

Screening auf Depression

Berichtsplan

Auftrag: S16-05
Version: 1.0
Stand: 07.08.2017

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Screening auf Depression

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

27.04.2017

Interne Auftragsnummer:

S16-05

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

Zu diesem Berichtsplan können alle interessierten Personen, Institutionen und Gesellschaften Stellungnahmen abgeben, die zu einer Überarbeitung des Berichtsplans führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG (www.iqwig.de), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

Schlagwörter: Reihenuntersuchung, Depressive Störung, Nutzenbewertung, Systematische Übersicht

Keywords: Mass Screening, Depressive Disorder, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis	vi
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 zur Screeningkette	4
4.1.1 Population.....	4
4.1.2 Prüf- und Vergleichsintervention	4
4.1.3 Patientenrelevante Endpunkte	4
4.1.4 Studientypen	5
4.1.5 Studiendauer	5
4.1.6 Tabellarische Darstellung der Kriterien für den Einschluss von Studien zur Screeningkette	5
4.2 Kriterien für den Einschluss von Studien zur Therapievorverlagerung	7
4.2.1 Population.....	7
4.2.2 Prüf- und Vergleichsintervention	7
4.2.3 Patientenrelevante Endpunkte	7
4.2.4 Studientypen	8
4.2.5 Studiendauer	8
4.2.6 Tabellarische Darstellung der Kriterien für den Einschluss von Studien zur Therapievorverlagerung	8
4.3 Kriterien für den Einschluss von systematischen Übersichten zur diagnostischen Güte	9
4.3.1 Population.....	9
4.3.2 Indextest	9
4.3.3 Referenztest	9
4.3.4 Zielgrößen	9
4.3.5 Studientypen	9
4.3.6 Tabellarische Darstellung der Kriterien für den Einschluss systematischer Übersichten.....	10
4.4 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	10

4.5 Informationsbeschaffung.....	11
4.5.1 Umfassende Informationsbeschaffung	11
4.5.1.1 Bibliografische Recherche.....	11
4.5.1.2 Öffentlich zugängliche Studienregister	12
4.5.1.3 Weitere Suchquellen.....	12
4.5.1.3.1 Systematische Übersichten.....	12
4.5.1.3.2 Anhörung.....	12
4.5.1.3.3 Autorenanfragen.....	12
4.5.2 Fokussierte Informationsbeschaffung	12
4.5.3 Selektion relevanter Studien und systematischer Übersichten.....	12
4.6 Informationsbewertung.....	13
4.6.1 Bewertung der Screening- und Therapiestudien	13
4.6.2 Bewertung von systematischen Übersichten	15
4.6.2.1 Bewertung von systematischen Übersichten zu Screeningstudien.....	15
4.6.2.2 Bewertung von systematischen Übersichten zur diagnostischen Güte.....	15
4.7 Informationssynthese und -analyse für Screening- und Therapiestudien.....	15
4.7.1 Gegenüberstellung der Ergebnisse der Einzelstudien	15
4.7.2 Metaanalysen.....	16
4.7.3 Aussagen zur Beleglage	16
4.7.4 Sensitivitätsanalysen	17
4.7.5 Subgruppenmerkmale und andere Effektmodifikatoren	17
4.8 Informationssynthese aus systematischen Übersichten zur diagnostischen Güte	18
4.8.1 Darstellung der Ergebnisse.....	18
4.8.2 Aussagen zum Nutzen des Screenings auf Basis gemeinsamer Betrachtung von Studien zur Therapievorverlagerung und der diagnostischen Güte von Screeningtests	18
5 Literatur	20
Anhang A – Offenlegung potenzieller Interessenkonflikte von Betroffenen	22

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Einschluss von Studien zur Screeningkette	6
Tabelle 2: Übersicht über die Kriterien für die Berücksichtigung von systematischen Übersichten zur Screeningkette.....	6
Tabelle 3: Übersicht über die Kriterien für den Einschluss von Studien zur Therapievorverlagerung	8
Tabelle 4: Übersicht über die Kriterien für den Einschluss systematischer Übersichten, die sich auf die diagnostische Güte beziehen.....	10
Tabelle 5: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit.....	17

Abkürzungsverzeichnis

Abkürzung	Bedeutung
CIDI	Composite International Diagnostic Interview
DIPS	Diagnostisches Interview bei psychischen Störungen
DSM	Diagnostisches und statistisches Manual psychischer Störungen
DSM-5	Diagnostisches und statistisches Manual psychischer Störungen, 5. Auflage
G-BA	Gemeinsamer Bundesausschuss
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Revision 10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Ausgabe)
ICH	International Conference of Harmonization
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
SCID	Structured Clinical Interview for DSM Disorders (strukturiertes klinisches Interview für DSM)
SÜ	systematische Übersicht

1 Hintergrund

Nach Definition der Weltgesundheitsorganisation versteht man unter einer Depression eine „psychische Störung, die durch Traurigkeit, Interesselosigkeit und Verlust an Genussfähigkeit, Schuldgefühle und geringes Selbstwertgefühl, Schlafstörungen, Appetitlosigkeit, Müdigkeit und Konzentrationsschwächen gekennzeichnet sein kann“ [1].

Für die Entstehung einer Depression werden multifaktorielle Erklärungskonzepte angenommen [2]: Jeder Depression liegt eine Mischung aus fehlangepassten, situationsbedingten und körperlichen Ursachen zugrunde [3]. Die Diagnosestellung einer Depression erfolgt nach den international anerkannten Kriterien des ICD-10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme) oder des DSM-5 (Diagnostisches und statistisches Manual psychischer Störungen) [3]. Standardisierte diagnostische Interviews gelten als Goldstandard, um diese Kriterien regelhaft umzusetzen, da sie alle Aspekte vollständig abdecken – angefangen bei der Erfassung und Beurteilung der Symptome bis hin zu ihrer algorithmischen und differenzialdiagnostischen Ableitung [3].

Die Depression kann als einzelne depressive Episode, als rezidivierende depressive Störung mit 2 bzw. mehreren depressiver Episoden oder als anhaltend depressive Störung auftreten. [3]. In diesem Bericht wird ausschließlich die unipolare Depression berücksichtigt.

Durch die Depression kommt es für die Betroffenen oft zu einer massiven Beeinträchtigung ihrer Lebensführung: Depressive Störungen führen zu einer starken Beeinträchtigung der körperlichen und psychischen Befindlichkeit und gehen mit einer hohen Mortalität, insbesondere durch Suizide, einher. Darüber hinaus wirken sich depressive Störungen auf die sozialen Beziehungen und die Arbeitsfähigkeit der Betroffenen aus [2].

Als primäre Behandlungsstrategien der depressiven Störung stehen neben der aktiv-abwartenden Begleitung durch „watchful waiting“ oder „niederschwellige psychosoziale Interventionen“ bei einer leichten depressiven Episode eine medikamentöse oder psychotherapeutische Behandlung sowie eine Kombinationstherapie zur Verfügung [2].

Aktuelle repräsentative Daten aus der 1. Welle der „Studie zur Gesundheit Erwachsener in Deutschland“ (DEGS) [4] zur Verbreitung der Depression in der Allgemeinbevölkerung im Alter zwischen 18 und 79 Jahren zeigen eine Lebenszeitprävalenz von insgesamt 11,6 % für eine diagnostizierte Depression, wobei Frauen mit 15,4 % fast doppelt so häufig eine jemals diagnostizierte Depression wie Männer mit 7,8 % angeben. Die 12-Monats-Prävalenz auf Basis der DEGS-Daten liegt bei 6,0 %, wobei auch hier die Prävalenz bei Frauen etwa doppelt so hoch ist wie bei den Männern (8,1 % versus 3,8 %) [4]. Für die Dysthymie ergibt sich eine 12-Monats-Prävalenz von 2 % [5].

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung eines systematischen, im Rahmen der hausärztlichen Versorgung erbrachten Screenings auf Depression im Vergleich zu keinem Screening hinsichtlich patientenrelevanter Endpunkte.

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat am 27.04.2017 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung des Screenings auf Depression beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 08.06.2017 Betroffene zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen konsultiert. Die Selbstangaben der Betroffenen zu potenziellen Interessenkonflikten sind in Anhang A dargestellt.

Der vorliegende Berichtsplan (Version 1.0) wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekanntgegeben. Stellungnahmen können alle interessierten Personen, Institutionen und Gesellschaften abgeben. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen. Diese sind ebenfalls auf der Website des IQWiG in einem Leitfaden dargelegt. Kommt es durch die Anhörung zu einer Änderung der Berichtsmethodik, wird ein überarbeiteter Berichtsplan erstellt und veröffentlicht. Die Dokumentation der Anhörung zum Berichtsplan wird auf der Website des IQWiG veröffentlicht.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Der Vorbericht wird gegebenenfalls zusätzlich einem externen Review unterzogen. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 8 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation der Anhörung zum Vorbericht veröffentlicht.

Der genaue Zeitplan für alle Arbeitsschritte zur Berichterstellung für dieses Projekt ist auf der Website des IQWiG dargelegt.

4 Methoden

Der Nutzen von Screeningmaßnahmen lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien der gesamten Screeningkette – bestehend aus Screeningtest mit anschließender Diagnostik und einer sich gegebenenfalls anschließenden Behandlung – unter Beachtung patientenrelevanter Endpunkte bewerten. Idealerweise werden Personen einer Gruppe randomisiert zugeteilt [6]. Ist eine Nutzenbewertung auf Basis randomisierter Studien nicht möglich, werden im Rahmen dieser Fragestellung auch nicht randomisierte Interventionsstudien hinzugezogen.

Liegen vergleichende Interventionsstudien der Screeningkette für die Nutzenbewertung nicht oder in nicht ausreichender Quantität und Qualität vor, wird geprüft, ob eine Vorverlagerung der Therapie bei screeningdetektierten Personen mit Depression einen Nutzen zeigt. Für die Bewertung des Nutzens eines Screenings werden den Ergebnissen darüber hinaus die gesundheitsbezogenen Konsequenzen für falsch-positive und falsch-negative Befunde gegenübergestellt. Hierfür erfolgt eine Einschätzung der diagnostischen Güte von Screeningtests auf Depression.

4.1 Kriterien für den Einschluss von Studien zur Screeningkette

4.1.1 Population

In die Bewertung werden Studien mit erwachsenen Personen aufgenommen, bei denen aktuell keine depressive Störung bekannt ist. Nicht betrachtet werden Studien, die über ihre Einschlusskriterien Personen einschließen, die durch einen Risikofaktor für eine depressive Störung wie zum Beispiel eine onkologische Erkrankung oder ein kritisches Lebensereignis charakterisiert werden.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet ein Screening auf Depression. Die zur Verfügung stehenden Optionen für die Behandlung einer Depression müssen in beiden Studienarmen vergleichbar sein, also unabhängig davon, ob die betreffende Person über ein Screening oder auf anderem Weg identifiziert wurde.

Als Vergleichsintervention gilt kein Screening.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität,
- Morbidität: Häufigkeit und Symptome der depressiven Störung,
- gesundheitsbezogene Lebensqualität,
- allgemeines und soziales Funktionsniveau,
- unerwünschte Ereignisse.

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

4.1.4 Studientypen

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

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

Für den zu erstellenden Bericht werden primär RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen. Es wird erwartet, dass zu dieser Fragestellung zwar keine RCTs, wohl aber hochwertige Kohortenstudien (beispielsweise prospektiver Vergleich zweier Regionen) vorliegen. Daher werden gegebenenfalls auch nicht randomisierte prospektiv geplante Interventionsstudien der gesamten Screeningkette mit zeitlich paralleler Kontrollgruppe zur Nutzenbewertung herangezogen.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.1.6 Tabellarische Darstellung der Kriterien für den Einschluss von Studien zur Screeningkette

Zur Identifizierung von Primärstudien zur Screeningkette dienen systematische Übersichten, die die unten angeführten Kriterien erfüllen. Für den aktuellen Zeitraum, der in den Übersichten nicht abgedeckt ist, wird eine bibliografische Recherche nach Primärstudien durchgeführt.

In der folgenden Tabelle sind die Kriterien aufgelistet, die Interventionsstudien der Screeningkette erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Einschluss von Studien zur Screeningkette

Einschlusskriterien	
S1	erwachsene Personen, bei denen aktuell keine depressive Störung bekannt ist (siehe auch Abschnitt 4.1.1)
S2	Prüfintervention: Screening auf Depression (siehe auch Abschnitt 4.1.2)
S3	Vergleichsintervention: kein Screening auf Depression (siehe auch Abschnitt 4.1.2)
S4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
S5	RCTs und nicht randomisierte prospektiv geplante Interventionsstudien der gesamten Screeningkette mit zeitlich paralleler Kontrollgruppe
S6	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang ein Bericht über die Studie, der den Kriterien des CONSORT- [7] oder TREND-Statements [8] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind. CONSORT: Consolidated Standards of Reporting Trials; RCT: randomisierte kontrollierte Studie	

Die Kriterien für die Berücksichtigung der systematischen Übersichten orientieren sich an den Kriterien für den Einschluss von Studien zur Screeningkette (siehe auch Abschnitt 4.1) und beziehen sich auf die in den systematischen Übersichten für die Studienselektion verwendeten Einschlusskriterien.

Tabelle 2: Übersicht über die Kriterien für die Berücksichtigung von systematischen Übersichten zur Screeningkette

Einschlusskriterien	
SSÜ1	Als Zielpopulation wird in den Einschlusskriterien der SÜ definiert: erwachsene Personen, bei denen aktuell keine depressive Störung bekannt ist
SSÜ 2	Als Prüfintervention wird in den Einschlusskriterien der SÜ definiert: Screening auf Depression
SSÜ 3	Als Vergleichsintervention wird in den Einschlusskriterien der SÜ definiert: kein Screening auf Depression
SSÜ4	Als Einschlusskriterien der SÜ werden patientenrelevante Endpunkte definiert wie in Abschnitt 4.1.3 formuliert
SSÜ5	Systematische Übersichten von RCTs und nicht randomisierten prospektiv geplanten Interventionsstudien mit zeitlich parallelen Kontrollgruppen
SSÜ6	Publikationszeitraum: letzte 5 Jahre
SSÜ7	Publikationssprache: Englisch oder Deutsch
SSÜ8	Vollpublikation verfügbar und keine Mehrfachpublikation ohne relevante Zusatzinformation ^a
a: Als Vollpublikation gilt in diesem Zusammenhang ein Bericht, der den Kriterien des PRISMA-Statements [9] genügt. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT: randomisierte kontrollierte Studie; SÜ: systematische Übersicht	

4.2 Kriterien für den Einschluss von Studien zur Therapievorverlagerung

Falls Studien zur Screeningkette nicht oder nicht in ausreichender Quantität und Qualität vorliegen, werden Therapiestudien herangezogen, die einen früheren mit einem späteren Behandlungsbeginn vergleichen.

4.2.1 Population

Die Zielpopulation bilden erwachsene Personen mit einer im Screening detektierten depressiven Störung. Es handelt sich hierbei also um Personen, die sich nicht wegen depressiver Symptome an eine Ärztin oder einen Arzt gewendet haben und für die auch von ärztlicher Seite keine entsprechende Verdachtsdiagnose vorlag. Im Anschluss an das Screening ist eine Diagnosesicherung gemäß DSM oder ICD erforderlich. Hierzu kann unter anderem das Diagnostische Interview bei psychischen Störungen (DIPS), das Composite International Diagnostic Interview (CIDI) oder das Structured Clinical Interview for DSM Disorders (SCID) herangezogen werden. Nicht betrachtet werden Studien mit Personen, die durch einen Risikofaktor für eine depressive Störung wie zum Beispiel eine onkologische Erkrankung oder ein kritisches Lebensereignis charakterisiert werden.

4.2.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention umfasst eine Behandlung mit den in der aktuellen S3-Leitlinie genannten Therapieverfahren zur Behandlung einer depressiven Störung. Dazu zählen unter anderem eine medikamentöse Behandlung, eine psychotherapeutische Behandlung, komplexe Behandlungskonzepte und niederschwellige psychosoziale Interventionen wie zum Beispiel angeleitete Selbsthilfe oder technologiegestützte psychosoziale Interventionen [2].

Als Vergleichsintervention gilt eine Therapie der Depression unter Anwendung desselben Therapieverfahrens wie in der Prüfintervention, wobei dieses später als in der Prüfintervention eingeleitet wurde.

Die in den Studien eingesetzten Prüf- und Vergleichsinterventionen müssen einander hinsichtlich aller Anwendungsmerkmale (z. B. Begleitbehandlung, Therapiedauer) entsprechen und bei einer medikamentösen Behandlung im Rahmen des für Deutschland gültigen Zulassungsstatus durchgeführt werden.

4.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität,
- Morbidität: Häufigkeit und Symptome der depressiven Störung,
- gesundheitsbezogene Lebensqualität,
- allgemeines und soziales Funktionsniveau,
- unerwünschte Ereignisse.

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

4.2.4 Studientypen

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

Für alle unter 4.2.2 genannten Interventionen und alle unter 4.2.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar. Für den zu erstellenden Bericht werden daher ausschließlich RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

4.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.2.6 Tabellarische Darstellung der Kriterien für den Einschluss von Studien zur Therapievorverlagerung

In der folgenden Tabelle sind die Kriterien aufgelistet, die Therapiestudien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 3: Übersicht über die Kriterien für den Einschluss von Studien zur Therapievorverlagerung

Einschlusskriterien	
T1	Patientinnen und Patienten mit einer im Screening detektierten Depression (siehe auch Abschnitt 4.2.1)
T2	Prüfintervention: jede zur Behandlung einer depressiven Störung eingesetzte Intervention (siehe auch Abschnitt 4.2.2)
T3	Vergleichsintervention: später eingeleitete Behandlung der Depression unter Anwendung desselben Therapieverfahrens wie in der Prüfintervention eingesetzt (siehe auch Abschnitt 4.2.2)
T4	patientenrelevante Endpunkte wie in Abschnitt 4.2.3 formuliert
T5	RCTs
T6	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [10] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [7] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und -ergebnissen nicht vertraulich sind. CONSORT: Consolidated Standards of Reporting Trials; ICH: International Conference of Harmonization; RCT: randomisierte kontrollierte Studie	

4.3 Kriterien für den Einschluss von systematischen Übersichten zur diagnostischen Güte

Wenn sich in den Studien zur Screeningkette oder gegebenenfalls in Studien zur Therapievorverlagerung ein Nutzen zeigt, wird die diagnostische Güte von bis zu 5 der in den Studien am häufigsten eingesetzten Screeningtests berichtet. Dies erfolgt ausschließlich auf Basis systematischer Übersichten. Aus diesem Vorgehen können sich die folgenden Limitationen ergeben: Möglicherweise sind in der einzuschließenden systematischen Übersicht auch Studien eingeschlossen, die die diagnostische Güte in einer Risiko-Population für eine Depression erheben; oder es sind auch Studien eingeschlossen, die die üblichen Anforderungen an den Einschluss von Studien zur Bewertung der diagnostischen Güte nicht vollständig erfüllen (zum Beispiel die prospektive Festlegung eines Schwellenwerts für Depression; der konsekutive Einschluss der teilnehmenden Personen sowie eine zeitnahe Untersuchung aller Personen mit Referenz- und Indextest in Unkenntnis des Ergebnisses des jeweils anderen Tests). Deshalb dienen die Angaben zu Sensitivität und Spezifität des jeweiligen Screeningtests lediglich der Orientierung.

4.3.1 Population

Die interessierende Zielpopulation bilden erwachsene Personen, bei denen aktuell keine depressive Störung diagnostiziert ist (siehe Abschnitt 4.2.1).

4.3.2 Indextest

Indextests in den einzuschließenden systematischen Übersichten sind Screeningtests, die sich in der hausärztlichen Praxis für den routinemäßigen Einsatz bei allen Patientinnen und Patienten eignen. Es werden bis zu 5 Screeningtests im Rahmen dieses Berichts untersucht, die in den Therapiestudien am häufigsten eingesetzt wurden.

4.3.3 Referenztest

Referenztest in den einzuschließenden systematischen Übersichten ist die Diagnose einer depressiven Störung über ein diagnostisches Interview zur Erfassung von psychischen Störungen gemäß DSM und ICD zeitnah zum Indextest.

4.3.4 Zielgrößen

Zielgröße in den einzuschließenden systematischen Übersichten sind Sensitivität und Spezifität im Hinblick auf die Diagnose einer depressiven Störung.

4.3.5 Studientypen

Für jeden Screeningtest werden systematische Übersichten, in denen die diagnostische Güte auf Basis von Testgütestudien im Querschnittsdesign ermittelt wurde, eingeschlossen. Die systematischen Übersichten müssen methodisch adäquat und der Fragestellung entsprechend angemessen durchgeführt worden sein. Mindestanforderungen sind eine systematische Recherche in bibliografischen Datenbanken sowie eine Qualitätsbewertung der

Primärliteratur. Je Screeningtest wird aus den eingeschlossenen systematischen Übersichten nur jeweils eine als Quelle für summarische Daten zur diagnostischen Güte herangezogen. Wie diese Auswahl erfolgt, wird in Abschnitt 4.6.2 beschrieben.

4.3.6 Tabellarische Darstellung der Kriterien für den Einschluss systematischer Übersichten

In der folgenden Tabelle sind die Kriterien aufgelistet, die systematische Übersichten erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 4: Übersicht über die Kriterien für den Einschluss systematischer Übersichten, die sich auf die diagnostische Güte beziehen

Einschlusskriterien	
D1	Als Zielpopulation wird in den Einschlusskriterien der SÜ definiert: erwachsene Personen, bei denen aktuell keine depressive Störung diagnostiziert ist (siehe auch Abschnitt 4.3.1)
D2	Als Indextest wird in den Einschlusskriterien der SÜ definiert: ein in Therapiestudien eingesetzter Screeningtest (siehe auch Abschnitt 4.3.2)
D3	Als Referenztest wird in den Einschlusskriterien der SÜ definiert: Diagnose einer depressiven Störung über ein standardisiertes diagnostisches Interview zur Erfassung von psychischen Störungen (siehe auch Abschnitt 4.3.3)
D4	Als Zielgröße wird in den Einschlusskriterien der SÜ definiert: diagnostische Güte im Hinblick auf die Diagnose einer Depression (siehe auch Abschnitt 4.3.4)
D5	Systematische Übersichten von Studien zur Testgüte, die als Mindestvoraussetzung eine systematische Recherche in bibliografischen Datenbanken sowie eine Qualitätsbewertung der Primärliteratur durchgeführt haben (siehe auch Abschnitt 4.3.5)
D6	Publikationszeitraum: letzte 5 Jahre
D7	Publikationssprache: Englisch oder Deutsch
a: Als Vollpublikation gilt in diesem Zusammenhang ein Bericht, der den Kriterien des PRISMA-Statements [11] genügt. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SÜ: systematische Übersicht	

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

Für das Einschlusskriterium „Population“ reicht es in den Studien zur Screeningkette und in den Therapiestudien aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen dieses Einschlusskriterium bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

Ebenfalls eingeschlossen werden Studien zur Screeningkette und Therapiestudien, die zu mindestens 80 % die Einschlusskriterien für die Prüfintervention beziehungsweise die Vergleichsintervention erfüllen.

4.5 Informationsbeschaffung

Studien und systematische Übersichten zur Screeningkette

Mit dem Ziel einer möglichst effizienten Ergebnisgewinnung und Nutzung bereits vorliegender wissenschaftlicher Evidenz auf höchster Evidenzstufe wird in einem 1. Schritt eine fokussierte Informationsbeschaffung nach systematischen Übersichten zur Screeningkette durchgeführt (siehe Abschnitt 4.5.2). Ziel ist es, hochwertige und aktuelle systematische Übersichten zu finden, aus denen Primärstudien zur Screeningkette identifiziert und anschließend bezüglich der spezifischen Einschlusskriterien des Berichtes bewertet werden. Bei diesem Vorgehen wird von der verwendeten systematischen Übersicht allein das Rechercheergebnis, nicht aber die Bewertung der eingeschlossenen Primärstudien oder die Datenextraktion übernommen. In einem 2. Schritt erfolgt dann eine Aktualisierung der Informationsbeschaffung für den Zeitraum, der nicht durch die systematischen Übersichten abgedeckt wird (siehe Abschnitt 4.5.1).

Studien zur Therapievorverlagerung

Falls Studien zur Screeningkette nicht oder nicht in ausreichender Quantität und Qualität vorliegen, werden Studien zur Therapievorverlagerung ohne Begrenzung des Zeitraumes umfassend recherchiert (siehe Abschnitt 4.5.1).

Systematische Übersichten zur diagnostischen Güte

Mit dem Ziel einer möglichst effizienten Ergebnisgewinnung und Nutzung bereits vorliegender wissenschaftlicher Evidenz wird die Bewertung der diagnostischen Güte relevanter Screeningtests auf Grundlage hochwertiger systematischer Übersichten durchgeführt. Hierfür wird eine fokussierte Informationsbeschaffung nach systematischen Übersichten zur diagnostischen Güte durchgeführt (siehe Abschnitt 4.5.2).

4.5.1 Umfassende Informationsbeschaffung

4.5.1.1 Bibliografische Recherche

Die systematische Recherche nach relevanten Studien beziehungsweise Dokumenten wird in folgenden bibliografischen Datenbanken durchgeführt:

- Suche nach Primärstudien in den Datenbanken MEDLINE, Embase, Cochrane Central Register of Controlled Trials, PsycINFO,
- Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE und Embase parallel zur Suche nach relevanter Primärliteratur sowie Suche in den Datenbanken Cochrane Database of Systematic Reviews und Health Technology Assessment Database.

4.5.1.2 Öffentlich zugängliche Studienregister

Die folgenden öffentlich zugänglichen Studienregister werden durchsucht:

- U.S. National Institutes of Health. ClinicalTrials.gov,
- World Health Organization. International Clinical Trials Registry Platform Search Portal,
- European Medicines Agency. EU Clinical Trials Register (für Therapievorverlagerung).

4.5.1.3 Weitere Suchquellen

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien beziehungsweise Informationen zu relevanten Studien zu ermitteln, werden weitere Quellen berücksichtigt.

4.5.1.3.1 Systematische Übersichten

Systematische Übersichten werden hinsichtlich weiterer relevanter Studien beziehungsweise Dokumente gesichtet.

4.5.1.3.2 Anhörung

Im Anschluss an die Veröffentlichungen des Berichtsplans und des Vorberichts erfolgt eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen kann. Relevante Informationen aus diesen Anhörungen werden im Rahmen der Nutzenbewertung berücksichtigt.

4.5.1.3.3 Autorenanfragen

Es werden Anfragen an Autorinnen und Autoren gestellt, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten lassen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen sind.

4.5.2 Fokussierte Informationsbeschaffung

Es erfolgt eine Suche in den Datenbanken MEDLINE, Cochrane Database of Systematic Reviews und Health Technology Assessment Database. Zudem werden die Websites von HTA-Agenturen wie NICE oder AHRQ nach systematischen Übersichten durchsucht.

4.5.3 Selektion relevanter Studien und systematischer Übersichten

Selektion relevanter Studien beziehungsweise Dokumente aus den Ergebnissen der bibliografischen Recherche

Die durch die Suche in bibliografischen Datenbanken identifizierten und zu screenenden Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 1 bis Tabelle 4) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Reviewerinnen oder Reviewer unabhängig voneinander. Diskrepanzen werden durch

Diskussion zwischen beiden aufgelöst. Konferenzabstracts werden im Rahmen der Nutzenbewertung nicht berücksichtigt.

Selektion relevanter Studien beziehungsweise Dokumente aus weiteren Suchquellen

Informationen aus den folgenden Suchquellen werden von 2 Reviewerinnen oder Reviewern unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

- öffentlich zugängliche Studienregister.

Informationen aus den folgenden Suchquellen werden von einer Reviewerin oder einem Reviewer auf Studien gesichtet, die diese dann in Bezug auf ihre Relevanz bewertet; eine 2. Reviewerin oder ein 2. Reviewer überprüft den gesamten Prozess inklusive der Bewertungen:

- identifizierte systematische Übersichten,
- im Rahmen der Anhörung zum Berichtsplan und zum Vorbericht eingereichte Informationen.

Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den Beiden aufgelöst. Konferenzabstracts werden im Rahmen der Nutzenbewertung nicht berücksichtigt.

Selektion relevanter systematischer Übersichten der fokussierten Informationsbeschaffung

Die durch die Suche in bibliografischen Datenbanken identifizierten Treffer sowie die Ergebnisse der Websites von HTA-Agenturen werden von einer Reviewerin oder einem Reviewer auf systematische Übersichten gesichtet und diese dann hinsichtlich ihrer Relevanz bewertet. Eine 2. Reviewerin oder ein 2. Reviewer überprüft den gesamten Prozess inklusive der Bewertungen.

4.6 Informationsbewertung

Die Bewertung der Informationen der eingeschlossenen Studien hängt stark von den verfügbaren Angaben und der Qualität der jeweiligen Publikationen und weiterer Informationsquellen ab. Alle für die Nutzenbewertung relevanten Ergebnisse werden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft. Auf Grundlage der Ergebnissicherheit wird für jedes Ergebnis endpunktspezifisch eine zugehörige Aussagesicherheit abgeleitet.

4.6.1 Bewertung der Screening- und Therapiestudien

Datenextraktion

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

Bewertung des Verzerrungspotenzials der Ergebnisse

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

A: Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (bei randomisierten Studien)
- Verdeckung der Gruppenzuteilung (bei randomisierten Studien)
- zeitliche Parallelität der Gruppen (bei nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen beziehungsweise Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung der Patientin oder des Patienten sowie der behandelnden Person (bei randomisierten Studien)
- ergebnisgesteuerte Berichterstattung

B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisgesteuerte Berichterstattung

Für randomisierte Studien wird anhand dieser Aspekte das Verzerrungspotenzial zusammenfassend als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts wird zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter (A) aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft. Falls diese Einstufung als „hoch“ erfolgt, wird das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet. Ansonsten finden die unter (B) genannten endpunktspezifischen Aspekte Berücksichtigung.

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

Für nicht randomisierte vergleichende Studien wird in der Regel keine zusammenfassende Bewertung der Verzerrungsaspekte durchgeführt, da die Ergebnisse dieser Studien aufgrund der fehlenden Randomisierung generell ein hohes Verzerrungspotenzial besitzen.

4.6.2 Bewertung von systematischen Übersichten

4.6.2.1 Bewertung von systematischen Übersichten zu Screeningstudien

Es werden systematische Übersichten herangezogen, die die Einschlusskriterien des Berichtes erfüllen und die eine hohe Qualität bei der Informationsbeschaffung gewährleisten.

4.6.2.2 Bewertung von systematischen Übersichten zur diagnostischen Güte

Die Entscheidung, welche systematische Übersicht als Quelle für die Darstellung der diagnostischen Güte des jeweiligen Screeningtests herangezogen wird, wird auf Basis der Bewertung der Qualität getroffen.

Damit eine systematische Übersicht verwendet werden kann, muss sie gewisse qualitative Mindestvoraussetzungen erfüllen; die Bewertung der Qualität erfolgt auf Basis des Oxman-Guyatt-Index [12], bei dem 9 Aspekte bewertet werden; dazu gehören unter anderem die Qualität der Informationsbeschaffung, die Studienselektion und die Evidenzsynthese. Allenfalls kleinere Mängel („minor flaws“) dürfen auftreten; ein Mindestscore von 5 Punkten des Oxman-Guyatt-Index muss gegeben sein [6]. Eine eigene Qualitätsbewertung der in die systematischen Übersichten eingeschlossenen Primärstudien erfolgt nicht.

Als Quelle für die im Bericht darzustellenden Daten zu Sensitivität und Spezifität dienen je Screeningtest die Ergebnisse der Übersicht, die die höchste Qualität aufweist. Bei mehreren gleichwertigen Übersichtsarbeiten wird geprüft, welche am geeignetsten ist Aussagen für die interessierende Population (siehe Abschnitt 4.3.1) zu machen.

4.7 Informationssynthese und -analyse für Screening- und Therapiestudien

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

4.7.1 Gegenüberstellung der Ergebnisse der Einzelstudien

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

In bestimmten Fällen werden einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt beziehungsweise nicht in die Nutzenbewertung einbezogen. Dies trifft insbesondere zu, wenn viele Patientinnen und Patienten nicht in der Auswertung enthalten sind. Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist. In der Literatur werden zum Teil bereits Auswertungen, in denen 20 % der Patientinnen und Patienten nicht berücksichtigt werden, als nicht mehr aussagekräftig betrachtet [13].

Ausnahmen von dieser Regel werden zum Beispiel dann gemacht, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [14].

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patientinnen und Patienten zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.7.2 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Für Metaanalysen werden im Fall von mindestens 5 Studien Modelle mit zufälligen Effekten nach der Methode von Knapp-Hartung unter Verwendung des Heterogenitätsschätzers nach Paule-Mandel [15] eingesetzt. Eine mögliche Heterogenität wird dabei anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität [16] geschätzt. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ($p \geq 0,05$ für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität wird nur das Prädiktionsintervall dargestellt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.7.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.7.5). Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Ansonsten erfolgt nur eine qualitative Zusammenfassung. In bestimmten Situationen und mit besonderer Begründung können Alternativen wie z. B. Bayes'sche Verfahren, generalisierte lineare Modelle eingesetzt werden. Bei binären Daten kommt insbesondere das Beta-Binomial-Modell infrage [17].

4.7.3 Aussagen zur Beleglage

Die Aussagen zur Beleglage, die entsprechend des unten stehenden Vorgehens getroffen werden, beziehen sich entweder auf den Nutzen des Screenings auf Depression (wenn Studien zur Screeningkette vorliegen) oder aber auf den Nutzen einer Therapievorverlagerung bei einer im Screening detektierten Depression. Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die Aussagesicherheit richtet sich nach der Anzahl verfügbarer Studien, der qualitativen und quantitativen Sicherheit ihrer Ergebnisse sowie der Homogenität der Ergebnisse bei mehreren Studien. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie zu beurteilen.

Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit. Ergebnisse nicht randomisierter vergleichender Studien haben eine geringe qualitative Ergebnissicherheit. Die regelhaft abzuleitende Aussagesicherheit ist Tabelle 5 zu entnehmen.

Tabelle 5: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			homogen	heterogen		
			Metaanalyse statistisch signifikant	gleichgerichtete Effekte ^a		
				deutlich	mäßig	nein
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–

a: Gleichgerichtete Effekte liegen vor, wenn trotz Heterogenität eine deutliche oder mäßige Richtung der Effekte erkennbar ist.

4.7.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sind Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel der Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder der Wahl des Effektmaßes. Derartige Sensitivitätsanalysen erfolgen unabhängig von gegebenenfalls weiteren Analysen, mit denen die Ergebnissicherheit eines beobachteten Effekts bewertet wird.

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

4.7.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.7.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf

einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Bei binären Daten werden Subgruppenanalysen nur durchgeführt, wenn mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Alter (älter als 65 Jahre),
- Geschlecht,
- Schweregrad der Depression,
- Screeningtest.

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

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen und Patienten eingeschränkt werden.

4.8 Informationssynthese aus systematischen Übersichten zur diagnostischen Güte

Für jeden der untersuchten Screeningtests werden die Charakteristika der systematischen Übersicht, die als Quelle für die Daten zur diagnostischen Güte genutzt wurde, beschrieben. Auch das Verzerrungspotenzial der eingeschlossenen Studien wird berücksichtigt. Weiterhin wird dargestellt, inwieweit sich die Population aus den in der systematischen Übersicht eingeschlossenen Studien eignet, Aussagen zur diagnostischen Güte des jeweiligen Screeningtests auf die hier interessierende Population zu übertragen.

4.8.1 Darstellung der Ergebnisse

Für jeden relevanten Screeningtest werden Sensitivität und Spezifität aus der jeweils zugrunde liegenden systematischen Übersicht dargestellt.

4.8.2 Aussagen zum Nutzen des Screenings auf Basis gemeinsamer Betrachtung von Studien zur Therapievorverlagerung und der diagnostischen Güte von Screeningtests

Allein auf Basis der diagnostischen Güte wird keine Nutzensaussage abgeleitet. Der Nutzen des Screenings kann durch die Gegenüberstellung der gesundheitsbezogenen Konsequenzen, die sich aus falsch-positiven und falsch-negativen Befunden ergeben können zusammen mit

einer Aussage zum Nutzen einer Therapievorverlagerung bei Patientinnen und Patienten mit einer im Screening detektierten Depression abgeleitet werden. Die Aussagesicherheit bezüglich des Nutzens des Screenings berücksichtigt dann sowohl die Aussagesicherheit bezüglich der Therapievorverlagerung als auch das Verzerrungspotenzial und die konkrete Passung der jeweils zugrunde liegenden systematischen Übersicht bezüglich der diagnostischen Güte.

5 Literatur

1. WHO-Regionalbüro für Europa. Definition einer Depression [online]. [Zugriff: 16.05.2017]. URL: <http://www.euro.who.int/de/health-topics/noncommunicable-diseases/pages/news/news/2012/10/depression-in-europe/depression-definition>.
2. Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde, Bundesärztekammer, Kassenärztliche Bundesvereinigung, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. S3-Leitlinie/Nationale VersorgungsLeitlinie: unipolare Depression; Langfassung; Version 5 [online]. 11.2015 [Zugriff: 13.06.2017]. URL: <http://www.leitlinien.de/mdb/downloads/nvl/depression/depression-2aufl-vers5-lang.pdf>.
3. Wittchen HU, Jacobi F, Klose M, Ryl L. Depressive Erkrankungen. Berlin: RKI; 2010. (Gesundheitsberichterstattung des Bundes; Band 51). URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsT/depression.pdf?__blob=publicationFile.
4. Busch MA, Maske UE, Ryl L, Schlack R, U. H. Prävalenz von depressiver Symptomatik und diagnostizierter Depression bei Erwachsenen in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013; 5(6): 733–739.
5. Jacobi F, Höfler M, Strehle S, Mack S, Gerschler A, Scholl L et al. Psychische Störungen in der Allgemeinbevölkerung: Studie zur Gesundheit Erwachsener in Deutschland und ihr Zusatzmodul Psychische Gesundheit (DEGS1-MH). Nervenarzt 2014; 85(1): 77–87.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 5.0. Köln: IQWiG; 2017. URL: <https://www.iqwig.de/download/Allgemeine-Methoden-Version-5-0.pdf>.
7. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869.
8. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004; 94(3): 361-366.
9. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009; 339: b2535.
10. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3; current step 4 version [online]. 30.11.1995 [Zugriff: 17.02.2017]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.

11. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009; 62(10): 1006-1012.
12. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991; 44(11): 1271-1278.
13. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
14. Lange S. The all randomized/full analysis set (ICH E9): may patients be excluded from the analysis? *Drug Inf J* 2001; 35(3): 881-891.
15. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O 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; (Suppl 1): 25-27.
16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.
17. Kuss O. Statistical methods for meta-analyses including information from studies without any events: add nothing to nothing and succeed nevertheless. *Stat Med* 2015; 34(7): 1097-1116.

Anhang A – Offenlegung potenzieller Interessenkonflikte von Betroffenen

Im Folgenden sind die potenziellen Interessenkonflikte der konsultierten Betroffenen zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“.

Potenzielle Interessenkonflikte der konsultierten Betroffenen

Name; Institution	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7
Susanne Fiege; Deutsche DepressionsLiga e. V.	ja	nein	ja	nein	ja	nein	nein
Francoise Margue; Deutsche DepressionsLiga e. V	ja	nein	nein	nein	ja	nein	nein

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ (Version 11/2016) wurden folgende 7 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, Hersteller von Medizinprodukten oder einem industriellen Interessenverband angestellt, für diese selbständig oder ehrenamtlich tätig bzw. sind oder waren Sie freiberuflich in eigener Praxis tätig? (Zu den oben genannten Einrichtungen zählen beispielsweise auch Kliniken, Einrichtungen der Selbstverwaltung, Fachgesellschaften, Auftragsinstitute)

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten (z. B. als Gutachter, Sachverständiger, Mitglied eines Advisory Boards, Mitglied eines Data Safety Monitoring Boards (DSMB) oder Steering Committees)?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband Honorare erhalten (z. B. für Vorträge, Schulungstätigkeiten, Stellungnahmen oder Artikel)?

Frage 4: Haben Sie oder haben die von Ihnen unter Frage 1 genannten Einrichtungen innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung z. B. für Forschungsaktivitäten, die Durchführung klinischer Studien, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten? (Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.)

Frage 5: Haben Sie oder haben die von Ihnen unter Frage 1 genannten Einrichtungen innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren für Fortbildungen / Kongresse) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband? (Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.)

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist? Besitzen Sie Patente für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt oder eine medizinische Methode oder Gebrauchsmuster für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt?

Frage 7: Sind oder waren Sie jemals an der Erstellung einer Leitlinie oder Studie beteiligt, die eine mit diesem Projekt vergleichbare Thematik behandelt/e? Gibt es sonstige Umstände, die aus Sicht eines unvoreingenommenen Betrachters als Interessenkonflikt bewertet werden können (z. B. Aktivitäten in gesundheitsbezogenen Interessengruppierungen bzw. Selbsthilfegruppen, politische, akademische, wissenschaftliche oder persönliche Interessen)?