Medikamenten nicht aus Sicherheitsgründen unterbrochen wurde</p> <p>b. u. a. Diabetes mellitus, Bluthochdruck, Dyslipidämie, Niereninsuffizienz, früherer Myokardinfarkt, Schlaganfall, transitorisch ischämische Attacke</p> <p>c. Myokardinfarkt, Hospitalisierung wegen instabiler Angina Pectoris, koronare oder zerebrale Revaskularisierung, Schlaganfall, Hospitalisierung wegen transitorischer ischämischer Attacke, periphere vaskuläre Erkrankung, Diabetes mellitus mit mikro- oder makrovaskulärer Erkrankung</p> <p>eClcr: geschätzte Kreatinin-Clearance; eGFR: geschätzte glomeruläre Filtrationsrate; NYHA: New York Heart Association</p>		

Tabelle 6: Charakterisierung der Studienpopulationen (mehreseitige Tabelle)

Studie Gruppe	N	Alter [Jahre] MW (SD)	Geschlecht [w / m] %	BMI MW (SD)	Vorliegen von Tophi n (%)	Kardio-vaskuläre Erkrankung n (%)	Diabetes mellitus n (%)	Hyperlipidämie n (%)	Hypertonie n (%)	Leichte oder moderate Niereninsuffizienz ^a n (%)
Becker 2010										
Allopurinol	756	52,9 (11,7)	6,2 / 93,8	32,7 (6,2)	148 (19,6)	k. A.	110 (14,6)	335 (44,3)	k. A. (53)	k. A. (66,3)
Febuxostat 40 mg	757	52,5 (11,7)	4,6 / 95,4	32,9 (6,4)	166 (21,9)	k. A.	89 (11,8)	299 (39,5)		k. A. (63,3)
Febuxostat 80 mg	756	53,0 (11,7)	6,1 / 93,9	32,9 (6,4)	163 (21,6)	k. A.	113 (14,9)	308 (40,7)		k. A. (66,5)
Desideri 2022										
Allopurinol	98	60,5 (10,4)	18 / 82	31,5 (4,9)	k. A.	k. A. ^b	k. A.	k. A.	k. A.	k. A.
Febuxostat	98	58,7 (10,8)	17 / 83	31,7 (5,5)	k. A.	k. A. ^b	k. A.	k. A.	k. A.	k. A.
Mackenzie 2020 (FAST)										
Allopurinol	3065	70,9 (6,5)	15,0 / 85,0	31,2 (5,3)	329 (10,7)	1008 ^c (32,9) ^c	719 (23,5)	1729 (56,4)	2619 (85,4)	483 (15,8)
Febuxostat	3063	71,0 (6,4)	14,5 / 85,5	31,0 (5,1)	299 (9,8)	1038 ^c (33,9) ^c	661 (21,6)	1819 (59,4)	2575 (84,1)	504 (16,5)
O'Dell 2022										
Allopurinol	468	62,9 (11,8)	1,5 / 98,5	33,6 (6,6)	81 (17,3)	k. A. (30,1)	k. A. (35,0)	k. A.	k. A. (78,0)	k. A. (87,8)
Febuxostat	472	61,3 (12,9)	1,7 / 98,3	33,7 (6,6)	71 (15,0)	k. A. (23,5)	k. A. (31,6)	k. A.	k. A. (74,8)	k. A. (87,3)
White 2018 (CARES)										
Allopurinol	3092	65,0 ^d (58 bis 71) ^d	16,2 / 83,8	33,4 (6,9)	650 (21,0)	3092 ^e (100) ^e	1699 (54,9)	2702 (87,4)	2851 (92,2)	3092 (100)
Febuxostat	3098	64,0 ^d (58 bis 71) ^d	15,9 / 84,1	33,6 (7,0)	668 (21,6)	3098 ^e (100) ^e	1710 (55,2)	2678 (86,4)	2864 (92,4)	3098 (100)

Tabelle 6: Charakterisierung der Studienpopulationen (mehreseitige Tabelle)

Studie Gruppe	N	Alter [Jahre] MW (SD)	Geschlecht [w / m] %	BMI MW (SD)	Vorliegen von Tophi n (%)	Kardio-vaskuläre Erkrankung n (%)	Diabetes mellitus n (%)	Hyperlipidämie n (%)	Hypertonie n (%)	Leichte oder moderate Niereninsuffizienz ^a n (%)
<p>a. Patientinnen und Patienten mit eGFR ≥ 30 ml/min und < 90 ml/min (Becker 2010, Mackenzie 2020 [FAST] und O'Dell 2022) bzw. Patientinnen und Patienten mit eCLcr ≥ 30 ml/min und < 90 ml/min (White 2018 [CARES]); in allen Studien wurden Patientinnen und Patienten mit einer schweren Niereninsuffizienz (eGFR bzw. eCLcr < 30 ml/min) ausgeschlossen (siehe Tabelle 5)</p> <p>b. kardiovaskulärer Risiko-Score gemäß Joint Task Force of the European Society of Cardiology und anderen Fachgesellschaften $7,34 \pm 3,72$ in der Allopurinol-Gruppe und $6,81 \pm 3,24$ in der Febuxostat-Gruppe</p> <p>c. Myokardinfarkt, Schlaganfall, transitorische ischämische Attacke, akutes koronares Syndrom, koronare Revaskularisierung, Angina Pectoris, Herzinsuffizienz</p> <p>d. Median (Spannweite)</p> <p>e. Myokardinfarkt, Hospitalisierung wegen instabiler Angina Pectoris, koronare oder zerebrale Revaskularisierung, Schlaganfall, Hospitalisierung wegen transitorischer ischämischer Attacke, periphere vaskuläre Erkrankung, Diabetes mellitus mit mikro- oder makrovaskulärer Erkrankung</p> <p>BMI: Body-Mass-Index; eGFR: geschätzte glomeruläre Filtrationsrate; k. A.: keine Angabe; m: männlich; MW: Mittelwert; N: Anzahl randomisierter (bzw. eingeschlossener) Patientinnen und Patienten; n: Anzahl Personen mit Merkmal; RCT: randomisierte kontrollierte Studie; SD: Standardabweichung; w: weiblich</p>										

5.3 Übersicht der Endpunkte

Aus den 5 eingeschlossenen Studien wurden Daten zu den kritischen Endpunkten extrahiert. Die Tabelle 7 zeigt die Übersicht der dargestellten Daten zu den von der Leitlinienkoordination definierten Endpunkten aus diesen Studien.

Tabelle 7: Matrix der einbezogenen Endpunkte

Studie	Kritische Endpunkte		
	Gesamtmortalität	MACE	Weitere unerwünschte Ereignisse
Becker 2010	•	•	•
Desideri 2022	•	–	•
Mackenzie 2020 (FAST)	•	•	•
O'Dell 2022	•	•	•
White 2018 (CARES)	•	•	•
●: Daten werden im Evidenzbericht dargestellt. –: Daten wurden für den Evidenzbericht nicht herangezogen oder in der Studie nicht berichtet. MACE: Major adverse cardiovascular Events			

5.4 Ergebnisse zu Endpunkten

5.4.1 Ergebnisse zum kritischen Endpunkt Gesamtmortalität

Tabelle 8: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für die Gesamtmortalität (Zeit-bis-Ereignis-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	Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]			HR [95 %-KI]	RD in %-Punkten [95 %-KI]	
							[I]	[C]				
Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES])												
Interpretation der Effektschätzung: Ein Hazard Ratio kleiner als 1 bzw. eine negative absolute Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
RCT; 2 [29,33]	schwerwiegend ^d	sehr schwerwiegend ^e	nicht schwerwiegend	unentdeckt	nicht schwerwiegend	keine	263 / 3065 199 / 3092 k. A.	222 / 3063 243 / 3098 k. A.	k. A. ^f	1,19 [1,00; 1,42] 0,82 [0,68; 0,99]	k. A. ^f	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. Es gab sonstige Aspekte (unzureichende Nachbeobachtung) des Verzerrungspotenzials in White 2018 (CARES).</p> <p>e. Bedeutsame Heterogenität und Effektschätzungen zeigen in unterschiedliche Richtungen, daher erfolgte keine Darstellung eines gepoolten Effekts.</p> <p>f. Aufgrund der sehr schwerwiegenden Inkonsistenz besteht große Unsicherheit bezüglich des Therapieeffekts. Hier kann die Risikodifferenz keinen Beitrag zur Abwägung von Vor- oder Nachteilen der Therapie leisten und wird daher (sowie auch das Basisrisiko) nicht geschätzt.</p> <p>C: Vergleichsgruppe; HR: Hazard Ratio; I: Prüflnterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Tabelle 9: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für die Gesamtmortalität (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]	
Nachbeobachtung 26 bis 72 Wochen												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
RCT; 3 [22,26,31]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	11/1322	10/2083	0,13	1,30 [0,55; 3,06]	0,04 [-0,06; 0,27]	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. In O'Dell 2022 war die adäquate Erzeugung der Randomisierungssequenz unklar. In O'Dell 2022 und Desideri 2022 war die Verdeckung der Gruppenzuteilung unklar und das ITT-Prinzip nicht adäquat umgesetzt.</p> <p>e. Die Effektschätzung beruht auf 2 Studien. Die 3. Studie (0 Ereignisse in beiden Gruppen) zeigt ebenfalls keinen statistisch signifikanten Unterschied.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; ITT: Intention to treat; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

5.4.1.1 Metaanalysen

Allopurinol vs. Febuxostat - HR
 Gesamtmortalität
 Modell mit festem Effekt - Inverse Varianz (zur Darstellung der Gewichte)

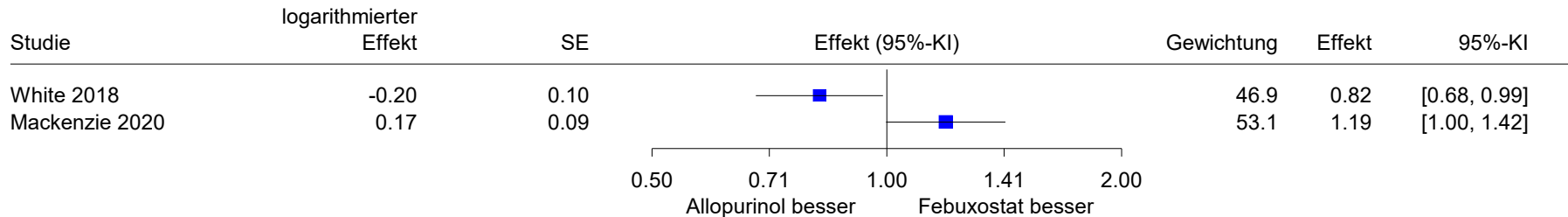


Abbildung 2: Metaanalyse, Allopurinol versus Febuxostat, Gesamtmortalität, Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES]); Effektmaß: HR

Allopurinol vs. Febuxostat
 Gesamtmortalität
 Modell mit festem Effekt - Mantel-Haenszel

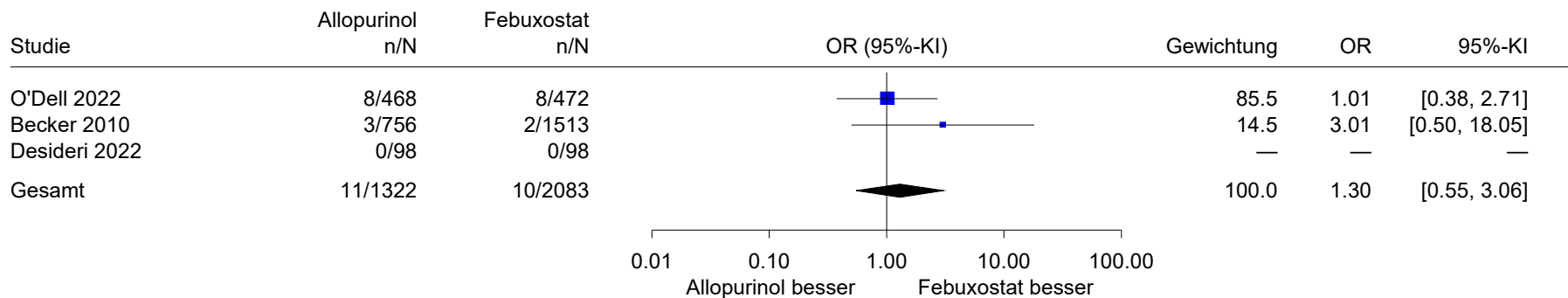


Abbildung 3: Metaanalyse, Allopurinol versus Febuxostat, Gesamtmortalität, Nachbeobachtung 26 bis 72 Wochen; Effektmaß: OR

5.4.2 Ergebnisse zum kritischen Endpunkt MACE

Tabelle 10: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für MACE (Zeit-bis-Ereignis-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	Publikationsbias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		HR [95 %-KI]	RD in %-Punkten [95 %-KI]	
							Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]					
Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES])												
Interpretation der Effektschätzung: Ein Hazard Ratio kleiner als 1 bzw. eine negative absolute Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
RCT; 2 [29,33]	schwerwiegend ^d	nicht schwerwiegend	nicht schwerwiegend	unentdeckt	nicht schwerwiegend	keine	556 / 6157 k. A.	552 / 6161 k. A.	8,96	1,01 [0,90; 1,14]	0,09 [-0,86; 1,19]	moderat
<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. Es gab sonstige Aspekte (unzureichende Nachbeobachtung) des Verzerrungspotenzials in White 2018 (CARES). In Mackenzie 2020 (FAST) war die Verblindung nicht gegeben.</p> <p>C: Vergleichsgruppe; HR: Hazard Ratio; I: Prüflinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; MACE: Major adverse cardiovascular Events; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Tabelle 11: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für MACE (binäre Analyse)

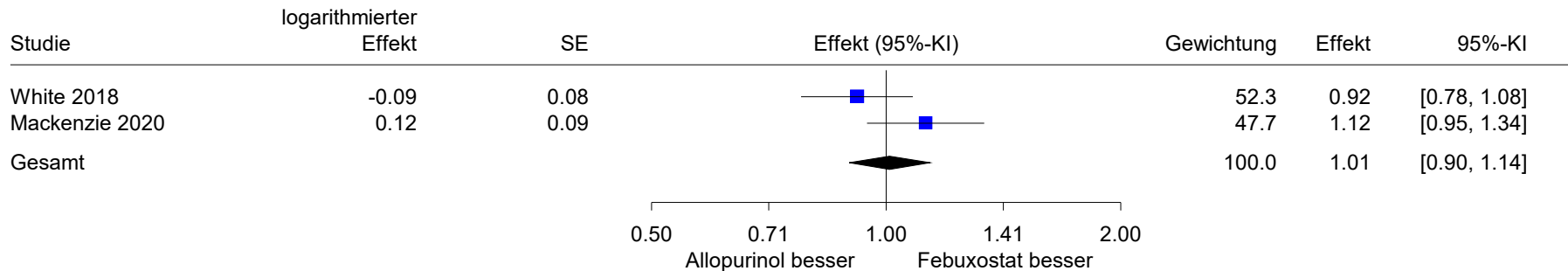
Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Nachbeobachtung 26 bis 72 Wochen												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
RCT; 2 [22,31]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	schwer-wiegend ^e	keine	13/1224	13/1985	1,16	1,18 [0,54; 2,56]	0,21 [-0,53; 1,75]	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. In O'Dell 2022 waren die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung unklar. Zudem war das ITT-Prinzip 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>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; ITT: Intention to treat; KI: Konfidenzintervall; MACE: Major adverse cardiovascular Events; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

5.4.2.1 Metaanalysen

Allopurinol vs. Febuxostat - HR

MACE

Modell mit festem Effekt - Inverse Varianz



Heterogenität: $Q=2.76$, $df=1$, $p=0.097$, $I^2=63.8\%$

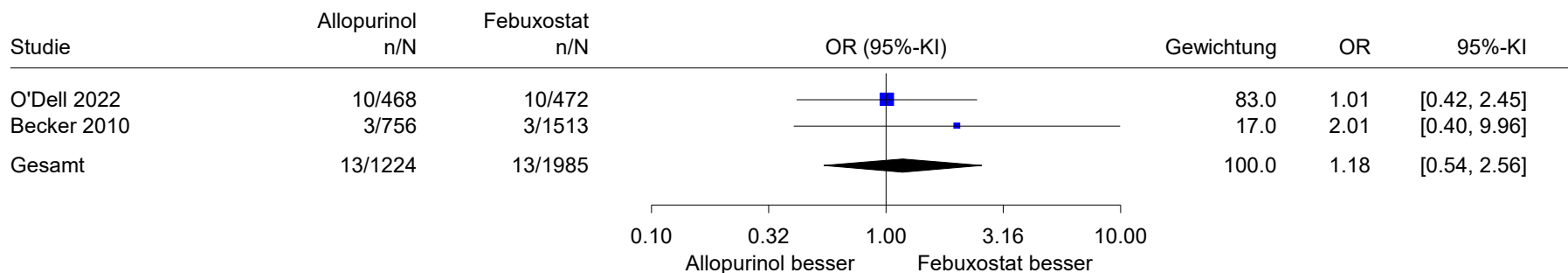
Gesamteffekt: Z-Score=0.17, $p=0.864$

Abbildung 4: Metaanalyse, Allopurinol versus Febuxostat, MACE, Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES]); Effektmaß: HR

Allopurinol vs. Febuxostat

MACE

Modell mit festem Effekt - Mantel-Haenszel



Heterogenität: $Q=0.54$, $df=1$, $p=0.462$, $I^2=0\%$

Gesamteffekt: Z-Score=0.41, $p=0.679$

Abbildung 5: Metaanalyse, Allopurinol versus Febuxostat, MACE, Nachbeobachtung 26 bis 72 Wochen; Effektmaß: OR

5.4.3 Ergebnisse zum kritischen Endpunkt weitere unerwünschte Ereignisse

Tabelle 12: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für weitere unerwünschte Ereignisse – SUEs (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	Publikationsbias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
Nachbeobachtung median 1467 Tage (Mackenzie 2020 [FAST]) bzw. median 32 Monate (White 2018 [CARES])												
RCT; 2 [29,33]	schwerwiegend ^d	sehr schwerwiegend ^e	nicht schwerwiegend	unentdeckt	nicht schwerwiegend	keine	995/3092 1812/3050	1046/3098 1720/3001	k. A. ^f	0,93 [0,84; 1,03] 1,09 [0,98; 1,21]	k. A. ^f	sehr niedrig
Nachbeobachtung 26 bis 72 Wochen												
RCT; 3 [22,26,31]	schwerwiegend ^g	nicht schwerwiegend	nicht schwerwiegend	unentdeckt	nicht schwerwiegend	keine	369/1322	330/2083	10	1,57 [1,25; 1,96]	5 [2; 8]	moderat
<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. In White 2018 (CARES) war das ITT-Prinzip nicht adäquat umgesetzt. In Mackenzie 2020 (FAST) war die Verblindung nicht gegeben.</p> <p>e. Bedeutsame Heterogenität und Effektschätzungen zeigen in unterschiedliche Richtungen, daher erfolgte keine Darstellung eines gepoolten Effekts.</p> <p>f. Aufgrund der sehr schwerwiegenden Inkonsistenz besteht große Unsicherheit bezüglich des Therapieeffekts. Hier kann die Risikodifferenz keinen Beitrag zur Abwägung von Vor- oder Nachteilen der Therapie leisten und wird daher (sowie auch das Basisrisiko) nicht geschätzt.</p> <p>g. In O'Dell 2022 war die adäquate Erzeugung der Randomisierungssequenz unklar. In O'Dell 2022 und Desideri 2022 waren die Verdeckung der Gruppenzuteilung unklar und das ITT-Prinzip nicht adäquat umgesetzt. In Desideri 2022 war die Verblindung nicht gegeben.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; ITT: Intention to treat; k. A.: keine Angabe; KI: Konfidenzintervall; SUE: schwerwiegendes unerwünschtes Ereignis; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

Tabelle 13: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für weitere unerwünschte Ereignisse – Gichtanfälle (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	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Allopurinol-Gruppe.												
Nachbeobachtung median 1467 Tage												
RCT; 1 [29]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	1044/3050	1017/3001	34	1,02 [0,91; 1,13]	0 [-2; 3]	moderat
Nachbeobachtung 36 Wochen												
RCT; 1 [26]	schwer-wiegend ^e	nicht zutreffend	nicht schwer-wiegend	unentdeckt	sehr schwer-wiegend ^f	keine	8/98	6/98	6	1,36 [0,45; 4,09]	2 [-6; 10]	sehr niedrig
Nachbeobachtung innerhalb von 24 Wochen (Woche 49 bis Woche 72)												
RCT; 1 [31]	schwer-wiegend ^g	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	135/370	165/379	44	0,75 [0,56; 1,00]	-7 [-14; 0]	moderat

Tabelle 13: Evidenzprofil für den Vergleich Allopurinol versus Febuxostat für weitere unerwünschte Ereignisse – 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	Indirektheit	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 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. In Mackenzie 2020 (FAST) war die Verblindung nicht gegeben.</p> <p>e. In Desideri 2022 war die Verdeckung der Gruppenzuteilung unklar, die Verblindung nicht gegeben sowie die Umsetzung des ITT-Prinzips nicht adäquat.</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>g. In O'Dell 2022 waren die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung unklar. Zudem war das ITT-Prinzips nicht adäquat umgesetzt.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; ITT: Intention to treat; KI: Konfidenzintervall; SUE: schwerwiegendes unerwünschtes Ereignis; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

5.4.3.1 Metaanalysen

Allopurinol vs. Febuxostat
 SUE

Modell mit festem Effekt - Mantel-Haenszel (zur Darstellung der Gewichte)

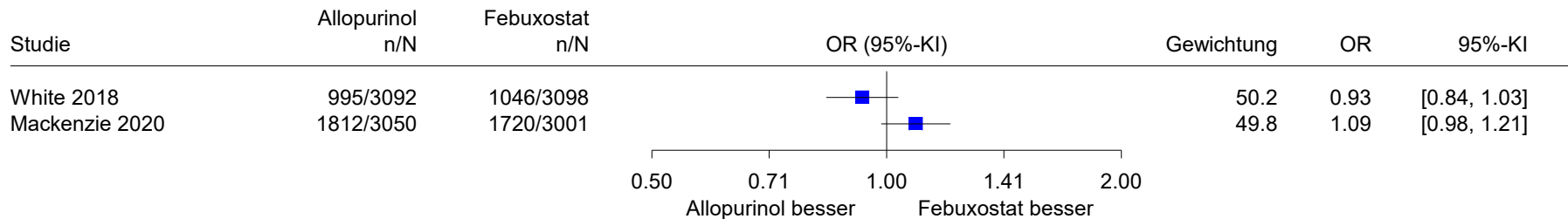


Abbildung 6: Metaanalyse, Allopurinol versus Febuxostat, SUE, Nachbeobachtung maximal 6 Jahre; Effektmaß: OR

Allopurinol vs. Febuxostat
 SUE

Modell mit festem Effekt - Mantel-Haenszel

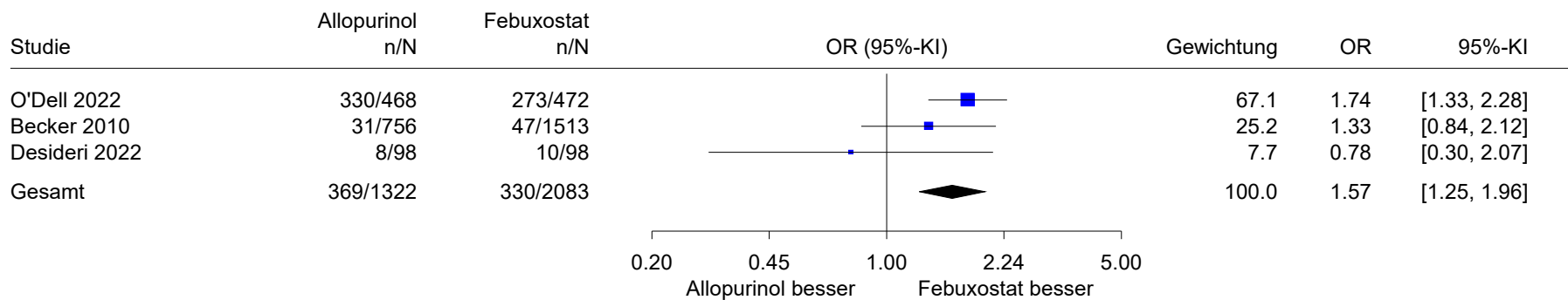


Abbildung 7: Metaanalyse, Allopurinol versus Febuxostat, SUE, Nachbeobachtung 28 bis 72 Wochen; Effektmaß: OR

6 Zusammenfassung und Anmerkungen zu den Ergebnissen

6.1 Zusammenfassung der Ergebnisse

Für den Vergleich von Allopurinol mit Febuxostat wurden insgesamt 5 Studien identifiziert. 4 Studien lieferten verwertbare Daten für alle 3 kritischen Endpunkte. 1 Studie lieferte Daten für 2 Endpunkte (Gesamtmortalität und weitere unerwünschte Ereignisse). Aus 2 Studien mit insgesamt etwa 12000 Patientinnen und Patienten lagen Ergebnisse für eine mediane Nachbeobachtung von 1467 Tagen (ca. 48 Monate; Mackenzie 2020 [FAST]) bzw. im Median 32 Monaten vor (White 2018 [CARES]). Aus 3 weiteren Studien mit insgesamt etwa 3400 Patientinnen und Patienten lagen Ergebnisse für eine Nachbeobachtung von 26 bis 72 Wochen vor (6 Monate bis ca. 17 Monate; Becker 2010, Desideri 2022 und O'Dell 2022).

Langzeitergebnisse (Mackenzie 2020 [FAST] und White 2018 [CARES])

In diesen beiden Studien weisen die Ergebnisse für die Gesamtmortalität, MACE und schwerwiegende unerwünschte Ereignisse (SUEs) jeweils in verschiedene Richtungen. Während sich für den Endpunkt MACE bei der metaanalytischen Zusammenfassung kein statistisch signifikanter Unterschied zeigte, war die Heterogenität zwischen den Studienergebnissen für die Endpunkte Gesamtmortalität und SUEs so groß (keine bzw. kaum Überschneidung der Konfidenzintervalle), dass die Ergebnisse nicht sinnvoll gepoolt werden konnten. Für Gichtanfälle lag nur in Mackenzie 2020 (FAST) ein verwertbares Ergebnis vor; hier zeigte sich kein statistisch signifikanter Unterschied. In der Betrachtung dieser Analysen kann somit weder ein Vorteil noch ein Nachteil für Febuxostat ausgeschlossen werden. Das Vertrauen in die Ergebnisse für die Gesamtmortalität und SUEs ist aufgrund der schwerwiegenden Studienlimitationen sowie der sehr schwerwiegenden Inkonsistenz sehr gering. Das Vertrauen in die Effektschätzungen für MACE und Gichtanfälle ist dagegen allein aufgrund der schwerwiegenden Studienlimitationen mäßig.

Kurzzeitergebnisse (Becker 2010, Desideri 2022 und O'Dell 2022)

Für die Gesamtmortalität (alle 3 Studien) und MACE (Becker 2010 und O'Dell 2022) zeigte sich jeweils kein statistisch signifikanter Unterschied. Für SUEs zeigte sich ein statistisch signifikanter Unterschied zugunsten von Febuxostat. Hinsichtlich der Gichtanfälle ergab sich kein klares Bild: In O'Dell 2022 zeigte sich ein knapp statistisch signifikanter Unterschied zugunsten von Allopurinol, während in Desideri 2022 bei einer sehr unpräzisen Schätzung kein statistisch signifikanter Unterschied gefunden wurde. Das Vertrauen in die Effektschätzungen ist für die Endpunkte Gesamtmortalität und MACE aufgrund der schwerwiegenden Studienlimitationen und schwerwiegender fehlender Genauigkeit begrenzt. Das Vertrauen in die Effektschätzung für SUEs ist aufgrund der schwerwiegenden Studienlimitationen mäßig. Das Vertrauen in die Effektschätzung für Gichtanfälle ist für die Ergebnisse aus O'Dell aufgrund schwerwiegender Studienlimitationen mäßig. Für die Ergebnisse aus der Studie Desideri hingegen ist das Vertrauen in die Effektschätzung für Gichtanfälle sehr gering, da hier neben

schwerwiegenden Studienlimitationen auch eine sehr schwerwiegende fehlende Genauigkeit vorliegt.

6.2 Anmerkungen zu den Ergebnissen

Interpretation der heterogenen Langzeitergebnisse von Mackenzie 2020 (FAST) und White 2018 (CARES)

Bei der Interpretation der heterogenen Langzeitergebnisse von Mackenzie 2020 (FAST) und White 2018 (CARES) ist zu berücksichtigen, dass sich die Studien hinsichtlich mehrerer Aspekte deutlich unterscheiden, insbesondere hinsichtlich der Population, des Anteils der Therapieabbrecher sowie des Anteils der Studienabbrecher:

- In White 2018 (CARES) wiesen alle Patientinnen und Patienten eine kardiovaskuläre Komorbidität auf, in Mackenzie 2020 (FAST) zu Studienbeginn dagegen nur ca. ein Drittel der Patientinnen und Patienten.
- In Mackenzie 2020 (FAST) wurden Patientinnen und Patienten mit schwerer Herzinsuffizienz (NYHA III oder IV) ausgeschlossen, in White 2018 (CARES) nicht.
- In White 2018 (CARES) betrug der Anteil der Patientinnen und Patienten, die die Therapie vorzeitig abbrachen, in beiden Gruppen 55 %, in Mackenzie 2020 (FAST) dagegen in der Febuxostat-Gruppe ca. 32 % und in der Allopurinol-Gruppe ca. 17 %.
- In White 2018 (CARES) betrug der Anteil der Patientinnen und Patienten, die die Studie vorzeitig abbrachen, in beiden Gruppen ca. 45 %, in Mackenzie 2020 (FAST) dagegen in beiden Gruppen ca. 6 %.

Anmerkungen zur Zulassung von Allopurinol und Febuxostat

In den eingeschlossenen Studien wurden für Allopurinol und Febuxostat teilweise Dosierungen verwendet, die in Deutschland nicht zugelassen sind. Zugelassen sind in Deutschland für Allopurinol Dosierungen von 100 mg pro Tag bis 800 mg pro Tag [37], für Febuxostat die Dosierungen 80 mg pro Tag und 120 mg pro Tag [38]. In Mackenzie 2020 (FAST) war auch eine Allopurinol-Dosis von 900 mg pro Tag erlaubt. Jedoch wurden mit dieser Dosis weniger als 5 % der Patientinnen und Patienten behandelt. In den 3 US-amerikanischen Studien White 2018 (CARES), Becker 2010 und O'Dell 2022 wurden jeweils mindestens 50 % der Patientinnen und Patienten mit 40 mg Febuxostat pro Tag behandelt. In den USA ist diese Dosierung für Febuxostat zugelassen [39,40].

Aufgrund der Ergebnisse von White 2018 (CARES) haben die Zulassungsinhaber von febuxostathaltigen Arzneimitteln in Abstimmung mit der Europäischen Arzneimittel-Agentur (EMA) und dem Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) in Deutschland in einem Rote-Hand-Brief gewarnt, dass eine Therapie mit Febuxostat im Vergleich zu Allopurinol bei Patientinnen und Patienten mit Gicht und bestehender kardiovaskulären

Komorbidität mit einem erhöhten Risiko bezüglich der Gesamtmortalität sowie der kardiovaskulären Mortalität einhergeht [41]. Zudem wurde in diesem Rote-Hand-Brief darauf hingewiesen, dass bei Patientinnen bzw. Patienten mit bestehender schwerer kardiovaskulärer Erkrankung (z. B. Myokardinfarkt, Schlaganfall oder instabile Angina Pectoris) die Behandlung mit Febuxostat vermieden werden sollte, es sei denn, es bestehen keine anderen Therapiemöglichkeiten.

Aus gleichem Anlass hat die Food and Drug Administration (FDA) 2019 in den USA die Packungsbeilage von Febuxostat mit der gleichen Warnung versehen [35,42]. Zudem wurde die Zulassung von Febuxostat auf Patientinnen und Patienten eingeschränkt, deren Gicht mit Allopurinol nicht effektiv behandelt werden kann und / oder bei denen unter Therapie mit Allopurinol schwere Nebenwirkungen auftreten. Die FDA hat diese Warnung in der Packungsbeilage von Febuxostat auch nach der Publikation der Ergebnisse von FAST, die im Kontrast zu den Ergebnissen von CARES stehen, beibehalten.

Anmerkungen zur Zulassung von Colchicin für die Gichtanfallprophylaxe

Colchicin ist in Deutschland nur dann für die Gichtanfallprophylaxe zugelassen, wenn nicht steroidale Antirheumatika (NSAR) kontraindiziert sind oder nicht vertragen werden [43]. Dem entgegen wurde in allen eingeschlossenen Studien für die Gichtanfallprophylaxe 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.

Darstellung weiterer unerwünschter Ereignisse

Die Darstellung der Ergebnisse zu weiteren unerwünschten Ereignissen wurde begrenzt auf SUEs sowie das krankheitsspezifische unerwünschte Ereignis Gichtanfälle. Die Begrenzung auf SUEs und Gichtanfälle erfolgte, da weitere spezifische UE in der Abwägung von Vor- und Nachteilen der Interventionen angesichts der großen Heterogenität in schwerwiegenderen Endpunkten (Gesamtmortalität und SUEs) und der daraus folgenden großen Unsicherheit bezüglich des Therapieeffekts nur eine nachgeordnete Rolle spielen.

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. S3-Leitlinie Diagnostik und Therapie der Gicht; angemeldetes Leitlinienvorhaben [online]. 2021 [Zugriff: 21.02.2023]. URL: <https://register.awmf.org/de/leitlinien/detail/060-005#anmeldung>.
3. Schünemann H, Brożek J, Guyatt G et al. GRADE Handbook [online]. 2013 [Zugriff: 19.02.2021]. URL: <https://gdt.grade.pro.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. Balshem 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 Gesundheitswes* 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. 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>.
21. 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>.
22. 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>.

23. 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>.
24. 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>.
25. Takeda. Efficacy and Safety of Oral Febuxostat in Participants With Gout (CONFIRMS) [online]. 2012 [Zugriff: 03.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT00430248>.
26. 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>.
27. Menarini International Operations Luxembourg. Intensive Urate Lowering Therapy of Febuxostat Compared to Allopurinol on Cardiovascular Risk in Patients With Gout (FORWARD) [online]. 2019 [Zugriff: 03.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT02500641>.
28. Menarini International Operations Luxembourg. The Effect of Intensive Urate Lowering Therapy (ULT) with Febuxostat in Comparison with Allopurinol on Cardiovascular Risk in Patients with Gout Using Surrogate Markers: a Randomized, Controlled Trial (Acronym: the FORWARD Trial) [online]. [Zugriff: 03.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005567-33.
29. 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).
30. University of Dundee. FAST- Febuxostat versus Allopurinol Streamlined Trial; A prospective, randomised, open-label, blinded endpoint (PROBE) clinical trial evaluating long term cardiovascular safety of febuxostat in comparison with allopurinol in patients with chronic symptomatic hyperuricaemia [online]. [Zugriff: 03.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-001883-23.
31. 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>.

32. VA Office of Research and Development. CSP594 Comparative Effectiveness in Gout: Allopurinol vs. Febuxostat [online]. 2022 [Zugriff: 03.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT02579096>.
33. 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>.
34. Takeda. Cardiovascular Safety of Febuxostat and Allopurinol in Participants With Gout and Cardiovascular Comorbidities (CARES) [online]. 2018 [Zugriff: 03.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01101035>.
35. Food and Drug Administration. FDA adds Boxed Warning for increased risk of death with gout medicine Uloric (febuxostat) [online]. 2019 [Zugriff: 21.06.2023]. URL: <https://www.fda.gov/drugs/drug-safety-and-availability/fda-adds-boxed-warning-increased-risk-death-gout-medicine-uric-febuxostat>.
36. Perk J, De Backer G, Gohlke H et al. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). Eur Heart J 2012; 33(13): 1635-1701. <https://dx.doi.org/10.1093/eurheartj/ehs092>.
37. Maucher IV. Allopurinol [online]. 2019 [Zugriff: 24.07.2023]. URL: https://www.gelbe-liste.de/wirkstoffe/Allopurinol_351.
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. Food and Drug Administration. ULORIC; Labeling-Package Insert [online]. 2023 [Zugriff: 20.07.2023]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/021856s015lbl.pdf.
40. Food and Drug Administration. Drug Approval Package; Uloric (Febuxostat) Tablets [online]. 2009 [Zugriff: 20.07.2023]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/021856s000toc.cfm.
41. Bundesinstitut für Arzneimittel und Medizinprodukte. Rote-Hand-Brief zu Adenuric (Febuxostat) und anderen febuxostathaltigen Arzneimitteln: Erhöhtes Risiko für kardiovaskulär bedingte Mortalität und Gesamtmortalität [online]. 2019 [Zugriff: 20.07.2023]. URL: <https://www.bfarm.de/SharedDocs/Risikoinformationen/Pharmakovigilanz/DE/RHB/2019/rhb-febuxostat.html>.

42. Food and Drug Administration. FDA Drug Safety Communication: FDA to evaluate increased risk of heart-related death and death from all causes with the gout medicine febuxostat (Uloric) [online]. 2018 [Zugriff: 21.06.2023]. URL: <https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-evaluate-increased-risk-heart-related-death-and-death-all-causes>.
43. 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>.
44. 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.
45. 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.mayocpiqo.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 B, Zhang P, Li X. Efficacy and safety of Febuxostat Versus Allopurinol in Hyperuricemic patients with or without Gout: A meta-analysis. *Neuroendocrinology Letters* 2020; 41(4): 195-204.
4. 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>.
5. Guan X, Zhang S, Liu J et al. Cardiovascular safety of febuxostat and allopurinol in patients with gout: A meta-analysis. *Front Pharmacol* 2022; 13: 998441. <https://dx.doi.org/10.3389/fphar.2022.998441>.
6. 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>.
7. Wang M, Zhang Y, Zhang M et al. The major cardiovascular events of febuxostat versus allopurinol in treating gout or asymptomatic hyperuricemia: a systematic review and meta-analysis. *Annals of Palliative Medicine* 2021; 10(10): 10327-10337. <https://dx.doi.org/10.21037/apm-21-1564>.
8. 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>.
9. 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. 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>.
2. 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>.
3. 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>.
4. 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>.
5. 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>.
6. 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>.
7. 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>.
8. 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>.

9. 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.
10. 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>.
11. 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.
12. 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>.
13. 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>.
14. 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, 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>.
2. 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>.
3. 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.

4. 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>.
5. 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>.
6. 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>.
7. 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>.
8. 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>.
9. 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>.
10. 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.e1127. <https://dx.doi.org/10.1016/j.amjmed.2012.05.025>.

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. <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 14: Kriterien des Verzerrungspotenzials zum Endpunkt Gesamtmortalität

Studie	Adäquate Erzeugung der Randomisierungssequenz	Adäquate Zuteilungsverdeckung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Becker 2010	ja	ja	unklar	unklar	unklar	ja
Desideri 2022	ja	unklar	nein	nein	unklar	ja
Mackenzie 2020 (FAST)	ja	ja	nein	ja	ja	ja
O'Dell 2022	unklar	unklar	unklar	nein	ja	ja
White 2018 (CARES)	ja	ja	unklar	ja	ja	nein ^a
<p>a. Es wurden Todesfälle von Patientinnen und Patienten, die die Studie abgebrochen hatten, nachrecherchiert. Die Methodik und die Vollständigkeit der Recherche wurden nicht beschrieben. Die Auswertung, in der nachrecherchierte Todesfälle berücksichtigt wurden, liefert ein HR von 0,92 [0,78; 1,06]. Dieses Ergebnis weicht in der Signifikanzaussage von der primär geplanten Auswertung ab.</p> <p>ITT: Intention to treat; OR: Odds Ratio</p>						

Tabelle 15: Kriterien des Verzerrungspotenzials zum Endpunkt MACE

Studie	Adäquate Erzeugung der Randomisierungssequenz	Adäquate Zuteilungsverdeckung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Becker 2010	ja	ja	unklar	unklar	unklar	ja
Mackenzie 2020 (FAST)	ja	ja	nein	ja	ja	ja
O'Dell 2022	unklar	unklar	unklar	nein	ja	ja
White 2018 (CARES)	ja	ja	unklar	ja	ja	nein ^a
<p>a. Für den Endpunkt MACE kann nicht ausgeschlossen werden, dass sich die Effektschätzung ändern würde (vergleichbar wie beim Endpunkt Gesamtmortalität), wenn die Patientinnen und Patienten berücksichtigt werden könnten, die die Studie vorzeitig abbrachen.</p> <p>ITT: Intention to treat; MACE: Major adverse cardiovascular Events</p>						

Tabelle 16: Kriterien des Verzerrungspotenzials zum Endpunkt weitere unerwünschte Ereignisse

Studie	Adäquate Erzeugung der Randomisierungssequenz	Adäquate Zuteilungsverdeckung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Becker 2010	ja	ja	unklar	unklar	unklar	ja
Desideri 2022	ja	unklar	nein	nein	unklar	ja
Mackenzie 2020 (FAST)	ja	ja	nein	ja	ja	ja
O'Dell 2022	unklar	unklar	unklar	nein	ja	ja
White 2018 (CARES)	ja	ja	unklar	nein	ja	ja
ITT: Intention to treat						

Anhang B Übersicht über die Ergebnisse der MACE-Einzelkomponenten

Tabelle 17: Ergebnisse – MACE-Einzelkomponenten

Studie MACE-Einzelkomponente	Intervention vs. Vergleich	
	Effektmaß: Punktschätzer	[95 %-KI]
Becker 2010		
Kardiovaskuläre Mortalität	OR: 10,03	[0,48; 209,18]
Nicht tödlicher Herzinfarkt	OR: 2,00	[0,13; 32,06]
Nicht tödlicher Schlaganfall	OR: 0,40	[0,02; 8,33]
Mackenzie 2020 (FAST)		
Kardiovaskuläre Mortalität	HR: 1,04	[0,81; 1,35]
Hospitalisierung wegen nicht tödlichen Herzinfarkts oder Biomarker-positiven akuten koronaren Syndroms	HR: 1,08	[0,83; 1,40]
Nicht tödlicher Schlaganfall	HR: 1,09	[0,80; 1,47]
O'Dell 2022		
Kardiovaskuläre Mortalität	OR: 3,04	[0,31; 29,32]
Nicht tödlicher Herzinfarkt	OR: 0,50	[0,09; 2,75]
Nicht tödlicher Schlaganfall	OR: 0,50	[0,05; 5,57]
Instabile Angina Pectoris	OR: 1,35	[0,30; 6,05]
White 2018 (CARES)		
Kardiovaskuläre Mortalität	HR: 0,75	[0,58; 0,97]
Nicht tödlicher Herzinfarkt	HR: 1,08	[0,83; 1,39]
Nicht tödlicher Schlaganfall	HR: 0,99	[0,71; 1,37]

HR: Hazard Ratio; KI: Konfidenzintervall; MACE: Major adverse cardiovascular Events; OR: Odds Ratio

Anhang C Suchstrategien

C.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 [44] – 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 [45] – 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

C.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]

C.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