

Medikamentöse Behandlung eines Gichtanfalls

Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht

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

EVIDENZBERICHT

Projekt: V22-06B

Version: 1.0

Stand: 22.09.2023

IQWiG-Berichte – Nr. 1636

DOI: 10.60584/V22-06B

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Medikamentöse Behandlung eines Gichtanfalls – Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht

Auftraggeber

Bundesministerium für Gesundheit

Datum des Auftrags

20.12.2022

Interne Projektnummer

V22-06B

DOI-URL

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

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. Medikamentöse Behandlung eines Gichtanfalls; Evidenzbericht zur S3-Leitlinie Diagnostik und Therapie der Gicht [online]. 2023 [Zugriff: TT.MM.JJJJ]. URL: <https://dx.doi.org/10.60584/V22-06B>.

Schlagwörter

Antiphlogistika – nichtsteroidale, Interleukin-1, Colchicin, Glukokortikoide, Gicht, Systematische Übersicht

Keywords

Anti-Inflammatory Agents – Non-Steroidal, Interleukin-1, Colchicine, Glucocorticoids, Gout, Systematic Review

Dieser Bericht wurde ohne die Beteiligung externer Sachverständiger erstellt.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Alexandra Korzeczek-Opitz
- Tatjana Hermanns
- Corinna Kiefer
- Anna Margraf
- Claudia Mischke
- Andrea Tasar

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vii
Abbildungsverzeichnis	ix
Abkürzungsverzeichnis.....	x
1 Hintergrund.....	1
2 Fragestellung.....	2
3 Projektverlauf.....	3
3.1 Zeitlicher Verlauf des Projekts	3
3.2 Spezifizierungen und Änderungen im Projektverlauf.....	3
4 Methodik gemäß Projektskizze	5
4.1 Tabellarische Darstellung der Kriterien für den Studieneinschluss	5
4.1.1 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf	6
4.1.2 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	6
4.2 Informationsbeschaffung.....	6
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten	6
4.2.2 Fokussierte Informationsbeschaffung von Studien	7
4.2.3 Anwendung von Limitierungen auf Datenbankebene	8
4.2.4 Selektion relevanter Studien	8
4.3 Informationsdarstellung und Synthese.....	9
4.3.1 Darstellung der Einzelstudien.....	9
4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial).....	10
4.3.3 Metaanalysen	12
4.3.4 Subgruppenmerkmale und andere Effektmodifikatoren.....	13
4.3.5 Bewertung der Qualität der Evidenz	13
4.3.5.1 Abwertung der Qualität der Evidenz	14
4.3.5.2 Aufwertung der Qualität der Evidenz	16
5 Ergebnisse	17
5.1 Ergebnisse der Informationsbeschaffung	17
5.2 Charakteristika der für die Evidenzdarstellung berücksichtigten Studien	19
5.3 Übersicht der Endpunkte	34
5.4 Ergebnisse zu Endpunkten	35
5.4.1 Ergebnisse zu dem Vergleich Colchicin vs. Placebo	35

5.4.1.1	Ergebnisse zum kritischen Endpunkt Schmerzen	35
5.4.1.2	Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse.....	37
5.4.2	Ergebnisse zu dem Vergleich NSAR vs. Colchicin	40
5.4.2.1	Ergebnisse zum kritischen Endpunkt Schmerzen	40
5.4.2.2	Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse.....	43
5.4.2.3	Ergebnisse zum wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse.....	48
5.4.3	Ergebnisse zu dem Vergleich NSAR vs. Glukokortikoide.....	49
5.4.3.1	Ergebnisse zum kritischen Endpunkt Schmerzen	49
5.4.3.2	Ergebnisse zum wichtigen Endpunkt Funktionsfähigkeit	52
5.4.3.2.1	Metaanalysen.....	53
5.4.3.3	Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse.....	54
5.4.3.4	Ergebnisse zum wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse.....	55
5.4.3.4.1	Metaanalysen.....	56
5.4.4	Ergebnisse zu dem Vergleich IL-1-Inhibitoren vs. Glukokortikoide	57
5.4.4.1	Ergebnisse zum kritischen Endpunkt Schmerzen	57
5.4.4.1.1	Metaanalysen.....	58
5.4.4.2	Ergebnisse zum wichtigen Endpunkt Funktionsfähigkeit	59
5.4.4.2.1	Metaanalysen.....	61
5.4.4.3	Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse.....	62
5.4.4.3.1	Metaanalysen.....	65
6	Zusammenfassung und Anmerkungen zu den Ergebnissen	66
6.1	Zusammenfassung der Ergebnisse	66
6.2	Anmerkungen zu den Ergebnissen	67
7	Literatur	69
8	Studienlisten	74
8.1	Liste der identifizierten systematischen Übersichten.....	74
8.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	74
Anhang A	Bewertung der Qualität der Informationsbeschaffung der Basis-SÜ aus der fokussierten Recherche	75
Anhang B	Bewertung der Qualität der Ergebnisse der Basis-SÜ aus der fokussierten Recherche	76
Anhang C	Kriterien des Verzerrungspotenzials.....	77
Anhang D	Suchstrategien	79
D.1	Bibliografische Datenbanken	79
D.2	Studienregister.....	81

D.3 Weitere Informationsquellen und Suchtechniken 82

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	5
Tabelle 2: Berücksichtigte systematische Übersichten.....	17
Tabelle 3: Studienpool der eingeschlossenen Studien	19
Tabelle 4: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien	20
Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien.....	21
Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien.....	25
Tabelle 7: Charakterisierung der Studienpopulationen.....	32
Tabelle 8: Matrix der einbezogenen Endpunkte.....	34
Tabelle 9: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen).....	35
Tabelle 10: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten).....	37
Tabelle 11: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen)	38
Tabelle 12: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen).....	40
Tabelle 13: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den kritischen Endpunkt Schmerzen (stetige Daten) (aus NICE 2022 übernommen)	42
Tabelle 14: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten).....	43
Tabelle 15: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen)	44
Tabelle 16: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse (binäre Daten)	48
Tabelle 17: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen).....	49
Tabelle 18: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (stetige Daten) (aus NICE 2022 übernommen)	51
Tabelle 19: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt Funktionsfähigkeit (stetige Daten).....	52
Tabelle 20: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten).....	54
Tabelle 21: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse (binäre Daten)	55
Tabelle 22: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (stetige Daten)	57

Tabelle 23: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt Funktionsfähigkeit (stetige Daten).....	59
Tabelle 24: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten).....	62
Tabelle 25: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen).....	63
Tabelle 26: Bewertung der Informationsbeschaffung zur systematischen Übersicht NICE 2022	75
Tabelle 27: Bewertung der Ergebnisse zur systematischen Übersicht NICE 2022	76
Tabelle 28: Kriterien des Verzerrungspotenzials zum Endpunkt Schmerzen	77
Tabelle 29: Kriterien des Verzerrungspotenzials zum Endpunkt Funktionsfähigkeit	77
Tabelle 30: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse – SUE	78
Tabelle 31: Kriterien des Verzerrungspotenzials zum Endpunkt Therapieabbruch aufgrund von UE	78

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung	18
Abbildung 2: Metaanalyse, NSAR vs. Glukokortikoide, Endpunkt Funktionsfähigkeit – Kurzzeiteffekt, Effektmaß: Hedges' g	53
Abbildung 3: Metaanalyse, NSAR vs. Glukokortikoide, Endpunkt Therapieabbruch aufgrund von unerwünschten Ereignissen – Kurzzeiteffekt, Effektmaß: OR	56
Abbildung 4: Metaanalyse, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt Schmerzen – Kurzzeiteffekt, Effektmaß: OR	58
Abbildung 5: Forest Plot, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt Funktionsfähigkeit – Langzeiteffekt, Effektmaß: Mittelwertdifferenz.....	61
Abbildung 6: Metaanalyse, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt unerwünschte Ereignisse – Langzeiteffekt, Effektmaß: OR.....	65

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
BMG	Bundesministerium für Gesundheit
DGRh	Deutsche Gesellschaft für Rheumatologie e. V.
DVG	Digitale-Versorgung-Gesetz
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IL-1-Inhibitoren	Interleukin-1-Inhibitoren
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KI	Konfidenzintervall
MWD	Mittelwertdifferenz
NICE	National Institute for Health and Care Excellence
NSAR	nicht steroidale Antirheumatika
OR	Odds Ratio
PICO	Population, Intervention, Comparison, Outcome
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	absolute Risikodifferenz
SMD	Standardized Mean Difference (standardisierte Mittelwertdifferenz)
SÜ	systematische Übersicht
SUE	schwerwiegendes unerwünschtes Ereignis

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 medikamentösen Behandlung mit nicht steroidal Antirheumatika (NSAR), Interleukin(IL)-1-Inhibitoren, Glukokortikoiden oder Colchicin im Vergleich zu Placebo oder einer medikamentösen Behandlung mit einer anderen Wirkstoffgruppe (NSAR, IL-1-Inhibitoren, Glukokortikoiden oder Colchicin) bei erwachsenen Patientinnen und Patienten mit einem Gichtanfall.

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 medikamentösen Behandlung in der Prüf- und Vergleichsintervention wurde folgende Spezifizierung vorgenommen: Voraussetzung für die Evidenzdarstellung von Arzneimitteln ist zudem die Marktverfügbarkeit in Deutschland.
- Abschnitt 4.2.1: Die Bewertung der Qualität der Informationsbeschaffung einer hochwertigen und aktuellen systematischen Übersicht (SÜ), deren Informationsbeschaffung als Grundlage für die Evidenzdarstellung verwendet werden kann (Basis-SÜ), erfolgte mit A Measurement Tool to Assess Systematic Reviews (AMSTAR) 2 [3]. Die Bewertung erfolgte durch 1 Person und wurde von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.
- Abschnitt 4.2.2: Als weitere Informationsquelle wurden die Webseiten des IQWiG und Gemeinsamen Bundesausschusses (G-BA) durchsucht.
- Abschnitt 4.2.3: Die fokussierte Informationsbeschaffung von SÜs wurde auf das Publikationsdatum ab 01/2020 eingeschränkt.

- Abschnitt 4.3.1: Die Basis-SÜ wurde auch als Grundlage für die Evidenzdarstellung verwendet. Hierzu wurden Ergebnisse zu Studieninformationen und Endpunkten übernommen. Vorab erfolgte eine Prüfung der methodischen Qualität der Basis-SÜ mit AMSTAR 2. Die Bewertung erfolgt durch 1 Person und wurde von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst. Aus der Basis-SÜ übernommene Studien und Informationen wurden im vorliegenden Bericht kursiv dargestellt.
- 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) [4] und ist konform zu den Allgemeinen Methoden [5]. 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 einem Gichtanfall (M10.09G)
E2	Prüfintervention: medikamentöse Behandlung mit NSAR ^a , IL-1-Inhibitoren ^a , Glukokortikoiden ^a oder Colchicin ^a
E3	Vergleichsintervention: Placebo oder medikamentöse Behandlung mit einer anderen Wirkstoffgruppe (NSAR ^a , IL-1-Inhibitoren ^a , Glukokortikoide ^a oder Colchicin ^a)
E4	kritischer Endpunkt: Schmerzen wichtige Endpunkte: Funktionsfähigkeit (z. B. HAQ-DI, SF-36), unerwünschte Ereignisse, Therapieabbruch aufgrund von unerwünschten Ereignissen
E5	Studientyp: RCTs
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar ^b
<p>a. 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.</p> <p>b. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [6] oder ein Bericht über die Studie, der den Kriterien des CONSORT- Statements [7] 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; HAQ-DI: Health Assessment Questionnaire; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie; SF-36: 36-item Short-Form Health Survey</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 SÜs in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), der International Health Technology Assessment (HTA) Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

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

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).

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

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

- Autorenanfragen

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 [5] 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 [8]. Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen [5] 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.

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

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

Das Vorgehen zur Bewertung des Einflusses der Studienlimitationen auf die berichtsrelevanten Ergebnisse wird in Abschnitt 4.3.2 endpunktspezifisch pro Studie beschrieben. Nach dieser Bewertung werden die Informationen zu den ausgewählten Endpunkten in Evidenzprofilen zusammengeführt und analysiert [9]. 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 [10] angegeben. Sollte das Ergebnis des CSZ-Tests [11] 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 [12] 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 Der Simonian-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 [13]. Als Ergebnis wird der gemeinsame Effekt mit Konfidenzintervall im Forest Plot dargestellt. Da die Heterogenität im Fall sehr weniger (2 bis 4) Studien nicht verlässlich geschätzt werden kann, wird im Fall von 2 Studien aus pragmatischen Gründen ein Modell mit festem Effekt verwendet, sofern keine deutlichen Gründe dagegen sprechen. Bei 3 und 4 Studien erfolgt vorzugsweise eine Metaanalyse nach der Methode von Knapp und Hartung, sofern das KI informativ ist und kein Widerspruch zu einer qualitativen Zusammenfassung besteht. Andernfalls wird ein Modell mit festem Effekt verwendet, wenn keine deutlichen Gründe dagegensprechen. Ist ein Modell mit festem Effekt nicht vertretbar, bleibt es bei der Schätzung nach Knapp und Hartung.

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

4.3.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In den Evidenzbericht werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-regressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter.

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

4.3.5 Bewertung der Qualität der Evidenz

Alle für den Evidenzbericht relevanten Ergebnisse werden hinsichtlich einer Beeinflussung durch Faktoren, die zu einer Abwertung oder ggf. Aufwertung der Qualität der Evidenz führen können, überprüft. Für jeden für die Entscheidung kritischen und wichtigen Endpunkt wird eine studienübergreifende Aussage zur Qualität der Evidenz bezüglich des jeweiligen Vertrauens in die Effektschätzung getroffen. Hierzu erfolgt eine Einteilung der Qualität der Evidenz entsprechend der 4 Stufen der GRADE-Guideline in hoch, moderat, niedrig und sehr niedrig [14,15]:

- 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 [14]. 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 [16].

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 [17].

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 [18].

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 [19]. 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) [20].

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 [20].

C: Confounding

Die Qualität der Evidenz wird bei einer möglichen Verstärkung des beobachteten Effektes unter Berücksichtigung von potenziellen Confoundern oder Verzerrungsquellen regelhaft aufgewertet [20].

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 28.04.2023) und die Suche in Studienregistern (letzte Suche am 27.03.2023) sind in Anhang D dargestellt.

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

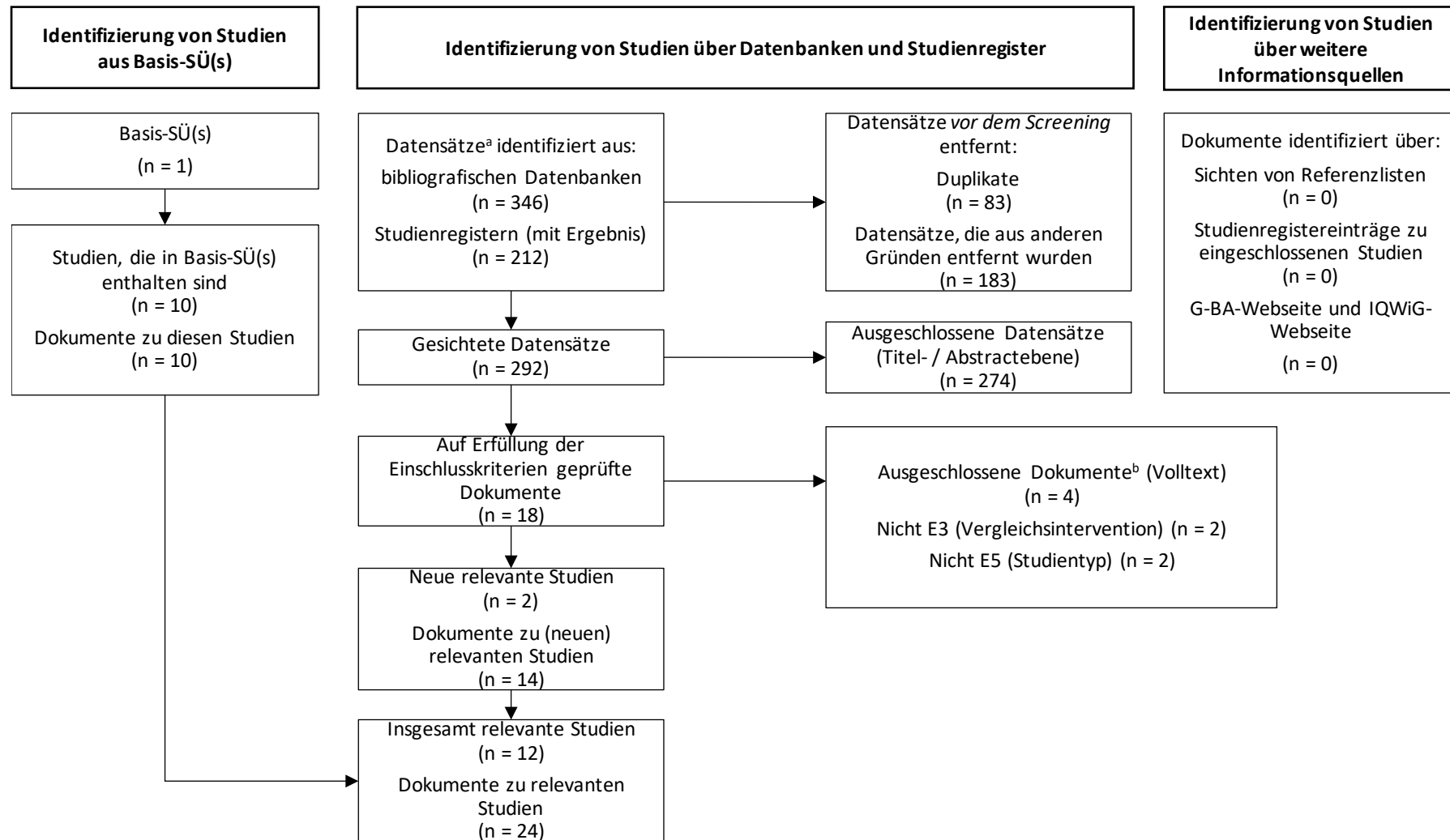
Autorenanfragen wurden nicht gestellt.

Von den 5 identifizierten SÜs (siehe Abschnitt 8.1) wurde 1 (siehe Tabelle 2) als Basis-SÜ berücksichtigt. Die Bewertung der Qualität der Informationsbeschaffung dieser Basis-SÜ findet sich in Anhang A. Die Bewertung der Qualität der Ergebnisse dieser Basis-SÜ findet sich in Anhang B.

Tabelle 2: Berücksichtigte systematische Übersichten

Dokument	Vollpublikation
NICE 2022	ja [21]
NICE: National Institute for Health and Care Excellence	

In der Basis-SÜ wurden 10 relevante Studien identifiziert, deren Ergebnisse zu den Studieninformationen und zu den Endpunkten Schmerzen und unerwünschte Ereignisse übernommen wurden. 2 Studien wurden aufgrund einer nicht passenden Prüfintervention (nicht medikamentöses Verfahren bzw. in Deutschland nicht zugelassener IL-1-Inhibitor-Wirkstoff) nicht übernommen. Es erfolgte eine ergänzende Suche nach Studien in bibliografischen Datenbanken für den Zeitraum (ab Juli 2021), der nicht durch die berücksichtigte Basis-SÜ abgedeckt war. Zu den Wirkstoffen Acemetacin und Phenylbutazon, welche nicht in der Basis-SÜ berücksichtigt wurden, fand eine systematische Recherche für den gesamten relevanten Zeitraum statt. Für alle weiteren Informationsquellen wurde ohne zeitliche Beschränkung (siehe Anhang D.2 und Anhang D.3) gesucht.



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 12 relevante Studien identifiziert (siehe Tabelle 3).

Tabelle 3: Studienpool der eingeschlossenen Studien

Vergleich Studie	Verfügbare Dokumente	
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern
Colchicin vs. Placebo		
<i>Terkeltaub 2010</i>	ja [22]	ja [23] / ja [23]
NSAR vs. Colchicin		
<i>Liu 2019</i>	ja [24]	nein
<i>Roddy 2020</i>	Ja [25]	ja [26] / ja [26]
NSAR vs. Glukokortikoide		
<i>Janssens 2008</i>	ja [27]	nein
<i>Man 2007</i>	ja [28]	nein
<i>Rainer 2016</i>	ja [29]	nein
<i>Xu 2016</i>	ja [30]	nein
IL-1-Inhibitoren vs. Glukokortikoide		
<i>So 2010</i>	ja [31,32]	ja [33,34] / ja [33,34]
<i>Schlesinger 2012a</i>	ja [35]	ja [36-38] / ja [36-38]
<i>Schlesinger 2012b</i>	ja [35]	ja [39-41] / ja [39-41]
NCT01356602	nein	ja [42,43] / ja [42,43]
NCT01362608	nein	ja [44,45] / ja [44,45]
Studien in kursiver Schrift wurden in der Basis-SÜ NICE 2022 eingeschlossen.		

5.2 Charakteristika der für die Evidenzdarstellung berücksichtigten Studien

In Tabelle 4 werden die Charakteristika der für die Evidenzdarstellung berücksichtigten Studien dargestellt. Die in den jeweiligen Studien untersuchten Interventionen werden in Tabelle 5 charakterisiert. Die wesentlichen Ein- und Ausschlusskriterien für Patientinnen und Patienten in den berücksichtigten Studien sowie die Charakteristika der jeweiligen Studienpopulationen sind in Tabelle 6 und Tabelle 7 aufgeführt.

Tabelle 4: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien

Studie	Studien- design	Patientenzahl N	Intervention	Vergleich	Ort und Zeitraum der Durchführung	Dauer der Nachbeobachtung
Colchicin vs. Placebo						
<i>Terkeltaub 2010</i>	<i>RCT</i>	<i>I: 74^a; C:59</i>	<i>Colchicin</i>	<i>Placebo</i>	<i>USA; 2007 bis 2008</i>	<i>7 Tage</i>
NSAR vs. Colchicin						
<i>Liu 2019</i>	<i>RCT</i>	<i>I: 61; C: 61</i>	<i>Etoricoxib</i>	<i>Colchicin</i>	<i>China; k. A.</i>	<i>10 Tage</i>
<i>Roddy 2020</i>	<i>RCT</i>	<i>I: 200; C: 199</i>	<i>Naproxen</i>	<i>Colchicin</i>	<i>Vereinigtes Königreich; 2014 bis 2015</i>	<i>7 Tage</i>
NSAR vs. Glukokortikoide						
<i>Janssens 2008</i>	<i>RCT</i>	<i>I: 60; C: 60</i>	<i>Naproxen</i>	<i>Prednisolon</i>	<i>Niederlande; 2004 bis 2006</i>	<i>3 Wochen</i>
<i>Man 2007</i>	<i>RCT</i>	<i>I: 46; C: 44</i>	<i>Indometacin</i>	<i>Prednisolon</i>	<i>Hong Kong (China); 2003 bis 2004</i>	<i>14 Tage</i>
<i>Rainer 2016</i>	<i>RCT</i>	<i>I: 208; C:208</i>	<i>Indometacin</i>	<i>Prednisolon</i>	<i>Hong Kong (China); 2010 bis 2012</i>	<i>14 Tage</i>
<i>Xu 2016</i>	<i>RCT</i>	<i>I₁: 46; I₂: 45 C: 41</i>	<i>I1: Etoricoxib I2: Indometacin</i>	<i>Prednisolon</i>	<i>China; 2015 bis 2015</i>	<i>4 Tage</i>
IL-1-Inhibitoren vs. Glukokortikoide						
<i>So 2010</i>	<i>RCT</i>	<i>I: 28; C: 57</i>	<i>Canakinumab</i>	<i>Triamcinolon</i>	<i>Schweiz; 2008 bis 2009</i>	<i>7 Tage</i>
<i>Schlesinger 2012a</i>	<i>RCT</i>	<i>I: 115; C: 115</i>	<i>Canakinumab</i>	<i>Triamcinolon</i>	<i>USA; 2009 bis 2010</i>	<i>24 Wochen</i>
<i>Schlesinger 2012b</i>	<i>RCT</i>	<i>I: 112; C: 114</i>	<i>Canakinumab</i>	<i>Triamcinolon</i>	<i>USA; 2009 bis 2010</i>	<i>24 Wochen</i>
<i>NCT01356602</i>	<i>RCT</i>	<i>I: 133^b; C: 132</i>	<i>Canakinumab</i>	<i>Triamcinolon</i>	<i>international; 2011 bis 2012</i>	<i>12 Wochen</i>
<i>NCT01362608</i>	<i>RCT</i>	<i>I: 67; C: 69</i>	<i>Canakinumab</i>	<i>Triamcinolon</i>	<i>Argentinien, China, Polen, Singapur; 2011 bis 2015</i>	<i>12 Wochen</i>
<p>a. Der 3. Studienarm (Hochdosis Colchicin) wird hier nicht dargestellt, da dieser nicht der Arzneimittelzulassung in Deutschland entspricht.</p> <p>b. Der 3. Studienarm (Lyophilisat) wird hier nicht dargestellt, da dieser nicht der Arzneimittelzulassung in Deutschland entspricht.</p> <p>C: Vergleichsgruppe; I: Interventionsgruppe; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; k. A.: keine Angabe; N: Anzahl randomisierter Patientinnen und Patienten; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie</p>						

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien (mehrsseitige Tabelle)

Studie	Intervention	Vergleich	Begleittherapie
Colchicin vs. Placebo			
<i>Terkeltaub 2010</i>	<p>Wirkstoff: Colchicin</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ initial: 1,2 mg ▪ 0,6 mg nach 1 Stunde ▪ Gesamtdosis: 1,8 mg <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ 1 Kapsel alle 5 Stunden <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 24 Stunden 	<p>Wirkstoff: Placebo</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ initial: 2 Kapseln ▪ 1 Kapsel alle 6 Stunden <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 24 Stunden 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ NSAR (z. B. Indometacin) <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ kontinuierliche Einnahme (z. B. Allopurinol)
NSAR vs. Colchicin			
<i>Liu 2019</i>	<p>Wirkstoff: Etoricoxib</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ Tag 1–5: 120 mg/Tag ▪ Tag 6–10: 60 mg/Tag <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 10 Tage 	<p>Wirkstoff: Colchicin</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ Tag 1–5: 3x0,5 mg/Tag ▪ Tag 6–10: 0,5 mg/Tag <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 10 Tage 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ k. A. <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ k. A.
<i>Roddy 2020</i>	<p>Wirkstoff: Naproxen</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ initial: 750 mg ▪ 250 mg alle 8 Stunden <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 7 Tage 	<p>Wirkstoff: Colchicin</p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ oral <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ 500 µg alle 8 Stunden <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ 4 Tage 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ k. A. <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ k. A.

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien (mehreseitige Tabelle)

Studie	Intervention	Vergleich	Begleittherapie
NSAR vs. Glukokortikoide			
<i>Janssens 2008</i>	<p>Wirkstoff: <i>Naproxen</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>2 x 500 mg/Tag</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>1 Kapsel/Tag</i> <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p>Wirkstoff: <i>Prednisolon</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>35 mg/Tag</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>2x1 Kapsel/Tag</i> <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ <i>keine</i> <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ <i>k. A.</i>
<i>Man 2007</i>	<p>Wirkstoff: <i>Indometacin</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 50 mg</i> ▪ <i>Tag 1–2: 50 mg alle 8 Stunden</i> ▪ <i>Tag 3–5: 25 mg alle 8 Stunden</i> <p><u>zusätzliches Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 75 mg (25 mg/ml) Diclofenac intramuskulär</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 6 Tabletten</i> ▪ <i>Tag 1–5: 6 Tabletten/Tag</i> <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p>Wirkstoff: <i>Prednisolon</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 30 mg (6 5mg Tabletten)</i> ▪ <i>Tag 1–5: 30 mg/Tag (6 5mg Tabletten)</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 2 Tabletten</i> ▪ <i>initial: 3 ml Injektion, intramuskulär</i> <p>Tag 1–5: Placebo <u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ <i>initial: 1 g Paracetamol</i> ▪ <i>Tag 1–5: 1 g Paracetamol alle 6 Stunden nach Bedarf</i> <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ <i>k. A.</i>
<i>Rainer 2016</i>	<p>Wirkstoff: <i>Indometacin</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>Tag 1-2: 150 mg/Tag als 3x50 mg (je 2 25 mg Tabletten)</i> ▪ <i>Tag 3–5: 3x25 mg/Tag</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>Tag 1–2: 6 Tabletten/Tag</i> ▪ <i>Tag 3–5: 6 Tabletten/Tag</i> <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p>Wirkstoff: <i>Prednisolon</i></p> <p><u>Anwendungsart:</u></p> <ul style="list-style-type: none"> ▪ <i>oral</i> <p><u>Dosierung:</u></p> <ul style="list-style-type: none"> ▪ <i>Tag 1–2: 30 mg/Tag (je 3 10 mg Tabletten)</i> ▪ <i>Tag 1–3: 30 mg/Tag</i> <p><u>Placebo:</u></p> <ul style="list-style-type: none"> ▪ <i>Tag 1–2: 3x2 Tabletten/Tag</i> ▪ <i>Tag 3–5: 3x1 Tablette/Tag</i> <p><u>Dauer:</u></p> <ul style="list-style-type: none"> ▪ <i>5 Tage</i> 	<p><u>Schmerzmittel:</u></p> <ul style="list-style-type: none"> ▪ <i>1 g Paracetamol alle 6 Stunden nach Bedarf</i> <p><u>harnsäuresenkende Therapie:</u></p> <ul style="list-style-type: none"> ▪ <i>k. A.</i>

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien (mehreseitige Tabelle)

Studie	Intervention	Vergleich	Begleittherapie
<i>Xu 2016</i>	Wirkstoff: <i>Etoricoxib (I₁), Indometacin (I₂)</i> <u>Anwendungsart:</u> ▪ k. A. <u>Dosierung:</u> ▪ <i>Etoricoxib: 120 mg/Tag</i> ▪ <i>Indometacin: 3x 50 mg/Tag</i> <u>Dauer:</u> ▪ <i>4 Tage</i>	Wirkstoff: <i>Prednisolon</i> <u>Anwendungsart:</u> ▪ k. A. <u>Dosierung:</u> ▪ <i>35 mg/Tag</i> <u>Dauer:</u> ▪ <i>4 Tage</i>	<u>Schmerzmittel:</u> ▪ <i>keine</i> <u>harnsäuresenkende Therapie:</u> ▪ <i>keine</i>
IL-1-Inhibitoren vs. Glukokortikoide			
<i>So 2010</i>	Wirkstoff: <i>Canakinumab</i> <u>Anwendungsart:</u> ▪ <i>subkutan</i> <u>Dosierung:</u> ▪ <i>150 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, intramuskulär</i>	Wirkstoff: <i>Triamcinolon</i> <u>Anwendungsart:</u> ▪ <i>intramuskulär</i> <u>Dosierung:</u> ▪ <i>40 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, subkutan</i>	<u>Schmerzmittel / Glukokortikoide:</u> ▪ <i>Paracetamol, Codein oder Prednisolon</i> <u>harnsäuresenkende Therapie:</u> ▪ <i>Allopurinol</i>
<i>Schlesinger 2012a</i>	Wirkstoff: <i>Canakinumab</i> <u>Anwendungsart:</u> ▪ <i>subkutan</i> <u>Dosierung:</u> ▪ <i>150 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, intramuskulär</i>	Wirkstoff: <i>Triamcinolon</i> <u>Anwendungsart:</u> ▪ <i>intramuskulär</i> <u>Dosierung:</u> ▪ <i>40 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, intramuskulär</i>	<u>Schmerzmittel:</u> ▪ <i>ja, keine näheren Angaben</i> <u>harnsäuresenkende Therapie:</u> ▪ <i>kontinuierliche Einnahme bisheriger Medikation</i>
<i>Schlesinger 2012b</i>	Wirkstoff: <i>Canakinumab</i> <u>Anwendungsart:</u> ▪ <i>subkutan</i> <u>Dosierung:</u> ▪ <i>150 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, intramuskulär</i>	Wirkstoff: <i>Triamcinolon</i> <u>Anwendungsart:</u> ▪ <i>intramuskulär</i> <u>Dosierung:</u> ▪ <i>40 mg als Einmaldosis</i> <u>Placebo:</u> ▪ <i>1 Injektion, intramuskulär</i>	<u>Schmerzmittel:</u> ▪ <i>ja, keine näheren Angaben</i> <u>harnsäuresenkende Therapie:</u> ▪ <i>kontinuierliche Einnahme bisheriger Medikation</i>

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien
(mehreseitige Tabelle)

Studie	Intervention	Vergleich	Begleittherapie
NCT01356602	Wirkstoff: Canakinumab <u>Anwendungsart:</u> ▪ subkutan <u>Dosierung:</u> ▪ 150 mg als Einmaldosis bei Gichtanfall <u>Placebo:</u> ▪ 1 Injektion, intramuskulär ▪ 1 Injektion, subkutan	Wirkstoff: Triamcinolon <u>Anwendungsart:</u> ▪ intramuskulär <u>Dosierung:</u> ▪ 40 mg als Einmaldosis bei Gichtanfall <u>Placebo:</u> ▪ 2 Injektionen, subkutan	<u>Schmerzmittel:</u> ▪ ja, keine näheren Angaben <u>harnsäuresenkende Therapie:</u> ▪ k. A.
NCT01362608	Wirkstoff: Canakinumab <u>Anwendungsart:</u> ▪ subkutan <u>Dosierung:</u> ▪ 150 mg als Einmaldosis <u>Placebo:</u> ▪ 1 Injektion, intramuskulär	Wirkstoff: Triamcinolon <u>Anwendungsart:</u> ▪ intramuskulär <u>Dosierung:</u> ▪ 40 mg als Einmaldosis <u>Placebo:</u> ▪ 1 Injektion, subkutan	<u>Schmerzmittel:</u> ▪ ja, keine näheren Angaben <u>harnsäuresenkende Therapie:</u> ▪ k. A.
IL-1-Inhibitoren: Interleukin-1-Inhibitoren; k. A.: keine Angabe; NSAR: nicht steroidale Antirheumatika			

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Colchicin vs. Placebo		
<i>Terkeltaub 2010</i>	<ul style="list-style-type: none"> ▪ Alter: ≥ 18 Jahre ▪ Patienten und postmenopausale Patientinnen ▪ Gicht diagnostiziert nach ACR-Klassifikationskriterien ▪ ≥ 2 Gichtanfälle innerhalb der letzten 12 Monate vor Einschluss in die Studien^a 	<ul style="list-style-type: none"> ▪ k. A.
NSAR vs. Colchicin		
<i>Liu 2019</i>	<ul style="list-style-type: none"> ▪ Alter: 18 bis 70 Jahre ▪ akuter Gichtanfall (< 48 Stunden) ▪ neue Diagnose einer akuten Gichtarthritis ▪ keine Einnahme von Colchicin, NSAR oder Glukokortikoiden ▪ keine Einnahme von Medikamenten, die den Harnsäurestoffwechsel beeinflussen, in den letzten 3 Monaten ▪ keine weiteren Krankheiten, die den Harnsäurestoffwechsel beeinflussen 	<ul style="list-style-type: none"> ▪ wiederholtes oder intermittierendes Auftreten von akuter Gichtarthritis oder akutes Auftreten Symptomen einer chronischen Gicht ▪ chronische Gicht mit Tophi, rheumatoide Arthritis, traumatische Arthritis oder andere Arten von Arthritis ▪ gichtartige Nephropathie ▪ Diabetes mellitus ▪ Bluthochdruck ▪ schwere Dyslipidämie ▪ schwere Leberschäden ▪ Alanin-Aminotransferase oder Aspartat-Aminotransferase > 2-fache Obergrenze des Normalwerts ▪ schwere kardiovaskuläre und zerebrale Erkrankung ▪ bösartige Erkrankung

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
<i>Roddy 2020</i>	<ul style="list-style-type: none"> ▪ Alter: ≥ 18 Jahre ▪ akuter Gichtanfall ▪ klinische Diagnose der Gicht (ohne Gelenkpunktion, Bluttests, Bildgebung oder anderen diagnostischen Kriterien) 	<ul style="list-style-type: none"> ▪ instabiler medizinischer Zustand (z. B. ischämische Herzerkrankung, eingeschränkte Leberfunktion) ▪ chronische Nierenerkrankung im Stadium 4/5 (eGFR / Kreatinin-Clearance < 30 ml/min) ▪ kürzliche Operation ▪ gastrointestinale Blutung ▪ Anamnese von Magengeschwüren ▪ Einnahme von Antikoagulantien ▪ Allergie gegen Acetylsalicylsäure oder NSAR ▪ frühere Unverträglichkeit von Naproxen oder niedrig dosiertem Colchicin ▪ Kontraindikationen für Studienmedikamente, die in der Zusammenfassung der Produktmerkmale beschrieben sind ▪ Einnahme von Naproxen oder Colchicin in den letzten 24 Stunden ▪ potentiell vulnerable Patientinnen und Patienten ▪ Teilnahme an der CONTACT-Studie während eines früheren Gichtanfalls oder Teilnahme an einer anderen klinischen Studie in den letzten 90 Tagen oder an einer anderen Forschungsarbeit in den letzten 30 Tagen
NSAR vs. Glukokortikoide		
<i>Janssens 2008</i>	<ul style="list-style-type: none"> ▪ Diagnose der monoartikulären Gichtarthritis mit nachgewiesenen Natriumuratkristallen in der Synovialflüssigkeit des betroffenen Gelenks 	<ul style="list-style-type: none"> ▪ instabiler Zustand (Angina pectoris, Myokardinfarkt, manifeste Herzinsuffizienz, schwere Niereninsuffizienz, Nierentransplantation oder Krebs) ▪ chronische rheumatische Erkrankung ▪ derzeitige Einnahme von Antikoagulantien ▪ vorangegangene Erkrankung des oberen Gastrointestinaltraktes

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Man 2007	<ul style="list-style-type: none"> ▪ Alter: > 17 Jahre ▪ Diagnose einer akuten Arthritis, die auf Gicht hindeutet. Definiert als Schmerzen und Wärme in einem Gelenk und dem vorstellig werden innerhalb von 3 Tagen nach dem Auftreten der Schmerzen sowie Aufweisen von mindestens 1 der folgenden Eigenschaften: <ul style="list-style-type: none"> ▫ Beteiligung des Großzehengrundgelenks ▫ Beteiligung des Knie- oder Sprunggelenkes und kristallhaltiges Aspirat ▫ Gichtarthritis mit Tophi oder Gicht diagnostiziert nach Gelenkpunktion 	<ul style="list-style-type: none"> ▪ klinischer Verdacht auf Sepsis oder eine andere Gelenkerkrankung ▪ Komorbidität, die die Studiendurchführung beeinträchtigen könnte ▪ Demenz, Verwirrung ▪ akute gastrointestinale Symptome ▪ Niereninsuffizienz (Serumkreatinin $\geq 200 \mu\text{mol/l}$) ▪ Blutungsstörung ▪ Allergie gegen eines der Studienmedikamente ▪ keine Diagnose der Gicht nach Gelenkpunktion ▪ Einnahme von Warfarin
Rainer 2016	<ul style="list-style-type: none"> ▪ Alter: ≥ 18 Jahre ▪ Gichtsymptome innerhalb der letzten 3 Tage ▪ Diagnose eines akuten Gichtanfalls anhand folgender Kriterien: <ul style="list-style-type: none"> ▫ plötzliches Auftreten (< 6-12 Stunden) von starken Schmerzen, Schwellungen, Empfindlichkeit und Rötung eines betroffenen Gelenks ▫ Mindestens 1 der folgenden klinischen Befunde trifft zu: <ul style="list-style-type: none"> - Beteiligung des Großzehengrundgelenks (Podagra) - Beteiligung des Knie-, Fuß-, Hand- oder Ellbogengelenks mit Tophi, Gicht diagnostiziert nach Gelenkpunktion, Hyperurikämie oder ≥ 1 bisheriger Gichtanfälle ▪ Diagnose der Gicht nach Gelenkpunktion, sofern keines der Kriterien zur Bestimmung der Gicht in Knie-, Fuß-, Hand- oder Ellenbogengelenken zutrifft 	<ul style="list-style-type: none"> ▪ Einnahme von Kortikosteroiden oder Indometacin innerhalb von 24 Stunden vor der Rekrutierung ▪ Anamnese von Blutungsstörungen oder Einnahme von Antikoagulantien ▪ Allergie gegen eines der Studienmedikamente ▪ Verdacht auf septische Arthritis oder eine andere Gelenkerkrankung (z. B. rheumatoide Arthritis) ▪ Fehlen von Natriumuratkristallen in der Gelenkflüssigkeit nach Gelenkpunktion ▪ instabiler kardialer Zustand (Angina pectoris, akuter Myokardinfarkt, Herzinsuffizienz) ▪ signifikante Komorbidität, die die Studiendurchführung beeinträchtigen könnte (Demenz, Verwirrung, aktive gastrointestinale Symptome) ▪ Serumkreatinin $> 200 \mu\text{mol/l}$ ($> 2,26 \text{ mg/dl}$) ▪ eGFR $< 30 \text{ ml/min/1,73 m}^2$

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
<i>Xu 2016</i>	<ul style="list-style-type: none"> ▪ <i>Alter: ≥ 18 Jahre</i> ▪ <i>Gichtanfall (< 72 Stunden)</i> ▪ <i>Gicht diagnostiziert nach ACR-Klassifikationskriterien</i> ▪ <i>moderate Schmerzen (≥ 2 auf einer 5-Punkte-Likert-Skala) im Indexgelenk zu Studienbeginn</i> ▪ <i>Indexgelenk wurde als das Gelenk definiert, das zum Zeitpunkt der Randomisierung am stärksten schmerzte</i> 	<ul style="list-style-type: none"> ▪ <i>chronische Gichtarthritis</i> ▪ <i>polyartikuläre Gicht mit Beteiligung von mehr als vier Gelenken</i> ▪ <i>Verdachtsdiagnose einer Gelenkinfektion oder einer anderen Gelenkerkrankung</i> ▪ <i>koronare Herzkrankheit</i> ▪ <i>Herzinsuffizienz</i> ▪ <i>gastrointestinale Blutung</i> ▪ <i>Magengeschwür</i> ▪ <i>entzündliche Darmerkrankung</i> ▪ <i>bösartiger Tumor</i> ▪ <i>Anamnese von Operationen am Verdauungstrakt</i> ▪ <i>Einnahme von NSAR oder Kortikosteroiden in den 72 Stunden zu Studienbeginn</i> ▪ <i>abnormale Leberfunktion mit Transaminasewerten > 2-fache Obergrenze des Normalwerts</i> ▪ <i>Niereninsuffizienz mit Serumkreatinin > 200 $\mu\text{mol/l}$</i>

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
IL-1-Inhibitoren vs. Glukokortikoide		
So 2010	<ul style="list-style-type: none"> ▪ Alter: 18 bis 80 Jahre ▪ akuter Gichtanfall (≤ 5 Tage) ▪ ≥ 1 vorangegangener Gichtanfall nach ACR-Klassifikationskriterien ▪ Schmerzen (≥ 50 mm auf einer 0-100 mm VAS) zu Studienbeginn ▪ NSAR und / oder Colchicin sind wirkungslos oder kontraindiziert für die Erkrankung der Patientin oder des Patienten ▪ BMI ≤ 40 kg/m² ▪ Patientinnen und Patienten, die eine uratsenkende Therapie erhalten, müssen ein stabiles Dosierungsschema einhalten und sollen dieses Schema während der Studiendauer beibehalten 	<ul style="list-style-type: none"> ▪ Anwendung einer der folgenden Wirkstoffe bzw. Wirkstoffe aus einer der folgenden Wirkstoffklassen vor dem Screening: Ibuprofen, Paracetamol, Acetylsalicylsäure, Diclofenac, Naproxen, Cyclooxygenase-2-Hemmer, andere NSAR, systemische oder intraartikuläre Kortikosteroide, Anakinra, Rilonacept, Tumornekrosefaktor-Hemmer ▪ nicht kontinuierliche Einnahme von > 1 Dosis (0,6 mg) Colchicin in den 24 Stunden vor dem Screening ▪ Arzneimittelallergien ▪ rheumatoide, infektiöse/septische oder andere entzündliche Arthritis ▪ schwere Nierenfunktionseinschränkung ▪ idiopathische thrombozytopenische Purpura ▪ akute oder wiederkehrende Infektion bei Studienbeginn ▪ aktive Lungenerkrankung ▪ latente Tuberkulose mit erforderlicher Antibiotikaeinnahme ▪ Risikofaktoren für Tuberkulose ▪ Kontraindikation für eine intramuskuläre Injektion ▪ Spende oder Verlust von ≥ 400 ml Blut in den 8 Wochen vor der Behandlung ▪ Lebendimpfstoff in den 3 Monaten vor Studienbeginn ▪ jede chirurgische oder zugrunde liegende hepatische, hämatologische, pulmonale, infektiöse oder gastrointestinale Erkrankung, welche das Immunsystem beeinträchtigt und / oder die Patientin oder den Patienten einem inakzeptablen Risiko aussetzt, wenn er eine immunmodulatorische Therapie erhält ▪ langes-QT-Syndrom oder QTc > 450 msec für Männer und > 470 msec für Frauen

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
<i>Schlesinger 2012a</i>	<ul style="list-style-type: none"> ▪ Alter: 18 bis 85 Jahre ▪ akuter Gichtanfall (≤ 5 Tage) ▪ Gicht diagnostiziert nach ACR-Klassifikationskriterien ▪ ≥ 3 Gichtanfälle innerhalb der letzten 12 Monate ▪ Schmerzen (≥ 50 mm auf einer 0-100 mm VAS) zu Studienbeginn ▪ Kontraindikation, Unverträglichkeit oder mangelnde Wirksamkeit von NSAR und / oder Colchicin ▪ $BMI \leq 45 \text{ kg/m}^2$ ▪ Patientinnen und Patienten mit uratsenkender Therapie mussten vor der Randomisierung mindestens 2 Wochen lang ein stabiles Dosierungsschema einhalten und sollten dieses während der Studie beibehalten 	<ul style="list-style-type: none"> ▪ Anwendung bestimmter Schmerzmittel oder Biologika (einschließlich Kortikosteroiden, Narkotika, Paracetamol/Acetaminophen, Ibuprofen, Colchicin, IL-Blocker, Tumornekrosefaktor-Hemmer) vor Studienbeginn ▪ rheumatoide Arthritis ▪ infektiöse / septische Arthritis ▪ andere akute entzündliche Arthritis ▪ vorangegangene bösartige Erkrankung ▪ aktive, chronische oder wiederkehrende Infektion, einschließlich Tuberkulose oder HIV-Infektion oder Hepatitis-B- oder C-Infektion
<i>Schlesinger 2012b</i>	<ul style="list-style-type: none"> ▪ siehe Schlesinger 2012a 	<ul style="list-style-type: none"> ▪ siehe Schlesinger 2012a

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
NCT01356602	<ul style="list-style-type: none"> ▪ Alter: 18 bis 85 Jahre ▪ ≥ 3 Gichtanfälle innerhalb der letzten 12 Monate ▪ Kontraindikation, Unverträglichkeit oder mangelnde Wirksamkeit von NSAR und / oder Colchicin ▪ BMI ≤ 45 kg/m² 	<ul style="list-style-type: none"> ▪ Verwendung folgender Therapien: Kortikosteroiden, Narkotika, topische Eis- / Kältepackungen, dauerhafte Opiatbehandlung, NSAR, Colchicin) ▪ Hämodialyse ▪ Lebendimpfstoff in den 3 Monaten vor der ersten Dosis ▪ Spende oder Verlust von ≥ 400 ml Blut in den 3 Monaten vor der ersten Dosis ▪ Chemotherapie-, Blei- oder Transplantations-induzierte Gicht ▪ Vorliegen einer anderen entzündlichen Arthritis wie z. B. rheumatoide Arthritis ▪ Erkrankung oder schwerwiegende medizinische Probleme (z. B. HIV, Hepatitis, Tuberkulose und andere Infektionen/ Krankheiten), die zu einem inakzeptablen immunologischen Risiko bei Erhalt der Studientherapie führen ▪ schwerwiegende kardiovaskuläre Erkrankung (z. B. unkontrollierter Bluthochdruck) ▪ schwerwiegende medizinische Erkrankungen (z. B. unkontrollierter Diabetes mellitus, Schilddrüsenerkrankung) ▪ bösartige Erkrankung in den letzten 5 Jahren
NCT01362608	<ul style="list-style-type: none"> ▪ Alter: 18 bis 85 Jahre ▪ akuter Gichtanfall (≤ 5 Tage) ▪ Erfüllung der Diagnosekriterien für akute Arthritis bei primärer Gicht ▪ ≥ 3 Gichtanfälle innerhalb der letzten 12 Monate vor Studienstart ▪ Kontraindikation (absolut oder relativ), Unverträglichkeit, mangelnde Wirksamkeit von NSAR und / oder Colchicin 	<ul style="list-style-type: none"> ▪ Einnahme von Antibiotika gegen latente Tuberkulose z. B. Isoniazid ▪ refraktäre Herzinsuffizienz ▪ instabile Herzrhythmusstörungen oder symptomatische koronare Ischämie ▪ sekundäre Gicht, Chemotherapie-, Blei- oder Transplantations-induzierte Gicht ▪ rheumatoide Arthritis ▪ Nachweis / Verdacht einer infektiösen / septischen Arthritis oder einer anderen akuten entzündlichen Arthritis
<p>a. Untersucht wurde in der Studie der Effekt der Studienmedikation bei Auftreten eines akuten Gichtanfalls. Die eingeschlossenen verblindeten Patientinnen und Patienten erhielten einen Blister der Studienmedikation, welche bei Auftreten eines Gichtanfalls eingenommen werden sollte.</p> <p>ACR: American College of Rheumatology; BMI: Body-Mass-Index; eGFR: geschätzte glomeruläre Filtrationsrate (estimated GFR); HIV: Humanes Immundefizienz-Virus; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; k. A.: keine Angabe; NSAR: nicht steroidale Antirheumatika; QTc: korrigierte QT-Zeiten; VAS: visuelle Analogskala mit keinen Schmerzen bei 0mm und unerträglichen Schmerzen bei 100mm</p>		

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
------------------	---------------------------------	---------------------------------

Tabelle 7: Charakterisierung der Studienpopulationen (mehreseitige Tabelle)

Studie Gruppe	Intervention		Vergleich
Colchicin vs. Placebo			
Terkeltaub 2010			
N	74		59
Alter [Jahre] MW (SD)	51,4 (11,79)		51,2 (11,36)
Geschlecht [w / m] %	3 / 97		7 / 93
NSAR vs. Colchicin			
Liu 2019			
N	53		52
Alter [Jahre] MW (SD)	44 (9)		43 (9)
Geschlecht [w / m] n	4 / 49		5 / 47
Roddy 2020			
N	200		199
Alter [Jahre] MW (SD)	58,7 (14,4)		60 (13,4)
Geschlecht [w / m] n	27 / 173		25 / 174
NSAR vs. Glukokortikoide			
Janssens 2008			
N	60		60
Alter [Jahre] MW (SD)	57,7 (13,4)		57,3 (12,2)
Geschlecht [w / m] n	7 / 53		6 / 54
Man 2007			
N	46		44
Alter [Jahre] MW (SD)	66 (16)		64 (15)
Geschlecht [w / m] n	7 / 39		9 / 35
Rainer 2016			
N	208		208
Alter [Jahre] MW (SD)	64,37 (16,01)		65,91 (24,95)
Geschlecht [w / m] n	44 / 164		63 / 145
Xu 2016^a			
N	46	45	41
Alter [Jahre] MW (SD)	44,43 (15,08)	43,81 (12,29)	44,03 (15,37)
Geschlecht [w / m] %	0 / 100	3 / 97	0 / 100
IL-1-Inhibitoren vs. Glukokortikoide			
So 2010			
N	28		57

Tabelle 6: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien (mehreseitige Tabelle)

Vergleich Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
	Alter [Jahre] MW (SD) Geschlecht [w / m] n	50,6 (15,38) 0 / 28
		52,4 (11,55) 2 / 55
Schlesinger 2012a		
N	113	115
Alter [Jahre] MW (SD)	54,0 (11,2)	54,6 (10,7)
Geschlecht [w / m] %	11 / 89	6 / 94
Schlesinger 2012b		
N	112	114
Alter [Jahre] MW (SD)	50,6 (12,1)	52,6 (12,3)
Geschlecht [w / m] %	11 / 89	8 / 92
NCT01356602		
N	133	133
Alter [Jahre] MW (SD)	53,4 (11,21)	53,7 (11,33)
Geschlecht [w / m] %	11 / 89	8 / 92
NCT01362608		
N	67	69
Alter [Jahre] MW (SD)	50,2 (11,98)	49,7 (11,64)
Geschlecht [w / m] %	1 / 99	3 / 97
<p>a. Die Studie untersucht 2 Prüfinterventionen. Dargestellt werden der Reihenfolge nach die Prüfinterventionsgruppe Etoricoxib und die Prüfinterventionsgruppe Indometacin.</p> <p>IL-1-Inhibitoren: Interleukin-1-Inhibitoren; k. A.: keine Angabe; m: männlich; MW: Mittelwert; N: Anzahl randomisierter (bzw. eingeschlossener) Patientinnen und Patienten; n: Anzahl der Patientinnen und Patienten mit Merkmal; NSAR: nicht steroidale Antirheumatika; SD: Standardabweichung; w: weiblich</p>		

5.3 Übersicht der Endpunkte

Aus 12 Studien wurden Daten zu kritischen und wichtigen Endpunkten extrahiert. Die Tabelle 8 zeigt die Übersicht der dargestellten Daten zu den von der Leitlinienkoordination definierten Endpunkten aus den berücksichtigten Studien.

Tabelle 8: Matrix der einbezogenen Endpunkte

Studie	Kritische Endpunkte	Wichtige Endpunkte		
	Schmerzen	Funktionsfähigkeit	unerwünschte Ereignisse	Therapieabbruch aufgrund von unerwünschten Ereignissen
Colchicin vs. Placebo				
<i>Terkeltaub 2010</i>	● ^a	–	● ^b	–
NSAR vs. Colchicin				
<i>Liu 2019</i>	● ^a	–	● ^b	●
<i>Roddy 2020</i>	● ^a	–	● ^b	–
NSAR vs. Glukokortikoide				
<i>Janssens 2008</i>	● ^a	●	– ^c	–
<i>Man 2007</i>	–	–	● ^{c, d}	●
<i>Rainer 2016</i>	● ^a	–	● ^{c, d}	●
<i>Xu 2016</i>	–	●	– ^c	–
IL-1-Inhibitoren vs. Glukokortikoide				
<i>So 2010</i>	●	●	● ^b	–
<i>Schlesinger 2012a</i>	●	●	● ^b	–
<i>Schlesinger 2012b</i>	●	●	● ^b	–
NCT01356602	●	–	● ^d	–
NCT01362608	●	–	● ^d	–
<p>●: Daten werden im Evidenzbericht dargestellt. –: Daten wurden für den Evidenzbericht nicht herangezogen oder in der Studie nicht berichtet. a. Ergebnisse zu diesem Endpunkt wurden aus der NICE 2022 entnommen. b. Es werden SUEs aus den Studien sowie unerwünschte Ereignisse aus der NICE 2022 dargestellt. c. Ergebnisse zu einzelnen gastrointestinalen unerwünschten Ereignissen werden in der NICE 2022 berichtet. d. Es werden SUEs aus den Studien dargestellt.</p> <p>IL-1-Inhibitoren: Interleukin-1-Inhibitoren; NSAR: nicht steroidale Antirheumatika; SUE: schwerwiegendes unerwünschtes Ereignis</p>				

5.4 Ergebnisse zu Endpunkten

5.4.1 Ergebnisse zu dem Vergleich Colchicin vs. Placebo

5.4.1.1 Ergebnisse zum kritischen Endpunkt Schmerzen

Tabelle 9: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
Anteil an Patientinnen und Patienten (≥ 50 % Reduktion der Schmerzintensität auf einer VAS seit Studienbeginn) – Kurzeiteffekt (< 2 Wochen)^d												
Interpretation der Effektschätzung: Ein RR größer als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^e	<i>schwer-wiegend^f</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^g	<i>nicht schwer-wiegend</i>	<i>keine</i>	31 / 74	10 / 58	172 pro 1000 ^h	2,43 [1,30; 4,54]	247 mehr pro 1000 [52 mehr bis 610 mehr]	<i>moderat</i>

Tabelle 9: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. 32 Stunden nach Gabe der 1. Dosis</p> <p>e. in NICE 2022 berücksichtigte Studie: <i>Terkeltaub 2010</i></p> <p>f. Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte (Seite 126, Tabelle 18 Fußnote a [21]).</p> <p>g. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkt > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>h. eigene Berechnung</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko; VAS: visuelle Analogskala</p>												

5.4.1.2 Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse

Tabelle 10: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten)

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	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
SUE – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [22]	– ^c	– ^c	– ^c	– ^c	– ^c	– ^c	0 / 74	0 / 59	– ^c	– ^c	– ^c	– ^c
<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. nicht dargestellt, da keine Ereignisse in der Prüfinterventions- und Vergleichsgruppe auftraten</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; SUE: schwerwiegendes unerwünschtes Ereignis</p>												

Tabelle 11: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
unerwünschte Ereignisse – gastrointestinal (Diarrhö und Erbrechen) – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüflnterventionsgruppe.												
RCT, 1 [21] ^d	<i>schwer-wiegend^e</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>sehr schwer-wiegend^g</i>	<i>keine</i>	19 / 74	12 / 59	203 pro 1000 ^h	1,26 [0,67; 2,39]	53 mehr pro 1000 [67 weniger bis 283 mehr]	<i>sehr niedrig</i>
<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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. in NICE 2022 berücksichtigte Studie: <i>Terkeltaub 2010</i></p> <p>e. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte</i> (Seite 126, Tabelle 18 Fußnote a [21]).</p> <p>f. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>g. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn das 95 %-KI eine MID-Schwelle überschreitet. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn das 95 %-KI beide MID-Schwellen überschreitet. Etablierte MID-Schwellen für die kontinuierliche VAS: Verbesserungen ≥ 10 Punkte auf einer 1-100 Skala; Standard MID-Schwellen nach GRADE wurden für alle anderen Endpunkte verwendet. Für binäre Endpunkte betragen die MID-Schwellen RR von 0,8 und 1,25. Für kontinuierliche Endpunkte wurde die MID-Schwelle mit 0,5 x SD zu Studienbeginn berechnet</i> (Seite 132, Tabelle 19 Fußnote b [21]).</p> <p>h. eigene Berechnung</p>												

Tabelle 11: Evidenzprofil für den Vergleich Colchicin vs. Placebo für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
C: Vergleichsgruppe; GRADE: Grading of Recommendations Assessment, Development and Evaluation; I: Prüflinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; MID: Relevanzschwelle für 1 kleinen Effekt (minimal important difference); N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko; SD: Standardabweichung; VAS: visuelle Analogskala												

5.4.2 Ergebnisse zu dem Vergleich NSAR vs. Colchicin

5.4.2.1 Ergebnisse zum kritischen Endpunkt Schmerzen

Tabelle 12: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
vollständige Schmerzauflösung – Kurzeiteffekt (< 2 Wochen)^d												
Interpretation der Effektschätzung: Ein RR größer als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^e	<i>sehr schwer-wiegend^f</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^g	<i>nicht schwer-wiegend</i>	<i>keine</i>	115 / 170	116 / 174	667 pro 1000 ^h	1,01 [0,88; 1,18]	7 mehr pro 1000 [80 weniger bis 120 mehr]	niedrig
vollständige Schmerzauflösung – mittelfristiger Effekt (2 bis 6 Wochen)ⁱ												
Interpretation der Effektschätzung: Ein RR größer als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^e	<i>sehr schwer-wiegend^f</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^g	<i>nicht schwer-wiegend</i>	<i>keine</i>	130 / 170	130 / 174	747 pro 1000 ^h	1,02 [0,91; 1,15]	15 mehr pro 1000 [67 weniger bis 112 mehr]	niedrig

Tabelle 12: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. <i>nach 7 Tagen</i></p> <p>e. in NICE 2022 berücksichtigte Studie: <i>Roddy 2020</i></p> <p>f. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte</i> (Seite 139, Tabelle 20 Fußnote a [21]).</p> <p>g. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>h. eigene Berechnung</p> <p>i. <i>nach 4 Wochen</i></p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko</p>												

Tabelle 13: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den kritischen Endpunkt Schmerzen (stetige Daten) (aus NICE 2022 übernommen)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Gliederschmerz (Veränderung des Punktwertes^c) – Kurzeiteffekt (< 2 Wochen)^d										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.										
RCT, 1 [21] ^e	nicht schwer-wiegend	nicht schwer-wiegend	nicht schwer-wiegend	k. A. ^f	nicht schwer-wiegend	keine	53; 52	0,06 [-0,28; 0,40]	k. A. ^g	hoch
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>c. Die verwendete Skala wurde in der NICE 2022 nicht angegeben.</p> <p>d. nach 10 Tagen 10</p> <p>e. in NICE 2022 berücksichtigte Studie: Liu 2019</p> <p>f. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>g. nicht in NICE 2022 angegeben</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie</p>										

5.4.2.2 Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse

Tabelle 14: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten)

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]	
SUE – mittelfristiger Effekt (2 bis 6 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [25]	nicht schwerwiegend	nicht zutreffend	nicht schwerwiegend	unentdeckt	sehr schwerwiegend ^d	keine	2 / 200	1 / 199	1	2,00 [0,18; 22,24]	1 [-2; 3]	niedrig
<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 C zu entnehmen.</p> <p>d. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfinterventionsgruppe ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; SUE: schwerwiegendes unerwünschtes Ereignis</p>												

Tabelle 15: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
unerwünschte Ereignisse – gastrointestinal (Erbrechen) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^d	<i>schwer-wiegend^e</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>sehr schwer-wiegend^g</i>	<i>keine</i>	2 / 53	1 / 52	19 pro 1000 ^h	1,96 [0,18; 20,99]	18 mehr pro 1000 [16 weniger bis 384 mehr]	<i>sehr niedrig</i>
unerwünschte Ereignisse – gastrointestinal (Diarrhö) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 2 [21] ⁱ	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>nicht schwer-wiegend</i>	<i>keine</i>	32 / 223	69 / 226	305 pro 1000 ^h	0,47 [0,33; 0,68] ^j	162 weniger pro 1000 [205 weniger bis 98 weniger]	<i>hoch</i>

Tabelle 15: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
unerwünschte Ereignisse – gastrointestinal (Übelkeit und / oder Erbrechen) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^k	nicht schwer-wiegend	nicht schwer-wiegend	nicht schwer-wiegend	k. A. ^f	schwer-wiegend ^g	keine	21 / 170	30 / 174	172 pro 1000 ^h	0,72 [0,43; 1,20] ^j	48 weniger pro 1000 [98 weniger bis 34 mehr]	moderat
unerwünschte Ereignisse – gastrointestinal (Dyspepsie) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^k	nicht schwer-wiegend	nicht schwer-wiegend	nicht schwer-wiegend	k. A. ^f	sehr schwer-wiegend ^g	keine	20 / 170	20 / 174	115 pro 1000 ^h	1,02 [0,57; 1,83] ^j	2 mehr pro 1000 [49 weniger bis 95 mehr]	niedrig

Tabelle 15: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
unerwünschte Ereignisse – gastrointestinal (Abdominalschmerzen) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^k	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>sehr schwer-wiegend^g</i>	<i>keine</i>	16 / 170	16 / 174	92 pro 1000 ^h	1,02 [0,53; 1,98] ^j	2 mehr pro 1000 [43 weniger bis 90 mehr]	<i>niedrig</i>
unerwünschte Ereignisse – gastrointestinal (Obstipation) – Kurzeffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^k	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>nicht schwer-wiegend</i>	<i>keine</i>	30 / 170	67 / 174	385 pro 1000 ^h	0,46 [0,31; 0,67] ^j	208 weniger pro 1000 [266 weniger bis 127 weniger]	<i>hoch</i>

Tabelle 15: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. in NICE 2022 berücksichtigte Studie: <i>Liu 2019</i></p> <p>e. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte</i> (Seite 139, Tabelle 20 Fußnote a [21]).</p> <p>f. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>g. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn das 95 %-KI eine MID-Schwelle überschreitet. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn das 95 %-KI beide MID-Schwellen überschreitet. Etablierte MID-Schwellen für die kontinuierliche VAS: Verbesserungen ≥ 10 Punkte auf einer 1-100 Skala; Standard MID-Schwellen nach GRADE wurden für alle anderen Endpunkte verwendet. Für binäre Endpunkte betragen die MID-Schwellen RR von 0,8 und 1,25. Für kontinuierliche Endpunkte wurde die MID-Schwelle mit $0,5 \times SD$ zu Studienbeginn berechnet</i> (Seite 132, Tabelle 19 Fußnote b [21]).</p> <p>h. eigene Berechnung</p> <p>i. in NICE 2022 berücksichtigte Studien: <i>Liu 2019 und Roddy 2020</i></p> <p>j. Beim mittelfristigen Effekt (2 bis 6 Wochen) berichtete die NICE 2022 für Roddy 2020 keinen statistisch signifikanten Unterschied.</p> <p>k. in NICE 2022 berücksichtigte Studie: <i>Roddy 2020</i></p> <p>C: Vergleichsgruppe; GRADE: Grading of Recommendations Assessment, Development and Evaluation; I: Prüflinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; MID: Relevanzschwelle für 1 kleinen Effekt (minimal important difference); N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko; SD: Standardabweichung;</p>												

5.4.2.3 Ergebnisse zum wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse

Tabelle 16: Evidenzprofil für den Vergleich NSAR vs. Colchicin für den wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse (binäre Daten)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Therapieabbruch aufgrund unerwünschter Ereignisse – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüflnterventionsgruppe.												
RCT, 1 [24]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	sehr schwer-wiegend ^e	keine	1 / 61	4 / 61	7	0,24 [0,03; 2,19]	-5 [-14; 3]	sehr niedrig
<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 C zu entnehmen.</p> <p>d. Die Verdeckung der Gruppenzuteilung war unklar.</p> <p>e. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüflnterventionsgruppe ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüflnterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

5.4.3 Ergebnisse zu dem Vergleich NSAR vs. Glukokortikoide

5.4.3.1 Ergebnisse zum kritischen Endpunkt Schmerzen

Tabelle 17: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
klinisch signifikante Änderung des Schmerzpunktwertes (13mm auf einer 100mm VAS) im Ruhezustand – Kurzeiteffekt (< 2 Wochen)^d												
Interpretation der Effektschätzung: Ein RR größer als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^e	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>schwer-wiegend^g</i>	<i>keine</i>	111 / 208 ^h	101 / 208 ^h	486 pro 1000 ⁱ	1,10 [0,91; 1,33] ^h	48 mehr pro 1000 [53 weniger bis 133 mehr] ^h	<i>moderat</i>
klinisch signifikante Änderung des Schmerzpunktwertes (13mm auf einer 100mm VAS) während Aktivität – Kurzeiteffekt (< 2 Wochen)^d												
Interpretation der Effektschätzung: Ein RR größer als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [21] ^e	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^f	<i>nicht schwer-wiegend</i>	<i>keine</i>	151 / 208 ^h	160 / 208 ^h	769 pro 1000 ⁱ	0,94 [0,84; 1,05] ^h	44 weniger pro 1000 [138 weniger bis 36 mehr] ^h	<i>hoch</i>

Tabelle 17: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (binäre Daten) (aus NICE 2022 übernommen) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publi-kations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
<p>a. Basisrisiko geschätzt durch (medianes) Risiko der Vergleichsgruppe 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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. <i>nach 14 Tagen</i></p> <p>e. in NICE 2022 berücksichtigte Studie: <i>Rainer 2016</i></p> <p>f. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>g. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn das 95 %-KI eine MID-Schwelle überschreitet. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn das 95 %-KI beide MID-Schwellen überschreitet. Etablierte MID-Schwellen für die kontinuierliche VAS: Verbesserungen ≥ 10 Punkte auf einer 1-100 Skala; Standard MID-Schwellen nach GRADE wurden für alle anderen Endpunkte verwendet. Für binäre Endpunkte betragen die MID-Schwellen RR von 0,8 und 1,25. Für kontinuierliche Endpunkte wurde die MID-Schwelle mit $0,5 \times SD$ zu Studienbeginn berechnet (Seite 132, Tabelle 19 Fußnote b [21]).</i></p> <p>h. Die Prüfinterventions- und Vergleichsgruppe wurden in der NICE 2022 der jeweils anderen Gruppe zugeordnet dargestellt (Prüfinterventionsgruppe: Glukokortikoide; Vergleichsgruppe: NSAR).</p> <p>i. eigene Berechnung</p> <p>C: Vergleichsgruppe; GRADE: Grading of Recommendations Assessment, Development and Evaluation; I: Prüfinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; MID: Relevanzschwelle für 1 kleinen Effekt (minimal important difference); N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko; SD: Standardabweichung; VAS: visuelle Analogskala</p>												

Tabelle 18: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (stetige Daten) (aus NICE 2022 übernommen)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Schmerzen (VAS 0-100mm) zu 90h nach Behandlung – Kurzeiteffekt (< 2 Wochen)										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.										
RCT, 1 [21] ^c	<i>schwer-wiegend^d</i>	<i>nicht schwer-wiegend</i>	<i>nicht schwer-wiegend</i>	k. A. ^e	<i>schwer-wiegend^f</i>	<i>keine</i>	59; 59	-3,9 [11,57 weniger bis 3,77 mehr] ^g	k. A. ^h	<i>moderatⁱ</i>
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>c. in NICE 2022 berücksichtigte Studie: <i>Janssens 2008</i></p> <p>d. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte</i> (Seite 132, Tabelle 19 Fußnote a [21]).</p> <p>e. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>f. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn das 95 %-KI eine MID-Schwelle überschreitet. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn das 95 %-KI beide MID-Schwellen überschreitet. Etablierte MID-Schwellen für die kontinuierliche VAS: Verbesserungen ≥ 10 Punkte auf einer 1–100-Skala; Standard MID-Schwellen nach GRADE wurden für alle anderen Endpunkte verwendet. Für binäre Endpunkte betragen die MID-Schwellen RR von 0,8 und 1,25. Für kontinuierliche Endpunkte wurde die MID-Schwelle mit $0,5 \times SD$ zu Studienbeginn berechnet</i> (Seite 132, Tabelle 19 Fußnote b [21]).</p> <p>g. Die Prüfinterventions- und Vergleichsgruppe wurden in der NICE 2022 der jeweils anderen Gruppe zugeordnet dargestellt (Prüfinterventionsgruppe: Glukokortikoide; Vergleichsgruppe: NSAR).</p> <p>h. nicht in NICE 2022 angegeben</p> <p>i. In der NICE 2022 wurden keine Gründe zur einfachen anstatt der zweifachen Abwertung der Qualität der Evidenz angegeben.</p> <p>C: Vergleichsgruppe; GRADE: Grading of Recommendations Assessment, Development and Evaluation; I: Prüfinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; MID: Relevanzschwelle für 1 kleinen Effekt (minimal important difference); N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie; SD: Standardabweichung; VAS: visuelle Analogskala (mit keinen Schmerzen bei 0 mm und unerträglichen Schmerzen bei 100 mm)</p>										

5.4.3.2 Ergebnisse zum wichtigen Endpunkt Funktionsfähigkeit

Tabelle 19: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt Funktionsfähigkeit (stetige Daten)

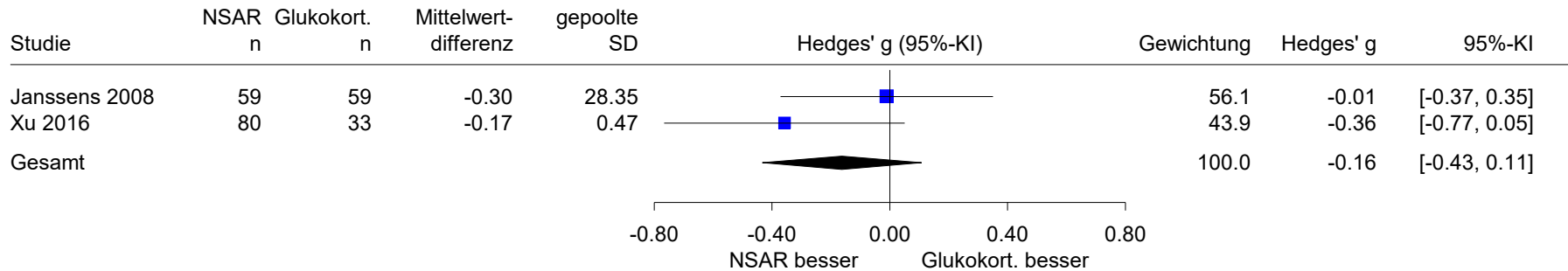
Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert differenz [95 %-KI]	Hedges'g [95 %-KI]	
Veränderung der Beweglichkeit der Gelenke (verschiedene-Skalen) – Kurzezeiteffekt (< 2 Wochen)^c										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.										
RCT, 2 [27,30]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	139; 92	– ^e	–0,16 [–0,43; 0,11]	moderat
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang C zu entnehmen.</p> <p>c. 90 Stunden[27] bzw. 4 Tage nach Studieneinschluss [30]</p> <p>d. Die Verdeckung der Gruppenzuteilung war in 1 Studie unklar, die Studie war nicht verblindet und das ITT-Prinzip war nicht adäquat umgesetzt.</p> <p>e. Berechnung der Mittelwertdifferenz nicht sinnvoll, da verschiedene Skalen verwendet wurden</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; RCT: randomisierte kontrollierte Studie</p>										

5.4.3.2.1 Metaanalysen

NSAR vs. Glukokortikoide

Funktionsfähigkeit

Modell mit festem Effekt - Inverse Varianz



Heterogenität: $Q=1.56$, $df=1$, $p=0.212$, $I^2=35.9\%$

Gesamteffekt: Z-Score=-1.18, $p=0.238$

Abbildung 2: Metaanalyse, NSAR vs. Glukokortikoide, Endpunkt Funktionsfähigkeit – Kurzeiteffekt, Effektmaß: Hedges' g

5.4.3.3 Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse

Tabelle 20: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten)

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]	
SUE (Indometacin + Diclofenac vs. Prednisolon) – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [28]	nicht schwerwiegend	nicht zutreffend	nicht schwerwiegend	unentdeckt	nicht schwerwiegend	keine	7 ^d / 46	0 / 44	0	16,90 [0,94; 305,47]	15 [4; 28]	hoch
SUE (Indometacin vs. Prednisolon bei Patientinnen und Patienten ohne vorangegangene obere gastrointestinale Blutungen^e) – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 1 [29]	– ^f	– ^f	– ^f	– ^f	– ^f	– ^f	0 / 208	0 / 208	– ^f	– ^f	– ^f	– ^f
<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 C zu entnehmen.</p> <p>d. Bei 5 der 7 aufgetretenen SUEs handelte es sich um obere gastrointestinale Blutungen und bei den weiteren 2 um Magen- und Zwölffingerdarmgeschwüre.</p> <p>e. Dieses Ausschlusskriterium wird zusätzlich in der Diskussion der Publikation genannt und führt zu einem Unterschied zwischen den Studienpopulationen von Man 2007 und Rainer 2016. Die Studienautoren dieser Studie waren auch in Man 2007 beteiligt.</p> <p>f. nicht dargestellt, da keine Ereignisse in der Prüfinterventions- und Vergleichsgruppe auftraten</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; SUE: schwerwiegendes unerwünschtes Ereignis</p>												

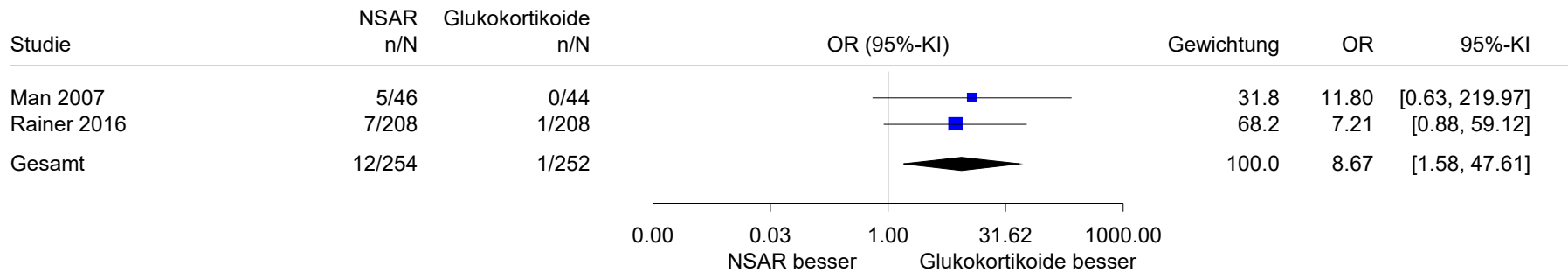
5.4.3.4 Ergebnisse zum wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse

Tabelle 21: Evidenzprofil für den Vergleich NSAR vs. Glukokortikoide für den wichtigen Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse (binäre Daten)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirekt-heit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD in %-Punkten [95 %-KI]	
Therapieabbruch aufgrund unerwünschter Ereignisse – Kurzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüflnterventionsgruppe.												
RCT, 2 [28,29]	schwer-wiegend ^d	nicht schwer-wiegend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	12 / 254	1 / 252	0,2	8,67 [1,58; 47,61]	2 [0,1; 10]	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 C zu entnehmen.</p> <p>d. In 1 Studie war die Verdeckung der Gruppenzuteilung unklar.</p> <p>C: Vergleichsgruppe; I: Prüflnterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; NSAR: nicht steroidale Antirheumatika; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz</p>												

5.4.3.4.1 Metaanalysen

NSAR vs. Glukokortikoide
 Therapieabbruch aufgrund UE
 Modell mit festem Effekt - Mantel-Haenszel



Heterogenität: $Q=0.07$, $df=1$, $p=0.788$, $I^2=0\%$

Gesamteffekt: Z-Score=2.49, $p=0.013$

Abbildung 3: Metaanalyse, NSAR vs. Glukokortikoide, Endpunkt Therapieabbruch aufgrund von unerwünschten Ereignissen – Kurzeiteffekt, Effektmaß: OR

5.4.4 Ergebnisse zu dem Vergleich IL-1-Inhibitoren vs. Glukokortikoide

5.4.4.1 Ergebnisse zum kritischen Endpunkt Schmerzen

Tabelle 22: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den kritischen Endpunkt Schmerzen (stetige Daten)

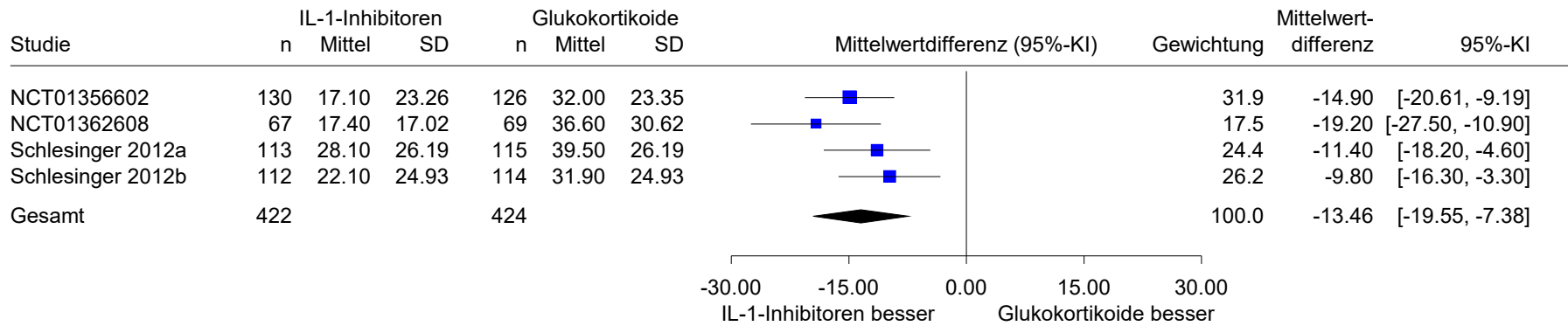
Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
Schmerzen (VAS 100mm) – Kurzeiteffekt (< 2 Wochen)^c										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.										
RCT, 4 [35,42,44]	schwer-wiegend ^d	nicht schwer-wiegend	schwer-wiegend ^e	unentdeckt	nicht schwer-wiegend	keine	422; 424	-13,46 [-19,55; -7,38] ^f	- ^g	niedrig
Schmerzveränderung in % (VAS 100mm) – Kurzeiteffekt (< 2 Wochen)^h										
Interpretation der Effektschätzung: Eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.										
RCT, 1 [32]	schwer-wiegend ⁱ	nicht zutreffend	schwer-wiegend ^e	unentdeckt	nicht schwer-wiegend	keine	27; 56	-17,90 [-30,86; -4,94] ^j	- ^g	niedrig
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang C zu entnehmen.</p> <p>c. 72 Stunden nach Gabe des Medikaments</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung war in allen 4 Studien unklar. 2 Studien waren nicht verblindet.</p> <p>e. Die Anwendungsart in der Vergleichsgruppe weicht von der deutschen Zulassung ab: es erfolgte eine intramuskuläre statt einer intraartikulären Injektion.</p> <p>f. Die Effektschätzung und die 95 %-KI der Studien Schlesinger 2012a und Schlesinger 2012b wurden aus der NICE 2022 entnommen.</p> <p>g. für die Bewertung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant</p> <p>h. 7 Tage nach Gabe des Medikaments</p> <p>i. Die Studie war nicht verblindet.</p> <p>j. Die Punktschätzungen, Standardabweichungen und Anzahl der Patientinnen und Patienten pro Gruppe der Studie So 2011 wurden aus der NICE 2022 entnommen.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; KI: Konfidenzintervall; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; SD: Standardabweichung; VAS: visuelle Analogskala (mit keinen Schmerzen bei 0mm und unerträglichen Schmerzen bei 100mm)</p>										

5.4.4.1.1 Metaanalysen

IL-1-Inhibitoren vs. Glukokortikoide

Schmerzen

Modell mit zufälligen Effekten - Knapp und Hartung



Heterogenität: $Q=3.65$, $df=3$, $p=0.301$, $I^2=17.9\%$

Gesamteffekt: $Z\text{-Score}=-7.04$, $p=0.006$, $\text{Tau(Paule-Mandel)}=1.724$

Abbildung 4: Metaanalyse, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt Schmerzen – Kurzeiteffekt, Effektmaß: OR

5.4.4.2 Ergebnisse zum wichtigen Endpunkt Funktionsfähigkeit

Tabelle 23: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt Funktionsfähigkeit (stetige Daten) (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
SF-36, körperliche Funktionsfähigkeit – Kurzzeiteffekt (< 2 Wochen)^c										
Interpretation der Effektschätzung: Eine positive Effektschätzung bedeutet einen Vorteil für die Prüflnterventionsgruppe.										
RCT, 1 [32]	schwer-wiegend ^d	nicht zutreffend	schwer-wiegend ^e	unentdeckt	nicht schwer-wiegend	keine	27; 56	6,40 [2,10; 10,70]	- ^f	niedrig
SF-36, körperliche Funktionsfähigkeit – Langzeiteffekt (> 6 Wochen)^g										
Interpretation der Effektschätzung: Eine positive Effektschätzung bedeutet einen Vorteil für die Prüflnterventionsgruppe.										
RCT, 3 [35]	schwer-wiegend ^h	sehr schwer-wiegend ⁱ	schwer-wiegend ^e	unentdeckt	schwer-wiegend ^j	keine	96; 92	0,28 [-0,50; 1,06]	0,10 [-0,18; 0,39]	sehr niedrig
[35]							83; 81	2,71 [1,85; 3,57]	0,96 [0,64; 1,29]	
[32]							27; 56	5,70 [1,83; 9,57]	0,57 [0,10; 1,03]	

Tabelle 23: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt Funktionsfähigkeit (stetige Daten) (mehrsseitige Tabelle)

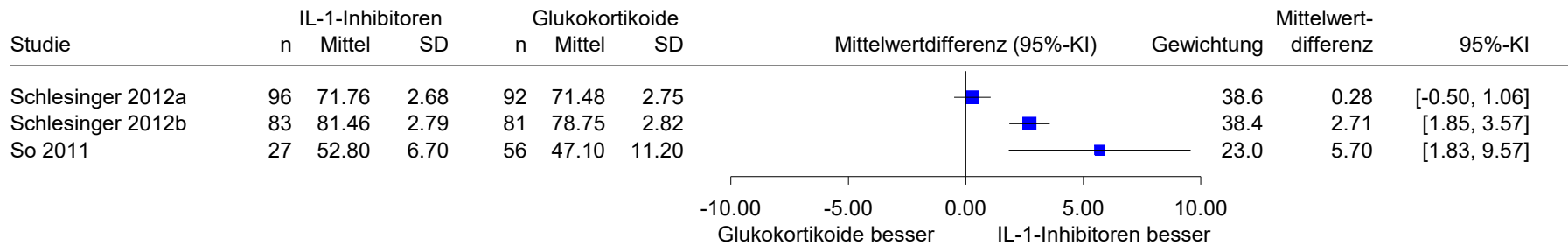
Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt		Qualität der Evidenz ^a
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges'g [95 %-KI]	
<p>a. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.</p> <p>b. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang C zu entnehmen.</p> <p>c. 7 Tage nach Gabe des Medikaments</p> <p>d. Die Studie war nicht verblindet. Die Berichterstattung war nicht ergebnisunabhängig.</p> <p>e. Die Anwendungsart in der Vergleichsgruppe weicht von der deutschen Zulassung ab: es erfolgte eine intramuskuläre statt einer intraartikulären Injektion.</p> <p>f. für die Bewertung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant</p> <p>g. 8 Wochen [32] bzw. 12 Wochen [35] nach Gabe des Medikaments</p> <p>h. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung war in 2 Studien unklar. Die Studien waren nicht verblindet. Das ITT-Prinzip war in 2 Studien nicht adäquat umgesetzt und die Berichterstattung war in 1 Studie nicht ergebnisunabhängig.</p> <p>i. Bedeutsame Heterogenität ($p < 0,001$; $I^2 = 90,8 \%$) und Effektschätzungen überlappen teilweise nicht, daher erfolgte keine Darstellung eines gepoolten Effekts.</p> <p>j. Die Vereinigung der 95 %-KIs des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zuungunsten der Prüfintervention noch mittelgroße Effekte zugunsten der Prüfintervention ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; SF-36: 36-item Short-Form Health Survey</p>										

5.4.4.2.1 Metaanalysen

IL-1-Inhibitoren vs. Glukokortikoide - Langzeiteffekt

Funktionsfähigkeit

Modell mit zufälligen Effekten - Knapp und Hartung (zur Darstellung der Gewichte)



Heterogenität: $Q=21.64$, $df=2$, $p<0.001$, $I^2=90.8\%$

Abbildung 5: Forest Plot, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt Funktionsfähigkeit – Langzeiteffekt, Effektmaß: Mittelwertdifferenz

5.4.4.3 Ergebnisse zum wichtigen Endpunkt unerwünschte Ereignisse

Tabelle 24: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten)

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]	
SUE – Langzeiteffekt (> 6 Wochen)												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative Effektschätzung bedeutet einen Vorteil für die Prüfinterventionsgruppe.												
RCT, 5 [32,35,42,44]	schwer-wiegend ^d	nicht schwer-wiegend	schwer-wiegend ^e	unentdeckt	schwer-wiegend ^f	keine	24 / 453	15 / 488	2	1,67 [0,78; 3,61]	1 [-0,4; 4]	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 C zu entnehmen.</p> <p>d. Die adäquate Erzeugung der Randomisierungssequenz und die Verdeckung der Gruppenzuteilung war in 4 Studien unklar. 3 Studien waren nicht verblindet und in 4 Studien war das ITT-Prinzip nicht adäquat umgesetzt.</p> <p>e. Die Anwendungsart in der Vergleichsgruppe weicht von der deutschen Zulassung ab: es erfolgte eine intramuskuläre statt einer intraartikulären Injektion.</p> <p>f. Das 95 %-KI des OR überdeckt 1 und 2. Somit können weder Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfinterventionsgruppe ausgeschlossen werden.</p> <p>C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; N: Anzahl der Studien; OR: Odds Ratio; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; SUE: schwerwiegendes unerwünschtes Ereignis</p>												

Tabelle 25: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
Gesamtrate unerwünschte Ereignisse – Kurzzeiteffekt (< 2 Wochen)												
Interpretation der Effektschätzung: Ein RR kleiner als 1 bedeutet einen Vorteil für die Prüflnterventionsgruppe.												
RCT, 3 [21] ^d	<i>schwer-wiegend^e</i>	<i>nicht schwer-wiegend</i>	<i>schwer-wiegend^f</i>	k. A. ^g	<i>schwer-wiegend^h</i>	keine	158 / 253	145 / 286	507 pro 1000 ⁱ	1,20 [1,03; 1,39]	101 mehr pro 1000 [15 mehr bis 198 mehr]	<i>sehr niedrig^j</i>
<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 Bewertung der Studienlimitationen pro Studie pro Endpunkt ist dem Anhang D der NICE 2022 zu entnehmen.</p> <p>d. in NICE 2022 berücksichtigte Studien: <i>So 2010 und Schlesinger 2012a und Schlesinger 2012b</i></p> <p>e. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein hohes Verzerrungspotential hatte. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn die Mehrheit der Studien im Evidenzkörper ein sehr hohes Verzerrungspotenzial hatte</i> (Seite 143, Tabelle 21 Fußnote a [21]).</p> <p>f. Die Anwendungsart in der Vergleichsgruppe weicht von der deutschen Zulassung ab: es erfolgte eine intramuskuläre statt einer intraartikulären Injektion.</p> <p>g. Laut Studienprotokoll der NICE 2022 wurde der Publikationsbias erst dann untersucht, wenn der Evidenzkörper für einen Endpunkte > 5 Studien beinhaltete (Appendix A, Seite 51).</p> <p>h. <i>Die Qualität der Evidenz wurde um 1 Stufe abgewertet, wenn das 95 %-KI eine MID-Schwelle überschreitet. Die Qualität der Evidenz wurde um 2 Stufen abgewertet, wenn das 95 %-KI beide MID-Schwellen überschreitet. Etablierte MID-Schwellen für die kontinuierliche VAS: Verbesserungen ≥ 10 Punkte auf einer 1-100 Skala; Standard MID-Schwellen nach GRADE wurden für alle anderen Endpunkte verwendet. Für binäre Endpunkte betragen die MID-Schwellen RR von 0,8 und 1,25. Für kontinuierliche Endpunkte wurde die MID-Schwelle mit 0,5 x SD zu Studienbeginn berechnet</i> (Seite 132, Tabelle 19 Fußnote b [21]).</p> <p>i. eigene Berechnung</p> <p>j. von NICE 2022 abweichende Bewertung, da die Indirektheit mit schwerwiegend bewertet wurde</p>												

Tabelle 25: Evidenzprofil für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide für den wichtigen Endpunkt unerwünschte Ereignisse (binäre Daten) (aus NICE 2022 übernommen) (mehrseitige Tabelle)

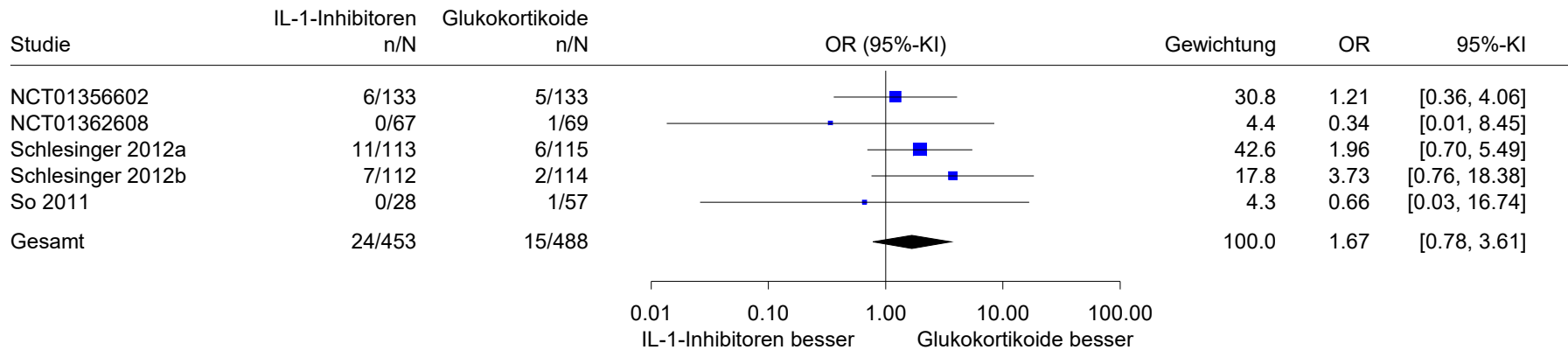
Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		RR [95 %-KI]	RD [95 %-KI]	
C: Vergleichsgruppe; GRADE: Grading of Recommendations Assessment, Development and Evaluation; I: Prüflinterventionsgruppe; k. A.: keine Angabe; KI: Konfidenzintervall; IL-1-Inhibitoren: Interleukin-1-Inhibitoren; N: Anzahl der Studien; RCT: randomisierte kontrollierte Studie; RD: absolute Risikodifferenz; RR: relatives Risiko; SD: Standardabweichung; VAS: visuelle Analogskala												

5.4.4.3.1 Metaanalysen

IL-1-Inhibitoren vs. Glukokortikoide

SUE

Modell mit zufälligen Effekten - Knapp und Hartung



Heterogenität: $Q=2.61$, $df=4$, $p=0.626$, $I^2=0\%$

Gesamteffekt: $Z\text{-Score}=1.86$, $p=0.137$, $\text{Tau(Paule-Mandel)}=0$

Abbildung 6: Metaanalyse, IL-1-Inhibitoren vs. Glukokortikoide, Endpunkt unerwünschte Ereignisse – Langzeiteffekt, Effektmaß: OR

6 Zusammenfassung und Anmerkungen zu den Ergebnissen

6.1 Zusammenfassung der Ergebnisse

Für die untersuchte Fragestellung zum Vergleich der Wirkstoffgruppen NSAR, IL-1-Inhibitoren, Glukokortikoide und Colchicin untereinander oder mit Placebo bei erwachsenen Patientinnen und Patienten mit einem Gichtanfall wurden insgesamt 12 RCTs identifiziert. Die Ergebnisse von je 10 Studien wurden für den kritischen Endpunkt Schmerzen und den wichtigen Endpunkt unerwünschte Ereignisse im EVB berücksichtigt. Zum wichtigen Endpunkt Funktionsfähigkeit gab es keine verwertbaren Daten für die Vergleiche Colchicin vs. Placebo und NSAR vs. Colchicin. Zum wichtigen Endpunkt Therapieabbruch aufgrund von unerwünschten Ereignissen gab es keine verwertbaren Daten für den Vergleich Colchicin vs. Placebo und IL-1-Inhibitoren vs. Glukokortikoide.

1 Studie untersuchte den Vergleich Colchicin vs. Placebo und berichtete Kurzzeiteffekte. Für den Endpunkt Schmerzen zeigten statistisch signifikant mehr Patientinnen und Patienten, die Colchicin erhielten, eine Reduktion der Schmerzintensität um mindestens 50 %. Das Vertrauen in diese Effektschätzung war mäßig (moderate Qualität der Evidenz). In beiden Gruppen traten keine SUEs auf, daher wurde die Qualität der Evidenz nicht bewertet. Für gastrointestinale unerwünschte Ereignisse (Diarrhö und Erbrechen) war die Effektschätzung nicht statistisch signifikant und von sehr niedriger Qualität der Evidenz.

Der Vergleich NSAR vs. Colchicin wurde in 2 Studien untersucht. Für die Endpunkte Schmerzen, Therapieabbruch aufgrund unerwünschter Ereignisse und unerwünschte Ereignisse in Form von SUE, Erbrechen, Übelkeit und / oder Erbrechen, Abdominalschmerzen, Dyspepsie waren die Effekte nicht statistisch signifikant (kurz- und mittelfristige Effekte mit sehr niedriger bis hoher Qualität der Evidenz). In Bezug auf den Endpunkt Schmerz als auch die genannten unerwünschten Ereignisse kann somit basierend auf den 2 Studien weder von einem Vorteil noch von einem Nachteil für NSAR gegenüber Colchicin gesprochen werden. Demgegenüber traten unter Colchicin statistisch signifikant mehr unerwünschte Ereignisse in Form von Diarrhö und Obstipation auf (Kurzzeiteffekte mit hoher Qualität der Evidenz).

4 Studien untersuchten den Vergleich NSAR vs. Glukokortikoide. In je 2 Studien zeigte sich kein statistisch signifikanter Unterschied zwischen beiden Wirkstoffgruppen für die Endpunkte Schmerzen und Funktionsfähigkeit. Das Vertrauen in diese Ergebnisse ist mäßig bis hoch. Für SUE und Therapieabbruch aufgrund unerwünschter Ereignisse lagen Ergebnisse aus jeweils 2 Studien vor. Beim Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse traten statistisch signifikant weniger Ereignisse bei Patientinnen und Patienten auf, die Glukokortikoide erhielten. Für den Endpunkt SUE traten in einer Studie keine Ereignisse auf, in der zweiten Studie traten auch hier statistisch signifikant weniger Ereignisse in der

Glukokortikoid-Gruppe auf. Das Vertrauen in diese Ergebnisse ist ebenfalls mäßig bis hoch. Für diesen Vergleich berichteten die Studien für alle Endpunkte Kurzzeiteffekte.

Für den Vergleich IL-1-Inhibitoren vs. Glukokortikoide wurden 5 Studien identifiziert. Die Qualität der Evidenz wurde zu allen Endpunkten mit niedrig bis sehr niedrig bewertet. Der Endpunkt Schmerzen war bei Patientinnen und Patienten in Folge einer IL-1-Inhibitoren-Behandlung statistisch signifikant weniger stark ausgeprägt. Zum Endpunkt körperliche Funktionsfähigkeit zeigte sich beim Kurzzeiteffekt ebenfalls ein statistisch signifikanter Vorteil der IL-1-Inhibitoren, die Ergebnisse zu Langzeiteffekten waren jedoch zu heterogen, um eine gepoolte Effektschätzung angeben zu können. Der Unterschied zwischen den Gruppen zu SUE war nicht statistisch signifikant, der Unterschied zwischen den Gruppen zur Gesamtrate unerwünschter Ereignisse zeigte einen knapp statistisch signifikanten Vorteil für die Glukokortikoide. Die Evidenz umfasst größtenteils Kurzzeiteffekte, lediglich für SUEs und körperliche Funktionsfähigkeit lagen auch Langzeiteffekte vor.

6.2 Anmerkungen zu den Ergebnissen

In der Gruppe der NSAR wurden die Wirkstoffe Indometacin, Etoricoxib und Naproxen untersucht. Es fehlen jedoch Studien zu weiteren NSAR-Wirkstoffen, z. B. Acemetacin und Phenylbutazon, die für die Behandlung von Gichtanfällen zugelassen sind. Für die Gruppe der Glukokortikoide wurde nur Evidenz für die Wirkstoffe Prednisolon und Triamcinolon dargestellt, da keine Studien mit anderen Wirkstoffen identifiziert wurden. Vorteile bzw. Nachteile für Glukokortikoide und IL-1-Inhibitoren im Vergleich zu Placebo und Colchicin sind nicht direkt abzuleiten, da Studien mit direkten Vergleichen mit Placebo oder Colchicin nicht identifiziert wurden.

In allen Studien zum Vergleich NSAR vs. Glukokortikoide wurde Prednisolon eingesetzt. Prednisolon ist in Deutschland nicht für die Behandlung eines Gichtanfalls zugelassen. Die Evidenzdarstellung für Prednisolon erfolgte auf expliziten Wunsch der Leitliniengruppe aufgrund der klinischen Relevanz des Wirkstoffes (orale Alternative zu NSAR).

Außer in Roddy 2020 und Janssens 2008 wich die Anwendung der Prüf- und / oder Vergleichsinterventionen in den Studien von den Angaben der deutschen Zulassung ab. Geringfügige Abweichungen in der Anwendung von Colchicin oder NSAR-Wirkstoffen wurden in Terkeltaub 2010, Liu 2019, Man 2007, Rainer 2016 und Xu 2016 registriert. In Man 2007 wurde zudem Diclofenac zusätzlich zu Indometacin eingesetzt, was laut deutscher Zulassung nicht empfohlen wird. In allen Studien zum Vergleich IL-1-Inhibitoren vs. Glukokortikoide wurde Triamcinolon intramuskulär statt intraartikulärer angewandt. Diese Abweichung wurde als schwerwiegende Indirektheit bewertet, da unklar ist, inwiefern sich eine zulassungskonforme Anwendung auf die Evidenz auswirkt.

Die Darstellung der Ergebnisse zu unerwünschten Ereignissen wurde für den Vergleich NSAR vs. Glukokortikoide auf die Darstellung von SUEs begrenzt, da sich in der Gesamtschau auf Basis der SUEs und dem Endpunkt Therapieabbruch aufgrund unerwünschter Ereignisse bereits ein Vorteil für die Glukokortikoide zeigte. In den zugehörigen Studien wurde in Rainer 2016 Indometacin und in Man 2007 Indometacin und Diclofenac als Prüfintervention eingesetzt. In der Studie Rainer 2016 traten keine SUEs auf, wohingegen in der Studie von Man 2007 statistisch signifikant mehr SUEs (alle gastrointestinal) in der NSAR-Gruppe auftraten. Sowohl die Unterschiede in der Prüfintervention als auch die in der Studienpopulation, welche bei Man 2007 Risikopatientinnen und -patienten mit vorangegangenen oberen gastrointestinalen Blutungen umfasste, wohingegen diese bei Rainer 2016 ausgeschlossen wurden, könnten zu diesen unterschiedlichen Ergebnissen beigetragen haben. Die in der NICE 2022 dargestellten Ergebnisse zu einzelnen gastrointestinalen unerwünschten Ereignissen basieren ebenfalls zum Großteil auf Man 2007 und Rainer 2016, und zeigen größtenteils in dieselbe Effektrichtung.

7 Literatur

1. Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG). Bundesgesetzblatt Teil 1 2019; (49): 2562-2584.
2. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Diagnostik und Therapie der Gicht; angemeldetes Leitlinienvorhaben [online]. 2021 [Zugriff: 21.02.2023]. URL: <https://register.awmf.org/de/leitlinien/detail/060-005>.
3. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358: j4008. <https://dx.doi.org/10.1136/bmj.j4008>.
4. Schünemann H, Brożek J, Guyatt G et al. GRADE Handbook [online]. 2013 [Zugriff: 19.02.2021]. URL: <https://gdt.gradepro.org/app/handbook/handbook.html>.
5. 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>.
6. 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.
7. 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>.
8. 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>.
9. 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>.
10. 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).
11. 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).
12. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.

13. 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.
14. 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>.
15. 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>.
16. 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>.
17. 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>.
18. 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>.
19. 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>.
20. 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>.
21. National Institute for Health and Care Excellence. Gout: diagnosis and management; [D] Evidence reviews for pharmacological and non-pharmacological interventions for managing gout flares [online]. 2022 [Zugriff: 26.05.2023]. URL: <https://www.nice.org.uk/guidance/ng219/evidence/d-pharmacological-and-nonpharmacological-interventions-for-managing-gout-flares-pdf-11080521329>.
22. Terkeltaub RA, Furst DE, Bennett K et al. High versus low dosing of oral colchicine for early acute gout flare: Twenty-four-hour outcome of the first multicenter, randomized, double-blind, placebo-controlled, parallel-group, dose-comparison colchicine study. *Arthritis Rheum* 2010; 62(4): 1060-1068. <https://dx.doi.org/10.1002/art.27327>.
23. Takeda. MPC-004 for the Treatment of an Acute Gout Flare [online]. 2012 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT00506883>.

24. Liu P, Chen Y, Wang B et al. Expression of microRNAs in the plasma of patients with acute gouty arthritis and the effects of colchicine and etoricoxib on the differential expression of microRNAs. *Arch Med Sci* 2019; 15(4): 1047-1055.
<https://dx.doi.org/10.5114/aoms.2018.75502>.
25. Roddy E, Clarkson K, Blagojevic-Bucknall M et al. Open-label randomised pragmatic trial (CONTACT) comparing naproxen and low-dose colchicine for the treatment of gout flares in primary care. *Ann Rheum Dis* 2020; 79(2): 276-284.
<https://dx.doi.org/10.1136/annrheumdis-2019-216154>.
26. Keele University. Colchicine Or Naproxen Treatment for ACute gouT [online]. 2019 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01994226>.
27. Janssens HJ, Janssen M, van de Lisdonk EH et al. Use of oral prednisolone or naproxen for the treatment of gout arthritis: a double-blind, randomised equivalence trial. *Lancet* 2008; 371(9627): 1854-1860. [https://dx.doi.org/10.1016/S0140-6736\(08\)60799-0](https://dx.doi.org/10.1016/S0140-6736(08)60799-0).
28. Man CY, Cheung IT, Cameron PA et al. Comparison of oral prednisolone/paracetamol and oral indomethacin/paracetamol combination therapy in the treatment of acute goutlike arthritis: a double-blind, randomized, controlled trial. *Ann Emerg Med* 2007; 49(5): 670-677.
<https://dx.doi.org/10.1016/j.annemergmed.2006.11.014>.
29. Rainer TH, Cheng CH, Janssens HJ et al. Oral Prednisolone in the Treatment of Acute Gout: A Pragmatic, Multicenter, Double-Blind, Randomized Trial. *Ann Intern Med* 2016; 164(7): 464-471. <https://dx.doi.org/10.7326/M14-2070>.
30. Xu L, Liu S, Guan M et al. Comparison of Prednisolone, Etoricoxib, and Indomethacin in Treatment of Acute Gouty Arthritis: An Open-Label, Randomized, Controlled Trial. *Med Sci Monit* 2016; 22: 810-817. <https://dx.doi.org/10.12659/msm.895749>.
31. Schlesinger N, De Meulemeester M, Pikhlak A et al. Canakinumab relieves symptoms of acute flares and improves health-related quality of life in patients with difficult-to-treat Gouty Arthritis by suppressing inflammation: results of a randomized, dose-ranging study. *Arthritis Res Ther* 2011; 13(2): R53. <https://dx.doi.org/10.1186/ar3297>.
32. So A, De Meulemeester M, Pikhlak A et al. Canakinumab for the treatment of acute flares in difficult-to-treat gouty arthritis: Results of a multicenter, phase II, dose-ranging study. *Arthritis Rheum* 2010; 62(10): 3064-3076. <https://dx.doi.org/10.1002/art.27600>.
33. Novartis. Targeted Dose Finding of Canakinumab (ACZ885) for Management of Acute Flare in Refractory or Contraindicated Gout Patients [online]. 2012 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT00798369>.

34. Novartis Pharma Services. An adaptive dose-ranging, multi-center, single-blind, double-dummy, active-controlled trial to determine the target dose of canakinumab (ACZ885) in the treatment of acute flares in gout patients who are refractory or contraindicated to NSAIDs and/or colchicine [online]. [Zugriff: 08.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-004666-61.
35. Schlesinger N, Alten RE, Bardin T et al. Canakinumab for acute gouty arthritis in patients with limited treatment options: results from two randomised, multicentre, active-controlled, double-blind trials and their initial extensions. *Ann Rheum Dis* 2012; 71(11): 1839-1848. <https://dx.doi.org/10.1136/annrheumdis-2011-200908>.
36. Novartis Pharma Services. A randomized, controlled extension study of CACZ885H2356 on the treatment and prevention of gout flares in patients with frequent flares for whom NSAIDs and/ or colchicine are contraindicated, not tolerated or ineffective [online]. [Zugriff: 11.07.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-017503-29.
37. Novartis Pharma Services. A randomized, controlled study of ACZ885 (canakinumab) on the treatment and prevention of gout flares in patients with frequent flares for whom NSAIDs and/or colchicine are contraindicated, not tolerated or ineffective [online]. [Zugriff: 08.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-015018-23.
38. Novartis Pharmaceuticals. Canakinumab in the Treatment of Acute Gout Flares and Prevention of New Flares in Patients Unable to Use Non-steroidal Anti-inflammatory Drugs (NSAIDs) and/or Colchicine Including a 12 Weeks Extension and an Open-label 48 Weeks Extension Study [online]. 2014 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01029652>.
39. Novartis Pharma Services. A randomized, controlled study of ACZ885 (canakinumab) on the treatment and prevention of gout flares in patients with frequent flares for whom NSAIDs and/ or colchicine are contraindicated, not tolerated or ineffective [online]. [Zugriff: 11.07.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-017802-35.
40. Novartis Pharma Services. A controlled extension study of CACZ885H2357 on the treatment and prevention of gout flares in patients with frequent flares for whom NSAIDs and/ or colchicine are contraindicated, not tolerated or ineffective [online]. [Zugriff: 11.07.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-018913-32.

41. Novartis Pharmaceuticals. Canakinumab in the Treatment of Acute Gout Flares and Prevention of New Flares in Patients Unable to Use Non-steroidal Anti-inflammatory Drugs (NSAIDs) and/or Colchicines Including a 12 Week Extension and a 1 Year Open-label Extension Study [online]. 2014 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01080131>.
42. Novartis Pharma Services. A randomized, double-blind, active-controlled study of canakinumab (ACZ885) pre-filled syringes or reconstituted lyophilizate versus triamcinolone acetonide for treating acute gouty arthritis flares in frequently flaring patients [online]. [Zugriff: 08.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024173-39.
43. Novartis Pharmaceuticals. Safety and Efficacy of Canakinumab Prefilled Syringes in Frequently Flaring Acute Gouty Arthritis Patients [online]. 2014 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01356602>.
44. Novartis Pharma Services. A randomized, double-blind, double-dummy, active controlled study of ACZ885 (canakinumab) on the treatment and prevention of gout flares in patients with frequent flares, for whom NSAIDs and/or colchicine are contraindicated, not tolerated or ineffective [online]. [Zugriff: 08.05.2023]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024172-26.
45. Novartis Pharmaceuticals. Safety & Efficacy of Canakinumab (ACZ885) in Patients With Frequent Flares for Whom Nonsteroidal Anti-Inflammatory Drug (NSAIDs) and/ or Colchicine Are Contraindicated, Not Tolerated or Ineffective [online]. 2017 [Zugriff: 08.05.2023]. URL: <https://ClinicalTrials.gov/show/NCT01362608>.
46. 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.
47. 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. Arnold DD, Yalamanoglu A, Boyman O. Systematic Review of Safety and Efficacy of IL-1-Targeted Biologics in Treating Immune-Mediated Disorders. *Front Immunol* 2022; 13: 888392. <https://dx.doi.org/10.3389/fimmu.2022.888392>.
2. McKenzie BJ, Wechalekar MD, Johnston RV et al. Colchicine for acute gout. *Cochrane Database Syst Rev* 2021; 8: CD006190. <https://dx.doi.org/10.1002/14651858.CD006190.pub3>.
3. National Institute for Health and Care Excellence. Gout: diagnosis and management; [D] Evidence reviews for pharmacological and non-pharmacological interventions for managing gout flares [online]. 2022 [Zugriff: 26.05.2023]. URL: <https://www.nice.org.uk/guidance/ng219/evidence/d-pharmacological-and-nonpharmacological-interventions-for-managing-gout-flares-pdf-11080521329>.
4. van Durme CM, Wechalekar MD, Landewe RB et al. Non-steroidal anti-inflammatory drugs for acute gout. *Cochrane Database Syst Rev* 2021; 12: CD010120. <https://dx.doi.org/10.1002/14651858.CD010120.pub3>.
5. Zeng L, Qasim A, Neogi T et al. Efficacy and Safety of Pharmacologic Interventions in Patients Experiencing a Gout Flare: A Systematic Review and Network Meta-Analysis. *Arthritis Care Res (Hoboken)* 2021; 73(5): 755-764. <https://dx.doi.org/10.1002/acr.24402>.

8.2 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

Nicht E3

1. Dorfler H. Treatment of acute attacks of gout with acemetacin. *Die Therapiewoche* 1985; 35(11): 1161-1166.
2. Yu TF, Burns JJ, Gutman AB. Results of a clinical trial of G-28315, a sulfoxide analog of phenylbutazone, as a uricosuric agent in gouty subjects. *Arthritis Rheum* 1958; 1(6): 532-543. <https://dx.doi.org/10.1002/art.1780010606>.

Nicht E5

1. Sturge RA, Scott JT, Hamilton EB et al. Multi-centre trial of naproxen and phenylbutazone in acute gout. *Adv Exp Med Biol* 1977; 76B: 290-296. https://dx.doi.org/10.1007/978-1-4684-3285-5_44.
2. Sturge RA, Scott JT, Hamilton EB et al. Multicentre trial of naproxen and phenylbutazone in acute gout. *Ann Rheum Dis* 1977; 36(1): 80-82. <https://dx.doi.org/10.1136/ard.36.1.80>.

Anhang A Bewertung der Qualität der Informationsbeschaffung der Basis-SÜ aus der fokussierten Recherche

Die Bewertung der Qualität der Informationsbeschaffung der Basis-SÜ erfolgte mit AMSTAR 2, Items 1, 4 und 5 [3].

Tabelle 26: Bewertung der Informationsbeschaffung zur systematischen Übersicht NICE 2022 [21]

1. Did the research questions and inclusion criteria for the review include the components of PICO?		
For Yes: <ul style="list-style-type: none"> ✓ Population ✓ Intervention ✓ Comparator group ✓ Outcome 	Optional (recommended) Timeframe for follow-up	<ul style="list-style-type: none"> ✓ Yes No
4. Did the review authors use a comprehensive literature search strategy?		
For Partial Yes (all the following): <ul style="list-style-type: none"> ✓ searched at least 2 databases (relevant to the research question) ✓ provided key words and/or search strategy ✓ justified publication restrictions (eg, language) 	For Yes, should also have (all the following): <ul style="list-style-type: none"> ✓ searched the reference lists/ bibliographies of included studies ✓ searched trial/study registries ✓ included/consulted content experts in the field ✓ where relevant, searched for grey literature ✓ conducted search within 24 months of completion of the review 	<ul style="list-style-type: none"> ✓ Yes Partial Yes No
5. Did the review authors perform study selection in duplicate?		
For Yes, either ONE of the following: <ul style="list-style-type: none"> at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include ✓ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 per cent), with the remainder selected by one reviewer 		<ul style="list-style-type: none"> ✓ Yes No
✓: Kriterium erfüllt		

Anhang B Bewertung der Qualität der Ergebnisse der Basis-SÜ aus der fokussierten Recherche

Die Bewertung der Qualität der Ergebnisse der Basis-SÜ erfolgte mit AMSTAR 2 [3].

Tabelle 27: Bewertung der Ergebnisse zur systematischen Übersicht NICE 2022 [21]

Item	Bewertung
1. Did the research questions and inclusion criteria for the review include the components of PICO?	yes
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	yes
3. Did the review authors explain their selection of the study designs for inclusion in the review?	yes
4. Did the review authors use a comprehensive literature search strategy?	yes
5. Did the review authors perform study selection in duplicate?	yes
6. Did the review authors perform data extraction in duplicate?	yes
7. Did the review authors provide a list of excluded studies and justify the exclusions?	yes
8. Did the review authors describe the included studies in adequate detail?	yes
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	yes
10. Did the review authors report on the sources of funding for the studies included in the review?	yes
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	yes
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or the other evidence synthesis?	yes
13. Did the review authors account for RoB in individual studies when interpreting / discussing the results of the review?	yes
14. Did the review authors provide a satisfactory explanation, for and discussion of, any heterogeneity observed in the results of the review?	no
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	yes
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	yes
Rating overall confidence in the results of the review	high

Anhang C Kriterien des Verzerrungspotenzials

Für die Evidenzprofile, die aus der NICE 2022 entnommen wurden, ist die Bewertung der Kriterien der Verzerrungspotenzials pro Studie pro Endpunkt dem Anhang D der NICE 2022 zu entnehmen.

Tabelle 28: Kriterien des Verzerrungspotenzials zum Endpunkt Schmerzen

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
So 2010	ja	ja	nein	ja	ja	ja
Schlesinger 2012a	unklar	unklar	nein	ja	ja	ja
Schlesinger 2012b	unklar	unklar	nein	ja	ja	ja
NCT01356602	unklar	unklar	ja	ja	ja	ja
NCT01362608	unklar	unklar	ja	ja	ja	ja
ITT: Intention to treat						

Tabelle 29: Kriterien des Verzerrungspotenzials zum Endpunkt Funktionsfähigkeit

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Janssens 2008	ja	ja	ja	ja	unklar	ja
Xu 2016	ja	unklar	nein	nein	unklar	ja
So 2010	ja	ja	nein	ja	nein	ja
Schlesinger 2012a	unklar	unklar	nein	nein	unklar	ja
Schlesinger 2012b	unklar	unklar	nein	nein	unklar	ja
ITT: Intention to treat						

Tabelle 30: Kriterien des Verzerrungspotenzials zum Endpunkt unerwünschte Ereignisse – SUE

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Terkeltaub 2010	unklar	unklar	ja	ja	unklar	ja
Roddy 2020	ja	ja	nein	ja	unklar	ja
Man 2007	ja	ja	unklar	ja	unklar	ja
Rainer 2016	ja	unklar	unklar	ja	unklar	ja
So 2010	ja	ja	nein	ja	ja	ja
Schlesinger 2012a	unklar	unklar	nein	nein	ja	ja
Schlesinger 2012b	unklar	unklar	nein	nein	ja	ja
NCT01356602	unklar	unklar	ja	nein	ja	ja
NCT01362608	unklar	unklar	ja	nein	ja	ja
ITT: Intention to treat						

Tabelle 31: Kriterien des Verzerrungspotenzials zum Endpunkt Therapieabbruch aufgrund von UE

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte
Liu 2019	ja	unklar	unklar	ja	unklar	ja
Man 2007	ja	ja	unklar	ja	unklar	ja
Rainer 2016	ja	unklar	unklar	ja	unklar	ja
ITT: Intention to treat						

Anhang D Suchstrategien

D.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 [46] – 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 April 27, 2023

Es wurde folgender Filter übernommen:

- RCT: Lefebvre [47] – 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	randomized controlled trial.pt.
8	controlled clinical trial.pt.
9	(randomized or placebo or randomly).ab.
10	clinical trials as topic.sh.
11	trial.ti.
12	or/7-11
13	exp animals/ not humans.sh.
14	12 not 13
15	and/6,14
16	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/
17	hi.fs. or case report.mp.
18	or/16-17
19	15 not 18
20	19 and (english or german or multilingual or undetermined).lg.
21	20 and 20210701:3000.(dt).
22	(Acemetacin* or Phenylbutazon*).mp.
23	and/6,20,22
24	21 or 23

2. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials: Issue 4 of 12, April 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	#6 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
#8	#7 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)))
#9	#8 with Cochrane Library publication date from Jul 2021 to present, in Trials
#10	Acemetacin* or Phenylbutazon*
#11	#8 and #10
#12	#9 or #11 in Trials

D.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 flare* [Trials with Results]

3. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

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

Suchstrategie
gout [with Results only]

D.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