



IQWiG-Berichte – Nr. 1224

Kognitive Verhaltenstherapie für Insomnie

**Evidenzbericht zur S3-Leitlinie Nicht
erholsamer Schlaf / Schlafstörungen –
Insomnie bei Erwachsenen**

Evidenzbericht

Auftrag: V20-07B
Version: 1.0
Stand: 25.10.2021

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Kognitive Verhaltenstherapie für Insomnie – Evidenzbericht zur S3-Leitlinie Nicht
erholsamer Schlaf / Schlafstörungen – Insomnie bei Erwachsenen

Auftraggeber

Bundesministerium für Gesundheit

Datum des Auftrags

18.12.2020

Interne Auftragsnummer

V20-07B

Anschrift des Herausgebers

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

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Dieser Bericht wurde ohne die Beteiligung externer Sachverständiger erstellt.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Marc Krause
- Tatjana Herrmanns
- Fabian Lotz
- Claudia Mischke
- Julia Ströhlein
- Sarah Thys

Schlagwörter

Ein- und Durchschlafstörungen, Kognitive Therapie, Systematische Übersicht

Keywords

Sleep Initiation and Maintenance Disorders, Cognitive Behavioral Therapy, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abbildungsverzeichnis	vii
Abkürzungsverzeichnis.....	viii
1 Hintergrund	1
2 Fragestellung	2
3 Projektverlauf	3
3.1 Zeitlicher Verlauf des Projekts	3
4 Methoden	4
4.1 Kriterien für den Einschluss von Studien.....	4
4.1.1 Population.....	4
4.1.2 Prüf- und Vergleichsintervention	4
4.1.3 Endpunkte.....	4
4.1.4 Studientypen	4
4.1.5 Studiendauer	5
4.1.6 Publikationssprache	5
4.1.7 Publikationszeitraum	5
4.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss	5
4.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllten	5
4.2 Informationsbeschaffung.....	6
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten.....	6
4.2.2 Fokussierte Informationsbeschaffung von Studien	6
4.2.3 Selektion relevanter Studien.....	7
4.3 Informationsdarstellung und Synthese	7
4.3.1 Darstellung der Einzelstudien.....	7
4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial).....	9
4.3.3 Metaanalysen	9
4.3.4 Bewertung der Qualität der Evidenz	10
4.3.4.1 Abwertung der Qualität der Evidenz	10
4.3.4.2 Aufwertung der Qualität der Evidenz	12
5 Ergebnisse.....	13
5.1 Ergebnisse der Informationsbeschaffung	13
5.2 Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien	15

5.3	Übersicht der Endpunkte	26
5.4	Ergebnisse zu Endpunkten.....	27
5.4.1	Ergebnisse zum Endpunkt Krankheitsschwere	27
5.4.2	Ergebnisse zum Endpunkt Schlafeffizienz	29
5.4.2.1	Metaanalysen	33
5.4.2.2	Subgruppenmerkmale und andere Effektmodifikatoren.....	34
5.4.3	Ergebnisse zum Endpunkt Einschlaf latenz	35
5.4.4	Ergebnisse zum Endpunkt Wachzeit nach dem Einschlafen.....	39
5.4.5	Ergebnisse zum Endpunkt Schlafqualität.....	41
5.4.6	Ergebnisse zum Endpunkt Schlafdauer	42
5.4.6.1	Metaanalysen	46
5.4.6.2	Subgruppenmerkmale und andere Effektmodifikatoren.....	47
6	Anmerkungen zu den Ergebnissen	48
7	Literatur	49
8	Studienlisten	52
8.1	Liste der identifizierten systematischen Übersichten	52
8.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	52
Anhang A	Bewertung der Studienlimitationen.....	56
Anhang B	Suchstrategien.....	59
B.1	Bibliografische Datenbanken	59
B.2	Studienregister	61

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	5
Tabelle 2: Studienpool der Evidenzdarstellung	15
Tabelle 3: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien	16
Tabelle 4: Charakterisierung der Interventionen in den eingeschlossenen Studien.....	18
Tabelle 5: Ein- / Ausschlusskriterien für Patientinnen und Patienten in den Studien.....	22
Tabelle 6: Charakterisierung der Studienpopulationen.....	25
Tabelle 7: Matrix der Endpunkte	26
Tabelle 8: Evidenzprofil zum Endpunkt Krankheitsschwere (Insomnia Severity Index < 15) – Interventionsende	27
Tabelle 9: Evidenzprofil zum Endpunkt Krankheitsschwere (Insomnia Severity Index) – Interventionsende.....	28
Tabelle 10: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Schlaftagebuch – Interventionsende.....	29
Tabelle 11: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Schlaftagebuch – Nachbeobachtungszeitpunkt.....	30
Tabelle 12: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Polysomnografie – Interventionsende.....	31
Tabelle 13: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Polysomnografie – Nachbeobachtungszeitpunkt.....	32
Tabelle 14: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Schlaftagebuch – Interventionsende.....	35
Tabelle 15: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Schlaftagebuch – Nachbeobachtungszeitpunkt.....	36
Tabelle 16: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Polysomnografie – Interventionsende.....	37
Tabelle 17: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Polysomnografie – Nachbeobachtungszeitpunkt.....	38
Tabelle 18: Evidenzprofil zum Endpunkt Wachzeit nach dem Einschlafen (Minuten) – Schlaftagebuch – Interventionsende	39
Tabelle 19: Evidenzprofil zum Endpunkt Wachzeit nach dem Einschlafen (Minuten) – Polysomnografie – Interventionsende	40
Tabelle 20: Evidenzprofil zum Endpunkt Schlafqualität (Pittsburgh Schlafqualitäts-Index) – Interventionsende.....	41
Tabelle 21: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Schlaftagebuch – Interventionsende.....	42
Tabelle 22: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Schlaftagebuch – Nachbeobachtungszeitpunkt.....	43
Tabelle 23: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Polysomnografie – Interventionsende.....	44

Tabelle 24: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Polysomnografie – Nachbeobachtungszeitpunkt	45
Tabelle 25: Studienlimitationen zum Endpunkt Krankheitsschwere	56
Tabelle 26: Studienlimitationen zum Endpunkt Schlaffeffizienz.....	56
Tabelle 27: Studienlimitationen zum Endpunkt Einschlaf latenz	57
Tabelle 28: Studienlimitationen zum Endpunkt Wachzeit nach dem Einschlafen	57
Tabelle 29: Studienlimitationen zum Endpunkt Schlafqualität.....	58
Tabelle 30: Studienlimitationen zum Endpunkt Schlafdauer.....	58

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung	14
Abbildung 2: Metaanalyse, Schlaffeffizienz in % (Schlafstagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	33
Abbildung 3: Metaanalyse, Schlaffeffizienz in % (Polysomnografie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	33
Abbildung 4: Subgruppenanalyse nach Wirkstoffgruppe, Schlaffeffizienz in % (Schlafstagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	34
Abbildung 5: Subgruppenanalyse nach Wirkstoffgruppe, Schlaffeffizienz in % (Polysomnografie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	34
Abbildung 6: Metaanalyse, Schlafdauer in Minuten (Schlafstagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	46
Abbildung 7: Metaanalyse, Schlafdauer in Minuten (Polysomnografie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD.....	46
Abbildung 8: Subgruppenanalyse nach Wirkstoffgruppe, Schlafdauer in Minuten (Schlafstagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD	47
Abbildung 9: Subgruppenanalyse nach Wirkstoffgruppe, Schlafdauer in Minuten (Polysomnografie) – Interventionsende, KVT-I vs. Pharmakotherapie ; Effektmaß: MWD	47

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
BMG	Bundesministerium für Gesundheit
DGSM	Deutsche Gesellschaft für Schlafforschung und Schlafmedizin e. V.
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders – version IV (Diagnostisches und Statistisches Manual Psychischer Störungen -Version IV)
DVG	Digitale-Versorgung-Gesetz
GRADE	Grading of Recommendations Assessment, Development and Evaluation (Bewertung von Empfehlungen, Entwicklung und Auswertung)
ICSD-3	International Classification of Sleep Disorders - version 3 (Internationale Klassifikation der Schlafstörungen - Version 3)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KVT-I	Kognitive Verhaltenstherapie für Insomnie
MWD	Mittelwertdifferenz
OR	Odds Ratio
PICO	Population-Intervention-Comparison-Outcome (Akronym für Studienteilnehmer, Interventionsgruppe, Vergleichsgruppe, Endpunkt)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	absolute Risikodifferenz
SMD	standardisierte Mittelwertdifferenz

1 Hintergrund

Auf Basis des am 19.12.2019 in Kraft getretenen Digitale-Versorgung-Gesetzes (DVG) kann die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) dem Bundesministerium für Gesundheit (BMG) Leitlinien vorschlagen, bei denen das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) gemäß SGB V (§§ 139a Abs. 3 Nr. 3, 139b Abs. 6) deren Entwicklung oder Aktualisierung mit Evidenzrecherchen unterstützen soll [1].

Der vorliegende Auftrag umfasst die Beantwortung von Fragestellungen zur Aktualisierung der interdisziplinären S3-Leitlinie „Nicht erholsamer Schlaf/Schlafstörungen – Insomnie bei Erwachsenen“ (Registernummer der AWMF: 063-003) [2].

2 Fragestellung

Folgende Fragestellung wurde untersucht:

- Darstellung von Evidenz bezüglich der Effekte einer Behandlung mit kognitiver Verhaltenstherapie für Insomnie (KVT-I) im Vergleich zu Benzodiazepinen oder Benzodiazepinrezeptoragonisten bei erwachsenen Patientinnen und Patienten mit einer diagnostizierten nichtorganischen Insomnie.

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Das IQWiG wurde am 18.12.2020 vom BMG beauftragt, die Deutsche Gesellschaft für Schlafforschung und Schlafmedizin e. V. (DGSM) bei der Aktualisierung einer interdisziplinären S3-Leitlinie „Nicht erholsamer Schlaf / Schlafstörungen – Insomnie bei Erwachsenen“ [2] zu unterstützen.

Insgesamt wurden 2 Population-Intervention-Comparison-Outcome(PICO)-Fragen von den Leitlinienkoordinatoren in Abstimmung mit Patientenvertreterinnen und Patientenvertretern und mit Beratung durch die AWMF formuliert, zu denen das IQWiG jeweils einen Evidenzbericht erstellt hat. Die Projektbearbeitung im IQWiG begann mit dem Kick-off-Treffen, an dem die Leitlinienkoordinatorinnen und -koordinatoren, eine Ansprechpartnerin der AWMF und Ansprechpartnerinnen und -partner des IQWiG teilnahmen. Das Kick-off-Treffen hat am 17.03.2021 stattgefunden. Diese Evidenzberichte wurden an die Koordinatorinnen und Koordinatoren der Leitlinienerstellung sowie an die zuständige Ansprechpartnerin für die Leitlinie bei der AWMF jeweils nach Fertigstellung übermittelt.

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

4 Methoden

Die PICO-Fragestellung wurde durch die Darstellung von Evidenzprofilen beantwortet. Die Erstellung der Evidenzprofile erfolgte auf Grundlage der methodischen Vorgaben von Grading of Recommendations Assessment, Development and Evaluation (GRADE) [3] und ist konform zu den Allgemeinen Methoden [4].

4.1 Kriterien für den Einschluss von Studien

4.1.1 Population

In die Evidenzdarstellung wurden Studien mit erwachsenen Patientinnen und Patienten mit einer diagnostizierten nicht organischen Insomnie aufgenommen.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention stellte die Behandlung mit KVT-I dar.

Als Vergleichsintervention galten Benzodiazepine oder Benzodiazepinrezeptoragonisten entsprechend der Zulassung in Deutschland.

4.1.3 Endpunkte

Die Darstellung der Evidenzgrundlage für die Anwendung KVT-I erfolgte anhand von Endpunkten, die durch die Leitliniengruppe festgelegt worden sind:

- Krankheitsschwere (Insomnia Severity Index)
- Schlaffeffizienz
- Einschlaf latenz
- Wachzeit nach dem Einschlafen
- morgendliches Erwachen
- Schlafqualität
- Schlafdauer

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 Evidenzdarstellung einer medizinischen Intervention.

Für alle in Abschnitt 4.1.2 genannten Interventionen und alle in 4.1.3 genannten Endpunkte war eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

Für den zu erstellenden Evidenzbericht flossen daher RCTs als relevante wissenschaftliche Literatur in die Evidenzdarstellung ein.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer bestand keine Einschränkung.

4.1.6 Publikationssprache

Die Publikation musste in deutscher oder englischer Sprache verfasst sein.

4.1.7 Publikationszeitraum

Hinsichtlich des Publikationszeitraums bestand keine Einschränkung.

4.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 sind die Kriterien aufgelistet, die Studien erfüllen mussten, um in die Evidenzdarstellung eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Population: Erwachsene Patientinnen und Patienten mit einer diagnostizierten nicht organischen Insomnie (siehe auch Abschnitt 4.1.1)
E2	Prüfintervention: Behandlung mit KVT-I (siehe auch Abschnitt 4.1.2)
E3	Vergleichsintervention: Behandlung mit Benzodiazepinen oder Benzodiazepinrezeptoragonisten (siehe auch Abschnitt 4.1.2)
E4	Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	Studientypen: RCTs, wie in Abschnitt 4.1.4 formuliert
E6	Publikationssprache: Deutsch oder Englisch (siehe auch Abschnitt 4.1.6)
E7	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [5] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [6] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; KVT-I: Kognitive Verhaltenstherapie für Insomnie; RCT: Randomized controlled trial	

4.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllten

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reichte es aus, wenn bei mindestens ca. 80 % der eingeschlossenen Patientinnen und Patienten diese Kriterien erfüllt waren. Lagen für solche Studien Subgruppenanalysen für Patientinnen und Patienten vor, die die Einschlusskriterien erfüllten, sollte auf diese Analysen

zurückgegriffen werden. Studien, bei denen die Einschlusskriterien E1, E2 und E3 bei weniger als ca. 80 % erfüllt waren, konnten ggf. unberücksichtigt bleiben.

4.2 Informationsbeschaffung

4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung der Projektskizze erfolgte eine Recherche nach systematischen Übersichten in der Datenbank MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), HTA Database, sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ). Die Suche wurde hinsichtlich des Publikationsdatums ab 2011 eingeschränkt. Die Suche fand am 05.05.2021 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang B.

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

Es wurde geprüft, ob mindestens eine hochwertige und aktuelle systematische Übersicht infrage kam, deren Informationsbeschaffung als Grundlage für die Evidenzdarstellung verwendet werden konnte. Dafür erfolgte eine Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersicht(en). Konnte mindestens eine diesbezüglich hochwertige und aktuelle Übersicht identifiziert werden, wurden die zugrunde liegenden Studien beziehungsweise Dokumente von 1 Person auf ihre Relevanz für die vorliegende Evidenzdarstellung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der Studienlimitationen bzw. der Studienqualität der eingeschlossenen Studien oder die Datenextraktion aus den systematischen Übersichten wurden nicht übernommen.

Die finale Entscheidung, welche systematische(n) Übersicht(en) für die Evidenzdarstellung herangezogen wurden, erfolgte nach Erstellung der Projektskizze anhand der darin festgelegten Kriterien.

4.2.2 Fokussierte Informationsbeschaffung von Studien

Bei einer systematischen Recherche nach relevanten Studien beziehungsweise Dokumenten wurden folgende primäre und weitere Informationsquellen sowie Suchtechniken berücksichtigt:

Primäre Informationsquellen

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

Weitere Informationsquellen und Suchtechniken

- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter systematischer Übersichten

4.2.3 Selektion relevanter Studien

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

Die in bibliografischen Datenbanken identifizierten Treffer wurden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien bewertet. Als potenziell relevant erachtete Dokumente wurden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgten durch 2 Personen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

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

- Studienregister
- Referenzlisten identifizierter systematischer Übersichten

Die identifizierten Studien wurden dann auf ihre Relevanz geprüft. Der gesamte Prozess wurde anschließend von einer 2. Person überprüft. Sofern in einem der genannten Selektionsschritte Diskrepanzen auftraten, wurden diese jeweils durch Diskussion zwischen beiden aufgelöst.

4.3 Informationsdarstellung und Synthese

4.3.1 Darstellung der Einzelstudien

Alle für den Evidenzbericht notwendigen Informationen wurden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Ergaben sich im Abgleich der Informationen aus unterschiedlichen Dokumenten zu einer Studie (aber auch aus multiplen

Angaben zu einem Aspekt innerhalb eines Dokumentes selbst) Diskrepanzen, die auf die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wurde dies an den entsprechenden Stellen des Berichts dargestellt.

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

Die Ergebnisse zu den in den Studien berichteten und von den Leitlinienkoordinatoren festgelegten Endpunkten werden im Bericht vergleichend beschrieben.

Die Bewertung des Einflusses der Studienlimitationen auf die berichtsrelevanten Ergebnisse wird unter 4.3.2 endpunktspezifisch pro Studie beschrieben. Anschließend wurden die Informationen zu den ausgewählten Endpunkten in Evidenzprofilen zusammengeführt und analysiert [7]. Wenn möglich wurden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in den Abschnitten 4.3.3 und 4.3.4 beschriebenen Verfahren eingesetzt.

Für binäre Daten wurde als relativer Effektschätzer primär das Odds Ratio (OR) herangezogen. Dabei wurde im Fall von 0 Ereignissen in einem Studienarm bei der Berechnung von Effekt und KI der Korrekturfaktor 0,5 in beiden Studienarmen verwendet. Für absolute Effekte wurde vorrangig die absolute Risikodifferenz (RD) verwendet. Das Basisrisiko wurde in der Regel durch den Median des Risikos der Kontrollgruppe in den eingeschlossenen Einzelstudien bestimmt. Auf Grundlage dieses angenommenen Basisrisikos wurde mit Hilfe des relativen Gesamtschätzers der Metaanalyse das absolute Risiko in der Interventionsgruppe und die Risikodifferenz errechnet. Zur Abschätzung der quantitativen Ergebnisunsicherheit wurden ergänzend die oberen und unteren Konfidenzgrenzen der RD basierend auf den Konfidenzintervallgrenzen des relativen Gesamtschätzers der Metaanalyse berechnet. Wurde keine Metaanalyse durchgeführt, erfolgte eine Berechnung der RD pro Studie durch die Risiken in den beiden Behandlungsgruppen; als zugehöriges Konfidenzintervall wurde standardmäßig das KI nach der Wilson-Score-Methode [8] angegeben. Sollte das Ergebnis des CSZ-Tests [9] qualitativ nicht zum Ergebnis des KI nach der Wilson-Score-Methode, sondern zum Wald-KI passen, wurde dieses angegeben.

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

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

Die Ergebnisse konnten ggf. auch dann unberücksichtigt bleiben, wenn der Unterschied der Anteile nicht berücksichtigter Personen zwischen den Gruppen sehr groß war. Als Orientierung für die Entscheidung diente ein Unterschied von 15 Prozentpunkten.

4.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial)

Der Einfluss möglicher Studienlimitationen auf die Ergebnisse wurde endpunktspezifisch pro Studie insbesondere anhand folgender Kriterien bewertet:

- adäquate Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisunabhängige Berichterstattung

4.3.3 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien wurden mittels Forest Plots zusammenfassend dargestellt. Anschließend wurde die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [10] untersucht. Ergab der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wurde davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgte die Metaanalyse mithilfe des Modells mit zufälligen Effekten nach der Methode von Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [11]. Als Ergebnis wurde der gemeinsame Effekt inklusive Konfidenzintervall dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, wurden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Sind in einer Metaanalyse nur 2 Studien enthalten, so wurde aus pragmatischen Gründen regelhaft ein Modell mit festem Effekt gewählt, wenn keine deutlichen Gründe dagegensprechen. Bei 3 und 4 Studien erfolgte vorzugsweise eine Metaanalyse nach der Methode von Knapp und Hartung, sofern das Konfidenzintervall informativ ist und kein Widerspruch zu einer qualitativen Zusammenfassung bestand. Andernfalls wurde ein Modell mit festem Effekt verwendet, wenn keine deutlichen Gründe dagegensprochen. War ein Modell mit festem Effekt nicht vertretbar, blieb es bei der Schätzung nach Knapp und Hartung.

Ergab der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wurde untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählten methodische und klinische Faktoren, sogenannte Effektmodifikatoren. Konnte die Heterogenität durch diese Faktoren erklärt werden, wurde das Poolen der Effekte ggf. auf Subgruppen beschränkt [12]. Konnte die Heterogenität nicht erklärt werden, so wurde, sofern möglich, ebenfalls ein gemeinsamer (gepoolter) Effekt berechnet. Dabei ist zu beachten, dass ein gepoolter Effekt bei bedeutsamer Heterogenität möglicherweise nicht sinnvoll interpretierbar ist. Von der Berechnung eines gepoolten Effekts wurde abgesehen, falls sich die Konfidenzintervalle der

eingehenden Studien nur gering oder gar nicht überlappten und gleichzeitig die Effektschätzungen der Studien in unterschiedliche Richtungen wiesen.

4.3.4 Bewertung der Qualität der Evidenz

Alle für den Evidenzbericht relevanten Ergebnisse wurden hinsichtlich der Beeinflussung durch Studienlimitationen, des Risikos von Publikationsbias, der Genauigkeit und Konsistenz der Effekte, der Übertragbarkeit und ggf. hinsichtlich möglicher Aspekte zur Aufwertung der Qualität der Evidenz überprüft. Für jeden Endpunkt wurde eine endpunktbezogene studienübergreifende Aussage zur Qualität der Evidenz in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen und diese in eine der Stufen hohe, moderate, niedrige und sehr niedrige Qualität der Evidenz entsprechend der GRADE-Guideline eingeteilt [13,14]:

- Eine hohe Qualität der Evidenz bedeutet, dass der wahre Effekt sehr sicher nahe bei dem Effektschätzer liegt.
- Eine moderate Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich nahe bei dem Effektschätzer liegt, aber die Möglichkeit besteht, dass er relevant verschieden ist.
- Eine niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt durchaus relevant verschieden zum Effektschätzer sein kann.
- Eine sehr niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich relevant verschieden vom Effektschätzer ist.

Die Bewertung erfolgte durch 2 Personen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Ergebnissen aus RCTs wird im 1. Bewertungsschritt eine hohe, Ergebnissen aus Beobachtungsstudien eine niedrige Qualität der Evidenz attestiert [13]. Von dieser rein designbedingten Einschätzung ausgehend, wurde bei Studienlimitationen, einem Risiko von Publikationsbias, fehlender Genauigkeit der Effekte, inkonsistenten Effekten oder eingeschränkter Übertragbarkeit die Qualität der Evidenz abgewertet und bei (sehr) großen Effekten, einer Dosis-Wirkungs-Beziehung oder, wenn die Berücksichtigung aller potenziellen Confounder zu einer Effekterhöhung / Verstärkung des beobachteten Ergebnisses führen würde, aufgewertet. Unterstützend für diese Bewertungsschritte können Sensitivitätsanalysen durchgeführt werden. Alle Ab- und Aufwertungen der Qualität der Evidenz wurden durch Fußnoten gekennzeichnet.

4.3.4.1 Abwertung der Qualität der Evidenz

A: Studienlimitationen

Die Qualität der Evidenz wurde aufgrund von schwerwiegenden oder sehr schwerwiegenden Studienlimitationen um 1 oder 2 Stufen abgewertet, wenn starke Limitierungen in einem oder mehreren der in Abschnitt 4.3.2 genannten Kriterien das Vertrauen in die Effektschätzung

beeinträchtigen. Die endpunktbezogene studienübergreifende Bewertung der Studienlimitationen erfolgte unter Berücksichtigung des Einflusses der einzelnen Studien auf die Effektschätzung für jeden Endpunkt. Gründe für eine Abwertung wurden transparent beschrieben [15].

B: Inkonsistente (heterogene) Effekte

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

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

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

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

C: Indirektheit

Die Qualität der Evidenz wurde bei starker oder sehr starker Indirektheit um 1 oder 2 Stufen abgewertet. Indirektheit konnte auf Abweichungen zwischen dem PICO und den Einschlusskriterien der Studien basieren [16].

D: Publikationsbias

Die Qualität der Evidenz sollte um 1 oder 2 Stufen abgewertet werden, wenn ein Publikationsbias wahrscheinlich oder sehr wahrscheinlich war.

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

E: Fehlende Genauigkeit der Effekte

Die Qualität der Evidenz wurde wegen geringer oder sehr geringer Genauigkeit der Effekte um 1 oder 2 Stufen abgewertet. Maßgeblich hierfür war neben der Lage und Breite des

95 %-Konfidenzintervalls auch die Wahl des Metaanalysemodells (zufällige Effekte oder fester Effekt) (siehe Abschnitt 4.3.3). Außerdem konnten sehr kleine Fallzahlen zu einer Abwertung wegen fehlender Genauigkeit führen [18]. Bei ausreichender Patientenzahl wurde regelhaft nicht abgewertet, wenn das Konfidenzintervall eines gepoolten Effekts (relativ oder absolut im Fall binärer Daten bzw. Mittelwertdifferenz oder standardisierte Mittelwertdifferenz im Fall stetiger Daten) einen signifikanten Unterschied anzeigte.

Andernfalls erfolgte eine Abwertung wegen fehlender Präzision, falls das Konfidenzintervall (des relativen Effekts bei binären Daten bzw. Hedges' g bei stetigen Daten) einen großen Effekt und gleichzeitig auch den Nulleffekt überdeckte. Bei binären Daten wurde unter einem großen Effekt eine Halbierung oder eine Verdopplung (0,5 oder 2) des relativen Effekts verstanden. Bei stetigen Daten wurde von einem großen Effekt ausgegangen, wenn das Konfidenzintervall für eine standardisierte Mittelwertdifferenz die Grenze 0,5 oder $-0,5$ umfasste. In diesen Fällen erfolgte eine Abwertung wegen fehlender Genauigkeit um 1 Stufe. Eine Abwertung um 2 Stufen wurde in Fällen in Betracht gezogen, in denen die zuvor genannten Konfidenzintervallgrenzen beide, d. h. z. B. im Fall binärer Daten sowohl 0,5 als auch 2 im Konfidenzintervall enthalten waren.

4.3.4.2 Aufwertung der Qualität der Evidenz

A: Große Effekte

Die Qualität der Evidenz sollte in methodisch guten, nicht randomisierten Interventionsstudien bei großen bzw. sehr großen Effekten um 1 oder 2 Stufen aufgewertet werden. Eine Aufwertung um 1 Stufe ist bei großen Effekten möglich, wenn beispielsweise das relative Risiko zwischen 2 und 5 bzw. zwischen 0,2 und 0,5 liegt, ohne dass plausible Confounder diesen Effekt verringern könnten. Eine Aufwertung um 1 weitere Stufe sollte bei sehr großen Effekten erfolgen, wenn beispielsweise das relative Risiko über 5 bzw. unter 0,2 liegt, keine relevanten Verzerrungsrisiken vorliegen und die Effekte präzise (mit hinreichend geringen Konfidenzintervallen) geschätzt werden konnten [19].

B: Dosis-Wirkungs-Beziehung

Die Qualität der Evidenz sollte bei einer Dosis-Wirkungs-Beziehung aufgewertet werden [19].

C: Confounding

Die Qualität der Evidenz sollte bei einer möglichen Verstärkung des beobachteten Ergebnisses bei Berücksichtigung aller potenziellen Confounder oder Verzerrungsquellen aufgewertet werden [19].

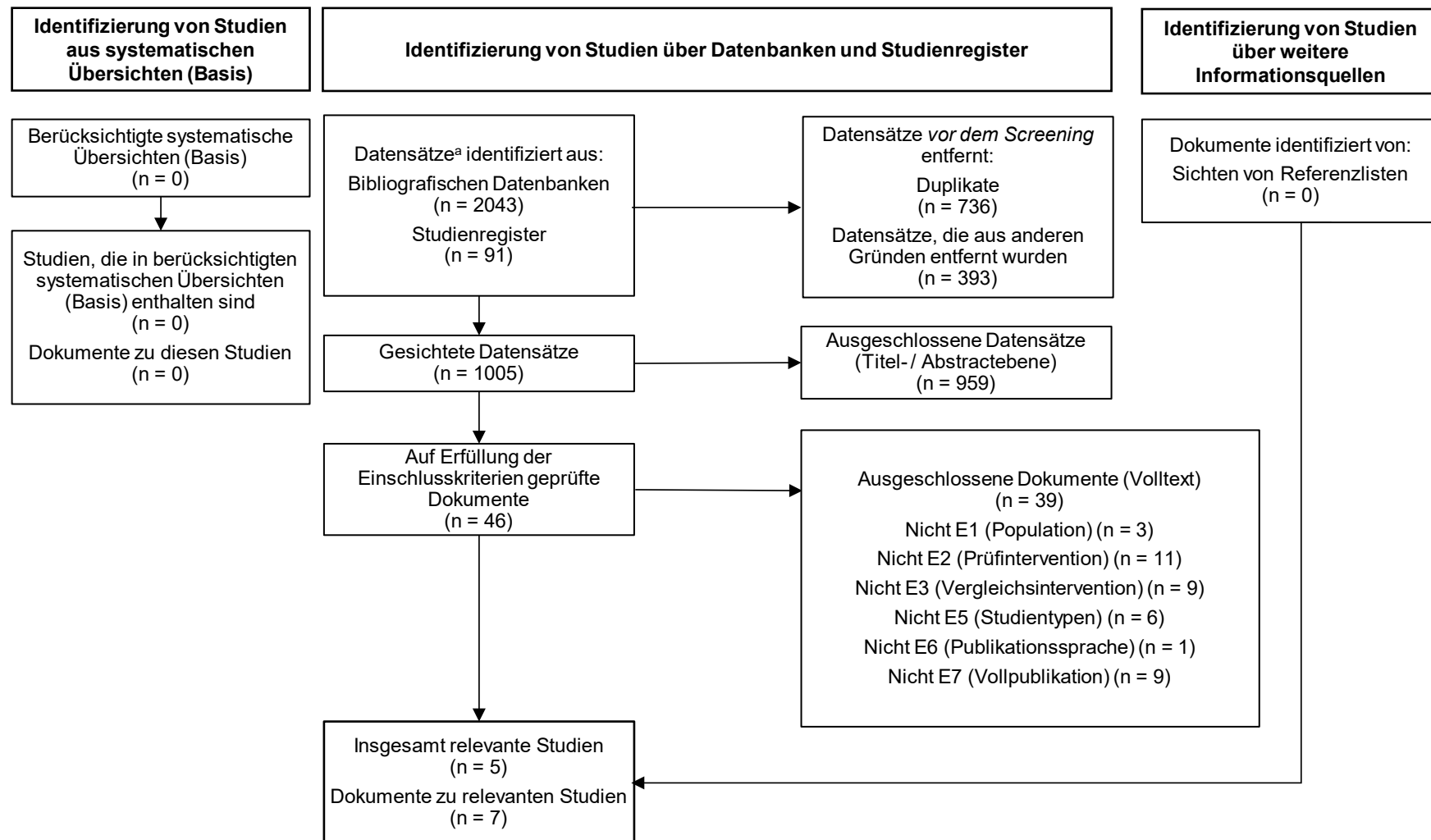
5 Ergebnisse

5.1 Ergebnisse der Informationsbeschaffung

Von den 8 identifizierten systematischen Übersichten (siehe Abschnitt 8.1) wurde keine systematische Übersicht als ausreichend aktuell und hochwertig bewertet und zum Zweck der Identifizierung von Studien berücksichtigt.

Abbildung 1 zeigt das Ergebnis der Informationsbeschaffung inklusive Studienselektion gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken (letzte Suche am 28.05.2021) und der Suche in Studienregistern (letzte Suche am 02.06.2021) sind in Anhang B dargestellt.

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



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

Abbildung 1: Ergebnis der Informationsbeschaffung

Resultierender Studienpool

Durch die verschiedenen Rechenschritte konnten insgesamt 5 relevante Studien identifiziert werden (siehe auch Tabelle 2).

Tabelle 2: Studienpool der Evidenzdarstellung

Studie	Verfügbare Dokumente	
	Vollpublikation (in Fachzeitschriften)	Ergebnisbericht aus Studienregistern
Jacobs 2004	ja [20]	nein
Morin 1999	ja [21,22]	nein
Pchelina 2019	ja [23]	nein
Sivertsen 2006	ja [24,25]	nein
Wu 2006	ja [26]	nein

5.2 Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien

Studiendesign und Studienpopulationen

In diesem Abschnitt werden die 5 eingeschlossenen Studien einschließlich der wesentlichen Ein- und Ausschlusskriterien für Patientinnen und Patienten sowie die Studienpopulationen charakterisiert (siehe Tabelle 3).

Dabei handelt es sich um 5 RCTs, die jeweils zwischen 30 und 41 Patientinnen und Patienten mit nicht organischer Insomnie eingeschlossen haben. Die Dauer der Prüfinerventionen lag in 4 der 5 RCTs zwischen 6 und 8 Wochen. Ausschließlich in Pchelina 2019 wurde eine Kurzversion der KVT-I untersucht, welche lediglich 2 Wochen dauerte und 2 Sitzungen umfasste. 2 RCTs verglichen eine KVT-I mit dem Benzodiazepin Temazepam [22,26]. 3 Studien nutzten einen Benzodiazepinrezeptoragonisten als Vergleichsintervention. Dabei verwendeten 2 Studien Zopiclon [23,25] und 1 Studie Zolpidem [20]. Die Nachbeobachtungszeit betrug, mit Ausnahme von Pchelina 2019, welche mit 2 Wochen vergleichsweise kurz war, 6 bis 24 Monate.

Tabelle 3: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien

Studie	Studiendesign	Patientenzahl (randomisiert) N	Intervention	Vergleich	Ort und Zeitraum der Durchführung	Interventionsdauer (N Therapiesitzungen), Nachbeobachtungsdauer
Benzodiazepine						
Morin 1999	RCT	I: 18; C: 20	KVT-I	Temazepam	k. A.	8 Wochen (8 Sitzungen), 24 Monate
Wu 2006	RCT	I: 19; C: 17	KVT-I	Temazepam	China; k. A.	8 Wochen (16 Sitzungen), 8 Monate
Benzodiazepinrezeptoragonisten						
Jacobs 2004	RCT	I: 15; C: 15	KVT-I	Zolpidem	USA; 1998 bis 2001	8 Wochen (5 Sitzungen), 12 Monate
Pchelina 2019	RCT (Cross-Over)	41	KVT-I	Zopiclon	Russland; k. A.	2 Wochen (2 Sitzungen), 2 Wochen
Sivertsen 2006	RCT	I: 18; C: 18	KVT-I	Zopiclon	Norwegen; 2004 bis 2005	6 Wochen (6 Sitzungen), 6 Monate
C: Vergleich; I: Prüfintervention; KVT-I: Kognitive Verhaltenstherapie für Insomnie; k. A.: keine Angaben; RCT: Randomized controlled Trial						

Charakterisierung der Prüf- und Vergleichsintervention

Eine Übersicht der Prüf- und Vergleichsinterventionen wird in Tabelle 4 dargestellt.

Die Prüfinerventionen waren hinsichtlich der Häufigkeit, der Interventionsdauer (mit Ausnahme von Pchelina 2019), des Zeitraums der Durchführung sowie der Inhalte vergleichbar. Alle Interventionen umfassten sowohl kognitive als auch verhaltenstherapeutische Komponenten. Wesentliche Unterschiede bestanden im Therapieformat der Prüfinerventionen. Während die KVT-I in Morin 1999 in Kleingruppen durchgeführt wurde, erfolgte in Jacobs 2004, Pchelina 2019 und Sivertsen 2006 eine KVT-I als Einzelbehandlung. In Wu 2006 wurde das Therapieformat nicht beschrieben.

Die Vergleichsinterventionen unterschieden sich aufgrund des breiten Einschlusskriteriums, welches sowohl alle in Deutschland zugelassenen Benzodiazepine als auch Benzodiazepinrezeptoragonisten umfasste. In 2 Studien [22,26] wurde das Benzodiazepin Temazepam untersucht. Die Dosierung und Art der Anwendung war in diesen beiden Studien vergleichbar. In den anderen Studien wurden die Benzodiazepinrezeptoragonisten Zolpidem [20] bzw. Zopiclon [23,25] als Vergleich verwendet, welche im Vergleich zu Temazepam kürzere Halbwertszeiten aufweisen [2].

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

Studie	Intervention	Vergleich
Benzodiazepine		
Morin 1999	KVT-I <u>Frequenz und N Therapiesitzungen (Dauer):</u> ▪ 1-mal/Woche ▪ 8 Therapiesitzungen (je 90 Minuten) <u>Interventionsdauer:</u> ▪ 8 Wochen <u>Therapieformat:</u> ▪ Kleingruppe (4-6 Patientinnen und Patienten) <u>Inhalt:</u> ▪ strukturierte, multifaktorielle Intervention, welche folgende Komponenten umfasst: ▫ Verhaltenstherapie (Schlafrestriktion, Stimuluskontrolle) ▫ kognitive Therapie ▫ edukative Therapie	Temazepam <u>Dosierung:</u> ▪ initial: ▫ 7,5 mg/Nacht ▪ schrittweise Erhöhung der Dosierung, abhängig vom Ansprechen auf die Behandlung und den Nebenwirkungen ▪ Max. Dosis: 30 mg/Nacht ▪ Empfehlung: mindestens 2- bis 3-mal/Woche <u>Art der Anwendung:</u> ▪ 1 Stunde vor dem Schlafengehen Dauer der Anwendung: k. A.
Wu 2006	KVT-I <u>Frequenz und N Therapiesitzungen (Dauer):</u> ▪ 2-mal/Woche ▪ 16 Therapiesitzungen <u>Interventionsdauer:</u> ▪ 8 Wochen <u>Therapieformat:</u> k. A. <u>Inhalt:</u> ▪ Verhaltenstherapeutische Komponenten: ▫ Stimuluskontrolle ▫ Schlafrestriktion ▫ Empfehlungen zur Schlafhygiene ▪ Kognitive Komponenten: ▫ Veränderung negativer Glaubenssätze und Einstellungen in Bezug auf den Schlaf	Temazepam <u>Dosierung:</u> ▪ Initial: ▫ 7,5 mg/Nacht ▪ schrittweise Erhöhung der Dosierung ▪ Max. Dosis: 30 mg/Nacht ▪ Absetzphase: ▫ 15 mg <u>Art der Anwendung:</u> ▪ 1 Stunde vor dem Schlafengehen <u>Dauer der Anwendung:</u> ▪ Initial: ▫ 1 Woche ▪ Max. 6 Wochen ▪ Absetzphase: ▫ 1 Woche

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

Studie	Intervention	Vergleich
Benzodiazepinrezeptoragonisten		
Jacobs 2004	KVT-I <u>Frequenz und N Therapiesitzungen (Dauer):</u> ▪ 1-mal / Woche (Woche 1–3) ▪ 4 Sitzungen (je 30 Minuten) + ▪ 1 telefonische Therapiesitzung (2 Wochen nach der letzten Therapiesitzung) <u>Interventionsdauer:</u> ▪ 8 Wochen <u>Therapieformat:</u> ▪ Einzelbehandlung <u>Inhalt:</u> ▪ schlaffokussierte, strukturierte, multifaktorielle Intervention, welche folgende Komponenten umfasst: ▫ kognitive Therapie ▫ Verhaltenstherapie ▫ Entspannung ▫ Edukation	Zolpidem <u>Dosierung:</u> ▪ Initial: ▫ 10 mg/Nacht ▪ Absetzphase: ▫ 5 mg/Nacht ▫ 5 mg/jede 2. Nacht <u>Art der Anwendung:</u> ▪ 30 Minuten vor dem Schlafengehen <u>Dauer der Anwendung:</u> ▪ Initial: ▫ 28 Tage ▪ Absetzphase: ▫ 7 Tage ▫ 7 Tage

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

Studie	Intervention	Vergleich
Pchelina 2019	KVT-I <u>Frequenz und N Therapiesitzungen (Dauer):</u> ▪ 1-mal / Woche ▪ 2 Sitzungen (je 60 Minuten) ▪ Audioentspannungstraining: ▫ 1-mal/Tag ▫ 32 Minuten <u>Interventionsdauer:</u> ▪ 2 Wochen <u>Therapieformat:</u> ▪ Einzelbehandlung ▪ Audioentspannungstraining: ▫ selbständiges Anhören der Aufnahmen <u>Inhalt:</u> ▪ Strukturierte Kurzintervention, welche folgende Komponenten umfasst: ▫ Besprechung von: - Mechanismen zur Schlafregulation - Gründe der Entstehung und Chronifizierung - Behandlungsstrategien ▫ Schlafrestriktion ▫ Stimuluskontrolle ▫ autogenes Entspannungstraining (Audio)	Zopiclon <u>Dosierung:</u> ▪ 7,5 mg/Nacht <u>Art der Anwendung:</u> ▪ 30 Minuten vor dem Schlafengehen <u>Dauer der Anwendung:</u> ▪ 2 Wochen

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

Studie	Intervention	Vergleich
Sivertsen 2006	KVT-I <u>Frequenz und N Therapiesitzungen (Dauer):</u> ▪ 1-mal / Woche ▪ 6 Therapiesitzungen (je 50 Minuten) <u>Interventionsdauer:</u> ▪ 6 Wochen <u>Therapieformat:</u> ▪ Einzelbehandlung <u>Inhalt:</u> ▪ Manualisierter Mehrkomponentenansatz, welcher folgende Komponenten umfasst: ▫ Edukation zur Schlafhygiene ▫ Schlafrestriktion ▫ Stimuluskontrolle ▫ kognitive Therapie ▫ progressive Entspannungstechniken	Zopiclon <u>Dosierung:</u> ▪ 7,5 mg/Nacht <u>Art und Dauer der Anwendung:</u> k. A.
k. A.: keine Angabe; KVT-I: Kognitive Verhaltenstherapie für Insomnie		

Ein- und Ausschlusskriterien für Patientinnen und Patienten in den Studien

Die Ein- und Ausschlusskriterien für Patientinnen und Patienten der Studien sind in Tabelle 5 dargestellt. Die wesentlichen Kriterien in Bezug auf die Population waren in allen Studien vergleichbar. Alle Studien erforderten die Diagnosestellung einer chronischen Insomnie nach Diagnostic and Statistical Manual of Mental Disorders Version 5 (DSM-IV), der International Classification of Sleep Disorders Version 3 (ICSD-3) oder sie schlossen Patientinnen und Patienten ein, die den dort festgelegten Kriterien entsprachen. Unterschiede gab es bezüglich des Alters. Während generell erwachsene Patientinnen und Patienten eingeschlossen wurden, beschränkten sich die Studien Morin 1999 und Sivertsen 2006 per Einschlusskriterium auf Patientinnen und Patienten im Alter von ≥ 55 Jahren.

Ausgeschlossen wurden Patientinnen und Patienten mit einer organischen Ursache der Insomnie sowie solche mit schwerwiegenden komorbiden physischen oder psychischen Erkrankungen. Darüber hinaus wurden Patientinnen und Patienten ausgeschlossen, welche zu Studienbeginn eine psychotherapeutische oder medikamentöse Behandlung der Insomnie erhielten. Weiterhin wurden in der Regel Patientinnen und Patienten ausgeschlossen, die regelmäßig in Nachtschichten arbeiten oder wenn bei Patientinnen eine Schwangerschaft vorlag.

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

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Benzodiazepine		
Wu 2006	▪ ICSD und DSM-IV-Kriterien für primäre und chronische Insomnie: ▫ Ein- und Durchschlafstörungen ▫ Einschlaf latenz und / oder Wachzeit nach dem Einschlafen von ≥ 30 Minuten pro Nacht ▫ Probleme treten mindestens 3 Nächte pro Woche über mindestens 6 Monate auf ▫ subjektiver Distress und funktionelle Beeinträchtigung im Alltag	▪ Hinweise darauf, dass die Insomnie in direktem Zusammenhang mit einer Erkrankung oder den Nebenwirkungen von Medikamenten steht ▪ Vorliegen einer Schlafapnoe oder einer PLM im Schlaf ▪ regelmäßige Einnahme von Hypnotika oder anderen Psychopharmaka, mit der fehlenden Bereitschaft oder Möglichkeit diese Medikamenteneinnahme einzustellen ▪ Teilnahme an einer Psychotherapie ▪ schwere Depression oder eine andere schwere psychische Störung ▪ unbewältigte persönliche oder psychische Probleme (außer Schlafangst) in den letzten 6 Monaten
Morin 1999	▪ Alter: ≥ 55 Jahre ▪ Insomnie ≥ 6 Monate ▪ Einschlaf- oder Durchschlafstörungen, definiert als Einschlaf latenz und / oder Aufwachen nach dem Einschlafen von mehr als 30 Minuten pro Nacht in mindestens 3 Nächten pro Woche ▪ mindestens 1-mal täglich Beeinträchtigung während der Wachzeit (z.B. Müdigkeit, Funktionsstörung, Stimmungsschwankungen), welche auf die Insomnie zurückzuführen ist	▪ Hinweise darauf, dass die Insomnie in direktem Zusammenhang mit einer Erkrankung oder den Nebenwirkungen von Medikamenten steht ▪ Vorliegen von Schlafapnoe (Apnoe-Hypopnoe-Index > 15) oder PLM im Schlaf (Myoklonischer Index mit Erregung > 15) ▪ regelmäßige Einnahme eines Hypnotikums oder eines anderen Psychopharmakons mit der fehlenden Bereitschaft oder der Unwilligkeit, das Medikament abzusetzen ▪ psychotherapeutische Behandlung ▪ Vorliegen einer schweren Depression oder anderer schwerer psychopathologischer Störungen auf der Grundlage einer kurzen Selbstauskunft (z. B. Brief Symptom Inventory) und DSM-III-R ▪ kognitive Störung, die durch einen Wert von ≤ 23 im Mini-Mental-Status-Test nachgewiesen wird

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

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Benzodiazepinrezeptoragonisten		
Jacobs 2004	- Alter: 25–64 Jahre - Insomnie ≥ 6 Monaten - DSM-IV-Kriterien für primäre und chronische Insomnie - Schlaflatenz von mindestens 1 Stunde ≥ 3-mal/Woche - die wöchentlichen Zeiten sind anhand von Schlaftagebüchern belegt - mindestens 1-mal tägliche Beeinträchtigung (z. B. Müdigkeit, beeinträchtigte Stimmung oder Leistung), die auf die Insomnie zurückzuführen ist	- Einnahme von verschreibungspflichtigen oder rezeptfreien Schlafmitteln oder die fehlende Bereitschaft oder Möglichkeit, die Einnahme dieser Medikamente ≥ 4 Wochen vor Beginn der Studie einzustellen - gesundheitliche Probleme, die eine direkte Ursache für die Insomnie sein könnten - Behandlung aufgrund von Depression, Alkohol- oder Drogenmissbrauch oder Psychose - frühere Diagnose von Schlafapnoe oder PLM oder frühere KVT-I - Aufnahme einer Psychotherapie in den letzten 6 Monaten - Frauen im gebärfähigem Alter, die schwanger sind, stillen oder nicht verhüten - Schichtarbeit
Pchelina 2019	ICSD-3 Kriterien für chronische Insomnie	- Einnahme von Medikamenten, die den Schlaf beeinflussen und die fehlende Bereitschaft oder Möglichkeit, die Einnahme dieser Medikamente mindestens 1 Wochen vor Studienbeginn und während der Studie einzustellen - Vorgeschichte des Missbrauchs von Medikamenten, Alkohol oder Betäubungsmitteln - diagnostizierte endogene psychische Störung - Demenz - Schwangerschaft oder stillend - Schichtarbeit - Krankheiten, die den Schlaf beeinträchtigen: - moderat-schweres obstruktives Apnoe Syndrom, Restless-Legs-Syndrom, PLM, starke Schmerzen - schwere chronische Erkrankungen oder somatische Erkrankungen in Exazerbation oder Dekompensation, die eine, Studienteilnahme verhindern.

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

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Sivertsen 2006	- Alter: ≥ 55 Jahre - Insomnie ≥ 3 Monaten - Erfüllung der Kriterien des DSM-IV für Insomnie, einschließlich Einschlaf- und Durchschlafstörungen und / oder frühmorgendliches Erwachen ohne die Möglichkeit wieder einzuschlafen, Beeinträchtigungen der Leistungsfähigkeit am Tag	- Einnahme von Hypnotika in den letzten 4 Wochen - Einnahme von Antidepressiva oder Antipsychotika - Anzeichen einer Demenz oder einer anderen schwerwiegenden kognitiven Störung, definiert durch einen Wert von ≤ 23 beim Mini-Mental-Status-Test - Vorliegen einer schweren depressiven Störung oder einer anderen schweren psychischen Störung, die durch eine klinische Beurteilung auf der Grundlage des strukturierten klinischen Interviews für DSM-IV festgestellt wurde - Vorliegen einer Schlafapnoe definiert als Apnoe-Hypopnoe-Index > 15 oder PLM im Schlaf (PLM-Index mit arousal > 15) - Arbeit in Nachtschichten und keine Bereitschaft oder Möglichkeit dieses Arbeitsmodell aufzugeben - fehlende Bereitschaft oder Möglichkeit die Einnahme von Hypnotika vor der Studienteilnahme einzustellen - eine schwere somatische Erkrankung, die eine kontinuierliche Teilnahme an der Studie verhindert
DSM: Diagnostic and Statistical Manual; ICSD: Internationale Klassifikation der Schlafstörungen; KVT-I: kognitive Verhaltenstherapie; PLM: Periodic limb movement (periodische Beinbewegungen)		

Charakterisierung der Studienpopulation

Keine Studie berichtete verwertbare Daten zur Krankheitsschwere zu Studienbeginn. Andere wesentliche Charakteristika der Studienpopulation unterschieden sich zwischen den meisten Studien geringfügig. Die relevantesten Unterschiede zeigten sich in den Studien Morin 1999 und Sivertsen 2006 im Vergleich zu den übrigen Studien. Hier war das mittlere Alter, sowie die Krankheitsdauer, höher als in den anderen Studien. Der Frauenanteil war in allen Studien, außer in Wu 2006 und Pchelina 2019, etwas geringer als der Männeranteil. Der Anteil der Studienabbrecher war in allen Studien in der Vergleichsgruppe höher.

Tabelle 6: Charakterisierung der Studienpopulationen

Subgruppe Studie	Intervention	Vergleich
Benzodiazepine		
Morin 1999		
N	18	17
Alter [Jahre] MW (SD)	64,4 (7,5)	64,1 (6,4)
Geschlecht [w / m] %	28 / 72	47 / 53
Dauer der Erkrankung [Jahre] (SD)	16,2 (14,8)	15,4 (15,5)
Studienabbrecher n	0	3
Wu 2006		
N	19 ^a	20 ^a
Alter [Jahre] MW (SD)	38 (12) ^a	38 (12) ^a
Geschlecht [w / m] %	53 / 47 ^a	53 / 47 ^a
Dauer der Erkrankung	k. A.	k. A.
Studienabbrecher n	0	3
Benzodiazepinrezeptoragonisten		
Jacobs 2004		
N	15	15
Alter [Jahre] MW (SD)	47,1 (8,1)	45,4 (9,3)
Geschlecht [w / m] %	33 / 67	27 / 73
Bildung [Jahre] (SD)	15,8 (3,2)	15,9 (2,8)
Dauer der Erkrankung [Jahre] (SD)	10,2 (9,1)	9,8 (7,8)
Studienabbrecher n	1	2
Pchelina 2019		
N	41	
Alter [Jahre] MW [Min; Max]	54 [29; 80]	
Geschlecht [w / m] %	67 / 33	
Dauer der Erkrankung [Jahre] (SD)	k. A.	
Studienabbrecher n	3 ^b	
Sivertsen 2006		
N	18	18
Alter [Jahre] MW (SD)	59,8 (4,3)	61,3 (6,9)
Geschlecht [w / m] %	39 / 61	38 / 62
Dauer der Erkrankung [Jahre] (SD)	15,8 (2–43)	13,7 (3–35)
Studienabbrecher n	2	5
a. Die Angaben beziehen sich auf die Gesamtpopulation der Studie. b. Alle Studienabbrüche erfolgten während der Vergleichsintervention (Zopiclon). k. A.: keine Angabe; m: männlich; MW: Mittelwert; N: Anzahl randomisierter (bzw. eingeschlossener) Patientinnen und Patienten; n: Anzahl Patientinnen und Patienten mit Ereignis; RCT: randomisierte kontrollierte Studie; SD: Standardabweichung; w: weiblich		

5.3 Übersicht der Endpunkte

Aus 5 Studien konnten Daten zu Endpunkten extrahiert werden. Die Tabelle 7 zeigt die Übersicht der verfügbaren Daten zu Endpunkten aus den eingeschlossenen Studien. In der Studie Pchelina 2019 wurden zwar Ergebnisse zu den Endpunkten Schlafeffizienz, Einschlaf latenz, Wachzeit nach dem Einschlafen sowie zur Schlafdauer erhoben und berichtet, allerdings wurden diese nur für nicht relevante Subgruppen dargestellt. Da somit keine vollständige Darstellung der Ergebnisse über die gesamte Population erfolgte, waren die Daten nicht verwertbar. Zum Endpunkt morgendliches Erwachen wurden in keiner Studie Daten berichtet. In Morin 1999 erfolgte trotz geplanter Erhebung keine Darstellung der Ergebnisse für die Endpunkte morgendliches Erwachen sowie Einschlaf latenz.

Tabelle 7: Matrix der Endpunkte

Studie	Endpunkte						
	Krankheitsschwere	Schlafeffizienz	Einschlaf latenz	Wachzeit nach dem Einschlafen	morgendliches Erwachen	Schlafqualität	Schlafdauer
Benzodiazepine							
Morin 1999	●	●	x	●	x	–	●
Wu 2006	–	●	●	–	–	–	●
Benzodiazepinrezeptoragonisten							
Jacobs 2004	–	●	●	–	–	–	●
Pchelina 2019	●	○	○	○	–	●	○
Sivertsen 2006	–	●	–	–	–	–	●
●: Daten wurden berichtet und waren verwertbar. ○: Daten wurden berichtet, aber waren nicht für den Evidenzbericht verwertbar. x: Daten wurden trotz geplanter Erhebung nicht berichtet. –: Es wurden keine Daten berichtet (keine weiteren Angaben) / Der Endpunkt wurde nicht erhoben.							

5.4 Ergebnisse zu Endpunkten

5.4.1 Ergebnisse zum Endpunkt Krankheitsschwere

Tabelle 8: Evidenzprofil zum Endpunkt Krankheitsschwere (Insomnia Severity Index < 15) – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patienten und Patientinnen mit Ereignis / Anzahl der Patientinnen und Patienten		Basisrisiko ^a in %	Effekt ^b		Qualität der Evidenz ^c
Studien-design; N	Studien-limitationen ^d	Inkonsistenz	Indirektheit	Publikationsbias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	Absolute Risiko-differenz [95 %-KI] in Prozentpunkten	
<i>Interpretation des Effektschätzers: Ein positiver absoluter Effekt bzw. ein Odds Ratio größer als 1 bedeuten einen Vorteil für die Interventionsgruppe.</i>												
Benzodiazepine												
RCT, 1 [22]	sehr schwerwiegend ^e	nicht zutreffend	keine	unentdeckt	schwerwiegend ^f	keine	14/18	10/17	59	2,45 [0,56; 10,68]	19 [-11; 45] ^g	sehr niedrig
a. Medianes Basisrisiko der Vergleichsgruppen der eingeschlossenen Studien. b. eigene Berechnung c. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. d. Die Bewertung der Studienlimitationen pro Studie und pro Endpunkt ist Anhang A zu entnehmen. e. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet. f. Das 95 %-KI des OR überdeckt 1 und 2. Somit können weder Effekte zuungunsten, noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet. g. eigene Berechnung von Effekt und KI nach der Wilson-Score-Methode (asymptotisch) C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio												

Tabelle 9: Evidenzprofil zum Endpunkt Krankheitsschwere (Insomnia Severity Index) – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepinrezeptoragonisten										
RCT (Crossover), 1 [23]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	38	−0,1 [−2,69; 2,49] ^f	−0,02 [−0,47; 0,43] ^f	sehr niedrig
a. eigene Berechnung b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen. d. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung war nicht adäquat. Die Umsetzung des ITT-Prinzips und die ergebnisunabhängige Berichterstattung waren unklar. Das sonstige Verzerrungspotenzial war aufgrund weniger und nicht übereinstimmender Angaben zu Baselinewerten der Patientinnen und Patienten unklar. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet. e. Das 95 %- KI von Hedges' g überdeckt annähernd sowohl −0,5 als auch 0,5. Es können weder annähernd große Effekte zugunsten der Prüfintervention noch annähernd große Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet. f. Die Angaben in der Publikation waren nicht ausreichend, um die Abhängigkeit der Daten, die aus dem Cross-Over-Design resultiert, adäquat zu berücksichtigen. Die Ergebnisse wurden wie bei einer Studie mit Parallelgruppen-Design ausgewertet, was zu einem zu breiten KI führen kann. C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

5.4.2 Ergebnisse zum Endpunkt Schlaffeffizienz

Da sich in der Metaanalyse eine signifikante Heterogenität zeigte (siehe Abschnitt 5.4.2.1), welche sich durch eine Subgruppenanalyse nach Medikamentengruppen erklären ließ (siehe Abschnitt 5.4.2.2), werden die Ergebnisse für Benzodiazepine und Benzodiazepinrezeptoragonisten separat dargestellt.

Tabelle 10: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Schlafstagebuch – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 2 [22,26]	sehr schwer-wiegend ^d	keine	keine	unentdeckt	schwer-wiegend ^e	keine	37; 34	0,16 [-3,69; 4,01]	-0,07 [-0,54; 0,40]	sehr niedrig
Benzodiazepinrezeptoragonisten										
RCT, 2 [20,25]	schwer-wiegend ^f	keine	schwer-wiegend ^g	unentdeckt	keine	keine	32; 27	12,03 [4,37; 19,68]	— ^h	niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren in beiden Studien unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren in beiden Studien nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder große Effekte zuungunsten der Prüfintervention noch Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
f. Die Randomisierung war in Sivertsen 2006 unklar. Die Verdeckung der Gruppenzuteilung war in beiden Studien unklar und die Verblindung war nicht adäquat. Die Umsetzung des ITT-Prinzips war in Jacobs 2004 nicht adäquat. Die ergebnisunabhängige Berichterstattung war in Sivertsen 2006 nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
h. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 11: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Schlafstagebuch – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)										
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	keine	keine	19; 17	7,30 [0,24; 14,36]	— ^e	niedrig
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 1 Monat)										
RCT, 1 [20]	schwer-wiegend ^f	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^g	keine	13; 11	5,4 [−7,00; 17,80]	0,36 [−0,45; 1,17]	niedrig
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 6 Monate)										
RCT, 1 [25]	sehr schwer-wiegend ^h	nicht zutreffend	schwer-wiegend ⁱ	unentdeckt	keine	keine	18; 16	9,30 [0,59; 18,01]	— ^e	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung war unklar. Die Verblindung und Umsetzung des ITT-Prinzips war nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant										
f. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung und Umsetzung des ITT-Prinzips war nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
g. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
h. Die Randomisierung und Verdeckung der Gruppenzuteilung war unklar. Die Verblindung, die Umsetzung des ITT-Prinzips sowie die ergebnisunabhängige Berichterstattung war nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
i. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 12: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Polysomnografie – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 2 [22]	sehr schwer-wiegend ^d	sehr schwer-wiegend ^e	keine	unentdeckt	sehr schwer-wiegend ^f	keine	18; 16	1,28 [−3,83; 6,39]	0,17 [−0,51; 0,84]	sehr niedrig
[26]							19; 17	−6,90 [−12,86; −0,94]	−0,75 [−1,43; −0,07]	
Benzodiazepinrezeptoragonisten										
RCT, 2 [20,25]	schwer-wiegend ^g	keine	schwer-wiegend ^h	unentdeckt	keine	keine	31; 28	6,83 [2,23; 11,42]	− ⁱ	niedrig
a. eigene Berechnung b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen. d. Die Randomisierung und Verdeckung der Gruppenzuteilung war in beiden Studien unklar. Die Verblindung und Umsetzung des ITT-Prinzips war in beiden Studien nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet. e. bedeutsame Heterogenität und Effektschätzer zeigen in unterschiedliche Richtungen, daher erfolgte keine Darstellung eines gepoolten Effektes. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet. f. Die Vereinigung der beiden 95 %-KIs der Hedges' g überdeckt −0,5 und 0,5. Somit können weder große Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Es erfolgte keine zusätzliche Abwertung. g. Die Randomisierung war in Sivertsen 2006 unklar. Die Verdeckung der Gruppenzuteilung war in beiden Studien unklar und die Verblindung war nicht adäquat. Die ergebnisunabhängige Berichterstattung war in Sivertsen 2006 nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet. h. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet. i. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 13: Evidenzprofil zum Endpunkt Schlaffeffizienz (%) – Polysomnografie – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b		
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]			
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.												
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)												
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	keine	keine	19; 17	15,70 [9,26; 22,14]	— ^e	niedrig		
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 6 Monate)												
RCT, 1 [25]	sehr schwer-wiegend ^f	nicht zutreffend	schwer-wiegend ^g	unentdeckt	keine	keine	18; 16	8,20 [2,54; 13,86]	— ^e	sehr niedrig		
a. eigene Berechnung												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.												
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
e. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant												
f. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung, die Umsetzung des ITT-Prinzips sowie die ergebnisunabhängige Berichterstattung waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien												

5.4.2.1 Metaanalysen

KVT-I vs. PCT - Behandlungsende

Schlafeffizienz - Schlaftagebuch

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

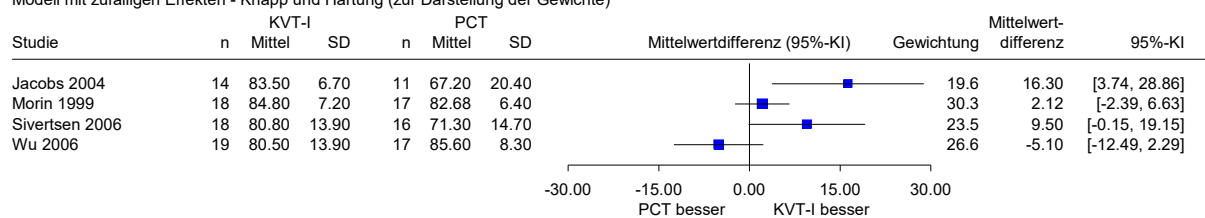


Abbildung 2: Metaanalyse, Schlafeffizienz in % (Schlaftagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

KVT-I vs. PCT - Behandlungsende

Schlafeffizienz - Polysomnographie

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

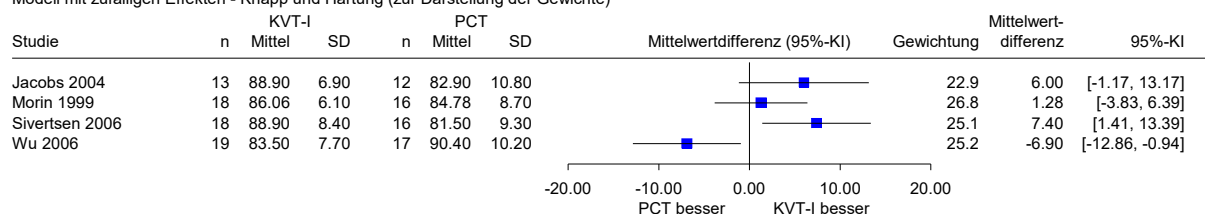
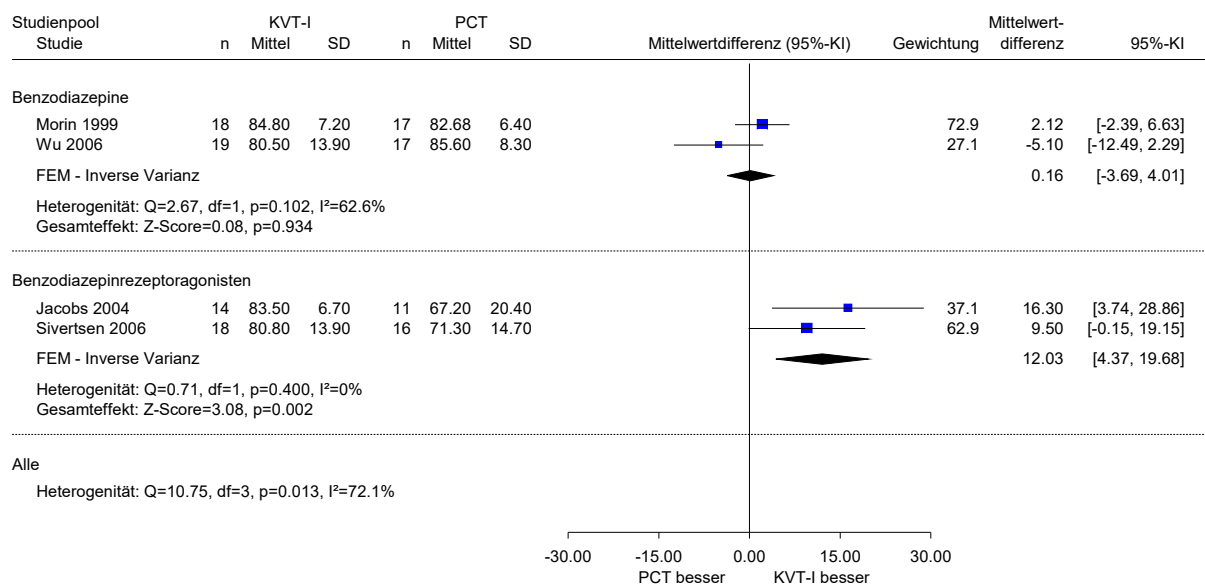


Abbildung 3: Metaanalyse, Schlafeffizienz in % (Polysomnographie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

5.4.2.2 Subgruppenmerkmale und andere Effektmodifikatoren

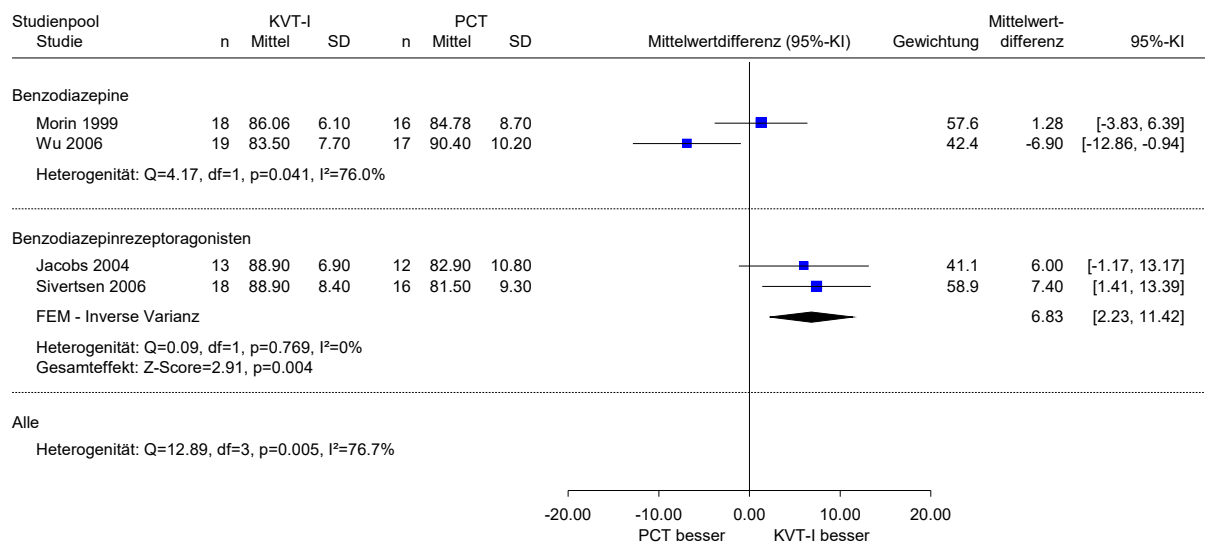
KVT-I vs. PCT - Behandlungsende
Schlafeffizienz - Schlaftagebuch



Heterogenität zwischen Studienpools: $Q=7.37$, $df=1$, $p=0.007$, $I^2=86.4\%$

Abbildung 4: Subgruppenanalyse nach Wirkstoffgruppe, Schlafeffizienz in % (Schlaftagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

KVT-I vs. PCT - Behandlungsende
Schlafeffizienz - Polysomnographie



Heterogenität zwischen Studienpools: $Q=8.63$, $df=1$, $p=0.003$, $I^2=88.4\%$

Abbildung 5: Subgruppenanalyse nach Wirkstoffgruppe, Schlafeffizienz in % (Polysomnographie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

5.4.3 Ergebnisse zum Endpunkt Einschlaf latenz

Tabelle 14: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Schlaf tagebuch – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	schwer-wiegend ^e	keine	keine	19; 17	12,00 [2,52; 21,48]	— ^f	sehr niedrig
Benzodiazepinrezeptoragonisten										
RCT, 1 [20]	schwer-wiegend ^g	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^h	keine	14; 13	–24,60 [–53,11; 3,91]	–0,66 [–1,44; 0,12]	niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. In Morin 1999 wurde der Endpunkt zwar erhoben jedoch nicht berichtet.										
f. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant										
g. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung war nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
h. Das 95 %-KI des Hedges' g überdeckt –0,5 und 0. Somit können weder große Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 15: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Schlafstagebuch – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b		
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]			
<i>Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.</i>												
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)												
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	schwer-wiegend ^e	keine	keine	19; 17	-26,00 [-39,17; -12,83]	- ^f	sehr niedrig		
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 1 Monat)												
RCT, 1 [20]	schwer-wiegend ^g	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^h	keine	13; 11	-21,20 [-56,77; 14,37]	-0,49 [-1,31; 0,33]	niedrig		
a. eigene Berechnung												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.												
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
e. In Morin 1999 wurde der Endpunkt zwar erhoben jedoch nicht berichtet.												
f. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant												
g. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
h. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder große Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet..												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien												

Tabelle 16: Evidenzprofil zum Endpunkt Einschlaflatenz (Minuten) – Polysomnografie – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^b	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	19; 17	5,60 [−3,82; 15,02]	0,39 [−0,27; 1,06]	sehr niedrig
Benzodiazepinrezeptoragonisten										
RCT, 1 [20]	schwer-wiegend ^f	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^g	keine	13; 12	−19,00 [−43,70; 5,70]	−0,62 [−1,42; 0,19]	niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder große Effekte zuungunsten der Prüfintervention noch Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
f. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
g. Das 95 %-KI des Hedges' g überdeckt −0,5 und 0. Somit können weder große Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 17: Evidenzprofil zum Endpunkt Einschlaf latenz (Minuten) – Polysomnografie – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)										
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	keine	keine	19; 17	-18,70 [-33,52; -3,88]	- ^e	niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

5.4.4 Ergebnisse zum Endpunkt Wachzeit nach dem Einschlafen

Tabelle 18: Evidenzprofil zum Endpunkt Wachzeit nach dem Einschlafen (Minuten) – Schlafstagebuch – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 1 [22]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	18; 17	-7,19 [-19,75; 5,37]	-0,38 [-1,05; 0,28]	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und die Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder große Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet..										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 19: Evidenzprofil zum Endpunkt Wachzeit nach dem Einschlafen (Minuten) – Polysomnografie – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 1 [22]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	sehr schwer-wiegend ^e	keine	18; 16	-2,65 [-22,12; 16,82]	-0,09 [-0,77; 0,58]	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0,5. Somit können weder große Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

5.4.5 Ergebnisse zum Endpunkt Schlafqualität

Tabelle 20: Evidenzprofil zum Endpunkt Schlafqualität (Pittsburgh Schlafqualitäts-Index) – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein negativer Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepinrezeptoragonisten										
RCT (Crossover), 1 [23]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	38	-1,10 [-3,07; 0,87] ^f	-0,25 [-0,70; 0,20] ^f	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Verdeckung der Gruppenzuteilung war unklar. Die Verblindung war nicht adäquat. Die Umsetzung des ITT-Prinzips und die ergebnisunabhängige Berichterstattung waren unklar. Das sonstige Verzerrungspotenzial war unklar. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder große Effekte zugunsten der Prüfintervention noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
f. Die Angaben in der Publikation waren nicht ausreichend, um die Abhängigkeit der Daten, die aus dem Cross-Over-Design resultiert, adäquat zu berücksichtigen. Die Ergebnisse wurden wie bei einer Studie mit Parallelgruppen-Design ausgewertet, was zu einem zu breiten KI führen kann.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

5.4.6 Ergebnisse zum Endpunkt Schlafdauer

Da sich in der Metaanalyse eine signifikante Heterogenität zeigte (siehe Abschnitt 5.4.6.1), welche sich durch eine Subgruppenanalyse nach Medikamentengruppen erklären ließ (siehe Abschnitt 5.4.6.2), werden die Ergebnisse für Benzodiazepine und Benzodiazepinrezeptoragonisten jeweils separat dargestellt.

Tabelle 21: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Schlafstagebuch – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine										
RCT, 2 [22,26]	sehr schwer-wiegend ^d	keine	keine	unentdeckt	keine	keine	37; 34	-37,24 [-62,91; -11,56]	- ^e	niedrig
Benzodiazepinrezeptoragonisten										
RCT, 2 [20,25]	schwer-wiegend ^f	keine	schwer-wiegend ^g	unentdeckt	schwer-wiegend ^h	keine	32; 29	-9,73 [-43,57; 24,11]	-0,14 [-0,65; 0,36]	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren in beiden Studien unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren in beiden Studien nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant										
f. Die Randomisierung war in Sivertsen 2006 unklar. Die Verdeckung der Gruppenzuteilung war in beiden Studien unklar und die Verblindung war nicht adäquat. Die Umsetzung des ITT-Prinzips war in Jacobs 2004 nicht adäquat. Die ergebnisunabhängige Berichterstattung war in Sivertsen 2006 nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
h. Das 95 %-KI des Hedges' g überdeckt -0,5 und 0. Somit können weder große Effekte zuungunsten der Prüfintervention noch Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 22: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Schlaftagebuch – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]	
Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.										
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)										
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	19; 17	13,40 [−29,11; 55,91]	0,21 [−0,45; 0,87]	sehr niedrig
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 6 Monate)										
RCT, 1 [25]	sehr schwer-wiegend ^f	nicht zutreffend	schwer-wiegend ^g	unentdeckt	schwer-wiegend ^e	keine	18; 16	16,10 [−29,06; 61,26]	0,24 [−0,43; 0,92]	sehr niedrig
a. eigene Berechnung										
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.										
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.										
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
e. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
f. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung, die Umsetzung des ITT-Prinzips, sowie die ergebnisunabhängige Berichterstattung waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.										
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.										
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien										

Tabelle 23: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Polysomnografie – Interventionsende

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b		
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]			
<i>Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.</i>												
Benzodiazepine												
RCT, 2 [22,26]	sehr schwer-wiegend ^d	keine	keine	unentdeckt	keine	keine	37; 33	-31,32 [-51,72; -10,93]	- ^e	niedrig		
Benzodiazepinrezeptoragonisten												
RCT, 2 [20,25]	schwer-wiegend ^f	keine	schwer-wiegend ^g	unentdeckt	schwer-wiegend ^h	keine	31; 28	28,62 [-4,25; 61,48]	0,46 [- 0,06; 0,98]	sehr niedrig		
a. eigene Berechnung												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.												
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren in beiden Studien unklar und die Verblindung war nicht adäquat. In beiden Studien war die Umsetzung des ITT-Prinzips nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
e. für die Einschätzung der fehlenden Genauigkeit nicht erforderlich, da Mittelwertdifferenz signifikant												
f. Die Randomisierung war in Sivertsen 2006 unklar. Die Verdeckung der Gruppenzuteilung war in beiden Studien unklar. Die Verblindung war in beiden Studien nicht adäquat. Die Umsetzung des ITT-Prinzips war in Jacobs 2004 nicht adäquat. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
h. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien												

Tabelle 24: Evidenzprofil zum Endpunkt Schlafdauer (Minuten) – Polysomnografie – Nachbeobachtungszeitpunkt

Faktoren der Qualität der Evidenz							Anzahl der Patientinnen und Patienten I; C	Effekt ^a		Qualität der Evidenz ^b		
Studien-design, N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren		Mittelwert-differenz [95 %-KI]	Hedges' g [95 %-KI]			
<i>Interpretation des Effektschätzers: Ein positiver Effektschätzer bedeutet einen Vorteil für die Interventionsgruppe.</i>												
Benzodiazepine (Nachbeobachtungszeitpunkt 3 Monate)												
RCT, 1 [26]	sehr schwer-wiegend ^d	nicht zutreffend	keine	unentdeckt	schwer-wiegend ^e	keine	19; 17	24,90 [−6,65; 56,45]	0,52 [−0,14; 1,19]	sehr niedrig		
Benzodiazepinrezeptoragonisten (Nachbeobachtungszeitpunkt 6 Monate)												
RCT, 1 [25]	sehr schwer-wiegend ^f	nicht zutreffend	schwer-wiegend ^g	unentdeckt	schwer-wiegend ^e	keine	18; 16	32,90 [−10,61; 76,41]	0,52 [−0,17; 1,20]	sehr niedrig		
a. eigene Berechnung												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang A zu entnehmen.												
d. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung und Umsetzung des ITT-Prinzips waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
e. Das 95 %-KI des Hedges' g überdeckt 0 und 0,5. Somit können weder Effekte zuungunsten der Prüfintervention noch große Effekte zugunsten der Prüfintervention ausgeschlossen werden. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
f. Die Randomisierung und Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung, die Umsetzung des ITT-Prinzips sowie die ergebnisunabhängige Berichterstattung waren nicht adäquat. Die Qualität der Evidenz wurde um 2 Kategorien abgewertet.												
g. Abweichung von der Zulassung in Deutschland: Die verwendete Dosierung in Sivertsen 2006 betrug 7,5 mg pro Tag. Für die untersuchte Population der älteren Patientinnen und Patienten wird jedoch eine initiale maximale Tagesdosis von 3,75 mg empfohlen. Die Qualität der Evidenz wurde um 1 Kategorie abgewertet.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien												

5.4.6.1 Metaanalysen

KVT-I vs. PCT - Behandlungsende

Schlafdauer - Schlaftagebuch

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

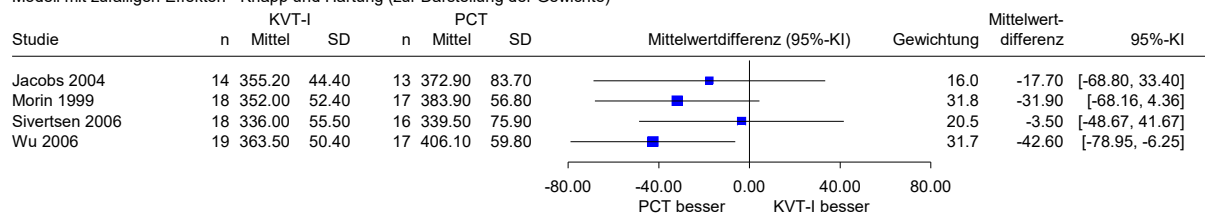


Abbildung 6: Metaanalyse, Schlafdauer in Minuten (Schlaftagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

KVT-I vs. PCT - Behandlungsende

Schlafdauer - Polysomnographie

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

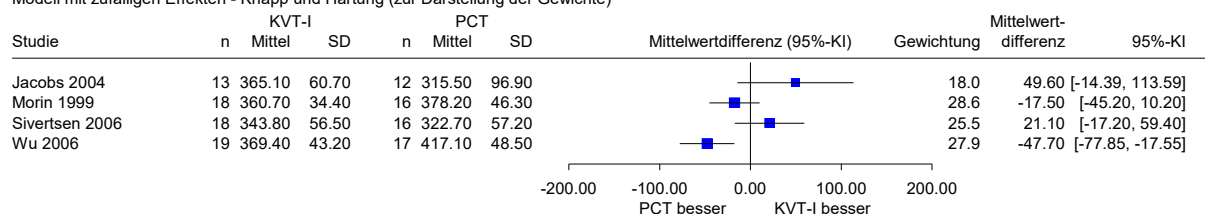


Abbildung 7: Metaanalyse, Schlafdauer in Minuten (Polysomnographie) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

5.4.6.2 Subgruppenmerkmale und andere Effektmodifikatoren

KVT-I vs. PCT - Behandlungsende
Schlafdauer - Schlaftagebuch

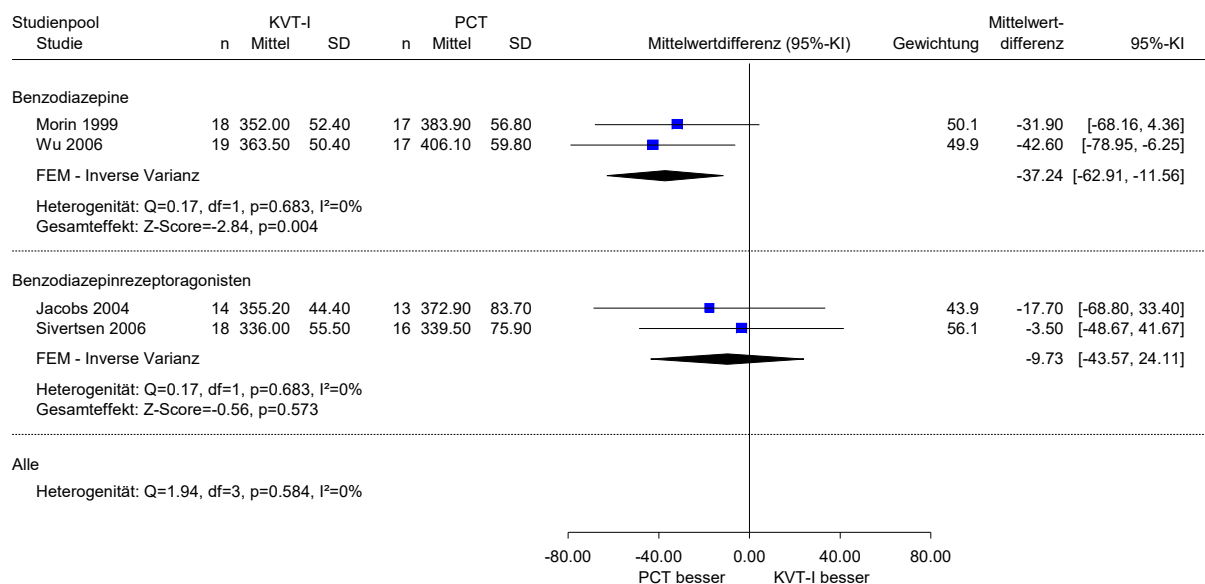


Abbildung 8: Subgruppenanalyse nach Wirkstoffgruppe, Schlafdauer in Minuten (Schlaftagebuch) – Interventionsende, KVT-I vs. Pharmakotherapie; Effektmaß: MWD

KVT-I vs. PCT - Behandlungsende
Schlafdauer - Polysomnographie

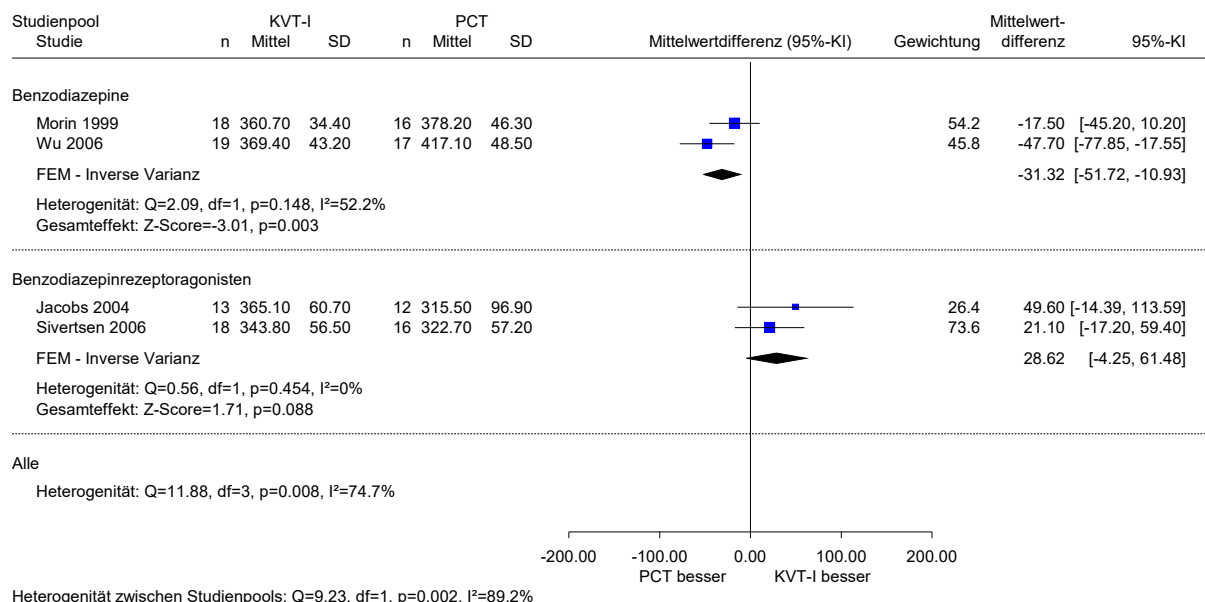


Abbildung 9: Subgruppenanalyse nach Wirkstoffgruppe, Schlafdauer in Minuten (Polysomnographie) – Interventionsende, KVT-I vs. Pharmakotherapie ; Effektmaß: MWD

6 Anmerkungen zu den Ergebnissen

Es konnten insgesamt 5 RCTs zur Behandlung einer nicht organischen Insomnie mittels KVT-I im Vergleich zu Benzodiazepinen bzw. Benodiazepinrezeptoragonisten identifiziert werden. In 2 der 5 Studien wurde das Benzodiazepin Temazepam als Vergleichsintervention verwendet [22,26], in 1 Studie wurde der Benodiazepinrezeptoragonist Zolpidem [20] und bei 2 Studien Zopiclon [25] als Vergleich untersucht.

Die Ergebnisse wiesen für die Endpunkte Schlaffeffizienz und Schlafdauer eine erhebliche Heterogenität auf, welche zum Großteil durch Subgruppenanalysen nach Wirkstoffgruppen erklärt werden konnte. Aufgrund dieser Effektmodifikation wurden für alle Endpunkte getrennte Effektschätzer für die Vergleiche KVT-I vs. Benzodiazepine und KVT-I vs. Benzodiazepinrezeptoragonisten dargestellt. Eine mögliche Erklärung für die signifikant längere Schlafdauer und höhere Schlaffeffizienz unter der Gabe von Benzodiazepinen könnte die höhere Halbwertszeit und die damit einhergehende längere Wirkungsdauer des Benzodiazepins Temazepam im Vergleich zu den Benzodiazepinrezeptoragonisten Zolpidem und Zopiclon sein.

Für alle Studien erfolgte eine Prüfung, ob die Dosierung, Art und Dauer der Anwendung der Vergleichsinterventionen mit der Zulassung in Deutschland konform sind. Eine relevante Abweichung zur Zulassung in Deutschland zeigte sich lediglich in einer Studie. In Sivertsen 2006 erhielten die Patientinnen und Patienten von Beginn an 7,5 mg Zopiclon pro Tag anstatt der für die Population der älteren Patientinnen und Patienten empfohlenen initialen Dosis von 3,75 mg. Aus diesem Grund erfolgte in den Evidenzprofilen an entsprechenden Stellen eine Abwertung aufgrund von Indirektheit.

Obwohl alle Studien Daten zu Nachbeobachtungzeitpunkten berichteten, ist die Evidenz für die anhaltende Wirkung nach Interventionsende stark limitiert. Dies ist darauf zurückzuführen, dass der Anteil der Patientinnen und Patienten, für die keine Daten vorliegen, im Verlauf der Studien deutlich steigt, sodass der Anteil der berücksichtigten Patientinnen und Patienten zu den Nachbeobachtungszeitpunkten weniger als 70 % betrug bzw. der Unterschied zwischen den Gruppen 15%-Punkte überstieg. Das Erreichen dieser Grenzen wird zusätzlich durch die geringe Studiengröße verstärkt. Daher war der Großteil der berichteten Ergebnisse der Nachbeobachtungen nicht verwertbar und wurde in diesem Evidenzbericht nicht berücksichtigt.

7 Literatur

1. Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG). Bundesgesetzblatt Teil 1 2019; (49): 2562-2584.
2. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Nicht erholsamer Schlaf/Schlafstörungen - Insomnie bei Erwachsenen [online]. 2017 [Zugriff: 03.05.2021]. URL: <https://www.awmf.org/leitlinien/detail/ll/063-003.html>.
3. Schünemann H, Brożek J, Guyatt G et al. GRADE handbook [online]. 2013 [Zugriff: 19.02.2021]. URL: <https://gdt.gradepro.org/app/handbook/handbook.html>.
4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Entwurf für Version 6.1 [online]. 2021 [Zugriff: 30.08.2021]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_entwurf-fuer-version-6-1.pdf.
5. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 19.02.2021]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.
6. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. <https://dx.doi.org/10.1136/bmj.c869>.
7. Guyatt GH, Oxman AD, Santesso N et al. GRADE guidelines: 12. Preparing summary of findings tables—binary outcomes. J Clin Epidemiol 2013; 66(2): 158-172. <https://dx.doi.org/10.1016/j.jclinepi.2012.01.012>.
8. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 1998; 17(8): 873-890. [https://dx.doi.org/10.1002/\(sici\)1097-0258\(19980430\)17:8<873::aid-sim779>3.0.co;2-i](https://dx.doi.org/10.1002/(sici)1097-0258(19980430)17:8<873::aid-sim779>3.0.co;2-i).
9. Andrés AM, Mato AS. Choosing the optimal unconditioned test for comparing two independent proportions. Comput Stat Data Anal 1994; 17(5): 555-574. [https://dx.doi.org/https://doi.org/10.1016/0167-9473\(94\)90148-1](https://dx.doi.org/https://doi.org/10.1016/0167-9473(94)90148-1).
10. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.
11. Veroniki AA, Jackson D, Viechtbauer W et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015; (10 Suppl 1): 25-27.
12. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. J Clin Epidemiol 2011; 64(12): 1294-1302. <https://dx.doi.org/10.1016/j.jclinepi.2011.03.017>.

13. Balshem H, Helfand M, Schunemann HJ et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64(4): 401-406.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.015>.
14. Meerpohl JJ, Langer G, Perleth M et al. GRADE-Leitlinien: 3. Bewertung der Qualität der Evidenz (Vertrauen in die Effektschätzer). *Z Evid Fortbild Qual Gesundhwes* 2012; 106(6): 449-456. <https://dx.doi.org/10.1016/j.zefq.2012.06.013>.
15. Guyatt GH, Oxman AD, Vist G et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol* 2011; 64(4): 407-415.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.017>.
16. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011; 64(12): 1303-1310.
<https://dx.doi.org/10.1016/j.jclinepi.2011.04.014>.
17. Guyatt GH, Oxman AD, Montori V et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol* 2011; 64(12): 1277-1282.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.011>.
18. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines 6. Rating the quality of evidenced—imprecision. *J Clin Epidemiol* 2011; 64(12): 1283-1293.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.012>.
19. Guyatt GH, Oxman AD, Sultan S et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12): 1311-1316.
<https://dx.doi.org/10.1016/j.jclinepi.2011.06.004>.
20. Jacobs GD, Pace-Schott EF, Stickgold R et al. Cognitive behavior therapy and pharmacotherapy for insomnia: a randomized controlled trial and direct comparison. *Arch Intern Med* 2004; 164(17): 1888-1896. <https://dx.doi.org/10.1001/archinte.164.17.1888>.
21. Morin CM, Blais F, Savard J. Are changes in beliefs and attitudes about sleep related to sleep improvements in the treatment of insomnia? *Behav Res Ther* 2002; 40(7): 741-752.
[https://dx.doi.org/10.1016/s0005-7967\(01\)00055-9](https://dx.doi.org/10.1016/s0005-7967(01)00055-9).
22. Morin CM, Colecchi C, Stone J et al. Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. *JAMA* 1999; 281(11): 991-999.
<https://dx.doi.org/10.1001/jama.281.11.991>.
23. Pchelina PV, Tabidze AA, Poluektov MG. A Comparative Study of the Efficacy of Cognitive Behavioral Therapy and Zopiclone in Chronic Insomnia. *Neuroscience and Behavioral Physiology* 2019; 49(1): 38-44. <https://dx.doi.org/10.1007/s11055-018-0688-z>.
24. Omvik S, Sivertsen B, Pallesen S et al. Daytime functioning in older patients suffering from chronic insomnia: treatment outcome in a randomized controlled trial comparing CBT with Zopiclone. *Behav Res Ther* 2008; 46(5): 623-641.
<https://dx.doi.org/10.1016/j.brat.2008.02.013>.

25. Sivertsen B, Omvik S, Pallesen S et al. Cognitive behavioral therapy vs zopiclone for treatment of chronic primary insomnia in older adults: a randomized controlled trial. JAMA 2006; 295(24): 2851-2858. <https://dx.doi.org/10.1001/jama.295.24.2851>.

26. Wu R, Bao J, Zhang C et al. Comparison of sleep condition and sleep-related psychological activity after cognitive-behavior and pharmacological therapy for chronic insomnia. Psychother Psychosom 2006; 75(4): 220-228. <https://dx.doi.org/10.1159/000092892>.

27. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

28. Lefebvre C, Glanville J, Briscoe S et al. Cochrane Handbook for Systematic Reviews of Interventions; Version 6; Technical Supplement to Chapter 4: Searching for and selecting studies [online]. 2019 [Zugriff: 19.02.2021]. URL: <https://training.cochrane.org/handbook/version-6/chapter-4-tech-suppl>.

8 Studienlisten

8.1 Liste der identifizierten systematischen Übersichten

1. Benz F, Knoop T, Ballesio A et al. The efficacy of cognitive and behavior therapies for insomnia on daytime symptoms: A systematic review and network meta-analysis. Clin Psychol Rev 2020; 80: 101873. <https://dx.doi.org/10.1016/j.cpr.2020.101873>.
2. Brasure M, MacDonald R, Fuchs E et al. Management of Insomnia Disorder; AHRQ Publication No. 15(16)-EHC027-EF [online]. 2015 [Zugriff: 29.06.2021]. URL: <https://www.ncbi.nlm.nih.gov/books/n/cer159/pdf/>.
3. Ballesio A, Bacaro V, Vacca M et al. Does cognitive behaviour therapy for insomnia reduce repetitive negative thinking and sleep-related worry beliefs? A systematic review and meta-analysis. Sleep Med Rev 2021; 55: 101378. <https://dx.doi.org/10.1016/j.smr.2020.101378>.
4. Cheung JMY, Jarrin DC, Ballot O et al. A systematic review of cognitive behavioral therapy for insomnia implemented in primary care and community settings. Sleep Med Rev 2019; 44: 23-36. <https://dx.doi.org/10.1016/j.smr.2018.11.001>.
5. Herbert V, Kyle SD, Pratt D. Does cognitive behavioural therapy for insomnia improve cognitive performance? A systematic review and narrative synthesis. Sleep Med Rev 2018; 39: 37-51. <https://dx.doi.org/10.1016/j.smr.2017.07.001>.
6. Kwon CY, Lee B, Cheong MJ et al. Non-pharmacological Treatment for Elderly Individuals With Insomnia: A Systematic Review and Network Meta-Analysis. Front Psychiatry 2020; 11: 608896. <https://dx.doi.org/10.3389/fpsy.2020.608896>.
7. Mitchell MD, Gehrman P, Perlis M et al. Comparative effectiveness of cognitive behavioral therapy for insomnia: a systematic review. BMC Fam Pract 2012; 13: 40. <https://dx.doi.org/10.1186/1471-2296-13-40>.
8. Wang YY, Yang Y, Rao WW et al. Cognitive behavioural therapy monotherapy for insomnia: A meta-analysis of randomized controlled trials. Asian J Psychiatr 2020; 49: 101828. <https://dx.doi.org/10.1016/j.ajp.2019.10.008>.

8.2 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

Nicht EI

1. Deroose SF, Rozema E, Chen A et al. A population health approach to insomnia using internet-based cognitive behavioral therapy for insomnia. J Clin Sleep Med 2021. <https://dx.doi.org/10.5664/jcsm.9280>.
2. Koffel E, Kuhn E, Petsoulis N et al. A randomized controlled pilot study of CBT-I Coach: Feasibility, acceptability, and potential impact of a mobile phone application for patients in cognitive behavioral therapy for insomnia. Health Informatics J 2018; 24(1): 3-13. <https://dx.doi.org/10.1177/1460458216656472>.

3. Taylor DJ, Dietch JR, Pruiksma K et al. Developing and Testing a Web-Based Provider Training for Cognitive Behavioral Therapy of Insomnia. *Mil Med* 2021; 186(Suppl 1): 230-238. <https://dx.doi.org/10.1093/milmed/usaa359>.

Nicht E2

1. Arnedt JT, Conroy DA, Mooney A et al. Telemedicine versus face-to-face delivery of cognitive behavioral therapy for insomnia: a randomized controlled noninferiority trial. *Sleep* 2021; 44(1). <https://dx.doi.org/10.1093/sleep/zsaa136>.

2. Blom K, Tarkian Tillgren H, Wiklund T et al. Internet-vs. group-delivered cognitive behavior therapy for insomnia: A randomized controlled non-inferiority trial. *Behav Res Ther* 2015; 70: 47-55. <https://dx.doi.org/10.1016/j.brat.2015.05.002>.

3. Kjorstad K, Sivertsen B, Vedaa O et al. The Effect of Reducing Insomnia Severity on Work- and Activity-Related Impairment. *Behavioral Sleep Medicine* 2020: 1-11. <https://dx.doi.org/10.1080/15402002.2020.1799792>.

4. Lancee J, Effting M, van der Zweerde T et al. Cognitive processes mediate the effects of insomnia treatment: evidence from a randomized wait-list controlled trial. *Sleep Med* 2019; 54: 86-93. <https://dx.doi.org/10.1016/j.sleep.2018.09.029>.

5. Lancee J, van Straten A, Morina N et al. Guided Online or Face-to-Face Cognitive Behavioral Treatment for Insomnia: A Randomized Wait-List Controlled Trial. *Sleep* 2016; 39(1): 183-191. <https://dx.doi.org/10.5665/sleep.5344>.

6. McClusky HY, Milby JB, Switzer PK et al. Efficacy of behavioral versus triazolam treatment in persistent sleep-onset insomnia. *Am J Psychiatry* 1991; 148(1): 121-126. <https://dx.doi.org/10.1176/ajp.148.1.121>.

7. Miranda J, Sidani S, Fredericks S et al. Sleep and cardiovascular effects of behavioural therapies for insomnia. *British Journal of Cardiac Nursing* 2017; 12(10): 488-495. <https://dx.doi.org/10.12968/bjca.2017.12.10.488>.

8. Morin CM, Edinger JD, Beaulieu-Bonneau S et al. Effectiveness of Sequential Psychological and Medication Therapies for Insomnia Disorder: A Randomized Clinical Trial. *JAMA Psychiatry* 2020; 77(11): 1107-1115. <https://dx.doi.org/10.1001/jamapsychiatry.2020.1767>.

9. Salomonsson S, Santoft F, Lindsater E et al. Stepped care in primary care - guided self-help and face-to-face cognitive behavioural therapy for common mental disorders: a randomized controlled trial. *Psychol Med* 2018; 48(10): 1644-1654. <https://dx.doi.org/10.1017/s0033291717003129>.

10. Taylor DJ, Peterson AL, Pruiksma KE et al. Impact of cognitive behavioral therapy for insomnia disorder on sleep and comorbid symptoms in military personnel: a randomized clinical trial. *Sleep* 2018; 41(6). <https://dx.doi.org/10.1093/sleep/zsy069>.

11. Taylor DJ, Peterson AL, Pruiksma KE et al. Internet and In-Person Cognitive Behavioral Therapy for Insomnia in Military Personnel: A Randomized Clinical Trial. *Sleep* 2017; 40(6). <https://dx.doi.org/10.1093/sleep/zsx075>.

Nicht E3

1. Freeman D, Sheaves B, Goodwin GM et al. The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. *Lancet Psychiatry* 2017; 4(10): 749-758. [https://dx.doi.org/10.1016/s2215-0366\(17\)30328-0](https://dx.doi.org/10.1016/s2215-0366(17)30328-0).

2. Jansson M, Linton SJ. Cognitive-behavioral group therapy as an early intervention for insomnia: a randomized controlled trial. *Journal of Occupational Rehabilitation* 2005; 15(2): 177-190. <https://dx.doi.org/10.1007/s10926-005-1217-9>.

3. Krieger T, Urech A, Duss SB et al. A randomized controlled trial comparing guided internet-based multi-component treatment and internet-based guided sleep restriction treatment to care as usual in insomnia. *Sleep Med* 2019; 62: 43-52. <https://dx.doi.org/10.1016/j.sleep.2019.01.045>.

4. Lancee J, van den Bout J, van Straten A et al. Internet-delivered or mailed self-help treatment for insomnia?: a randomized waiting-list controlled trial. *Behav Res Ther* 2012; 50(1): 22-29. <https://dx.doi.org/10.1016/j.brat.2011.09.012>.

5. Morin CM, Azrin NH. Behavioral and cognitive treatments of geriatric insomnia. *J Consult Clin Psychol* 1988; 56(5): 748-753. <https://dx.doi.org/10.1037//0022-006x.56.5.748>.

6. Steurer J. Insomnie: Eine einfache Verhaltenstherapie verbessert den Schlaf. *Praxis* 2012; 101(5): 343-344. <https://dx.doi.org/10.1024/1661-8157/a000854>.

7. Van der Zweerde T, Lancee J, Slottje P et al. Nurse-Guided Internet-Delivered Cognitive Behavioral Therapy for Insomnia in General Practice: Results from a Pragmatic Randomized Clinical Trial. *Psychother Psychosom* 2020; 89(3): 174-184. <https://dx.doi.org/10.1159/000505600>.

8. Vincent N, Walsh K. Hyperarousal, sleep scheduling, and time awake in bed as mediators of outcome in computerized cognitive-behavioral therapy (cCBT) for insomnia. *Behav Res Ther* 2013; 51(3): 161-166. <https://dx.doi.org/10.1016/j.brat.2012.12.003>.

9. Wong KY, Chung KF, Au CH. Low-Intensity Cognitive Behavioral Therapy for Insomnia as the Entry of the Stepped-Care Model in the Community: A Randomized Controlled Trial. *Behavioral Sleep Medicine* 2021; 19(3): 378-394. <https://dx.doi.org/10.1080/15402002.2020.1764000>.

Nicht E5

1. Balzer K, Hesse K, Eisold U et al. [Better health for body and soul]. *Pflege Z* 2013; 66(11): 652-655.

2. Epstein DR. A behavioral intervention to enhance the sleep-wake patterns of older adults with insomnia. University of Arizona; 1994.

3. Krystal AD. The treatment of primary insomnia. *CNS Spectr* 2009; 14(12 Suppl 13): 6-10. <https://dx.doi.org/10.1017/s1092852900003953>.
4. Peter L, Reindl R, Zauter S et al. Effectiveness of an Online CBT-I Intervention and a Face-to-Face Treatment for Shift Work Sleep Disorder: A Comparison of Sleep Diary Data. *Int J Environ Res Public Health* 2019; 16(17). <https://dx.doi.org/10.3390/ijerph16173081>.
5. Unnewehr S, Winter C, Kraemer S et al. Wirksamkeit einer verhaltenstherapeutischen Gruppenbehandlung bei Patienten mit primären Insomnien. *Verhaltenstherapie* 1997; 7(2): 68-80. <https://dx.doi.org/10.1159/000259012>.
6. Vincent N, Walsh K, Lewycky S. Sleep locus of control and computerized cognitive-behavioral therapy (cCBT). *Behav Res Ther* 2010; 48(8): 779-783. <https://dx.doi.org/10.1016/j.brat.2010.05.006>.

Nicht E6

1. Azimi M, Moradi A, Hasani J. Effectiveness of cognitive behavioral therapy for insomnia (traditional and Internet-based) on everyday memory of people with insomnia and comorbid depression. *Advances in Cognitive Science* 2019; 20(4): 20-34.

Nicht E7

1. Behaviour therapy for chronic insomnia. *Prescrire Int* 2001; 10(51): 23.
2. Cognitive therapy superior to zopiclone for insomnia. *J Fam Pract* 2006; 55(10): 845.
3. Dekker K, Benjamins JS, Maksimovic T et al. Combined Internet-Based Cognitive-Behavioral and Chronobiological Intervention for Insomnia: A Randomized Controlled Trial. *Psychother Psychosom* 2020; 89(2): 117-118. <https://dx.doi.org/10.1159/000503570>.
4. Morgan K, Gregory P, Tomeny M et al. Self-help CBT-I in the management of insomnia symptoms associated with chronic disease in older adults: a randomized controlled trial. *J Sleep Res* 2010; 19(Suppl s2): 79.
5. Morin CM, Colecche C, Stone J et al. Cognitive behavior therapy, temazepam, or both improved short-term outcomes for older adults with persistent insomnia. *Evidence Based Medicine* 1999; 4(5): 146.
6. Pchelina P, Poluektov M, Berger T et al. Effectiveness and Cost-Effectiveness of Internet-Based Cognitive Behavioral Therapy for Insomnia in Clinical Settings. *Front Psychiatr* 2020; 11: 838. <https://dx.doi.org/10.3389/fpsy.2020.00838>.
7. Pimlott NJ. Pharmacologic or behavioural therapy for elderly people's insomnia. Which is better? *Can Fam Physician* 2000; 46: 1430-1432.
8. Shaughnessy AF. CBT Effective for Chronic Insomnia. *Am Fam Physician* 2016; 93(1): 60.
9. CBT may be more effective than zopiclone for insomnia in the elderly. *Brown University Geriatric Psychopharmacology Update* 2006; 10(9): 1-4.

Anhang A Bewertung der Studienlimitationen

Tabelle 25: Studienlimitationen zum Endpunkt Krankheitsschwere

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Morin 1999	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
Pchelina 2019	adäquat	unklar	nicht adäquat	unklar	unklar	unklar ^a
a. Wenige und nicht übereinstimmende Angaben zu Baselinewerten der Patientinnen und Patienten. ITT: Intention to treat						

Tabelle 26: Studienlimitationen zum Endpunkt Schlaffeffizienz

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Jacobs 2004	adäquat	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
Morin 1999	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
Sivertsen 2006	unklar	unklar	nicht adäquat	adäquat / nicht adäquat ^a	nicht adäquat	adäquat
Wu 2006	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
a. Für den subjektiven und den objektiven Endpunkt zum Nachbeobachtungzeitpunkt 6 Monate, ist die Differenz der Anteile der Patientinnen und Patienten, welche in der Analyse unberücksichtigt bleiben, groß. ITT: Intention to treat						

Tabelle 27: Studienlimitationen zum Endpunkt Einschlaf latenz

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Jacobs 2004	adäquat	unklar	nicht adäquat	adäquat / nicht adäquat ^a	unklar	adäquat
Wu 2006	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
a. Für den subjektiven Endpunkt zum Nachbeobachtungzeitpunkt 1 Monat, sowie für den objektiven Endpunkt nach Interventionsende ist der Anteil der Patientinnen und Patienten welche in der Analyse unberücksichtigt bleiben groß. ITT: Intention to treat						

Tabelle 28: Studienlimitationen zum Endpunkt Wachzeit nach dem Einschlafen

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Morin 1999	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
a. Für den subjektiven und den objektiven Endpunkt zum Nachbeobachtungzeitpunkt 6 Monate, ist die Differenz der Anteile der Patientinnen und Patienten welche in der Analyse unberücksichtigt bleiben groß. ITT: Intention to treat						

Tabelle 29: Studienlimitationen zum Endpunkt Schlafqualität

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Pchelina 2019	adäquat	unklar	nicht adäquat	unklar	unklar	unklar ^a
a. Wenige und nicht übereinstimmende Angaben zu Baselinewerten der Patientinnen und Patienten. ITT: Intention to treat						

Tabelle 30: Studienlimitationen zum Endpunkt Schlafdauer

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Jacobs 2004	adäquat	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
Morin 1999	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
Sivertsen 2006	unklar	unklar	nicht adäquat	adäquat / nicht adäquat ^a	nicht adäquat	adäquat
Wu 2006	unklar	unklar	nicht adäquat	nicht adäquat	unklar	adäquat
a. Für den subjektiven und den objektiven Endpunkt zum Nachbeobachtungzeitpunkt 6 Monate, ist die Differenz der Anteile der Patientinnen und Patienten welche in der Analyse unberücksichtigt bleiben groß. ITT: Intention to treat						

Anhang B Suchstrategien

B.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. PubMed

Suchoberfläche: NLM

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [27] – High specificity strategy (adaptiert für PubMed)

#	Searches
1	insomnia
2	cognitive behavior therapy
3	"The Cochrane database of systematic reviews"[Journal] OR search[tiab] OR MEDLINE[tiab] OR systematic review[tiab] OR "meta-analysis"[pt:noexp]
4	#1 AND #2 AND #3
5	#4 AND 2011:2021[DP]
6	#5 AND (english[LA] OR german[LA])

2. Health Technology Assessment Database

Suchoberfläche: INAHTA

#	Searches
1	insomnia
2	(sleep*)[title]
3	#2 OR #1
4	cognitive AND behavio* AND therapy
5	#4 AND #3
6	(cognitive AND behavio* AND therapy) AND (((sleep*)[title]) OR (insomnia)) FROM 2011 TO 2021

Suche nach Primärstudien

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) 1946 to May 27, 2021

Es wurde folgender Filter übernommen:

- RCT: Lefebvre [28] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision)

#	Searches
1	"Sleep Initiation and Maintenance Disorders"/
2	insomnia*.ti,ab.
3	or/1-2
4	Cognitive Behavioral Therapy/
5	((behavioral or behavior or behavioural or behaviour) adj3 (therap* or treatment* or intervention*)).ti,ab.
6	CBT.ti,ab.
7	or/4-6
8	Randomized Controlled Trial.pt.
9	Controlled Clinical Trial.pt.
10	(randomized or placebo or randomly).ab.
11	Clinical Trials as Topic/
12	trial.ti.
13	or/8-12
14	13 not (exp animals/ not humans.sh.)
15	and/3,7,14
16	15 not (comment or editorial).pt.
17	16 and (english or german).lg.

Suchoberfläche: Ovid

- Ovid MEDLINE(R) Epub Ahead of Print and In-Process, In-Data-Review & Other Non-Indexed Citations May 27, 2021

#	Searches
1	insomnia*.ti,ab.
2	((behavioral or behavior or behavioural or behaviour) adj3 (therap* or treatment* or intervention*)).ti,ab.
3	CBT.ti,ab.
4	or/2-3
5	(clinical trial* or random* or placebo).ti,ab.
6	trial.ti.
7	or/5-6
8	and/1,4,7
9	8 not (comment or editorial).pt.
10	9 and (english or german).lg.

2. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials: Issue 4 of 12, April 2021

#	Searches
#1	[mh ^"Sleep Initiation and Maintenance Disorders"]
#2	insomnia*:ti,ab
#3	#1 or #2
#4	[mh ^"Cognitive Behavioral Therapy"]
#5	((behavioral or behavior or behavioural or behaviour) NEAR/3 (therap* or treatment* or Intervention*)):ti,ab
#6	CBT:ti,ab
#7	#4 or #5 or #6
#8	#3 and #7
#9	#8 not (*clinicaltrial*gov* or *who*trialsearch* or *clinicaltrialsregister*eu* or *anzctr*org*au* or *trialregister*nl* or *irct*ir* or *isrctn* or *controlled*trials*com* or *drks*de*):so
#10	#9 not ((language next (afr or ara or aze or bos or bul or car or cat or chi or cze or dan or dut or es or est or fin or fre or gre or heb or hrv or hun or ice or ira or ita or jpn or ko or kor or lit or nor or peo or per or pol or por or pt or rom or rum or rus or slo or slv or spa or srp or swe or tha or tur or ukr or urd or uzb)) not (language near/2 (en or eng or english or ger or german or mul or unknown))) in Trials

B.2 Studienregister

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

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

Suchstrategie
insomnia AND (Cognitive Behavioral Therapy OR CBT) AND AREA[ResultsFirstSubmitDate] NOT MISSING

2. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <http://apps.who.int/trialsearch>
- Eingabeoberfläche: Standard Search / Advanced Search

Suchstrategie
insomnia AND behavior* OR insomnia AND behaviour* OR insomnia AND CBT (with Results only)