

Arthroskopie des Kniegelenks bei Gonarthrose

Vorbericht (vorläufige Nutzenbewertung)

Auftrag: N11-01
Version: 1.0
Stand: 02.09.2013

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Arthroskopie des Kniegelenks bei Gonarthrose

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

26.07.2011

Interne Auftragsnummer:

N11-01

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
Im Mediapark 8 (KölnTurm)
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Nr. 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Bei dem vorliegenden Vorbericht handelt es sich um eine vorläufige Nutzenbewertung. Zu diesem Vorbericht können Stellungnahmen abgegeben werden, die zu einer Ergänzung und / oder Überarbeitung des Berichts führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG (www.iqwig.de), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vii
Abbildungsverzeichnis	ix
Abkürzungsverzeichnis.....	x
Kurzfassung	xi
1 Hintergrund	1
1.1 Definition des Krankheitsbildes	1
1.2 Epidemiologie und individuelle Krankheitslast	1
1.3 Ursache der Erkrankung	2
1.4 Verlauf der Erkrankung	2
1.5 Klassifikationen der Gonarthrose	2
1.6 Diagnostische Verfahren	3
1.7 Derzeit übliche Therapiepraxis.....	3
1.8 Therapieempfehlungen von Fachgesellschaften oder anderen Institutionen.....	4
2 Ziel der Untersuchung.....	6
3 Projektbearbeitung.....	7
3.1 Zeitlicher Verlauf des Projekts.....	7
3.2 Dokumentation der Änderungen im Projektverlauf	7
4 Methoden	9
4.1 Kriterien für den Einschluss von Studien in die Untersuchung	9
4.1.1 Population.....	9
4.1.2 Prüf- und Vergleichsintervention	9
4.1.3 Patientenrelevante Endpunkte	10
4.1.4 Studientypen	10
4.1.5 Studiendauer	11
4.1.6 Tabellarische Übersicht über die Kriterien für den Studieneinschluss	11
4.1.7 Einschluss von Studien, die vorgenannte Kriterien nicht vollständig erfüllen	11
4.2 Informationsbeschaffung.....	12
4.2.1 Bibliografische Literaturrecherche	12
4.2.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien.....	12
4.2.2.1 Systematische und nicht systematische Übersichten	12
4.2.2.2 Öffentlich zugängliche Studienregister	12
4.2.2.3 Durch den G-BA übermittelte Unterlagen	12

4.2.2.4	Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen .	13
4.2.2.5	Informationen aus der Anhörung	13
4.2.3	Selektion relevanter Studien.....	13
4.3	Informationsbewertung	14
4.4	Informationssynthese und -analyse	15
4.4.1	Gegenüberstellung der Ergebnisse der Einzelstudien	15
4.4.2	Meta-Analysen	16
4.4.3	Sensitivitätsanalyse	16
4.4.4	Subgruppenmerkmale und andere Effektmodifikatoren	17
4.5	Änderungen der Methodik.....	17
5	Ergebnisse.....	18
5.1	Ergebnisse der Informationsbeschaffung	18
5.1.1	Bibliografische Literaturrecherche	18
5.1.2	Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien.....	19
5.1.2.1	Systematische und nicht systematische Übersichten	19
5.1.2.2	Öffentlich zugängliche Studienregister.....	20
5.1.2.3	Durch den G-BA übermittelte Unterlagen	20
5.1.2.4	Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen .	21
5.1.2.5	Informationen aus der Anhörung	21
5.1.3	Resultierender Studienpool	21
5.2	Charakteristika der in die Bewertung eingeschlossenen Studien	22
5.2.1	Studiendesign und Studienpopulationen	24
5.2.2	Einschätzung des Verzerrungspotenzials auf Studienebene	39
5.3	Ergebnisse zu patientenrelevanten Endpunkten	41
5.3.1	Arthroskopie versus keine aktive Vergleichsintervention.....	41
5.3.1.1	Symptomatik der Gonarthrose	42
5.3.1.1.1	Schmerz.....	42
5.3.1.1.2	Körperliche Funktion	55
5.3.1.1.3	Globale Bewertung der Symptomatik.....	68
5.3.1.2	Gesundheitsbezogene Lebensqualität	73
5.3.1.3	Unerwünschte Therapiewirkungen	76
5.3.2	Arthroskopie versus aktive Behandlung.....	79
5.3.2.1	Symptomatik der Gonarthrose	79
5.3.2.1.1	Schmerz.....	80
5.3.2.1.2	Körperliche Funktion	85
5.3.2.1.3	Globale Bewertung der Symptomatik.....	90
5.3.2.2	Gesundheitsbezogene Lebensqualität	96

5.3.2.3	Unerwünschte Therapiewirkung	96
5.3.3	Zusammenfassung der Beleglage	98
6	Diskussion.....	102
6.1	Ergebnisse	102
6.2	Vergleich der Studien untereinander	102
6.3	Unerwünschte Therapiewirkungen.....	103
6.4	Verzerrungspotenzial.....	104
6.5	Studien mit „Placebo-Chirurgie“	104
6.6	Abgrenzung zu anderen Interventionen und Indikationen	105
6.7	Vergleich mit anderen Übersichtsarbeiten und Leitlinien.....	106
7	Fazit.....	108
8	Liste der eingeschlossenen Studien	109
9	Literatur	110
	Anhang A – Suchstrategien	120
	Anhang B – Liste der ausgeschlossenen Dokumente mit Ausschlussgründen	123
	Anhang C – Liste der in den Stellungnahmen an den G-BA genannten Referenzen	129
	Anhang D – Liste der gesichteten systematischen und nichtsystematischen Übersichten	132
	Anhang E – Ergebnisparameter und Messinstrumente	134
	Anhang F – Sensitivitätsanalysen: Auswertung zusätzlicher Skalen	138

Tabellenverzeichnis

	Seite
Tabelle 1: Behandlungsmöglichkeiten bei Gonarthrose	4
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss.....	11
Tabelle 3: In Studienregistern identifizierte relevante Studien.....	20
Tabelle 4: In Studienregistern identifizierte Studien unklarer Relevanz	20
Tabelle 5: Übersicht zu Autorenanfragen	21
Tabelle 6: Studienpool der Nutzenbewertung.....	22
Tabelle 7: Übersicht der in den Studien beschriebenen Vergleiche.....	23
Tabelle 8: Eingeschlossene Studien – Übersicht.....	26
Tabelle 9: Charakterisierung der Interventionen in den eingeschlossenen Studien.....	28
Tabelle 10: Ein- / Ausschlusskriterien für Patienten in den Studien.....	32
Tabelle 11: Charakterisierung der Studienpopulation.....	35
Tabelle 12: Bewertung des Verzerrungspotenzials auf Studienebene	40
Tabelle 13: Übersicht zur Extraktion von berichtsrelevanten Daten	41
Tabelle 14: Übersicht der verwendeten Instrumente zur Erfassung des Schmerzes beim Vergleich keine aktive Vergleichsintervention	42
Tabelle 15: Bewertung des Verzerrungspotenzials auf Endpunktebene: Schmerz, Vergleich keine aktive Vergleichsintervention	44
Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention	45
Tabelle 17: Subgruppenanalysen (Interaktionstest) zum Endpunkt Schmerz.....	51
Tabelle 18: Übersicht der verwendeten Instrumente zur Erfassung der körperlichen Funktion, Vergleich keine aktive Vergleichsintervention.....	55
Tabelle 19: Bewertung des Verzerrungspotenzials auf Endpunktebene: körperliche Funktion, Vergleich keine aktive Vergleichsintervention.....	58
Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention	59
Tabelle 21: Ergebnisse zum Endpunkt körperliche Funktion , Vergleich keine aktive Vergleichsintervention – Responderanalysen	64
Tabelle 22: Subgruppenanalysen körperliche Funktion.....	66
Tabelle 23: Übersicht der verwendeten Instrumente zur Erfassung der globalen Bewertung der Symptomatik beim Vergleich keine aktive Vergleichsintervention.....	68
Tabelle 24: Bewertung des Verzerrungspotenzials auf Endpunktebene: globale Bewertung der Symptomatik beim Vergleich keine aktive Vergleichsintervention	70
Tabelle 25: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich keine aktive Vergleichsintervention.....	71
Tabelle 26: Subgruppenanalysen globale Bewertung	73

Tabelle 27: Übersicht der verwendeten Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität, Vergleich keine aktive Vergleichsintervention	73
Tabelle 28: Bewertung des Verzerrungspotenzials auf Endpunktebene: gesundheitsbezogene Lebensqualität, Vergleich keine aktive Vergleichsintervention	75
Tabelle 29: Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität, Vergleich keine aktive Vergleichsintervention.....	75
Tabelle 30: Bewertung des Verzerrungspotenzials auf Endpunktebene: unerwünschte Therapiewirkungen, Vergleich keine aktive Vergleichsintervention.....	77
Tabelle 31: Ergebnisse zum Endpunkt unerwünschten Therapiewirkungen, Vergleich keine aktive Vergleichsintervention.....	78
Tabelle 32: Übersicht der verwendeten Instrumente zur Erfassung des Schmerz, Vergleich aktive Vergleichsintervention	80
Tabelle 33: Bewertung des Verzerrungspotenzials auf Endpunktebene: Schmerz, Vergleich aktive Vergleichsintervention.....	82
Tabelle 34: Ergebnisse zum Endpunkt Schmerz, Vergleich aktive Vergleichsintervention....	83
Tabelle 35: Ergebnisse zum Endpunkt Schmerz, Vergleich aktive Vergleichsintervention – Responderanalysen.....	84
Tabelle 36: Übersicht der verwendeten Instrumente zur Erfassung der körperlichen Funktion, Vergleich aktive Vergleichsintervention	85
Tabelle 37: Bewertung des Verzerrungspotenzials auf Endpunktebene: körperliche Funktion, Vergleich aktive Vergleichsintervention	87
Tabelle 38: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich aktive Vergleichsintervention	88
Tabelle 39: Übersicht der verwendeten Instrumente zur Erfassung der globalen Bewertung, Vergleich aktive Vergleichsintervention	90
Tabelle 40: Bewertung des Verzerrungspotenzials auf Endpunktebene: globale Bewertung, Vergleich aktive Vergleichsintervention	93
Tabelle 41: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich aktive Vergleichsintervention	94
Tabelle 42: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich aktive Vergleichsintervention – Responderanalysen	95
Tabelle 43: Bewertung des Verzerrungspotenzials auf Endpunktebene: unerwünschte Therapiewirkungen, Vergleich aktive Vergleichsintervention	97
Tabelle 44: Ergebnisse zum Endpunkt unerwünschte Therapiewirkungen, Vergleich aktive Vergleichsintervention	97
Tabelle 45: Landkarte der Beleglage in Bezug auf die patientenrelevanten Endpunkte.....	101
Tabelle 46: Ergebnisparameter und Messinstrumente, die in den ausgewerteten Studien verwendet und für den vorliegenden Bericht berücksichtigt wurden	134

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der bibliografischen Literaturrecherche und des Literaturscreenings	19
Abbildung 2: Übersicht über die relevanten Studien aus den verschiedenen Quellen der Informationsbeschaffung und daraus resultierender Studienpool	22
Abbildung 3: Meta-Analysen zum Schmerz, Vergleich keine aktive Vergleichsintervention	50
Abbildung 4: Subgruppenanalyse zur Art des arthroskopischen Eingriffs beim Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention	53
Abbildung 5: Subgruppenanalyse zur Art der Kontrollbehandlung beim Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention	54
Abbildung 6: Meta-Analysen zur körperlichen Funktion, Vergleich keine aktive Vergleichsintervention	65
Abbildung 7: Subgruppenanalyse zur Art der Kontrollbehandlung beim Endpunkt körperlichen Funktion, Vergleich keine aktive Vergleichsintervention	67
Abbildung 8: Meta-Analyse zur globalen Bewertung der Symptomatik der Gonarthrose; Vergleich keine aktive Vergleichsintervention	72
Abbildung 9: Auswertung zusätzlicher Skalen zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention	138
Abbildung 10: Auswertung zusätzlicher Skalen zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention	139

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ACR	American College of Rheumatism
AIMS-2 / AIMS	Arthritis Impact Measurement Scales
AS	Arthroskopie
EULAR	European League against Rheumatism
FU	Follow-up
G-BA	Gemeinsamer Bundesausschuss
HADS	hospital anxiety and depression scale
ICD-10-GM	Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision, German Modification
IQR	interquartile range
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	intention to treat
KOOS	Knee Injury and Osteoarthritis Outcome Scale
KSCRS	Knee Society Clinical Rating System
KSPS	Knee-Specific Pain Scale
LOCF	last observation carried forward
MACTAR	McMaster-Toronto Arthritis Patient Preference Disability Questionnaire
MID	Minimal Important Difference
MW	Mittelwert
NSAID	non-steroidal anti-inflammatory drugs
PFS	physical functioning scale
RCT	randomized controlled trial (randomisierte kontrollierte Studie)
SD	Standardabweichung
SF-36	Short Form (36) Health Survey
SMD	Standardized Mean Difference
VAS	visuelle Analogskala
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

Kurzfassung

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 26.07.2011 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung der therapeutischen Arthroskopie des Kniegelenks bei Gonarthrose beauftragt.

Fragestellung

Ziel der vorliegenden Untersuchung ist die Nutzenbewertung der therapeutischen Arthroskopie mit Lavage und gegebenenfalls Debridement im Vergleich zu keiner Behandlung oder einer Scheinbehandlung oder im Vergleich zu einer jeweils anderen Behandlung einschließlich Physiotherapie, Arzneimitteltherapie, Nahrungsergänzungsmitteln, physikalischer Therapie oder Osteotomie, bei Patienten mit Gonarthrose hinsichtlich patientenrelevanter Endpunkte.

Methoden

Es wurden randomisierte kontrollierte Studien (RCTs) eingeschlossen, die eine Behandlung mit einer Kniegelenksarthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement im Vergleich zu keiner zusätzlichen Behandlung, einer Scheinbehandlung oder einer anderen aktiven Behandlung untersuchten. Die Studien mussten zudem Patienten mit einer primären oder sekundären Gonarthrose eingeschlossen haben und patientenrelevante Endpunkte wie Symptomatik der Gonarthrose (insbesondere Schmerz), körperliche Funktion, gesundheitsbezogene Lebensqualität und unerwünschte Therapiewirkungen untersucht haben.

Hierzu wurde eine systematische Literaturrecherche in den folgenden Datenbanken durchgeführt: MEDLINE, EMBASE und Cochrane Central Register of Controlled Trials (Clinical Trials). Außerdem erfolgte eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanten Primärstudien sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments). Die systematischen Übersichten wurden hinsichtlich weiterer relevanter Studien durchsucht. Die Suche fand am 10.01.2012 statt.

Die Selektion relevanter Studien wurde je nach Suchquelle von 2 Reviewern unabhängig voneinander durchgeführt oder von einem Reviewer durchgeführt und durch einen zweiten überprüft. Eine Studie wurde nach dem Recherchezeitpunkt identifiziert und eingeschlossen.

Die Datenextraktion erfolgte in standardisierte Tabellen von einem Reviewer und wurde von einem zweiten Reviewer überprüft. Nach einer Bewertung des Verzerrungspotenzials wurden die Ergebnisse der einzelnen Studien, nach Endpunkten geordnet, beschrieben.

Ergebnisse

Die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement wurde insgesamt in 10 Studien untersucht. In 5 dieser Studien wurde als Vergleich keine aktive Vergleichsintervention (Scheinarthroskopie, diagnostische Arthroskopie, keine zusätzliche Behandlung) eingesetzt, in 5 Studien eine aktive Vergleichsintervention (intraartikuläre Injektion von Glukokortikoiden, intraartikuläre Injektion von Hyaluronsäure, Lavage oder orale Gabe von NSAIDs). Im Folgenden werden zunächst die Ergebnisse der Studien, in denen keine aktive Vergleichsintervention eingesetzt wurde, gemeinsam dargestellt. Danach erfolgt für jede der aktiven Vergleichsinterventionen eine separate Darstellung der Ergebnisse.

Arthroskopische Eingriffe im Vergleich zu keiner aktiven Vergleichsintervention

Es lagen 5 Studien vor, die die therapeutische Arthroskopie mit einer Scheinarthroskopie, einer diagnostischen Arthroskopie oder keiner zusätzlichen Behandlung verglichen. Die Ergebnisse dieser Studien wurden gemeinsam betrachtet. Es muss allerdings darauf hingewiesen werden, dass eine Studie lediglich eine kleine Pilotstudie war, deren Ergebnisse in dieser gemeinsamen Betrachtung kein Gewicht hatte.

Für die Endpunkte Schmerz und körperliche Funktion lagen jeweils Daten aus allen Studien vor, allerdings zeigten sich häufig heterogene Ergebnisse. Zusammenfassend zeigte sich kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Für die globale Bewertung der Symptomatik der Gonarthrose lagen nur Daten aus 2 Studien vor. Die Ergebnisse waren teilweise heterogen und es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Für die gesundheitsbezogene Lebensqualität lagen nur Daten aus einer Studie vor, ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie konnte nicht abgeleitet werden.

Für den Endpunkt unerwünschte Therapiewirkungen war die Datenlage insgesamt unzureichend, sodass keine eindeutige Aussage zum möglichen Schaden der therapeutischen Arthroskopie gemacht werden konnte.

Insgesamt konnte für die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen im Vergleich zu keiner aktiven Vergleichsintervention festgestellt werden.

Arthroskopische Eingriffe im Vergleich zu einer aktiven Vergleichsintervention

Insgesamt lagen 5 Studien vor, die die therapeutische Arthroskopie mit einer aktiven Intervention verglichen. Es lag jeweils eine Studie für den Vergleich therapeutische Arthroskopie versus intraartikuläre Injektion von Glukokortikoiden, Lavage oder orale Gabe

von NSAIDs vor. 2 Studien verglichen die therapeutische Arthroskopie mit der intraartikulären Injektion von Hyaluronsäure.

Arthroskopische Eingriffe versus Glukokortikoide

In der einzigen Studie, die die therapeutische Arthroskopie mit der intraartikulären Injektion von Glukokortikoiden verglich, lagen für die Endpunkte Schmerz und körperliche Funktion jeweils statistisch signifikante Effekte zugunsten der therapeutischen Arthroskopie vor. Die Konfidenzintervalle der Schätzer lagen jedoch jeweils nicht vollständig oberhalb der Irrelevanzschwelle, womit ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte. Es konnte daher auch kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie für diese Endpunkte abgeleitet werden.

Für den Endpunkt globale Bewertung der Symptomatik der Gonarthrose wurde in der eingeschlossenen Studie die Verbesserung der Kniesymptome basierend auf der Einschätzung der Patienten dargestellt. Für den Nachbeobachtungszeitpunkt 3 Monate lag das OR zugunsten der therapeutischen Arthroskopie bei 3,49 (95 %-KI: [1,73; 7,07]) und für den Zeitpunkt 6 Monate bei 4,07 (95 %-KI: [1,99; 8,33]). Aufgrund des hohen Verzerrungspotenzials auf Studien- und Endpunktebene konnte ein Anhaltspunkt jedoch kein Hinweis oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden.

Die Datenlage zum Endpunkt unerwünschte Therapiewirkungen war in der Studie unzureichend, sodass ein Schaden der therapeutischen Arthroskopie oder der intraartikulären Injektion von Glukokortikoiden nicht eingeschätzt werden konnte.

Für den Endpunkt gesundheitsbezogene Lebensqualität lagen keine Daten vor.

Arthroskopische Eingriffe versus Lavage

Für diesen Vergleich konnte eine Studie eingeschlossen werden, die Angaben zu den Endpunkten Schmerz, körperliche Funktion und der globalen Bewertung der Symptomatik der Gonarthrose machte. Für keinen Endpunkt konnte ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden. Für die Endpunkte gesundheitsbezogene Lebensqualität und unerwünschte Therapiewirkungen lagen aus der Studie keine Angaben vor.

Arthroskopische Eingriffe versus Hyaluronsäure

Der Vergleich therapeutische Arthroskopie versus Hyaluronsäure wurde in 2 Studien untersucht, allerdings lagen nur für den Endpunkt globale Bewertung der Symptomatik der Gonarthrose zum Zeitpunkt 12 Monate Daten aus beiden Studien vor. Für diesen Endpunkt waren die Ergebnisse beider Studien heterogen, es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Die Endpunkte Schmerz und körperliche Funktion wurden nur in einer der beiden Studien untersucht. Hier zeigte sich kein signifikanter Effekt zugunsten der therapeutischen

Arthroskopie, somit konnte für diese Endpunkte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Für die Endpunkte gesundheitsbezogene Lebensqualität und unerwünschte Therapiewirkungen lagen keine Daten vor.

Arthroskopische Eingriffe versus orale Gabe von NSAIDs

Für den Vergleich der therapeutischen Arthroskopie mit der oralen Gabe von NSAIDs lagen Daten aus einer Studie vor, die Angaben zu dem Endpunkt globale Bewertung der Symptomatik der Gonarthrose zu dem Zeitpunkt 36 Monate machte. Es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden. Für den Endpunkt unerwünschte Therapiewirkungen lagen unzureichende Angaben vor, sodass eine verlässliche Aussage über einen möglichen Schaden der therapeutischen Arthroskopie nicht gemacht werden konnte.

Für den Endpunkt gesundheitsbezogene Lebensqualität lagen keine Daten vor.

Fazit

Der Nutzen der therapeutischen Arthroskopie (mit Lavage und gegebenenfalls zusätzlichem Debridement) zur Behandlung einer Gonarthrose ist nicht belegt. Im Vergleich zu keiner aktiven Vergleichsintervention zeigte sich für keinen patientenrelevanten Endpunkt ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie. In den Vergleichen zur Lavage, oralen Gabe von NSAIDs oder der intraartikulären Injektion von Hyaluronsäure zeigte sich ebenfalls für keinen Endpunkt ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie. Lediglich in dem Vergleich zur intraartikulären Injektion von Glukokortikoiden ergab sich bei dem Endpunkt globale Bewertung der Symptomatik der Gonarthrose ein Anhaltspunkt für einen Nutzen der therapeutischen Arthroskopie.

Schlagwörter: Arthroskopie, Osteoarthrose – Knie, Nutzenbewertung, systematische Übersicht

Keywords: Arthroscopy, Osteoarthritis – Knee, Benefit Assessment, Systematic Review

1 Hintergrund

1.1 Definition des Krankheitsbildes

Die Gonarthrose (Kniegelenkarthrose) ist gekennzeichnet durch ein progressives Gelenkversagen, welches durch Strukturveränderungen, Schmerzen und Funktionsstörungen charakterisiert ist. Die Strukturveränderungen werden in der Regel in konventionellen Röntgenbildern sichtbar und bestehen aus der Gelenkspaltverschmälerung als indirekter Hinweis auf einen Knorpelverlust sowie Osteophyten und der Achsenfehlstellung.

Hierbei können 1, 2 oder alle 3 Anteile des Kniegelenks, nämlich mediales Kompartiment, laterales Kompartiment und Femoropatellargelenk, betroffen sein. Eine Arthrose führt in der Regel auch zu subchondralen Veränderungen am Knochen und Verdickungen an der Gelenkkapsel und an der Synovialis. Die periartikuläre Muskulatur ist häufig atrophiert. Die häufigsten Symptome sind Schmerzen und Gelenksteifheit. Funktionell zeigt sich eine verminderte Beweglichkeit, wodurch die täglichen Aktivitäten und die Lebensqualität erheblich eingeschränkt sein können. Neben belastungsabhängigen Gelenkschmerzen können im weiteren Verlauf auch Ruheschmerzen und Durchschlafstörungen auftreten. Die Lebensqualität und oft auch die Erwerbsfähigkeit der Betroffenen können erheblich beeinträchtigt werden [1]. Im höheren Alter kann auch eine unabhängige Lebensführung gefährdet sein.

Gemäß einer Festlegung des American College of Rheumatism¹ (ACR) sollte die Diagnose einer Gonarthrose gestellt werden, wenn folgende Kriterien erfüllt sind [2]: Knieschmerzen + Osteophyten + mindestens eines der folgenden 3 Kriterien: Alter über 50 Jahre, morgendliche Gelenksteifheit von weniger als 30 Minuten oder Knirschen bei Bewegung. Die European League against Rheumatism (EULAR) hat kürzlich ähnliche Empfehlungen für die Diagnose der Gonarthrose konsentiert. 3 Symptome (chronische Knieschmerzen, Morgensteifigkeit und Funktionseinschränkung) und 3 Krankheitszeichen (spürbares Reiben, eingeschränkte Beweglichkeit und Osteophyten) wurden als wichtigste Diagnosekriterien identifiziert [3].

1.2 Epidemiologie und individuelle Krankheitslast

Laut den Ergebnissen des Bundes-Gesundheitssurveys für das Jahr 2009 erkrankten in Deutschland 17,3 % aller Männer und 26,6 % aller Frauen im Laufe ihres Lebens an Arthrose, vorwiegend der Hüftgelenke und Kniegelenke (Lebenszeitprävalenz) [4]. In einem Alter von über 60 Jahren sind laut der WHO weltweit 9,6 % der Männer und 18,0 % der Frauen an einer symptomatischen Arthrose erkrankt [5].

Die Arthrose ist auch in anderen Ländern eine sehr häufige Erkrankung. Für das Jahr 2005 wurde geschätzt, dass 27 Millionen Menschen in den USA an einer klinisch relevanten Arthrose erkrankt waren. Ausgehend von einer Gesamtbevölkerung von etwa 300 Millionen Menschen lag die Prävalenz zu diesem Zeitpunkt bei 9 % [6].

¹ frühere Bezeichnung: American Rheumatism Association

Je nach Ätiologie sind häufig beide Knie von einer Gonarthrose betroffen. Bei deutschen Patienten, die eine Kniegelenksendoprothese erhielten, wurde in 87 % der Fälle eine bilaterale Gonarthrose und in 34 % eine generalisierte Arthrose, also ein Befall weiterer Gelenke, beobachtet [7].

Im Jahr 2009 wurden in deutschen Krankenhäusern 73 975 Männer und 131 716 Frauen wegen einer Gonarthrose behandelt [8]. Hierbei wurden bei vollstationären Patienten 366 182 arthroskopische Interventionen am Kniegelenk durchgeführt [9], davon 18 % bei Patienten mit einer Gonarthrose als Hauptdiagnose (ICD-10-GM-Codes der Kategorie M17).

1.3 Ursache der Erkrankung

Es werden primäre und sekundäre Formen der Arthrose unterschieden. Falls keine spezifische Ursache bekannt ist, wird von einer primären Arthrose ausgegangen. Zu den möglichen Ursachen einer sekundären Arthrose gehören Gelenkverletzungen [10], angeborene Fehlstellungen des Gelenkes und lokale oder generalisierte Krankheiten wie zum Beispiel bakterielle Kniegelenksinfektionen, rheumatoide Arthritis oder Hämophilie [11]. Zu Risikofaktoren für die Entwicklung einer Gonarthrose gehören beispielsweise Alter, Geschlecht, Ethnizität, genetische Prädisposition und Übergewicht [12,13].

1.4 Verlauf der Erkrankung

Die Arthrose ist sowohl hinsichtlich der strukturellen Schäden als auch der Symptome eine im Regelfall chronisch fortschreitende Erkrankung [14]. Es lässt sich auf individueller Basis nicht sicher vorhersagen, ob der Patient eine Progression zeigen wird. Bei den Symptomen nehmen die belastungsabhängigen Schmerzen zu, und es können Ruheschmerzen und Nachtschmerzen auftreten. Mit Fortschreiten der Erkrankung kommt es zu einer Verminderung der Beweglichkeit (Extension / Flexion), einer zunehmenden Achsenfehlstellung und einer Instabilität im Kniegelenk. Damit verbunden nehmen die Einschränkungen bei den Aktivitäten des täglichen Lebens, zum Beispiel Treppensteigen oder Aufheben eines Gegenstandes vom Boden, zu.

1.5 Klassifikationen der Gonarthrose

Für die Einteilung des Schweregrads der Gonarthrose gibt es eine Reihe von Klassifikationen. Zur radiologischen Beurteilung werden meist die Stadieneinteilungen nach der Kellgren- und Lawrence-Klassifikation oder auch nach Ahlbäck 1968 empfohlen und verwendet [15]. Zusätzlich wird häufig die Einteilung nach Altman 1987 zitiert [16]. Schiphof 2008 weist darauf hin, dass die Kellgren- und Lawrence-Klassifikation des Öfteren modifiziert angewendet wird [17]. Darüber hinaus wird das Ausmaß arthrosebedingter Knorpelschädigungen auch mittels offener makroskopischer Inspektion oder Arthroskopie nach Outerbridge 1961 eingeteilt.

1.6 Diagnostische Verfahren

Eine Anamnese und eine körperliche Untersuchung werden bei Patienten mit Beschwerden zuerst durchgeführt. Falls die Ergebnisse auf eine Gonarthrose deuten, erfolgen weitere Untersuchungen, wie z. B. eine Röntgenaufnahme, um Osteophyten nachzuweisen [14]. Mithilfe der konventionellen Röntgenaufnahme kann zudem der Schweregrad einer Gonarthrose anhand sekundärer Veränderungen ergänzend eingeschätzt werden: beispielsweise Verengung des Gelenkspaltes, subchondrale Sklerosierung und Zystenbildung. Da die konventionelle Röntgenaufnahme eine Erkrankung nicht verlässlich nachweisen kann, ist sie insbesondere nicht für die Darstellung eines frühen Krankheitsstadiums geeignet [18,19]. In ausgewählten Fragestellungen kann zur Differenzierung von anderen Erkrankungen (zum Beispiel der avaskulären Knochennekrose) eine Magnetresonanztomografie sinnvoll sein. Biochemische Marker gelten momentan als nicht geeignet, relevante diagnostische oder prognostische Zusatzinformationen zu liefern [14].

1.7 Derzeit übliche Therapiepraxis

Es besteht ein breites Spektrum an Behandlungsmöglichkeiten. Die Optionen sind aber nicht alle hinsichtlich ihrer Wirksamkeit belegt. Es werden die operativen von den nicht operativen Behandlungsmöglichkeiten unterschieden [15,18]. Eine beispielhafte Übersicht erfolgt in Tabelle 1.

Bei der medikamentösen Behandlung werden nichtsteroidale Entzündungshemmer (non-steroidal anti-inflammatory drugs, NSAIDs) mit antiphlogistischer und analgetischer Wirkung wie beispielsweise Ibuprofen und Analgetika wie beispielsweise Paracetamol (Acetaminophen) als Monotherapie oder in Kombination als Schmerzmittel der ersten Wahl für die Behandlung des chronischen Schmerzes bei Osteoarthritis eingesetzt [20,21].

Eine weitere Behandlungsmöglichkeit ist Lavage. Während eine blind durchgeführte Lavage (nicht arthroskopisch) zu den nicht operativen Behandlungsmöglichkeiten zählt, ist eine Lavage unter Sicht (arthroskopisch) ein operativer Eingriff. In beiden Fällen ist es das Ziel, mittels der Infundierung von physiologischer Kochsalzlösung das Gelenk zu spülen [22]. Dabei sollen frei im Gelenk liegende Partikel wie beispielsweise Knorpel, Kalziumpyrophosphatkristalle und freie Synovia entfernt werden. Beim Debridement wird mit zusätzlichen Instrumenten versucht, raue Knorpeloberflächen und flottierende Meniskusteile abzutragen. Dabei sollte das Debridement immer eine Lavage beinhalten, um frei liegende sowie durch das Debridement entstehende Partikel zu entfernen.

Tabelle 1: Behandlungsmöglichkeiten bei Gonarthrose

Kategorie		Beispiele
Nicht operativ	Nicht-medikamentös	Aufklärung über Erkrankung, Risikofaktoren und therapeutische Optionen Gewichtsreduktion physiotherapeutische Anwendungen (z. B. Bewegungstherapien, Elektrostimulation, Ultraschall) Akupunktur orthopädische Hilfen (z. B. Schuheinlagen) Nahrungsergänzungsmittel (z. B. Chondroitin, Glukosamine, S-Adenosyl-L-Methionin, Vitamine) intraartikuläre Lavage intraartikuläre Glukokortikosteroide
	Medikamentös (oral)	Paracetamol (Acetaminophen) NSAID (z. B. ASS, Ibuprofen) Opioide
Operativ		Arthroskopie (ggfs. mit Lavage und ggf. zusätzlich mit Debridement (das Debridement sollte immer eine Lavage beinhalten)) Umstellungsosteotomie Synovektomie Knorpeltechniken (z. B. Pridie-Bohrung) partieller oder totaler Kniegelenkersatz (Endoprothese)
ASS: Acetylsalicylsäure; NSAID: non-steroidal anti-inflammatory drugs		

Weitere operative Behandlungsmöglichkeiten lassen sich in gelenkerhaltende und gelenkersetzende Eingriffe unterteilen. Als gelenkerhaltende Operationen sind primär die Synovektomie, die Umstellungsosteotomie und die Pridie-Bohrung zu nennen. Zu den gelenkersetzenden Eingriffen werden hauptsächlich die unikondyläre Knieendoprothese sowie die bikondyläre Oberflächenersatzprothese (auch als Knieendoprothese bezeichnet) gezählt [15].

Die Behandlung der Gonarthrose lässt sich als Pyramide mit Therapiestufen verstehen, die in Abhängigkeit vom Schweregrad der Erkrankung hinzugefügt werden [23]. Die untere Stufe enthält einfache Empfehlungen zur Selbsthilfe. Das Addieren weiterer, komplexer Stufen wie z. B. medikamentöse und intraartikuläre Behandlungsmöglichkeiten und letztendlich operativer Verfahren (siehe Tabelle 1) ist abhängig vom Schweregrad der Erkrankung.

1.8 Therapieempfehlungen von Fachgesellschaften oder anderen Institutionen

Die *European League Against Rheumatism (EULAR)* hat zuletzt im Jahr 2003 evidenzbasierte Empfehlungen zur Behandlung der Gonarthrose veröffentlicht [24]. Auch die *EULAR* hat Empfehlungen in Stufen formuliert, die – inhaltlich betrachtet – den Empfehlungen des *Osteoarthritis Research Society International (OARSI) Treatment Guidelines Committee* weitgehend entsprechen [25]. Evidenzbasierte deutsche Leitlinien zum Thema sind auf den Seiten der Arbeitsgemeinschaften wissenschaftlicher medizinischer Fachgesellschaften (AWMF) nicht abrufbar. Eine Leitlinie *Gonarthrose* der Deutschen Gesellschaft für

Orthopädie und Orthopädische Chirurgie (DGOOC) wird derzeit aktualisiert und soll im Dezember 2013 fertiggestellt werden [26]. Das *American College of Rheumatism (ACR)* hat im April 2012 evidenzbasierte Empfehlungen zur medikamentösen und nichtmedikamentösen Behandlung der Gonarthrose veröffentlicht [24]. Die arthroskopische Behandlung wird in den Empfehlungen nicht erwähnt. Gemäß den internationalen Leitlinien gelten folgende wesentlichen Therapieprinzipien:

- Eine Kombination von medikamentöser und nichtmedikamentöser Behandlung ist optimal.
- Die nichtmedikamentöse Behandlung sollte Schulung, Sport, orthopädische Hilfen und Gewichtsreduktion einschließen.
- Paracetamol ist das bevorzugte orale Langzeitanalgetikum und sollte zuerst verwendet werden.
- NSAID sollten verwendet werden, wenn Paracetamol nicht ausreichend wirkt.
- Opioidanalgetika sind eine Alternative, wenn NSAID nicht wirken oder kontraindiziert sind.
- Die intraartikuläre Injektion von Kortikosteroiden ist indiziert bei starken Knieschmerzen, insbesondere, wenn sie von Entzündungszeichen und Erguss begleitet sind („aktivierte Arthrose“).
- Ein Gelenkersatz sollte bei Patienten mit fortbestehenden Schmerzen und Funktionseinschränkung in Erwägung gezogen werden.

2 Ziel der Untersuchung

Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung der therapeutischen Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement im Vergleich zu keiner Behandlung oder einer Scheinbehandlung oder im Vergleich zu einer jeweils anderen Behandlung einschließlich Physiotherapie, Arzneimitteltherapie, Nahrungsergänzungsmitteln, physikalischer Therapie oder Osteotomie

bei Patienten mit Gonarthrose hinsichtlich patientenrelevanter Endpunkte.

3 Projektbearbeitung

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 26.07.2011 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung der therapeutischen Arthroskopie des Kniegelenks bei Gonarthrose beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurde am 06.12.2011 ein Patientenvertreter der Deutschen Rheuma-Liga zur Festlegung patientenrelevanter Zielgrößen konsultiert.

Der vorläufige Berichtsplan in der Version 1.0 vom 09.01.2012 wurde am 18.01.2012 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 15.02.2012 konnten schriftliche Stellungnahmen eingereicht werden. Unklare Aspekte aus den schriftlichen Stellungnahmen zum vorläufigen Berichtsplan wurden am 27.03.2012 in einer wissenschaftlichen Erörterung mit den Stellungnehmenden diskutiert. Die Dokumentation und Würdigung der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht.

Im Anschluss an die Anhörung wurde ein überarbeiteter Berichtsplan (Version 1.0 vom 14.06.2012) publiziert.

Der vorliegende Vorbericht wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf der Website des IQWiG in einem entsprechenden Leitfaden dargelegt sind. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichts führen. Im Anschluss an diese Anhörung wird der Abschlussbericht erstellt. Dieser Bericht wird an den G-BA übermittelt und 8 Wochen später auf der Website des IQWiG veröffentlicht.

3.2 Dokumentation der Änderungen im Projektverlauf

Berichtsplan im Vergleich zum vorläufigen Berichtsplan

- Im Abschnitt 1.2 wurden Angaben zur Häufigkeit arthroskopischer Eingriffe eingefügt, um den Unterschied zwischen arthroskopischen Einzelinterventionen und der Arthroskopie insgesamt deutlich zu machen.
- In Abschnitt 1.5 wurde die Indikationsstellung zur Behandlung einer Gonarthrose präzisiert.

- In Kapitel 2 wurde das Ziel der Untersuchung dahin gehend konkretisiert, dass mit der Prüfintervention eine Kniegelenksarthroskopie mittels Lavage und gegebenenfalls zusätzlichem Debridement gemeint ist. Ferner wurden im Abschnitt 4.1.2 das Debridement und die diagnostische Arthroskopie begrifflich präziser eingegrenzt.
- In Abschnitt 4.1.1 wurde detaillierter erläutert, unter welchen Bedingungen eine Gonarthrosedefinition als hinreichend ähnlich im Vergleich zur ACR-Definition gelten kann.
- Der Endpunkt „Notwendigkeit eines operativen Gelenkersatzes“ wurde im Abschnitt 4.1.3 gestrichen.

Vorbericht im Vergleich zum Berichtsplan

- Aktualisierungen und redaktionelle Änderungen im gesamten Kapitel 1.
- In Abschnitt 4.1.3 wurde der Endpunkt „Symptomatik der Gonarthrose“ eingeführt und die patientenrelevanten Endpunkte Schmerz und körperliche Funktion wurden diesem zugeordnet. Durch die Einführung dieses neuen Endpunkts konnte auch die globale Bewertung der Symptomatik diesem Endpunkt zugeordnet und dargestellt werden.
- In Abschnitt 4.4.3 wurde ergänzt, dass Sensitivitätsanalysen auch dann durchgeführt wurden, falls für einen Endpunkt verschiedene Instrumente in den Studien eingesetzt wurden.

4 Methoden

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

Gegenstand der Nutzenbewertung sind solche therapeutischen arthroskopischen Eingriffe bei primärer oder sekundärer Gonarthrose, die eine oder mehrere therapeutische Maßnahmen aus dem Bereich des Debridement (Entfernung krankhaften oder störenden Gewebes / Materials) oder der Gelenkspülung (Lavage) sowie Eingriffe an der Synovialis, den Gelenkknorpeln und Menisken umfassen [27]. Das Debridement umfasst die Entfernung von Meniskusanteilen, Knorpelanteilen der Gelenkflächen, Gelenkhaut (Synovia) und freien Gelenkkörpern und sollte eine zusätzliche Lavage zur Entfernung freier Partikel beinhalten.

Nicht vom Antrag umfasst werden solche Eingriffe an der Synovialis, den Gelenkknorpeln und Menisken, die aufgrund nachgewiesener gravierender Veränderungen (beispielsweise Meniskusriss) an den genannten Strukturen oder wegen freier Gelenkkörper („Gelenkblockade“ als führendes Symptom) durchgeführt werden, sofern die Symptome zuverlässig auf diese Veränderungen zurückzuführen und damit durch eine mechanische Intervention zu beeinflussen sind.

4.1.1 Population

In die Untersuchung wurden Patienten mit primärer oder sekundärer Arthrose des Kniegelenks primär gemäß den Kriterien für die Diagnose einer Gonarthrose des American College of Rheumatism (ACR) [2] eingeschlossen. Alternativ konnten auch andere ähnliche Definitionen einer Gonarthrose zum Patienteneinschluss herangezogen werden, sofern diese auf den klassischen Kriterien einer Arthrose (Schmerzen, Gelenksteifigkeit, Knirschen oder Reiben im Gelenk und / oder Osteophyten) beruhten.

Nicht eingeschlossen wurden Studien mit Patienten, bei denen die präoperativen Symptome oder die intraoperativen Befunde anzeigten, dass gravierende, primär nicht arthrotische Veränderungen an den Kniebinnenstrukturen (zum Beispiel Meniskusriss, freier Gelenkkörper) die Symptomatik dominierten. Dies war insbesondere dann anzunehmen, wenn eine mechanische Beeinträchtigung (zum Beispiel Kniegelenkblockade) angegeben wurde.

4.1.2 Prüf- und Vergleichsintervention

Für den vorliegenden Bericht wurden folgende Prüfinderventionen betrachtet:

- arthroskopische Eingriffe am Kniegelenk, die eine oder mehrere therapeutische Maßnahmen aus dem Bereich des Debridements oder der Lavage sowie Eingriffe an der Synovialis, den Gelenkknorpeln und Menisken umfassen.

Für den vorliegenden Bericht wurde folgende Vergleichsinterventionen betrachtet:

- keine therapeutische Intervention,
- Scheinarthroskopie,
- diagnostische Arthroskopie,
- Lavage (ohne Arthroskopie),
- nicht operative Behandlung (medikamentöse und nichtmedikamentöse, siehe Beispiele in Tabelle 1),
- operative Behandlungen, die nicht vom Antrag umfasst werden (offen chirurgisch oder arthroskopisch).

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung wurden folgende patientenrelevante Endpunkte betrachtet:

- Symptomatik der Gonarthrose
 - Schmerz
 - körperliche Funktion
 - globale Bewertung (zum Beispiel der WOMAC-Globalscore als Kombination der Scores zu Schmerzen, Gelenksteifigkeit und körperlicher Funktion)
- gesundheitsbezogene Lebensqualität
- unerwünschte Therapiewirkungen (zum Beispiel Mortalität oder Kniegelenkinfektion)

Subjektive Endpunkte (z. B. gesundheitsbezogene Lebensqualität) wurden nur dann berücksichtigt, wenn sie mit validen Messinstrumenten (z. B. validierten Skalen) erfasst wurden.

4.1.4 Studientypen

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle unter 4.1.2 genannten Interventionen und alle unter 4.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

In den zu erstellenden Bericht wurden daher ausschließlich RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung eingeschlossen.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer bestand keine Einschränkung.

4.1.6 Tabellarische Übersicht über die Kriterien für den Studieneinschluss

Die folgende Tabelle zeigt die Kriterien für den Einschluss von Studien in die Bewertung.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Patienten mit Gonarthrose (siehe auch Abschnitt 4.1.1)
E2	Behandlung durch einen arthroskopischen Eingriff am Kniegelenk (siehe auch Abschnitt 4.1.2)
E3	Vergleichsintervention: keine therapeutische Intervention Scheinarthroskopie diagnostische Arthroskopie Lavage (ohne Arthroskopie) Behandlung mit nicht operativen Verfahren Behandlung mit anderen operativen Verfahren, die nicht vom Antrag umfasst werden (offen oder arthroskopisch) (siehe auch Abschnitt 4.1.2)
E4	patientenrelevante Endpunkte (siehe auch Abschnitt 4.1.3)
E5	randomisierte kontrollierte Studien (siehe auch Abschnitt 4.1.4)
E6	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT-Statements [28] genügt und eine Bewertung der Studie ermöglicht.	

4.1.7 Einschluss von Studien, die vorgenannte Kriterien nicht vollständig erfüllen

Für das Einschlusskriterium E1 (Population) reichte es aus, wenn es bei mindestens 80 % der eingeschlossenen Patienten erfüllt war. Lagen für solche Studien entsprechende Subgruppenanalysen vor, wurde auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1 bei weniger als 80 % erfüllt war, wurden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorlagen.

Ebenfalls eingeschlossen wurden Studien, die zu mindestens 80 % das Einschlusskriterium E2 erfüllten (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und zu mindestens 80 % das Einschlusskriterium E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie).

4.2 Informationsbeschaffung

4.2.1 Bibliografische Literaturrecherche

Die systematische Literaturrecherche nach relevanten Studien wurde in folgenden bibliografischen Datenbanken durchgeführt:

- Suche nach Primärstudien in den Datenbanken MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (Clinical Trials)
- Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanter Primärliteratur sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments)

Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Suche fand am 10.01.2012 statt.

4.2.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien zu ermitteln, wurden weitere Quellen berücksichtigt. Diese Rechercheergebnisse wurden anschließend auf weitere relevante Studien und Studienunterlagen untersucht (siehe Abschnitt 4.2.3).

4.2.2.1 Systematische und nicht systematische Übersichten

Relevante systematische und nicht systematische Übersichten wurden hinsichtlich weiterer relevanter Publikationen bzw. Studien gesichtet.

4.2.2.2 Öffentlich zugängliche Studienregister

Die folgenden öffentlich zugänglichen Studienregister wurden durchsucht:

- U.S. National Institutes of Health. ClinicalTrials.gov [online]. URL: <http://www.clinicaltrials.gov>
- World Health Organization. International Clinical Trials Registry Platform Search Portal [online]. URL: <http://apps.who.int/trialsearch>

Die Suche in öffentlich zugänglichen Studienregistern fand am 03.02.2012 statt.

4.2.2.3 Durch den G-BA übermittelte Unterlagen

Der G-BA erhielt Stellungnahmen zum Projekt, in denen auf Literatur zum Thema verwiesen wurde. Diese Referenzen wurden in den Screeningprozess einbezogen.

4.2.2.4 Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen

Es war geplant, Anfragen an die Autoren der Publikationen zu stellen, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten ließen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen waren und keine Sponsoren, i. d. R. die Hersteller einer der geprüften Therapieoptionen, zu einer Studie identifizierbar waren.

4.2.2.5 Informationen aus der Anhörung

Im Anschluss an die Veröffentlichung des vorläufigen Berichtsplans erfolgte eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen konnte. Relevante Informationen aus dieser Anhörung konnten in die Nutzenbewertung einfließen.

4.2.3 Selektion relevanter Studien

Selektion relevanter Publikationen aus den Ergebnissen der bibliografischen Literaturrecherche

Die durch die Suche in bibliografischen Datenbanken identifizierten Zitate wurden in einem ersten Schritt anhand ihres Titels und, sofern vorhanden, Abstracts auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 2) bewertet. Als potenziell relevant erachtete Publikationen wurden in einem zweiten Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgten durch 2 Reviewer unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen den beiden Reviewern aufgelöst.

Selektion relevanter Studien aus weiteren Suchquellen

Informationen aus den folgenden Suchquellen wurden von 2 Reviewern unabhängig voneinander auf ihre Relevanz bewertet:

- öffentlich zugängliche Studienregister
- durch den G-BA übermittelte Unterlagen

Informationen aus den folgenden Suchquellen wurden von einem Reviewer auf Studien gesichtet, der diese dann auf ihre Relevanz bewertete; ein zweiter Reviewer überprüfte den gesamten Prozess inklusive der Bewertungen

- im Rahmen der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht eingereichte Informationen.

Die identifizierten relevanten systematischen und nicht systematischen Übersichten wurden nach weiteren potenziell relevanten Studien durchsucht, deren Relevanz von 2 Reviewern unabhängig voneinander geprüft wurde.

Sofern in einem der genannten Selektionsschritte Diskrepanzen auftraten, wurden diese jeweils durch Diskussion zwischen den beiden Reviewern aufgelöst.

4.3 Informationsbewertung

Die Bewertung der Informationen der eingeschlossenen Studien hängt stark von den verfügbaren Angaben und der Qualität der jeweiligen Publikationen und weiterer Informationsquellen ab. Alle für die Nutzenbewertung relevanten Ergebnisse wurden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft.

Datenextraktion

Alle für die Nutzenbewertung notwendigen Informationen wurden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert.

Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wurde für jede in die Nutzenbewertung eingeschlossene Studie bewertet, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu wurden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet:

A: Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des Behandlers
- ergebnisgesteuerte Berichterstattung

B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung

Das Verzerrungspotenzial wurde als *niedrig* oder *hoch* eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts wurde zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als *niedrig* oder *hoch* eingestuft. Falls diese Einstufung als *hoch* erfolgte, wurde das Verzerrungspotenzial für den Endpunkt in der

Regel auch als *hoch* bewertet. Ansonsten fanden die unter B genannten endpunktspezifischen Aspekte Berücksichtigung.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als *hoch* führte nicht zum Ausschluss aus der Nutzenbewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

4.4 Informationssynthese und -analyse

Die Informationen wurden einer Informationssynthese und -analyse unterzogen. Wenn möglich wurden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Werkzeuge eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgte darüber hinaus in jedem Fall.

4.4.1 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten wurden im Bericht vergleichend beschrieben.

Die in den Studien berichteten jeweiligen Endpunkte wurden den im Berichtsplan definierten patientenrelevanten Endpunkten (siehe Abschnitt 4.1.3) zugeordnet.

In einigen Studien lagen Ergebnisse aus mehreren Instrumenten zu einem patientenrelevanten Endpunkt vor. In den Ergebnistabellen sind diese Ergebnisse vollständig für jedes Instrument dargestellt. Wenn in den Studien mehrere Instrumente zu einem Endpunkt erhoben wurden, wurde das Instrument, das in den Studien am häufigsten verwendet wurde, primär in der Meta-Analyse berücksichtigt. Diese wurde im Ergebnisteil dargestellt. Die alternativen Instrumente wurden im Rahmen von Sensitivitätsanalysen in weiteren Meta-Analysen betrachtet und im Anhang F dargestellt.

Ergebnisse flossen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basierten, das heißt, wenn der Anteil der Patienten ohne jegliche Berücksichtigung in der Auswertung (Nichtberücksichtigungsanteil) größer als 30 % ist. In der Literatur werden zum Teil bereits Nichtberücksichtigungsanteile größer als 20 Prozentpunkte als nicht mehr aussagekräftig betrachtet [29].

Ausnahmen von dieser Regel können zum Beispiel dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war.

Die Ergebnisse wurden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Nichtberücksichtigungsanteile zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.4.2 Meta-Analysen

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren, wurden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung wurden primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Meta-Analysen erfolgten in der Regel auf Basis von Modellen mit zufälligen Effekten [30]. In begründeten Ausnahmefällen wurden Modelle mit festen Effekten eingesetzt. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorlagen, wurden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

Für kontinuierliche Variablen wurde die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g , als Effektmaß eingesetzt. Bei binären Variablen wurden Meta-Analysen primär anhand des Odds Ratio durchgeführt. In begründeten Ausnahmefällen kommen auch andere Effektmaße zum Einsatz. Bei kategorialen Variablen wurde ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und von den verfügbaren Daten verwendet [31].

Die Effektschätzer und Konfidenzintervalle aus den Studien wurden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgte die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität [32]. War die Heterogenität der Studienergebnisse nicht bedeutsam ($p \geq 0,2$ für Heterogenitätstest), wurde der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität wurden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt. Außerdem wurde untersucht, welche Faktoren diese Heterogenität möglicherweise erklären konnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.4.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.4).

4.4.3 Sensitivitätsanalyse

Zur Einschätzung der Robustheit der Ergebnisse waren Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeiträume oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien *hoch* und *niedrig* sollte für Sensitivitätsanalysen verwendet werden. Sensitivitätsanalysen wurden auch durchgeführt, falls für einen Endpunkt verschiedene Instrumente in den Studien eingesetzt wurden.

Das Ergebnis der Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufter Effekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen Nutzen attestiert wird.

4.4.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.4.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen bestand hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren wurden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgten eigene Analysen in Form von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Es wurden folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einbezogen:

- Geschlecht,
- Alter,
- Erkrankung (Schweregrad, primäre oder sekundäre Arthrose),
- Art des therapeutisch-arthroskopischen Eingriffs (Debridement, Lavage),
- Art der Behandlung in der Kontrollgruppe.

Ergaben sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren, wurden diese ebenfalls begründet einbezogen.

Bei Identifizierung möglicher Effektmodifikatoren erfolgte gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen.

4.5 Änderungen der Methodik

Berichtsplan im Vergleich zum vorläufigen Berichtsplan

- Der patientenrelevante Endpunkt „Notwendigkeit eines operativen Gelenkersatzes“ wurde im Abschnitt 4.1.3 gestrichen.

Vorbericht im Vergleich zum Berichtsplan

- In Abschnitt 4.1.3 wurde der Endpunkt „Symptomatik der Gonarthrose“ als Oberbegriff eingeführt und die patientenrelevanten Endpunkte Schmerz und körperliche Funktion diesem zugeordnet. Durch die Einführung dieses neuen Endpunkts konnte auch die globale Bewertung der Symptomatik diesem Endpunkt zugeordnet und dargestellt werden.
- In Abschnitt 4.4.1 wurde ergänzend beschrieben, dass in manchen Studien derselbe Endpunkt unter Verwendung mehrerer Instrumente untersucht wurde. Für diese Fälle wurde ein Instrument für die primär im Ergebnisteil dargestellte Meta-Analyse ausgewählt. Die zusätzlich erhobenen Instrumente für diesen Endpunkt wurden in weiteren Meta-Analysen im Rahmen einer Sensitivitätsanalyse betrachtet.

5 Ergebnisse

5.1 Ergebnisse der Informationsbeschaffung

5.1.1 Bibliografische Literaturrecherche

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche nach Studien in den bibliografischen Datenbanken und des Literaturscreenings gemäß den Kriterien zum Studieneinschluss.

Es wurden insgesamt 1113 Treffer gefunden, nach Ausschluss von 286 Duplikaten ergab sich eine Gesamtzahl von 827 zu screenenden Treffern.

718 Treffer wurden von beiden Reviewern nach Konsentierung zunächst diskrepanter Einschätzungen übereinstimmend im Rahmen des Titel- und Abstractscreenings als nicht relevant ausgeschlossen. Aus der bibliografischen Literaturrecherche verblieben damit 109 potenziell relevante Treffer, die im Volltext gesichtet wurden. Hiervon wurden 85 aufgrund fehlender Relevanz sicher ausgeschlossen. Die Zitate der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B. Bei weiteren 15 Treffern handelte es sich um relevante systematische und nicht systematische Übersichten, welche im Hinblick auf relevante Studien gescreent wurden (siehe Abschnitt 5.1.2.1).

Die verbliebenen 9 Publikationen zu 9 Studien erfüllten nach übereinstimmender Einschätzung beider Reviewer die für diesen Vorbericht definierten Kriterien zum Studieneinschluss.

Im Rahmen der kontinuierlichen Beobachtung der Literatur zum Thema wurde eine zusätzlich relevante Studie identifiziert, die aufgrund der zeitlichen Beschränkung nicht über andere Rechenschritte identifiziert wurde. Eine systematische Aktualisierung der Informationsbeschaffung erfolgt für die Erstellung des Abschlussberichts.

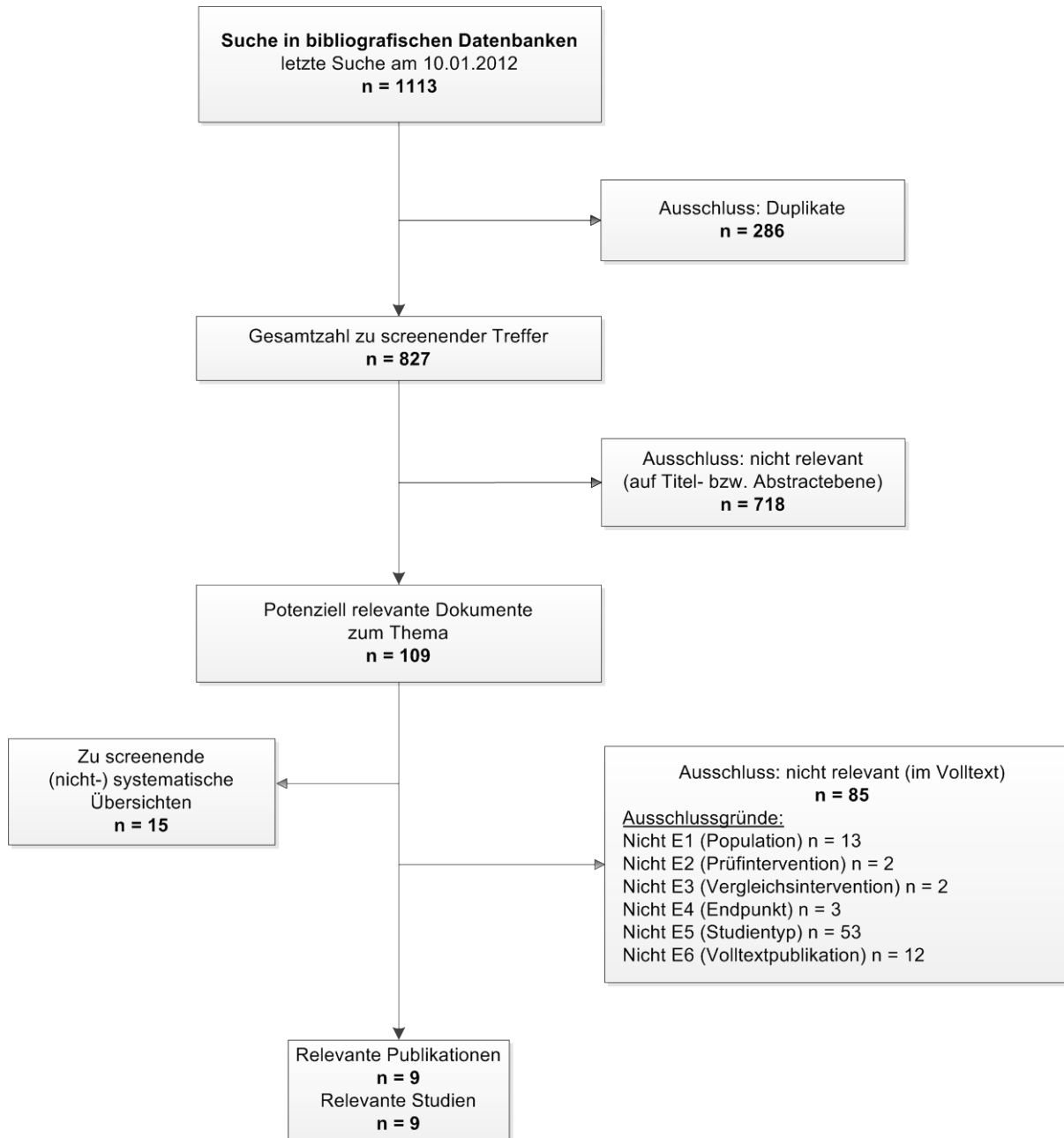


Abbildung 1: Ergebnis der bibliografischen Literaturrecherche und des Literaturscreenings

5.1.2 Weitere Suchquellen zur Identifikation von zusätzlichen publizierten und nicht publizierten Studien bzw. Informationen zu relevanten Studien

5.1.2.1 Systematische und nicht systematische Übersichten

Im Rahmen der bibliografischen Literaturrecherche wurden 15 relevante systematische und nicht systematische Übersichten identifiziert. Deren Sichtung ergab keine weiteren relevanten Publikationen, die nicht über andere Rechenschritte identifiziert wurden. Die entsprechenden Zitate finden sich in Anhang D.

5.1.2.2 Öffentlich zugängliche Studienregister

Durch die Suche nach weiteren publizierten und nicht publizierten Studien in den in Abschnitt 4.2.2.2 genannten öffentlich zugänglichen Studienregistern wurden 2 relevante Einträge identifiziert.

Tabelle 3: In Studienregistern identifizierte relevante Studien

Studienregister ID	Studie	Studienregister	Ergebnisbericht in Studienregister vorhanden
NCT00158431	Kirkley 2008	Clinicaltrials.gov [33]	Nein
NCT00562822	Kirkley 2008	Clinicaltrials.gov [34]	Nein
ID: Identifizierungsnummer; NCT: National Clinical Trial			

Diese beiden Einträge bezogen sich beide auf die Studie Kirkley 2008, welche bereits durch andere Rechenschritte identifiziert wurde.

Es wurden weiter 2 Registereinträge unklarer Relevanz identifiziert (siehe Tabelle 4). In dem Registereintrag zu Mohsen fanden sich Angaben dazu, dass die Studie vorzeitig gestoppt wurde, da die Finanzierung nicht sichergestellt war. Der Registereintrag zu Giffin wurde nicht aktualisiert, daher erfolgte am 03.09.2012 eine Anfrage nach weiterführenden Informationen, die bis zum Zeitpunkt der Erstellung des vorliegenden Vorberichts unbeantwortet geblieben ist.

Tabelle 4: In Studienregistern identifizierte Studien unklarer Relevanz

Studienregister ID	Studie	Studienregister	Status	Ergebnisbericht in Studienregister vorhanden
ISRCTN82192623	Mohsen	ISRCTN [35]	Gestoppt aufgrund fehlender Finanzierung	Nein
NCT00284622	Giffin	Clinicaltrials.gov [36]	Nicht eindeutig, Eintrag nicht mehr aktualisiert	Nein
ID: Identifizierungsnummer; ISRCTN: International Standard Randomised Controlled Trial Number ; NCT: National Clinical Trial;				

5.1.2.3 Durch den G-BA übermittelte Unterlagen

Insgesamt wurden 169 Referenzen vom G-BA mit Auftragserteilung an das IQWiG weitergeleitet [37].

131 Dokumente wurden von beiden Reviewern nach Konsentierung zunächst diskrepanter Einschätzungen übereinstimmend im Rahmen des Titel- und Abstractscreenings als nicht relevant ausgeschlossen.

28 weitere Dokumente wurden von beiden Reviewern im Rahmen der Volltextsichtung nach Konsentierung zunächst diskrepanter Einschätzungen als nicht relevant ausgeschlossen. Die Zitate der ausgeschlossenen Dokumente finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B.

Bei weiteren 4 Dokumenten handelte es sich um relevante systematische und nicht systematische Übersichten. Sofern diese nicht bereits im Rahmen der bibliografischen Literaturrecherche identifiziert worden waren, wurden diese Dokumente im Hinblick auf relevante Studien gescreent (siehe Abschnitt 5.1.2.1).

6 Dokumente (6 Studien) erfüllten nach übereinstimmender Einschätzung beider Reviewer die für diesen Vorbericht definierten Kriterien für den Studieneinschluss. Alle Studien waren auch über die bibliografische Literaturrecherche (siehe Abschnitt 5.1.1) identifiziert worden.

5.1.2.4 Zusätzliche Informationen zu relevanten Studien aus Autorenanfragen

Für die vorliegende Bewertung wurden für 2 Studien Autorenanfragen versendet. Eine Übersicht zu den Autorenanfragen befindet sich in Tabelle 5. Die Informationen aus den eingegangenen Antworten sind in die Studienbewertung eingeflossen (siehe Abschnitt 5.2.2).

Tabelle 5: Übersicht zu Autorenanfragen

Studie	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
Kalunian 2000	Frage 1: Die Randomisierung erfolgte laut Publikation mithilfe eines „random number generator“, wie wurde die verdeckte Gruppenzuteilung sichergestellt? Frage 2: Basiert die Datenauswertung nach 1 Jahr auf allen 90 Patienten?	ja	Antwort 1: Bei Einschluss der Patienten wurde das Koordinierungszentrum vom Arzt angerufen, um die entsprechende Gruppenzuteilung zu erfragen. Antwort 2: Alle Patienten wurden zum Zeitpunkt 1 Jahr nachuntersucht.
Moseley 2002	Waren die Briefumschläge, die die Gruppenzuteilungen enthielten, opak?	ja	Die Umschläge waren opak.

5.1.2.5 Informationen aus der Anhörung

Im Rahmen der Anhörung zum vorläufigen Berichtsplan wurden keine zusätzlichen relevanten Studien genannt.

5.1.3 Resultierender Studienpool

Die nachfolgende Abbildung 2 fasst die Ergebnisse der durchsuchten Quellen zusammen. Für jede Quelle wird die Anzahl der dort als relevant identifizierten Studien dargestellt, wie sie auch in den vorangegangenen Abschnitten beschrieben wurden. Zudem wird dargestellt, wie groß der daraus resultierende Studienpool für die Nutzenbewertung ist.

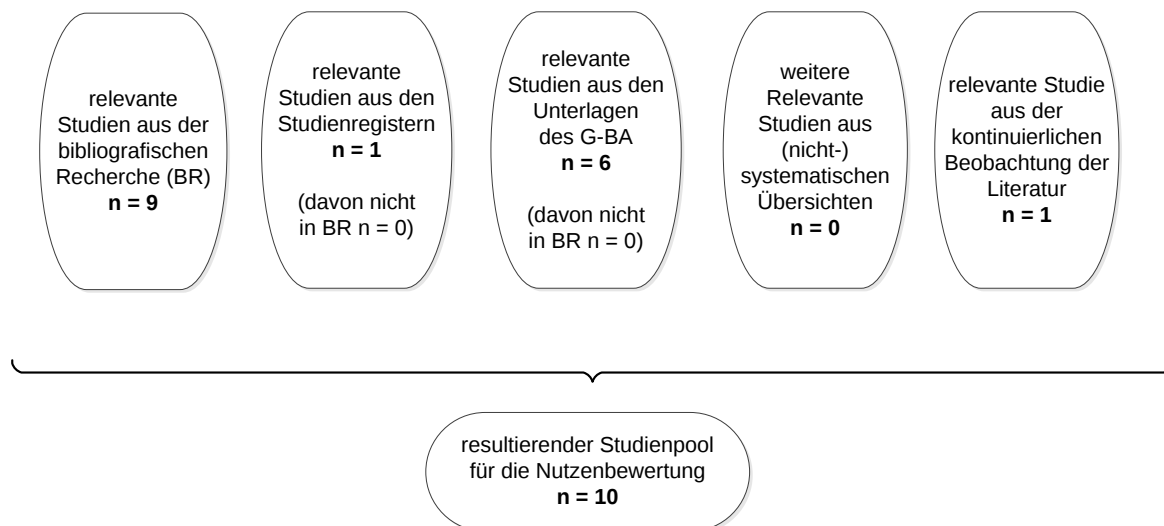


Abbildung 2: Übersicht über die relevanten Studien aus den verschiedenen Quellen der Informationsbeschaffung und daraus resultierender Studienpool

Durch die verschiedenen Suchschritte konnten insgesamt 10 relevante Studien (11 Publikationen) identifiziert werden (siehe auch Tabelle 6).

Tabelle 6: Studienpool der Nutzenbewertung

Studie	Vollpublikation (in öffentlich zugänglichen Fachzeitschriften)
Arden 2008	[38]
Chang 1993	[39]
Forster 2003	[40]
Kalunian 2000	[41]
Kang 2005	[42]
Katz 2013	[43-45]
Kirkley 2008	[46]
Merchan 1993	[47]
Moseley 1996	[48]
Moseley 2002	[49]

5.2 Charakteristika der in die Bewertung eingeschlossenen Studien

Die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement wurde insgesamt in 10 Studien untersucht. In 5 dieser Studien wurde als Vergleich keine aktive Vergleichsintervention eingesetzt, in 5 Studien eine aktive Vergleichsintervention (siehe Tabelle 7). Die Studiencharakteristika werden für alle Studien gemeinsam dargestellt, bei der Ergebnisdarstellung werden die Studien gemäß Vergleichsintervention (keine aktive versus aktive Vergleichsintervention) dargestellt.

Tabelle 7: Übersicht der in den Studien beschriebenen Vergleiche

Studie	Prüfintervention Arthroskopie mit ...	Vergleichskategorie	
		keine aktive Vergleichsintervention	aktive Vergleichsintervention
Arden 2008	Lavage		Glukokortikoide ^a
Chang 1993	Debridement ^b		Lavage (ohne Arthroskopie)
Forster 2003	Lavage		Hyaluronsäure ^a
Kalunian 2000	Lavage	diagnostische Arthroskopie	
Kang 2005	Debridement ^b		Hyaluronsäure ^a
Katz 2013	Debridement ^c	keine zusätzliche Behandlung	
Kirkley 2008	Debridement ^b	keine zusätzliche Behandlung	
Merchan 1993	Debridement ^c		oral NSAID
Moseley 1996	Lavage	Scheinarthroskopie	
	Debridement ^b		
Moseley 2002	Lavage	Scheinarthroskopie	
	Debridement ^b		
a: intraartikulär b: Debridement schließt Lavage mit ein c: In der Publikation wird eine Lavage nicht explizit erwähnt, es wird jedoch davon ausgegangen, dass eine solche durchgeführt wurde. NSAID: non-steroidal anti-inflammatory drugs			

5.2.1 Studiendesign und Studienpopulationen

In Tabelle 8 und Tabelle 9 sind die Angaben zu den eingeschlossenen Studien und zu den in den Studien eingesetzten Interventionen dargestellt.

Unter den Studien befand sich eine Pilotstudie (Moseley 1996), die mit nur wenigen Patienten die Durchführbarkeit einer Studie in dem entsprechenden Zentrum untersuchen sollte. Die Studie Moseley 2002 basiert auf dieser Vorarbeit, da aber völlig unabhängige Patientenkollektive untersucht wurden, wurden beide Studien separat betrachtet.

Von allen 10 randomisierten Studien waren 2 Studien (Moseley 1996 und Moseley 2002) doppelblind (Patient, Endpunkterheber), 4 Studien einfach verblindet (Endpunkterheber) und die verbleibenden 4 Studien wurden nicht verblindet durchgeführt, bzw. bei der Studie Kang 2005 lagen keine Angaben zur Verblindung vor. Beide doppelblinden Studien, 2 Studien, die einfach verblindet waren (Kalunian 2000 und Kirkley 2008) und die unverblindete Studie Katz 2013, verglichen die therapeutische Arthroskopie mit keiner aktiven Vergleichsintervention. Unter den Studien, die die therapeutische Arthroskopie mit einer aktiven Vergleichsintervention verglichen, waren 3 nicht verblindete Studien und 2 einfach verblindete Studien.

Die Studiendauer betrug zwischen 6 und 36 Monaten. Nur in 2 Studien (Arden 2008 und Moseley 1996) betrug die Studiendauer 6 Monate, in 5 Studien 12 Monate (Chang 1993, Forster 2003, Kalunian 2000, Kang 2005 und Katz 2013), in 2 Studien (Kirkley 2008 und Moseley 2002) 24 Monate und in einer Studie (Merchan 1993) 36 Monate, wobei die mittlere Nachbeobachtungsdauer dieser Studie zwischen 23 und 25 Monaten lag.

Jeweils 3 Studien wurden monozentrisch (Kirkley 2008, Moseley 1996 und Moseley 2002) oder in 2 Zentren (Arden 2008, Chang 1993 und Kang 2005) durchgeführt. In jeweils einer Studie (Kalunian 2000) waren 4 bzw. 7 (Katz 2013) Zentren beteiligt, zu den verbleibenden 2 Studien fanden sich hierzu keine Angaben in der Publikation.

8 Studien zeigten einen 2-armigen Vergleich, 2 Studien (Moseley 1996 und Moseley 2002) einen 3-armigen Vergleich. In den beiden Studien von Moseley wurde sowohl die therapeutische Arthroskopie nur mit Lavage als auch die therapeutische Arthroskopie mit Lavage und Debridement untersucht und mit der Vergleichsintervention – einer Scheinarthroskopie – verglichen. Die übrigen 8 Studien untersuchten als Prüfintervention immer eine dieser Varianten – also entweder therapeutische Arthroskopie nur mit Lavage oder mit zusätzlichem Debridement – und verglichen diese mit einer Kontrollgruppe. Es wurden entweder keine aktiven Vergleichsinterventionen (diagnostische Arthroskopie, keine zusätzliche Behandlung oder Scheinarthroskopie) oder aktive Vergleichsinterventionen (intraartikuläre Injektion von Glukokortikoiden oder Hyaluronsäure, Lavage ohne Arthroskopie oder NSAIDs per os) eingesetzt. Neben den in Tabelle 9 detailliert aufgeführten Prüf- und Vergleichsinterventionen wurde in einigen Studien eine gewisse Basistherapie in beiden Studienarmen beschrieben, die z. B. aus unterstützender Physiotherapie oder

ergänzender medikamentöser Behandlung mit Analgetika bestehen konnte. Die Informationen zu einer Basistherapie wurden nur in einigen Studien transparent dargestellt (Chang 1993, Katz 2013 und Kirkley 2008), in den anderen Studien nur erwähnt oder sind anzunehmen. Daher wurde die Basistherapie im vorliegenden Bericht nicht weiter zu den Prüf- und Vergleichsinterventionen aufgeführt.

In 2 Studien (Forster 2003 und Katz 2013) wechselte ein nicht unerheblicher Anteil der randomisierten Patienten von der Vergleichsintervention zur therapeutischen Arthroskopie. In der Studie Forster 2003 betrug dieser Anteil 26 % der Patienten, in der Studie Katz 2013 33 % der Patienten.

Tabelle 8: Eingeschlossene Studien – Übersicht

Studie	Studiendesign / Interventionen	Studien-dauer Monate	Randomisierte Patienten	Zentren (Ort); Zeitraum	Relevante Zielkriterien
Arden 2008	RCT, einfach blind (Endpunkterheber), parallel	6	Gesamtzahl: 150	2 (UK); n. g.	primär: Schmerz sekundär: globale Bewertung der Symptomatik der Gonarthrose, Schmerz, körperliche Funktion, gesundheitsbezogene Lebensqualität, unerwünschte Therapiewirkungen
	AS + Lavage		71		
	intraartikulär Glukokortikoide		79		
Chang 1993	RCT, einfach blind (Endpunkterheber), parallel	12	Gesamtzahl: 34	2 (USA); n. g.	globale Bewertung der Symptomatik der Gonarthrose, Schmerz, körperliche Funktion, gesundheitsbezogene Lebensqualität
	AS + Lavage + Debridement		19		
	Lavage ohne AS		15		
Forster 2003	RCT, nicht verblindet, parallel	12	Gesamtzahl: 38	n. g. (UK); n. g.	globale Bewertung der Symptomatik der Gonarthrose, Schmerz, körperliche Funktion
	AS + Lavage		19		
	intraartikulär Hyaluronsäure		19		
Kalunian 2000	RCT, einfach blind, parallel	12	Gesamtzahl: 90	4 (USA); n. g.	primär: globale Bewertung der Symptomatik der Gonarthrose sekundär: Schmerz, körperliche Funktion
	AS + Lavage		41		
	diagnostische AS		49		
Kang 2005	RCT (Verblindung: n. g.), parallel	12	Gesamtzahl: 69	2 (China); Rekrutierung 01 / 2001 bis 03 / 2003	globale Bewertung der Symptomatik der Gonarthrose
	AS + Lavage + Debridement		32		
	intraartikulär Hyaluronsäure		37		
Katz 2013	RCT, nicht verblindet, parallel	12	Gesamtzahl 351	7 (USA); Rekrutierung 06 / 2008 bis 08 / 2011	primär: körperliche Funktion sekundär: Schmerz, körperliche Funktion, unerwünschte Therapiewirkungen
	AS + Lavage ^a + Debridement		174		
	keine zusätzliche Behandlung		177		

(Fortsetzung)

Tabelle 8: Eingeschlossene Studien – Übersicht (Fortsetzung)

Studie	Studiendesign / Interventionen	Studien-dauer Monate	Randomisierte Patienten	Zentren (Ort); Zeitraum	Relevante Zielkriterien
Kirkley 2008	RCT, einfach verblindet (Endpunkterheber), parallel	24	Gesamtzahl: 188	1 (Kanada); 01 / 1999 bis 08 / 2007	primär: globale Bewertung der Symptomatik der Gonarthrose sekundär: Schmerz, körperliche Funktion, gesundheitsbezogene Lebensqualität, unerwünschte Therapiewirkungen
	AS + Lavage + Debridement		94		
	keine zusätzliche Behandlung		94		
Merchan 1993	RCT, nicht verblindet, parallel	36	Gesamtzahl: 80	n. g. (Spanien); Rekrutierung 01 / 1988 bis 12 / 1990	primär: globale Bewertung der Symptomatik der Gonarthrose, unerwünschte Therapiewirkungen
	AS + Lavage ^a + Debridement	(mittleres FU: 25)	40		
	oral NSAID	(mittleres FU: 23)	40		
Moseley 1996	RCT, Pilotstudie, doppelblind, parallel	6	Gesamtzahl: 10	1 (USA); 1992	Schmerz, körperliche Funktion, unerwünschte Therapiewirkungen
	AS + Lavage + Debridement		2		
	AS + Lavage		3		
	Schein-AS		5		
Moseley 2002	RCT, doppelblind, parallel	24	Gesamtzahl: 180	1 (USA); Rekrutierung 10 / 1995 bis 09 / 1998	primär: Schmerz sekundär: Schmerz, körperliche Funktion
	AS + Lavage + Debridement		59		
	AS + Lavage		61		
	Schein-AS		60		
a: In der Publikation wird eine Lavage nicht explizit erwähnt, es wird jedoch davon ausgegangen, dass eine solche durchgeführt wurde. Abkürzungen: AS: Arthroskopie; FU: Follow-up; n. g.: nicht genannt					

Tabelle 9: Charakterisierung der Interventionen in den eingeschlossenen Studien

Studie	Prüfintervention	Vergleichsintervention
Arden 2008	Arthroskopie + Lavage Lokalanästhesie Arthroskopie mit kleinkalibrigen Instrumenten (3,2 mm Durchmesser) Spülung mit bis zu 1 Liter 0,9 % Kochsalzlösung	Intraartikuläre Injektion von Glukokortikoiden Injektion von 40 mg Triamcinolon in die Gelenkhöhle
Chang 1993	Arthroskopie + Lavage + Debridement Allgemeinnarkose kontinuierliche Spülung mit Kochsalzlösung Reinigung (Debridement) von gerissenem Meniskus und Entnahme von Meniskus- oder Kreuzband-Fragmenten Entnahme von proliferierender Synovia Extraktion von losen Fragmenten des Gelenkknorpels	Lavage (ohne Arthroskopie) Lokalanästhesie Spülung mit einem Gesamtvolumen von 1 Liter Kochsalzlösung Portionen mit einem Volumen von 40 bis 120 ml wurden in das Knie injiziert und wieder aus dem Knie aspiriert.
Forster 2003	Arthroskopie + Lavage allgemeine oder spinale Anästhesie Spülung mit mindestens 2 Liter 0,9 % Kochsalzlösung Debridement der Gelenkflächen oder Menisci wurde nach Entscheidung des Chirurgen vorgenommen große Knorpellappen oder Meniskusabrisse wurden entfernt, stabile degenerierte Menisci wurden intakt gelassen.	Intraartikuläre Injektion von Hyaluronsäure Jeder Patient erhielt im Abstand von einer Woche 5 Arthrozentesen und intraartikuläre Injektionen von jeweils 20 mg Hyaluronsäure.
Kalunian 2000	Arthroskopie + Lavage Lokalanästhesie Arthroskopie mit kleinkalibrigen Instrumenten (2,7 bzw. 1,7 mm Durchmesser je nach Zentrum) Spülung des Gelenks mit 3 Liter Kochsalzlösung	Diagnostische Arthroskopie Lokalanästhesie Arthroskopie mit kleinkalibrigen Instrumenten (2,7 bzw. 1,7 mm Durchmesser je nach Zentrum) Spülung des Gelenks mit 0,25 Liter Kochsalzlösung

(Fortsetzung)

Tabelle 9: Charakterisierung der Interventionen in den eingeschlossenen Studien (Fortsetzung)

Studie	Prüfintervention	Vergleichsintervention
Kang 2005	Arthroskopie + Lavage + Debridement Lokalanästhesie oder spinale Anästhesie Spülung des Gelenks mit 3 Liter Kochsalzlösung Falls indiziert: Glättung von rauem Gelenkknorpel, Abtragen von Synovia Entnahme loser Gewebeteile, Abtrennen von Osteophyten Abtrennen degenerierter Meniskusfragmente, Lösen von peripatellären Verwachsungen	Intraartikuläre Injektion von Hyaluronsäure 5-mal Arthrozentesen und intraartikuläre Injektion von 2 ml Hyaluronsäure jeweils im Abstand von einer Woche
Katz 2013	Arthroskopie + Lavage + Debridement Regional- oder Allgemeinanästhesie Keine Angaben zur Lavage. Es wird jedoch davon ausgegangen, dass eine solche durchgeführt wurde Partielle Meniskektomie durch Abtrennen beschädigter Meniskusanteile bis auf einen stabilen Rand. Entfernung loser Knochen- und Knorpelfragmente Kein Eindringen in den subchondralen Knochen	Keine zusätzliche Behandlung
Kirkley 2008	Arthroskopie + Lavage + Debridement Allgemeinnarkose Spülung des Gelenks mit mindestens 1 Liter Kochsalzlösung Synovektomie oder Debridement oder Exzision gerissener degenerativer Meniskusfragmente oder Exzision von Gelenkknorpeln oder Exzision von Knorpellappen oder Knochenanbauten, die eine vollständige Streckung einschränkten keine Abrasio des Knorpels oder Mikrofrakturierung	Keine zusätzliche Behandlung
Merchan 1993	Arthroskopie + Lavage + Debridement Keine Angabe zur Narkose Keine Angaben zur Lavage. Es wird jedoch davon ausgegangen, dass eine solche durchgeführt wurde Glättung der Gelenkhaut, Entnahme von degenerativen Meniskusanteilen, Osteophyten und losen Gewebeteilen, Teilweise Glättung von rauem Gelenkknorpel	Oral NSAID Nichtsteroidale entzündungshemmende Medikamente

(Fortsetzung)

Tabelle 9: Charakterisierung der Interventionen in den eingeschlossenen Studien (Fortsetzung)

Studie	Prüfintervention	Vergleichsintervention
Moseley 1996	Prüfgruppe 1: Arthroskopie + Lavage + Debridement Allgemeinnarkose 3 Stichinzisionen, Einführung eines Arthroskops, einer Spülkanüle und chirurgischer Instrumente, Spülung mit mindestens 10 Liter Kochsalzlösung, Glättung von rauem Gelenkknorpel, Entnahme loser Gewebeteile, Abtrennen degenerierter Meniskusfragmente Prüfgruppe 2: Arthroskopie + Lavage Allgemeinnarkose 3 Stichinzisionen, Einführung eines Arthroskops und einer Spülkanüle, Spülung mit mindestens 10 Liter Kochsalzlösung	Scheinarthroskopie Sedo-/Neuroleptanalgesie (intravenöse Gabe eines Benzodiazepins, von Droperidol oder beiden Substanzen, jeweils in Kombination mit Fentanyl oder Sufentanil), 3 Stichinzisionen Kein Einführen von Instrumenten
Moseley 2002	Prüfgruppe 1: Arthroskopie + Lavage + Debridement Allgemeinnarkose Spülung mit mindestens 10 Liter Kochsalzlösung Glättung von rauem Gelenkknorpel Entnahme loser Gewebeteile Abtrennen degenerierter Meniskusfragmente Prüfgruppe 2: Arthroskopie + Lavage Allgemeinnarkose Spülung mit mindestens 10 Liter Kochsalzlösung Abtrennen instabiler Teile nach Meniskusriss	Scheinarthroskopie Intravenöse kurz wirkende Beruhigungsmittel zusammen mit einem Opioid 3 Stichinzisionen Kein Einführen von Instrumenten
NSAID: non-steroidal anti-inflammatory drugs		

In Tabelle 10 und Tabelle 11 finden sich Angaben zu den Ein- / Ausschlusskriterien und zu Charakterisierungen der Studienpopulation.

In allen Studien wurde eine symptomatische Gonarthrose und / oder eine radiografische Veränderung als Einschlusskriterium definiert. Bei der radiografischen Veränderung wurde in 5 Studien die Kellgren- und Lawrence-Klassifikation herangezogen. In 2 Studien (Chang 1993, Kang 2005) wurden Patienten mit Grad 1 bis 3 eingeschlossen, in der Studie Kalunian 2000 Patienten mit Grad 0 bis 2, in der Studie Katz 2013 Patienten mit Grad 0 bis 3 und in der Studie Kirkley 2008 Patienten mit Grad 2 bis 4 der Kellgren- und Lawrence-Klassifikation.

In 8 Studien wurde das Alter als Einschlusskriterium verwendet, allerdings waren die Kriterien nicht einheitlich. 5 Studien schlossen Patienten ab jeweils 18 Jahren (Kirkley 2008), ab 20 Jahren (Chang 1993), ab 40 Jahren (Kalunian 2000), ab 45 Jahren (Katz 2013) oder ab 50 Jahren (Merchan 1993) ein. Die Studie Arden 2008 gab einen Alterszeitraum von 40 bis 90 Jahren an, die Studien Moseley 1996 und Moseley 2002 schlossen Patienten unter 70 bzw. unter 75 Jahren ein.

Unterschiede bestanden auch bezüglich der Spezifikation von Ausschlussgründen. So wurden in der Studie Moseley 1996 nur Krankheiten, für die eine allgemeine Narkose ein erhöhtes Risiko darstellt, als Ausschlussgrund definiert. In den Studien Arden 2008, Chang 1993, Forster 2003, Merchan 1993 und Moseley 2002 wurden vor allem frühere Behandlungen am Knie, mechanische Symptome, Fehlstellungen, lang anhaltende Schmerzen oder radiografisch festgestellte hohe Schweregrade der Gonarthrose als Ausschlusskriterien angegeben. In den Studien Kalunian 2000, Katz 2013 und Kirkley 2008 wurden zudem noch weitere Ausschlusskriterien genannt.

Insgesamt wurden 1190 Patienten in allen Studien eingeschlossen (Median: 85), die Spannweite je Studie reichte von 10 Patienten (Moseley 1996) bis 351 Patienten (Katz 2013).

Das mittlere Alter in den Studien lag zwischen 46 (Moseley 1996) und 66 Jahren (Arden 2008), im Median bei 59 Jahren. Die Studien unterschieden sich bezüglich des Anteils Männer und Frauen in der Studienpopulation. Die Patientenkollektive der Studien Moseley 1996 und Moseley 2002 bestanden (fast) nur aus männlichen Patienten (siehe Tabelle 11). In den Studien Chang 1993 und Merchan 1993 lag der Anteil Männer an der Studienpopulation bei ca. 30 %, in den Studien Kirkley 2008 und Arden 2008 lag dieser zwischen 30 % und 40 %. In den Studien Kalunian 2000, Kang 2005 und Katz 2013 war das Verhältnis von Männern und Frauen beinahe ausgeglichen mit einem Männeranteil von gut 45 %. Insgesamt lag der Männeranteil im Median bei 43 %.

Die Dauer der Symptome vor der Therapie wurde nicht in allen Studien angegeben, in den 4 Studien, die dazu Angaben machten, variierte diese stark. So lag die Symptombdauer in der Studie Kalunian 2000 mit ca. 32 Monaten am unteren Ende des Spektrums und in der Studie Kang 2005 mit ca. 70 Monaten am oberen Ende des Spektrums, allerdings waren die Standardabweichungen oder angegebenen Spannweiten so groß bzw. weit, dass eine eindeutige Abgrenzung zwischen den Studien nicht möglich war.

Tabelle 10: Ein- / Ausschlusskriterien für Patienten in den Studien

Studie (Vergleich)	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Arden 2008 (Glukokortikoide)	- Alter zwischen 40 und 90 Jahren - klinische Diagnose einer Gonarthrose und Knieschmerzen an den meisten Tagen des vorangegangenen Monats - radiografisch nachweisbare Veränderungen, die mit einer Gonarthrose vereinbar sind	- begleitende Knieentzündung oder Kristallarthropathien - Verletzungen am Knie innerhalb der vorangegangenen 6 Monate - frühere Operationen am Knie - Injektionen in das Kniegelenk innerhalb der vorangegangenen 3 Monate - Symptomatische Coxarthrose
Chang 1993 (Lavage)	- Alter über 20 Jahre - Knieschmerzen, die trotz konservativer medikamentöser Therapie und Rehabilitationsmaßnahmen länger als 3 Monate andauerten und Arbeit, Sport oder die Selbstversorgung in für den Patienten unakzeptabler Weise einschränkten - radiografisch nachweisbare Veränderungen Grad 1, 2 oder 3 nach Kellgren- und Lawrence-Klassifikation - Einverständnis, an Nachbeobachtungen nach 3 und 12 Monaten teilzunehmen	- radiografisch nachweisbare Veränderungen Grad 4 nach der Kellgren- und Lawrence-Klassifikation - Knie-Totalendoprothese - Krankheit, die die Untersuchung der Funktionstüchtigkeit des Knies beeinflussen konnte oder die Arthroskopie verhindern konnte wie zum Beispiel schwere periphere arterielle Verschlusskrankheit mit Claudicatio intermittens oder Herzkrankheit - geplante Knieoperationen innerhalb von 6 Monaten nach Studieneintritt
Forster 2003 (Hyaluronsäure)	- symptomatische Gonarthrose - radiografisch nachweisbare Veränderungen, die mit einer Gonarthrose vereinbar sind - Geeignet für lokale und generelle Anästhesie	- Gelenkinjektionen innerhalb der letzten 6 Monate - frühere arthroskopische Operation - mechanische Symptome - Überempfindlichkeit gegen Hyaluronsäure
Kalunian 2000 (diagnostische Arthroskopie)	- Alter mindestens 40 Jahre - nicht befriedigende Schmerzreduktion mittels einer mindestens 6-wöchigen Vorbehandlung einschließlich isometrischer Übungen, mindestens 2 verschiedene NSAIDs über einen Behandlungszeitraum von mindestens 3 Wochen pro Arzneimittel, sofern der Patient die Physiotherapie durchführen und die Arzneimittel vertragen konnte - normale oder minimal auffällige Befunde in Röntgenaufnahmen, Kellgren- und Lawrence-Klassifikation Grad 0 bis 2 - Erfüllung der Diagnosekriterien gemäß American College of Rheumatology - Einverständnis, an Nachfolgeuntersuchungen teilzunehmen	- Deutlich auffällige Befunde in Röntgenaufnahmen, Kellgren- und Lawrence-Klassifikation Grad 3 bis 4 - Intraartikuläre Glukokortikoid-Injektion innerhalb der letzten 4 Wochen vor Beginn der Studie - Erkrankungen der Wirbelsäule, der Hüften, der Fußgelenke oder der Füße - schwerwiegende andere Krankheiten mit einem Risiko, durch die Studienteilnahme verschlimmert zu werden - Body-Mass-Index größer als 35 kg/m² - Unverträglichkeit gegenüber Narkosemittel - Drogenmissbrauch

(Fortsetzung)

Tabelle 10: Ein- / Ausschlusskriterien für Patienten in den Studien (Fortsetzung)

Studie (Vergleich)	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Kang 2005 (Hyaluron- säure)	▪ milde oder mittelgradige Gonarthrose (Grad 1, 2 oder 3 nach Kellgren- und Lawrence-Klassifikation)	-
Katz 2013 (keine zusätzliche Behandlung)	▪ Alter mindestens 45 Jahre ▪ Symptome eines Meniskusrisses über mindestens 4 Wochen, für die medikamentöse Behandlung, Physiotherapie und / oder Aktivitätsreduktion erforderlich war ▪ nachweisbare Gonarthrose im MRT (Osteophyten oder bis zum Knochen reichender Knorpeldefekt) oder im konventionellen Röntgenbild (Osteophyten oder Gelenkspaltverschmälerung in Stehendaufnahme) ▪ Meniskusriss (bis an die Oberfläche reichender Riss im MRT) ▪ Radiografisch nachweisbare Veränderungen, Kellgren- und Lawrence-Klassifikation Grad 0 bis 3	▪ Chronisch blockiertes Knie ▪ Grad 4 nach Kellgren-Lawrence-Klassifikation ▪ Entzündliche Arthritis oder Chondrokalzinose ▪ Injektion von Hyaluronsäure in den letzten 4 Wochen in das Index-Knie ▪ Bilateraler symptomatischer Meniskusriss ▪ Vorhergehende Operation am Index-Knie ▪ Kontraindikation gegen Operation oder Physikalische Behandlung
Kirkley 2008 (keine zusätzliche Behandlung)	▪ Alter mindestens 18 Jahre ▪ primäre oder sekundäre Gonarthrose ▪ Radiografische Veränderungen Grad 2, 3 oder 4 gemäß der modifizierten Kellgren- und Lawrence-Klassifikation	▪ Grad 4 nach Kellgren- und Lawrence-Klassifikation; Arthrose in 2 Kompartimenten bei Personen in einem Alter von über 60 Jahren ▪ frühere Arthroskopie wegen einer Gonarthrose ▪ intraartikuläre Injektion von Kortikosteroiden innerhalb der letzten 3 Monate ▪ frühere große Knieverletzung ▪ große Meniskusrisse ▪ Arthritis mit Nachweis von erhöhten Entzündungswerten ▪ mehr als 5-Grad-Varus- oder Valgus-Fehlstellung ▪ schwerwiegende neurologische Einschränkungen ▪ schwerwiegende Krankheiten mit einer Lebenserwartung von weniger als 2 Jahren oder hohem intraoperativem Risiko ▪ Schwangerschaft ▪ Risiko, dass die Zusammenarbeit über einen Zeitraum von 2 Jahren nicht eingehalten wird

(Fortsetzung)

Tabelle 10: Ein- / Ausschlusskriterien für Patienten in den Studien (Fortsetzung)

Studie (Vergleich)	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Merchan 1993 (oral NSAID)	▪ Patienten in einem Alter über 50 Jahre ▪ schmerzhafte Gonarthrose ▪ geringgradige radiografische Veränderungen (minimale Gelenkspaltverschmälerung und kleine Osteophyten)	▪ frühere chirurgische Maßnahmen ▪ mehr als 15-Grad-Varus- oder Valgus-Fehlstellung ▪ Arthrose des Femoropatellargelenks ▪ Schmerzen über einen Zeitraum von länger als 6 Monaten ▪ Körpergewicht über 85 kg bei Männern und über 70 kg bei Frauen
Moseley 1996 (Schein-arthroskopie)	▪ Alter unter 70 Jahre ▪ symptomatische Gonarthrose trotz einer mindestens 6-monatigen nicht operativen Behandlung einschließlich nichtsteroidaler antiphlogistischer medikamentöser Behandlung ▪ mindestens mittelmäßige Knieschmerzen mit einem durchschnittlichen Score von mindestens 4 auf einer Skala von 0 bis 10 über den Verlauf einer Woche	▪ Krankheiten, für die eine allgemeine Narkose ein erhöhtes Risiko darstellten
Moseley 2002 (Schein-arthroskopie)	▪ Alter höchstens 75 Jahre ▪ Gonarthrose gemäß Definition American College of Rheumatology ▪ mittelmäßig starke Knieschmerzen Grad 4 oder höher auf einer Skala von 0 bis 10 trotz maximaler medikamentöser Therapie für mindestens 6 Monate	▪ Schweregrad von 9 oder höher auf einer Skala von 0 bis 12 gemäß radiografischen Zeichen bewertet von 0 bis 4 und addiert für 3 Kompartimente ▪ Kniearthroskopie innerhalb der vorangegangenen 2 Jahre ▪ schwerwiegende Fehlstellung ▪ schwerwiegende Krankheiten

Tabelle 11: Charakterisierung der Studienpopulation

Studie Gruppe	N ^a	Alter, Jahre MW (SD)	Anteil Männer (%)	Gewicht, kg MW (SD)	BMI MW (SD)	Radiografischer Schweregrad ^b in %	Dauer der Symptome ^c , Monate, MW (SD)	Abbruch vor Therapie (N)	Abbruch nach Therapie (N)	Therapie- wechsel (N)
Arden 2008										
AS: Lavage	71	64,9 (9,7)	26,8	83,3 (15,2)	n. g.	0-1: 16,9 2: 67,6 3-4: 11,3 n. g. 4,2	M 60 (IQR 24; 120)	0	6	n. g.
Gluko- kortikoide	79	67,7 (9,1)	41,8	87,2 (13,6)	n. g.	0-1: 10,1 2: 68,4 3-4: 20,3 n. g. 1,2	M 54 (IQR 29; 120)	0	8	n. g.
Chang 1993										
AS: Debridement	19	61 (11)	28	n. g.	n. g.	1: 22 2: 28 3: 50	51 (51)	1	4	3
Lavage	15	65 (13)	29	n. g.	n. g.	1: 14 2: 36 3: 50	53 (57)	1	1	2
Forster 2003										
AS: Lavage	19	63 (n. g.)	n. g.	n. g.	n. g.	n. g.	n. g.	2	2	6 ^d
Hyaluronsäure	19	60 (n. g.)	n. g.	n. g.	n. g.	n. g.	n. g.	0	2	1 ^d
Kalunian 2000										
AS: Lavage	41	60,9 (R: 41-88)	46	n. g.	n. g.	0-2: 100	30,0 (R: 2-120)	n. g.	n. g.	n. g.
Diagnostische AS	49	58,3 (R: 40-85)	47	n. g.	n. g.	0-2: 100	34,4 (R: 2-120)	n. g.	n. g.	n. g.

(Fortsetzung)

Tabelle 11: Charakterisierung der Studienpopulation (Fortsetzung)

Studie Gruppe	N ^a	Alter, Jahre MW (SD)	Anteil Männer (%)	Gewicht, kg MW (SD)	BMI MW (SD)	Radiografischer Schweregrad ^b in %		Dauer der Symptome ^c , Monate, MW (SD)	Abbruch vor Therapie (N)	Abbruch nach Therapie (N)	Therapie- wechsel (N)
Kang 2005											
AS: Debridement	32	n. g. (R: 42-65)	47	n. g.	n. g.	0:	0 ^e	66 (R: 6-252)	n. g.	n. g.	n. g.
						1:	9				
						2:	17				
						3:	15				
Hyaluronsäure	37	53 (R: 45-67)	43	n. g.	n. g.	0:	0 ^e	76 (R: 18-312)	n. g.	n. g.	n. g.
						1:	14				
						2:	18				
						3:	16				
Katz 2013 ^f											
AS + Debridement	174	59 (7,9)	44,1	n. g.	30,0 (6,1)	0:	21,1 ^g	n. g.	n. g. ^h	18 ^h	8 ⁱ
						1:	16,1				
						2:	23,0				
						3:	28,0				
keine zusätzliche Behandlung	177	57,8 (6,8)	42,6	n. g.	30,0 (6,1)	0:	21,3 ^g	n. g.	n. g. ^h	13 ^h	51 bzw.59 ^j
						1:	20,7				
						2:	23,1				
						3:	23,1				
Kirkley 2008											
AS: Debridement	94	58,6 (10,2)	41	91,3 (17,3)	31,6 (6,7)	2:	46	47,1 (69,4)	2	4	6
						3:	49				
						4:	5				
keine zusätzliche Therapie	94	60,6 (9,9)	33	84,9 (17,9)	30,2 (6,3)	2:	42	40,1 (72,6)	8	6	0
						3:	53				
						4:	5				

(Fortsetzung)

Tabelle 11: Charakterisierung der Studienpopulation (Fortsetzung)

Studie Gruppe	N ^a	Alter, Jahre MW (SD)	Anteil Männer (%)	Gewicht, kg MW (SD)	BMI MW (SD)	Radiografischer Schweregrad ^b in %		Dauer der Symptome ^c , Monate, MW (SD)	Abbruch vor Therapie (N)	Abbruch nach Therapie (N)	Therapie- wechsel (N)
Merchan 1993											
AS: Debridement	40	57,1 (n. g.)	20 ^k	n. g.	n. g.	n. g.		n. g.	0	5	0
oral NSAID	40	56,9 (n. g.)	34 ^k	n. g.	n. g.	n. g.		n. g.	0	2	0
Moseley 1996 ^l											
AS: Lavage	2	46,4	100 (für alle)	n. g.	n. g.	n. g.		n. g.	0	0	0
AS: Debridement	3	(R: 30-67) (für alle)		n. g.	n. g.	n. g.		n. g.	0	1	0
Schein-AS	5			n. g.	n. g.	n. g.		n. g.	0	0	0
Moseley 2002											
AS: Lavage	59	51,2 (10,5)	88,5	n. g.	n. g.	mild: mod: schw:	27,9 ^m 45,9 26,2	n. g.	n. g.	n. g.	n. g.
AS: Debridement	61	53,6 (12,2)	96,6	n. g.	n. g.	mild: mod: schw:	30,5 ^m 45,8 23,7	n. g.	n. g.	n. g.	n. g.
Schein-AS	60	52,0 (11,1)	93,3	n. g.	n. g.	mild: mod: schw:	28,3 46,7 25,0	n. g.	n. g.	n. g.	n. g.

(Fortsetzung)

Tabelle 11: Charakterisierung der Studienpopulation (Fortsetzung)

a: Anzahl randomisierter Patienten

b: Kellgren- und Lawrence-Klassifikation des radiografischen Schweregrads der Gonarthrose [50]: Grad 0: normal; Grad 1: verdächtige Verschmälerung des Gelenkspaltes und mögliche Anbauten durch irreguläre Knochenformationen (Osteophyten) am Gelenkrand; Grad 2: Osteophyten und mögliche Verschmälerung des Gelenkspaltes; Grad 3: vielfache mittelgroße Osteophyten und sichere Verschmälerung des Gelenkspaltes, geringe Sklerosierung und mögliche Verformung der Knochenkontur; Grad 4: große Osteophyten, ausgeprägte Verschmälerung des Gelenkspaltes, ausgeprägte Sklerosierung und sichere Verformung der Knochenkontur.

c: Dauer der Symptome vor der Therapie

d: 6 Patienten in der Hyaluron-Gruppe hatten nach den geplanten 5 Injektionen auch eine arthroskopische Intervention (n = 5) oder einen totalen Gelenkersatz (n = 1); 1 Patient in der AS-Gruppe hatte vor Ablauf der Studie einen totalen Gelenkersatz.

e: Zahlen beziehen sich auf die Anzahl von Knien und nicht auf % der Patienten; 41 Knie in der Interventions- und 48 Knie in der Kontrollgruppe.

f: Studiencharakteristika beziehen auf 161 bzw. 169 Patienten

g: Bei 11,8 % der Patienten wurde kein Wert erhoben.

h: Unklare Aussage, ob die Patienten vor, während oder nach der Therapie bzw. dem 6 Monatszeitpunkt von der Studie ausgeschlossen wurden.

i: In verschiedenen Tabellen der Publikation werden unterschiedliche Zahlen genannt (8 bzw. 9).

j: Anzahl Patienten mit Therapiewechsel nach 6 bzw. 12 Monaten.

k: Zahlen beziehen sich auf ausgewertete Patienten (AS: N = 35, NSAID: N = 38).

l: Die Gesamtzahl der Teilnehmer zum Zeitpunkt 0 Monate war = 10, zum Zeitpunkt 3 und 6 Monate n = 9, weil ein Patient der Gruppe Lavage wegen einer psychiatrischen Erkrankung ausfiel; die Gruppenzuordnung kann indirekt aus der Tabelle 3 des Artikels entnommen werden.

m: Schweregrad der Gonarthrose gemäß einer Skala von 0 bis 12. Ausschluss der Patienten mit Schweregrad von 9 bis 12. Einteilung der eingeschlossenen Patienten (0 bis 8) in mild, moderat und schwer, ohne Angabe der Zuordnung zu den Skalenwerten.

AS: Arthroskopie; BMI: Body-Mass-Index; FU: Follow-up; IA: intraartikuläre Injektion; IQR: interquartile range; M: Median; mod: moderat; MW: Mittelwert; n. g.: nicht genannt; NSAID: non-steroidal anti-inflammatory drug; R: range; schw: schwer; SD: Standardabweichung

5.2.2 Einschätzung des Verzerrungspotenzials auf Studienebene

Um darzustellen, mit welcher Sicherheit sich ein Effekt oder das Fehlen eines Effektes aus einer Studie ableiten lässt, erfolgte eine Bewertung des Verzerrungspotenzials der in den Studien berichteten Ergebnisse. Die Angaben zur Bewertung des Verzerrungspotenzials auf Endpunktebene werden in den Abschnitten zu den einzelnen Endpunkten präsentiert.

Tabelle 12 beinhaltet die Bewertung des Verzerrungspotenzials der 10 eingeschlossenen Studien. Für 3 Studien (Kalunian 2000, Katz 2013 und Moseley 2002) war das Verzerrungspotenzial auf Studienebene niedrig. Für diese Einschätzung wurden Informationen aus den Autorenanfragen berücksichtigt (siehe Abschnitt 5.1.2.4). Die übrigen 7 Studien hatten ein hohes Verzerrungspotenzial. Bei diesen 7 Studien war die Verdeckung der Gruppenzuteilung unklar, außerdem war die adäquate Erzeugung der Randomisierungssequenz mit Ausnahme bei der Studie Kirkley 2008 bei diesen Studien ebenfalls unklar. Autorenanfragen hätten diese Unklarheiten gegebenenfalls beseitigen können und damit die Einstufung des studienbezogenen Verzerrungspotenzials ändern können. Da allerdings bei 6 dieser 7 Studien die Patienten nicht verblindet wurden und die relevanten Endpunkte in diesem Bericht (Symptomatik der Gonarthrose und die gesundheitsbezogene Lebensqualität) durch Patientenberichtete Instrumente erhoben wurden, lag in diesen Fällen ein hohes Verzerrungspotenzial auf Endpunktebene vor. Eine Autorenanfrage hätte diese Einschätzung nicht ändern können, daher wurde auch für die Einschätzung des Verzerrungspotenzials auf Studienebene auf eine Autorenanfrage verzichtet. Bei der siebten Studie handelte es sich um die Pilotstudie Moseley 1996, wo aufgrund der minimalen Patientenzahl auf eine Autorenanfrage verzichtet wurde.

Tabelle 12: Bewertung des Verzerrungspotenzials auf Studienebene

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisgesteuerte Berichterstattung unwahrscheinlich	Fehlen sonstiger Aspekte, die Verzerrungen verursachen können	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
Arden 2008	unklar	unklar	nein	nein	ja	ja	hoch ^a
Chang 1993	unklar	unklar	nein	nein	ja	ja	hoch ^a
Forster 2003	unklar	unklar	nein	nein	ja	nein ^b	hoch ^a
Kalunian 2000	ja	ja ^b	ja	nein	ja	nein ^b	niedrig
Kang 2005	unklar	unklar	nein	nein	ja	ja	hoch ^a
Katz 2013	ja	ja	nein	nein	ja	ja	niedrig
Kirkley 2008	ja	unklar	nein	nein	ja	ja	hoch ^a
Merchan 1993	unklar	unklar	nein	nein	ja	ja	hoch ^a
Moseley 1996	unklar	unklar	ja	nein / ja ^c	unklar ^d	ja	hoch ^a
Moseley 2002	ja	ja ^e	ja	nein / ja ^c	ja	ja	niedrig
a: unklare Randomisierung b: unklarer Patientenfluss c: Der die Operation durchführende Behandler war nicht verblindet, der postoperative Behandler war verblindet. d: Hinweis, dass nur ein Teil der Ergebnisse berichtet wird e: Beurteilung basiert auf zusätzlichen Informationen aus einer Autorenanfrage.							

5.3 Ergebnisse zu patientenrelevanten Endpunkten

In Tabelle 13 werden Informationen zu der Extraktion von berichtsrelevanten Daten zu den patientenrelevanten Endpunkten für alle eingeschlossenen Studien dargestellt. Hierbei wurde berücksichtigt, ob es überhaupt extrahierte Daten für die entsprechenden Endpunkte in der Studie zum jeweils gleichen Zeitpunkt gab und ob diese in einer Meta-Analyse berücksichtigt wurden.

Die Ergebnisse zu den patientenrelevanten Endpunkten wurden – je nach Verfügbarkeit – für die Zeitpunkte 3, 6, 12, 24 und 36 Monate extrahiert und bewertet. Aufgrund der Datenlage in den Studien, wurden dabei primär die Werte zu den einzelnen Zeitpunkten im Rahmen von Meta-Analysen verwendet. Veränderungen im Vergleich zu Baseline wurden – sofern verfügbar – ebenfalls in den Ergebnistabellen dargestellt.

Tabelle 13: Übersicht zur Extraktion von berichtsrelevanten Daten

Vergleich	berichtsrelevante Daten zum Endpunkt				
Studie	Globale Bewertung der Symptomatik	Schmerz	Körperliche Funktion	Gesundheitsbezogene Lebensqualität	Unerwünschte Therapie-wirkung
Keine aktive Vergleichsintervention					
Kalunian 2000	●	●	●	-	-
Katz 2013	-	●	●	-	○
Kirkley 2008	●	●	●	○	○
Moseley 1996	-	○	○	-	-
Moseley 2002	-	●	●	-	○
Aktive Vergleichsintervention					
Arden 2008	○	○	○	○	○
Chang 1993	○	○	○	○	-
Forster 2003	○	○	○	-	-
Kang 2005	○	-	-	-	-
Merchan 1993	○	-	-	-	○
-: keine berichtsrelevanten Daten extrahiert; ○: für den Bericht extrahierte Daten; ●: für den Bericht extrahierte Daten, die in einer Meta-Analyse berücksichtigt wurden					

5.3.1 Arthroskopie versus keine aktive Vergleichsintervention

In 5 Studien wurde die therapeutische Arthroskopie mit Lavage und gegebenenfalls mit zusätzlichem Debridement mit keiner aktiven Vergleichsintervention verglichen. Unter der Bezeichnung „keine aktive Vergleichsintervention“ wurden 3 verschiedene Vergleichsinterventionen zusammengefasst. In 2 Studien (Moseley 1996 und 2002) wurde eine Scheinarthroskopie, in der Studie Kalunian 2000 eine diagnostische Arthroskopie und in 2 Studien (Katz 2013 und Kirkley 2008) keine zusätzliche Behandlung durchgeführt. Die Studien wurden trotz der unterschiedlichen Vergleichsintervention gemeinsam bewertet, was

auch die gemeinsame Auswertung im Rahmen einer Meta-Analyse bedeuten konnte. Die Studie Moseley 1996 wurde nicht in Meta-Analysen berücksichtigt, da keine Streuungsmaße angegeben waren und diese auch nicht durch interne Angaben oder anhand von Angaben aus anderen Studien sinnvoll geschätzt oder ersetzt werden konnten. Aufgrund der geringen Patientenzahl (n = 10) und einem daraus resultierenden extrem geringen Gewicht der Studie ist davon auszugehen, dass die Nicht-Berücksichtigung der Studie die Ergebnisse der Meta-Analysen nicht nennenswert beeinflusst. Die Ergebnisse dieser Studie (Moseley 1996) wurden in diesem Bericht demnach nur tabellarisch dargestellt.

5.3.1.1 Symptomatik der Gonarthrose

Für den Endpunkt Symptomatik der Gonarthrose konnten aus den eingeschlossenen Studien Daten zum Schmerz, zur körperlichen Funktion und zur globalen Bewertung der Symptomatik extrahiert werden. Diese Daten wurden getrennt voneinander dargestellt.

5.3.1.1.1 Schmerz

Zum Endpunkt Schmerz lagen aus 5 Studien entsprechende Daten vor, die dort verwendeten Instrumente werden in Tabelle 14 dargestellt.

Tabelle 14: Übersicht der verwendeten Instrumente zur Erfassung des Schmerzes beim Vergleich keine aktive Vergleichsintervention

Studien	Verwendete Instrumente
Kalunian 2000	WOMAC-Subskala Schmerz VAS Schmerz
Katz 2013	KOOS Subskala Schmerz
Kirkley 2008	WOMAC-Subskala Schmerz
Moseley 1996	Durchschnittliche Schmerzintensität (Einschätzung Patient)
Moseley 2002	AIMS-2-Subskala Schmerz SF-36-Subskala Schmerz KSPS
AIMS: Arthritis impact measurement scales; KOOS: Knee Injury and Osteoarthritis Outcome Scale; KSPS: knee-specific pain scale; SF-36: Short Form (36) Health Survey; VAS: visuelle Analogskala; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index	

In den Studien Kalunian 2000 und Kirkley 2008 wurde die WOMAC-Subskala zum Schmerz erhoben, bei Kalunian 2000 zusätzlich auch noch der vom Patienten empfundene Schmerz, dargestellt auf einer visuellen Analogskala. Die Studie Katz 2013 verwendete die Schmerz-Subskala des KOOS. In der Studie Moseley 1996 wurde der Endpunkt Schmerz anhand eines Fragebogens erfasst. Die durchschnittlichen Schmerzintensität der vorangegangenen Woche angegeben durch die Patienten wurde als Ergebnis in diesem Bericht dargestellt (siehe auch Abschnitt 5.3.1). In der Studie Moseley 2002 wurde neben der AIMS-2-Subskala zum Schmerz auch die Subskala des SF-36 zum Schmerz und ein selbst entwickeltes Instrument KSPS verwendet.

Verzerrungspotenzial der Ergebnisse zum Schmerz

Die Studien Kalunian 2000 und Moseley 2002 hatten ein niedriges Verzerrungspotenzial auf Endpunktebene. Die Studien Katz 2013 und Kirkley 2008 hatten ein hohes Verzerrungspotenzial, die nicht verblindete Endpunkterhebung sowie die Umsetzung des ITT-Prinzips waren nicht adäquat. Die Studie Moseley 1996 hatte auf Endpunktebene keine weiteren Einschränkungen, die das Verzerrungspotenzial erhöhten, allerdings musste aufgrund der unklaren Randomisierung und der unklaren Verdeckung der Gruppenzuteilung sowie der fehlenden Verblindung (siehe Tabelle 15) auch auf Endpunktebene von einem hohen Verzerrungspotenzial ausgegangen werden.

Ergebnisse zum Schmerz

In Tabelle 16 sind die Ergebnisse aus allen 5 Studien zum Endpunkt Schmerz dargestellt. Abbildung 3 zeigt die Meta-Analysen der Ergebnisse zum Endpunkt Schmerz jeweils für die Zeitpunkte 3, 6, 12 und 24 Monate.

In den Meta-Analysen wurden die Instrumente WOMAC Schmerz (Kalunian 2000; Kirkley 2008), KOOS Schmerz (Katz 2013) und das Instrument AIMS-2 Schmerz (Moseley 2002) verwendet (für die Ergebnisse der zusätzlich erhobenen Instrumente der Studien Kalunian 2000 und Moseley 2002 siehe den Abschnitt zu Sensitivitätsanalysen). In den Meta-Analysen zu den Zeitpunkten 3, 6 und 12 Monate zeigte sich eine bedeutsame Heterogenität ($p < 0,2$), sodass kein gemeinsamer Schätzer berechnet wurde. Die Ergebnisse waren nicht gleichgerichtet. In der Meta-Analyse zu dem Zeitpunkt 24 Monate konnte ein gemeinsamer Schätzer berechnet werden. Für diesen Zeitpunkt lagen allerdings nur die Daten aus 2 Studien (Kirkley 2008 und Moseley 2002) vor. Ein statistisch signifikanter Effekt zugunsten der therapeutischen Arthroskopie oder der Vergleichsintervention konnte nicht festgestellt werden.

In der Studie Moseley 2002 wurde zusätzlich noch das Instrument KSPS eingesetzt. Auch hier ergaben sich zu keinem Zeitpunkt signifikante Unterschiede zwischen den Behandlungsgruppen.

Zusammenfassend zeigte sich kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Tabelle 15: Bewertung des Verzerrungspotenzials auf Endpunktebene: Schmerz, Vergleich keine aktive Vergleichsintervention

Studie	Verzerrungspotenzial Studienebene	Verblindung Endpunkt- erheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungs- potenzial beeinflussen	Verzerrungspotenzial für diesen Endpunkt
Kalunian 2000	niedrig	ja	ja ^a	ja	ja	niedrig
Katz 2013	niedrig	nein	nein ^b	ja	ja	hoch ^c
Kirkley 2008	hoch	nein ^d	nein ^e	ja	ja	hoch ^c
Moseley 1996	hoch	ja	ja	ja	ja	hoch ^f
Moseley 2002	niedrig	ja	ja	ja	ja	niedrig
a: Bewertung basiert auf zusätzlichen Informationen aus einer Autorenanfrage. b: Bis zu 20 % der Daten wurden nicht in der Auswertung berücksichtigt. c: keine Verblindung, keine adäquate ITT-Auswertung. d: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung der Schmerzen erfolgte durch unverblindeten Patienten begleitet durch verblindeten Behandler. e: Eine unterschiedliche Anzahl von Patienten wurde in den Auswertungen zu den verschiedenen Zeitpunkten nicht berücksichtigt. f: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung						

Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention

Studie	Instru- ment (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Kalunian 2000	WOMAC Schmerz (0-20)	0	41	9,4 [8,8; 10,0]		49	8,8 [8,3; 9,3]			
		3	n. g.	6,4 [5,8; 7,0]	n. g.	n. g.	5,7 [5,2; 6,2]	n. g.	n. g.	n. g.
		12	n. g.	5,6 [5,0; 6,2]	4,2 [-0,9; 9,4]	n. g.	6,6 [6,1; 7,1]	2,3 [-0,1; 4,7]	n. g.	n. g. (n. g.); 0,04 ^a
	VAS Schmerz (0-10)	0	41	3,7 [3,4; 4,0]		49	3,6 [3,3; 3,9]			
		3	n. g.	2,9 [2,6; 3,2]	n. g.	n. g.	2,7 [2,4; 3,0]	n. g.	n. g.	n. g.
		12	n. g.	2,3 [1,9; 2,7]	1,47 [-1,2; 4,1]	n. g.	3,3 [3,0; 3,6]	0,12 [0; 0,3]	n. g.	n. g. (n. g.); 0,02 ^a
Katz 2013	KOOS Schmerz (0-100) ^b	0	159	46 (15,5)		167	47,2 (16,4)			
		3	155 ^c	23 ^c [20; 26]		154 ^c	31 ^c [28; 34]			
		6	148 ^c	21,1 [18,3; 23,9]	24,2 [21,3; 27,1]	153 ^c	25,2 [22,4; 28,0]	21,3 [18,4; 24,2]	n. g.	2,9 [-1,2; 7,0]; n. g.
		12	142 ^c	19,1 [16,4; 21,9]	26,8 [23,7; 30,0]	145 ^c	19,3 [16,6; 22,0]	27,3 [24,1; 30,4]	n. g.	-0,4 [-4,8; 4,0]; n. g.
Kirkley 2008	WOMAC Schmerz (0-500)	0	92	239 (105)		86	214 (122)			
		3	90	141 (109)	n. g.	80	172 (124)	n. g.	n. g.	n. g.
		6	90	143 (113)	n. g.	73	155 (118)	n. g.	n. g.	n. g.
		12	80	155 (125)	n. g.	77	147 (116)	n. g.	n. g.	n. g.
		24	88	168 (134)	n. g.	80	185 (132)	n. g.	n. g. (n. g.); 0,14 ^d	n. g.

(Fortsetzung)

Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instru- ment (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Moseley 2002	AIMS-2 Schmerz (0-100) ^e	Prüfintervention Lavage								
		0	61	59,3 (16,7)		59	59,5 (18,5)			
		3	59	53,7 (23,1)	n. g.	56	50,1 (21,3)	n. g.	-3,6 [-11,8; 4,6]; 0,39	n. g.
		6	59	54,8 (21,6)	n. g.	57	50,0 (20,7)	n. g.	-4,8 [-12,5; 3,0]; 0,23	n. g.
		12	57	57,8 (23,5)	n. g.	54	53,6 (22,1)	n. g.	-4,2 [-12,8; 4,4]; 0,34	n. g.
		24	56	56,7 (24,1)	n. g.	55	52,5 (25,1)	n. g.	-4,2 [-13,5; 5,1]; 0,37	n. g.
		Prüfintervention Lavage + Debridement								
		0	58	59,3 (22,2)		59	59,5 (18,5)			
		3	58	49,9 (21,7)	n. g.	56	50,1 (21,3)	n. g.	0,3 [-7,7; 8,2]; 0,95	n. g.
		6	55	52,0 (20,8)	n. g.	57	50,0 (20,7)	n. g.	-2,0 [-9,8; 5,7]; 0,60	n. g.
		12	51	53,3 (25,4)	n. g.	54	53,6 (22,1)	n. g.	0,3 [-8,9; 9,5]; 0,95	n. g.
		24	53	54,0 (23,3)	n. g.	55	52,5 (25,1)	n. g.	-1,5 [-10,8; 7,7]; 0,75	n. g.

(Fortsetzung)

Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instru- ment (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Moseley 2002	SF-36 Schmerz (0-100) ^e	Prüfintervention Lavage								
		0	61	37,4 (15,9)		60	37,8 (17,6)			
		3	59	47,1 (21,1)	n. g.	56	46,9 (24,9)	n. g.	-0,2 [-8,7; 8,3]; 0,97	n. g.
		6	59	46,0 (22,0)	n. g.	57	46,3 (26,4)	n. g.	0,3 [-8,6; 9,2]; 0,95	n. g.
		12	57	42,8 (21,2)	n. g.	54	43,6 (24,8)	n. g.	0,8 [-7,9; 9,4]; 0,86	n. g.
		24	57	44,4 (22,4)	n. g.	55	42,3 (24,2)	n. g.	-2,1 [-10,9; 6,6]; 0,63	n. g.
		Prüfintervention Lavage + Debridement								
		0	59	38,9 (19,3)		60	37,8 (17,6)			
		3	58	46,8 (21,9)	n. g.	56	46,9 (24,9)	n. g.	0,1 [-8,6; 8,8]; 0,98	n. g.
		6	55	45,1 (20,6)	n. g.	57	46,3 (26,4)	n. g.	1,2 [-7,7; 10,0]; 0,80	n. g.
		12	51	44,5 (24,3)	n. g.	54	43,6 (24,8)	n. g.	-1,0 [-10,5; 8,5]; 0,84	n. g.
		24	52	45,0 (23,0)	n. g.	55	42,3 (24,2)	n. g.	-2,7 [-11,8; 6,4]; 0,56	n. g.

(Fortsetzung)

Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instru- ment (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Moseley 2002	KSPS (0-100)	Prüfintervention Lavage								
		0	61	63,9 (15,7)		60	65,0 (13,9)			
		3	59	53,1 (20,7)	n. g.	56	48,8 (21,5)	n. g.	-4,3 [-12,1; 3,5]; 0,28	n. g.
		6	59	53,2 (22,6)	n. g.	57	47,6 (20,7)	n. g.	-5,6 [-13,6; 2,4]; 0,17	n. g.
		12	57	54,8 (19,8)	n. g.	53	48,9 (21,9)	n. g.	-6,0 [-13,8; 1,9]; 0,14	n. g.
		24	55	53,7 (23,7)	n. g.	55	51,6 (23,7)	n. g.	-2,1 [-11,1; 6,9]; 0,64	n. g.
		Prüfintervention Lavage + Debridement								
		0	58	62,9 (17,6)		60	65,0 (13,9)			
		3	58	49,3 (22,0)	n. g.	56	48,8 (21,5)	n. g.	-0,5 [-8,6; 7,5]; 0,89	n. g.
		6	56	50,0 (21,0)	n. g.	57	47,6 (20,7)	n. g.	-2,3 [-10,1; 5,4]; 0,55	n. g.
		12	50	51,7 (22,4)	n. g.	53	48,9 (21,9)	n. g.	-2,9 [-11,5; 5,8]; 0,51	n. g.
		24	53	51,4 (23,2)	n. g.	55	51,6 (23,7)	n. g.	0,2 [-8,8; 9,2]; 0,96	n. g.

(Fortsetzung)

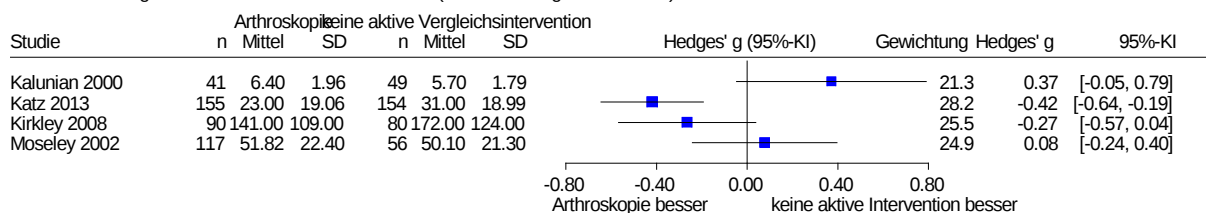
Tabelle 16: Ergebnisse zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instru- ment (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)		
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert	
Moseley 1996	Durch- schnittliche Schmerz- intensität (1-10) ^f	Prüfintervention Lavage									
		0	3	7,5 (n. g.)		5	6,8 (n. g.)				
		3	2	6,0 (n. g.)	n. g.	5	4,8 (n. g.)	n. g.	n. g.	n. g.	
		6	2	5,5 (n. g.)	n. g.	5	5,6 (n. g.)	n. g.	n. g.	n. g.	
		Prüfintervention Lavage + Debridement									
		0	2	7,0 (n. g.)		5	6,8 (n. g.)				
		3	2	4,5 (n. g.)	n. g.	5	4,8 (n. g.)	n. g.	n. g.	n. g.	
		6	2	4,5 (n. g.)	n. g.	5	5,6 (n. g.)	n. g.	n. g.	n. g.	
		<i>kursiv:</i> aus Grafik abgelesen a: Werte aus einer adjustierten Analyse (unabhängige Variable: Behandlungsgruppe, Kovariablen: Baseline-Wert und Schwellungen) b: Laut Literatur definiert bei dem Instrument KOOS ein Wert von 0 extreme Knie-Probleme. In der Studie Katz 2013 wird hingegen definiert, dass ein höherer Wert mehr Schmerzen bedeutet. c: Die teilweise einer Grafik entnommenen Werte konnten anhand der Veröffentlichung nicht uneingeschränkt nachvollzogen werden. Die Baseline-Werte entsprachen nicht den Werten der eingeschlossenen Patienten, sondern waren vermutlich um Patienten verringert, welche die Studie nicht abschlossen. Zusätzlich wurden in der Grafik geringere Zahlen an Patienten zur Baseline angegeben. d: adjustiert nach Baseline-Wert und Schweregrad e: Ursprünglicher Skalenbereich transformiert auf den Bereich 0-100 f: Likert-Skala:1 = kein Schmerz bis 10 = starker Schmerz AIMS: Arthritis impact measurement scale; KI: Konfidenzintervall; KOOS: Knee Injury and Osteoarthritis Outcome Scale; KSPS: knee-specific pain scale; MW: Mittelwert; N: Anzahl ausgewerteter Patienten; n. g.: nicht genannt; SD: Standardabweichung; SF-36: Short Form (36) Health Survey; VAS: visuelle Analogskala; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index									

Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 3 Monate

Schmerz

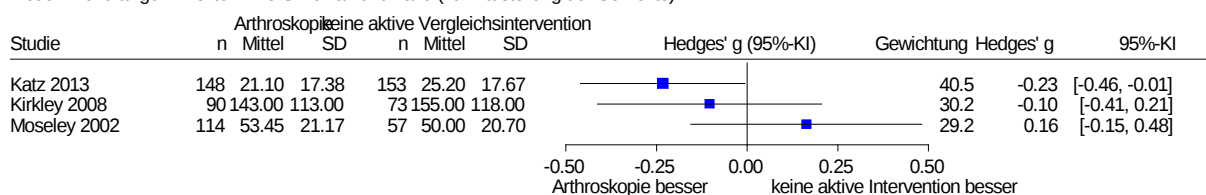
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 6 Monate

Schmerz

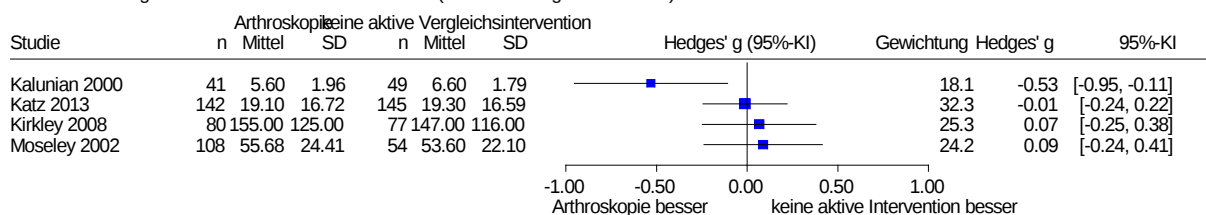
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 12 Monate

Schmerz

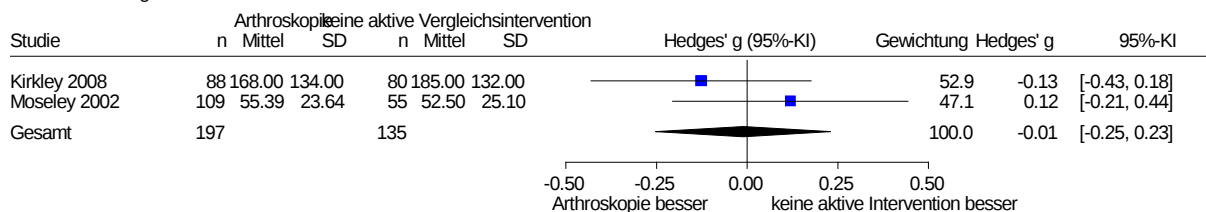
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 24 Monate

Schmerz

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 6, 12 und 24 Monate. Primäre Verwendung folgender Skalen: WOMAC Schmerz (Kalunian 2000; Kirkley 2008), KOOS Schmerz (Katz 2013), AIMS Schmerz (Moseley 2002). Abkürzungen: KI:

Konfidenzintervall; SD: standard deviation

Abbildung 3: Meta-Analysen zum Schmerz, Vergleich keine aktive Vergleichsintervention

Sensitivitätsanalysen

Die in den Studien Kalunian 2000 und Moseley 2002 zusätzlich eingesetzten Skalen VAS Schmerz und SF-36-Subskala Schmerz wurden in einer weiteren Meta-Analyse untersucht (siehe Anhang F). Für die Studie Kirkley 2008 wurde auch in dieser Meta-Analyse das Instrument WOMAC Schmerz eingesetzt, für die Studie Katz 2013 das Instrument KOOS-Schmerz, da in diesen Studien kein relevantes alternatives Instrument zur Schmerzmessung erhoben wurde.

In den Meta-Analysen zu den Zeitpunkten 3 und 12 Monate zeigte sich ebenfalls eine bedeutsame Heterogenität, sodass kein gemeinsamer Schätzer berechnet wurde. Die Ergebnisse waren auch bei Verwendung der alternativen Instrumente nicht gleichgerichtet. Zu den Zeitpunkten 6 und 24 Monate wurde in den Meta-Analysen mit den alternativen Instrumenten ein gemeinsamer Schätzer berechnet. Es konnte aber auch hier kein signifikanter Effekt zugunsten der therapeutischen Arthroskopie oder der Vergleichsintervention festgestellt werden. Unabhängig von der Verwendung der in den Studien erhobenen Instrumente zeigte sich somit zu keinem Zeitpunkt ein signifikanter Vorteil der therapeutischen Arthroskopie.

Auch die alleinige Betrachtung der Studien mit niedrigem Verzerrungspotenzial (Kalunian 2000 und Moseley 2002) führte zu qualitativ identischen Ergebnissen.

Insgesamt ergibt sich auch auf Basis der Sensitivitätsanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Subgruppenanalysen und weitere Effektmodifikatoren

Für die potenziellen Effektmodifikatoren Geschlecht, Alter und Erkrankung (Schweregrad, primäre oder sekundäre Arthrose) lagen keine Daten vor. Für die potenziellen Effektmodifikatoren „Art des arthroskopischen Eingriffs“ und „Art der Behandlung in der Kontrollgruppe“ wurden Interaktionstests durchgeführt (Tabelle 17).

Tabelle 17: Subgruppenanalysen (Interaktionstest) zum Endpunkt Schmerz

Effektmodifikator	p-Wert Interaktionstest			
Zeitpunkt in Monaten	3	6	12	24
Art des arthroskopischen Eingriffs Lavage vs. Lavage plus Debridement	0,002	0,131	0,599	0,370
Art der Behandlung in der Kontrollgruppe keine zusätzliche Behandlung vs. Scheinarthroskopie vs. diagnostische Arthroskopie	0,001	0,061	0,047	0,227

Für den potenziellen Effektmodifikator „Art des arthroskopischen Eingriffs“ ergab sich ein Beleg für eine Effektmodifikation für den Zeitpunkt 3 Monate ($p = 0,002$). Eine getrennte Betrachtung der Subgruppenergebnisse zeigte zwar numerisch gegenläufige Effekte, jedoch lag nur in der Subgruppe Debridement + Lavage ein statistisch signifikanter Effekt zugunsten der Arthroskopie vor (Abbildung 4). Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied und keine Responderanalysen zu einer validierten MID (minimal important difference) vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz ($SMD = \text{standardised mean difference}$; in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwertes überschritt allerdings die Irrelevanzschwelle von $-0,2$, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Für den Zeitpunkt 6 Monate ergab sich ein Hinweis für eine Effektmodifikation ($p = 0,131$). Eine getrennte Betrachtung der Subgruppenergebnisse zeigte zwar numerisch gegenläufige Effekte, jedoch waren diese in beiden Subgruppen nicht statistisch signifikant (Abbildung 4). Die Interaktionstests für die übrigen Zeitpunkte (12 und 24 Monate) waren nicht statistisch signifikant.

Für den potenziellen Effektmodifikator „Art der Behandlung in der Kontrollgruppe“ ergab der Interaktionstest einen Beleg für eine Effektmodifikation für die Zeitpunkte 3 Monate ($p = 0,001$) und 12 Monate ($p = 0,047$). Für den Zeitpunkt 6 Monate ergab sich ein Hinweis auf eine Effektmodifikation ($p = 0,061$). Die Subgruppe Scheinarthroskopie bestand aus der Studie Moseley 2002, die Subgruppe diagnostische Arthroskopie aus der Studie Kalunian 2000 und die Subgruppe keine zusätzliche Behandlung aus den Studien Katz 2013 und Kirkley 2008.

Aufgrund der Heterogenität der Subgruppen war eine Zusammenfassung einzelner Subgruppen nicht angezeigt. Eine getrennte Betrachtung der 3 Subgruppenergebnisse zum Zeitpunkt 3 Monate zeigte zwar numerisch gegenläufige Effekte in den Subgruppen, diese waren jedoch in den Subgruppen Scheinarthroskopie und diagnostische Arthroskopie nicht statistisch signifikant. In der Subgruppe keine zusätzliche Behandlung zeigte sich ein statistisch signifikanter Effekt (Abbildung 5). Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied und keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwertes überschritt allerdings die Irrelevanzschwelle von $-0,2$, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Für den Zeitpunkt 6 Monate lagen nur Ergebnisse der Subgruppen Scheinarthroskopie und keine zusätzliche Behandlung vor. In der Subgruppe Scheinarthroskopie zeigte sich kein statistisch signifikanter Effekt. In der Subgruppe keine zusätzliche Behandlung ergab sich ein statistisch signifikanter Effekt (Abbildung 5). Auch hier überschritt die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwerts die Irrelevanzschwelle von $-0,2$, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Bei der Prüfung der Homogenität einzelner Subgruppen für den Zeitpunkt 12 Monate zeigte sich eine homogene Datenlage für die Studien Katz 2013, Kirkley 2008 und Moseley 2002 (Heterogenitätstest: $p = 0,863$). Eine getrennte Betrachtung der resultierenden Subgruppen (diagnostische Arthroskopie vs. keine zusätzliche Behandlung oder Scheinarthroskopie) ergab keinen signifikanten Effekt für die Subgruppe mit den Studien Katz 2013, Kirkley 2008 und Moseley 2002. Die Studie Kalunian 2000 (Subgruppe diagnostische Arthroskopie) zeigte einen statistisch signifikanten Effekt (Abbildung 5). Auch hier überschritt die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwertes die Irrelevanzschwelle von

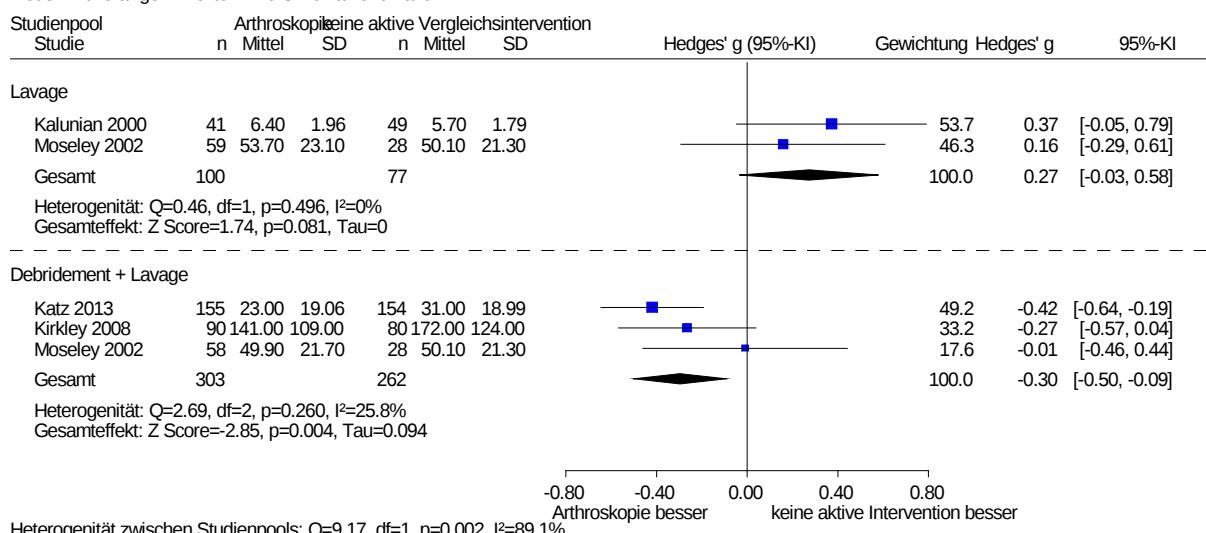
-0,2, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte. Der Interaktionstest für den Zeitpunkt 24 Monate war nicht statistisch signifikant.

Insgesamt ergibt sich auf Basis der Subgruppenanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Arthroskopie vs. keine aktive Vergleichsintervention - Art des arthroskopischen Eingriffs, 3 Monate

Schmerz

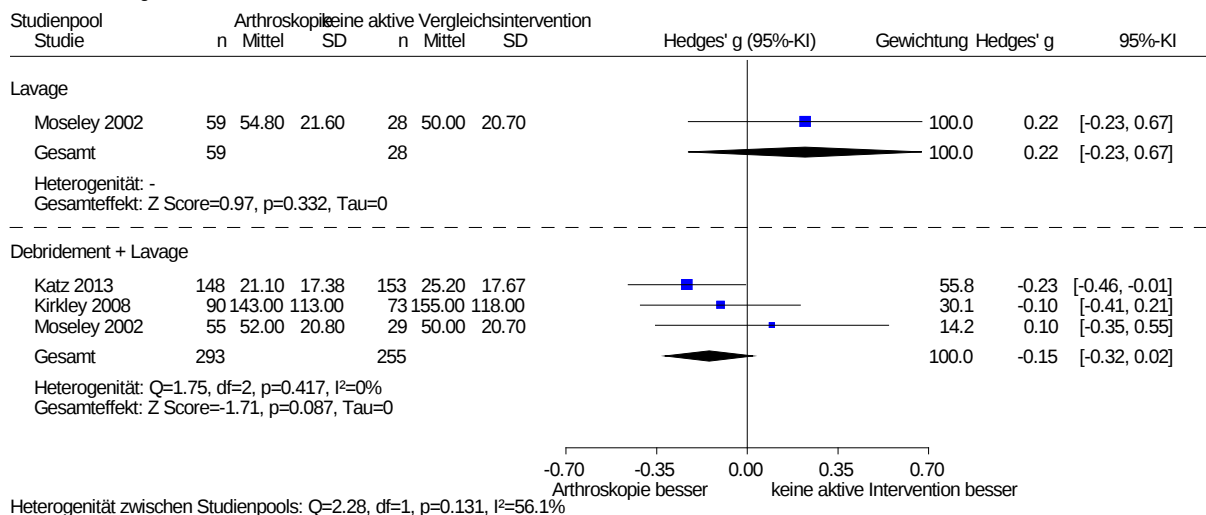
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - Art des arthroskopischen Eingriffs, 6 Monate

Schmerz

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkt: 3 und 6 Monate. Primäre Verwendung folgender Skalen: WOMAC Schmerz (Kalunian 2000; Kirkley 2008), KOOS Schmerz (Katz 2013), AIMS Schmerz (Moseley 2002). Da im Rahmen dieser Auswertungen in der Studie Moseley 2002 dieselbe Kontrollgruppe bei 2 Vergleichen einging, wurde die Anzahl der Patienten in der Kontrollgruppe jeweils halbiert.

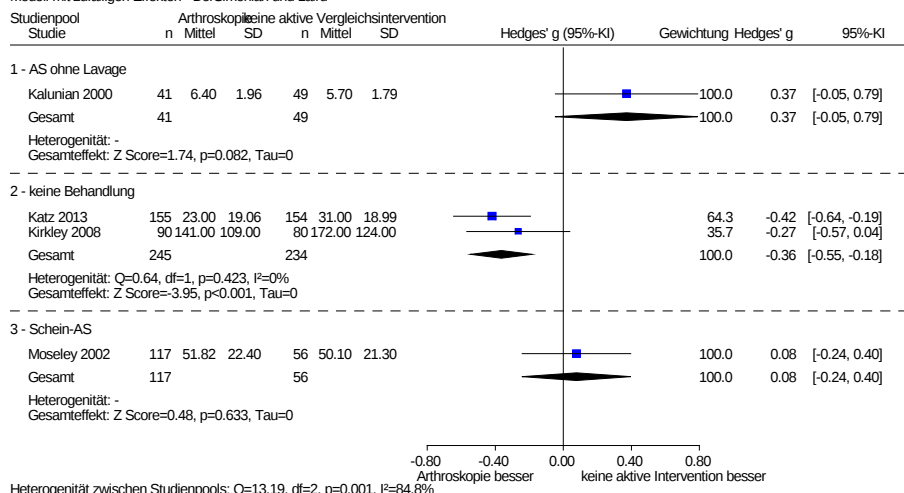
KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 4: Subgruppenanalyse zur Art des arthroskopischen Eingriffs beim Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention

Arthroskopie vs. keine aktive Vergleichsintervention - Art der Kontrollbehandlung, 3 Monate

Schmerz

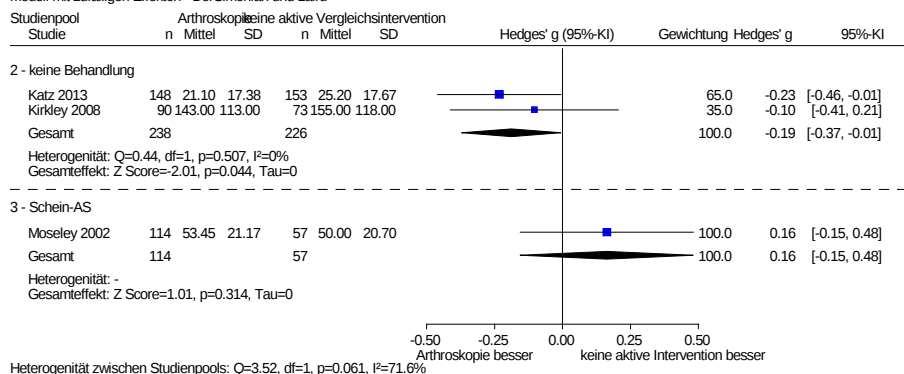
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - Art der Kontrollbehandlung, 6 Monate

Schmerz

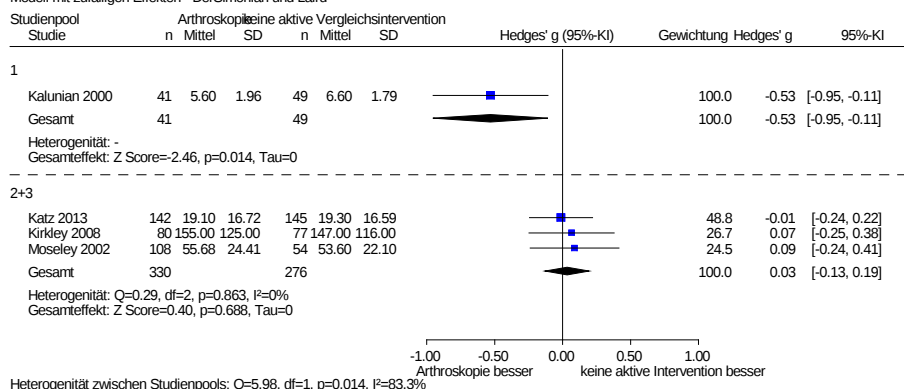
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - Art der Kontrollbehandlung, 12 Monate

Schmerz

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 6 und 12 Monate. Primäre Verwendung folgender Skalen: WOMAC Schmerz (Kalunian 2000; Kirkley 2008), KOOS Schmerz (Katz 2013), AIMS Schmerz (Moseley 2002).

AS: Arthroskopie; KI: Konfidenzintervall; SD: Standardabweichung; Subgruppen: 1. Diagnostische Arthroskopie, 2. Keine zusätzliche Behandlung, 3. Scheinarthroskopie

Abbildung 5: Subgruppenanalyse zur Art der Kontrollbehandlung beim Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention

5.3.1.1.2 Körperliche Funktion

Alle 5 eingeschlossenen Studien berichteten Daten zur körperlichen Funktion. Die verwendeten Instrumente sind in Tabelle 18 dargestellt.

Tabelle 18: Übersicht der verwendeten Instrumente zur Erfassung der körperlichen Funktion, Vergleich keine aktive Vergleichsintervention

Studien	Verwendete Instrumente
Kalunian 2000	WOMAC-Subskala körperliche Funktion WOMAC-Subskala Steifigkeit
Katz 2013	WOMAC-Subskala körperliche Funktion SF-36-Subskala körperliche Funktion
Kirkley 2008	WOMAC-Subskala körperliche Funktion WOMAC-Subskala Steifigkeit MACTAR-Subskala körperliche Funktion
Moseley 1996	15 Meter Gehen
Moseley 2002	AIMS-2-Subskala körperliche Funktion SF-36-Subskala körperliche Funktion PFS
AIMS: Arthritis impact measurement scales; MACTAR: McMaster-Toronto arthritis patient preference disability questionnaire; PFS: physical functioning scale; SF-36: Short Form (36) Health Survey; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index	

In den Studien Kalunian 2000 und Kirkley 2008 wurden die WOMAC-Subskalen zur körperlichen Funktion und zur Steifigkeit erhoben, bei Kirkley 2008 zusätzlich auch noch die MACTAR-Subskala zur körperlichen Funktion. Die Studie Katz 2013 erhob Daten mit der WOMAC-Subskala körperliche Funktion und der SF36-Subskala körperliche Funktion. In der Studie Moseley 1996 wurde die Zeit, die die Patienten benötigten um 15 Meter zu gehen dargestellt, und in dem vorliegenden Bericht wird dies wiedergegeben (siehe auch Abschnitt 5.3.1). In der Studie Moseley 2002 wurden neben der AIMS-2-Subskala zur körperlichen Funktion auch die Subskala des SF-36 zur körperlichen Funktion und ein selbst entwickeltes Instrument PFS (physical functioning scale), das objektive Funktionseinschätzungen beinhaltet, verwendet. Da die Erfassung des Symptoms Steifigkeit sich konzeptionell unterscheidet von der Erfassung der körperlichen Funktion im Allgemeinen, wurden die Ergebnisse der unterschiedlichen Subskalen nicht gemeinsam in einer Meta-Analyse betrachtet.

Verzerrungspotenzial der Ergebnisse zur körperlichen Funktion

Die Angaben zum Verzerrungspotenzial auf Endpunktebene befinden sich in Tabelle 19. Die Studien Kalunian 2000 und Moseley 2002 hatten ein niedriges Verzerrungspotenzial auf Endpunktebene. Die Studien Katz 2013 und Kirkley 2008 hatten auch auf Endpunktebene ein hohes Verzerrungspotenzial. Die nicht verblindete Endpunkterhebung sowie die Umsetzung des ITT-Prinzips waren nicht immer adäquat. Die Studie Moseley 1996 hatte auf

Endpunktebene keine weiteren Einschränkungen, die das Verzerrungspotenzial erhöhten, allerdings musste aufgrund der unklaren Randomisierung und der Verdeckung der Gruppenzuteilung sowie der fehlenden Verblindung (siehe Tabelle 12) auch auf Endpunktebene von einem hohen Verzerrungspotenzial ausgegangen werden.

Ergebnisse zur körperlichen Funktion

In Tabelle 20 sind die Ergebnisse zum Endpunkt körperliche Funktion dargestellt. Abbildung 6 zeigt die Meta-Analysen der Ergebnisse zum Endpunkt körperliche Funktion jeweils für die Zeitpunkte 3, 6, 12 und 24 Monate.

In der Meta-Analyse wurde die WOMAC-Subskala körperliche Funktion (Kalunian 2000, Katz 2013, Kirkley 2008) und die AIMS-2-Subskala körperliche Funktion (Moseley 2002) dargestellt (für die zusätzlich erhobenen Instrumente der Studien Katz 2013, Kirkley 2008 und Moseley 2002 siehe den Abschnitt zu Sensitivitätsanalysen). In den Meta-Analysen zu den Zeitpunkten 3, 6 und 12 Monate zeigte sich eine bedeutsame Heterogenität ($p < 0,2$), sodass kein gemeinsamer Schätzer berechnet wurde. Die Ergebnisse waren nicht gleichgerichtet. In der Meta-Analyse zu dem Zeitpunkt 24 Monate konnte ein gemeinsamer Schätzer berechnet werden, allerdings lagen nur die Daten aus 2 Studien (Kirkley 2008 und Moseley 2002) vor. Ein statistisch signifikanter Effekt zugunsten der therapeutischen Arthroskopie oder der Vergleichsintervention konnte nicht festgestellt werden.

In der Studie Katz 2013 wurde für das Instrument WOMAC-Subskala körperliche Funktion zum Zeitpunkt 6 Monate auch eine Responderanalyse, basierend auf einem validierten Responsekriterium, in der Publikation berechnet, bei der sich ein statistisch signifikanter Vorteil der therapeutischen Arthroskopie im Vergleich zu keiner zusätzlichen Behandlung zeigte.

In der Studie Moseley 2002 wurde zusätzlich noch das Instrument PFS eingesetzt, bei dem die körperliche Funktion durch den Arzt erfasst wurde. Bei Moseley 2002 ergaben sich für die Prüfintervention Lavage zu keinem Zeitpunkt statistisch signifikante Unterschiede zur Scheinbehandlung. Bei der Prüfintervention Lavage und Debridement ergab sich für den Zeitpunkt 12 Monate ein signifikanter Gruppenunterschied zugunsten der Prüfintervention. Dieser Effekt bestätigte sich allerdings nicht zu anderen Zeitpunkten. Auch in der Studie Moseley 1996 wurden Funktionseinschätzungen erhoben.

In der Studie Kirkley 2008 wurde zudem auch noch die Subskala körperliche Funktion des Instruments MACTAR erhoben. Hier zeigte sich zum Zeitpunkt 24 Monate kein signifikanter Unterschied zwischen den Behandlungsgruppen.

In den Studien Kalunian 2000 und Kirkley 2008 wurde zudem die WOMAC-Subskala zur Steifigkeit erhoben. In beiden Studien ergaben sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen.

Da es sich bei dem statistisch signifikanten Effekt der Responderanalyse der Studie Katz 2013 um ein Einzelergebnis handelt, das in der Gesamtschau aller Studien keine Bestätigung findet, wurde hieraus kein Vorteil der Prüfintervention abgeleitet.

Zusammenfassend ergibt sich kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Tabelle 19: Bewertung des Verzerrungspotenzials auf Endpunktebene: körperliche Funktion, Vergleich keine aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisun- abhängige Bericht- erstattung	Fehlen sonstiger Aspekte, die das Verzerrungs- potenzial beeinflussen	Verzerrungspotenzial für diesen Endpunkt
Kalunian 2000	niedrig	ja	ja ^a	ja	ja	niedrig
Katz 2013	niedrig	nein	ja / nein ^b	ja	ja	hoch ^c
Kirkley 2008	hoch	nein ^d	nein ^e	ja	ja	hoch ^f
Moseley 1996	hoch	ja	ja	ja	ja	hoch ^g
Moseley 2002	niedrig	ja	ja / nein ^h	ja	ja	niedrig / hoch ⁱ
a: Bewertung basiert auf zusätzlichen Informationen aus einer Autorenanfrage. b: Bis zu 20 % fehlende Werte; zu den Zeitpunkten 6 und 12 Monate Sensitivitätsanalysen mit Ersetzung fehlender Werte. c: keine Verblindung, teilweise keine adäquate ITT-Auswertung. d: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung erfolgt durch unverblindeten Patienten begleitet von verblindetem Endpunkterheber. e: unklare oder fehlende ITT-Auswertung f: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung g: hohes Verzerrungspotenzial auf Studienebene h: Für die Instrument: AIMS-2-Subskala körperliche Funktion und SF-36-Subskala körperliche Funktion war das ITT-Prinzip adäquat umgesetzt. Für das Instrument PFS wurden 25 % der Werte nicht in der Auswertung berücksichtigt, es lag keine adäquate Umsetzung des ITT-Prinzips vor. i: Für die Erhebung des Endpunkts mit den Instrumenten AIMS und SF 36 lag ein niedriges Verzerrungspotenzial vor, für die Erhebung mit dem Instrument PFS lag aufgrund inadäquater ITT-Analyse ein hohes Verzerrungspotenzial vor. AIMS-2: Arthritis Impact Measurement Scales ; ITT: intention to treat; SF-36: Short Form (36) Health Survey						

Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention

Studie	Instrument (Bereich)	Zeitpunkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p- Wert	Änderung zu Baseline 95 %- KI]; p-Wert
Kalunian 2000	WOMAC körperliche Funktion (0-68)	0	41	27 [28; 29]		49	28 [25; 31]			
		3	n. g.	23 [20; 26]	n. g.	n. g.	21 [18; 23]	n. g.	n. g.	n. g.
		12	n. g.	19 [17; 21]	9,9 [4,9; 13,0]	n. g.	22 [20; 24]	6,1 [2,8; 9,4]	n. g.	n. g. (n. g.); 0,15 ^a
	WOMAC Steifigkeit (0-8)	0	41	3,6 [3,3; 3,9]		49	3,8 [3,5; 4,1]			
		3	n. g.	2,9 [2,6; 3,2]	n. g.	n. g.	2,8 [2,5; 3,1]	n. g.	n. g.	n. g.
		12	n. g.	2,6 [2,3; 2,9]	1,2 [-1,6; 4,0]	n. g.	3,1 [2,8; 3,4]	0,7 [-0,5; 1,9]	n. g.	n. g. (n. g.); 0,22 ^a
Katz 2013 ^b	WOMAC körperliche Funktion (0-100)	0	159	37,1 (17,9)		167	37,5 (18,3)			
		3	154	16,0 [14,0; 18,0]		155	22,0 [19,0; 25,0]			
		6	147	14,7 [12,0; 17,5]	20,9 [17,9; 23,9]	153	19 [16,3; 21,7]	18,5 [15,6; 21,5]	n. g.	2,4 [-1,8; 6,5]; n. g. ^c
		12	140	13,7 [11,2; 16,2]	23,5 [20,5; 26,5]	144	14,5 [12,0; 16,9]	22,8 [19,8; 25,8]	n. g.	0,7 [-3,5; 4,9]; n. g. ^c
	SF-36 körperliche Funktion (0-100)	0	161	44,3 (23,7)		169	43,3 (23,3)			
		6	n. g.	69,2 [65,2; 73,2]	24,2 [20,3; 28,0]	n. g.	66,1 [62,1; 70,1]	23,1 [19,2; 27,0]	n. g.	1,1 (-4,4; 6,6)
		12	n. g.	69,0 [64,6; 73,4]	25,0 [20,9; 29,1]	n. g.	71,4 [67,0; 75,7]	28,1 [24,0; 32,1]	n. g.	-3,0 [-8,8; 2,7]

(Fortsetzung)

Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p- Wert	Änderung zu Baseline 95 %- KI]; p-Wert
Kirkley 2008	WOMAC körperliche Funktion (0-1700)	0	92	830 (355)		86	726 (397)			
		3	90	522 (341)	n. g.	80	568 (369)	n. g.	n. g.	n. g.
		6	90	551 (382)	n. g.	73	520 (368)	n. g.	n. g.	n. g.
		12	80	570 (417)	n. g.	77	513 (370)	n. g.	n. g.	n. g.
		24	88	612 (448)	n. g.	80	623 (439)	n. g.	n. g. (n. g.); 0,26 ^d	n. g.
	WOMAC Steifigkeit (0-200)	0	92	117 (50)		86	103 (48)			
		3	90	80 (54)	n. g.	80	84 (53)	n. g.	n. g.	n. g.
		6	90	86 (53)	n. g.	73	82 (47)	n. g.	n. g.	n. g.
		12	80	85 (56)	n. g.	77	81 (51)	n. g.	n. g.	n. g.
		24	88	93 (60)	n. g.	80	88 (51)	n. g.	n. g. (n. g.); 1,00 ^d	n. g.
	MACTAR (0-500)	0	92	348 (85)		86	320 (99)			
		3	90	257 (108)	n. g.	80	249 (109)	n. g.	n. g.	n. g.
		6	90	234 (118)	n. g.	73	246 (115)	n. g.	n. g.	n. g.
		12	80	232 (128)	n. g.	77	225 (117)	n. g.	n. g.	n. g.
		24	88	238 (146)	n. g.	80	244 (133)	n. g.	n. g. (n. g.); 0,58 ^d	n. g.

(Fortsetzung)

Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %- KI]; p-Wert	Änderung zu Baseline 95 %- KI]; p-Wert
Moseley 2002	PFS (Sekunden) ^e	Prüfintervention Lavage								
		0	59	50,0 (14,3)		59	48,5 (14,5)			
		3	55	48,8 (21,0)	n. g.	54	47,3 (16,0)	n. g.	-1,4 [-8,5; 5,6]; 0,69	n. g.
		6	52	49,4 (20,4)	n. g.	54	47,0 (13,0)	n. g.	-2,4 [-9,0; 4,1]; 0,47	n. g.
		12	54	50,4 (17,6)	n. g.	49	45,6 (10,2)	n. g.	-4,9 [-10,6; 0,9]; 0,09	n. g.
		24	50	53,2 (21,6)	n. g.	44	47,7 (12,0)	n. g.	-5,5 [-12,8; 1,8]; 0,13	n. g.
		Prüfintervention Lavage + Debridement								
		0	58	52,1 (20,2)		59	48,5 (14,5)			
		3	56	49,5 (17,4)	n. g.	54	47,3 (16,0)	n. g.	-2,2 [-8,5; 4,1]; 0,49	n. g.
		6	54	49,8 (17,4)	n. g.	54	47,0 (13,0)	n. g.	-2,8 [-8,7; 3,1]; 0,34	n. g.
		12	47	52,5 (20,3)	n. g.	49	45,6 (10,2)	n. g.	-6,9 [-13,3; -0,4]; 0,04	n. g.
		24	44	52,6 (16,4)	n. g.	44	47,7 (12,0)	n. g.	-4,9 [-11,0; 1,2]; 0,11	n. g.
	SF-36 körperliche Funktion (0-100)	Prüfintervention Lavage								
		0	61	44,4 (22,8)		60	46,8 (22,5)			
		3	59	52,9 (26,7)	n. g.	56	52,4 (23,5)	n. g.	-0,5 [-9,8; 8,8]; 0,92	n. g.
		6	59	53,4 (27,6)	n. g.	57	48,4 (25,9)	n. g.	-5,0 [-14,8; 4,9]; 0,32	n. g.
		12	57	50,0 (28,0)	n. g.	54	49,3 (24,5)	n. g.	-0,6 [-10,5; 9,3]; 0,90	n. g.
		24	57	50,9 (27,3)	n. g.	54	49,0 (27,2)	n. g.	-1,9 [-12,2; 8,3]; 0,71	n. g.

(Fortsetzung)

Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %- KI]; p-Wert	Änderung zu Baseline 95 %- KI]; p-Wert
Moseley 2002	SF-36 körperliche Funktion (0-100) (Fortsetzung)	Prüfintervention Lavage + Debridement								
		0	59	42,2 (22,4)		60	46,8 (22,5)			
		3	58	49,6 (24,2)	n. g.	56	52,4 (23,5)	n. g.	2,8 [-6,0; 11,7]; 0,53	n. g.
		6	55	51,0 (25,9)	n. g.	57	48,4 (25,9)	n. g.	-2,6 [-12,3; 7,1]; 0,60	n. g.
		12	50	47,3 (27,1)	n. g.	54	49,3 (24,5)	n. g.	2,0 [-8,0; 12,1]; 0,69	n. g.
		24	52	47,9 (26,6)	n. g.	54	49,0 (27,2)	n. g.	1,1 [-9,3; 11,5]; 0,83	n. g.
	AIMS-2 körperliche Funktion (0-100)	Prüfintervention Lavage								
		0	61	55,5 (25,0)		60	58,5 (21,4)			
		3	59	47,9 (30,1)	n. g.	56	49,9 (21,6)	n. g.	2,0 [-7,7; 11,7]; 0,68	n. g.
		6	59	48,7 (31,6)	n. g.	57	49,1 (25,8)	n. g.	0,4 [-10,2; 11,1]; 0,94	n. g.
		12	57	49,6 (29,1)	n. g.	54	49,4 (25,5)	n. g.	-0,1 [-10,4; 10,2]; 0,98	n. g.
		24	56	51,1 (28,3)	n. g.	55	53,8 (27,5)	n. g.	2,7 [-7,8; 13,2]; 0,61	n. g.
		Prüfintervention Lavage + Debridement								
		0	58	59,3 (26,0)		60	58,5 (21,4)			
		3	58	53,5 (28,6)	n. g.	56	49,9 (21,6)	n. g.	-3,6 [-13,0; 5,8]; 0,45	n. g.
		6	55	52,5 (28,7)	n. g.	57	49,1 (25,8)	n. g.	-3,4 [-13,6; 6,8]; 0,51	n. g.
		12	51	56,4 (28,4)	n. g.	54	49,4 (25,5)	n. g.	-7,0 [-17,4; 3,4]; 0,19	n. g.
		24	53	56,4 (29,4)	n. g.	55	53,8 (27,5)	n. g.	-2,6 [-13,4; 8,2]; 0,64	n. g.

(Fortsetzung)

Tabelle 20: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention (Fortsetzung)

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)		
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %- KI]; p-Wert	Änderung zu Baseline [95 %- KI]; p-Wert	
Moseley 1996	15 Meter Gehen (Sekunden)	Prüfintervention Lavage									
		0	3	11,0 (n. g.)		5	11,5 (n. g.)				
		3	2	11,5 (n. g.)	n. g.	5	11,4 (n. g.)	n. g.	n. g.	n. g.	
		6	2	11,5 (n. g.)	n. g.	5	12,0 (n. g.)	n. g.	n. g.	n. g.	
		Prüfintervention Lavage + Debridement									
		0	2	12,0 (n. g.)		5	11,5 (n. g.)				
		3	2	9,0 (n. g.)	n. g.	5	11,4 (n. g.)	n. g.	n. g.	n. g.	
		6	2	10,5 (n. g.)	n. g.	5	12,0 (n. g.)	n. g.	n. g.	n. g.	
		<i>kursiv:</i> aus Grafik abgelesen									
a: Werte aus einer adjustierten Analyse (unabhängige Variable: Behandlungsgruppe, Kovariablen: Baseline-Wert und Schwellungen)											
b: Die teilweise einer Grafik entnommenen Werte konnten anhand der Veröffentlichung nicht uneingeschränkt nachvollzogen werden. Die Baseline-Werte entsprachen nicht den Werten der eingeschlossenen Patienten, sondern waren vermutlich um Patienten verringert, welche die Studie nicht abschlossen. Zusätzlich wurden in der Grafik geringere Zahlen an Patienten zur Baseline angegeben.											
c: ergänzend wurden Ergebnisse aus Sensitivitätsanalysen mit Ersetzung fehlender Werte berichtet: 3,7 [-0,9; 7,1] (6 Monate), 0,8 [-3,3; 4,9] (12 Monate)											
d: adjustiert nach Baseline-Wert und radiologischem Schweregrad											
e: Zeit in Sekunden für das Gehen von 30 Meter ebener Strecke und das Hinauf- und Hinabsteigen einer Treppe											
AIMS: Arthritis impact measurement scale; KI: Konfidenzintervall; MACTAR: McMaster-Toronto arthritis patient preference disability questionnaire; MW: Mittelwert; N: Anzahl ausgewerteter Patienten; n. g.: nicht genannt; SD: Standardabweichung; PFS: physical functioning scale; SF-36: Short Form (36) Health Survey; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index											

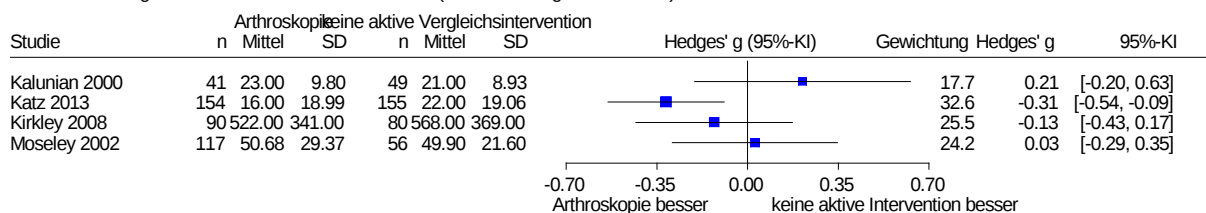
Tabelle 21: Ergebnisse zum Endpunkt körperliche Funktion , Vergleich keine aktive Vergleichsintervention – Responderanalysen

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention		Vergleichsintervention		Gruppenunterschied Effektschätzer [95 %-KI]; p-Wert
			N	Patienten mit Response n (%)	N	Patienten mit Response n (%)	
Katz 2013 (keine zusätzliche Behandlung)	WOMAC körperliche Funktion ^a	6	161	108 (67)	169	74 (44)	n. g., [n. g.]; 0,001
a: Responder-Definition: Verbesserung von ≥ 8 Punkten KI: Konfidenzintervall; N:Anzahl Patienten in Gruppe; n:Anzahl Patienten mit Response; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index							

Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 3 Monate

Körperliche Funktion

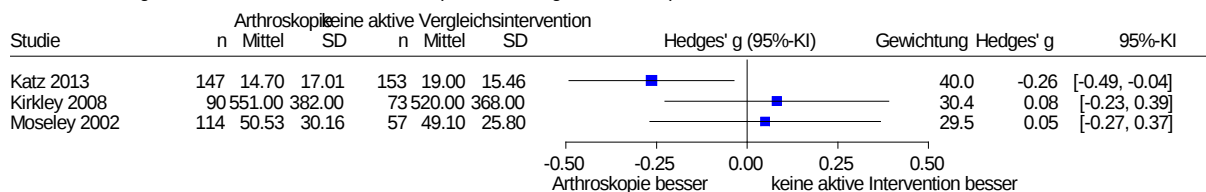
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 6 Monate

Körperliche Funktion

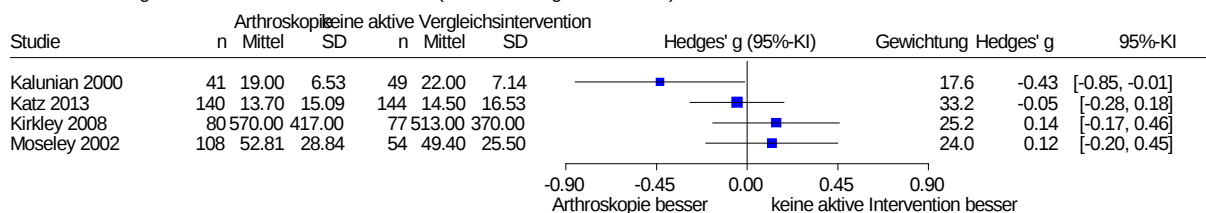
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 12 Monate

Körperliche Funktion

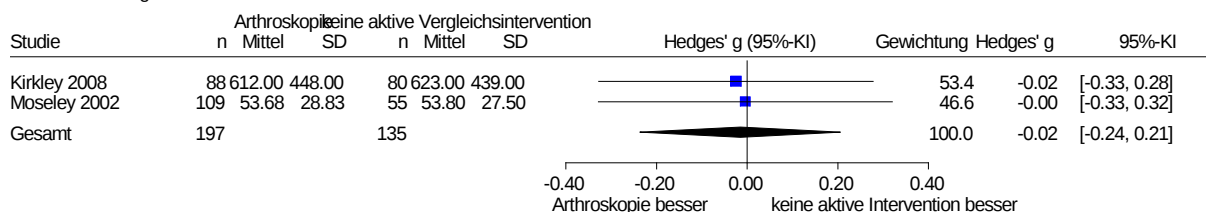
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 24 Monate

Körperliche Funktion

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 6, 12, 24 Monate. Primäre Verwendung folgender Skalen: WOMAC körperliche Funktion (Kalunian 2000; Katz 2013, Kirkley 2008), AIMS-2 körperliche Funktion (Moseley 2002).

KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 6: Meta-Analysen zur körperlichen Funktion, Vergleich keine aktive Vergleichsintervention

Sensitivitätsanalysen

Die in den Studien Katz 2013 und Moseley 2002 zusätzlich eingesetzte Subskala des SF-36 zur körperlichen Funktion wurde in einer zusätzlichen Meta-Analyse untersucht (siehe Anhang F). Bei Verwendung dieses alternativen Instruments zeigte sich ein identisches Bild wie bei den zuvor beschriebenen Meta-Analysen mit den primär verwendeten Instrumenten

AIMS-2-Subskala körperliche Funktion und WOMAC-Subskala körperliche Funktion. Nur für die Zeitpunkte 6 und 24 Monate konnte ein gemeinsamer Schätzer berechnet werden. Hierbei konnte kein signifikanter Effekt zugunsten der Prüf- oder Vergleichsintervention festgestellt werden.

Auch die alleinige Betrachtung der Studien mit niedrigem Verzerrungspotenzial (Kalunian 2000, Moseley 2002) führte zu qualitativ identischen Ergebnissen.

Insgesamt ergibt sich auch auf Basis der Sensitivitätsanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Subgruppenanalysen und weitere Effektmodifikatoren

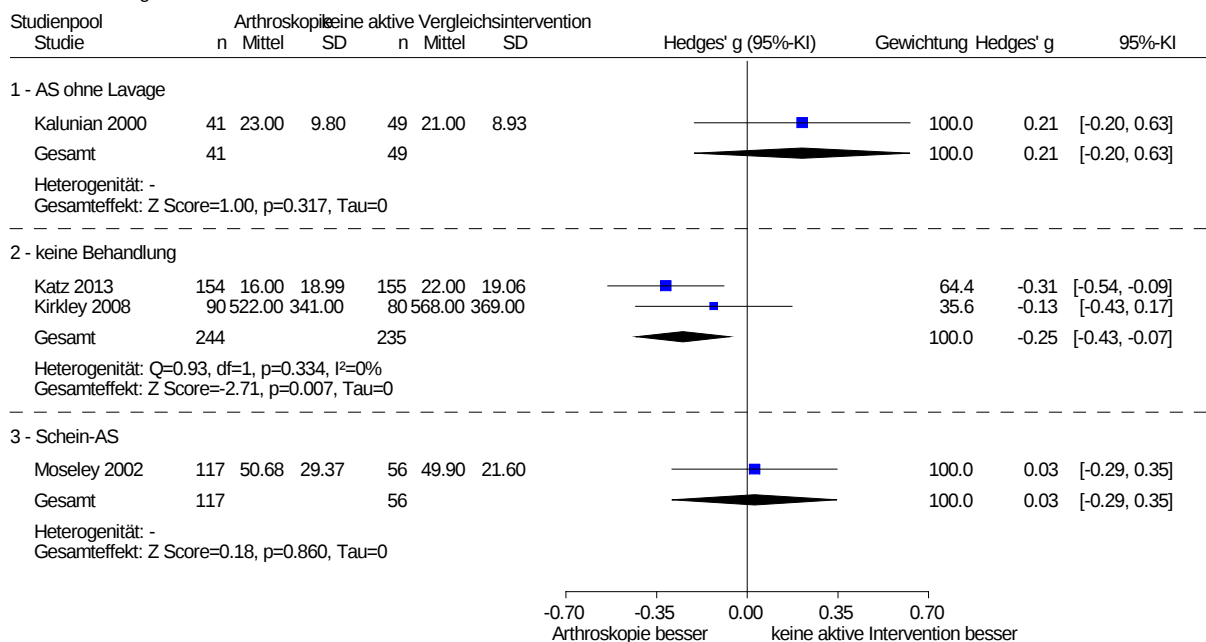
Für die Effektmodifikatoren Geschlecht, Alter und Erkrankung (Schweregrad, primäre oder sekundäre Arthrose) lagen keine Daten vor. Für die potenziellen Effektmodifikatoren „Art des arthroskopischen Eingriffs“ und „Art der Behandlung in der Kontrollgruppe“ wurden Interaktionstests durchgeführt (Tabelle 22).

Tabelle 22: Subgruppenanalysen körperliche Funktion

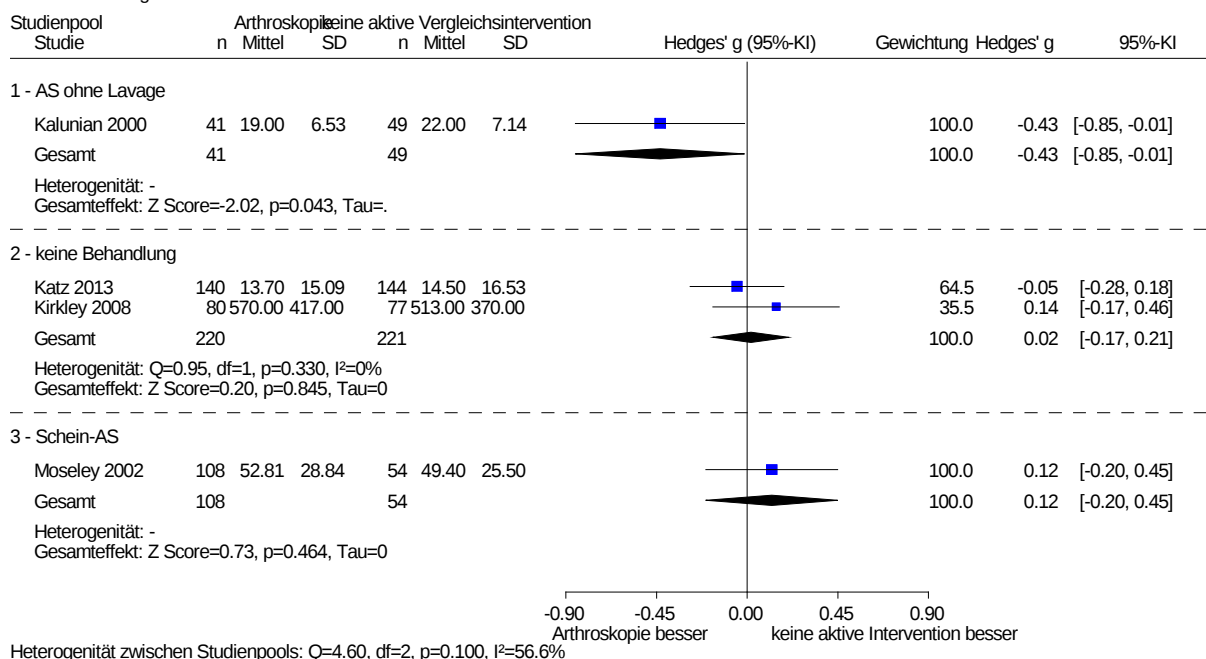
Effektmodifikator	p-Wert Interaktionstest			
	3	6	12	24
Zeitpunkt in Monaten				
Art des arthroskopischen Eingriffs Lavage vs. Lavage plus Debridement	0,210	0,864	0,248	0,694
Art der Behandlung in der Kontrollgruppe keine zusätzliche Behandlung vs. Schein-AS vs. diagnostische AS	0,074	0,507	0,100	0,928
AS: Arthroskopie				

Für den potenziellen Effektmodifikator „Art des arthroskopischen Eingriffs“ ergab sich für keinen Zeitpunkt ein Hinweis auf eine Effektmodifikation.

Arthroskopie vs. keine aktive Vergleichsintervention - Art der Kontrollbehandlung, 3 Monate
Körperliche Funktion
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - Art der Kontrollbehandlung, 12 Monate
Körperliche Funktion
Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 12 Monate. Primäre Verwendung folgender Skalen: WOMAC körperliche Funktion (Kalunian 2000; Katz 2013, Kirkley 2008), AIMS-2 körperliche Funktion (Moseley 2002).

AS: Arthroskopie; KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 7: Subgruppenanalyse zur Art der Kontrollbehandlung beim Endpunkt körperlichen Funktion, Vergleich keine aktive Vergleichsintervention

Für den potenziellen Effektmodifikator „Art der Behandlung in der Kontrollgruppe“ ergab der Interaktionstest einen Hinweis auf eine Effektmodifikation für die Zeitpunkte 3 und 12 Monate ($p = 0,074$ und $p = 0,100$). Die Interaktionstests für die übrigen Zeitpunkte (6 und 24 Monate) waren nicht statistisch signifikant.

Für den Zeitpunkt 3 Monate ergab sich in den Subgruppen diagnostische Arthroskopie und Scheinarthroskopie kein statistisch signifikanter Effekt. In der Subgruppe keine zusätzliche Behandlung zeigte sich ein statistisch signifikanter Effekt (Abbildung 7). Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied und keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwertes überschritt allerdings die Irrelevanzschwelle von $-0,2$, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Für den Zeitpunkt 12 Monate ergab sich in den Subgruppen keine zusätzliche Behandlung und Scheinarthroskopie kein statistisch signifikanter Effekt. In der Subgruppe diagnostische Arthroskopie ergab sich ein statistisch signifikanter Effekt (Abbildung 7). Auch hier überschritt die obere Grenze des 95 %-Konfidenzintervalls des standardisierten Mittelwertes die Irrelevanzschwelle von $-0,2$, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Insgesamt ergibt sich auf Basis der Subgruppenanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

5.3.1.1.3 Globale Bewertung der Symptomatik

Tabelle 23: Übersicht der verwendeten Instrumente zur Erfassung der globalen Bewertung der Symptomatik beim Vergleich keine aktive Vergleichsintervention

Studien	Verwendete Instrumente
Kalunian 2000	WOMAC global
Kirkley 2008	WOMAC global
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index	

Die Studien Kalunian 2000 und Kirkley 2008 stellten Daten zur globalen Bewertung der Symptomatik der Gonarthrose dar. Beide Studien verwendeten hierzu das Instrument WOMAC, allerdings in einer unterschiedlichen Version (Kirkley 2008 die VAS- und Kalunian 2000 die Likert-Version), wodurch die Effekte auf unterschiedlich breiten Skalenbereichen abgetragen wurden. Für die Meta-Analysen wurde daher die gewichtete Mittelwertdifferenz als Effektmaß verwendet.

Verzerrungspotenzial der Ergebnisse zur globalen Bewertung der Symptomatik der Gonarthrose

Das Verzerrungspotenzial auf Endpunktebene war bei Kalunian 2000 niedrig. Bei Kirkley 2008 zeigte sich auf Endpunktebene ein hohes Verzerrungspotenzial. Da die Erfassung des Endpunkts anhand eines durch den Patienten ausgefüllten Instruments (WOMAC) erfolgte und der Patient nicht verblindet war, lag keine verblindete Endpunkterhebung vor. Des Weiteren war die Umsetzung des ITT-Prinzips nicht adäquat, die Behandlungsgruppen unterschieden sich teilweise erheblich in den Raten fehlender Werte. Eine Übersicht der Bewertung des Verzerrungspotenzials zeigt Tabelle 24.

Ergebnisse zum globalen Bewertung der Symptomatik der Gonarthrose

In Tabelle 25 sind die Ergebnisse zum Endpunkt globale Bewertung der Symptomatik der Gonarthrose dargestellt.

Abbildung 8 zeigt die Meta-Analysen der Ergebnisse zum Endpunkt globale Bewertung der Symptomatik jeweils für den Zeitpunkt 3 Monate und 12 Monate nach Therapiebeginn, da nur für diese Messzeitpunkte ein Ergebnis vorlag. In beiden Meta-Analysen zeigte sich eine bedeutsame Heterogenität ($p < 0,2$). Sowohl für den Zeitpunkt 3 Monate als auch für den Zeitpunkt 12 Monate lag eine qualitative Heterogenität vor. Ein gemeinsamer Schätzer wurde daher nicht berechnet. Zum Zeitpunkt 3 Monate waren die Ergebnisse gegenläufig, aber nicht statistisch signifikant. Bei dem Zeitpunkt 12 Monate lagen ebenfalls gegenläufige Effekte vor, hier zeigte die Studie Kalunian 2000 allerdings einen statistisch signifikanten Vorteil zugunsten der therapeutischen Arthroskopie. Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied und keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Obwohl der Effekt in der Studie Kalunian 2000 statistisch signifikant war, lag das 95 %-Konfidenzintervall nicht vollständig unterhalb der Irrelevanzschwelle von $-0,2$. Ein irrelevanter Effekt konnte somit nicht sicher ausgeschlossen werden. Daher konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie für diesen Endpunkt abgeleitet werden.

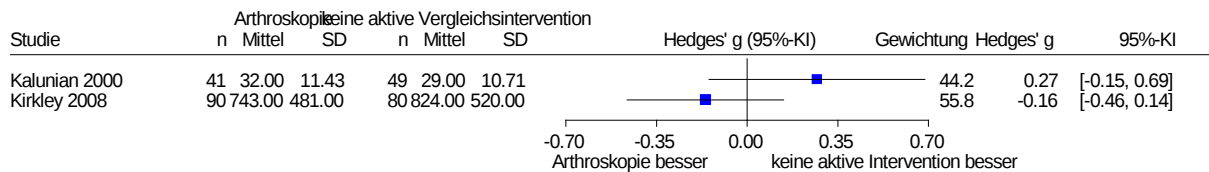
Tabelle 24: Bewertung des Verzerrungspotenzials auf Endpunktebene: globale Bewertung der Symptomatik beim Vergleich keine aktive Vergleichsintervention

Studie	Verzerrungspotenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungs- potenzial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Kalunian 2000	niedrig	ja	ja ^a	ja	ja	niedrig
Kirkley 2008	hoch	nein ^b	nein ^c	ja	ja	hoch ^d
a: Bewertung basiert aus zusätzlichen Informationen aus einer Autorenanfrage. b: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung erfolgt durch unverblindeten Patienten begleitet vom verblindeten Endpunkterheber. c: unklare oder fehlende ITT-Auswertung d: hohes Verzerrungspotenzial auf Studienebene, keine verblindete Endpunkterhebung, unklare ITT-Auswertung ITT: intention to treat						

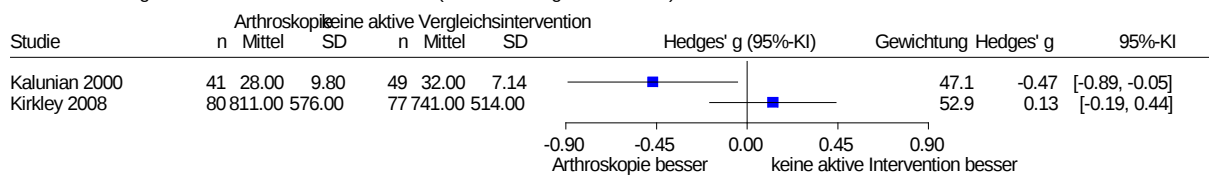
Tabelle 25: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich keine aktive Vergleichsintervention

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Kalunian 2000	WOMAC Global (0-96)	0	41	41,1 [37; 43]		49	40,7 [38; 43]			
		3	n. g.	32 [28; 35]	n. g.	n. g.	29 [26;32]	n. g.	n. g.	n. g.
		12	n. g.	28 [25; 31]	15,5 [7,7; 23,4]	n. g.	32 [30; 34]	8,9 [4,9; 13,0]	n. g.	n. g. [n. g.]; 0,10 ^a
Kirkley 2008	WOMAC Global (0-2400)	0	92	1187 (483)		86	1043 (542)			
		3	90	743 (481)	n. g.	80	824 (520)	n. g.	n. g. [n. g.]; 0,293 ^b	n. g.
		6	90	780 (518)	n. g.	73	757 (510)	n. g.	n. g. [n. g.]; 0,777 ^b	n. g.
		12	80	811 (576)	n. g.	77	741 (514)	n. g.	n. g. [n. g.]; 0,425 ^b	n. g.
		24	88	874 (624)	n. g.	80	897 (583)	n. g.	-23 [-208; 161]; 0,22 ^c	n. g.
<i>kursiv:</i> aus Grafik abgelesen a: Werte aus einer adjustierten Analyse (unabhängige Variable: Behandlungsgruppe, Kovariablen: Baseline-Wert und Schwellungen) b: t-Test (eigene Berechnung) c: adjustiert nach Baseline-Wert und radiologischem Schweregrad KI: Konfidenzintervall; Mon: Monat; MW: Mittelwert; N: Anzahl ausgewerteter Patienten; n. g.: nicht genannt; SD: Standardabweichung; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index										

Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 3 Monate
Globale Bewertung
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - primäre Scores, 12 Monate
Globale Bewertung
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Zeitpunkte: 3, 12 Monate. Primäre Verwendung folgender Skalen: WOMAC Global (Kalunian 2000; Kirkley 2008).

KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 8: Meta-Analyse zur globalen Bewertung der Symptomatik der Gonarthrose; Vergleich keine aktive Vergleichsintervention

Sensitivitätsanalysen

Die Studie mit niedrigem Verzerrungspotenzial (Kalunian 2000) und die Studie mit hohem Verzerrungspotenzial (Kirkley 2008) wiesen heterogene Ergebnisse auf. In keiner der beiden Studien zeigte sich jedoch ein relevanter Effekt.

Insgesamt ergibt sich auf Basis der Sensitivitätsanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Subgruppenanalysen

Von den in Abschnitt 4.4.4 spezifizierten Effektmodifikatoren lagen lediglich Daten für die Effektmodifikatoren „Art des arthroskopischen Eingriffs“ und „Art der Behandlung in der Kontrollgruppe“ für die Zeitpunkte 3 Monate und 12 Monate basierend auf den Studien Kalunian 2000 und Kirkley 2008 vor. Die zugehörigen Interaktionstests sind in Tabelle 26 dargestellt. Eine Studie entspricht jeweils einer Subgruppe bzgl. beider Effektmodifikatoren (d. h. Kalunian 2000 [Lavage, diagnostische Arthroskopie], Kirkley 2008 [Lavage + Debridement, Scheinarthroskopie]).

Zum Zeitpunkt 3 Monate sind die Ergebnisse beider Studien nicht signifikant, zum Zeitpunkt 12 Monate zeigt sich ein signifikanter Effekt bei Kalunian 2000 (für eine grafische Einschätzung siehe Abbildung 8). Da auch hier keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied oder keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Da das

obere Ende des Konfidenzintervalls die Irrelevanzschwelle von $-0,2$ überschritt, konnte ein irrelevanter Effekt nicht sicher ausgeschlossen werden.

Tabelle 26: Subgruppenanalysen globale Bewertung

Effektmodifikator	p-Wert Interaktionstest			
Zeitpunkt in Monaten	3	6	12	> 12
Art des arthroskopischen Eingriffs Lavage vs. Lavage plus Debridement	0,101	-	0,026	-
Art der Behandlung in der Kontrollgruppe keine zusätzliche Behandlung vs. Scheinarthroskopie vs. diagnostische Arthroskopie	0,101	-	0,026	-
-: keine Daten				

In der Studie Kirkley 2008 wurde die Subgruppe der leichteren oder schwereren Arthroseschweregrade untersucht. In dieser Post-hoc-Subgruppenanalyse fanden sich sowohl bei den 78 Patienten mit einem radiologischen Schweregrad 2 nach der Kellgren- und Lawrence-Klassifikation als auch bei den 99 Patienten mit einem Schweregrad von 3 oder 4 keine mittel- oder langfristigen Unterschiede im WOMAC-Gesamtscore. Ohne dass ein statistischer Interaktionstest angegeben wurde, schlussfolgerten die Autoren, dass es keinen Zusammenhang zwischen dem Schweregrad der Erkrankung und dem Effekt einer Arthroskopie gebe. Somit zeigen die verfügbaren Daten keine unterschiedlichen Behandlungseffekte hinsichtlich des potenziellen Effektmodifikators Schweregrad der Arthrose.

Insgesamt ergibt sich auf Basis der Subgruppenanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

5.3.1.2 Gesundheitsbezogene Lebensqualität

Nur in der Studie Kirkley 2008 wurden Instrumente zur Ermittlung der gesundheitsbezogenen Lebensqualität eingesetzt (siehe Tabelle 27).

Tabelle 27: Übersicht der verwendeten Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität, Vergleich keine aktive Vergleichsintervention

Studien	Verwendete Instrumente
Kirkley 2008	SF-36 Standard gamble utility score
SF-36: Short Form (36) Health Survey	

Verzerrungspotenzial der Ergebnisse zur gesundheitsbezogenen Lebensqualität

Das Verzerrungspotenzial des Endpunktes war aufgrund der fehlenden Verblindung des Endpunkterhebers und einer inadäquaten Umsetzung des ITT-Prinzips als hoch einzuschätzen (Tabelle 28).

Ergebnisse zur gesundheitsbezogenen Lebensqualität

In Tabelle 29 sind die Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität dargestellt. Für beide Instrumente war zu keinem Zeitpunkt ein signifikanter Unterschied zwischen der Prüf- und der Vergleichsintervention festzustellen.

Basierend auf dieser Studienlage konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie für den Endpunkt gesundheitsbezogene Lebensqualität abgeleitet werden.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Es wurden keine Subgruppenanalysen durchgeführt.

Tabelle 28: Bewertung des Verzerrungspotenzials auf Endpunktebene: gesundheitsbezogene Lebensqualität, Vergleich keine aktive Vergleichsintervention

Studie	Verzerrungspotenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT-Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungs- potenzial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Kirkley 2008	hoch	nein ^a	nein ^b	ja	ja	hoch ^c
a: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung erfolgte durch unverblindeten Patienten begleitet vom verblindeten Endpunkterheber. b: unklare oder fehlende ITT-Auswertung c: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung						

Tabelle 29: Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität, Vergleich keine aktive Vergleichsintervention

Studie	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p- Wert	Änderung zu Baseline [95 %- KI]; p-Wert
Kirkley 2008	SF-36 körperliche Funktions- fähigkeit (0-100)	0	92	33,8 (7,6)		86	33,9 (8,6)			
		3	90	38,7 (9,0)	n. g.	80	37,7 (10,2)	n. g.	1,0 [-1,9; 3,9]; 0,498 ^a	n. g.
		6	90	38,7 (9,3)	n. g.	73	38,1 (10,2)	n. g.	0,6 [-2,4; 3,6]; 0,695 ^a	n. g.
		12	80	38,3 (10,7)	n. g.	77	37,7 (10,0)	n. g.	0,6 [-2,7; 3,9]; 0,717 ^a	n. g.
		24	88	37,0 (11,4)	n. g.	80	37,2 (10,6)	n. g.	-0,2 [-3,6; 3,2]; 0,93 ^b	n. g.
	Standard gamble utility score	0	92	0,79 (0,23)		86	0,81 (0,20)			
		3	90	0,81 (0,21)	n. g.	80	0,80 (0,22)	n. g.	0,01 [-0,06; 0,08]; 0,762 ^a	n. g.
		6	90	0,84 (0,20)	n. g.	73	0,81 (0,22)	n. g.	0,03 [-0,04; 0,10]; 0,364 ^a	n. g.
		12	80	0,82 (0,21)	n. g.	77	0,86 (0,16)	n. g.	-0,04 [-0,10; 0,02]; 0,183 ^a	n. g.
		24	88	0,87 (0,18)	n. g.	80	0,86 (0,16)	n. g.	0,01 [-0,04; 0,06] ^c ; 0,87 ^b	n. g.
	a: t-Test (eigene Berechnung) b: adjustiert nach Baseline-Wert und radiologischem Schweregrad c: eigene Berechnung KI: Konfidenzintervall; MW: Mittelwert; N: Anzahl ausgewerteter Patienten; n. g.: nicht genannt; SD: standard deviation (Standardabweichung); Short Form (36) Health Survey									

5.3.1.3 Unerwünschte Therapiewirkungen

In den Studien Katz 2013, Kirkley 2008 und Moseley 2002 wurden Angaben zu unerwünschten Therapiewirkungen gemacht.

Verzerrungspotenzial der Ergebnisse zu unerwünschten Therapiewirkungen

In der Studie Kirkley 2008 wurden nur Angaben zur Mortalität gemacht. Hier war das Verzerrungspotenzial auf Endpunktebene als hoch einzuschätzen. Neben einer inadäquaten Umsetzung des ITT-Prinzips, in diesem Fall war das Problem der erhebliche Unterschied bei der Rate der fehlenden Werte zwischen den Behandlungsgruppen, war die verdeckte Gruppenzuteilung nicht eindeutig in der Studie beschrieben.

In den Studien Moseley 2002 und Katz 2013 wurden Angaben zur Mortalität in den Behandlungsgruppen gemacht. Für die Erhebung der Mortalität lag in beiden Studien ein niedriges Verzerrungspotenzial auf Endpunktebene vor. Für die Erhebung weiterer unerwünschter Therapiewirkungen in der Studie Katz 2013 waren eine ergebnisunabhängige Berichterstattung des Endpunkts sowie eine adäquate Verblindung der Endpunkterhebung nicht gewährleistet. Daher wurde für diese Ergebnisse ein hohes Verzerrungspotenzial konstatiert (siehe Tabelle 30).

Ergebnisse zu unerwünschten Therapiewirkungen

Die Angaben zur Mortalität sind in Tabelle 31 dargestellt. Die Angaben der übrigen berichteten unerwünschten Therapiewirkungen waren nicht eindeutig den entsprechenden Behandlungsgruppen zuzuordnen und werden daher im Folgenden lediglich im Text beschrieben.

In der Studie Moseley 2002 wurde für einen Patienten eine Rötung der Inzisionsstelle berichtet, für einen weiteren Patienten eine Wadenschwellung ohne Thrombose. Für beide Patienten konnte aus den Angaben der Publikation nicht nachvollzogen werden, zu welcher Gruppe diese zugeteilt waren.

Insgesamt war die Datenlage zum Endpunkt unerwünschte Therapiewirkungen unzureichend, sodass keine eindeutige Aussage zum Schaden der Interventionen getroffen werden konnte.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Es wurden keine Subgruppenanalysen durchgeführt.

Tabelle 30: Bewertung des Verzerrungspotenzials auf Endpunktebene: unerwünschte Therapiewirkungen, Vergleich keine aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungspotenzial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Katz 2013	niedrig	nein	ja	ja / unklar ^a	ja	niedrig / hoch ^b
Kirkley 2008	hoch	ja	unklar ^c	ja	ja	hoch ^d
Moseley 2002	niedrig	ja	ja	ja	ja	niedrig
a: Für die Erfassung der Mortalität lag kein Problem vor, die Auswahl der übrigen unerwünschten Therapiewirkungen bleibt unklar, daher sind diese Daten als potenziell ergebnisgesteuert anzusehen. b: Für die Daten zur Mortalität lag ein niedriges Verzerrungspotenzial vor, die übrigen Daten zu unerwünschten Therapiewirkungen haben wegen fehlender Verblindung und potenziell ergebnisgesteuerter Berichterstattung ein hohes Verzerrungspotenzial. c: unklare oder fehlende ITT-Auswertung d: hohes Verzerrungspotenzial auf Studienebene, unklare ITT-Auswertung						

Tabelle 31: Ergebnisse zum Endpunkt unerwünschten Therapiewirkungen, Vergleich keine aktive Vergleichsintervention

Studie	Unerwünschte Therapiewirkung	Prüfintervention		Vergleichsintervention		Gruppenunterschied
		N	Anzahl (%) ^a	N	Anzahl (%) ^a	RR [95 %-KI]; p-Wert ^b
Katz 2013	Mortalität ^c	174	1 (0,6)	177	1 (0,6)	Peto OR 1,02 [0,06; 16,33]; > 0,999
	SUE (gesamt):		3 (1,7)		2 (1,1)	1,53 [0,26; 9,02]; 0,668
	▪ Kardiovaskulär		▪ 2 (1,7)		▪ 2 (1,1)	
	▪ Hypoxämie		▪ 1 (0,6)		▪ 0	
	UE (gesamt):		15 (8,6)		13 (7,3)	1,17 [0,58; 2,39]; 0,710
	▪ Muskuloskeletal		▪ 7 (4,0)		▪ 6 (3,4)	
	▪ Kardiovaskulär		▪ 6 (3,5)		▪ 3 (1,7)	
	▪ Schmerzen in Rücken, Hüfte oder Fuß		▪ 2 (1,6)		▪ 4 (2,3)	
	Kniegelenkersatz		5 (2,9)		3 (1,7)	1,70 [0,41; 6,99]; 0,497
Kirkley 2008	Mortalität	94	1 (1,0)	94	0 (0,0)	Peto OR 1,96 [0,39; 9,78]; 0,428 ^b
Moseley 2002	Mortalität	120 ^d	0	60	0	
a: Anzahl Patienten mit Ereignis b: eigene Berechnung, unbedingter exakter Test (CSZ-Methode nach Martin Andres 1994 [51]) c: Verstorbene Patienten sind in den Zahlen der SUE enthalten d: Patienten beider Behandlungsgruppen (Lavage und Lavage + Debridement) zusammengefasst SUE: Schwerwiegende Unerwünschte Ereignisse; UE nicht schwerwiegende unerwünschte Ereignisse						

5.3.2 Arthroskopie versus aktive Behandlung

In 5 Studien wurde die therapeutische Arthroskopie mit einer aktiven Vergleichsintervention verglichen. In der Gruppe der aktiven Vergleichsinterventionen wurden eine Lavage ohne arthroskopischen Eingriff (Chang 1993), die intraartikuläre Injektion von Glukokortikoiden (Arden 2008), die intraartikuläre Injektion von Hyaluronsäure (Forster 2003 und Kang 2005) sowie die orale Gabe von NSAIDs (Merchan 1993) zusammengefasst. Die Studien wurden allerdings aufgrund der unterschiedlichen Wirkmechanismen der Vergleichsinterventionen nicht gemeinsam bewertet. Es erfolgt lediglich eine gemeinsame Darstellung in diesem Kapitel, die Aussagen zu einem Nutzen der therapeutischen Arthroskopie im Vergleich zu der jeweiligen speziellen aktiven Vergleichsintervention basieren nur auf den entsprechenden Studien.

5.3.2.1 Symptomatik der Gonarthrose

Für den Endpunkt Symptomatik der Gonarthrose konnten aus den eingeschlossenen Studien Daten zum Schmerz, zur körperlichen Funktion und zur globalen Bewertung der Symptomatik extrahiert werden. Diese Daten wurden getrennt voneinander dargestellt.

5.3.2.1.1 Schmerz

Aus 3 Studien konnten Daten zum Schmerz extrahiert werden, die verwendeten Instrumente sind in Tabelle 32 dargestellt.

Tabelle 32: Übersicht der verwendeten Instrumente zur Erfassung des Schmerz, Vergleich aktive Vergleichsintervention

Studien	Vergleichsgruppe	Verwendete Instrumente
Arden 2008	Glukokortikoide	WOMAC-Subskala Schmerz
Chang 1993	Lavage	AIMS-Subskala Schmerz schmerzhafte Kniegelenke
Forster 2003	Hyaluronsäure	VAS Schmerz
AIMS: Arthritis impact measurement scales; VAS: visuelle Analogskala; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index		

In der Studie Arden 2008 wurde die WOMAC-Subskala zum Schmerz eingesetzt. In der Studie Chang 1993 wurde neben der AIMS-Subskala zum Schmerz eine Einschätzung des Arztes über die Schmerzempfindlichkeit der Kniegelenke vorgenommen. In der Studie Forster 2003 wurden Daten zum Schmerz anhand einer visuellen Analogskala erhoben.

Verzerrungspotenzial der Ergebnisse zum Schmerz

Das Verzerrungspotenzial auf Endpunktebene war bei allen 3 Studien hoch. In allen Studien war die Verblindung der Endpunkterheber bezüglich der Therapie nicht gegeben. In der Studie Forster 2003 war das ITT-Prinzip nicht adäquat umgesetzt worden.

Ergebnisse zum Schmerz

In Tabelle 34 und Tabelle 35 sind die Ergebnisse zum Endpunkt Schmerz dargestellt. Die Studien Arden 2008, Chang 1993 und Forster 2003 berichteten die Veränderungen zu Baseline für unterschiedliche Zeitpunkte (Tabelle 34). In der Studie Chang 1993 erfolgte auch eine zusätzliche Responderanalyse, die ebenfalls dargestellt wird (Tabelle 35).

Vergleich Arthroskopie versus Glukokortikoide

In der Studie Arden 2008 wurde die Subskala Schmerz des WOMAC erhoben. Hier zeigte sich zu den Zeitpunkten 3 und 6 Monate jeweils ein signifikanter Unterschied zwischen den Behandlungsgruppen. Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied oder keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Da das 95 %-Konfidenzintervall jeweils nicht vollständig unterhalb der Irrelevanzschwelle von $-0,2$ lag, konnte ein irrelevanter Effekt somit nicht sicher ausgeschlossen werden. Da diese Studie zudem eine mäßige qualitative Ergebnissicherheit hatte, konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie für diesen Endpunkt abgeleitet werden.

Vergleich Arthroskopie versus Lavage

Für den Vergleich therapeutische Arthroskopie (mit Lavage und Debridement) versus Lavage lagen Daten aus der Studie Chang 1993 vor. Für diesen Endpunkt wurde die Subskala Schmerz des AIMS erfasst. In der Studie wurden zu den Zeitpunkten 3 und 12 Monate Daten erhoben, es konnte aber zu keinem Zeitpunkt ein statistisch signifikanter Unterschied zwischen der Prüf- und der Vergleichsintervention festgestellt werden. Zusätzlich wurde die Bewertung der schmerzhaften Kniegelenke, eingeschätzt durch den Arzt, in einer Responderanalyse untersucht. Zu keinem Zeitpunkt zeigte sich ein signifikanter Unterschied zwischen den Behandlungsgruppen.

Somit konnte für den Endpunkt Schmerz kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Lavage festgestellt werden.

Vergleich Arthroskopie versus Hyaluronsäure

Für den Vergleich therapeutische Arthroskopie versus Hyaluronsäure konnten Daten aus der Studie Forster 2003 verwendet werden. Zu keinem Zeitpunkt (3, 6 und 12 Monate) lag ein signifikanter Unterschied zwischen der therapeutischen Arthroskopie- und der Hyaluronsäure-Gruppe vor.

Es konnte kein Anhaltspunkt für; Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Hyaluronsäure für den Endpunkt Schmerz festgestellt werden.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Für die in Abschnitt 4.4.4 beschriebenen Effektmodifikatoren und möglichen Subgruppen lagen keine Daten zur Berechnung von Subgruppenanalysen vor.

In der Studie Arden 2008 (Arthroskopie versus Glukokortikoide) wurden für den Endpunkt Schmerz Subgruppenergebnisse zu den radiologischen Schweregraden 0 bis 1, 2 und 3 bis 4 nach der Kellgren- und Lawrence-Klassifikation zum Zeitpunkt 6 Monate berichtet. Der Interaktionstest (Cochran's Q-Statistik, eigene Berechnung basierend auf Angaben in der Publikation [Tabelle 1, Abbildung 4 und im Text in Arden 2008]) ergab einen nicht statistisch signifikanten p-Wert von 0,453. Somit zeigen die verfügbaren Daten keine unterschiedlichen Behandlungseffekte hinsichtlich des potenziellen Effektmodifikators radiologischer Schweregrad.

Insgesamt ergibt sich auf Basis der Subgruppenanalysen kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Tabelle 33: Bewertung des Verzerrungspotenzials auf Endpunktebene: Schmerz, Vergleich aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungs- potenzial beeinflussen	Verzerrungspotenzial für diesen Endpunkt
Arden 2008	hoch	nein ^a	ja	ja	ja	hoch ^b
Chang 1993	hoch	nein ^a	ja	ja	ja	hoch ^b
Forster 2003	hoch	nein	nein ^c	ja	ja	hoch ^d
a: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung der Schmerzen erfolgte durch den bezüglich der Therapie unverblindeten Patienten begleitet durch verblindeten Behandler. b: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes c: Eine unklare Anzahl von Patienten wurde nicht in der Auswertung berücksichtigt. d: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung						

Tabelle 34: Ergebnisse zum Endpunkt Schmerz, Vergleich aktive Vergleichsintervention

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline (SD)	N	MW (SD) [95 %-KI]	Änderung zu Baseline (SD)	Messwert [95 %- KI]; p-Wert	Änderung zu Baseline [95 %- KI]; p-Wert
Arden 2008 (Gluko- kortikoide)	WOMAC- Schmerz (0-500)	0	71	254 (88)		79	247 (97)			
		3	68 ^a	n. g.	-79 (106)	73 ^a	n. g.	-44 (96)	n. g.	-0,35 [-0,67; -0,02] ^b ; < 0,05
		6	65 ^a	n. g.	-75 (114)	71 ^a	n. g.	-19 (99)	n. g.	-0,52 [-0,85; -0,198] ^b ; < 0,01
Chang 1993 (Lavage)	AIMS- Schmerz (0-10)	0	18	6,5 (2,0)		14	6,1 (2,1)			
		3	18	5,0 (n. g.)	n. g.	14	5,4 (n. g.)	n. g.	-0,4 [-1,6; 0,9]; > 0,05 ^c	n. g.
		12	18	5,3 (n. g.)	n. g.	14	5,0 (n. g.)	n. g.	0,3 [-1,1; 1,8]; > 0,05 ^c	n. g.
Forster 2003 (Hyaluron- säure)	VAS Schmerz (0-10)	0	19	7,5 (n. g.)		19	7,6 (n. g.)			
		3	n. g.	6,0 (n. g.)	n. g.	n. g.	6,0 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
		6	n. g.	6,2 (n. g.)	n. g.	n. g.	5,4 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
		12	n. g.	5,7 (n. g.)	n. g.	n. g.	5,7 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
a: Die Nummern beziehen sich auf die Anzahl der Patienten mit verfügbaren Daten. Da die Autoren die sogenannte „last observation carried forward“ (LOCF) angewandt haben, ist die Analyse nach dem „Intention-to-treat“-Prinzip verlaufen. b: Hedges’ g (eigene Berechnung) c: Kovarianzanalyse, adjustiert nach Baseline-Werten AIMS: Arthritis Impact Measurement Scale; MV: Mittelwert; SD: Standardabweichung; VAS: visuelle Analogskala; WOMAC: Western Ontario and McMaster Universities Arthritis Index										

Tabelle 35: Ergebnisse zum Endpunkt Schmerz, Vergleich aktive Vergleichsintervention – Responderanalysen

Studie (Vergleich)	Instrument (Bereich)	Zeitpunkt (Mon.)	Prüfintervention		Vergleichsintervention		Gruppenunterschied Effektschätzer [95 %-KI]; p-Wert
			N	Patienten mit Response n (%)	N	Patienten mit Response n (%)	
Chang 1993 (Lavage)	Schmerzhafte Kniegelenke ^a	3	18	8 (47)	14	5 (38)	RD ^b : 9 [-26; 44]; > 0,05
		12	18	8 (47)	14	4 (31)	RD ^b : 16 [-18; 50]; > 0,05
a: Responder-Definition: % mit Verbesserung der Schmerzempfindlichkeit von ≥ 1 Punkt (auf der 4-Punkte-Skala, Anzahl der Patienten mit Response anhand der Prozentangaben selbst berechnet b: in % RD: Risikodifferenz							

5.3.2.1.2 Körperliche Funktion

In 3 Studien lagen Daten zum Endpunkt körperliche Funktion vor, die verwendeten Instrumente sind in Tabelle 36 dargestellt.

Tabelle 36: Übersicht der verwendeten Instrumente zur Erfassung der körperlichen Funktion, Vergleich aktive Vergleichsintervention

Studien	Vergleichsgruppe	Verwendete Instrumente
Arden 2008	Glukokortikoide	WOMAC-Subskala Funktion WOMAC-Subskala Steifigkeit 50 Meter Gehen 10 Treppenstufen Hinauf- und Hinabsteigen
Chang 1993	Lavage	AIMS-Subskala Funktion AIMS-Subskala Mobilität 15 Meter Gehen
Forster 2003	Hyaluronsäure	KSCRS
AIMS: Arthritis impact measurement scales; KSCRS: Knee Society Clinical Rating System Function Score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index		

Die Studie Arden 2008 verwendete 2 Subskalen des WOMAC, die Subskala zur Funktion und die Subskala zur Steifigkeit. Zusätzlich zu diesem Instrument wurde die Zeit bestimmt, die ein Patient benötigte, um 50 Meter zu gehen oder eine bestimmte Anzahl Treppen zu steigen. In der Studie Chang 1993 war das Vorgehen ähnlich, hier wurde von dem Instrument AIMS jeweils die Subskala Funktion und die Subskala Mobilität erfasst und zusätzlich die Zeit, die die Patienten benötigten, um 15 Meter zu gehen. Die Studie Forster 2003 verwendete das Instrument KSCRS, bei dem neben der Bewertung der Mobilität des Patienten auch klinische Bewertungen des Knies eingehen.

Verzerrungspotenzial der Ergebnisse zum Endpunkt körperliche Funktion

Bei allen 3 Studien fehlte die Verblindung des Endpunkterhebers, in der Studie Forster 2003 war darüber hinaus die adäquate Umsetzung des ITT-Prinzips fraglich. Zusätzlich erhielten in dieser Studie 26 % der Patienten in der Hyalurongruppe eine arthroskopische Intervention (siehe Tabelle 37). Insgesamt lag bei allen Studien ein hohes Verzerrungspotenzial auf Endpunktebene vor.

Ergebnisse zum Endpunkt körperliche Funktion

In Tabelle 38 sind die Ergebnisse zum Endpunkt körperliche Funktion dargestellt. Da jeweils nur eine Studie zu einem spezifischen aktiven Vergleich vorlag, war die Durchführung einer Meta-Analyse nicht möglich.

Vergleich Arthroskopie versus Glukokortikoide

Dieser Vergleich wurde nur in der Studie Arden 2008 untersucht. Hier wurden die Subskalen körperliche Funktion und Steifigkeit des WOMAC sowie die Zeit untersucht, die die

Patienten benötigten, um 50 Meter zu gehen oder 10 Treppenstufen zu steigen. Für beide Subskalen des WOMAC lag zu den Zeitpunkten 3 und 6 Monaten ein statistisch signifikanter Unterschied zwischen der Prüf- und der Vergleichsintervention vor. Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied oder keine Responderanalysen zu einer validierten MID vorlagen, wurde auf die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) für die Relevanzbewertung zurückgegriffen. Da das 95 %-Konfidenzintervall jeweils nicht vollständig oberhalb der Irrelevanzschwelle von 0,2 lag, konnte ein irrelevanter Effekt somit nicht sicher ausgeschlossen werden.

Es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zu Glukokortikoiden bei der körperlichen Funktion abgeleitet werden.

Vergleich Arthroskopie versus Lavage

Für den Vergleich therapeutische Arthroskopie (mit Lavage und Debridement) versus Lavage lagen Daten aus der Studie Chang 1993 vor. Für diesen Endpunkt wurden die Subskalen Funktion und Mobilität des AIMS erfasst. Zusätzlich wurde die Zeit erfasst, die ein Patient benötigt, um 15 Meter zu gehen. In der Studie wurden zu den Zeitpunkten 3 und 12 Monate nach Baseline Daten erhoben, es konnte aber für keines der verwendeten Instrumente zu diesen Zeitpunkten ein statistisch signifikanter Unterschied zwischen der Prüf- und der Vergleichsintervention festgestellt werden.

Für den Endpunkt körperliche Funktion konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Lavage festgestellt werden.

Vergleich Arthroskopie versus Hyaluronsäure

Für den Vergleich therapeutische Arthroskopie versus Hyaluronsäure lagen aus der Studie Forster 2003 Daten des Knee Society Clinical Rating System Function Score (KSCRS) vor. Es zeigte sich zu keinem Zeitpunkt (3, 6 und 12 Monate) ein signifikanter Unterschied zwischen der Arthroskopie- und der Hyaluronsäure-Gruppe.

Es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Hyaluronsäure für den Endpunkt körperliche Funktion festgestellt werden.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Es wurden keine Subgruppenanalysen durchgeführt.

Tabelle 37: Bewertung des Verzerrungspotenzials auf Endpunktebene: körperliche Funktion, Vergleich aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungspotenzial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Arden 2008	hoch	nein ^a	ja	ja	ja	hoch ^b
Chang 1993	hoch	nein ^a	ja	ja	ja	hoch ^b
Forster 2003	hoch	nein	unklar ^c	ja	ja	hoch ^d
a: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung erfolgt durch den bezüglich der Therapie unverblindeten Patienten begleitet von verblindetem Endpunkterheber. b: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes c: Anzahl Patienten in Auswertung unklar, unklarer Umgang mit fehlenden Werten d: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung ITT: intention to treat						

Tabelle 38: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich aktive Vergleichsintervention

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Arden 2008 (Gluko- kortikoide)	WOMAC Funktion (0-1700)	0	71	853 (313)		79	831 (341)			
		3	68 ^a	n. g.	255 (333)	73 ^a	n. g.	149 (297)	n. g.	0,34 [0,01; 0,66] ^b ; < 0,05
		6	65 ^a	n. g.	219 (356)	71 ^a	n. g.	69 (329)	n. g.	0,44 [0,11; 0,76] ^b ; < 0,01
	WOMAC Steifigkeit (0-200)	0	71	123 (38)		79	112 (44)			
		3	68 ^a	n. g.	43 (51)	73 ^a	n. g.	23 (49)	n. g.	0,40 [0,06; 0,73] ^b ; < 0,05
		6	65 ^a	n. g.	34 (58)	71 ^a	n. g.	10 (49)	n. g.	0,45 [0,11; 0,79] ^b ; < 0,01
	50 Meter gehen	0	71	50 (45; 61) ^c		79	50 (45; 61) ^c			
		3	68 ^a	n. g.	2 (-2; 7) ^c	73 ^a	n. g.	0 (-5; 5) ^c	n. g.	n. g.
		6	65 ^a	n. g.	2 (-3; 8) ^c	71 ^a	n. g.	0 (-5; 5) ^c	n. g.	n. g.
	Treppen steigen	0	71	9 (7; 12) ^c		79	9 (7; 13) ^c			
		3	68 ^a	n. g.	1 (-1; 2) ^c	73 ^a	n. g.	0 (-1; 2) ^c	n. g.	n. g.
		6	65 ^a	n. g.	1 (-1; 2) ^c	71 ^a	n. g.	0 (-1; 2) ^c	n. g.	n. g.

(Fortsetzung)

Tabelle 38: Ergebnisse zum Endpunkt körperliche Funktion, Vergleich aktive Vergleichsintervention (Fortsetzung)

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)		
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %-KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert	
Chang 1993 (Lavage)	AIMS Funktion	0	18	2,3 (1,6)		14	1,7 (1,0)				
		3	18	1,5 (n. g.)	n. g.	14	2,0 (n. g.)	n. g.	-0,5 [-1,2; 0,3]; > 0,05 ^d	n. g.	
		12	18	1,7 (n. g.)	n. g.	14	2,0 (n. g.)	n. g.	0,3 [-1,1; 0,5]; > 0,05 ^d	n. g.	
	AIMS Mobilität	0	18	6,9 (2,0)		14	5,3 (2,1)				
		3	18	5,0 (n. g.)	n. g.	14	6,3 (n. g.)	n. g.	-1,3 [-3,0; 0,4]; > 0,05 ^d	n. g.	
		12	18	4,8 (n. g.)	n. g.	14	6,2 (n. g.)	n. g.	-1,4 [-3,3; 0,4]; > 0,05 ^d	n. g.	
	15 Meter gehen	0	18	14,9 (4,3)		14	15,0 (4,7)				
		3	18	14,2 (n. g.)	n. g.	14	15,0 (n. g.)	n. g.	-0,8 [-2,8; 1,2]; > 0,05 ^d	n. g.	
		12	18	13,9 (n. g.)	n. g.	14	14,1 (n. g.)	n. g.	-0,2 [-2,8; 2,3]; > 0,05 ^d	n. g.	
	Forster 2003 (Hyaluron- säure)	KSCRS (0-100)	0	19	45 (n. g.)		19	65 (n. g.)			
			3	n. g.	45 (n. g.)	n. g.	n. g.	65 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
			6	n. g.	45 (n. g.)	n. g.	n. g.	80 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
		12	n. g.	55 (n. g.)	n. g.	n. g.	90 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.	
a: Die Nummern beziehen sich auf die Anzahl der Patienten mit verfügbaren Daten, fehlende Daten wurden mittels „last observation carried forward“ (LOCF) ersetzt. b: Hedges’ g (eigene Berechnung) c: Median (Interquartilsabstand) d: Kovarianzanalyse, adjustiert nach Baseline-Werten AIMS: Arthritis Impact Measurement Scale; KI: Konfidenzintervall; KSCRS: Knee Society Clinical Rating System Function Score, MW: Mittelwert; SD: Standardabweichung; WOMAC: Western Ontario and McMaster Universities Arthritis Index											

5.3.2.1.3 Globale Bewertung der Symptomatik

Alle 5 Studien stellten Daten zur globalen Bewertung der Symptomatik der Gonarthrose dar. Die eingesetzten Instrumente unterschieden sich allerdings in allen Studien (siehe Tabelle 39).

Tabelle 39: Übersicht der verwendeten Instrumente zur Erfassung der globalen Bewertung, Vergleich aktive Vergleichsintervention

Studien	Vergleichsgruppe	Verwendete Instrumente
Arden 2008	Glukokortikoide	Verbesserung der Kniesymptome (Einschätzung Patient)
Chang 1993	Lavage	AIMS Subskala (VAS) allgemeines Wohlbefinden
Forster 2003	Hyaluronsäure	Lequesne-Index
Kang 2005	Hyaluronsäure	Lysholm Score
Merchan 1993	oral NSAID	Hospital for Special Surgery Knee Rating Score
AIMS: arthritis impact measurement scale; NSAID: non-steroidal anti-inflammatory drugs; VAS: visuelle Analogskala		

In der Studie Arden 2008 wurde die Verbesserung der Kniesymptome durch den Patienten eingeschätzt. Die Studie Chang 1993 erhob die globale Bewertung der Symptomatik anhand der AIMS-Subskala allgemeines Wohlbefinden, die mit einer visuellen Analogskala gemessen wurde. In der Studie Forster 2003 wurde der Lequesne Index eingesetzt, in der Studie Kang 2005 der Lysholm Score. In der Studie Merchan 1993 wurde der Hospital for Special Surgery Knee Rating Score verwendet.

Verzerrungspotenzial der Ergebnisse zur globalen Bewertung der Symptomatik der Gonarthrose

Das Verzerrungspotenzial auf Endpunktebene war bei allen 5 Studien hoch. In 4 Studien war die Verblindung der Endpunkterheber nicht gegeben, bei einer Studie war die Einschätzung nicht möglich (Merchan 1993). Die adäquate Umsetzung des ITT-Prinzips war nur in 2 Studien (Arden 2008 und Chang 1993) feststellbar, bei den übrigen 3 Studien blieb dies aus den Publikationen unklar. Bei der Studie Kang 2005 wurden bei 29 % aller Patienten Daten für beide Knie ausgewertet, ohne dies in der Auswertung adäquat zu berücksichtigen (siehe Tabelle 40).

Ergebnisse zur globalen Bewertung der Symptomatik der Gonarthrose

In Tabelle 41 und Tabelle 42 sind die Ergebnisse zum Endpunkt globale Bewertung der Symptomatik der Gonarthrose dargestellt. Die Studien Chang 1993, Forster 2003, Kang 2005 und Merchan 1993 berichteten die Veränderungen zu Baseline für unterschiedliche Zeitpunkte (Tabelle 41). Die Studie Arden 2008 berichtete eine Responderanalyse zu diesem Endpunkt (Tabelle 42).

Vergleich Arthroskopie versus Glukokortikoide

Dieser Vergleich wurde nur in der Studie Arden 2008 untersucht. Hier wurde die Verbesserung der Kniesymptome, eingeschätzt vom Patienten, in einer Responderanalyse dargestellt. Die Beurteilung der Verbesserung (Response) erfolgte durch den Patienten basierend auf einer 5-stufigen Likert-Skala, die von „viel schlechter“ bis „viel besser“ reichte (siehe hierzu auch Anhang E). Zu allen angegebenen Zeitpunkten (3 und 6 Monate) lag ein signifikanter Effekt zugunsten der therapeutischen Arthroskopie vor. Das OR hatte einen Wert von 3,49 (95 %-KI: [1,73; 7,07]) zum Zeitpunkt 3 Monate und von 4,07 (95 %-KI: [1,99; 8,33]) zum Zeitpunkt 6 Monate.

Insgesamt ergibt sich ein Anhaltspunkt für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur intraartikulären Injektion von Glukokortikoiden für den Endpunkt globale Bewertung der Symptomatik der Gonarthrose.

Vergleich Arthroskopie versus Lavage

Für den Vergleich therapeutische Arthroskopie (mit Lavage und Debridement) versus Lavage lagen Daten aus der Studie Chang 1993 vor. Für diesen Endpunkt wurde die Subskala globales Wohlbefinden des AIMS (gemessen anhand einer visuellen Analogskala) erfasst. In der Studie wurden zu den Zeitpunkten 3 und 12 Monate Daten erhoben, es konnte aber zu keinem Zeitpunkt ein statistisch signifikanter Unterschied zwischen der Prüf- und der Vergleichsintervention festgestellt werden.

Für die globale Bewertung der Symptomatik der Gonarthrose konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Lavage festgestellt werden.

Vergleich Arthroskopie versus Hyaluronsäure

Der Vergleich therapeutische Arthroskopie versus Hyaluronsäure wurde in 2 Studien untersucht. In der Studie Forster 2003 wurde für die globale Bewertung der Symptomatik der Gonarthrose der Lequesne-Index, in der Studie Kang 2005 der Lysholm Score erhoben. Bei der Studie Forster 2003 lag zu keinem Zeitpunkt (3, 6 und 12 Monate) ein signifikanter Unterschied zwischen der Arthroskopie- und der Hyaluronsäure-Gruppe vor. Die Studie Kang 2005 zeigte einen signifikanten Effekt zum Zeitpunkt 12 Monate. Da eine Berechnung der standardisierten Mittelwertsdifferenz aufgrund unklarer bzw. fehlender Angaben nicht möglich war, konnte eine Relevanz-Bewertung dieses signifikanten Effektes nicht durchgeführt werden.

Es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur Hyaluronsäure für diesen Endpunkt festgestellt werden.

Vergleich Arthroskopie versus orale Gabe von NSAID

Eine Studie (Merchan 1993) lieferte Daten zum Endpunkt globale Bewertung der Symptomatik der Gonarthrose gemessen anhand des Hospital for Special Surgery Knee Rating Scores. Die Nachbeobachtungszeit betrug zwischen 12 und 36 Monaten. Die gemessenen Daten beziehen sich auf die finale Erhebung des Endpunktes (LOCF-Auswertung zum Zeitpunkt 36 Monate). Es lag ein signifikanter Gruppenunterschied vor.

Da keine skalenspezifisch validierten bzw. etablierten Relevanzkriterien für den Gruppenunterschied und keine Responderanalysen zu einer validierten MID vorlagen, und zudem die Betrachtung der standardisierten Mittelwertdifferenz (SMD in Form von Hedges' g) aufgrund fehlender Streuungsangaben nicht möglich war, konnte die Relevanz des beobachteten Behandlungsunterschieds nicht beurteilt werden, sodass ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte.

Somit konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie im Vergleich zur oralen Gabe von NSAIDs für diesen Endpunkt festgestellt werden.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Es wurden keine Subgruppenanalysen durchgeführt.

Tabelle 40: Bewertung des Verzerrungspotenzials auf Endpunktebene: globale Bewertung, Vergleich aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnisunab- hängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungspotenzial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Arden 2008	hoch	nein ^a	ja	ja	ja	hoch ^b
Chang 1993	hoch	nein ^a	ja ^c	ja	ja	hoch ^b
Forster 2003	hoch	nein	unklar ^d	ja	ja	hoch ^e
Kang 2005	hoch	nein	unklar ^f	ja	nein ^g	hoch ^h
Merchan 1993	hoch	unklar ⁱ	unklar ^f	ja	ja	hoch ^j

a: Endpunkterheber zwar als verblindet beschrieben, aber Endpunktmessung erfolgt durch unverblindeten Patienten begleitet vom verblindeten Endpunkterheber.
b: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes
c: Je 1 Patient pro Behandlungsgruppe schied nach Randomisierung vor Behandlungsbeginn aus der Studie aus.
d: Anzahl Patienten in Auswertung unklar, unklarer Umgang mit fehlenden Werten
e: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung
f: keine Angaben zu fehlenden Werten
g: Bei 29 % aller Patienten werden Daten beider Knie ausgewertet ohne dies in der Auswertung adäquat zu berücksichtigen.
h: hohes Verzerrungspotenzial auf Studienebene, unverblindete Erhebung eines subjektiven Endpunktes, unklare ITT-Auswertung, abhängige Messungen
i: keine Angaben zur Verblindung der Endpunkterheber
j: hohes Verzerrungspotenzial auf Studienebene, unklare Verblindung, unklare ITT-Auswertung
ITT. intention to treat

Tabelle 41: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich aktive Vergleichsintervention

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention			Vergleichsintervention			Gruppenunterschied (Mittelwertsdifferenz)	
			N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	N	MW (SD) [95 %-KI]	Änderung zu Baseline [95 %-KI]	Messwert [95 %- KI]; p-Wert	Änderung zu Baseline [95 %-KI]; p- Wert
Chang 1993 (Lavage)	AIMS globales Wohlbe- finden (0-10)	0	18	4,6 (2,6)		14	4,6 (2,5)			
		3	18	3,4 (n. g.)	n. g.	14	3,6 (n. g.)	n. g.	-0,2 [-10,6; 13,8]; > 0,05 ^a	n. g.
		12	18	4,1 (n. g.)	n. g.	14	3,3 (n. g.)	n. g.	0,8 [-5,3; 21,2]; > 0,05 ^a	n. g.
Forster 2003 (Hyaluron- säure)	Lequesne- Index (0-24)	0	19	13 (n. g.)		19	10,5 (n. g.)			
		3	n. g.	13 (n. g.)	n. g.	n. g.	11 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
		6	n. g.	12 (n. g.)	n. g.	n. g.	9 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
		12	n. g.	10,5 (n. g.)	n. g.	n. g.	8 (n. g.)	n. g.	n. g. (n. g.); > 0,05	n. g.
Kang 2005 (Hyaluron- säure)	Lysholm Score	0	41 ^b	46,0 (1,9) ^c		48 ^d	46,4 (1,9) ^c			
		12	n. g.	81,5 (1,7) ^c	-32,4 (0,57)	n. g.	75,7 (1,6) ^c	-25,9 (0,35)	n. g.	n. g. (n. g.); < 0,05
Merchan 1993 (oral NSAID)	Hospital for Special Surgery Knee Rating Score	0	35	26,9 (n. g.)		38	29,9 (n. g.)			
		36 ^e	35	37,0 (n. g.)	10,1 (n. g.)	38	32,8 (n. g.)	2,9 (n. g.)	n. g. (n. g.); 0,022	n. g. (n. g.); 0,001
<i>kursiv:</i> aus Grafik abgelesen a: Kovarianzanalyse, adjustiert nach Baseline-Werten b: N = 41 Knie bei 32 Patienten c: Standardfehler d: N = 48 Knie bei 37 Patienten e: LOCF-Auswertung, mittlere Follow-up Dauer 23-24 Monate (Range 12-36) AIMS: Arthritis Impact Measurement Scale; KI: Konfidenzintervall; MW: Mittelwert; SD: Standardabweichung										

Tabelle 42: Ergebnisse zum Endpunkt globale Bewertung der Symptomatik, Vergleich aktive Vergleichsintervention – Responderanalysen

Studie (Vergleich)	Instrument (Bereich)	Zeit- punkt (Mon.)	Prüfintervention		Vergleichsintervention		Gruppenunterschied Effektschätzer [95 %-KI]; p-Wert
			N	Patienten mit Response n (%)	N	Patienten mit Response n (%)	
Arden 2008 (Gluko- kortikoide)	Kniesymptome ^a	3	68	49 (72)	73	31 (43)	OR: 3,49 [1,73; 7,07] ^b ; < 0,01
		6	65	41 (63)	71	21 (30)	OR: 4,07 [1,99; 8,33] ^b ; < 0,001
a: Prozentangaben aus Grafik abgelesen, eigene Berechnung der jeweiligen Anzahl aus Prozentangaben, Responder-Definition: Verbesserung der Symptome b: eigene Berechnung AIMS: Arthritis Impact Measurement Scale; KI: Konfidenzintervall; OR: Odds Ratio; RD: Risikodifferenz							

5.3.2.2 Gesundheitsbezogene Lebensqualität

Es lagen aus keiner Studie Daten zur gesundheitsbezogenen Lebensqualität vor.

5.3.2.3 Unerwünschte Therapiewirkung

Zum Endpunkt unerwünschte Therapiewirkungen lagen aus der Studie Arden 2008 und der Studie Merchan 1993 Daten vor.

Verzerrungspotenzial der Ergebnisse unerwünschte Therapiewirkungen

Das endpunktbezogene Verzerrungspotenzial war für beide Studien als hoch einzustufen (siehe Tabelle 43). Bei Arden 2008 war unklar, ob die Berichterstattung unabhängig vom Ergebnis geschah, da der Endpunkt nicht vorab festgelegt wurde. Bei Merchan 1993 war die verblindete Erhebung des Endpunkts unklar. Die Erhebung der Mortalitätsdaten wurde in der Studie vorab festgelegt, zu den weiteren berichteten unerwünschten Therapiewirkungen wurde diese Information nicht in der Publikation dargestellt. Eine ergebnisgesteuerte Berichterstattung konnte daher nicht ausgeschlossen werden.

Ergebnisse zu unerwünschten Therapiewirkungen

In Tabelle 44 sind die Ergebnisse zum Endpunkt unerwünschte Therapiewirkungen dargestellt.

Vergleich Arthroskopie versus Glukokortikoide

In der Studie Arden 2008 lag bezüglich vermehrter Knieschwellungen kein Unterschied zwischen den beiden Gruppen vor. Zudem wurde in der Studie berichtet, dass bei keinem Patienten eine septische Arthritis, eine Wundinfektion oder Fisteln festgestellt wurden.

Auf Grundlage der vorliegenden Daten konnte keine eindeutige Aussage zum Schaden der Interventionen getroffen werden.

Vergleich Arthroskopie versus orale Gabe von NSAID

Die Raten unerwünschter Therapiewirkungen (Mortalität, tiefe Venenthrombose, Hämarthros) ließen in der Studie Merchan 1993 keine Unterschiede zwischen den Gruppen erkennen.

Auf Grundlage der vorliegenden Daten konnte keine eindeutige Aussage zum Schaden der Interventionen getroffen werden.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Subgruppenanalysen

Es wurden keine Subgruppenanalysen durchgeführt.

Tabelle 43: Bewertung des Verzerrungspotenzials auf Endpunktebene: unerwünschte Therapiewirkungen, Vergleich aktive Vergleichsintervention

Studie	Verzerrungs- potenzial Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung ITT- Prinzip	Ergebnis- unabhängige Berichterstattung	Fehlen sonstiger Aspekte, die das Verzerrungspoten- zial beeinflussen	Verzerrungs- potenzial für diesen Endpunkt
Arden 2008	hoch	ja	ja	unklar ^a	ja	hoch ^b
Merchan 1993	hoch	unklar ^c	unklar ^d	unklar ^a	ja	hoch ^e
a: Endpunkt nicht vorab definiert, potenziell ergebnisgesteuert b: hohes Verzerrungspotenzial auf Studienebene, potenziell ergebnisgesteuerte Berichterstattung c: keine Angaben zur Verblindung d: unklare oder fehlende ITT-Auswertung e: hohes Verzerrungspotenzial auf Studienebene, unklare Verblindung, unklare ITT-Auswertung						

Tabelle 44: Ergebnisse zum Endpunkt unerwünschte Therapiewirkungen, Vergleich aktive Vergleichsintervention

Studie	Unerwünschte Therapiewirkung	Prüfintervention		Vergleichsintervention		Gruppenunterschied
		N	Anzahl (%) ^a	N	Anzahl (%) ^a	
Arden 2008 (Glukokortikoide)	vermehrte Knieschwellung	71	4 (5,6)	79	4 (5,1)	RR 1,11 [0,29; 4,28]; 0,917
Merchan 1993 (oral NSAID)	Mortalität	40	5 (12,5)	40	2 (5,0)	RR 2,50 [0,51; 12,14]; 0,271
	tiefe Venenthrombose	40	2 (5,0)	40	0 (0,0)	Peto OR 7,58 [0,47; 123,37]; 0,156
	Hämarthros	40	1 (2,5)	40	0 (0,0)	Peto OR 7,39 [0,15; 372,38]; 0,528
	oberflächliche Infektion	40	1 (2,5)	40	0 (0,0)	Peto OR 7,39 [0,15; 372,38]; 0,528
a: Anzahl Patienten mit Ereignis b: eigene Berechnung, unbedingter exakter Test (CSZ-Methode nach Martin Andres 1994 [51]) NSAID: non-steroidal anti-inflammatory drugs; OR: Odds Ratio; RR: relatives Risiko						

5.3.3 Zusammenfassung der Beleglage

Die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement wurde insgesamt in 10 Studien untersucht. In 5 dieser Studien wurde als Vergleich keine aktive Vergleichsintervention (Scheinarthroskopie, diagnostische Arthroskopie, keine zusätzliche Behandlung) eingesetzt, in 5 Studien eine aktive Vergleichsintervention (intraartikuläre Injektion von Glukokortikoiden oder Hyaluronsäure, Lavage oder orale Gabe von NSAIDs). Im Folgenden werden zunächst die Ergebnisse der Studien, in denen keine aktive Vergleichsintervention eingesetzt wurde, gemeinsam dargestellt. Danach erfolgt für jede der aktiven Vergleichsinterventionen eine separate Darstellung der Ergebnisse. Die Ergebnisse sind in der Landkarte der Beleglage (Tabelle 45) zusammenfassend grafisch dargestellt.

Arthroskopische Eingriffe im Vergleich zu keiner aktiven Vergleichsintervention

Es lagen 5 Studien vor, die die therapeutische Arthroskopie mit einer Scheinarthroskopie, einer diagnostischen Arthroskopie oder keiner zusätzlichen Behandlung verglichen. Die Ergebnisse dieser Studien wurden gemeinsam betrachtet. Es muss allerdings darauf hingewiesen werden, dass eine Studie lediglich eine kleine Pilotstudie war, deren Ergebnisse in dieser gemeinsamen Betrachtung kein Gewicht hatte.

Für die Endpunkte Schmerz und körperliche Funktion lagen jeweils Daten aus allen Studien vor, allerdings zeigten sich häufig heterogene Ergebnisse. Zusammenfassend zeigte sich kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie.

Für die globale Bewertung der Symptomatik der Gonarthrose lagen nur Daten aus 2 Studien vor. Die Ergebnisse waren teilweise heterogen und es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Für die gesundheitsbezogene Lebensqualität lagen nur Daten aus einer Studie vor, ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie konnte nicht abgeleitet werden.

Für den Endpunkt unerwünschte Therapiewirkungen war die Datenlage insgesamt unzureichend, sodass keine eindeutige Aussage zum möglichen Schaden der therapeutischen Arthroskopie gemacht werden konnte.

Insgesamt konnte für die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen im Vergleich zu keiner aktiven Vergleichsintervention festgestellt werden.

Arthroskopische Eingriffe im Vergleich zu einer aktiven Vergleichsintervention

Insgesamt lagen 5 Studien vor, die die therapeutische Arthroskopie mit einer aktiven Intervention verglichen. Es lag jeweils eine Studie für den Vergleich therapeutische Arthroskopie versus intraartikuläre Injektion von Glukokortikoiden, Lavage oder orale Gabe

von NSAIDs vor. 2 Studien verglichen die therapeutische Arthroskopie mit der intraartikulären Injektion von Hyaluronsäure.

Arthroskopische Eingriffe versus Glukokortikoide

In der einzigen Studie, die die therapeutische Arthroskopie mit der intraartikulären Injektion von Glukokortikoiden verglich, lagen für die Endpunkte Schmerz und körperliche Funktion jeweils statistisch signifikante Effekte zugunsten der therapeutischen Arthroskopie vor. Die Konfidenzintervalle der Schätzer lagen jedoch jeweils nicht vollständig oberhalb der Irrelevanzschwelle, womit ein irrelevanter Effekt nicht sicher ausgeschlossen werden konnte. Es konnte daher auch kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie für diese Endpunkte abgeleitet werden.

Für den Endpunkt globale Bewertung der Symptomatik der Gonarthrose wurde in der eingeschlossenen Studie die Verbesserung der Kniesymptome basierend auf der Einschätzung der Patienten dargestellt. Für den Nachbeobachtungszeitpunkt 3 Monate lag das OR zugunsten der therapeutischen Arthroskopie bei 3,49 (95 %-KI: [1,73; 7,07]) und für den Zeitpunkt 6 Monate bei 4,07 (95 %-KI: [1,99; 8,33]). Aufgrund des hohen Verzerrungspotenzials auf Studien- und Endpunktebene konnte ein Anhaltspunkt jedoch kein Hinweis oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden.

Die Datenlage zum Endpunkt unerwünschte Therapiewirkungen war in der Studie unzureichend, sodass ein Schaden der therapeutischen Arthroskopie oder der intraartikulären Injektion von Glukokortikoiden nicht eingeschätzt werden konnte.

Für den Endpunkt gesundheitsbezogene Lebensqualität lagen keine Daten vor.

Arthroskopische Eingriffe versus Lavage

Für diesen Vergleich konnte eine Studie eingeschlossen werden, die Angaben zu den Endpunkten Schmerz, körperliche Funktion und der globalen Bewertung der Symptomatik der Gonarthrose machte. Für keinen Endpunkt konnte ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden. Für die Endpunkte gesundheitsbezogene Lebensqualität und unerwünschte Therapiewirkungen lagen aus der Studie keine Angaben vor.

Arthroskopische Eingriffe versus Hyaluronsäure

Der Vergleich therapeutische Arthroskopie versus Hyaluronsäure wurde in 2 Studien untersucht, allerdings lagen nur für den Endpunkt globale Bewertung der Symptomatik der Gonarthrose zum Zeitpunkt 12 Monate Daten aus beiden Studien vor. Für diesen Endpunkt waren die Ergebnisse beider Studien heterogen, es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Die Endpunkte Schmerz und körperliche Funktion wurden nur in einer der beiden Studien untersucht. Hier zeigte sich kein signifikanter Effekt zugunsten der therapeutischen

Arthroskopie, somit konnte für diese Endpunkte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie abgeleitet werden.

Für die Endpunkte gesundheitsbezogene Lebensqualität und unerwünschte Therapiewirkungen lagen keine Daten vor.

Arthroskopische Eingriffe versus orale Gabe von NSAIDs

Für den Vergleich der therapeutischen Arthroskopie mit der oralen Gabe von NSAIDs lagen Daten aus einer Studie vor, die Angaben zu dem Endpunkt globale Bewertung der Symptomatik der Gonarthrose zu dem Zeitpunkt 36 Monate machte. Es konnte kein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie festgestellt werden. Für den Endpunkt unerwünschte Therapiewirkungen lagen unzureichende Angaben vor, sodass eine verlässliche Aussage über einen möglichen Schaden der therapeutischen Arthroskopie nicht gemacht werden konnte.

Für den Endpunkt gesundheitsbezogene Lebensqualität lagen keine Daten vor.

Tabelle 45: Landkarte der Beleglage in Bezug auf die patientenrelevanten Endpunkte

Vergleich Zeitpunkt (Monate)	Symptomatik der Gonarthrose			gesundheits- bezogene Lebensqualität	unerwünschte Therapiewirkungen
	Schmerz	körperliche Funktion	globale Bewertung		
Keine aktive Vergleichsintervention					
3	↑↓	↑↓	↑↓	(⇔)	
6	↑↓	↑↓ ^a	(⇔)	(⇔)	
12	↑↓	↑↓ ^a	↑↓	(⇔)	
24	⇔	⇔	(⇔)	(⇔)	
Gesamt	↑↓	↑↓	↑↓	(⇔)	(⇔)
Glukokortikoide					
3	(⇔) ^b	(⇔) ^b	↗	-	
6	(⇔) ^b	(⇔) ^b	↗	-	
12	-	-	-	-	
24	-	-	-	-	
Gesamt	(⇔)	(⇔)	↗		-
Lavage					
3	(⇔)	(⇔)	(⇔)	-	
6	-	-	-	-	
12	(⇔)	(⇔)	(⇔)	-	
24	-	-	-	-	
Gesamt	(⇔)	(⇔)	(⇔)		-
Hyaluronsäure					
3	(⇔)	(⇔)	(⇔)	-	
6	(⇔)	(⇔)	(⇔)	-	
12	(⇔)	(⇔)	↑↓ ^b	-	
24	-	-	-	-	
Gesamt	(⇔)	(⇔)	↑↓		-
NSAID					
3	-	-	-	-	
6	-	-	-	-	
12	-	-	-	-	
36	-	-	(⇔) ^b	-	
Gesamt			(⇔)		(⇔)
a: Da ein irrelevanter Effekt nicht ausgeschlossen werden konnte, liegt kein Anhaltspunkt vor, obwohl in einer Studie mit einer hohen Ergebnissicherheit ein statistisch signifikantes Ergebnis berichtet wurde.					
b: Da ein irrelevanter Effekt nicht ausgeschlossen werden konnte, liegt hier kein Anhaltspunkt vor.					
-: keine Daten zum Gruppenunterschied berichtet					
↗ Anhaltspunkt für einen (Zusatz-)Nutzen oder Anhaltspunkt für einen geringeren Schaden					
⇔ kein Anhaltspunkt, Hinweis oder Beleg, homogenes Ergebnis, gute Datenlage					
(⇔) kein Anhaltspunkt, Hinweis oder Beleg, homogenes Ergebnis, Datenlage aber unzureichend (z. B. nur eine kleine Studie)					
↑↓ kein Anhaltspunkt, Hinweis oder Beleg, heterogenes Ergebnis					

6 Diskussion

6.1 Ergebnisse

In den vorliegenden Bericht wurden 10 RCTs eingeschlossen. Der Nutzen der therapeutischen Arthroskopie wurde in 5 Studien im Vergleich zu keiner aktiven Vergleichsintervention und in 5 Studien im Vergleich zu anderen aktiven Interventionen untersucht. Es lag nur eine geringe Anzahl RCTs für jede Vergleichskategorie vor, bei denen zudem häufig ein hohes Verzerrungspotenzials festgestellt wurde und eine begrenzte Eignung der Daten für eine Meta-Analyse vorlag. Es ergab sich für keinen der betrachteten Endpunkte und für keinen der betrachteten Zeiträume ein Beleg oder ein Hinweis für einen Nutzen der therapeutischen Arthroskopie bei Gonarthrose.

Für den Vergleich therapeutische Arthroskopie versus intraartikuläre Injektion von Glukokortikoiden gab es einen Anhaltspunkt für einen Nutzen beim Endpunkt globale Bewertung der Symptomatik der Gonarthrose. Zu diesem Vergleich war allerdings nur eine Studie mit einem hohen Verzerrungspotenzial eingeschlossen worden. Der Endpunkt wurde in dieser unverblindeten Studie mit einem Instrument erhoben, das aus einer 5-stufigen Likert-Skala bestand, die von „viel schlechter“ bis „viel besser“ reichte. Da nicht bekannt ist, wie die Zwischenstufen der Skala bezeichnet waren (z.B. „etwas besser“ oder „ein wenig besser“), lässt sich nicht bestimmen, wie patientenrelevant die in der Studie beobachteten Zustandsbesserungen tatsächlich waren, auch wenn im Rahmen des vorliegenden Berichts von einer grundsätzlichen Patientenrelevanz dieser Verbesserungen ausgegangen wurde. Für die anderen Endpunkte (Schmerz, körperliche Funktion und gesundheitsbezogene Lebensqualität) ergab sich aus validierten Instrumenten zwar jeweils ein statistisch signifikantes Ergebnis zugunsten der therapeutischen Arthroskopie. Allerdings konnte für keinen Endpunkt zu keinem Zeitpunkt der Nachbeobachtung ein irrelevanter Effekt ausgeschlossen werden, denn die Konfidenzintervalle der beobachteten Schätzer schlossen einen Irrelevanzbereich immer mit ein. Anhaltspunkte aus weiteren Vergleichen gab es nicht.

Damit ist, unter zusammenfassender Berücksichtigung aktiver und inaktiver Vergleichsinterventionen, der Nutzen der therapeutischen Arthroskopie (mit Lavage und gegebenenfalls zusätzlichem Debridement) bei Gonarthrose insgesamt nicht belegt.

6.2 Vergleich der Studien untereinander

Erwartungsgemäß wurden die beiden Endpunkte Schmerz und körperliche Funktion in den eingeschlossenen Studien ausführlich untersucht. Relativ einheitlich war der Erhebungszeitraum zu diesen Endpunkten, die primär nach 6, 12 und teilweise auch 24 und 36 Monaten erfasst wurden. Weniger einheitlich war dagegen die Wahl der Messinstrumente. So wurden zur Schmerzerfassung hauptsächlich die Instrumente WOMAC, AIMS, KOOS und VAS eingesetzt. Zur Erfassung der körperlichen Funktion kamen vor allem die Instrumente WOMAC und AIMS zur Anwendung. Prinzipiell kann die metaanalytische Zusammenfassung unterschiedlich erfasster Daten problematisch sein, wenn beispielsweise die verschie-

denen Instrumente unterschiedlich stark auf Änderungen im zeitlichen Verlauf reagieren [52]. Da jedoch sowohl die Instrumente zur Schmerzerfassung als auch jene zur Erfassung der körperlichen Funktion untereinander gut korreliert sind [53-57], kann davon ausgegangen werden, dass die Instrumente in ausreichend vergleichbarer Weise den relevanten Endpunkt erfassen. Da auch die Änderungssensitivität der Instrumente grundsätzlich als ähnlich angenommen werden kann [58], erscheint das metaanalytische Zusammenfassen der Studienergebnisse zu diesen Endpunkten gerechtfertigt. Ferner ergaben die im Bericht durchgeführten Sensitivitätsanalysen keine Hinweise darauf, dass die Auswahl des einen statt des anderen Messinstruments einen relevanten Einfluss auf die Ergebnisse haben könnte.

Eher uneinheitlich wurden in den eingeschlossenen Studien die Vergleichsinterventionen ausgewählt (siehe Tabelle 7): Während als Prüfintervention allein die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement untersucht wurde, umfasste das Spektrum der Vergleichsinterventionen aktive und nicht aktive Vergleichsinterventionen, nämlich neben der nicht therapeutischen Arthroskopie auch verschiedene medikamentöse Behandlungsformen. Diese Vielfalt spiegelt wider, dass das Spektrum der Behandlungsoptionen bei Gonarthrose umfangreich ist. Auch wenn die Datengrundlage zur stadienabhängigen Wirksamkeit der diversen Therapieformen unterschiedlich gut ist, ist im Abgleich mit klinischen Leitlinien nicht erkennbar, dass in einer der eingeschlossenen Studien eine inadäquate Form von Vergleichsintervention durchgeführt wurde. Sofern Arzneimittel eingesetzt wurden, erfolgte dies zulassungskonform, auch wenn bei der Studie Arden 2008 die Angaben zum Vorliegen einer aktivierten Arthrose nicht ganz eindeutig waren. Entsprechend der Verschiedenheit der in den Studien untersuchten Vergleiche waren separate Meta-Analysen sinnvoll.

6.3 Unerwünschte Therapiewirkungen

Viele der eingeschlossenen Studien machten keine Angaben dazu, ob unerwünschte Therapiewirkungen aufgetreten waren. Es ist bekannt, dass Studien zu operativen im Vergleich zu pharmakologischen Interventionen weniger häufig und weniger einheitlich über Komplikationen (bzw. unerwünschte Therapiewirkungen) berichten [59,60]. Die Anzahl der in den Studien eingeschlossenen Patienten scheint jedoch auch insgesamt nicht ausreichend, um das Risiko von potenziell schwerwiegenden Komplikationen wie beispielsweise Gelenksinfektion, Hämarthros oder tiefe Venenthrombose bewerten zu können.

Analysen größerer Fallserien haben gezeigt, dass eine Arthroskopie etwa 1 % bis 2 % schwerwiegende Komplikationen (Gelenksinfektion, tiefe Venenthrombose, Nervenschaden etc.) nach sich zieht [61-70]. Auch wenn im vorliegenden Bericht zu unerwünschten Ereignissen kein Beleg, Hinweis oder Anhaltspunkt abgeleitet werden konnte, ist es offensichtlich, dass eine therapeutische Arthroskopie im Vergleich zu keiner Intervention ein höheres Risiko an schwerwiegenden Komplikationen bedingt. Dies gilt jedoch nicht für die aktiven Vergleichsinterventionen, mit denen die therapeutische Arthroskopie verglichen wurde, weil für diese nicht angenommen werden kann, dass sie frei von unerwünschten Ereignissen sind.

6.4 Verzerrungspotenzial

Das Verzerrungspotenzial von 7 der 10 eingeschlossenen Studien war bereits auf Studienebene hoch, denn bei diesen 7 Studien war die Verdeckung der Gruppenzuteilung unklar. Darüber hinaus waren in 6 der 7 Studien die Patienten nicht verblindet. Da auch 1 der 3 Studien mit adäquater Verdeckung der Gruppenzuteilung unverblindet durchgeführt wurde (Katz 2013), ergibt sich auf Endpunktebene eine Gruppe von 8 Studien mit hohem Verzerrungspotenzial. Allein die Studien Moseley 2002 und Kalunian 2000, die beide die Arthroskopie mit keiner aktiven Vergleichsintervention verglichen, hatten ein niedriges Verzerrungspotenzial. Die Patienten waren in beiden Studien verblindet. In der Studie Kalunian 2000 bekam die Kontrollgruppe eine diagnostische Arthroskopie, bei Moseley 2002 wurde in der Kontrollgruppe eine Scheinarthroskopie durchgeführt. Die Autoren der Studie Kalunian 2000 begründeten die Auswahl einer diagnostischen Arthroskopie mit ethischen Beweggründen, da sie es nicht für gerechtfertigt hielten, die Patienten einer Scheinarthroskopie zu unterziehen.

Das Fehlen von Daten im Rahmen der Nachbeobachtung und ein Wechseln der Patienten zwischen den Therapiearmen erschwerten in einzelnen Studien die adäquate Umsetzung des ITT-Prinzips bzw. die unvermengte Bewertung der Prüfintervention. Am deutlichsten trat dieses Problem in den Studien Forster 2003 und Katz 2013 zutage, weil dort innerhalb des ersten Jahres nach Randomisierung 26 % beziehungsweise 33 % der Patienten der Kontrollgruppe eine Arthroskopie erhielten. Eine solche Vermengung zwischen Interventions- und Kontrollgruppe erschwert den Nachweis eines Unterschiedes.

6.5 Studien mit „Placebo-Chirurgie“

Die Studien Moseley 1996 und Moseley 2002 ragen aus allen anderen Studien heraus, weil diese als einzige die therapeutische Arthroskopie gegen eine Scheinarthroskopie untersuchten. Die ältere Publikation beschreibt die Pilotstudie, in der die Durchführbarkeit dieses Designs getestet wurde. In der Studie Moseley 2002 wurden die tatsächlichen Ergebnisse der durchgeführten Studie veröffentlicht. Die Simulation einer arthroskopischen Operation, die eine Inzisionsnarbe am Knie hinterlässt, wurde von keiner weiteren Arbeitsgruppe durchgeführt. Die Studie löste eine Debatte aus, in der verschiedene methodische und auch ethische Argumente gegen die Studie vorgebracht wurden [71-80]:

- Es wurde kritisiert, dass in der Studie Moseley 2002 primär ein nicht validiertes Instrument zur Schmerzmessung verwendet wurde. Hierzu ist anzumerken, dass alle 3 in der Studie verwendeten Instrumente zur Schmerzmessung sehr ähnliche Ergebnisse lieferten. Nach Studiendurchführung wurde auch das von Moseley 2002 selbst entwickelte Instrument validiert [81]. In den vorliegenden Bericht fanden aber ohnehin primär Schmerzdaten Eingang, die mit anderen validierten und etablierten Messinstrumenten erhoben wurden.
- Aufgrund des sehr hohen Anteils männlicher Patienten wurde hinterfragt, ob die Ergebnisse der Studie Moseley 2002 auch auf Frauen mit Gonarthrose übertragbar seien.

Es gibt jedoch keinen Grund dafür anzunehmen, dass sich die Effekte der Arthroskopie relevant zwischen Frauen und Männern unterscheiden [82].

- Ob es statistisch korrekt war, wie Moseley 2002 nach der Überlegenheitsanalyse auch eine Äquivalenzanalyse durchzuführen, wurde ebenfalls bezweifelt [71]. Für den vorliegenden Bericht ergibt sich hieraus kein Problem, da allein die beobachteten Effektschätzer zu den relevanten klinischen Endpunkten im vorliegenden Bericht verwendet wurden.
- Schließlich wurde diskutiert, ob eine placebochirurgische Intervention wie in der Studie Moseley 2002 ethisch vertretbar sei. Dies wurde in einer Begleitpublikation gerechtfertigt [83]. Die American Medical Association (AMA) publizierte im Erscheinungsjahr der Studie von Moseley 2002 speziell für Design und Durchführung chirurgischer placebokontrollierter Studien gedachte ethische Richtlinien [84].

Zusammenfassend lässt sich feststellen, dass die Kritik an der Studie Moseley 2002 keine wesentlichen Schwächen der Studie offenkundig werden ließ. Daher muss dieser Studie auch im Rahmen der hier untersuchten Berichtsfragestellung eine besonders hohe Aussagekraft beigemessen werden. Weltweit führte die Studie Moseley 2002 zu der Erwartung, dass nach der Publikation der Studienergebnisse die Zahl arthroskopischer Eingriffe bei Gonarthrose absinken müsste, was sich jedoch nur teilweise bestätigen ließ [85-90].

6.6 Abgrenzung zu anderen Interventionen und Indikationen

Die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement war die im vorliegenden Bericht zu prüfende Intervention. Abzugrenzen hiervon ist die Lavage des Kniegelenks, die ohne Arthroskopie über 1 oder 2 Kniegelenkspunktionen erfolgt [91,92]. Studien, die die nicht arthroskopische Lavage des Kniegelenks mit anderen nicht arthroskopischen Behandlungsformen verglichen, wurden daher in den vorliegenden Bericht nicht eingeschlossen. Es wurde als sinnvoll erachtet, beide Methoden der Lavage getrennt zu betrachten, weil bei der Lavage mittels Punktion ohne Arthroskopie das Kniegelenk in der Regel mit deutlich weniger Spülflüssigkeit ausgewaschen wird als bei einer arthroskopischen Lavage. Die kleineren Zugänge und das Verwenden nur einer Punktion („tidal irrigation“) limitieren die Möglichkeiten der Spülung erheblich. Insgesamt aber ergab auch das Cochrane Review von Reichenbach 2010, dass die Lavage, ob mit oder ohne Arthroskopie, keinen Nutzen für Patienten mit Arthrose hinsichtlich der Endpunkte Schmerz und körperliche Funktion hat [22]. Somit würde auch ein zusätzlicher Einschluss von Studien zur nicht arthroskopischen Lavage in den vorliegenden Bericht die Schlussfolgerungen vermutlich kaum ändern.

Es ist klinisch oft schwierig, Patienten mit einer symptomatischen Gonarthrose von einem symptomatischen Meniskusriss abzugrenzen, insbesondere weil beide Krankheitsbilder oft auch gemeinsam vorkommen [93,94]. Bei Kirkley 2008 wurden Patienten mit klinischem Verdacht auf einen großen Meniskusriss von der Studienteilnahme ausgeschlossen. Hierfür diente der McMurray-Test, der Innenmeniskusläsionen mit hoher Spezifität erkennen lässt [95]. Da von den Studienteilnehmern bei Kirkley 2008 etwa die Hälfte gelegentliche

Blockierungen im Knie berichteten und von den Patienten in der Interventionsgruppe über 80 % eine Glättung oder Teilresektion des Meniskus erhielten, ist anzunehmen, dass ein relevanter Anteil der Studienpopulation aufgrund seiner Gonarthrose auch degenerative Meniskusschäden aufwies. Es ist allerdings fraglich, ob und in welcher Form diese degenerativen Meniskusschäden zu den Symptomen der Kniearthrose beitragen. In einer Kohortenstudie mit MRI fand sich bei symptomatischen Studienteilnehmern mit radiologisch gesicherter Kniearthrose in 63 % ein Meniskusriss, bei den asymptomatischen Studienteilnehmer in 60 % [94].

Felson 2010 ist der Meinung, dass bei Patienten, bei welchen anamnestisch kein Kniegelenkstrauma eruiert werden kann, eine Arthroskopie vermieden werden sollte, weil eine Gonarthrose sowohl mit als auch ohne degenerativen Meniskusschaden nicht von einer Arthroskopie profitiere [73]. Nicht alle der in den Bericht eingeschlossenen Studien machen Angaben dazu, wie viele der Studienpatienten einen Meniskusschaden aufwiesen. Wesentliche Bedeutung hat daher die Studie Katz 2013, weil dort alle Studienpatienten sowohl eine Gonarthrose als auch eine kernspintomografisch bestätigte degenerative Meniskusläsion aufweisen mussten [44,45]. Eine weitere ähnliche Studie von Herrlin et al. [96,97] konnte nicht in den Bericht eingeschlossen werden, weil dort auch ein Anteil von Patienten ohne Arthrose (Schweregrad 0 nach Ahlbäck [98]) eingeschlossen worden war. Eine weitere Studie zur arthroskopischen Meniskektomie bei degenerativem Meniskusschaden an Patienten ohne klinische oder radiologische Gonarthrose wurde kürzlich begonnen [99].

6.7 Vergleich mit anderen Übersichtsarbeiten und Leitlinien

Es gibt 2 Cochrane Reviews zu arthroskopischen Eingriffen bei Patienten mit Gonarthrose, eins zur Lavage mit 7 eingeschlossenen RCTs (567 Patienten) und eins zum Debridement mit 3 eingeschlossenen RCTs (271 Patienten) [22,100]. Das Review zur Lavage enthält zusätzlich zum Vergleich im vorliegenden Bericht auch die nicht arthroskopische Lavage als Prüfintervention [22]. Daher erfüllen lediglich 3 der 7 eingeschlossenen RCTs die Einschlusskriterien des vorliegenden Bericht: Kalunian 2000, Moseley 2002 und Moseley 1996. Das Review zum Debridement enthält auch Studien, die eine therapeutische Arthroskopie in der Vergleichsgruppe untersuchen [100]. Daher erfüllen lediglich 2 der 3 eingeschlossenen RCTs die Einschlusskriterien des vorliegenden Berichts: Chang 1993 und Moseley 2002. Beide Cochrane Reviews kommen zu ähnlichen Ergebnissen wie der vorliegende Bericht. Auch wenn die Prüfinerventionen anders gruppiert und zusammengefasst wurden, schlussfolgern beide Cochrane Reviews, dass sich für die jeweilige Gruppe von Interventionen kein patientenrelevanter Nutzen nachweisen lässt. Auch in der systematischen Übersicht von Avouac 2010 [101] konnte eine Wirksamkeit der Lavage mit und ohne Arthroskopie nicht nachgewiesen werden. Schließlich sah auch die systematische Übersicht von Samson 2007 [102] keinen Vorteil einer Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement bei Gonarthrose.

Das Cochrane Review zur Lavage [22] untersuchte auch die Frage, inwieweit die Art der Vergleichsbehandlung einen möglichen Therapieeffekt der Kniegelenkslavage in einer Studie beeinflusst. Hierbei zeigte sich, dass die Prüfintervention einen tendenziell größeren Effekt aufwies, wenn eine Studie keine Verblindung, also keinen Vergleich gegenüber einer Scheinintervention beinhaltete. Aufgrund des somit wahrscheinlichen Placebo-Effektes, der gerade bei invasiven Therapieverfahren besonders stark sei [103,104], sei daher zu erwarten, dass die in unverblindeten Studien gefundenen Therapieeffekte keiner Überprüfung in verblindeten Studien standhalten könnten. Dieser Sachverhalt hat für den vorliegenden Bericht eine hohe Bedeutung, da auch hier sowohl aktive als auch nicht aktive Vergleichsinterventionen vorkamen, wobei einzelne Studien aus der zweiten Gruppe im Rahmen der Verblindung doch eine aktive Intervention vortäuschten.

In den internationalen Leitlinien wurde nach der Publikation der Studien Moseley 2002 und Kirkley 2008 der klinische Stellenwert der therapeutischen Arthroskopie bei Gonarthrose skeptisch beurteilt. Die American Academy of Orthopaedic Surgeons (AAOS) riet 2008 explizit von arthroskopischer Lavage und arthroskopischem Debridement ab, sondern sah die Wirksamkeit der therapeutischen Arthroskopie auf Patienten mit Gonarthrose begrenzt, bei denen Symptome und Befunde auf einen Meniskusriss oder einen freien Gelenkkörper als Primärdiagnose hinweisen [105]. Die im März 2013 veröffentlichte Studie Katz et al. [44] konnte hier noch keine Berücksichtigung finden, stützt aber diese Empfehlung. In den Empfehlungen der Osteoarthritis Research Society International (OARSI) dagegen wird lediglich darauf hingewiesen, dass der Stellenwert von Kniegelenkslavage und arthroskopischem Debridement kontrovers sei [25]. In den Empfehlungen der Arzneimittelkommission der deutschen Ärzteschaft aus dem Jahr 2008 wird die therapeutische Arthroskopie als mögliche Therapieoption benannt, die bei frühen Krankheitsstadien und aktivierter Arthrose im Einzelfall vorteilhaft sein könne [106]. In den Empfehlungen des American College of Rheumatology (ACR) aus dem Jahr 2000 [107] und auch in den neuen Empfehlungen aus dem Jahr 2012 [24] zur pharmakologischen und nicht pharmakologischen Behandlung der Gonarthrose wird die therapeutische Arthroskopie hingegen nicht genannt. In der Leitlinie der Blue Cross Blue Shield of North Carolina aus dem Jahr 2012 wird die therapeutische Arthroskopie mit Lavage und gegebenenfalls zusätzlichem Debridement nicht als Behandlung der Gonarthrose empfohlen [108]. Das arthroskopische Debridement könne empfohlen werden, wenn bildgebende Untersuchungen darauf hinweisen, dass spezifische nicht arthrosebedingte anatomische Läsionen wie beispielsweise große Meniskusrisse und freie Gelenkkörper Symptome beim Patienten verursachen, unabhängig davon, ob eine Arthrose vorliegt [108].

Insgesamt unterstützen damit wesentliche internationale evidenz-basierte Leitlinien das Berichtsfazit dahingehend, dass die therapeutische Arthroskopie (mit Lavage und gegebenenfalls zusätzlichem Debridement) zur Behandlung einer Gonarthrose in ihrem Nutzen fraglich sei und daher im Regelfall nicht empfohlen werden könne.

7 Fazit

Der Nutzen der therapeutischen Arthroskopie (mit Lavage und gegebenenfalls zusätzlichem Debridement) zur Behandlung einer Gonarthrose ist nicht belegt. Im Vergleich zu keiner aktiven Vergleichsintervention zeigte sich für keinen patientenrelevanten Endpunkt ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie. In den Vergleichen zur Lavage, oralen Gabe von NSAIDs oder der intraartikulären Injektion von Hyaluronsäure zeigte sich ebenfalls für keinen Endpunkt ein Anhaltspunkt für, Hinweis auf oder Beleg für einen Nutzen der therapeutischen Arthroskopie. Lediglich in dem Vergleich zur intraartikulären Injektion von Glukokortikoiden ergab sich bei dem Endpunkt globale Bewertung der Symptomatik der Gonarthrose ein Anhaltspunkt für einen Nutzen der therapeutischen Arthroskopie.

8 Liste der eingeschlossenen Studien

Arden NK, Reading IC, Jordan KM, Thomas L, Platten H, Hassan A et al. A randomised controlled trial of tidal irrigation vs corticosteroid injection in knee osteoarthritis: the KIVIS study. *Osteoarthritis Cartilage* 2008; 16(6): 733-739.

Chang RW, Falconer J, Stulberg SD, Arnold WJ, Manheim LM, Dyer AR. A randomized, controlled trial of arthroscopic surgery versus closed-needle joint lavage for patients with osteoarthritis of the knee. *Arthritis Rheum* 1993; 36(3): 289-296.

Forster MC, Straw R. A prospective randomised trial comparing intra-articular Hyalgan injection and arthroscopic washout for knee osteoarthritis. *Knee* 2003; 10(3): 291-293.

Kalunian KC, Moreland LW, Klashman DJ, Brion PH, Concoff AL, Myers S et al. Visually-guided irrigation in patients with early knee osteoarthritis: a multicenter randomized, controlled trial. *Osteoarthritis Cartilage* 2000; 8(6): 412-418.

Kang JG, Wang ML, Zhang XN. Treatment of knee osteoarthritis with arthroscopic debridement and intra-articular sodium hyaluronate injection [Chinesisch]. *Journal of Jilin University Medicine Edition* 2005; 31(5): 802-805.

Katz JN, Brophy RH, Chaisson CE, De Chaves L, Cole BJ, Dahm DL et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med* 18.03.2013 [Epub ahead of print].

Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ et al. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008; 359(11): 1097-1107.

Merchan EC, Galindo E. Arthroscope-guided surgery versus nonoperative treatment for limited degenerative osteoarthritis of the femorotibial joint in patients over 50 years of age: a prospective comparative study. *Arthroscopy* 1993; 9(6): 663-667.

Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(2): 81-88.

Moseley JB Jr, Wray NP, Kuykendall D, Willis K, Landon G. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo-controlled trial; results of a pilot study. *Am J Sports Med* 1996; 24(1): 28-34.

9 Literatur

1. Altman RD. Early management of osteoarthritis. Am J Manag Care 2010; 16(Suppl Management): S41-S47.
2. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986; 29(8): 1039-1049.
3. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis 2010; 69(3): 483-489.
4. Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie »Gesundheit in Deutschland aktuell 2009«. Berlin: RKI; 2009. URL: http://www.gbe-bund.de/gbe10/owards.prc_show_pdf?p_id=13126&p_sprache=D.
5. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003; 81(9): 646-656.
6. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part II. Arthritis Rheum 2008; 58(1): 26-35.
7. Günther KP, Stürmer T, Sauerland S, Zeissig I, Sun Y, Kessler S et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998; 57(12): 717-723.
8. Statistisches Bundesamt. Diagnosedaten der Krankenhäuser ab 2000 (Fälle/Sterbefälle, Pflage tage, durchschnittliche Verweildauer); Gliederungsmerkmale: Jahre, Behandlungsort, Alter, Geschlecht, Verweildauer, ICD10; Jahr 2009, Behandlungsort Deutschland, alle Altersgruppen, beide Geschlechter, alle Fälle [online]. In: Gesundheitsberichterstattung des Bundes. [Zugriff: 31.01.2011]. URL: <http://www.gbe-bund.de>.
9. Statistisches Bundesamt. DRG-Statistik 2009: vollstationäre Patientinnen und Patienten in Krankenhäusern [unveröffentlicht].
10. Muthuri SG, McWilliams DF, Doherty M, Zhang W. History of knee injuries and knee osteoarthritis: a meta-analysis of observational studies. Osteoarthritis Cartilage 2011; 19(11): 1286-1293.
11. Statistisches Bundesamt. Arthrose: Kapitel 5.10 [online]. In: Gesundheitsberichterstattung des Bundes: Gesundheitsbericht für Deutschland. 1998 [Zugriff: 20.01.2011]. URL: http://www.gbe-bund.de/gbe10/abrechnung.prc_abr_test_logon?p_aid=96734606&p_uid=gastg&p_sprache=D&p_knoten=FID&p_suchstring=909.
12. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Rheum Dis Clin North Am 2008; 34(3): 515-529.

13. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 2010; 18(1): 24-33.
14. Bijlsma JW, Berenbaum F, Lافeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet* 2011; 377(9783): 2115-2126.
15. Deutsche Gesellschaft für Unfallchirurgie, Österreichische Gesellschaft für Unfallchirurgie. Endoprothese bei Gonarthrose [online]. 06.2009 [Zugriff: 08.11.2011]. URL: http://www.awmf.org/uploads/tx_szleitlinien/012-008l_S1_Endoprothese_bei_Gonarthrose_2009.pdf.
16. Altman RD, Fries JF, Bloch DA, Carstens J, Cooke TD, Genant H et al. Radiographic assessment of progression in osteoarthritis. *Arthritis Rheum* 1987; 30(11): 1214-1225.
17. Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Ann Rheum Dis* 2008; 67(7): 1034-1036.
18. Hunter DJ, Lo GH. The management of osteoarthritis: an overview and call to appropriate conservative treatment. *Rheum Dis Clin North Am* 2008; 34(3): 689-712.
19. Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *J Rheumatol* 2000; 27(6): 1513-1517.
20. Towheed TE, Maxwell L, Judd MG, Catton M, Hochberg MC, Wells G. Acetaminophen for osteoarthritis. *Cochrane Database Syst Rev* 2006; (1): CD004257.
21. Doherty M, Hawkey C, Goulder M, Gibb I, Hill N, Aspley S et al. A randomised controlled trial of ibuprofen, paracetamol or a combination tablet of ibuprofen/paracetamol in community-derived people with knee pain. *Ann Rheum Dis* 2011; 70(9): 1534-1541.
22. Reichenbach S, Rutjes AW, Nuesch E, Trelle S, Juni P. Joint lavage for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2010; (5): CD007320.
23. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol* 2006; 20(4): 721-740.
24. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res (Hoboken)* 2012; 64(4): 455-474.
25. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al. OARSI recommendations for the management of hip and knee osteoarthritis; part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008; 16(2): 137-162.
26. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Gonarthrose: Registernummer 033 - 004; angemeldetes Leitlinienvorhaben [online]. 21.06.2011 [Zugriff: 07.05.2013]. URL: <http://www.awmf.org/leitlinien/detail/anmeldung/1/ll/033-004.html>.

27. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Beratungsverfahrens: Bewertung der Arthroskopie des Kniegelenks bei Kniegelenksarthrose (Gonarthrose) gemäß § 137c SGB V [online]. 20.01.2011 [Zugriff: 19.12.2011]. URL: http://www.g-ba.de/downloads/40-268-1545/2011-01-20_SN-137c-Arthroskopie-Kniegelenk_TrG.pdf.
28. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869.
29. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet 2002; 359(9308): 781-785.
30. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7(3): 177-188.
31. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses [online]. In: Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions: version 5.1.0. 03.2011 [Zugriff: 13.06.2012]. URL: http://www.mrc-bsu.cam.ac.uk/cochrane/handbook/chapter_9/9_analysing_data_and_undertaking_meta_analyses.htm.
32. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327(7414): 557-560.
33. Fowler PJ. Surgery versus no surgery for OA of the knee [online]. In: ClinicalTrials.gov. 06.12.2007 [Zugriff: 02.02.2012]. URL: <http://ClinicalTrials.gov/show/NCT00158431>.
34. Fowler PJ. Surgery versus no surgery for osteoarthritis (OA) of the knee [online]. In: ClinicalTrials.gov. 21.11.2007 [Zugriff: 02.02.2012]. URL: <http://clinicaltrials.gov/ct2/show/study/NCT00562822>.
35. Mohsen A. Prospective randomised single blind trial comparing the effectiveness of combined arthroscopic washout and intra articular hyaluronan injection to intra articular hyaluronan injection and arthroscopic washout in isolation, for osteoarthritis of knee [online]. In: International Standard Randomised Controlled Trial Number Register. 28.09.2011 [Zugriff: 02.02.2012]. URL: <http://www.controlled-trials.com/ISRCTN82192623/ISRCTN82192623>.
36. Giffin JR. HTO with and without arthroscopy [online]. In: ClinicalTrials.gov. 28.01.2009 [Zugriff: 02.02.2012]. URL: <http://ClinicalTrials.gov/show/NCT00284622>.
37. Gemeinsamer Bundesausschuss. Bekanntmachung des Gemeinsamen Bundesausschusses über weitere Beratungsthemen zur Überprüfung gemäß §135 Absatz 1 und §137c Absatz 1 des Fünften Buches Sozialgesetzbuch (SGB V): Bewertung der Arthroskopie des Kniegelenks bei Gonarthrose. Bundesanzeiger 2011; (87): 2062.

38. Arden NK, Reading IC, Jordan KM, Thomas L, Platten H, Hassan A et al. A randomised controlled trial of tidal irrigation vs corticosteroid injection in knee osteoarthritis: the KIVIS study. *Osteoarthritis Cartilage* 2008; 16(6): 733-739.
39. Chang RW, Falconer J, Stulberg SD, Arnold WJ, Manheim LM, Dyer AR. A randomized, controlled trial of arthroscopic surgery versus closed-needle joint lavage for patients with osteoarthritis of the knee. *Arthritis Rheum* 1993; 36(3): 289-296.
40. Forster MC, Straw R. A prospective randomised trial comparing intra-articular Hyalgan injection and arthroscopic washout for knee osteoarthritis. *Knee* 2003; 10(3): 291-293.
41. Kalunian KC, Moreland LW, Klashman DJ, Brion PH, Concoff AL, Myers S et al. Visually-guided irrigation in patients with early knee osteoarthritis: a multicenter randomized, controlled trial. *Osteoarthritis Cartilage* 2000; 8(6): 412-418.
42. Kang JG, Wang ML, Zhang XN. Treatment of knee osteoarthritis with arthroscopic debridement and intra-articular sodium hyaluronate injection [Chinesisch]. *Journal of Jilin University Medicine Edition* 2005; 31(5): 802-805.
43. Katz JN. Comparing knee cartilage surgery versus standard physical therapy in treating people with a meniscal tear and osteoarthritis: full text view [online]. In: *Clinicaltrials.gov*. 21.02.2012 [Zugriff: 16.07.2013]. URL: <http://clinicaltrials.gov/ct2/show/NCT00597012>.
44. Katz JN, Brophy RH, Chaisson CE, De Chaves L, Cole BJ, Dahm DL et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med* 18.03.2013 [Epub ahead of print].
45. Katz JN, Chaisson CE, Cole B, Guermazi A, Hunter DJ, Jones M et al. The MeTeOR trial (Meniscal Tear in Osteoarthritis Research): rationale and design features. *Contemp Clin Trials* 2012; 33(6): 1189-1196.
46. Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ et al. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008; 359(11): 1097-1107.
47. Merchan EC, Galindo E. Arthroscope-guided surgery versus nonoperative treatment for limited degenerative osteoarthritis of the femorotibial joint in patients over 50 years of age: a prospective comparative study. *Arthroscopy* 1993; 9(6): 663-667.
48. Moseley JB Jr, Wray NP, Kuykendall D, Willis K, Landon G. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo-controlled trial; results of a pilot study. *Am J Sports Med* 1996; 24(1): 28-34.
49. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(2): 81-88.
50. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Ann Rheum Dis* 1957; 16(4): 494-502.

51. Martin Andres A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. *Comput Stat Data Anal* 1994; 17(5): 555-574.
52. Puhan MA, Soesilo I, Guyatt GH, Schünemann HJ. Combining scores from different patient reported outcome measures in meta-analyses: when is it justified? *Health Qual Life Outcomes* 2006; 4: 94.
53. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. *Clin Rheumatol* 2010; 29(7): 749-756.
54. Rosemann T, Szecsenyi J. Cultural adaptation and validation of a German version of the Arthritis Impact Measurement Scales (AIMS2). *Osteoarthritis Cartilage* 2007; 15(10): 1128-1133.
55. Lin YC, Davey RC, Cochrane T. Tests for physical function of the elderly with knee and hip osteoarthritis. *Scand J Med Sci Sports* 2001; 11(5): 280-286.
56. Maly MR, Costigan PA, Olney SJ. Determinants of self-report outcome measures in people with knee osteoarthritis. *Arch Phys Med Rehabil* 2006; 87(1): 96-104.
57. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988; 15(12): 1833-1840.
58. Gentelle-Bonnassies S, Le Claire P, Mezieres M, Ayral X, Dougados M. Comparison of the responsiveness of symptomatic outcome measures in knee osteoarthritis. *Arthritis Care Res* 2000; 13(5): 280-285.
59. Goldhahn S, Sawaguchi T, Audige L, Mundi R, Hanson B, Bhandari M et al. Complication reporting in orthopaedic trials: a systematic review of randomized controlled trials. *J Bone Joint Surg Am* 2009; 91(8): 1847-1853.
60. Martin RC 2nd, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg* 2002; 235(6): 803-813.
61. Maletis GB, Inacio MC, Reynolds S, Funahashi TT. Incidence of symptomatic venous thromboembolism after elective knee arthroscopy. *J Bone Joint Surg Am* 2012; 94(8): 714-720.
62. Paul J, Kirchhoff C, Imhoff AB, Hinterwimmer S. Infektion nach Arthroskopie. *Orthopade* 2008; 37(11): 1048-1055.
63. Bamford DJ, Paul AS, Noble J, Davies DR. Avoidable complications of arthroscopic surgery. *J R Coll Surg Edinb* 1993; 38(2): 92-95.
64. Williams JJ, Hulstyn MJ, Fadale PD, Lindy PB, Ehrlich MG, Cronan J et al. Incidence of deep vein thrombosis after arthroscopic knee surgery: a prospective study. *Arthroscopy* 1995; 11(6): 701-705.

65. Demers C, Marcoux S, Ginsberg JS, Laroche F, Cloutier R, Poulin J. Incidence of venographically proved deep vein thrombosis after knee arthroscopy. *Arch Intern Med* 1998; 158(1): 47-50.
66. Delis KT, Hunt N, Strachan RK, Nicolaides AN. Incidence, natural history and risk factors of deep vein thrombosis in elective knee arthroscopy. *Thromb Haemost* 2001; 86(3): 817-821.
67. Yacub JN, Rice JB, Dillingham TR. Nerve injury in patients after hip and knee arthroplasties and knee arthroscopy. *Am J Phys Med Rehabil* 2009; 88(8): 635-641.
68. Kane D, Veale DJ, FitzGerald O, Reece R. Survey of arthroscopy performed by rheumatologists. *Rheumatology (Oxford)* 2002; 41(2): 210-215.
69. Kieser C. A review of the complications of arthroscopic knee surgery. *Arthroscopy* 1992; 8(1): 79-83.
70. Bohensky MA, Desteiger R, Kondogiannis C, Sundararajan V, Andrianopoulos N, Bucknill A et al. Adverse outcomes associated with elective knee arthroscopy: a population-based cohort study. *Arthroscopy* 2013; 29(4): 716-725.
71. Jackson RW. Debate on the use of arthroscopic surgery for osteoarthritis of the knee. *Proc (Bayl Univ Med Cent)* 2003; 16(1): 27.
72. Laskin RS, Ohnsorge JAK. The role of arthroscopy: beneficial, placebo, or worse? *Orthopedics* 2005; 28(9): 975-976.
73. Felson DT. Arthroscopy as a treatment for knee osteoarthritis. *Best Pract Res Clin Rheumatol* 2010; 24(1): 47-50.
74. Johnson LL. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *Arthroscopy* 2002; 18(7): 683-686.
75. Chambers K, Schulzer M, Sobolev B. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *Arthroscopy* 2002; 18(7): 686-687.
76. Jackson RW. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(21): 1717.
77. Ewing W, Ewing JW. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(21): 1717.
78. Chambers KG, Schulzer M. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(21): 1717-1718.
79. Morse LJ. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(21): 1718.
80. Blacher RS. Arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347(21): 1718.

81. O'Malley KJ, Suarez-Almazor M, Aniol J, Richardson P, Kuykendall DH, Moseley JB Jr et al. Joint-specific multidimensional assessment of pain (J-MAP): factor structure, reliability, validity, and responsiveness in patients with knee osteoarthritis. *J Rheumatol* 2003; 30(3): 534-543.
82. Felson DT, Buckwalter J. Debridement and lavage for osteoarthritis of the knee. *N Engl J Med* 2002; 347(2): 132-133.
83. Horng S, Miller FG. Is placebo surgery unethical? *N Engl J Med* 2002; 347(2): 137-139.
84. Tenery R, Rakatansky H, Riddick FA Jr, Goldrich MS, Morse LJ, O'Bannon JM 3rd et al. Surgical "placebo" controls. *Ann Surg* 2002; 235(2): 303-307.
85. Holmes R, Moschetti W, Martin B, Tomek I, Finlayson S. Effect of evidence and changes in reimbursement on the rate of arthroscopy for osteoarthritis. *Am J Sports Med* 04.03.2013 [Epub ahead of print].
86. Howard D, Brophy R, Howell S. Evidence of no benefit from knee surgery for osteoarthritis led to coverage changes and is linked to decline in procedures. *Health Aff (Millwood)* 2012; 31(10): 2242-2249.
87. Harris IA, Madan NS, Naylor JM, Chong S, Mittal R, Jalaludin BB. Trends in knee arthroscopy and subsequent arthroplasty in an Australian population: a retrospective cohort study. *BMC Musculoskelet Disord* 2013; 14: 143.
88. Hawker G, Guan J, Judge A, Dieppe P. Knee arthroscopy in England and Ontario: patterns of use, changes over time, and relationship to total knee replacement. *J Bone Joint Surg Am* 2008; 90(11): 2337-2345.
89. Potts A, Harrast JJ, Harner CD, Miniaci A, Jones MH. Practice patterns for arthroscopy of osteoarthritis of the knee in the United States. *Am J Sports Med* 2012; 40(6): 1247-1251.
90. Bohensky MA, Sundararajan V, Andrianopoulos N, De Steiger RN, Bucknill A, Kondogiannis CM et al. Trends in elective knee arthroscopies in a population-based cohort, 2000-2009. *Med J Aust* 2012; 197(7): 399-403.
91. Bradley JD, Heilman DK, Katz BP, Gsell P, Wallick JE, Brandt KD. Tidal irrigation as treatment for knee osteoarthritis: a sham-controlled, randomized, double-blinded evaluation. *Arthritis Rheum* 2002; 46(1): 100-108.
92. Ravaud P, Moulinier L, Giraudeau B, Ayral X, Guerin C, Noel E et al. Effects of joint lavage and steroid injection in patients with osteoarthritis of the knee: results of a multicenter, randomized, controlled trial. *Arthritis Rheum* 1999; 42(3): 475-482.
93. Marx RG. Arthroscopic surgery for osteoarthritis of the knee? *N Engl J Med* 2008; 359(11): 1169-1170.
94. Englund M, Guermazi A, Gale D, Hunter DJ, Aliabadi P, Clancy M et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med* 2008; 359(11): 1108-1115.

95. Jackson JL, O'Malley PG, Kroenke K. Evaluation of acute knee pain in primary care. *Ann Intern Med* 2003; 139(7): 575-588.
96. Herrlin SV, Wange PO, Lapidus G, Hållander M, Werner S, Weidenhielm L. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. *Knee Surg Sports Traumatol Arthrosc* 2013; 21(2): 358-364.
97. Herrlin S, Hållander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007; 15(4): 393-401.
98. Ahlbäck S. Osteoarthritis of the knee: a radiographic investigation. *Acta Radiol Diagn (Stockh)* 1968; (Suppl 277): 7-72.
99. Sihvonen R, Paavola M, Malmivaara A, Järvinen TL. Finnish Degenerative Meniscal Lesion Study (FIDELITY): a protocol for a randomised, placebo surgery controlled trial on the efficacy of arthroscopic partial meniscectomy for patients with degenerative meniscus injury with a novel 'RCT within-a-cohort' study design. *BMJ Open* 2013; 3(3): e002510.
100. Laupattarakasem W, Laopaiboon M, Laupattarakasem P, Sumananont C. Arthroscopic debridement for knee osteoarthritis. *Cochrane Database Syst Rev* 2008; (1): CD005118.
101. Avouac J, Vicaut E, Bardin T, Richette P. Efficacy of joint lavage in knee osteoarthritis: meta-analysis of randomized controlled studies. *Rheumatology (Oxford)* 2010; 49(2): 334-340.
102. Samson DJ, Grant MD, Ratko TA, Bonnell CJ, Ziegler KM, Aronson N. Treatment of primary and secondary osteoarthritis of the knee: AHRQ publication no. 07-E012 [online]. 09.2007 [Zugriff: 16.12.2011]. (Evidence Report/Technology Assessment; Band 157). URL: <http://www.ahrq.gov/downloads/pub/evidence/pdf/oaknee/oaknee.pdf>.
103. Beecher HK. Surgery as placebo: a quantitative study of bias. *JAMA* 1961; 176(13): 1102-1107.
104. Johnson AG. Surgery as a placebo. *Lancet* 1994; 344(8930): 1140-1142.
105. American Academy of Orthopaedic Surgeons. Treatment of osteoarthritis of the knee (non-arthroplasty) [online]. 06.12.2008 [Zugriff: 18.05.2011]. URL: <http://www.aaos.org/research/guidelines/oakguideline.pdf>.
106. Arzneimittelkommission der deutschen Ärzteschaft. Degenerative Gelenkerkrankungen. *Arzneiverordnung in der Praxis* 2008; 35(Sonderheft 1 Therapieempfehlungen): 1-39.
107. Altman RD, Hochberg MC, Moskowitz RW, Schnitzer TJ. Recommendations for the medical management of osteoarthritis of the hip and knee [online]. 2000 [Zugriff: 05.09.2012]. URL: <http://www.rheumatology.org/practice/clinical/guidelines/oa-mgmt.asp>.

108. BlueCross BlueShield of North Carolina. Evidence-based guideline: arthroscopic debridement and lavage as treatment of knee osteoarthritis [online]. 02.2012 [Zugriff: 05.09.2012]. URL: http://www.bcbsnc.com/assets/services/public/pdfs/medicalpolicy/arthroscopic_debridement_and_lavage_as_treatment_of_knee_osteoarthritis.pdf.
109. Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc* 2006; 94(4): 451-455.
110. Lefebvre C, Manheimer E, Glanville J. Searching for studies. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. New York: Wiley; 2008. S. 95-150.
111. Meenan RF, Kazis LE, Anderson JJ. The stability of health status in rheumatoid arthritis: a five-year study of patients with established disease. *Am J Public Health* 1988; 78(11): 1484-1487.
112. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2: the content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum* 1992; 35(1): 1-10.
113. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee: validation; value in comparison with other assessment tests. *Scand J Rheumatol Suppl* 1987; 65: 85-89.
114. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med* 1982; 10(3): 150-154.
115. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function. *Arthritis Care Res (Hoboken)* 2011; 63(S11): S208-S228.
116. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop* 1985; (198): 43-49.
117. Wolfe F, Michaud K, Li T. Sleep disturbance in patients with rheumatoid arthritis: evaluation by medical outcomes study and visual analog sleep scales. *J Rheumatol* 2006; 33(10): 1942-1951.
118. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis* 2005; 64(1): 34-37.
119. Stauffer ME, Taylor SD, Watson DJ, Peloso PM, Morrison A. Definition of nonresponse to analgesic treatment of arthritic pain: an analytical literature review of the smallest detectable difference, the minimal detectable change, and the minimal clinically important difference on the pain visual analog scale. *Int J Inflam* 2011; 2011: 231926.

120. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop 1989; (248): 13-14.
121. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS): validation of a Swedish version. Scand J Med Sci Sports 1998; 8(6): 439-448.
122. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36); I: conceptual framework and item selection. Med Care 1992; 30(6): 473-483.
123. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001; 45(4): 384-391.
124. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire: an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987; 14(3): 446-451.
125. Lauterbach KW, Stock S, Brunner H (Ed). Gesundheitsökonomie: Lehrbuch für Mediziner und andere Gesundheitsberufe. Bern: Huber; 2006.
126. Schöffski O, Graf von der Schulenburg JM (Ed). Gesundheitsökonomische Evaluationen. Berlin: Springer; 2007.
127. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67(6): 361-370.

Anhang A– Suchstrategien

1. EMBASE

Suchoberfläche: Ovid

- EMBASE 1980 to 2012 January 09

Es wurden folgende Filter übernommen:

Systematic Review: Wong [109] – High specificity strategy;

RCT: Wong [109] – High sensitivity strategies

#	Searches
1	exp Arthroscopic Surgery/
2	arthroscop*.ti,ab.
3	or/1-2
4	Knee Osteoarthritis/
5	(osteoarthritis* or osteoarthro* or oa or degenerative joint disease*).ti,ab.
6	Joint Degeneration/
7	knee*.ti,ab.
8	Knee Arthroscopy/
9	5 and (6 or 7)
10	5 and 8
11	3 and (4 or 9)
12	10 or 11
13	(meta analysis or systematic review or MEDLINE).tw.
14	random*.tw.
15	clinical trial*.mp.
16	exp health care quality/
17	or/14-16
18	or/13,17
19	12 and 18

2. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) 1946 to December Week 4 2011
- Ovid MEDLINE(R) Daily Update January 09, 2012
- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations January 09, 2012

Es wurden folgende Filter übernommen:

Systematic Review: Wong [109] – Strategy minimizing difference between sensitivity and specificity;

RCT: Lefebvre [110] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision maximizing version (2008 revision)

#	Searches
1	Arthroscopy/
2	arthroscop*.ti,ab.
3	or/1-2
4	Osteoarthritis, Knee/
5	Osteoarthritis/
6	(osteoarthrit* or osteoarthro* or oa or degenerative joint disease*).ti,ab.
7	exp Knee Joint/
8	knee*.ti,ab.
9	(5 or 6) and (7 or 8)
10	3 and (4 or 9)
11	randomized controlled trial.pt.
12	controlled clinical trial.pt.
13	randomized.ab.
14	placebo.ab.
15	drug therapy.fs.
16	randomly.ab.
17	trial.ab.
18	groups.ab.
19	or/11-18
20	(animals not (humans and animals)).sh.
21	19 not 20
22	search*.tw.
23	meta analysis.mp,pt.
24	review.pt.
25	or/22-24
26	or/21,25
27	10 and 26

3. PubMed

Suchoberfläche: NLM

- PubMed - as supplied by publisher
- PubMed - in process

- PubMed – OLDMEDLINE
- PubMed - pubmednotmedline

Search	Most Recent Queries
#1	Search arthroscop*[tiab]
#2	Search osteoarthritis*[tiab] or osteoarthro*[tiab] or oa[tiab] or "degenerative joint disease*" [tiab]
#3	Search knee*[tiab]
#4	Search #1 AND #2 AND #3
#5	Search #4 not medline[sb]

4. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews (Cochrane Reviews), Issue 12, 2011
- Cochrane Central Register of Controlled Trials (Clinical Trials), Issue 4, 2011
- Database of Abstracts of Reviews of Effects (Other Reviews), Issue 4, 2011
- Health Technology Assessment Database (Technology Assessments), Issue 4, 2011

#	Searches
1	MeSH descriptor Arthroscopy explode all trees
2	arthroscop*:ti,ab
3	(#1 OR #2)
4	MeSH descriptor Osteoarthritis, Knee explode all trees
5	MeSH descriptor Osteoarthritis explode all trees
6	(osteoarthritis* or osteoarthro* or oa or degenerative joint disease*):ti,ab
7	MeSH descriptor Knee Joint explode all trees
8	knee*:ti,ab
9	((#5 OR #6) AND (#7 OR #8))
10	(#3 AND (#4 OR #9))

Anhang B – Liste der ausgeschlossenen Dokumente mit Ausschlussgründen

Nicht E1

1. Fabricant PD, Jokl P. Surgical outcomes after arthroscopic partial meniscectomy. *J Am Acad Orthop Surg* 2007; 15(11): 647-653.
2. Falah M, Nierenberg G, Soudry M, Hayden M, Volpin G. Treatment of articular cartilage lesions of the knee. *Int Orthop* 2010; 34(5): 621-630.
3. Garrett WE, Kaeding CC, ElAttrache NS, Xerogeanes JW, Hewitt MS, Skrepnik NV et al. Novel drug OMS103HP reduces pain and improves joint motion and function for 90 days after arthroscopic meniscectomy. *Arthroscopy* 2011; 27(8): 1060-1070.
4. Herrlin S, Hållander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007; 15(4): 393-401.
5. Kirkley A, Rampersaud R, Griffin S, Amendola A, Litchfield R, Fowler P. Tourniquet versus no tourniquet use in routine knee arthroscopy: a prospective, double-blind, randomized clinical trial. *Arthroscopy* 2000; 16(2): 121-126.
6. Papalia R, Del Buono A, Osti L, Denaro V, Maffulli N. Meniscectomy as a risk factor for knee osteoarthritis: a systematic review. *Br Med Bull* 2011; 99(1): 89-106.
7. Spahn G, Kahl E, Muckley T, Hofmann GO, Klinger HM. Arthroscopic knee chondroplasty using a bipolar radiofrequency-based device compared to mechanical shaver: results of a prospective, randomized, controlled study. *Knee Surg Sports Traumatol Arthrosc* 2008; 16(6): 565-573.
8. Spahn G, Klinger HM, Muckley T, Hofmann GO. Four-year results from a randomized controlled study of knee chondroplasty with concomitant medial meniscectomy: mechanical debridement versus radiofrequency chondroplasty. *Arthroscopy* 2010; 26(9 Suppl): S73-S80.
9. Van Oosterhout M, Sont JK, Bajema IM, Breedveld FC, Van Laar JM. Comparison of efficacy of arthroscopic lavage plus administration of corticosteroids, arthroscopic lavage plus administration of placebo, and joint aspiration plus administration of corticosteroids in arthritis of the knee: a randomized controlled trial. *Arthritis Rheum* 2006; 55(6): 964-970.
10. Van Oosterhout M, Sont JK, Bajema IM, Breedveld FC, Van Laar JM. Arthroscopic lavage plus corticosteroids is more effective than joint aspiration plus corticosteroids in patients with arthritis of the knee [Niederländisch]. *Ned Tijdschr Geneesk* 2008; 152(36): 1973-1980.
11. Vincken PWJ, Ter Braak BPM, Van Erkel AR, Bloem RM, Van Luijt PA, Coene LNJEM et al. Only MR can safely exclude patients from arthroscopy. *Skeletal Radiol* 2009; 38(10): 977-982.

12. Weale AE, Ackroyd CE, Mani GV, Winson IG. Day-case or short-stay admission for arthroscopic knee surgery: a randomised controlled trial. *Ann R Coll Surg Engl* 1998; 80(2): 146-149.

13. Wei N, Delauter SK, Erlichman MS. The holmium YAG laser in office based arthroscopy of the knee: comparison with standard interventional instruments in patients with arthritis. *J Rheumatol* 1997; 24(9): 1806-1808.

Nicht E2

1. Aydogan NH, Baydar ML, Atay T, Perktas I, Baykal BY, Ozmeric A. The effect of arthroscopic surgery and intraarticular drug injection to the antioxidation system and lipid peroxidation at osteoarthritis of knee. *Saudi Med J* 2008; 29(3): 397-402.

2. Davis A. Osteoarthritis year in review: rehabilitation and outcomes. *Osteoarthritis Cartilage* 2011; 19(Suppl 1): S6.

Nicht E3

1. Edelson R, Burks RT, Bloebaum RD. Short-term effects of knee washout for osteoarthritis. *Am J Sports Med* 1995; 23(3): 345-349.

2. Hubbard MJ. Articular debridement versus washout for degeneration of the medial femoral condyle: a five-year study. *J Bone Joint Surg Br* 1996; 78(2): 217-219.

Nicht E4

1. Campbell MK, Entwistle VA, Cuthbertson BH, Skea ZC, Sutherland AG, McDonald AM et al. Developing a placebo-controlled trial in surgery: issues of design, acceptability and feasibility. *Trials* 2011; 12: 50.

2. Chang RW, Falconer J, Stulberg SD, Arnold WJ, Dyer AR. Prerandomization: an alternative to classic randomization; the effects on recruitment in a controlled trial of arthroscopy for osteoarthritis of the knee. *J Bone Joint Surg Am* 1990; 72(10): 1451-1455.

3. Ou ZX, Jin JC, Huang D. Comparative study on effects of combined massage-smouldering-washing therapy and mini-invasive surgery in treating knee osteoarthritis [Chinesisch]. *Zhongguo Zhong Xi Yi Jie He Za Zhi* 2008; 28(10): 925-928.

Nicht E5

1. Arthroscopic surgery may not be best treatment for knee osteoarthritis. *Mayo Clin Womens Healthsource* 2009; 13(2): 3.

2. Agency for Healthcare Research and Quality. Three treatments for osteoarthritis of the knee: evidence shows lack of benefit comparative effectiveness review summary guides for clinicians [online]. 04.2009 [Zugriff: 12.04.2012]. URL: http://www.effectivehealthcare.ahrq.gov/ehc/products/89/134/20090408_OsteoKneeClinician.pdf.

3. Almazán A, Cruz F, Pérez F, Bravo C, Ibarra C. Arthroscopic treatment for osteoarthritis: knee and shoulder [Spanisch]. *Reumatol Clin* 2007; 3(Suppl 3): S44-S49.

4. Aly TA, Hafez K, Amin O. Arthrodiatasis for management of knee osteoarthritis. *Orthopedics* 2011; 34(8): e338-e343.
5. Başaran SH, Öneş HN, Peker G, Edipoğlu E, Avkan MC. The effect of arthroscopic debridement on life quality of mild and moderate gonarthrosis patients [Türkisch]. *Medical Journal of Bakirköy* 2011; 7(1): 20-26.
6. Bert JM. Role of abrasion arthroplasty and debridement in the management of osteoarthritis of the knee. *Rheum Dis Clin North Am* 1993; 19(3): 725-739.
7. Burks RT. Arthroscopy and degenerative arthritis of the knee: a review of the literature. *Arthroscopy* 1990; 6(1): 43-47.
8. Calvert GT, Wright RW. The use of arthroscopy in the athlete with knee osteoarthritis. *Clin Sports Med* 2005; 24(1): 133-152.
9. Canosa Sevillano R. Arthroscopic debridement and abrasion arthroplasty: current possibilities [Spanisch]. *Revista de Ortopedia y Traumatologia* 2002; 46(5): 451-457.
10. Clarke HD, Scott WN. The role of debridement: through small portals. *J Arthroplasty* 2003; 18(3 Suppl 1): 10-13.
11. Cooper MT, Miller MD. Primary articular cartilage procedures in the middle-aged knee. *Sports Med Arthrosc* 2003; 11(2): 112-121.
12. David T, Gambardella RA. Arthroscopic debridement of the arthritic knee: indications and results. *Curr Opin Orthop* 2000; 11(1): 9-13.
13. Fairweather J, Jawad ASM. Arthroscopic surgery and osteoarthritis of knee. *Saudi Med J* 2008; 29(9): 1361.
14. Feeley BT, Gallo RA, Sherman S, Williams RJ. Management of osteoarthritis of the knee in the active patient. *J Am Acad Orthop Surg* 2010; 18(7): 406-416.
15. Felson DT. Arthroscopy as a treatment for knee osteoarthritis. *Best Pract Res Clin Rheumatol* 2010; 24(1): 47-50.
16. Glinz W. Operative Möglichkeiten bei der Kniearthroskopie. *Aktuelle Probl Chir Orthop* 1991; 40: 64-76.
17. Glinz W. Arthroskopische Operationstechniken bei degenerativen Meniscus- und Gelenkflächenschäden des Kniegelenks. *Chirurg* 1993; 64(5): 371-378.
18. Hanssen AD, Stuart MJ, Scott RD, Scuderi GR. Surgical options for the middle-aged patient with osteoarthritis of the knee joint. *Instr Course Lect* 2001; 50: 499-511.
19. Hitzeman N, Masley C. Arthroscopic surgery for knee osteoarthritis. *Am Fam Physician* 2008; 78(3): 331-333.
20. Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR et al. Guidelines for the medical management of osteoarthritis; part II: osteoarthritis of the knee. *Arthritis Rheum* 1995; 38(11): 1541-1546.

21. Howell SM. The role of arthroscopy in treating osteoarthritis of the knee in the older patient. *Orthopedics* 2010; 33(9): 652.
22. Huang LK, Chen SR, Tang J, Peng LH, Xu T, Feng QL. Effectiveness comparison between knee replacement and arthroscopic debridement combined sodium hyaluronate injection for knee osteoarthritis [Chinesisch]. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2010; 14(13): 2345-2348.
23. Hunt SA, Jazrawi LM, Sherman OH. Arthroscopic management of osteoarthritis of the knee. *J Am Acad Orthop Surg* 2002; 10(5): 356-363.
24. Ioannidis G. The current role of arthroscopy in the treatment of osteoarthritis of the knee. *Curr Opin Orthop* 2004; 15(1): 31-36.
25. Jackson RW, Gilbert JE, Sharkey PF. Arthroscopic debridement versus arthroplasty in the osteoarthritic knee. *J Arthroplasty* 1997; 12(4): 465-469.
26. Karahan M. Diagnosis and management of unicompartmental knee osteoarthritis. *Marmara Medical Journal* 2003; 16(3): 228-235.
27. Kellett CF, Boscainos PJ, Gross AE. Surgical options for articular defects of the knee. *Expert Rev Med Devices* 2006; 3(5): 585-593.
28. Krishnan SP, Skinner JA. (i) Novel treatments for early osteoarthritis of the knee. *Curr Orthop* 2005; 19(6): 407-414.
29. Kuzmanova S, Andreev S, Atanassov A, Batalov A, Solakov P. Arthroscopic synovectomy, debridement and tidal lavage of the knee joint in osteoarthritis: long term results [Bulgarisch]. *Rheumatology* 2004; 12(4): 42-47.
30. Lane NE, Thompson JM. Management of osteoarthritis in the primary-care setting: an evidence-based approach to treatment. *Am J Med* 1997; 103(6 Suppl 1): S25-S30.
31. Laskin RS, Ohnsorge JAK. The role of arthroscopy: beneficial, placebo, or worse? *Orthopedics* 2005; 28(9): 975-976.
32. Livesley PJ, Doherty M, Needoff M, Moulton A. Arthroscopic lavage of osteoarthritic knees. *J Bone Joint Surg Br* 1991; 73(6): 922-926.
33. Lohmander LS, Roos EM. Clinical update: treating osteoarthritis. *Lancet* 2007; 370(9605): 2082-2084.
34. Lützner J, Kasten P, Günther KP, Kirschner S. Surgical options for patients with osteoarthritis of the knee. *Nat Rev Rheumatol* 2009; 5(6): 309-316.
35. Madry H, Grün UW, Knutsen G. Cartilage repair and joint preservation: medical and surgical treatment options. *Dtsch Arztebl Int* 2011; 108(40): 669-677.
36. Michael JWP, Schlüter-Brust KU, Eysel P. Epidemiologie, Ätiologie, Diagnostik und Therapie der Gonarthrose. *Dtsch Arztebl* 2010; 107(9): 152-162.

37. Moseley B. Arthroscopic surgery did not provide additional benefit to physical and medical therapy for osteoarthritis of the knee. *J Bone Joint Surg Am* 2009; 91(5): 1281.
38. Neverov VA, Lankin IV. Experience with treatment of gonarthrosis by the method of arthroscopic sanation of the joint [Russisch]. *Vestn Khir Im I I Grek* 2010; 169(4): 86-89.
39. Osborne RH, Chapman AB, McColl GJ. Management of osteoarthritis in older people. *Journal of Pharmacy Practice and Research* 2002; 32(4): 276-281.
40. Pedersen MS, Moghaddam AZ, Bak K, Koch JSS. The effect of bone drilling on pain in gonarthrosis. *Int Orthop* 1995; 19(1): 12-15.
41. Richmond J, Hunter D, Irrgang J, Jones MH, Snyder-Mackler L, Van Durme D et al. American Academy of Orthopaedic Surgeons clinical practice guideline on the treatment of osteoarthritis (OA) of the knee. *J Bone Joint Surg Am* 2010; 92(4): 990-993.
42. Ringdahl E, Pandit S. Treatment of knee osteoarthritis. *Am Fam Physician* 2011; 83(11): 1287-1292.
43. Risberg MA. Arthroscopic surgery provides no additional benefit over physiotherapy and medication for the treatment of knee osteoarthritis. *Aust J Physiother* 2009; 55(2): 137.
44. Rönn K, Reischl N, Gautier E, Jacobi M. Current surgical treatment of knee osteoarthritis. *Arthritis* 2011; 2011: Article ID 454873.
45. Seed SM, Dunican KC, Lynch AM. Osteoarthritis: a review of treatment options. *Formulary* 2009; 44(5): 143-151.
46. Shaffer B, Hanypsiak B. Nonoperative and arthroscopic approaches to the postmeniscectomy arthritic knee. *Arthroscopy* 2003; 19(10 Suppl 1): 122-128.
47. Spahn G, Wittig R. Short-term effects of different arthroscopic techniques in the treatment of chondral defects (shaving, coblation, and microfracturing). *European Journal of Trauma* 2002; 28(6): 349-354.
48. Stone KR. Response to randomized trial of arthroscopic surgery. *Arthroscopy* 2008; 24(11): 1313.
49. Stuart MJ. Arthroscopic management for degenerative arthritis of the knee. *AAOS Instructional Course Lectures* 1999; 48: 135-141.
50. Su JY, Chang JK, Lu YM, Lin SY. Arthroscopic debridement for osteoarthritis of the knee: a seven years follow-up study. *Gaoxiong Yi Xue Ke Xue Za Zhi* 1995; 11(12): 667-672.
51. Voelker R. Guideline provides evidence-based advice for treating osteoarthritis of the knee. *JAMA* 2009; 301(5): 475-476.
52. Yang KGA, Saris DBF, Dhert WJA, Verbout AJ. Osteoarthritis of the knee: current treatment options and future directions. *Curr Orthop* 2004; 18(4): 311-320.

53. Zeichen J, Hankemeier S, Knobloch K, Jagodzinski M. Die arthroskopische Meniskusteilresektion. Oper Orthop Traumatol 2006; 18(5-6): 380-392.

Nicht E6

1. Correction: "A randomized trial of arthroscopic surgery for osteoarthritis of the knee" (N Engl J Med 2008; 359(11): 1097-1107). N Engl J Med 2009; 361(20): 2004.
2. Bailey RE. Arthroscopic surgery ineffective for osteoarthritis of the knee. J Fam Pract 2002; 51(10): 813.
3. Chambers K, Schulzer M, Sobolev B. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. Arthroscopy 2002; 18(7): 686-687.
4. Fleiss DJ. Arthroscopic treatment of osteoarthritis of the knee: a prospective randomised, placebo controlled trial. Am J Sports Med 1997; 25(5): 724.
5. Franke H. Arthroskopie bei Kniegelenksarthrose als klinischer Standard obsolet! Forsch Komplementmed 2008; 15(6): 354-355.
6. Hartman GP. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo controlled trial. Am J Sports Med 1997; 25(5): 724.
7. Hodgkinson B, Merlin T, Graves S, Cleland L, Hiller JE. Therapeutic arthroscopy for the management of osteoarthritis of the knee: systematic review. Adelaide: Adelaide Health Technology Assessment; 2004.
8. Moseley JB, O'Malley K, Petersen NJ, Wray NP, Gillespie WJ. Arthroscopic surgery was not effective for relieving pain or improving function in osteoarthritis of the knee. Evid Based Med 2003; 8(2): 56.
9. Skonieczki DJ, Palmisano J, Losina E, Barraza AS, Chaisson CE, Koeth P et al. Factors associated with refusal to participate in a randomized controlled trial of surgery vs non-operative therapy for meniscal tear in knee osteoarthritis. Arthritis Rheum 2009; 60(Suppl 10): 1936.
10. Smith M, Wetherall M, Darby T, Esterman A, Roberts-Thomson P, Coleman M et al. Erratum: "A randomized placebo-controlled trial of arthroscopic lavage versus lavage plus intra-articular corticosteroids in the management of symptomatic osteoarthritis of the knee" (Rheumatology (Oxford) 2003; 42(1): 123-134). Rheumatology (Oxford) 2003; 42(6): 814.
11. Ward PJ, Ramos JL, Fernandez GN, Palazon J. A prospective randomised controlled trial of cannula versus arthroscopic lavage in patients with osteoarthritis of the knee. J Bone Joint Surg Br 1998; 80(Suppl 1): 46.
12. Wray NR, Moseley JB, O'Malley K. Arthroscopic treatment of osteoarthritis of the knee. J Bone Joint Surg Am 2003; 85(2): 381.

Anhang C – Liste der in den Stellungnahmen an den G-BA genannten Referenzen

Nicht E1

1. Herrlin S, Hållander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007; 15(4): 393-401.
2. Spahn G, Klinger HM, Mückley T, Hofmann GO. Four-year results from a randomized controlled study of knee chondroplasty with concomitant medial meniscectomy: mechanical debridement versus radiofrequency chondroplasty. *Arthroscopy* 2010; 26(9 Suppl 1): S73-S80.

Nicht E2

1. Bin SI, Lee SH, Kim CW, Kim TH, Lee DH. Results of arthroscopic medial meniscectomy in patients with grade IV osteoarthritis of the medial compartment. *Arthroscopy* 2008; 24(3): 264-268.
2. Bradley JD, Heilman DK, Katz BP, G'Sell P, Wallick JE, Brandt KD. Tidal irrigation as treatment for knee osteoarthritis: a sham-controlled, randomized, double-blinded evaluation. *Arthritis Rheum* 2002; 46(1): 100-108.
3. Dawes PT, Kirlew C, Haslock I. Saline washout for knee osteoarthritis: results of a controlled study. *Clin Rheumatol* 1987; 6(1): 61-63.
4. Englund M, Guermazi A, Roemer FW, Aliabadi P, Yang M, Lewis CE et al. Meniscal tear in knees without surgery and the development of radiographic osteoarthritis among middle-aged and elderly persons: the multicenter osteoarthritis study. *Arthritis Rheum* 2009; 60(3): 831-839.
5. Ike RW, Arnold WJ, Rothschild EW, Shaw HL. Tidal irrigation versus conservative medical management in patients with osteoarthritis of the knee: a prospective randomized study. *J Rheumatol* 1992; 19(5): 772-779.
6. Ravaud P, Moulinier L, Giraudeau B, Ayral X, Guerin C, Noel E et al. Effects of joint lavage and steroid injection in patients with osteoarthritis of the knee: results of a multicenter, randomized, controlled trial. *Arthritis Rheum* 1999; 42(3): 475-482.

Nicht E3

1. Aaron RK, Skolnick AH, Reinert SE, Ciombor DM. Arthroscopic débridement for osteoarthritis of the knee. *J Bone Joint Surg Am* 2006; 88(5): 936-943.
2. Al-Omran AS, Sadat-Ali M. Arthroscopic joint lavage in osteoarthritis of the knee: is it effective? *Saudi Med J* 2009; 30(6): 809-812.
3. Bernard J, Lemon M, Patterson MH. Arthroscopic washout of the knee: a 5-year survival analysis. *Knee* 2004; 11(3): 233-235.
4. Crevoisier X, Munzinger U, Drobny T. Arthroscopic partial meniscectomy in patients over 70 years of age. *Arthroscopy* 2001; 17(7): 732-736.

5. Dervin GF, Stiell IG, Rody K, Grabowski J. Effect of arthroscopic débridement for osteoarthritis of the knee on health-related quality of life. *J Bone Joint Surg Am* 2003; 85(1): 10-19.
6. Fond J, Rodin D, Ahmad S, Nirschl RP. Arthroscopic debridement for the treatment of osteoarthritis of the knee: 2- and 5-year results. *Arthroscopy* 2002; 18(8): 829-834.
7. Gross DE, Brenner SL, Esformes I, Gross ML. Arthroscopic treatment of degenerative joint disease of the knee. *Orthopedics* 1991; 14(12): 1317-1321.
8. Harwin SF. Arthroscopic debridement for osteoarthritis of the knee: predictors of patient satisfaction. *Arthroscopy* 1999; 15(2): 142-146.
9. Hubbard MJS. Articular debridement versus washout for degeneration of the medial femoral condyle: a five-year study. *J Bone Joint Surg Br* 1996; 78(2): 217-219.
10. Jackson RW, Dieterichs C. The results of arthroscopic lavage and debridement of osteoarthritic knees based on the severity of degeneration: a 4- to 6-year symptomatic follow-up. *Arthroscopy* 2003; 19(1): 13-20.
11. Smith MD, Wetherall M, Darby T, Esterman A, Slavotinek J, Roberts-Thomson P et al. A randomized placebo-controlled trial of arthroscopic lavage versus lavage plus intra-articular corticosteroids in the management of symptomatic osteoarthritis of the knee. *Rheumatology* 2003; 42(12): 1477-1485.
12. Suter LG, Fraenkel L, Losina E, Katz JN, Gomoll AH, Paltiel AD. Medical decision making in patients with knee pain, meniscal tear, and osteoarthritis. *Arthritis Rheum* 2009; 61(11): 1531-1538.

Nicht E4

1. Gibson JN, White MD, Chapman VM, Strachan RK. Arthroscopic lavage and debridement for osteoarthritis of the knee. *J Bone Joint Surg Br* 1992; 74(4): 534-537.

Nicht E5

1. Gonarthrose: arthroskopische Operation in Studie ohne Vorteile. *Deutsches Ärzteblatt* [online] 11.09.2008. URL: <http://www.aerzteblatt.de/v4/news/news.asp?id=33667>.
2. Bouillon B, Tiling T, Shafizadeh S. Stellenwert von arthroskopischem Debridement und Lavage bei der Therapie der Arthrose. *Dtsch Z Sportmed* 2003; 54(6): 215-217.
3. Dandy DJ. Arthroscopic debridement of the knee for osteoarthritis. *J Bone Joint Surg Br* 1991; 73(6): 877-878.
4. Marx RG. Arthroscopic surgery for osteoarthritis of the knee? *N Engl J Med* 2008; 359(11): 1169-1170.
5. Poehling GG. Degenerative arthritis arthroscopy and research. *Arthroscopy* 2002; 18(7): 683.

6. Rand JA. Role of arthroscopy in osteoarthritis of the knee. Arthroscopy 1991; 7(4): 358-363.

7. Su JY, Chang JK, Lu YM, Lin SY. Arthroscopic debridement for osteoarthritis of the knee: a seven years follow-up study. Gaoxiong Yi Xue Ke Xue Za Zhi 1995; 11(12): 667-672.

Anhang D – Liste der gesichteten systematischen und nichtsystematischen Übersichten

1. Allgood P. Arthroscopic lavage for knee osteoarthritis. London: Bazian; 2003. (STEER: Succinct and Timely Evaluated Evidence Reviews; Band 3(3)).
2. Avouac J, Vicaut E, Bardin T, Richette P. Efficacy of joint lavage in knee osteoarthritis: meta-analysis of randomized controlled studies. *Rheumatology (Oxford)* 2010; 49(2): 334-340.
3. Bazian. Arthroscopic lavage for osteoarthritis of the knee. *Evidence-Based Healthcare & Public Health* 2005; 9(3): 192-196.
4. Campbell MK, Skea ZC, Sutherland AG, Cuthbertson BH, Entwistle VA, McDonald AM et al. Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo-controlled trial (the KORAL study). *Health Technol Assess* 2010; 14(5): 1-180.
5. Laupattarakasem W, Laopaiboon M, Laupattarakasem P, Sumananont C. Arthroscopic debridement for knee osteoarthritis. *Cochrane Database Syst Rev* 2008; (1): CD005118.
6. Misso ML, Pitt VJ, Jones KM, Barnes HN, Piterman L, Green SE. Quality and consistency of clinical practice guidelines for diagnosis and management of osteoarthritis of the hip and knee: a descriptive overview of published guidelines. *Med J Aust* 2008; 189(7): 394-399.
7. Pendleton A, Arden N, Dougados M, Doherty M, Bannwarth B, Bijlsma JWJ et al. EULAR-Richtlinien für die Behandlung von Kniegelenksarthrose: Bericht der Arbeitsgruppe des Standing Committees for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Journal für Mineralstoffwechsel* 2003; 10(3): 23-31.
8. Reichenbach S, Rutjes AW, Nuesch E, Trelle S, Juni P. Joint lavage for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2010; (5): CD007320.
9. Reichenbach S, Rutjes AWS, Nüesch E, Trelle S, Jüni P. Arthroscopic lavage for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2008; (3): CD007320.
10. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R et al. Treatment of osteoarthritis of the knee (nonarthroplasty). *J Am Acad Orthop Surg* 2009; 17(9): 591-600.
11. Samson DJ, Grant MD, Ratko TA, Bonnell CJ, Ziegler KM, Aronson N. Treatment of primary and secondary osteoarthritis of the knee: AHRQ publication no. 07-E012 [online]. 09.2007 [Zugriff: 16.12.2011]. (Evidence Report/Technology Assessment; Band 157). URL: <http://www.ahrq.gov/downloads/pub/evidence/pdf/oaknee/oaknee.pdf>.
12. Medical Advisory Secretariat. Arthroscopic lavage and debridement for osteoarthritis of the knee: an evidence-based analysis. Toronto: MAS; 2005. (Ontario Health Technology Assessment Series; Band 5(12)). URL: http://www.health.gov.on.ca/english/providers/program/mas/tech/reviews/pdf/rev_lavdeb_090105.pdf.

13. Siparsky P, Ryzewicz M, Peterson B, Bartz R. Arthroscopic treatment of osteoarthritis of the knee: are there any evidence-based indications? Clin Orthop Relat Res 2007; 455: 107-112.
14. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al. OARSI recommendations for the management of hip and knee osteoarthritis; part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis Cartilage 2007; 15(9): 981-1000.
15. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al. OARSI recommendations for the management of hip and knee osteoarthritis; part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008; 16(2): 137-162.

Anhang E – Ergebnisparameter und Messinstrumente**Beschreibung der berücksichtigten Ergebnisparameter und Messinstrumente**

Tabelle 46: Ergebnisparameter und Messinstrumente, die in den ausgewerteten Studien verwendet und für den vorliegenden Bericht berücksichtigt wurden

Instrument	Erläuterung / Kommentar
Arthritis Impact Measurement Scales (AIMS)	Der AIMS-Fragebogen ist ein arthrosespezifisches Selbstbeurteilungsinstrument zur Messung des physischen, sozialen und emotionalen Wohlbefindens. Er besteht aus den 9 Subskalen Mobilität, körperliche Aktivität, Geschicklichkeit, Haushaltsaktivitäten, soziale Aktivitäten, Aktivitäten des täglichen Lebens, Schmerz, Depression und Ängstlichkeit. AIMS umfasst insgesamt 64 Items, darunter 45 spezifische Fragen zum Gesundheitszustand. Für jede Subskala ergibt sich ein Score von 0 bis 10, aus dem ein Gesamtscore von 0 bis 60 errechnet werden kann. Niedrigere Werte stehen dabei für einen besseren Gesundheitszustand [111,112].
Arthritis Impact Measurement Scales, Modifikation (AIMS-2)	Der AIMS-2 ist eine Modifikation von AIMS mit Erweiterung um die Teilbereiche Armfunktion, soziale Unterstützung und Arbeit. Das Instrument umfasst 12 Subskalen mit insgesamt 78 Items. Die Berechnung und Interpretation der Scores erfolgt analog AIMS, Daten zur Reliabilität sind beschrieben [112].
Hospital for Special Surgery Knee Rating Score (HSS)	Der Hospital for Special Surgery Knee Rating Score ist ein Selbst- und Fremdbeurteilungsinstrument zur Bewertung der Kniefunktion nach Knieoperationen und wurde in der Studie Merchan 1993 eingesetzt. Er besteht aus einem direkt vom Patienten berichteten (Schmerz, Instabilität, Gehhilfen und Gehstrecke) und einem durch den Arzt erhobenen (Streckung, Beugung und Erguss) Teil. Für diese 7 Dimensionen werden unterschiedlich hohe Scores vergeben. Der Gesamtscore beträgt maximal 100 (60 subjektiver, 40 objektiver Teil). Eine höhere Punktzahl steht für einen besseren Zustand.
Lequesne-Index	Der Lequesne-Index ist ein gebräuchliches, validiertes Instrument zur Messung von Schmerz und Funktionseinschränkungen bei Gonarthrose und Coxarthrose. Er ist zur interviewergestützten Anwendung entwickelt worden. Er umfasst die 3 Bereiche Schmerzen (5 Items), Gehstrecke (ein Item) und Aktivitäten des täglichen Lebens (4 Items). Für jeden Bereich kann ein Score von 0 (keine Schmerzen oder Einschränkungen) bis 8 vergeben werden (Gesamtscore: 0–24). Für Knie- und Hüftgelenk existieren unterschiedliche Versionen. Sie sind dem jeweils typischen Beschwerdebild angepasst und unterscheiden sich hinsichtlich der Items zu Aktivitäten des täglichen Lebens [113].
Lysholm Score	Der Lysholm Score ist ein Selbstbeurteilungsinstrument und dient der Messung von Instabilität, Funktions- und Bewegungseinschränkungen, Schmerzen sowie Schwellung. Er wird bei Patienten mit Knieband-Verletzungen und Knieinstabilität nach Bandoperationen verwendet [114,115]. Der Lysholm Score besteht aus 8 Items, für die unterschiedlich hohe Scores vergeben werden [116]. Der Gesamtscore reicht von 0 bis 100 (100 bedeutet „keine Einschränkungen oder Symptome“). Das Ergebnis wird den Kategorien exzellent (95–100), gut (84–94), ausreichend (65–83) und schlecht (≤ 64) zugeordnet. Ursprünglich war die Anwendung durch einen Untersucher intendiert. Reliabilität und Validität sind beschrieben [115].

(Fortsetzung)

Tabelle 46: Ergebnisparameter und Messinstrumente, die in den ausgewerteten Studien verwendet und für den vorliegenden Bericht berücksichtigt wurden (Fortsetzung)

Instrument	Erläuterung / Kommentar
Verbesserung der Knie Symptome (Einschätzung Patient)	In der Studie von Arden (2008) wurden die Veränderungen bezüglich der Kniesymptome nach der Intervention gegenüber dem vorherigen Zustand erfasst. Dazu wurde eine 5-Punkte-Likert-Skala verwendet („viel schlechter“ bis „viel besser“). Es wurde nicht beschrieben, wie die Kategorien der verwendeten Skala im Einzelnen und wie eine „Verbesserung“ definiert wurden. Für die Bewertung wurde davon ausgegangen, dass der neutrale Wert in der Mitte keine Verbesserung darstellte.
Durchschnittliche Schmerzintensität	Für den Endpunkt Schmerz wurden mittels eines Fragebogens die Beurteilung der durchschnittlichen Schmerzintensität der vorangegangenen Woche durch die Patienten erfasst.
Knee-Specific Pain Scale (KSPS, jetzt J-MAP)	Beim KSPS handelt es sich um ein kniespezifisches Selbstbeurteilungsinstrument zur Erfassung von Schmerz, das für die vorliegende Studie entwickelt wurde. Es umfasst insgesamt 12 Items, davon 4 Items zur Schmerzintensität, 3 zur Schmerzfrequenz und 5 zur schmerzbezogenen Zufriedenheit (<i>pain distastefulness</i>). Die Items haben unterschiedlich hohe Scores und werden nach einem bestimmten Algorithmus zu einem Gesamtscore zwischen 0 und 100 zusammengefasst. Höhere Werte stehen dabei für stärkere Schmerzen. Dieses Instrument wurde validiert [81].
Visuelle Analogskala (VAS) – Schmerz	Bei der visuellen Analogskala handelt es sich um eine Selbstbeurteilungsskala. Sie misst die Intensität eines Zustandes in der Einschätzung eines Patienten auf einem Kontinuum zwischen 0 (keine Beschwerden) und 100 (extreme Beschwerden). Der Patient markiert dafür auf einer visualisierten horizontalen Linie einen Punkt, der seiner subjektiven Einschätzung entspricht. Die Schmerz-Version der VAS ist gut und an unterschiedlichen Patientenkollektiven untersucht [117]. Eine minimale klinisch relevante Verbesserung (MCID) wurde in Studien validiert und liegt zwischen 11,1 und 19,9 [118,119]
Knee Society Clinical Rating System Function Score (KSCRS)	Der KSCRS Function Score ist ein Fremdbeurteilungsinstrument zur Einschätzung der Kniefunktion nach Arthroplastien. Er ist unterteilt in einen Score für das Kniegelenk (Schmerz, Bewegungsausmaß und Stabilität) und einen Score für die allgemeine physische Funktionsfähigkeit (Gehen und Treppensteigen). Der maximale Gesamtscore beträgt 200. Der Höchstwert von 100 für das Kniegelenk wird erreicht, wenn der Bewegungsumfang mindestens 125 Grad beträgt und das Knie keine Schmerzen, Instabilität und Achsenabweichung aufweist. Den Höchstwert für die physische Funktionsfähigkeit von 100 erreicht ein Patient, der eine unbegrenzte Strecke gehen und Treppen ohne Festhalten hinauf- und hinuntersteigen kann [120].
Knee Injury and Osteoarthritis Outcome Scale (KOOS)	Der Knee Injury and Osteoarthritis Outcome Scale ist ein frei verfügbares Selbstbeurteilungsinstrument zur Bewertung von Knieproblemen und wird häufig in Studien mit Fragestellungen zu Knieverletzungen und Osteoarthritis eingesetzt. Er besteht aus 5 Subskalen mit Verwendung von Likert-Skalen. In der Studie Katz 2013 wurde die Subskala Schmerz verwendet. Der KOOS ist validiert [121]. Der KOOS beinhaltet den WOMAC und dessen Werte können daraus berechnet werden. Ein „Minimal clinically Important Change“ von 8-10 wird als relevant angesehen.

(Fortsetzung)

Tabelle 46: Ergebnisparameter und Messinstrumente, die in den ausgewerteten Studien verwendet und für den vorliegenden Bericht berücksichtigt wurden (Fortsetzung)

Instrument	Erläuterung / Kommentar
Knee Injury and Osteoarthritis Outcome Scale (KOOS) – Schmerz	Der KOOS-Schmerz ist eine von 5 Subskalen des KOOS und besteht aus 9 Items. Es wird ein normalisierter Score zwischen 0 und 100 vergeben, wobei 0 für extreme Knieprobleme steht. In der Studie Katz 2013 wird hingegen definiert, dass ein höherer Wert mehr Schmerzen bedeutet.
Physical functioning scale (PFS)	Die physical functioning scale ist ein im Rahmen der Studie Moseley 2002 entwickeltes Instrument zur objektiven Erfassung der Gangfunktion. Dabei wird die Zeit gemessen, die ein Patient benötigt, um eine 30 Meter lange Gehstrecke und eine vorgegebene Anzahl Stufen treppauf und treppab zu bewältigen.
Short Form 36 Health Survey (SF-36)	Der SF-36 ist ein generisches Selbstbeurteilungsinstrument zur Messung der Lebensqualität. Die Skala mit 36 Items beinhaltet die 8 Subskalen Körperfunktion, körperliche Rollenfunktion, Schmerz, allgemeine Gesundheitswahrnehmung, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden. Für jede Dimension wird ein Score zwischen 0 und 100 vergeben, wobei ein höherer Wert für einen besseren Gesundheitszustand steht. Der SF-36 ist validiert [122].
Short Form 36 Health Survey (SF-36) – Körperfunktion	Der SF-36-Körperfunktion ist eine von 8 Subskalen des SF-36 und umfasst 10 Items.
Short Form 36 Health Survey (SF-36) – Schmerz	Der SF-36-Schmerz ist eine von 8 Subskalen des SF-36 und besteht aus 2 Items.
Western Ontario and McMaster Universities Arthritis Index (WOMAC)	Der WOMAC ist ein validiertes Selbstbeurteilungsinstrument zur Erfassung von körperlichen Funktionseinschränkungen bei Patienten mit Gonarthrose oder Coxarthrose. Dieser Index beinhaltet 3 Subskalen (Schmerz, Steifigkeit, körperliche Funktionsfähigkeit) mit insgesamt 24 Items. Es existieren 3 verschiedene Versionen des WOMAC, bei denen die Items entweder anhand einer visuellen Analogskala (VAS, 0–100), einer Likert-Skala (5 Kategorien) oder anhand einer numerischen Rating Skala (NRS, 0–10) beantwortet werden. Je nach Version reichen mögliche Gesamtscores von 0 bis 2400 (VAS), 0 bis 94 (Likert-Skala) oder 0 bis 240 (NRS). Höhere Werte stehen dabei für einen schlechteren Gesundheitszustand [57].
Western Ontario and McMaster Universities Arthritis Index (WOMAC) – Körperfunktion	Der WOMAC Körperfunktion ist eine von 3 Subskalen des WOMAC. Sie besteht aus 17 Items. Der Score dieser Subskala reicht von 0 bis 68 (Likert-Version), 0 bis 1700 (VAS-Version) oder 0-100 (normalisierte Daten) [115]. Für die Version der normalisierten Daten (0-10) wurde eine minimale klinisch relevante Verbesserung von 0,79 beschrieben [123]

(Fortsetzung)

Tabelle 46: Ergebnisparameter und Messinstrumente, die in den ausgewerteten Studien verwendet und für den vorliegenden Bericht berücksichtigt wurden (Fortsetzung)

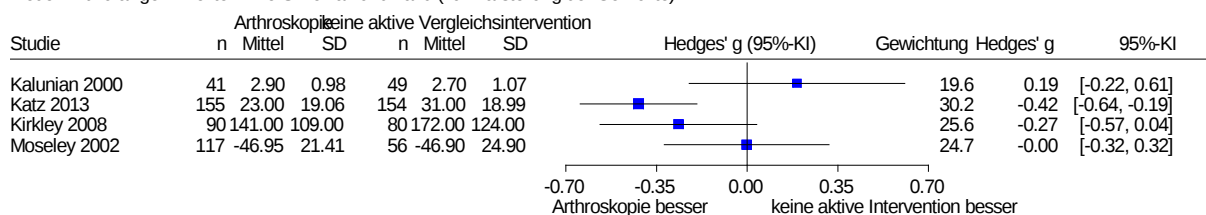
Instrument	Erläuterung / Kommentar
Western Ontario and McMaster Universities Arthritis Index (WOMAC) – Steifigkeit	Beim WOMAC Steifigkeit handelt es sich um eine von 3 Subskalen des WOMAC. Diese Subskala enthält 2 Items. Im Likert-Format liegen mögliche Scores bei 0 bis 8, im VAS-Format bei 0 bis 200 [115].
Western Ontario and McMaster Universities Arthritis Index (WOMAC) – Schmerz	Der WOMAC Schmerz ist eine von 3 Subskalen des WOMAC. Sie umfasst 5 Items. Im Likert-Format liegen mögliche Scores bei 0 bis 20 und bei der VAS-Skala bei 0 bis 500 [115].
McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR)	Der MACTAR ist ein interviewergestütztes Instrument zur Erfassung von Symptomen und Funktionseinschränkungen bei Patienten mit rheumatoider Arthritis und Arthrose. Der erste Abschnitt (priority function section) erfasst die krankheitsbedingt eingeschränkten Aktivitäten. Der Patient kann diese Aktivitäten nach seinen Präferenzen gewichten. Außerdem enthält das Instrument 4 globale Fragen zum physischen, sozialen und emotionalen Wohlbefinden. Es existieren 2 aufeinander aufbauende Versionen des MACTAR. Die Baseline-Version umfasst 19 Items. Für den ersten Abschnitt wird kein Score vergeben. Die globalen Fragen erhalten einen Score zwischen 5 und 26. Die Nachfolge-Version besteht aus einer variablen Itemzahl, abhängig von den Antworten hinsichtlich Veränderungen der Symptomatik. Der erste Abschnitt kann Scores von 15 bis 45 (bei Präferenzen) oder von 5 bis 15 (ohne Präferenzen) erhalten. Für die globalen Fragen wird ein Score zwischen 11 und 47 vergeben. Höhere Werte stehen für einen besseren Zustand [124]. (Kirkley verwendet den MACTAR abweichend vom hier beschriebenen Vorgehen mit einem Score von 0 bis 500 und umgekehrter Interpretation der Werte (höhere Werte stehen für größere Einschränkungen).)
Standard Gamble utility score	Das Standard-Gamble-Verfahren ist eine gebräuchliche Methode der nutzentheoretischen Lebensqualitätsmessung. Die Probanden werden dabei vor eine (hypothetische) Wahl zwischen 2 Möglichkeiten gestellt. Die eine Möglichkeit ist der Gesundheitszustand, der bewertet werden soll. Die Alternative ist eine mögliche medizinische Intervention mit unsicherem Ausgang. Die Intervention kann entweder mit einer Wahrscheinlichkeit von p zur Heilung oder mit einer Wahrscheinlichkeit von 1-p zum Tod führen. Die Wahrscheinlichkeiten werden so lange verändert, bis der Patient beide Alternativen als gleichwertig betrachtet [125]. Die Wahrscheinlichkeit p entspricht dem Nutzwert (utility score) für den zu bewertenden Gesundheitszustand. Ein Score von 0 steht für den schlechtesten vorstellbaren Gesundheitszustand (zumeist Tod), ein Score von 1 für bestmögliche Gesundheit [126].
Hospital Anxiety and Depression Scale (HADS) – Depression	Die Hospital Anxiety and Depression Scale ist ein Selbstbeurteilungsinstrument zur Erfassung von Angst und Depression bei nicht psychiatrischen Patienten im ambulanten Bereich [127]. Die Subskala Depression besteht aus 7 Items, aus denen ein Score von 0 bis 21 errechnet wird. Ein niedriger Score steht dabei für eine geringer ausgeprägte Symptomatik. Daten zu Gütekriterien sind vorhanden.
Hospital Anxiety and Depression Scale (HADS) – Ängstlichkeit	Die Subskala Ängstlichkeit des HADS besteht aus 7 Items. Die Berechnung der Scores erfolgt wie beim HADS Depression (s. o.). Daten zu Gütekriterien sind vorhanden [127].

Anhang F – Sensitivitätsanalysen: Auswertung zusätzlicher Skalen

Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 3 Monate

Schmerz

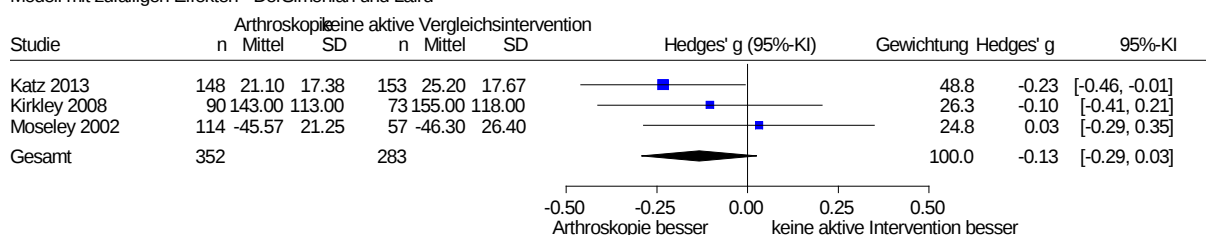
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 6 Monate

Schmerz

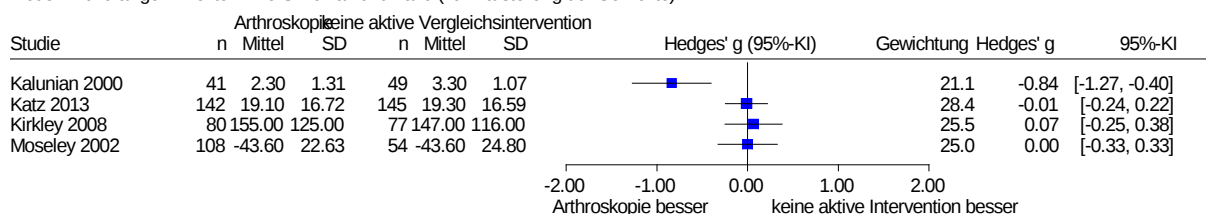
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 12 Monate

Schmerz

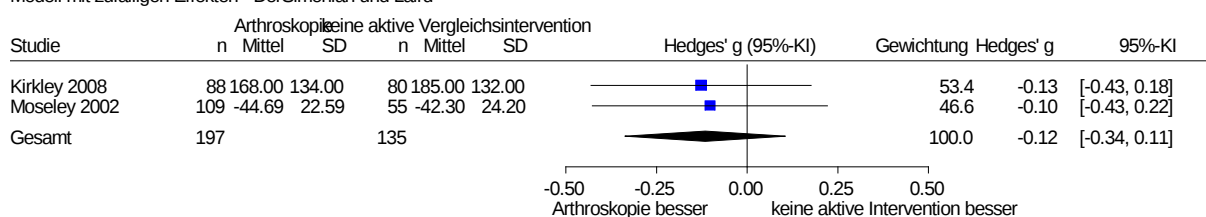
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 24 Monate

Schmerz

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 6, 12 und 24 Monate. Verwendung alternativer Scores (wenn möglich): VAS Schmerz (Kalunian 2000), KOOS Schmerz (Katz 2013), WOMAC Schmerz (Kirkley 2008), SF36 Schmerz (Moseley 2002).

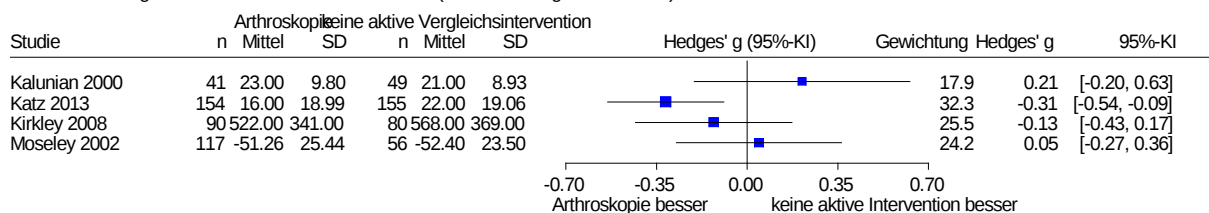
KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 9: Auswertung zusätzlicher Skalen zum Endpunkt Schmerz, Vergleich keine aktive Vergleichsintervention

Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 3 Monate

Körperliche Funktion

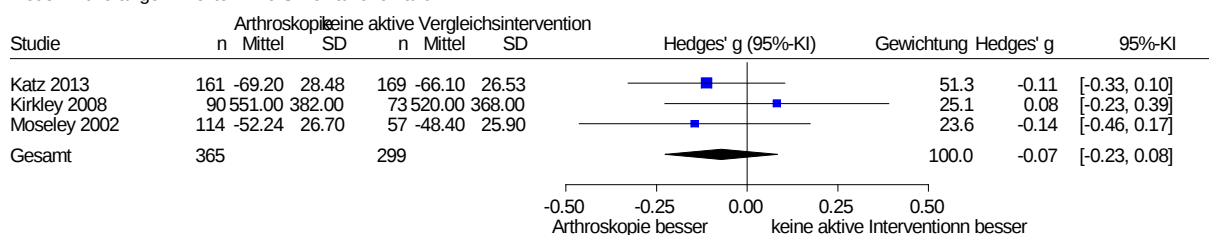
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 6 Monate

Körperliche Funktion

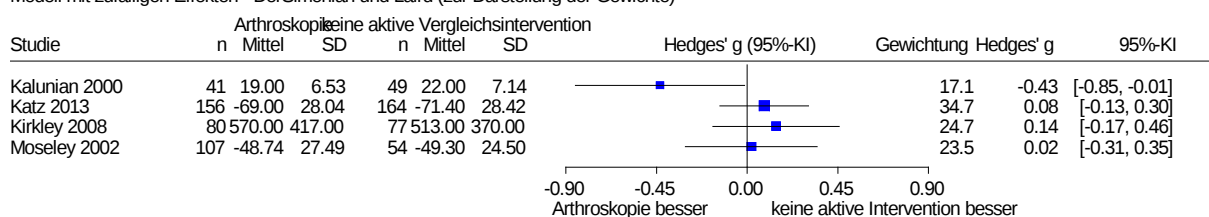
Modell mit zufälligen Effekten - DerSimonian und Laird



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 12 Monate

Körperliche Funktion

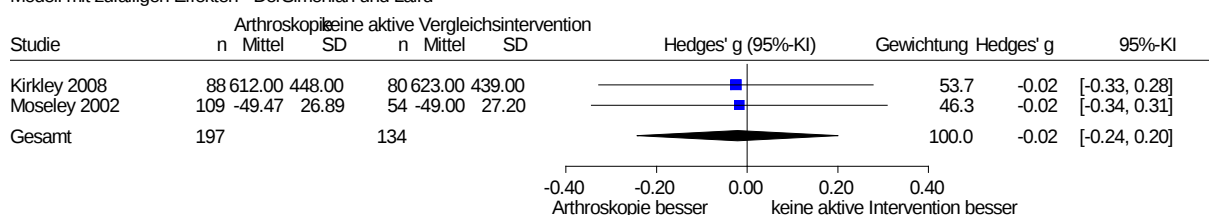
Modell mit zufälligen Effekten - DerSimonian und Laird (zur Darstellung der Gewichte)



Arthroskopie vs. keine aktive Vergleichsintervention - alternative Scores, 24 Monate

Körperliche Funktion

Modell mit zufälligen Effekten - DerSimonian und Laird



Zeitpunkte: 3, 6, 12 und 24 Monate. Verwendung alternativer Scores (wenn möglich): WOMAC körperliche Funktion (Kalunian 2000, Kirkley 2008), SF36 körperliche Funktion (Katz 2013, Moseley 2002).

KI: Konfidenzintervall; SD: Standardabweichung

Abbildung 10: Auswertung zusätzlicher Skalen zum Endpunkt körperliche Funktion, Vergleich keine aktive Vergleichsintervention