IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Cholinesterasehemmer bei Alzheimer Demenz – Ergänzungsauftrag: Rivastigmin-Pflaster und Galantamin

Berichtsplan

Auftrag A09-05
Version 1.0
Stand: 09.08.2010

Impressum

Herausgeber:
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:
Cholinesterasehemmer bei Alzheimer Demenz – Ergänzungsauftrag: Rivastigmin-Pflaster und Galantamin

Auftraggeber:
Gemeinsamer Bundesausschuss

Datum des Auftrags:
18.12.2009

Interne Auftragsnummer:
A09-05

Anschrift des Herausgebers:
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Dillenburger Str. 27
51105 Köln

Tel.: +49 221 35685-0
Fax: +49 221 35685-1
Berichte@iqwig.de
www.iqwig.de

Schlagwörter: Cholinesteraseinhibitoren, Galantamin, Rivastigmin, Alzheimer-Krankheit, Demenz, Nutzenbewertung, systematische Übersicht

Inhaltsverzeichnis

Seite

Tabellenverzeichnis ..... **iv**
Abkürzungsverzeichnis ..... **v**
1 Hintergrund **1**
1.1 Ergänzungsauftrag zum Thema Cholinesterasehemmer bei Alzheimer Demenz **1**
1.2 Ergebnisse des Abschlussberichts A05-19A **1**
1.3 Neue Applikationsform Rivastigmin transdermales Pflaster **2**
1.4 Therapieevaluation bei der Alzheimer Demenz **2**
2 Ziele der Untersuchung **4**
3 Projektbearbeitung **5**
3.1 Zeitlicher Verlauf des Projekts **5**
3.2 Dokumentation der Änderungen des Berichtsplans **5**
4 Methoden **6**
4.1 Kriterien für den Einschluss von Studien in die Untersuchung **6**
4.1.1 Population 6
4.1.2 Prüf- und Vergleichsintervention 6
4.1.3 Patientenrelevante Endpunkte 6
4.1.4 Studientypen 7
4.1.5 Studiendauer 8
4.1.6 Tabellarische Übersicht der Kriterien für den Studieneinschluss 8
4.1.7 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf 8
4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen 8
4.2 Informationsbeschaffung **9**
4.2.1 Selektion relevanter Studien 9
4.3 Informationsbewertung **9**
4.4 Informationssynthese und -analyse **11**
4.4.1 Gegenüberstellung der Ergebnisse der Einzelstudien 12
4.4.2 Meta-Analysen 12
4.4.3 Sensitivitätsanalyse 13
4.4.4 Subgruppenmerkmale und andere Effektmodifikatoren 13
5 Literaturverzeichnis **15**

Tabellenverzeichnis

Tabelle 1: Übersicht der Kriterien für den Studieneinschluss..... 8

Abkürzungsverzeichnis

Abkürzung	**Bedeutung**
ADAS-cog	Alzheimer's Disease Assessment Scale – cognitive subscale
ADL	Activities of Daily Living
CIBIC-plus	Clinician Interview-Based Impression of Change plus caregiver input
CONSORT	Consolidated Standards of Reporting Trials
DSM	Diagnostic and Statistical Manual of Mental Disorders
EMA	European Medicines Agency
G-BA	Gemeinsamer Bundesausschuss
ICD	International Classification of Diseases
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
MID	Minimal Important Difference
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association
RCT	randomised controlled trial (randomisierte kontrollierte Studie)
SMD	standardisierte Mittelwertdifferenz
TTS	transdermales therapeutisches System

1 Hintergrund

1.1 Ergänzungsauftrag zum Thema Cholinesterasehemmer bei Alzheimer Demenz

Im Februar 2007 hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) einen Abschlussbericht zum Thema Cholinesterasehemmer bei Alzheimer Demenz erstellt (Auftrag A05-19A) [1]. Dieser Abschlussbericht stellte den ersten Bericht aus der Themenreihe Alzheimer Demenz dar. Bei Fertigstellung der vierten und letzten Bewertung (Auftrag A05-19C: „Memantin bei Alzheimer Demenz“ [2]) lag die Literaturrecherche des Abschlussberichts A05-19A bereits etwa 3 Jahre zurück. Der Gemeinsame Bundesausschuss (G-BA) hatte das IQWiG daher mit einer Aktualisierungsrecherche in Form eines Rapid Reports (Auftrag A09-03) beauftragt, festzustellen, ob inzwischen substanziell neue Evidenz zu den Cholinesterasehemmern vorliegt. Diese „Aktualisierungsrecherche zum Bericht A05-19A (Cholinesterasehemmer bei Alzheimer Demenz)“ [3] stellte Aktualisierungsbedarf für 2 Fragestellungen fest.

Für den Vergleich Galantamin vs. Placebo bzw. keine Behandlung erschien es aufgrund einer relevanten Menge zusätzlicher, z. T. unpublizierter Daten möglich, dass eine erneute Bewertung dieser Fragestellung die hierzu im Abschlussbericht A05-19A getroffenen Aussagen ändert. Zudem stellte die Aktualisierungsrecherche fest, dass nach Fertigstellung des Abschlussberichts A05-19A eine neue Applikationsform von Rivastigmin (Pflaster zur transdermalen Applikation) zugelassen wurde. Eine Bearbeitung dieser neuen Fragestellung würde daher zu zusätzlichen Aussagen im Vergleich zu den im Abschlussbericht A05-19A getroffenen Aussagen führen.

Als Konsequenz aus den Ergebnissen des Rapid Reports A09-03 erteilte der G-BA dem IQWiG diesen Ergänzungsauftrag.

Zur Epidemiologie, Pathogenese, Diagnose und zu den Behandlungsmöglichkeiten der Alzheimer Demenz wird auf den Abschlussbericht A05-19A verwiesen.

1.2 Ergebnisse des Abschlussberichts A05-19A

Für die Cholinesterasehemmer Donepezil, Galantamin und Rivastigmin wurde ein Beleg für einen Nutzen im Bereich der kognitiven Leistungsfähigkeit bei Patienten mit Alzheimer Demenz gefunden. Dies galt für Donepezil über alle eingesetzten Dosierungen hinweg, für Galantamin und Rivastigmin nur bei mittleren und hohen Dosierungen. Für andere Dimensionen, wie z. B. Aktivitäten des täglichen Lebens (activities of daily living, ADL), begleitende psychopathologische Symptome, gesundheitsbezogene Lebensqualität und die Notwendigkeit einer vollstationären Pflege (Institutionalisierung) lag kein Beleg für einen Nutzen vor.

Für alle Wirkstoffe lag ein Beleg für einen dosisabhängigen Schaden vor, insbesondere im Bereich gastrointestinaler unerwünschter Ereignisse (wie z. B. Übelkeit, Erbrechen, Durchfall).

Weiterhin gab es keinen Beleg für einen Zusatznutzen für eine der 3 untersuchten Substanzen der jeweils anderen gegenüber [1].

1.3 Neue Applikationsform Rivastigmin transdermales Pflaster

Ein Rivastigmin-Pflaster zur transdermalen Applikation (transdermales therapeutisches System, TTS) ist seit 2007 in Deutschland in 2 verschiedenen Dosierungen zugelassen [4]: ein Matrix-Pflaster mit 9 mg Rivastigmin (5 cm^2), das pro 24 Stunden 4,6 mg Rivastigmin freisetzt und ein Matrix-Pflaster mit 18 mg Rivastigmin (10 cm^2), das pro 24 Stunden 9,5 mg Rivastigmin freisetzt. Es ist zugelassen zur symptomatischen Behandlung der leichten bis mittelschweren Alzheimer Demenz. Die Behandlung sollte mit 4,6 mg/24 Stunden begonnen werden und bei guter Verträglichkeit nach mindestens 4-wöchiger Behandlung auf 9,5 mg/24 Stunden erhöht werden. Die Vorgehensweise der Umstellung von Hartkapseln oder Lösung zum Einnehmen auf transdermale Pflaster ist in der Fachinformation beschrieben. Das Pflaster ist im Gegensatz zu den oralen Rivastigmin-Darreichungsformen, die zweimal täglich zu applizieren sind, nur einmal täglich anzuwenden. Es sollte auf saubere, trockene, unbehaarte, intakte, gesunde Haut im oberen oder unteren Rückenbereich, auf dem Oberarm oder Brustkorb aufgeklebt werden. Die in der Fachinformation am häufigsten berichteten unerwünschten Ereignisse sind, wie auch bei den oralen Darreichungsformen, gastrointestinaler Art [5].

1.4 Therapieevaluation bei der Alzheimer Demenz

Zur Therapieevaluation bei der Alzheimer Demenz stehen unterschiedliche Instrumente zur Verfügung. Auf diese Instrumente als auch auf die Besonderheiten der Erfassung von krankheitsbedingten Beeinträchtigungen bei Alzheimer-Patienten wurde bereits im Abschlussbericht A05-19A eingegangen [1].

Eine neuere systematische Übersicht von Molnar 2009 beschäftigt sich mit der Bewertung von Effekten von Antidementiva in randomisierten kontrollierten Studien (RCTs) [6]. Die klinische Relevanz der Effekte wurde in 46 % der untersuchten RCTs diskutiert. Allerdings erfolgten die meisten Bewertungen auf Grundlage von Expertenmeinungen, wie z. B. die als klinisch relevant eingestufte Veränderung um 4 Punkte auf der Alzheimer's Disease Assessment Scale – cognitive subscale (ADAS-cog) [7]. Häufig wurden auch globale Erhebungsskalen (z. B. Clinician's Interview-Based Impression of Change plus caregiver input, CIBIC-plus) zur Bewertung der klinischen Relevanz genutzt. Jedoch fehlt auch hier ein standardisiertes Vorgehen bei deren Verwendung für diesen Zweck. Empirische Methoden wie z. B. die Minimal Important Difference (MID) oder statistische Ansätze (z. B. Cohen's d) kamen in den untersuchten Studien gar nicht zum Einsatz. Diese Ergebnisse zeigen, dass auf

dem Gebiet der Relevanzbewertung klinischer Effekte im Bereich Alzheimer Demenz großer Forschungsbedarf besteht. Insbesondere sind Daten zu MIDs von zentraler Bedeutung, da sie die Einschätzung von Patienten und Betreuenden zu relevanten Endpunktveränderungen erfassen. Für die meisten Instrumente, die zur Therapieevaluation bei der Alzheimer Demenz eingesetzt werden, existieren bisher keine Angaben zu MIDs weder auf individueller Ebene noch auf Gruppenebene [6]. Bei der Nutzenbewertung im Bereich Demenz ist daher zu erwarten, dass die Beurteilung der Relevanz der beobachteten Effekte mithilfe statistischer Verfahren vorgenommen werden muss. Das entsprechende Vorgehen ist in Abschnitt 4.4 des vorliegenden Berichtsplans beschrieben.

2 Ziele der Untersuchung

Ziele der vorliegenden Untersuchung sind:

- die Aktualisierung der Bewertung des Nutzens von Galantamin

sowie

- die Bewertung des Nutzens und des Zusatznutzens von Rivastigmin als transdermales Pflaster

jeweils bei Patienten mit Alzheimer Demenz hinsichtlich patientenrelevanter Endpunkte.

3 Projektbearbeitung

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 18.12.2009 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung „Cholinesterasehemmer bei Alzheimer Demenz – Ergänzungsauftrag: Rivastigmin-Pflaster und Galantamin“ beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Im Rahmen des Auftrags A05-19A „Cholinesterasehemmer bei Alzheimer Demenz“ fand während der Erstellung des Berichtsplans ein Treffen mit Vertretern der Deutschen Alzheimer-Gesellschaft zur Festlegung patientenrelevanter Zielgrößen statt [1].

Der vorläufige Berichtsplan in der Version 1.0 vom 08.04.2010 wurde am 19.04.2010 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 18.05.2010 konnten schriftliche Stellungnahmen eingereicht werden. Unklare Aspekte aus den schriftlichen Stellungnahmen zum vorläufigen Berichtsplan wurden am 24.06.2010 in einer wissenschaftlichen Erörterung mit den Stellungnehmenden diskutiert. Die Dokumentation und Würdigung der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht.

Der vorliegende Berichtsplan beinhaltet die Änderungen, die sich aus der Anhörung ergeben haben.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Der Vorbericht wird 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 8 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation und Würdigung der Anhörung zum Vorbericht veröffentlicht.

3.2 Dokumentation der Änderungen des Berichtsplans

Im Vergleich zum vorläufigen Berichtsplan Version 1.0 haben sich folgende Änderungen bzw. Ergänzungen ergeben:

- In Abschnitt 4.4 wurde die Beschreibung des Vorgehens bei der Bewertung der Relevanz von Effekten, die mithilfe von (komplexen) Skalen erhoben wurden, sprachlich überarbeitet, um mögliche Missverständnisse zu vermeiden.

- In Abschnitt 4.4.1 wurde die Beschreibung des Vorgehens im Fall, dass Ergebnisse auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basieren, sprachlich überarbeitet, um mögliche Missverständnisse zu vermeiden.

4 Methoden

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

4.1.1 Population

Für die Nutzenbewertung werden Studien mit Patienten mit einer leichten oder mittelschweren Alzheimer Demenz, auch als Mischform mit zum Beispiel vaskulärer Demenz, berücksichtigt. Die Diagnosesicherung sollte dabei anhand allgemein akzeptierter Kriterien wie zum Beispiel derjenigen der ICD-10[1], des DSM-IV-TR[2] oder der NINCDS-ADRDA[3], wie in der einschlägigen Publikation der European Medicines Agency (EMA) beschrieben [8], erfolgt sein.

Studien, bei denen ausschließlich Patienten mit einer vaskulären Demenz, einer Demenz aufgrund eines Morbus Parkinson, einer Demenz aufgrund einer Lewy-Körper-Demenz, einer Demenz aufgrund einer Creutzfeld-Jakob-Erkrankung oder einer anderen seltenen Demenzform eingeschlossen wurden, werden nicht berücksichtigt.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfenden Interventionen sind die beiden Cholinesterasehemmer Galantamin als orale Applikationsform und Rivastigmin als Pflaster zur transdermalen Applikation.

Als Vergleichsinterventionen werden im Fall von Galantamin eine Placebobehandlung und keine Behandlung betrachtet (Bewertung des Nutzens). Im Fall von Rivastigmin werden eine Placebobehandlung bzw. keine Behandlung und jegliche andere in Deutschland zugelassene und verfügbare medikamentöse und nichtmedikamentöse Therapieoption bei der Alzheimer Demenz als Vergleichsinterventionen betrachtet (Bewertung des Nutzens und Zusatznutzens).

Die Anwendung der in den Studien eingesetzten Prüf- und Vergleichsinterventionen muss im Rahmen des für Deutschland gültigen Zulassungsstatus erfolgen.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte verwendet:

- Aktivitäten des täglichen Lebens (alltagspraktische Fähigkeiten),
- kognitive Leistungsfähigkeit,

[1] International Classification of Diseases
[2] Diagnostic and Statistical Manual of Mental Disorders
[3] National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association

- gesundheitsbezogene Lebensqualität,
- andere mit der Erkrankung verbundene Symptome, zum Beispiel Depression, Schlaf-Wach-Umkehr, Wahnhaftigkeit, Agitiertheit (begleitende psychopathologische Symptome),
- Notwendigkeit einer vollstationären Pflege (Institutionalisierung),
- Mortalität,
- unerwünschte Arzneimittelwirkungen.

Zusätzlich werden auch Zielgrößen verwendet, die eine Beurteilung angehörigenrelevanter Therapieziele ermöglichen:

- Lebensqualität der betreuenden Angehörigen und
- Betreuungsaufwand durch eine oder mehrere betreuende Person(en) oder Institution(en).

Ergänzend werden der interventions- und erkrankungsbedingte Aufwand, die Zufriedenheit der Patienten mit der Behandlung sowie Ergebnisse, die sich auf „das klinische Krankheitsstadium gemäß dem klinischen Eindruck“ (klinisches Globalurteil) beziehen, berücksichtigt und die entsprechenden zugehörigen Effekte dargestellt. Ein Nutzen oder Zusatznutzen kann sich allein auf Basis dieser Zielgrößen jedoch nicht ergeben. Dies gilt auch für die angehörigenrelevanten Therapieziele.

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

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 Abschnitt 4.1.2 genannten Interventionen und alle unter Abschnitt 4.1.3 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.

4.1.5 Studiendauer

Analog zu den Berichten A05-19A und A09-03 werden Studien mit einer Mindestdauer von 16 Wochen eingeschlossen.

4.1.6 Tabellarische Übersicht der Kriterien für den Studieneinschluss

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

Tabelle 1: Übersicht der Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Patienten mit einer leichten bis mittelschweren Alzheimer Demenz, auch als Mischform mit z. B. vaskulärer Demenz (siehe Abschnitt 4.1.1)
E2	Prüfintervention: Behandlung mit Galantamin oder Rivastigmin-Pflaster zur transdermalen Applikation (siehe Abschnitt 4.1.2)
E3	Vergleichsintervention: Behandlung mit Placebo oder keiner Behandlung im Falle des Galantamins und Behandlung mit Placebo oder keiner Behandlung oder einer anderen medikamentösen oder nichtmedikamentösen Therapieoption im Falle des Rivastigmin-Pflasters (siehe Abschnitt 4.1.2)
E4	Patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	Randomisierte kontrollierte Studien (RCTs)
E6	Studien mit einer Beobachtungsdauer von mindestens 16 Wochen
E7	Vollpublikation verfügbar[a]
a: Als Vollpublikation gilt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT-Statements [9] genügt und eine Bewertung der Studie ermöglicht.	

4.1.7 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf

Sofern sich im Projektverlauf Änderungen im Zulassungsstatus der zu bewertenden Interventionen ergeben, werden die Kriterien für den Studieneinschluss gegebenenfalls an die neuen Zulassungsbedingungen angepasst. Die jeweils vorgenommenen Änderungen werden im Vorbericht bzw. im Abschlussbericht explizit vermerkt.

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

Für das Einschlusskriterium E1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

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

4.2 Informationsbeschaffung

Sämtliche für den Abschlussbericht A05-19A sowie den Rapid Report A09-03 relevanten identifizierten Studien, Publikationen, Studienberichte und Studienregisterinformationen zu Galantamin sowie Rivastigmin-Pflaster, die die unter 4.1.6 genannten Einschlusskriterien erfüllen, werden für den vorliegenden Auftrag herangezogen und in die Bewertung einbezogen.

Da Rivastigmin-Pflaster zum Zeitpunkt der Erstellung des Berichts A05-19A noch nicht in Deutschland zugelassen war, wurde für den im Bericht A05-19A abgedeckten Recherchezeitraum eine Vorabrecherche nach relevanten Studien zu Rivastigmin-Pflaster in den Datenbanken MEDLINE und EMBASE durchgeführt. Diese Vorabrecherche ergab keine Publikationen zu relevanten Studien für diesen Zeitraum (letztes Recherchedatum A05-19A: 12.06.2006).

Im Projektverlauf wird eine Aktualisierungsrecherche für den Zeitraum ab September 2009 analog zur Recherche des Rapid Reports A09-03 in bibliografischen Datenbanken, öffentlich zugänglichen Studienregistern und öffentlich zugänglichen Dokumenten von Zulassungsbehörden durchgeführt. Dabei wird die Recherche auf Galantamin und Rivastigmin-Pflaster eingeschränkt.

Zudem sollen Unterlagen von Herstellerfirmen, Informationen von Autoren einzelner Publikationen (z. B. zur Frage nach nicht publizierten Teilaspekten) sowie im Rahmen der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht eingereichte Informationen zur Identifizierung publizierter und nicht publizierter Studien herangezogen werden.

4.2.1 Selektion relevanter Studien

Die Selektion relevanter Studien erfolgt durch 2 Reviewer unabhängig voneinander. Dazu wird das Ergebnis der Recherche in den oben genannten Quellen herangezogen.

4.3 Informationsbewertung

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

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
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des Behandlers
- Ergebnisgesteuerte Berichterstattung

B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene

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

Das Verzerrungspotenzial wird 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 i. d. R. 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.

4.4 Informationssynthese und -analyse

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

Bei der Betrachtung patientenrelevanter Endpunkte, die mithilfe von (komplexen) Skalen operationalisiert werden, ist es in besonderer Weise notwendig, neben der statistischen Signifikanz der Effekte die Relevanz der beobachteten Wirkungen der untersuchten Interventionen zu bewerten, da die Komplexität der Skalen häufig eine sinnvolle Interpretation geringfügiger Unterschiede erschwert.

Die Relevanzbewertung kann auf Basis von Mittelwertdifferenzen und Responderanalysen vorgenommen werden. Welches dieser Verfahren gewählt werden kann, hängt insbesondere von der Verfügbarkeit der Daten aus den Primärstudien ab.

Um skalen- und indikationsspezifischen Besonderheiten gerecht zu werden, soll die Relevanzbewertung primär auf Basis validierter (bzw. etablierter) Relevanzkriterien für die jeweilige Skala durchgeführt werden. Für die Betrachtung von Mittelwertdifferenzen ist dies z. B. eine validierte bzw. etablierte Relevanzschwelle für den Gruppenunterschied: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Relevanzschwelle, wird von einer relevanten Effektstärke ausgegangen. Bei der Betrachtung von Responderanalysen ist es erforderlich, dass bei diesen Analysen ein validiertes bzw. etabliertes Responsekriterium angewendet wurde (im Sinne einer individuellen Minimal Important Difference [MID]). Liegt bei einer solchen Auswertung ein statistisch signifikanter Unterschied der Anteile der Responder zwischen den Gruppen vor, wird dies als relevanter Effekt angesehen, da die Responsedefinition bereits eine Schwelle der Relevanz (nämlich die MID) beinhaltet.

Falls skalenspezifisch validierte bzw. etablierte Relevanzkriterien nicht vorliegen, muss auf ein allgemeines statistisches Maß zur Relevanzbewertung zurückgegriffen werden. In diesem Fall werden standardisierte Mittelwertdifferenzen (SMD in Form von Hedges' g) betrachtet. Als Relevanzschwelle wird dann 0,2 SMD verwendet: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Relevanzschwelle, wird von einer relevanten Effektstärke ausgegangen. Dies soll gewährleisten, dass der beobachtete Effekt hinreichend sicher mindestens als „klein" angesehen werden kann [10,11].

In manchen Fällen liegt zwar eine validierte MID bzw. ein etabliertes Responsekriterium vor, jedoch werden keine Responderanalysen (im obigen Sinne) berichtet. Dann kann ggf. eine andere Relevanzschwelle als 0,2 SMD in Abhängigkeit von dieser MID bzw. diesem Responsekriterium gewählt werden.

4.4.1 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 bzw. nicht in die Nutzenbewertung einbezogen. Dies trifft insbesondere zu, wenn viele Patienten nicht in der Auswertung enthalten sind. Ergebnisse fließen i. d. R. nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basieren, das heißt wenn der Anteil der Patienten ohne jegliche Berücksichtigung in der Auswertung (Nichtberücksichtigungsanteil) größer als 30 % ist. In der Literatur werden zum Teil bereits Nichtberücksichtigungsanteile größer als 20 Prozentpunkte als nicht mehr aussagekräftig betrachtet [12].

Ausnahmen von dieser Regel können zum Beispiel dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [13]. Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Nichtberücksichtigungsanteile zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.4.2 Meta-Analysen

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar sind, werden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung werden primär die Ergebnisse aus Intention-to-Treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Meta-Analysen erfolgen i. d. R. auf Basis von Modellen mit zufälligen Effekten [14]. In begründeten Ausnahmefällen werden Modelle mit festen Effekten eingesetzt. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, werden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet bzw. näherungsweise bestimmt.

Für kontinuierliche Variablen wird die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt. Bei binären Variablen werden Meta-Analysen primär anhand des Odds Ratios durchgeführt. In begründeten Ausnahmefällen kommen auch andere Effektmaße zum Einsatz. Bei kategorialen Variablen wird ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet [15].

Die Effektschätzer und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgt die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität [16]. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ($p > 0,2$ für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall

dargestellt. Bei bedeutsamer Heterogenität werden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.4.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.4).

4.4.3 Sensitivitätsanalyse

Zur Einschätzung der Robustheit der Ergebnisse sind Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, z. B. die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ oder „niedrig“ wird für Sensitivitätsanalysen verwendet.

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

4.4.4 Subgruppenmerkmale und andere Effektmodifikatoren

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

- Geschlecht
- Alter
- Schweregrad der Demenz
- Dosierung

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

Bei Identifizierung möglicher Effektmodifikatoren erfolgt ggf. eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines Zusatznutzens auf eine spezielle Subgruppe von Patienten eingeschränkt werden.

5 Literaturverzeichnis

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Cholinesterasehemmer bei Alzheimer Demenz: Abschlussbericht; Auftrag A05-19A [online]. 07.02.2007 [Zugriff: 03.03.2010]. (IQWiG-Berichte; Band 17). URL: http://www.iqwig.de/download/A05-19A_Abschlussbericht_Cholinesterasehemmer_bei_Alzheimer_Demenz.pdf.

2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Memantin bei Alzheimer Demenz: Abschlussbericht; Auftrag A05-19C [online]. 10.09.2009 [Zugriff: 03.03.2010]. (IQWiG-Berichte; Band 59). URL: http://www.iqwig.de/download/A05-19C_Abschlussbericht_Memantin_bei_Alzheimer_Demenz.pdf.

3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Aktualisierungsrecherche zum Bericht A05-19A (Cholinesterasehemmer bei Alzheimer Demenz): Rapid Report; Auftrag A09-03 [online]. 12.10.2009 [Zugriff: 04.03.2010]. (IQWiG-Berichte; Band 67). URL: http://www.iqwig.de/download/A09-03_Rapid_Report_Aktualisierungsrecherche_zu_A05-19A_Cholinesterasehemmer_bei_Alzheimer_Demenz.pdf.

4. European Medicines Agency. Exelon patch: European public assessment report (EPAR); scientific discussion [online]. 2007 [Zugriff: 23.03.2010]. URL: http://www.emea.europa.eu/humandocs/PDFs/EPAR/Exelon/Exelon-H-169-X-38-SD.pdf.

5. Novartis Pharma. Exelon transdermales Pflaster: Fachinformation [online]. 12.05.2008 [Zugriff: 23.03.2010]. URL: http://www.fachinfo.de.

6. Molnar FJ, Man-Son-Hing M, Fergusson D. Systematic review of measures of clinical significance employed in randomized controlled trials of drugs for dementia. J Am Geriatr Soc 2009; 57(3): 536-546.

7. Kreis J, Kaiser T, Wieseler B. If shrimps had lips, they could whistle. J Am Geriatr Soc 2009; 57(10): 1947.

8. European Medicines Agency. Guideline on medicinal products for the treatment of Alzheimer's disease and other dementias [online]. 24.07.2008 [Zugriff: 09.03.2010]. URL: http://www.emea.europa.eu/pdfs/human/ewp/055395en.pdf.

9. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869.

10. Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. Chichester: Wiley; 2007.

11. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Erlbaum; 1988.

12. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet 2002; 359(9308): 781-785.

13. 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.

14. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7(3): 177-188.

15. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.

16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327(7414): 557-560.