

IQWiG-Berichte · Jahr: 2007 Nr. 17

Cholinesterasehemmer bei Alzheimer Demenz

Abschlussbericht

Auftrag A05-19A
Version 1.0
Stand: 07.02.2007

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Thema:

Nutzenbewertung von Cholinesterasehemmern bei Patienten mit Alzheimer Demenz, auch im Vergleich mit anderen Therapieoptionen.

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

22.02.2005

Interne Auftragsnummer:

A05-19A als Teil des Auftrages A05-19 „Nutzenbewertung von Cholinesterasehemmern, ginkgohaltigen Präparaten und Memantin bei Patienten mit Demenz, auch im Vergleich untereinander“

Anschrift des Herausgebers:

IQWiG
Dillenburger Straße 27
51105 Köln

Telefon (02 21) 3 56 85-0
Telefax (02 21) 3 56 85-1
berichte@iqwig.de
www.iqwig.de

ISSN

1864-2500

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

Externe Sachverständige:

1. An der Berichterstellung unmittelbar beteiligt:

- Prof. Dr. med. Michael Hüll, Universitätsklinikum Freiburg, Freiburg
- Prof. Dr. med. Klaus Schmidtke, Universitätsklinikum Freiburg, Freiburg
- Dr. med. Dr. P.H. Dipl.-Volksw. Stefan Weinmann, Universität Ulm, Günzburg

2. Externes Review des Vorberichts:

- Prof. Dr. med. Peter Berlit, Alfried-Krupp-Krankenhaus, Essen

Im folgenden Text wurde bei der Angabe von Personenbezeichnungen jeweils die männliche Form angewandt. Dies erfolgte ausschließlich zur Verbesserung der Lesbarkeit.

Zu allen Dokumenten, auf die via Internet zugegriffen wurde und die entsprechend zitiert sind, ist das jeweilige Zugriffsdatum angegeben. Sofern diese Dokumente zukünftig nicht mehr über die genannte Zugriffsadresse verfügbar sein sollten, können sie im Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen eingesehen werden.

Der vorliegende Bericht soll wie folgt zitiert werden:

IQWiG. *Cholinesterasehemmer bei Alzheimer Demenz*. Abschlussbericht A05-19A. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); Februar 2007.

KURZFASSUNG

Hintergrund

Im Auftrag des Gemeinsamen Bundesausschusses (G-BA) wurde vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine Nutzenbewertung der Behandlung mit Cholinesterasehemmern bei Alzheimer Demenz durchgeführt.

Fragestellung

Ziele der vorliegenden Untersuchung waren

- die Nutzenbewertung einer langfristigen Behandlung mit einem Cholinesterasehemmer bei der Alzheimer Demenz im Vergleich zu einer Placebo-Gabe hinsichtlich patientenrelevanter Therapieziele sowie
- die Nutzenbewertung einer langfristigen Behandlung mit einem Cholinesterasehemmer bei der Alzheimer Demenz im Vergleich zu einer Behandlung mit einer anderen medikamentösen oder nichtmedikamentösen Therapieoption hinsichtlich patientenrelevanter Therapieziele.

Methoden

Die Bewertung wurde auf Grundlage vorhandener randomisierter kontrollierter Studien zu den oben genannten Fragestellungen durchgeführt. Hierzu wurde eine systematische Literaturrecherche in den bibliographischen Datenbanken MEDLINE, EMBASE und CENTRAL durchgeführt (Erfassung jeweils bis Juni 2006) und in der Datenbank CHID recherchiert. Darüber hinaus wurden Literaturverzeichnisse relevanter Sekundärpublikationen (systematische Übersichten, HTA-Berichte, Meta-Analysen) untersucht sowie die Hersteller von Cholinesterasehemmern bzgl. relevanter veröffentlichter oder unveröffentlichter Studien angefragt.

Eingeschlossen wurden randomisierte kontrollierte Studien, in denen Cholinesterasehemmer (Donepezil, Galantamin und Rivastigmin) bei Patienten mit Alzheimer Demenz untersucht wurden. Das Literaturscreening wurde von 2 Reviewern unabhängig voneinander durchgeführt.

Nach einer Bewertung der Studienqualität wurden die Ergebnisse der einzelnen Studien nach Therapievergleichen und Therapiezielen geordnet gegenübergestellt und beschrieben.

Die vorläufige Nutzenbewertung des IQWiG, der Vorbericht, wurde im Internet veröffentlicht und zur schriftlichen Stellungnahme freigegeben. Unklare Aspekte der schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung diskutiert. Im Anschluss an die Erörterung wurde der Abschlussbericht erstellt.

Ergebnisse

Aus den gesichteten Arbeiten wurden 54 Publikationen zu 33 Studien als relevant beurteilt. Von diesen konnten 48 Publikationen zu 27 Studien in die Nutzenbewertung eingeschlossen werden. 22 Studien waren Placebo-kontrolliert (Donepezil 12, Galantamin 6, Rivastigmin 4), 5 Studien beinhalteten direkte Vergleiche verschiedener Cholinesterasehemmer. Insgesamt wurden in diesen Studien 9883 Patienten untersucht. 11 der eingeschlossenen Publikationen enthielten gepoolte Analysen mehrerer Studien. Studien zum Vergleich mit anderen in Deutschland zugelassenen und verfügbaren medikamentösen und nichtmedikamentösen Interventionen konnten nicht identifiziert werden. Von den eingeschlossenen Studien wiesen 16 leichte und 11 grobe Mängel bezüglich der Studien- und Publikationsqualität auf.

Bis auf zwei Studien (beide zu Donepezil, Dauer jeweils etwa ein Jahr) waren alle Vergleiche mit Placebo lediglich auf einen Behandlungs- bzw. Beobachtungszeitraum von maximal 26 Wochen ausgerichtet. Auch wenn die länger währenden Studien keine grundsätzlich anderen Ergebnisse zeigten, lassen sich belastbare Aussagen im Wesentlichen nur für den 6-Monats-Zeitraum treffen. Immerhin drei der fünf Studien zum Vergleich von verschiedenen Substanzen untereinander hatten dem gegenüber eine Dauer von einem Jahr oder länger. Dafür litten diese Studien wiederum – bis auf eine zu Donepezil und Rivastigmin – neben der eingeschränkten Validität aufgrund eines unverblindeten Designs an zu geringen Fallzahlen, um Unterschiede entdecken, aber auch um die Vergleichbarkeit demonstrieren zu können.

Vergleiche mit Placebo

Bei allen Studien ist ein dosisabhängiger Effekt zu beobachten, wobei im Niedrigdosisbereich bei Galantamin und Rivastigmin im Gegensatz zu Donepezil jeweils keine bzw. nur eine unsichere Wirksamkeit sichtbar ist. Bei Galantamin besteht kein erkennbarer Unterschied zwischen einer Dosierung von 16 mg und 24 mg. Bezüglich der in den Studien berichteten Häufigkeiten von unerwünschten Ereignissen wird diese Dosis-Wirkungs-Beziehung bestätigt.

Für das Therapieziel **Aktivitäten des täglichen Lebens** zeigen sich für alle drei Substanzen bei mittlerer und/oder hoher Dosis Hinweise auf eine günstige Beeinflussung. Die anhand von Meta-Analysen gefundenen durchschnittlichen Effekte betragen etwa drei Scorepunkte auf der DAD- bzw. PDS-Skala für Galantamin bzw. Rivastigmin. Entsprechende Schätzwerte für Donepezil lassen sich nicht ausreichend sicher ableiten, da bei der diesbezüglichen Meta-Analyse von einer Überschätzung des Effekts ausgegangen werden muss. Nichtsdestotrotz kann auch für Donepezil ein Hinweis auf eine günstige Beeinflussung angenommen werden.

Im Hinblick auf die **begleitende Psychopathologie** lassen sich keine Hinweise auf eine günstige oder ungünstige Beeinflussung durch Donepezil und Rivastigmin ableiten; bei Donepezil aufgrund nicht überzeugender Daten, bei Rivastigmin wegen praktisch nicht vorhandener Daten. Bei Galantamin hingegen zeigt sich ein Hinweis auf einen positiven

Effekt, der allerdings mit etwa ein bis zwei Scorepunkten auf der NPI-Skala als äußerst gering einzustufen ist.

Für die drei Substanzen kann der Nutzen bezüglich einer im Vergleich zu Placebo günstigen Beeinflussung der **Kognition** reklamiert werden. Dieser beträgt auf der ADAS-cog-Skala ca. zwei (Donepezil 5 mg bzw. mit flexibler Dosierung) bis drei Scorepunkte (Donepezil 10 mg, Galantamin, Rivastigmin).

Zum Therapieziel **krankheitsbezogene Lebensqualität** liegen nur für Donepezil Daten aus zwei Studien vor, die keinen eindeutigen Hinweis auf einen günstigen oder ungünstigen Effekt liefern. Für Galantamin und Rivastigmin wurden dazu in den Studien überhaupt keine Daten berichtet.

Keinerlei (interpretierbare) Daten fanden sich auch zum Therapieziel **Notwendigkeit einer vollstationären Pflege** (Institutionalisierung).

Bei insgesamt in den Studien nur wenig berichteten Todesfällen lassen sich keine Hinweise auf eine günstige oder ungünstige Beeinflussung der **Mortalität** ableiten.

Bei allen Substanzen sind im höheren Dosisbereich größere Abbruchraten aufgrund **unerwünschter Ereignisse** zu verzeichnen. Es treten zudem häufiger solche unerwünschten Ereignisse auf, die im Einklang mit dem Wirkprinzip der Acetylcholinesterasehemmer stehen (z. B. Übelkeit, Erbrechen, Durchfall). Es findet sich kein Hinweis auf einen größeren Anteil von Patienten mit schwerwiegenden unerwünschten Ereignissen im Vergleich zu Placebo, wobei einschränkend bemerkt werden muss, dass die diesbezügliche Berichterstattung teilweise unzureichend war. Über eher selten bzw. erst nach längerer Einnahmedauer auftretende unerwünschte Ereignisse können aufgrund der Anlage der Studien und der Methodik der Berichterstattung kaum Aussagen getroffen werden.

Aus den vorliegenden Ergebnissen lassen sich für Donepezil keine Hinweise auf eine günstige oder ungünstige Beeinflussung der **Lebensqualität der betreuenden Angehörigen** erkennen. Bei Galantamin zeigt sich ein Hinweis auf einen – mit einer Größenordnung von einem Zehntel Standardabweichung allerdings als äußerst gering einzustufenden – positiven Effekt. Für Rivastigmin wurden dazu keinerlei Daten gefunden.

Es liegen Hinweise vor, dass zu Rivastigmin Daten bezüglich des Therapieziels **Betreuungsaufwand** in allen vier größeren Phase-III-Studien erhoben wurden. Diese Daten wurden allerdings bislang nicht publiziert, so dass hierzu keine Aussagen getroffen werden können. Die Daten für Donepezil sind aus vorwiegend methodischen Gründen nur wenig belastbar, so dass sich daraus ebenfalls kein Hinweis auf eine günstige Beeinflussung der Betreuungszeiten ergibt. Für Galantamin liegt aus einer Studie ein Hinweis auf einen diesbezüglich positiven Effekt vor.

Der **klinische Gesamteindruck** wird konsistent durch alle Substanzen gebessert.

Für Galantamin und Rivastigmin liegen Hinweise vor, dass bei schwerer beeinträchtigten Personen der Behandlungseffekt größer ist als bei weniger schwer beeinträchtigten Personen. Bezüglich Alter, Geschlecht und begleitende Komorbidität lassen sich keine entsprechend differenzierten Aussagen treffen.

Vergleiche untereinander

Eine quantitative Zusammenfassung (Meta-Analyse) der vergleichenden Ergebnisse zu einzelnen Zielkriterien war aufgrund der begrenzten Anzahl von Studien und der teils sehr unterschiedlichen Operationalisierungen nicht angemessen. Nur zwei der fünf Studien verfügten über eine für die Detektion von moderaten Unterschieden ausreichende Fallzahl.

Für den Vergleich von Donepezil und Galantamin liefern die beiden einbezogenen Studien keinen eindeutigen Hinweis auf eine Überlegenheit einer der beiden Substanzen bezogen auf die Beeinflussung der Aktivitäten des täglichen Lebens, der begleitenden Psychopathologie, der Kognition und der therapiebezogenen unerwünschten Ereignisse. Zur krankheitsbezogenen Lebensqualität der Patienten, zur Frage der Institutionalisierung und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden bzw. keine eindeutig interpretierbaren Daten berichtet.

Beim Vergleich von Donepezil mit Rivastigmin weisen die Daten aus einer Studie bezüglich der Beeinflussung der Aktivitäten des täglichen Lebens zwar auf eine geringfügige Überlegenheit von Rivastigmin hin (Effektschätzer in der Größenordnung von einem Zehntel Standardabweichung), allerdings bestehen aus methodischen Gründen Zweifel an der Validität dieser Daten. Im Hinblick auf die begleitende Psychopathologie, die Kognition und die Mortalität gibt es keinen eindeutigen Hinweis auf einen Unterschied zwischen den beiden Substanzen. Unter Rivastigmin zeigen sich deutlich höhere Raten beim Auftreten von unerwünschten Ereignissen, insbesondere bezüglich Übelkeit, Erbrechen, Appetitverlust und Gewichtsabnahme. Zur krankheitsbezogenen Lebensqualität der Patienten und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden Daten berichtet.

Zum Vergleich von Galantamin und Rivastigmin liegen lediglich Ergebnisse aus einem Drei-Gruppen-Vergleich mit sehr niedrigen Fallzahlen vor. Hier sind keine Unterschiede bezüglich der Beeinflussung psychopathologischer Befunde und des Auftretens unerwünschter Ereignisse zu erkennen. Zu den anderen Therapiezielen liegen keine Daten vor.

Insgesamt lässt sich aus den nicht vorhandenen bzw. allenfalls geringfügigen (und mit einiger Unsicherheit behafteten) Unterschieden bei den Wirksamkeitsparametern in den Vergleichsstudien kein Nachweis einer Überlegenheit einer der Wirkstoffe den anderen gegenüber ableiten. Die Ergebnisse können allerdings auch nicht im Sinne einer Vergleichbarkeit der Substanzen interpretiert werden, da alle Studien nicht erkennbar als

Äquivalenz- bzw. Nichtunterlegenheitsstudien mit einer A-priori-Definition von irrelevanten Unterschieden geplant waren.

Fazit

Die Cholinesterasehemmer Donepezil, Galantamin und Rivastigmin haben bei Patienten mit einer Alzheimer Demenz leichten bis mittleren Schweregrades einen Nutzen bezüglich des Therapieziels der kognitiven Leistungsfähigkeit. Für Donepezil gilt dies über alle eingesetzten Dosen hinweg, für Galantamin und Rivastigmin nur bei mittleren und hohen Dosen.

Weiterhin gibt es für alle drei Substanzen Hinweise auf einen Nutzen im Hinblick auf das Therapieziel der Aktivitäten des täglichen Lebens.

Für Galantamin gibt es darüber hinaus Hinweise auf einen Nutzen bezüglich der begleitenden psychopathologischen Symptome. Für Donepezil deuten die vorliegenden Daten auf keinen entsprechenden Nutzen, für Rivastigmin gibt es diesbezüglich keine Daten.

Für das Therapieziel der krankheitsbezogenen Lebensqualität liegen entweder ebenfalls keine Daten vor (Galantamin und Rivastigmin) oder sie liefern keinen Hinweis auf einen Nutzen (Donepezil).

Zur Vermeidung der Notwendigkeit einer vollstationären Pflege (Institutionalisierung) liegen keine interpretierbaren Daten vor.

Dem gegenüber steht für alle drei Wirkstoffe ein dosisabhängiger Schaden durch das Auslösen therapieassoziiierter unerwünschter Ereignisse. Eine Beeinflussung der Mortalität lässt sich aus den vorliegenden Daten nicht ableiten, allerdings waren die Studien auch nicht angelegt, dazu Aussagen zu treffen.

Während es im direkten Vergleich Hinweise auf einen Zusatznutzen durch Rivastigmin gegenüber Donepezil bezüglich der Aktivitäten des täglichen Lebens gibt, hat Rivastigmin andererseits auch ein größeres Schadenspotenzial. Zu den anderen beiden möglichen Vergleichen (Galantamin versus Donepezil bzw. Galantamin versus Rivastigmin) lassen sich keine Aussagen treffen. Insgesamt lässt sich kein eindeutiger Vorteil für eine der drei untersuchten Substanzen aus den vorliegenden Daten ableiten.

Die oben getroffenen Aussagen beziehen sich im Wesentlichen auf den den meisten Studien zugrunde liegenden Zeitraum von bis zu 6 Monaten. Für eine weitergehende Nutzen-Schaden-Abwägung wären direkte Vergleichsstudien zu anderen Behandlungsoptionen (andere Medikamente oder nichtmedikamentöse Behandlungsstrategien) wünschenswert.

Der Stellenwert der Cholinesterasehemmer gegenüber anderen medikamentösen und nichtmedikamentösen Behandlungsoptionen ist mangels Daten unklar.

Schlüsselwörter: Cholinesterasehemmer, Donepezil, Galantamin, Rivastigmin, Alzheimer Demenz, systematische Übersicht

INHALTSVERZEICHNIS

	Seite
KURZFASSUNG.....	v
TABELLENVERZEICHNIS	xiv
ABBILDUNGSVERZEICHNIS	xvi
ABKÜRZUNGSVERZEICHNIS.....	xix
1 HINTERGRUND.....	1
1.1 Beschreibung der zugrunde liegenden Erkrankung.....	1
1.2 Überblick zur Epidemiologie und Pathogenese der Alzheimer Demenz	2
1.3 Diagnose der Alzheimer Demenz.....	3
1.4 Möglichkeiten der Behandlung der Alzheimer Demenz.....	5
1.5 Therapieevaluation bei der Alzheimer-Demenz.....	6
2 ZIELE DER UNTERSUCHUNG	9
3 PROJEKTABLAUF.....	10
3.1 Verlauf des Projekts.....	10
3.2 Zusammenfassung der Änderungen im Vergleich zum Vorbericht.....	11
4 METHODEN	12
4.1 Kriterien für den Einschluss von Studien in die Untersuchung	12
4.1.1 Population.....	12
4.1.2 Interventionen und Vergleichsbehandlungen.....	12
4.1.3 Zielgrößen	13
4.1.4 Studientypen.....	14
4.1.5 Sonstige Studiencharakteristika	14
4.1.6 Ein-/Ausschlusskriterien	14
4.2 Informationsbeschaffung.....	15
4.2.1 Literaturrecherche	15
4.2.2 Suche nach weiteren publizierten und nicht publizierten Studien	16
4.2.3 Identifizierung relevanter Studien.....	16
4.2.4 Suche nach zusätzlichen Informationen zu relevanten Studien	17
4.2.5 Informationen aus der Anhörung zum Vorbericht	17
4.3 Informationsbewertung	18
4.4 Informationssynthese und -analyse	21
4.4.1 Charakterisierung der Studien.....	21

4.4.2	Gegenüberstellung der Ergebnisse der Einzelstudien	21
4.4.3	Meta-Analyse	23
4.4.4	Sensitivitätsanalyse	24
4.4.5	Subgruppenanalyse.....	24
4.5	Änderungen im Vergleich zum Berichtsplan.....	25
4.5.1	Änderungen während der Erstellung des Vorberichts.....	25
4.5.2	Änderungen nach Veröffentlichung des Vorberichts.....	25
5	ERGEBNISSE.....	26
5.1	Ergebnisse der Informationsbeschaffung	26
5.1.1	Ergebnis der Literaturrecherche	26
5.1.2	Anfrage an Hersteller	28
5.1.3	Anfrage an Autoren	29
5.1.4	Resultierender Studienpool	29
5.1.5	Informationen aus der Anhörung	35
5.2	Charakteristika der in die Bewertung eingeflossenen Studien	36
5.2.1	Donepezil	36
5.2.2	Galantamin	53
5.2.3	Rivastigmin	65
5.2.4	Vergleiche unterschiedlicher Cholinesterasehemmer	73
5.3	Ergebnisse zu Therapiezielen	84
5.3.1	Donepezil	84
5.3.2	Galantamin	114
5.3.3	Rivastigmin	135
5.3.4	Vergleiche unterschiedlicher Cholinesterasehemmer	149
5.3.5	Sensitivitätsanalyse	160
5.3.6	Subgruppenauswertungen	160
5.4	Zusammenfassung.....	163
6	DISKUSSION	169
7	FAZIT	185
8	LISTE DER EINGESCHLOSSENEN STUDIEN.....	186
9	LITERATUR	192
ANHANG A: SUCHSTRATEGIE		204
ANHANG B: LISTE DER IM VOLLTEXT ÜBERPRÜFTEN, ABER AUSGESCHLOSSENEN STUDIEN MIT AUSSCHLUSSGRÜNDEN		215

ANHANG C: LISTE DER GESCREENTEN SYSTEMATISCHEN ÜBERSICHTEN.....	224
ANHANG D: ÜBERBLICK ÜBER DIE IN DEN STUDIEN ERHOBENEN ERGEBNISMAßE.....	226
ANHANG E: KURZBESCHREIBUNG DER ERGEBNISMAßE.....	230
ANHANG F: BISLANG UNVERÖFFENTLICHTE INFORMATIONEN PHARMAZEUTISCHER UNTERNEHMEN.....	234
ANHANG G: ZUSÄTZLICHE META-ANALYSEN	235
ANHANG H: PROTOKOLL DER WISSENSCHAFTLICHEN ERÖRTERUNG	241
ANHANG I: STELLUNGNAHMEN	268

TABELLENVERZEICHNIS

Tabelle 1: ICD-10-Kriterien für eine Alzheimer Demenz	3
Tabelle 2: DSM-IV-Kriterien für eine Alzheimer Demenz	4
Tabelle 3: NINCDS-ADRDA Kriterien für eine Alzheimer Demenz	5
Tabelle 4: Studienpool	30
Tabelle 5: Eingeschlossene Studien zu Donepezil	40
Tabelle 6: Wesentliche Ein- und Ausschlusskriterien der Studien zu Donepezil	42
Tabelle 7: Charakterisierung der Studienpopulationen Donepezil	45
Tabelle 8: Studien- und Publikationsqualität Donepezil.....	50
Tabelle 9: Eingeschlossene Studien zu Galantamin.....	56
Tabelle 10: Wesentliche Ein- und Ausschlusskriterien der Studien zu Galantamin.....	57
Tabelle 11: Charakterisierung der Studienpopulationen Galantamin	59
Tabelle 12: Studien- und Publikationsqualität Galantamin.....	63
Tabelle 13: Eingeschlossene Studien zu Rivastigmin.....	67
Tabelle 14: Wesentliche Ein- und Ausschlusskriterien der Studien zu Rivastigmin.....	68
Tabelle 15: Charakterisierung der Studienpopulationen Rivastigmin	69
Tabelle 16: Studien- und Publikationsqualität Rivastigmin.....	72
Tabelle 17: Eingeschlossene direkte Vergleichsstudien zwischen Cholinesterasehemmern...	76
Tabelle 18: Wesentliche Ein- und Ausschlusskriterien der direkten Vergleichsstudien zwischen Cholinesterasehemmern.....	78
Tabelle 19: Charakterisierung der Studienpopulationen zu direkten Vergleichsstudien zwischen Cholinesterasehemmern.....	80
Tabelle 20: Studien- und Publikationsqualität Vergleichsstudien	83
Tabelle 21: Donepezil: Ergebnisse zu alltagspraktischen Fähigkeiten	85
Tabelle 22: Donepezil: Ergebnisse zur begleitenden Psychopathologie.....	89
Tabelle 23: Donepezil: Ergebnisse zu kognitiven Fähigkeiten	91
Tabelle 24: Donepezil: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen	97
Tabelle 25: Donepezil: unerwünschte Ereignisse	99
Tabelle 26: Donepezil: Ergebnisse zum klinischen Gesamteindruck	113

Tabelle 27: Galantamin: Ergebnisse zu alltagspraktischen Fähigkeiten.....	116
Tabelle 28: Galantamin: Ergebnisse zur begleitenden Psychopathologie	118
Tabelle 29: Galantamin: Ergebnisse zu kognitiven Fähigkeiten.....	120
Tabelle 30: Galantamin: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen	123
Tabelle 31: Galantamin: unerwünschte Ereignisse	125
Tabelle 32: Galantamin: Ergebnisse zur Lebensqualität der (betreuenden) Angehörigen.....	131
Tabelle 33: Galantamin: Ergebnisse zu CIBIC-plus ≤ 4 (Stabilität oder Verbesserung).....	134
Tabelle 34: Rivastigmin: Ergebnisse zu alltagspraktischen Fähigkeiten.....	136
Tabelle 35: Rivastigmin: Ergebnisse zu kognitiven Fähigkeiten.....	138
Tabelle 36: Rivastigmin: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen	141
Tabelle 37: Rivastigmin: unerwünschte Ereignisse	142
Tabelle 38: Rivastigmin: Ergebnisse zu CIBIC-plus ≤ 3 (Verbesserung)	148
Tabelle 39: Rivastigmin vs. Donepezil: Ergebnisse zu alltagspraktischen Fähigkeiten	152
Tabelle 40: Rivastigmin vs. Donepezil: Ergebnisse zu kognitiven Fähigkeiten.....	153
Tabelle 41: Rivastigmin vs. Donepezil: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen	156
Tabelle 42: Rivastigmin vs. Donepezil: unerwünschte Ereignisse	157
Tabelle 43: Anzahl eingeschlossener Studien und randomisierter Patienten.....	163
Tabelle 44: Zusammenfassung der Ergebnisse zu den Therapiezielen aus Placebo-kontrollierten Studien	165
Tabelle 45: Zusammenfassung der Ergebnisse zu den Therapiezielen aus Vergleichsstudien von Cholinesterasehemmern untereinander.....	167
Tabelle 46: Geschätzter und beobachteter Anteil Responder für den ADAS-cog bei Responsekriterium 4 Scorepunkte	171
Tabelle 47: Gegenüberstellung von Zielkriterien des Berichtsplan und in den Studien verwendeten Ergebnismaßen.....	227

ABBILDUNGSVERZEICHNIS

Abbildung 1: Bibliographische Literaturrecherche und sonstiges Literaturscreening, endgültiger Studienpool für die Nutzenbewertung.....	27
Abbildung 2: Donepezil: Meta-Analyse zu Aktivitäten des täglichen Lebens	87
Abbildung 3: Donepezil: Meta-Analyse zur begleitenden Psychopathologie	90
Abbildung 4: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten.....	93
Abbildung 5: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – niedrige Dosis	93
Abbildung 6: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – flexible Dosis	94
Abbildung 7: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – hohe Dosis.....	94
Abbildung 8: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – niedrige Dosis	101
Abbildung 9: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – flexible Dosis.....	101
Abbildung 10: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – hohe Dosis.....	102
Abbildung 11: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – niedrige Dosis	102
Abbildung 12: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – flexible Dosis.....	103
Abbildung 13: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – hohe Dosis	103
Abbildung 14: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – niedrige Dosis	104
Abbildung 15: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – flexible Dosis	104
Abbildung 16: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – hohe Dosis.....	105
Abbildung 17: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – niedrige Dosis.....	105
Abbildung 18: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – flexible Dosis.....	106
Abbildung 19: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – hohe Dosis	106
Abbildung 20: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – niedrige Dosis	106
Abbildung 21: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – flexible Dosis.....	107
Abbildung 22: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – hohe Dosis	107
Abbildung 23: Donepezil: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ – niedrige Dosis	108

Abbildung 24: Donepezil: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ – hohe Dosis..	108
Abbildung 25: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – niedrige Dosis	109
Abbildung 26: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – flexible Dosis	109
Abbildung 27: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – hohe Dosis.....	110
Abbildung 28: Galantamin: Meta-Analyse zu Aktivitäten des täglichen Lebens.....	117
Abbildung 29: Galantamin: Meta-Analyse zur begleitenden Psychopathologie	119
Abbildung 30: Galantamin: Meta-Analyse zu kognitiven Fähigkeiten	121
Abbildung 31: Galantamin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse	127
Abbildung 32: Galantamin: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen	127
Abbildung 33: Galantamin: Meta-Analyse zum Endpunkt „Übelkeit“	128
Abbildung 34: Galantamin: Meta-Analyse zum Endpunkt „Erbrechen“	128
Abbildung 35: Galantamin: Meta-Analyse zum Endpunkt „Diarrhoe“	129
Abbildung 36: Galantamin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“	129
Abbildung 37: Galantamin: Meta-Analyse zum Endpunkt „Schwindel“	130
Abbildung 38: Galantamin: Meta-Analyse zur Lebensqualität der (betreuenden) Angehörigen	132
Abbildung 39: Rivastigmin: Meta-Analyse zu alltagspraktischen Fähigkeiten.....	137
Abbildung 40: Rivastigmin: Meta-Analyse zu kognitiven Fähigkeiten.....	139
Abbildung 41: Rivastigmin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse	143
Abbildung 42: Rivastigmin: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen	143
Abbildung 43: Rivastigmin: Meta-Analyse zum Endpunkt „Übelkeit“.....	144
Abbildung 44: Rivastigmin: Meta-Analyse zum Endpunkt „Erbrechen“	144
Abbildung 45: Rivastigmin: Meta-Analyse zum Endpunkt „Diarrhoe“	145
Abbildung 46: Rivastigmin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“.....	145
Abbildung 47: Rivastigmin: Meta-Analyse zum Endpunkt „Schwindel“	146
Abbildung 48: Zusammenhang zwischen in den Placebo-kontrollierten Studien beobachtetem Effekt auf die Kognition und Differenz der Abbruchraten von Verum und Placebo.	173

Abbildung 49: Rivastigmin: Meta-Analyse zu alltagspraktischen Fähigkeiten (mit B351, feste und zufällige Effekte)	235
Abbildung 50: Rivastigmin: Meta-Analyse zu kognitiven Fähigkeiten (mit B351, feste und zufällige Effekte)	236
Abbildung 51: Rivastigmin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse (mit B351, feste und zufällige Effekte)	237
Abbildung 52: Rivastigmin: Meta-Analyse zum Endpunkt „Übelkeit“ (mit B351, feste und zufällige Effekte)	238
Abbildung 53: Rivastigmin: Meta-Analyse zum Endpunkt „Erbrechen“ (mit B351)	239
Abbildung 54: Rivastigmin: Meta-Analyse zum Endpunkt „Schwindel“ (mit B351)	239
Abbildung 55: Rivastigmin: Meta-Analyse zum Endpunkt „Diarrhoe“ (mit B351)	240
Abbildung 56: Rivastigmin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ (mit B351) ..	240

ABKÜRZUNGSVERZEICHNIS

Abkürzung	Bedeutung
ABS	Adaptive Behavior Scale
ACTS	Allocation of Caregiver Time Survey
AD	Alzheimer Demenz
ADAS-cog	Alzheimer's Disease Assessment Scale – cognitive subscale
ADCS-ADL	AD Cooperative Study Activities of Daily Living Inventory
AD-CVD	Alzheimer Demenz und gleichzeitig vorliegende zerebrovaskuläre Erkrankung
ADEAR	Alzheimer's Disease Education & Referral Center
ADFACS	AD Functional Assessment and Change Scale
ADL	Activities of Daily Living
ANCOVA	Kovarianzanalyse
ANOVA	Varianzanalyse
APP	Amyloid-Precursor-Protein
A β	β -Amyloid-Peptid
bADL	Basic Activities of Daily Living
BADLS	Bristol Activities of Daily Living Scale
BEHAVE-AD	Behavioral Pathology in Alzheimer's Disease Rating Scale
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BRDS	Blessed Roth Dementia Scale
CAS	Caregiver Activity Survey
CBS	Caregiving Burden Scale
CDR, CDR-SB	Clinical Dementia Rating Scale, Clinical Dementia Rating Scale – Sum of the Boxes
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CGIC	Clinical Global Impression of Change
CHID	Combined Health Information Database

Abkürzung	Bedeutung
CIBIC	Clinician's Interview-based Impression of Change
CIBIC-plus	Clinician's Interview-based Impression of Change with additional caregiver input
CMBT	Computerized Memory Battery Test
CMCS	Caregiver-rated Modified Crichton Scale
CSS	Caregiver Stress Scale
CT	Computertomographie
CVD	Zerebrovaskuläre Erkrankung
DAD	Disability Assessment for Dementia Scale
DARE	Database of Abstracts of Reviews of Effects
DEMQOL	Dementia Quality of Life Instrument
DMR	Dementia Questionnaire for Mentally Retarded Persons
DON	Donepezil
DRC	Diagnostic Research Criteria
DSM	Diagnostic and Statistical Manual of Mental Disorders
EMBASE	Excerpta Medica Database
EMA	The European Agency for the Evaluation of Medicinal Products
FAST	Functioning Assessment Staging Scale
FRS	Functional Rating Scale
GAL	Galantamin
GAL-PRC	Galantamin Prolonged Release
GAS	Goal Attainment Scale
G-BA	Gemeinsamer Bundesausschuss
GBS	Gottfries-Bråne-Steen Scale
GCP	Good Clinical Practice
GDS	Global Deterioration Scale
HTA	Health Technology Assessment
IADL, iADL	Instrumental Activities of Daily Living

Abkürzung	Bedeutung
ICD	International Classification of Diseases
IDDD	Interview for Deterioration in Daily Living Activities in Dementia
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-To-Treat
J-CGIC	Japanese-Clinical Global Impression of Change
k.A.	keine Angaben
KI	Konfidenzintervall
LJ	Lebensjahr
LOCF	Last Observation Carried Forward-Analyse
MCI	Mild Cognitive Impairment
MEDLINE	Medical Literature Analysis and Retrieval System Online
MENFIS	Mental Function Impairment Scale
MMSE (sMMSE)	(standardized) Mini Mental State Examination
MRT	Magnetresonanztomographie
N	Number
n.g.	nicht genannt
n.s.	nicht statistisch signifikant
NICE	National Institute for Health and Clinical Excellence
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer’s Disease and Related Disorders Association
NOSGER	Nurses’ Observation Scale for Geriatric Patients
NPI	Neuropsychiatric Inventory
NPI-D	Neuropsychiatric Inventory Caregiver Distress Scale
NPI-NH	Neuropsychiatric Interview – Nursing Home
OC	Observed Cases
p	p-Wert, Wahrscheinlichkeit
PDS	Progressive Deterioration Scale
PET	Positronenemissionstomographie

Abkürzung	Bedeutung
PGAS	Patient Global Assessment Scale
PP	Per Protokoll
PSMS	Physical Self-Maintenance Scale
QoL	Quality of Life
RCT	Randomized Controlled Trial
RIV	Rivastigmin
RUD	Resource Utilization in Dementia
SCGB	Screen for Caregiver Burden
SD	Standard Deviation
SEM	Standard Error of the Mean
SF-36	36-item Short-Form Health Survey
SIB	Severe Impairment Battery
SN	Stellungnahme(n)
UE	unerwünschtes Ereignis
VaD	Vaskuläre Demenz
VFA	Verband Forschender Arzneimittelhersteller e.V.
w	weiblich
WHO	World Health Organization
ZNS	Zentrales Nervensystem

1 HINTERGRUND

1.1 Beschreibung der zugrunde liegenden Erkrankung

Mit dem Begriff der Demenz wird eine chronische und meist progressive Fehlfunktion des Gehirns bezeichnet, die zu einer Verschlechterung des Gedächtnisses und anderer kognitiver Funktionen, zu einer Beeinträchtigung der Aktivitäten des täglichen Lebens und zu wechselnd ausgeprägten, begleitenden psychopathologischen Symptomen führt. Die Alzheimer Demenz (AD) stellt mit einem Anteil von ca. 60 % die häufigste alleinige Ursache eines demenziellen Syndroms dar. Daneben kommt die Alzheimer Demenz auch häufig als Mischform mit der vaskulär bedingten Demenz vor [1-3]. Andere Ursachen der Demenz (Morbus Pick, Lewy-Body-Demenz, Creutzfeld-Erkrankung u.a.) sowie reversible demenzielle Zustandsbilder im Gefolge anderer Erkrankungen sind deutlich seltener.

Das demenzielle Syndrom ist gekennzeichnet durch eine Störung in mehreren Bereichen (u. a. Gedächtnis, Denkvermögen, emotionale Kontrolle). Während speziell die Alzheimer Demenz meist zuerst wegen kognitiver Symptome, insbesondere der Merkfähigkeitsstörung auffällt, bestimmen die Beeinträchtigungen der alltagspraktischen Fähigkeiten, besonders die basalen Fähigkeiten, den Pflegeaufwand. Daneben spielen begleitende psychopathologische Symptome wie Apathie, Depression, Agitiertheit, Ängstlichkeit, Schlafstörung und paranoide Symptome eine wesentliche Rolle im Hinblick auf die Lebensqualität. Psychopathologische Symptome und Verhaltensänderungen, die gelegentlich auch schon in frühen Erkrankungsstadien auftreten, können sowohl die Betroffenen als auch die pflegenden Angehörigen in hohem Maß belasten.

Als Risikopersonen für die Entwicklung einer Demenz werden Menschen mit einer leichten kognitiven Beeinträchtigung, „Mild Cognitive Impairment“ (MCI), angesehen. Beim MCI sind die Gedächtnisleistungen beeinträchtigt, ohne dass das Kriterium einer mehr als geringfügigen Einschränkung der alltagspraktischen Fähigkeiten erfüllt ist, somit also noch nicht von einer Demenz gesprochen werden kann (siehe auch die Übersicht bei Petersen RC 2001 [4]).

Studien, die diese Personengruppe betreffen, sind, angesichts der laufenden Diskussionen um die Abgrenzung des Begriffes „MCI“, nicht Gegenstand des vorliegenden Berichtes. Für die Indikation „MCI“ sind Cholinesterasehemmer nicht zugelassen.

Eine Differenzierung zwischen einer Demenz vom Alzheimer Typ und einer vaskulären Demenz kann aufgrund klinischer Kriterien vorgenommen werden, die jedoch im Einzelfall zu einer Fehldiagnose führen können [5,6]. Bildgebende Verfahren zur Erkennung klinisch nicht manifester zerebraler Ischämien können hierbei hilfreich sein. Klinisch abgrenzbare Formen einer Demenz vom Alzheimer Typ oder einer vaskulären Demenz lassen sich

häufiger bei jüngeren Patienten finden, jenseits des 75. Lebensjahres sind dagegen Mischpathologien aus alzheimerartigen und vaskulären Läsionen häufig [7].

1.2 Überblick zur Epidemiologie und Pathogenese der Alzheimer Demenz

Das Auftreten einer Alzheimer Demenz ist nach Korrektur für die unterschiedliche Lebenserwartung bei Frauen etwas häufiger als bei Männern [8]. Andere in Studien bestätigte Risiken sind neben dem Alter und dem Geschlecht eine positive Familienanamnese und das Vorhandensein des E4-Allels des ApoE-Gens.

In einer Auswertung europäischer epidemiologischer Studien mit operationalisierten Diagnosekriterien fanden sich folgende Prävalenzraten für demenzielle Syndrome: 1 % (65- bis 69-Jährige), 4 % (70- bis 74-Jährige), 6 % (75- bis 79-Jährige), 13 % (80- bis 84-Jährige), 22 % (85- bis 89-Jährige) und 32 % (90- bis 94-Jährige) [9]. Die Erkrankung ist kontinuierlich progredient und mit einem erhöhten Sterblichkeitsrisiko verbunden. Die mittlere Überlebenszeit nach Diagnosestellung ist deutlich niedriger als bei der gleichaltrigen Allgemeinbevölkerung und ist mit derjenigen bei anderen schweren Erkrankungen im Alter vergleichbar [10]. Aufgrund der zunehmenden Bemühungen um eine Früherkennung der Alzheimer Demenz, des steigenden Bewusstseins in der Allgemeinbevölkerung und wachsender Angebote in Form von „Gedächtnissprechstunden“ kann eine scheinbare Verlängerung der Krankheitsdauer durch eine vorverlegte Diagnosestellung stattfinden [11,12]. Bei Diagnosen im Frühstadium findet sich in den ersten 6 Monaten meist nur eine geringe Krankheitsprogression.

In der Pathogenese der Alzheimer Demenz spielen der Metabolismus des Amyloid-Precursor-Proteins (APP) und sein Abbau zu β -Amyloid-Peptiden ($A\beta$) eine herausragende Rolle. Die Evidenz hierfür wurde anhand von Mutationen in den humanen Genen des Amyloidmetabolismus und Untersuchungen an transgenen Mäusen gewonnen [13-15]. Neben einem primären Defekt im APP-Stoffwechsel mit einer vermehrten Bildung von $A\beta$ besonders bei früh erkrankten, genetisch belasteten Patienten wird auch ein primärer Defekt im Abbau von $A\beta$ als Ursache bei den sporadischen Formen der Alzheimer Demenz diskutiert. Insbesondere im höheren Alter bzw. bei fortschreitender Pathologie scheinen vielfache Vernetzungen zu weiteren gewebeschädigenden Prozessen zu bestehen (oxidativer Stress, Zellschädigung durch glutamaterge Exzitotoxizität, entzündliche Vorgänge, Bildung von Neurofibrillen, neuronale Apoptose). Die Anfänge der neuropathologischen Veränderungen gehen den manifesten kognitiven Störungen um viele Jahre voraus.

Die Alzheimer Demenz wird in verschiedene Schweregrade eingeteilt, beispielsweise anhand des Mini Mental State Examination (MMSE)-Scores oder der Global Deterioration Scale (GDS). Die meisten Betroffenen werden im Stadium einer leichten bis mittelschweren Demenz diagnostiziert [16]. Eine Verschiebung zu einer früheren Diagnosestellung könnte aufgrund der oben angeführten Gründe (Veränderung des Krankheitsbewusstseins in der

Bevölkerung, Erwartung von zunehmenden Behandlungsmöglichkeiten, neue diagnostische Tests) in naher Zukunft stattfinden [11,12].

1.3 Diagnose der Alzheimer Demenz

Zur Diagnose einer Alzheimer Demenz sind zurzeit keine ausreichend sensitiven und spezifischen Surrogatparameter in der klinischen Routine verfügbar. Neben dem Nachweis des typischen klinischen Bildes erfolgt die Diagnose auch durch den Ausschluss möglicher Alternativursachen. Die Diagnose einer Alzheimer Demenz wird international anhand der ICD-10- (Tabelle 1) oder der verwandten DSM-IV-Kriterien (Tabelle 2) gestellt. Die ICD-9 und die DSM-III-R sind Vorläuferversionen der Klassifikationssysteme ICD-10 und DSM-IV.

Die ICD-10-Kriterien erfordern das Vorliegen eines demenziellen Syndroms mit schleichendem Beginn und langsamer Verschlechterung, den Ausschluss anderer neurologischer oder systemischer Ursachen für die Demenz (wie endokrinologischer Erkrankungen, Vitamin-B12-Mangel, Neurosyphilis oder Hydrozephalus) und die Abwesenheit eines plötzlichen Beginns, fokaler neurologischer Symptome wie Hemiparesen und Gesichtsfeldausfällen. Die definitive Diagnose der Alzheimer Demenz kann lediglich post mortem durch eine Autopsie des Gehirns erfolgen, im seltenen Fall einer dominant vererbten Form auch durch eine genetische Mutationsanalyse.

Tabelle 1: ICD-10-Kriterien für eine Alzheimer Demenz [17]

Vorliegen einer Demenz	1. Gedächtnisstörung
	2. weitere kognitive Störung
	3. dadurch bedingte Beeinträchtigung im Alltag
	Symptome nicht nur im Rahmen eines Delirs, ein Bestehen der Symptome > 6 Monate erhöht die diagnostische Sicherheit
Verlauf	schleichend, kein plötzlicher Beginn
Andere Erkrankungen	kein Hinweis auf/Ausschluss von z. B. endokrin-metabolischen Erkrankungen, Normaldruckhydrozephalus, subdurales Hämatom
Kein frühzeitiges Auftreten neurologischer Herdzeichen	z. B. Hemiparese, Ataxie, Hemianopsie
Verschlüsselungen	F00.0 AD früher Beginn (< 65. LJ) F00.1 AD später Beginn (≥ 65. LJ)

Die ICD-10 und DSM-IV definieren zwei Subtypen der Alzheimer Demenz: die Alzheimer Demenz mit frühem Beginn (Beginn der Demenz vor dem bzw. bis zum 65. Lebensjahr) und die Alzheimer Demenz mit spätem Beginn (erste klinische Symptome ab dem bzw. nach dem 65. Lebensjahr). Dieser Unterscheidung kommt bzgl. der Symptomatik, des Verlaufs und der Neuropathologie keine Bedeutung zu. Genetisch bedingte Formen der Alzheimer Demenz führen zumeist zu einem frühen Beginn, sind aber selten, auch bei Patienten mit einem frühen Beginn.

Tabelle 2: DSM-IV-Kriterien für eine Alzheimer Demenz [18]

Kriterium A1	Gedächtnisstörung
Kriterium A2	weitere kognitive Störung A2a Aphasie A2b Apraxie A2c Agnosie A2d Exekutive Störung
Kriterium B	Abfall der Leistungen von einem früher erreichten Niveau mit Behinderung der Alltagsaktivität
Kriterium C	schleichender Beginn und langsame Zunahme
Kriterium D	Die Leistungsminderung ist nicht bedingt durch: D1 andere ZNS-Erkrankungen D2 systemische Erkrankungen D3 Substanzmissbrauch
Kriterium E	Die Leistungsminderung ist nicht nur im Rahmen eines Delirs vorhanden
Kriterium F	Die Symptome lassen sich nicht durch eine andere psychiatrische Erkrankung erklären, z. B. eine Depression oder Schizophrenie
Verschlüsselungen	290.10 AD früher Beginn (≤ 65 . LJ) 290.0 AD später Beginn (> 65 . LJ)

Neben den ICD-10- und DSM-IV-Kriterien kommen in klinischen Studien am häufigsten die Kriterien der National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) [19] zur Anwendung. Danach wird die Diagnose in definitiv (bei klinischer Diagnose mit histologischer Bestätigung), wahrscheinlich („probable“, bei typischer klinischer Symptomatik ohne histologische Bestätigung) und möglich („possible“, bei atypischen Symptomen, Vorliegen einer im Prinzip möglichen alternativen Ursache, die aber im individuellen Fall nicht für die Demenzursache gehalten wird) unterteilt. Die Sensitivität und Spezifität für die Diagnose der wahrscheinlichen Alzheimer Demenz nach den Kriterien der NINCDS-ADRDA betragen etwa 0,65 bzw. 0,75 (Tabelle 3).

Tabelle 3: NINCDS-ADRDA Kriterien für eine Alzheimer Demenz [19]

<p style="text-align: center;"><u>Klinisch mögliche Alzheimer Demenz („possible AD“)</u></p> <ul style="list-style-type: none">• Demenzielles Syndrom mit atypischer Symptomatik oder atypischem Verlauf ohne erkennbare andere neurologische oder internistische Demenzursache• Demenzielles Syndrom mit gleichzeitig vorliegender anderer Erkrankung, die auch eine Demenz erzeugen kann, in diesem Fall aber nicht als entscheidende Ursache angesehen wird• progredientes Defizit in nur einem kognitiven Bereich
<p style="text-align: center;"><u>Klinisch wahrscheinliche Alzheimer Demenz („probable AD“)</u></p> <p>I. notwendige Voraussetzungen:</p> <ul style="list-style-type: none">• Zeichen einer Demenz in der klinischen Untersuchung und bei neuropsychologischen Tests (z. B. MMSE)• Defizite in zwei oder mehr kognitiven Bereichen• fortschreitende Verschlechterung des Gedächtnisses und anderer kognitiver Funktionen• keine Bewusstseinstörung• Beginn zwischen dem 40. und 90. Lebensjahr• Ausschluss einer anderen körperlichen oder neurologischen Krankheit, die für die Symptomatik verantwortlich gemacht werden kann <p>II. unterstützende Befunde:</p> <ul style="list-style-type: none">• fortschreitende Verschlechterung der Sprache (Aphasie), Motorik (Apraxie) und Wahrnehmung (Agnosie)• Beeinträchtigungen des Alltagslebens und Verhaltensänderungen• positive Familienanamnese für Alzheimer Demenz, besonders falls neuropathologisch gesichert• Normalbefund einer Liquoranalyse, unspezifische EEG-Veränderungen, CCT-gesicherte Progression einer zerebralen Atrophie <p>III. mit der Diagnose einer wahrscheinlichen Alzheimer-Demenz vereinbar:</p> <ul style="list-style-type: none">• Plateaus im Krankheitsverlauf• Begleitsymptome wie Depression, Schlaflosigkeit, Inkontinenz, Wahn, Verkennungen, Halluzinationen, „katastrophisierende Reaktionen“, Störungen des Sexualverhaltens, Gewichtsverlust• besonders bei fortgeschrittener Erkrankung: erhöhter Muskeltonus, Myoklonien, Gangstörungen, Krampfanfälle• normales CCT <p>IV. Befunde und anamnestische Angaben, die die Diagnose einer Alzheimer Demenz unwahrscheinlich erscheinen lassen:</p> <ul style="list-style-type: none">• plötzlicher Beginn (apoplexartig)• früh auftretende fokal-neurologische Ausfälle: Hemiparesen, Gesichtsfeldausfälle, Ataxien• früh auftretende Krampfanfälle und Gangstörungen

1.4 Möglichkeiten der Behandlung der Alzheimer Demenz

Die Möglichkeiten der Behandlung der Alzheimer Demenz umfassen verschiedene Bereiche, z. B. nach Cummings 2004 [20]:

- Aufbau von Aktivitäten zur Aufrechterhaltung und Förderung körperlicher und geistiger Gesundheit,
- Koordination der Zusammenarbeit zwischen Therapeuten, Familienangehörigen und anderen Pflegenden,
- Einsatz von Antidementiva, und zwar Acetylcholinesterasehemmern, NMDA(N-methyl-D-Aspartat)-Antagonisten und Ginkgo biloba in den Demenzstadien, für die jeweils eine Zulassung in Deutschland vorliegt,

- Nichtmedikamentöse Strategien und Psychopharmaka zur Behandlung von psychopathologischen Symptomen und Verhaltensstörungen.

In den 70er Jahren des 20. Jahrhunderts wurde in Post-mortem-Untersuchungen von Gehirnen von Patienten mit fortgeschrittener Alzheimer Demenz eine deutlich reduzierte Aktivität zentraler cholinergener Neuronen festgestellt [21]. Diese Befunde wurden mit den kognitiven Defiziten, der Schwere der Alzheimer Demenz [22] und psychischen Symptomen in Zusammenhang gebracht [23]. Zusammen mit der Beobachtung, dass zentral wirksame anticholinerge Substanzen zu einer Reduktion der kognitiven Funktionen führen, folgten daraufhin viele Bestrebungen, durch eine Intensivierung der cholinergen Neurotransmission die demenzielle Symptomatik im Rahmen einer Alzheimer Demenz zu behandeln. Während Acetylcholin-Vorläufersubstanzen als Nahrungsergänzungstoffe keinen Effekt zeigten und cholinerge Agonisten (cholinerg stimulierende Substanzen) bisher zum klinischen Einsatz nicht verfügbar sind, stehen von zahlreichen in klinischen Studien erprobten Cholinesterasehemmern in Deutschland drei Cholinesterasehemmer zur Verfügung (Donepezil, zugelassen 1996, Rivastigmin, zugelassen 1998, Galantamin, zugelassen 2001). Der 1995 in Deutschland unter strengen Auflagen zugelassene Cholinesterasehemmer Tacrin ist hier nicht mehr erhältlich. Die drei Substanzen sind in den meisten anderen europäischen Ländern sowie den USA ebenfalls zugelassen.

1.5 Therapieevaluation bei der Alzheimer-Demenz

Es bestehen deutliche Probleme, die Beeinträchtigungen und Bedürfnisse der von der Alzheimer Demenz Betroffenen zu erfassen. Zum einen können Beeinträchtigungen fluktuieren oder nur in bestimmten Situationen in Erscheinung treten, zum anderen kann der krankheitsbedingte Verlust kommunikativer Fähigkeiten die adäquate Äußerung von Bedürfnissen oder Präferenzen durch den Erkrankten beeinträchtigen. Ein weiteres spezielles Problem demenzieller Erkrankungen ist die bei einem Teil der Patienten fehlende Selbstwahrnehmung eigener Defizite und Hilfsbedürftigkeit. Diese krankheitsbedingt fehlende Selbstwahrnehmung kann nach Ansicht einiger Autoren ebenso zu einer von außen gesehen groben Fehleinschätzung möglicher Beeinträchtigungen und Bedürfnisse führen [24].

Die Auswirkungen der Erkrankung und die Wirksamkeit einer Therapie bei der Alzheimer Demenz können vom Betroffenen selbst (mit den oben erwähnten Einschränkungen), vom Kliniker und von den Angehörigen oder Pflegenden beurteilt werden. Die Bewertung der Ausprägung der Krankheitssymptome und der dadurch bedingten Beeinträchtigungen kann je nach Bereich und Beurteiler variieren [25].

Behandlungs- und Therapiemaßnahmen können angesichts der komplexen und mit zunehmender Krankheitsprogression sich verändernden Bedürfnislage in drei Bereiche gegliedert werden. Psychopharmakologische Therapien (Antidementiva oder andere

psychotrope Medikamente) zielen auf eine direkte Veränderung kognitiver oder psychologischer Leistungen des Demenzerkrankten. Psychologisch-verhaltenstherapeutische Interventionen setzen im Frühstadium der Erkrankung auf modifizierbare Verhaltensmuster oder Ersatzstrategien. Interventionen im Sozialbereich (Angehörigenberatung und -schulung, Tagesstätten, Pflegedienste, betreute Wohnformen) treten ab einem mittleren Erkrankungsstadium in den Vordergrund. Alle drei Interventionsebenen – die biologische (Antidementiva und Psychopharmaka), die psychologische (z. B. kognitive Strategien) und die soziale Ebene (z. B. Angehörigenarbeit, Pflegeunterstützung) – interagieren dabei in hohem Maße.

Immer noch wird kontrovers diskutiert, mit welchen Instrumenten die Wirksamkeit medikamentöser und nichtmedikamentöser Therapien bei der Alzheimer Demenz adäquat evaluiert werden kann [24,26,27]. In den meisten Therapiestudien wird das Behandlungsergebnis auf der Basis der kognitiven Funktion evaluiert, insbesondere mittels der ADAS-cog-Skala (Alzheimer's Disease Assessment Scale – cognitive subscale) [28] oder der MMSE-Skala (Mini Mental State Examination) [29]. Zusätzlich kommen globale Erhebungsskalen wie die CIBIC (Clinician's Interview-based Impression of Change) zur Anwendung [30,31]. Um der Komplexität der Erkrankung gerecht zu werden, werden auch Skalen zur Messung der alltagspraktischen Fähigkeiten (Activities of Daily Living: ADL [32]), oft noch zusätzlich unterteilt in basale ADL (bADL) und instrumentelle ADL (iADL) [33], eingesetzt. Instrumentelle ADL-Bereiche (Umgang mit Geld, Telefonieren etc.) verändern sich besonders zu Beginn der Erkrankung, wogegen basale ADL-Bereiche (z. B. Waschen, Anziehen) eng mit der zunehmenden Pflegebedürftigkeit im mittelschweren Krankheitsstadium verbunden sind. Eine weitere Skala, die die Einschränkungen der Aktivitäten des täglichen Lebens erfasst, ist die PDS (Progressive Deterioration Scale). Zur Messung der begleitenden Psychopathologie, die über den gesamten Krankheitsverlauf, meist aber erst ab einem mittleren Stadium in den Vordergrund tritt, werden spezielle Interviews mit den Pflegenden geführt (z. B. Neuropsychiatric Inventory, NPI [34]).

Während die ADAS-cog-Skala im Rahmen einer direkten Untersuchung des Patienten aufgrund des beobachtbaren Antwortverhaltens von einem erfahrenen Testleiter ausgefüllt wird, beruhen die Angaben in den ADL, der PDS und dem NPI auf Auskünften eines Angehörigen. Zu diesen häufig verwendeten Skalen gibt es Untersuchungen zur Validität und Reliabilität, die jedoch zum Teil leicht abweichende Ergebnisse erbrachten [35-38].

Es existieren bisher wenige Verfahren zur Messung der gesundheitsbezogenen Lebensqualität bei der Alzheimer Demenz, was mit den Schwierigkeiten zusammenhängt, verlässliche Informationen von den Betroffenen direkt zu erhalten. Der negative Einfluss der Erkrankung auf die Lebensqualität ist jedoch häufig beschrieben [39], ebenso die Fähigkeit der Betroffenen, über ihre subjektiven Erfahrungen zu berichten [40]. In einer neueren Studie wurden Instrumente zur Erfassung der Lebensqualität untersucht und ein neues

krankheitsspezifisches Instrument entwickelt und validiert [24]. Es zeigte sich, dass Betroffene in der Einschätzung der gesundheitsbezogenen Lebensqualität deutlich von ihren Angehörigen oder dem Pflegepersonal abweichen können. Die Übereinstimmung ist höher bei Betroffenen, die eng mit Angehörigen zusammenleben, und niedriger bei starker kognitiver Beeinträchtigung oder ausgeprägten affektiven Symptomen [41].

Der Einfluss der Alzheimer Demenz auf Angehörige und Pflegende steigt mit dem Fortschreiten der Erkrankung. Sowohl die Hilfsbedürftigkeit im Alltag als auch Verhaltensprobleme, Kommunikationsschwierigkeiten und eine mögliche soziale Stigmatisierung können für Angehörige sehr belastend sein. Die Intensität der Betreuung, insbesondere der Zeitaufwand, spiegelt jedoch nicht zwangsläufig die psychische und physische Belastung der Angehörigen wider. Der Einfluss der Erkrankung auf die Angehörigen kann z. B. mittels Lebensqualitätsskalen gemessen werden.

Die Literatur zur Alzheimer Demenz ist uneinheitlich bezüglich der Kriterien für die Wirksamkeit von Interventionen. Von Zulassungsbehörden werden oft die kognitive Leistungsfähigkeit, z. B. gemessen mit der ADAS-cog-Skala, der klinische Gesamteindruck und die alltagspraktischen Fähigkeiten als Wirksamkeitskriterien gefordert. Wünschenswert ist eine Änderung auf mehr als einer Erfassungsebene. In welchem Ausmaß eine Intervention ein solches Kriterium beeinflussen sollte, damit von einem Nutzen gesprochen werden kann, ist allerdings umstritten.

Idealerweise sollten alle Instrumente zur Wirksamkeitsbeurteilung akzeptable psychometrische Eigenschaften wie Reliabilität, Validität (Konstruktvalidität) und Veränderungssensitivität haben.

2 ZIELE DER UNTERSUCHUNG

Die Ziele der Untersuchung ergaben sich aus der Auftragsformulierung des Gemeinsamen Bundesausschusses (G-BA) sowie der Verfügbarkeit und dem Zulassungsstatus der Cholinesterasehemmer in Deutschland.

Ziele der vorliegenden Untersuchung waren

- die Nutzenbewertung einer langfristigen Behandlung mit einem Cholinesterasehemmer bei der Alzheimer Demenz im Vergleich zu einer Placebo-Gabe hinsichtlich patientenrelevanter Therapieziele

sowie

- die Nutzenbewertung einer langfristigen Behandlung mit einem Cholinesterasehemmer bei der Alzheimer Demenz im Vergleich zu einer Behandlung mit einer anderen medikamentösen oder nichtmedikamentösen Therapieoption hinsichtlich patientenrelevanter Therapieziele.

Unter „Cholinesterasehemmer“ sind dabei alle diejenigen Wirkstoffe dieser Wirkstoffgruppe zu verstehen, die in Deutschland zur Behandlung der Alzheimer Demenz zugelassen und erhältlich sind. Dies sind die Wirkstoffe:

- Donepezil
- Galantamin
- Rivastigmin

Die Nutzenbewertung beruht auf der Gegenüberstellung und Abwägung von erwünschten und unerwünschten Effekten der jeweiligen Substanzen (Nutzen-Schaden-Abwägung).

3 PROJEKTABLAUF

3.1 Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 22.02.2005 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Nutzenbewertung verschiedener zur Therapie einer Alzheimer Demenz zugelassener Medikamente beauftragt. Darunter fällt auch die Nutzenbewertung von Cholinesterasehemmern. Die Auftragskonkretisierung erfolgte mit dem G-BA vorab auf Grundlage eines Auftragsentwurfs am 02.02.2005.

In die Bearbeitung des Projekts wurden externe Sachverständige eingebunden, die an der Erstellung des Berichtsplans, an der Informationsbeschaffung und -bewertung sowie an der Erstellung des Vorberichts beteiligt waren.

Um in die Definition der patientenrelevanten Endpunkte für die Nutzenbewertung auch die Sicht von Angehörigen mit einzubeziehen, fand ein Treffen mit Vertretern der Deutschen Alzheimer-Gesellschaft statt. Bei den Vertretern der Deutschen Alzheimer-Gesellschaft handelt es sich um Angehörige von an Demenz Erkrankten. Eine direkte Befragung von Betroffenen erfolgte nicht. Die abschließende Beratung fand in der internen Projektgruppe im IQWiG statt.

Der Berichtsplan wurde am 02.06.2005 fertiggestellt und anschließend nach Weitergabe an den G-BA im Internet veröffentlicht. Am 19.06.2006 erfolgte die Veröffentlichung eines Amendments zum Berichtsplan in der Version vom 12.06.2006. Die vorläufige Bewertung, der Vorbericht, wurde am 08.09.2006 im Internet publiziert. Zu diesem Vorbericht konnten bis einschließlich 06.10.2006 Stellungnahmen von allen interessierten Personen, Institutionen und Gesellschaften, einschließlich Privatpersonen, Fachgesellschaften und Industrieunternehmen, abgegeben werden. Unklare Aspekte der schriftlichen Stellungnahmen wurden dann am 14.11.2006 in einer wissenschaftlichen Erörterung hinsichtlich ihrer Relevanz für den Abschlussbericht mit den Stellungnehmenden diskutiert. Ein Protokoll dieser wissenschaftlichen Erörterung liegt in Anhang H vor. Der Vorbericht wurde zusätzlich einem externen Review unterzogen.

Im Anschluss an die wissenschaftliche Erörterung erstellte das IQWiG den vorliegenden Abschlussbericht, der 8 Wochen nach Übermittlung an den G-BA im Internet veröffentlicht wird.

3.2 Zusammenfassung der Änderungen im Vergleich zum Vorbericht

Durch die Anhörung ergaben sich im Abschlussbericht folgende Änderungen im Vergleich zum Vorbericht:

- modifizierte Bewertung der Qualität von Studien aufgrund von zusätzlichen, von den Herstellern zur Verfügung gestellten Informationen sowie Korrekturen,
- Ergänzung einzelner Ergebnisse aufgrund von zusätzlichen, von den Herstellern zur Verfügung gestellten Informationen,
- Ergänzung einer meta-analytischen Zusammenfassung der Ergebnisse zur Lebensqualität der (betreuenden) Angehörigen für die Studienergebnisse zu Galantamin.

4 METHODEN

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

4.1.1 Population

Aufgrund der Epidemiologie der Alzheimer Demenz und des Zulassungsstatus der Cholinesterasehemmer sollten Patienten mit einer leichten oder mittelschweren Alzheimer Demenz und Patienten mit einer Mischform aus Alzheimer Demenz und einer vaskulären Schädigung des Gehirns in die Untersuchung einbezogen werden. Diese Definition erlaubte z. B. auch die gleichzeitige Diagnose einer vaskulären Demenz bei vorliegender Alzheimer Demenz. Die Diagnosesicherung musste anhand allgemein akzeptierter Kriterien (z. B. ICD-9, ICD-10, DSM-III-R, DSM-IV oder NINCDS-ADRDA), wie in der einschlägigen EMEA-Publikation beschrieben, erfolgen [42]. Für die Bestimmung des Schweregrades wurde die jeweilige Bezeichnung in der Studie herangezogen. Die verwendeten Einteilungen waren nicht zwangsläufig konsistent mit anderen Klassifikationsschemata und auch zwischen den Studien gab es geringe Abweichungen. Grob lässt sich klassifizieren, dass für die Bezeichnung „mittelschwere Demenz“ meist ein Punktwert im MMSE von mindestens 10 gefordert wurde und für die Bezeichnung „leichte Demenz“ mindestens 20 Punkte (bis maximal 26).

Studien, bei denen ausschließlich Patienten mit einer leichten kognitiven Störung (Mild Cognitive Impairment, MCI), einer vaskulären Demenz, einer Demenz aufgrund eines Morbus Parkinson, einer Demenz aufgrund einer Lewy Body Disease, einer Demenz aufgrund einer Creutzfeldt-Jakob-Erkrankung oder einer anderen seltenen Demenzform eingeschlossen wurden, wurden nicht berücksichtigt.

4.1.2 Interventionen und Vergleichsbehandlungen

Die berücksichtigten Interventionen waren die drei Cholinesterasehemmer Donepezil, Galantamin und Rivastigmin in jeglicher Darreichungsform.

Als Vergleichsinterventionen wurden eine Placebo-Behandlung und jegliche andere medikamentöse (einschließlich eines anderen Cholinesterasehemmers als des in der jeweiligen Interventionsgruppe verwendeten) und nichtmedikamentöse Therapieoptionen bei der Alzheimer Demenz betrachtet.

4.1.3 Zielgrößen

Als Zielgrößen für die Untersuchung werden Parameter verwendet, die eine Beurteilung folgender patientenrelevanter Therapieziele ermöglichen:

- Besserung bzw. Verhinderung der Beeinträchtigung der Aktivitäten des täglichen Lebens
- Besserung bzw. Normalisierung von begleitenden psychopathologischen Symptomen (z. B. Depression, Schlaf-Wach-Umkehr, Wahnhaftigkeit, Agitiertheit)
- Besserung bzw. Erhalt der kognitiven Leistungsfähigkeit
- Besserung bzw. Erhalt der krankheitsbezogenen Lebensqualität
- Vermeidung der Notwendigkeit einer vollstationären Pflege (Institutionalisierung)
- Reduktion der Mortalität
- Reduktion der (therapieassoziierten) unerwünschten Ereignisse

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

- Besserung bzw. Erhalt der Lebensqualität der (betreuenden) Angehörigen^a
- Reduktion der Höhe des Betreuungsaufwands durch eine oder mehrere betreuende Person/en oder Institution/en

Als ergänzende Informationen werden zudem Ergebnisse berichtet, die sich auf „die Besserung bzw. den Erhalt des klinischen Krankheitsstadiums gemäß dem klinischen Eindruck“ beziehen.

Ergebnisse zu angehörigenrelevanten Therapiezielen und Ergebnisse, die sich auf „die Besserung bzw. den Erhalt des klinischen Krankheitsstadiums gemäß dem klinischen Eindruck“ beziehen, fließen nicht primär in die Nutzenbewertung ein. Allerdings lassen sich gegebenenfalls Aussagen zum Zusammenhang zwischen Veränderungen dieser Größen und Veränderungen von patientenrelevanten Zielgrößen treffen.

^a Mit dem Begriff „Angehörige“ sind im Folgenden in aller Regel die direkten Angehörigen, aber auch andere betreuende Personen gemeint (im Sinne des englischen Begriffes „caregiver“).

4.1.4 Studientypen

Randomisierte kontrollierte Studien (RCT) liefern für die Bewertung des Nutzens einer medizinischen Intervention die zuverlässigsten Ergebnisse, weil sie, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet sind. Für alle unter 4.1.2 genannten Interventionen und für alle unter 4.1.3 genannten Therapieziele ist eine Evaluation im Rahmen von randomisierten, kontrollierten Studien möglich und praktisch durchführbar. Für den zu erstellenden Bericht wurden deshalb ausschließlich RCT als relevante wissenschaftliche Literatur für die Nutzenbewertung berücksichtigt.

4.1.5 Sonstige Studiencharakteristika

Für die Ersteinstellung von Medikamenten zur Pharmakotherapie kognitiver Störungen wird von der Arzneimittelkommission der Deutschen Ärzteschaft eine Kontrolle nach 12 Wochen empfohlen, um einen Therapieerfolg abschätzen zu können [43]. Von der EMEA wird eine Studiendauer von mindestens 24 Wochen für die Beurteilung der kurzfristigen Wirkungen von Cholinesterasehemmern als sinnvoll erachtet [42]. Um beiden Empfehlungen gerecht zu werden, wurde für den vorliegenden Bericht eine Mindestbeobachtungszeit von 16 Wochen festgelegt, da innerhalb dieses Zeitraumes ein Ansprechen auf die Therapie erwartet und ein mehr als kurzfristig anhaltender Effekt beobachtet werden kann.

4.1.6 Ein-/Ausschlusskriterien

In die Nutzenbewertung wurden alle Studien einbezogen, die alle nachfolgenden Einschlusskriterien und keines der nachfolgenden Ausschlusskriterien erfüllen.

Einschlusskriterien

E1	Patienten mit einer leichten oder mittelschweren Alzheimer Demenz, auch als Mischform mit z. B. vaskulärer Demenz. Diagnosesicherung gemäß EMEA [42] oder anhand allgemein akzeptierter Kriterien (z. B. ICD-9, ICD-10, DSM-III-R, DSM-IV oder NINCDS-ADRDA)
E2	Vergleich eines Cholinesterasehemmers (Donepezil, Galantamin, Rivastigmin) mit Placebo oder einer anderen medikamentösen oder nichtmedikamentösen Intervention (wie unter 4.1.2 beschrieben)
E3	Zielgrößen, die sich aus den in 4.1.3 formulierten Therapiezielen ableiten lassen
E4	Randomisierte kontrollierte Studien (RCT)

Ausschlusskriterien

A 1	Studien mit einer Beobachtungsdauer von weniger als 16 Wochen
A2	Studien, in die ausschließlich Patienten mit einer vaskulären Demenz, einer Demenz aufgrund eines Morbus Parkinson, einer Demenz aufgrund einer Lewy Body Disease, einer Demenz aufgrund einer Creutzfeldt-Jakob-Erkrankung oder einer anderen seltenen Demenzform eingeschlossen wurden
A3	Studien bzw. Publikationen, die lediglich Daten von unkontrollierten Open-label-follow-up-Phasen enthalten
A4	Doppelpublikationen ohne relevante Zusatzinformation
A5	Keine Vollpublikation ¹
1: Als Vollpublikation galt 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 [44] genügt und eine Bewertung der Studie ermöglichte.	

4.2 Informationsbeschaffung

Ziel der Informationsbeschaffung war es, publizierte und nicht publizierte Studien zu identifizieren, die zur Nutzenbewertung von Cholinesterasehemmern bei leichter bis mittelschwerer Alzheimer Demenz wesentliche Informationen liefern.

4.2.1 Literaturrecherche

Die Literaturrecherche nach relevanten, veröffentlichten Studien wurde in folgenden Quellen durchgeführt:

- bibliographische Datenbanken (MEDLINE; EMBASE; Cochrane Central Register of Controlled Trials [CENTRAL]; CHID via ADEAR [Combined Health Information Database via Alzheimer's Disease Education & Referral Center], www.alzheimers.org)
- Literaturverzeichnisse relevanter Sekundärpublikationen (systematische Übersichten, HTA-Berichte, Meta-Analysen)

Die Suchstrategien für die Suche in bibliographischen Datenbanken finden sich in Anhang A. Die Suche erfolgte in drei Schritten:

- Erstrecherche am 13.04.2005 (MEDLINE und EMBASE), 14.04.2005 (Cochrane-Datenbanken) und 25.04.2005 (CHID via ADEAR)

- Erste Nachrecherche am 03.11.2005 für den Zeitraum 4/2005 bis 10/2005 (alle Datenbanken)
- Zweite Nachrecherche am 12.06.2006 mittels einer überarbeiteten Suchstrategie in MEDLINE, EMBASE und Cochrane-Datenbanken. Die Datenbank CHID war zum Zeitpunkt der Nachrecherche nicht mehr verfügbar.

Die Suche nach relevanten Sekundärpublikationen (systematische Übersichten, HTA-Berichte, Meta-Analysen) erfolgte in den Datenbanken Medline und EMBASE parallel zur Suche nach relevanter Primärliteratur. Zusätzlich wurde eine Suche in den spezialisierten Datenbanken Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE) und HTA Database (HTA) vorgenommen (Erstrecherche und Nachrecherchen wie oben).

4.2.2 Suche nach weiteren publizierten und nicht publizierten Studien

Zur Identifizierung weiterer veröffentlichter und unveröffentlichter Studien wurde eine Anfrage an die Hersteller von Cholinesterasehemmern in Deutschland gestellt. Folgende Unternehmen wurden kontaktiert:

- Eisai GmbH, Frankfurt (Donepezil, Aricept®)
- Novartis Pharma GmbH, Nürnberg (Rivastigmin, Exelon®)
- Janssen-Cilag GmbH, Neuss (Galantamin, Reminyl®)

4.2.3 Identifizierung relevanter Studien

Titel- und Abstract-Screening der Ergebnisse der Recherche in bibliographischen Datenbanken

Die durch die Suche in bibliographischen Datenbanken identifizierten Zitate wurden anhand ihres Titels und, sofern vorhanden, Abstracts von zwei Reviewern unabhängig voneinander hinsichtlich ihrer Relevanz bewertet. Publikationen, die beide Reviewer als potenziell relevant erachteten, wurden anhand ihres Volltextes auf Relevanz geprüft. Zitate, die nur ein Reviewer als potenziell relevant einstufte, wurden von beiden Reviewern nochmals gesichtet und anschließend nach Diskussion entweder als irrelevant bezeichnet oder ebenfalls anhand ihres Volltextes auf Relevanz geprüft.

Überprüfung potenziell relevanter Volltexte

Die Überprüfung auf Relevanz anhand des Volltextes erfolgte wiederum von zwei Reviewern unabhängig voneinander. Als relevant wurden nach diesem Schritt folgende Studien bezeichnet:

- Studien, die von beiden Reviewern als relevant erachtet wurden
- Studien, die zunächst nur von einem der beiden Reviewer, aber nach anschließender Diskussion von beiden Reviewern als relevant erachtet wurden

Suche in Literaturverzeichnissen von Sekundärpublikationen

Die Literaturverzeichnisse relevanter Sekundärpublikationen wurden nach weiteren Primärpublikationen durchsucht. Die Volltexte der aus den Übersichtsarbeiten identifizierten Publikationen wurden von zwei Reviewern, wie oben beschrieben, bezüglich ihrer Relevanz bewertet.

Mehrfachveröffentlichungen

Die einzuschließenden Publikationen wurden daraufhin durchgesehen, ob sie Mehrfachveröffentlichungen ein und derselben Studie darstellten. Bei Mehrfachveröffentlichungen wurden alle Publikationen den entsprechenden Studien zugeordnet und alle Daten, die auswertbar waren und Informationen zu den in Abschnitt 4.1.3 aufgeführten Zielkriterien bereitstellten, herangezogen.

4.2.4 Suche nach zusätzlichen Informationen zu relevanten Studien

Die in der Literaturrecherche identifizierten Studien wurden ggf. durch weitere relevante Studien aus den in 4.2.2 beschriebenen Dokumenten ergänzt. In den nach 4.2.2 gefundenen Dokumenten wurde außerdem nach ergänzender Information zu bereits in der Literaturrecherche identifizierten Studien gesucht.

Darüber hinaus wurden Autoren von Publikationen und Sponsoren von Studien kontaktiert, wenn im Lauf der Bewertung Fragen zu eingeschlossenen Studien aufgeworfen wurden, die aus den Publikationen nicht beantwortet werden konnten.

4.2.5 Informationen aus der Anhörung zum Vorbericht

Im Anschluss an die Veröffentlichung des Vorberichts erfolgte eine schriftliche Anhörung mittels schriftlicher Stellungnahmen, die sich u. a. auch auf die Vollständigkeit der Informationsbeschaffung beziehen konnten. Relevante Informationen aus dieser Anhörung konnten in die Nutzenbewertung einfließen.

4.3 Informationsbewertung

Die Bewertung der eingeschlossenen Studien erfolgte anhand der zur Verfügung stehenden Informationen und hing damit stark von der Qualität der jeweiligen Publikation und weiterer Informationsquellen ab.

Die Bewertung erfolgte in drei Schritten:

- Extraktion der Daten
- Bewertung der Datenkonsistenz innerhalb der Publikation und zwischen den Informationsquellen (z. B. Publikation und Angaben in Zulassungsdokumenten)
- Bewertung der Studien- und Publikationsqualität

Datenextraktion

Die Extraktion der Daten publizierter Studien wurde anhand von standardisierten Datenextraktionsbögen vorgenommen. Zwei Gutachter führten die Datenextraktion unabhängig voneinander durch. Im Anschluss daran erfolgte ein Abgleich der extrahierten Daten und die Erarbeitung eines konsentierten Extraktionsbogens für jede Studie.

Angaben zu folgenden Aspekten der Studienqualität wurden systematisch extrahiert:

- Randomisierungsprozess und Verdeckung der Gruppenzuweisung (Allocation Concealment)

Der Randomisierungsprozess erhielt die Bezeichnung „unklar“, wenn lediglich der Begriff „randomisiert“ verwendet wurde, und „inadäquat“, wenn das Verfahren in der Publikation beschrieben, jedoch als nicht angemessen anzusehen war. Bei einer Bewertung als „adäquat“ lagen genauere Angaben zu einem angemessenen Vorgehen vor.

Bei der Bewertung der Gruppenzuteilung bedeutet „unklar“, dass keine Angaben dazu gemacht wurden, inwieweit die Gruppenzuteilung verdeckt erfolgte. Als „adäquat“ wurde sie bezeichnet, wenn ein angemessenes Vorgehen beschrieben wurde, und als „inadäquat“, wenn eine Vorgehensweise beschrieben wurde, die eindeutig als nicht adäquat zu bewerten war. Bei tatsächlich doppelblinden Studien kommt allerdings einer möglichst genauen Beschreibung der Verdeckung der Gruppenzuteilung vermutlich weniger zentrale Bedeutung zu als bei offenen Studien, wenngleich diesbezügliche Erläuterungen in jedem Fall wünschenswert und auch leicht realisierbar wären.

- Verblindung der Behandelnden, der Patienten und der Endpunkterhebung

Da der ganz überwiegende Anteil der Studien doppelblind durchgeführt wurde, wurde hier insbesondere jeweils dargestellt, ob in der Publikation Informationen dazu vorhanden waren, ob diejenigen Personen, die die Endpunkte bei den Patienten erhoben hatten, verblindet waren, auch und besonders im Hinblick auf andere Studienergebnisse und auf das Auftreten von Nebenwirkungen bei diesen Patienten.

- Fallzahlplanung

Die Fallzahlplanung wurde dann als „adäquat“ beurteilt, wenn das Zielkriterium, die Größe des erwarteten Effektes, die Teststärke, das Signifikanzniveau und die ermittelte Fallzahl benannt waren. Als wünschenswert wurden darüber hinaus Informationen zur erwarteten Variabilität in der Stichprobe, zum verwendeten statistischen Testverfahren sowie zur Ein- oder Zweiseitigkeit des Testes angesehen. Wenn die geforderten Angaben vorlagen, jedoch bei der Durchführung der Studie vom geplanten Vorgehen auf relevante Weise abgewichen wurde (bspw. die tatsächliche Fallzahl deutlich kleiner war als die geplante), dann wurde dies mit „(ja)“ gekennzeichnet. Wenn in der Publikation keine Angaben zu einer Fallzahlplanung gemacht wurden, wurde davon ausgegangen, dass keine durchgeführt worden war, und dies mit „nein“ beurteilt; bei Fehlen einiger Aspekte erfolgte die Bewertung „unklar“.

- Beschreibung der Studienabbrecher

Hier wurde jeweils beschrieben, ob sowohl die Anzahl der Studienabbrecher in den verschiedenen Gruppen als auch die jeweiligen Gründe für den Abbruch aus der Publikation hervorgingen.

- Abweichung vom ITT-Prinzip

Hier wurde beurteilt, ob eine wesentliche Verletzung des zugrunde liegenden Prinzips vorlag. Wurde ein kleiner Anteil von Patienten aus der Auswertung ausgeschlossen, so wurde dies im Rahmen dieses Berichtes nicht als wesentliche Verletzung angesehen. In den Fällen, wo eine „relevante“ Verletzung des ITT-Prinzips konstatiert wurde, lag die Rate der in den primären Auswertungen unberücksichtigt gebliebenen Patienten bei (gerundet) mindestens 11 % oder der Unterschied in den Nichtberücksichtigungsraten zwischen den verschiedenen Behandlungsgruppen betrug (gerundet) mindestens 5 Prozentpunkte. In diesen Fällen wurden die Anteile der nicht in die Auswertungen eingehenden Patienten (ITT-Auswertungen, soweit angegeben) dokumentiert, so dass die dargestellten Entscheidungen (relevante Abweichung vom ITT-Prinzip: ja/nein) jeweils nachvollziehbar sind. Die Feststellung einer relevanten Verletzung des ITT-Prinzips führte zur Abwertung der Studien- bzw. Publikationsqualität („grober Mangel“).

Überprüfung der Datenkonsistenz

An die Datenextraktion schloss sich ggf. ein Abgleich mit Informationen an, die durch die in den Abschnitten 4.2.2 und 4.2.4 beschriebene weiterführende Suche zu publizierten Studien gewonnen wurden. Sofern sich hieraus (aber auch aus multiplen Angaben zu einem Aspekt innerhalb der Publikation selbst) Diskrepanzen ergaben, die auf die Ergebnisse bzw. die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wurde dies an den entsprechenden Stellen im Ergebnisteil dargestellt.

Bewertung der Studien- und Publikationsqualität

Abschließend wurde unter Berücksichtigung der oben genannten und anderer, im Einzelfall dann jeweils dargestellter, Aspekte die Bewertung der Studien- und Publikationsqualität global, mittels eines vier Ausprägungen umfassenden Merkmals („biometrische Qualität“), durchgeführt. Mögliche Ausprägungen waren:

- keine erkennbaren Mängel
- leichte Mängel
- grobe Mängel
- unklar

Die Ausprägungen wurden vorab wie folgt definiert: „Leichte Mängel“ liegen dann vor, wenn davon ausgegangen wird, dass deren Behebung die Ergebnisse und damit die Gesamtaussage der Studie nicht wesentlich beeinflussen wird. Bei „grobe Mängel“ wäre die Gesamtaussage grundsätzlich infrage zu stellen, wenn die Mängel behoben würden. Eine Studie wird mit „keine erkennbaren Mängel“ beschrieben, wenn sie höchstens unwesentliche Mängel aufweist. „Unklar“ bedeutet, dass anhand der vorliegenden Unterlagen keine eindeutige Aussage über die biometrische Qualität der Studie getroffen werden kann.

Da wie oben beschrieben die Bewertung der Studienqualität unmittelbar durch die Qualität und Konsistenz der zur Verfügung stehenden Informationen beeinflusst wird, ist die Angabe „grobe Mängel“ nicht zwangsläufig eine Beschreibung der Qualität der Studie selber, sondern gegebenenfalls auch durch die Qualität der zugrunde liegenden Publikation(en) bedingt.

Diese Qualitätsklassifizierung sollte ggf. einer Sensitivitätsanalyse im Rahmen einer Meta-Analyse dienen.

4.4 Informationssynthese und -analyse

4.4.1 Charakterisierung der Studien

Im Bericht wurden die Studien anhand von Designcharakteristika (Studiendesign, Studiendauer, Ort und Zeitraum der Studiendurchführung, Zahl der randomisierten und ausgewerteten Patienten, relevante Zielkriterien) beschrieben. Darüber hinaus wurden die Testintervention(en) und die Vergleichsbehandlung(en) dargestellt. Die Studienpopulationen wurden durch demographische Daten (Alter, Geschlecht), durch Charakteristika der demenziellen Erkrankung (Alzheimer Demenz oder Mischformen, Krankheitsschwere) und durch die Zahl der Studienabbrecher beschrieben.

4.4.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien wurden nach Therapiezielen und Zielgrößen geordnet gegenübergestellt und vergleichend beschrieben. Dabei wurden folgende Zielgrößen berücksichtigt:

Therapieziel	Zielgröße
<i>Patientenrelevante Therapieziele:</i>	
Besserung bzw. Verhinderung der Beeinträchtigung der Aktivitäten des täglichen Lebens	<ul style="list-style-type: none"> - Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory (ADCS-ADL) - Alzheimer's Disease Functional Assessment and Change Scale (ADFACS) - Blessed-Roth Dementia Scale (BRDS) - Bristol Activities of Daily Living Scale (BADLS) - Caregiver-rated Modified Crichton Scale (CMCS) - Disability Assessment for Dementia Scale (DAD) - Instrumental Activities of Daily Living (IADL) - Interview for Deterioration in Daily Living in Dementia (IDDD) - Physical Self-Maintenance Scale (PSMS) - Progressive Deterioration Scale (PDS)
Besserung bzw. Normalisierung von begleitenden psychopathologischen Symptomen	<ul style="list-style-type: none"> - Neuropsychiatric Inventory (NPI) - Nurses' Observation Scale for Geriatric Patients (NOSGER) - Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD)

Therapieziel	Zielgröße
<i>Patientenrelevante Therapieziele:</i>	
Besserung bzw. Erhalt der kognitiven Leistungsfähigkeit	<ul style="list-style-type: none"> - Alzheimer's Disease Assessment Scale – cognitive subscale (ADAS-cog) - Mini Mental State Examination (MMSE) - Severe Impairment Battery (SIB)
Besserung bzw. Erhalt der krankheitsbezogenen Lebensqualität	<ul style="list-style-type: none"> - Quality of Life Scale (QoL Scale)
Vermeidung der Notwendigkeit einer vollstationären Pflege (Institutionalisierung)	<ul style="list-style-type: none"> - Zeitdauer bis zur Übersiedlung in ein Pflegeheim - Anteil von Patienten mit Übersiedlung in ein Pflegeheim nach bestimmter Behandlungsdauer
Reduktion der Mortalität	<ul style="list-style-type: none"> - Anteil verstorbener Patienten innerhalb eines bestimmten Zeitraums
Reduktion der (therapieassoziierten) unerwünschten Ereignisse	<ul style="list-style-type: none"> - Gesamtrate unerwünschter Ereignisse - Schwerwiegende unerwünschte Ereignisse - Studienabbruch wegen unerwünschter Ereignisse - Häufige unerwünschte Ereignisse
<i>Angehörigenrelevante Therapieziele:</i>	
Besserung bzw. Erhalt der Lebensqualität der (betreuenden) Angehörigen	<ul style="list-style-type: none"> - Caregiving Burden Scale (CBS) - Caregiver Stress Scale (CSS) - Neuropsychiatric Inventory Caregiver Distress Scale (NPI-D) - Screen for Caregiver Burden (SCGB)
Reduktion der Höhe des Betreuungsaufwands durch eine oder mehrere betreuende Person/en oder Institution/en	<ul style="list-style-type: none"> - Allocation of Caregiver Time Survey (ACTS) - Caregiver Activity Survey (CAS) - Resource Utilization in Dementia (RUD)
<i>Ergänzende Information zu:</i>	
Besserung bzw. Erhalt des klinischen Krankheitsstadiums	<ul style="list-style-type: none"> - Clinical Dementia Rating Scale (CDR) - Clinical Global Impression of Change Scale (CGIC) - Clinician's Interview-Based Impression of Change (mit Input des Betreuenden: CIBIC-plus) - Global Deterioration Scale (GDS) - Gottfries-Bråne-Steen-Skala (GBS)

Die Tabelle in Anhang D gibt einen Überblick über alle in den eingeschlossenen Studien verwendeten Zielgrößen, die sich einem der Therapieziele zuordnen lassen.

Wenn in den Studien zu einem Therapieziel (z. B. kognitive Fähigkeiten) mehrere Skalen berichtet wurden, wird in der folgenden Nutzenbewertung im Allgemeinen jeweils nur eine (vorzugsweise diejenige, die am häufigsten verwendet wurde) dargestellt und ausgewertet. In den Tabellen zur Beschreibung der eingeschlossenen Studien sind deshalb (zusätzlich zum jeweiligen primären Zielkriterium der Studie) nur diejenigen Zielkriterien aufgelistet, die zur Nutzenbewertung im vorliegenden Bericht herangezogen wurden.

Eine kurze Erläuterung aller ausgewerteten Skalen findet sich in Anhang E.

4.4.3 Meta-Analyse

Daten zu einer Zielgröße sollten in einer Meta-Analyse quantitativ zusammengefasst werden, sofern die Studienlage es inhaltlich und methodisch als sinnvoll erscheinen ließ. So erfolgte eine Meta-Analyse, wenn mindestens drei Studien mit validen statistischen Maßzahlen vorlagen oder zumindest hinreichend genau berechenbar oder anhand der Abbildungen einschätzbar waren. Abweichend davon wurden zur genaueren Beschreibung von Daten zu unerwünschten Ereignissen auch Meta-Analysen berechnet, wenn mindestens zwei Studien wie oben beschrieben vorlagen.

Im ersten Schritt wurde ein Modell mit festen Effekten verwendet. Bei Hinweisen auf eine mögliche Heterogenität der Einzelstudienresultate ($I^2 > 50\%$ bzw. $p < 0,2$ für Heterogenitätstest) erfolgte in einem zweiten Schritt die Berechnung anhand eines Modells mit zufälligen Effekten und es wurden diese Ergebnisse dargestellt. Relevante Abweichungen von den Ergebnissen bzgl. des Modells mit festen Effekten wurden im Text diskutiert.

Für die statistische Auswertung fanden primär die Ergebnisse aus Intention-To-Treat-Analysen, so wie sie in den Publikationen beschrieben waren, Verwendung. Bei Studien mit mehreren Interventionsgruppen unterschiedlicher Dosierung oder Einnahmehäufigkeit wurde beim Vergleich mit Placebo für die meta-analytische Zusammenfassung – soweit nicht anders gekennzeichnet – die Gruppe mit der höchsten Dosierung bzw. häufigeren Einnahme gewählt.

Für kontinuierliche Variablen wurde die standardisierte Mittelwertdifferenz (Cohen's d) als Effektmaß eingesetzt, um Scores verschiedener Skalen vergleichen zu können. Bei binären Variablen wurden Meta-Analysen anhand der absoluten Risikodifferenz sowie des Odds Ratios durchgeführt. In diesem Bericht werden allerdings nur die Ergebnisse für das Odds Ratio dargestellt, da diese im Vergleich zur absoluten Risikodifferenz wesentlich homogener waren.

Es ergab sich in keiner Meta-Analyse für binäre Daten der Fall, dass in beiden Behandlungsgruppen kein Ereignis beobachtet wurde. Trat in nur einer der

Behandlungsgruppen das Ereignis bei keinem Patienten ein, so wurde in der zugrunde liegenden Vierfeldertafel zu jeder Zellenhäufigkeit ein Korrekturfaktor von 0,5 addiert.

4.4.4 Sensitivitätsanalyse

Innerhalb der Meta-Analysen waren laut Berichtsplan Sensitivitätsanalysen bezüglich folgender Faktoren geplant:

- biometrische Qualitätsbewertung (siehe Abschnitt 4.3),
- in den Publikationen beschriebene Per-Protokoll-Auswertungen (versus ITT-Auswertungen), sofern möglich, und
- ein (statistisches) Modell mit festen Effekten versus ein Modell mit zufälligen Effekten

4.4.5 Subgruppenanalyse

Subgruppenanalysen waren – soweit durchführbar – für folgende Merkmale vorgesehen:

- Geschlecht
- Alter
- Schweregrad der Demenz
- Dosierungen der Cholinesterasehemmer
- Vorliegen unterschiedlicher Begleiterkrankungen
- Charakteristika, die für eine bedeutsame Heterogenität verantwortlich waren

Für die Merkmale Geschlecht, Alter, Schweregrad der Demenz und Vorliegen unterschiedlicher Begleiterkrankungen fanden sich nie bzw. fast nie differenzierte Analysen, so dass Subgruppenanalysen für diese Merkmale nicht möglich waren.

Für den Fall einer auffälligen Heterogenität (s. o.) in einer Meta-Analyse sollte gegebenenfalls eine Subgruppenanalyse für Merkmale, die möglicherweise die Heterogenität erzeugen, durchgeführt werden. In den Fällen, in denen in der vorliegenden Nutzenbewertung Heterogenität beobachtet wurde, fanden sich jedoch keine erklärenden Merkmale bzw. war die Zahl der Studien zu klein, so dass Subgruppenanalysen innerhalb dieses Studienpools nicht sinnvoll waren.

4.5 Änderungen im Vergleich zum Berichtsplan

4.5.1 Änderungen während der Erstellung des Vorberichts

Im Laufe der Bearbeitung des Projekts ergaben sich Änderungen und Ergänzungen des Vorgehens bei der Nutzenbewertung im Vergleich zu der im Berichtsplan einschließlich des am 19.06.2006 veröffentlichten Amendments dargestellten Methodik. Diese Änderungen betrafen im Wesentlichen folgende Punkte:

Inhaltliche Änderungen im Vergleich zum vorab geplanten Vorgehen

- Es erfolgte keine Kontaktaufnahme mit der Cochrane Dementia and Cognitive Improvement Group zur Identifikation weiterer Studien.

Änderungen ohne wesentliche inhaltliche Konsequenz (Spezifikationen des vorab geplanten Vorgehens)

- Die Datenquellen zur Suche nach systematischen Übersichten, Meta-Analysen und HTA-Berichten wurden explizit angegeben.
- Als Vergleichsinterventionen wurden eine Placebo-Behandlung und jegliche andere medikamentöse (einschließlich eines anderen Cholinesterasehemmers als des in der jeweiligen Interventionsgruppe verwendeten) und nichtmedikamentöse Therapieoptionen bei der Alzheimer Demenz betrachtet.
- Das Ausschlusskriterium „A5: Reine Abstract-Publikationen“ wurde dahingehend spezifiziert, dass ausschließlich Vollpublikationen berücksichtigt wurden („A5: Keine Vollpublikation“). Als Vollpublikation galt 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 [44] genügt und eine Bewertung der Studie ermöglichte.

4.5.2 Änderungen nach Veröffentlichung des Vorberichts

Die Stellungnahmen zum Vorbericht führten nicht zu einer Änderung der im Berichtsplan spezifizierten Methodik der Nutzenbewertung.

5 ERGEBNISSE

5.1 Ergebnisse der Informationsbeschaffung

5.1.1 Ergebnis der Literaturrecherche

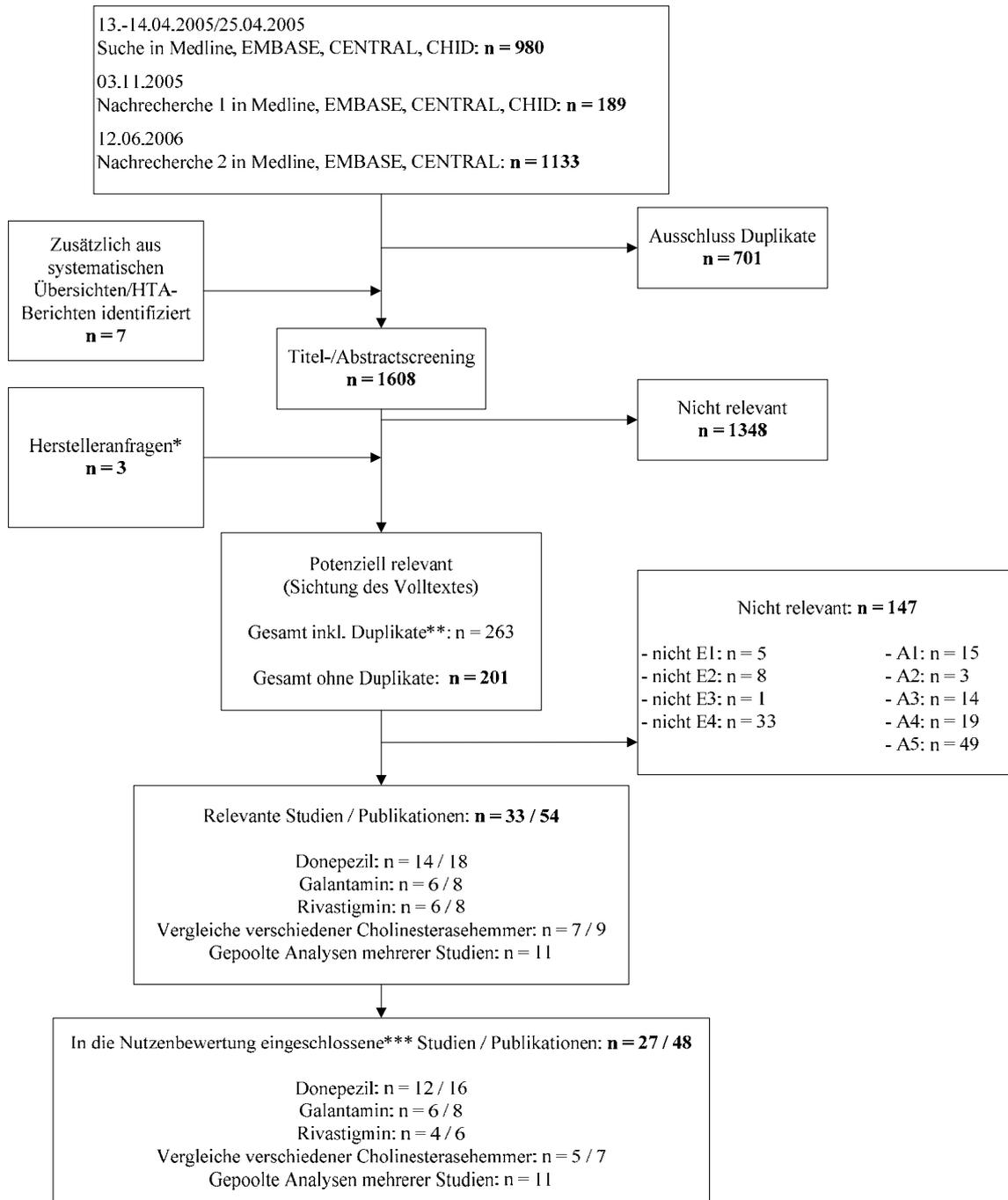
Das Ergebnis der systematischen Literaturrecherche nach Studien in den bibliographischen Datenbanken und des Literaturscreenings gemäß Ein-/Ausschlusskriterien ist in Abbildung 1 dargestellt.

Nach Ausschluss von 701 Duplikaten ergaben sich durch die Primärrecherche, die erste und die zweite Nachrecherche sowie die Suche in systematischen Übersichten und HTA-Berichten insgesamt 1608 Treffer. In der Datenbank CHID wurde mittels der Substanznamen recherchiert. Die in den Datenbanken MEDLINE, EMBASE und CENTRAL verwendeten Suchstrategien sind in Anhang A dargestellt. Im Rahmen der zweiten Nachrecherche zeigte sich, dass die verwendete Suchstrategie u. a. fehlerhafte Verknüpfungen und Feldbezeichnungen enthielt, so dass eine wesentliche Überarbeitung notwendig war. Deshalb wurde die zweite Nachrecherche mittels der überarbeiteten Suchstrategie (siehe Anhang A) für die Datenbanken MEDLINE, EMBASE und CENTRAL ohne Eingrenzung des Zeitraums durchgeführt. Die Datenbank CHID war zum Zeitpunkt der Nachrecherche nicht mehr verfügbar. Von den 1608 Treffern wurden insgesamt 1348 als nicht relevant beurteilt. Die Anfrage bei den Herstellern ergab Hinweise auf drei weitere potenziell relevante Studien.

Somit wurden insgesamt 263 potenziell relevante Referenzen ermittelt. Da die Primärrecherche und die zweite Nachrecherche einen gemeinsamen Zeitraum abdeckten, wurde eine Reihe von Referenzen durch beide Rechenschritte identifiziert. Abzüglich dieser Duplikate wurden 201 Publikationen im Volltext gesichtet, von denen 147 als nicht relevant ausgeschlossen wurden. Die Zitate der im Volltext gesichteten, nicht relevanten Publikationen finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B.

Aus den gesichteten Arbeiten wurden 54 Publikationen zu 33 Studien als relevant beurteilt. Von diesen konnten 48 Publikationen zu 27 Studien in die Nutzenbewertung eingeschlossen werden.

22 Studien waren Placebo-kontrolliert (Donepezil 12, Galantamin 6, Rivastigmin 4), 5 Studien beinhalteten direkte Vergleiche verschiedener Cholinesterasehemmer. 11 der eingeschlossenen Publikationen enthielten gepoolte Analysen mehrerer Studien.



* siehe Kapitel 5.1.2 „Anfrage an Hersteller“

** Duplikate ergaben sich dadurch, dass die zweite Nachrecherche den Zeitraum der Primärrecherche mit einschloss. Darüber hinaus wurde durch die Herstelleranfrage eine Studie identifiziert, die bereits in der bibliographischen Literaturrecherche identifiziert worden war (Tai 2000).

*** Für sechs Studien waren die Informationen für die Nutzenbewertung nicht ausreichend (s. Abschnitt 5.1.4).

Abbildung 1: Bibliographische Literaturrecherche und sonstiges Literaturscreening, endgültiger Studienpool für die Nutzenbewertung

5.1.2 Anfrage an Hersteller

Im Mai 2005 wurden bei der EISAI GmbH, dem Hersteller von Donepezil, bei der Janssen-Cilag GmbH, dem Hersteller von Galantamin, und bei der Novartis Pharma GmbH, dem Hersteller von Rivastigmin, Übersichten zu Studien mit Cholinesterasehemmern angefragt. Alle Unternehmen stellten daraufhin Listen von Studien zur Verfügung. Zwischenzeitlich ergab sich die Notwendigkeit, bezüglich der Übermittlung von Studieninformationen von Seiten pharmazeutischer Unternehmen an das IQWiG, zu einem geregelten, einheitlichen Vorgehen zu kommen. In den folgenden Monaten wurde eine entsprechende generelle Vereinbarung erarbeitet, die vor der Übersendung von Unterlagen zwischen dem IQWiG und den Unternehmen abgeschlossen werden soll. Die Vereinbarung für die Übermittlung und Nutzung von Daten wurde von allen Unternehmen unterzeichnet.

EISAI GmbH

Nach Unterzeichnung der Vereinbarung zur Übermittlung von Unterlagen stellte EISAI auf Anfrage des IQWiG den Expert Report des Zulassungsdossiers für Donepezil zur Verfügung und bestätigte die Vollständigkeit der eingereichten Studienlisten. Ein Review des Expert Reports und der Listen ergab, dass alle relevanten Studien veröffentlicht und bereits in der Literaturrecherche identifiziert waren.

Janssen-Cilag GmbH

Nach Abschluss der Vereinbarung zur Übermittlung von Unterlagen konnten die bereits im Vorfeld von Janssen-Cilag zur Verfügung gestellten Dokumente für die Nutzenbewertung des IQWiG genutzt werden. Dabei handelte es sich um einen Clinical Expert Report und einen Clinical Overview aus den Zulassungsverfahren von Galantamin, um Studienlisten sowie um Unterlagen, die Janssen-Cilag dem National Institute for Health and Clinical Excellence (NICE) in Großbritannien zur Nutzenbewertung von Galantamin vorgelegt hatte [45,46]. Weiterhin erhielt das IQWiG von Janssen-Cilag auf Nachfrage Erläuterungen zu statistischen Berechnungen aus den Studien zu Galantamin.

Ein Review der Unterlagen ergab keine Hinweise auf unveröffentlichte oder noch nicht identifizierte veröffentlichte Studien.

Informationen aus den Unterlagen, die in die Nutzenbewertung eingeflossen sind und nicht aus den Publikationen zu Galantamin, sondern aus dem NICE-Dossier oder von Janssen-Cilag direkt stammen, wurden im Bericht gesondert gekennzeichnet.

Novartis Pharma GmbH

Nach Unterzeichnung der Vereinbarung zur Übermittlung von Unterlagen konnten der von Novartis zur Verfügung gestellte Expert Report der Zulassung von Rivastigmin sowie eine

Liste der Studien nach Zulassung zur Identifizierung relevanter Studien genutzt werden. Das Review der Unterlagen ergab 3 Studien, die nicht oder nur als Abstract publiziert waren:

<u>Studie</u>	<u>Publikationsstatus</u>
B304	unpubliziert
B351	unpubliziert
Tai 2000	Abstractpublikation [47]

Zu den genannten Studien wurden im März 2006 die Studienberichte bei Novartis angefordert. Der Bericht zu B304 [48] wurde zur Verfügung gestellt, die Studie konnte deshalb in die Nutzenbewertung einfließen. Der Studienbericht zu B351 wurde bis Ende Juli 2006 nicht von Novartis eingereicht, daher konnte die Studie in der vorliegenden Nutzenbewertung nicht dargestellt und adäquat berücksichtigt werden. Zu B351 sind einzelne Ergebnisse in einem Cochrane Review veröffentlicht [49]. Diese Daten zeigen, dass das Ergebnis dieser Studie negativ war. Für den vorliegenden Vorbericht wurden – soweit möglich – Sensitivitätsanalysen unter Verwendung der Daten von B351 aus dem Cochrane Review durchgeführt, um den Einfluss dieser Studie auf die Gesamtaussage der Bewertung abzuschätzen. Die Ergebnisse der Sensitivitätsanalysen sind in den entsprechenden Kapiteln des Berichts und in Anhang G dargestellt.

Zur Studie von Tai 2000 lag bei Novartis kein Studienbericht vor. Auch diese Studie wurde deshalb nicht berücksichtigt.

Informationen, die aus den von Herstellern zur Verfügung gestellten, nicht öffentlich zugänglichen Dokumenten verwendet wurden, sind in den Tabellen durch Kursivschrift gekennzeichnet.

5.1.3 Anfrage an Autoren

Zu folgenden Studien wurden die Autoren angeschrieben und um Zusatzinformationen zu den identifizierten Volltextpublikationen gebeten: AD2000 (E-Mail am 07.07.2006 und 09.08.2006, schriftliche Anfrage am 09.08.2006), Karaman 2005 (E-Mail am 25.07.2006), Kim 2002 (E-Mail am 25.07.2006), Kemp 2003 (E-Mail am 26.07.2006), Thomas 2001 (E-Mail am 01.08.2006) und Wang 2001 (E-Mail am 31.07.2006).

Eine Antwort wurde lediglich von P. Kemp erhalten (E-Mail vom 31.07.2006).

Anfragen zu Studien, die von Firmen durchgeführt wurden, mit denen eine Vertraulichkeitsvereinbarung besteht, wurden direkt über die Hersteller abgewickelt.

5.1.4 Resultierender Studienpool

Die Tabelle 4 zeigt den aus den verschiedenen Suchschritten resultierenden Studienpool.

Tabelle 4: Studienpool

Testsubstanz(en) Studie	Zugeordnete Volltextpublikation	Identifiziert durch	Einschluss in den Bericht
Donepezil vs. Placebo			
AD 2000	Courtney C et al. Lancet 2004; 363: 2105-2115 [50]	Bibliographische Literaturrecherche	nein ^(a)
Burns 1999	Burns et al. Dement Geriatr Cogn Disord 1999; 10: 237-244 [51]	Bibliographische Literaturrecherche	ja
Gauthier 2002	Gauthier S et al. Curr Med Res Opin 2002; 18: 347-354 [52] Feldman H et al. Neurology 2001; 57: 613-620 [53] Feldman H et al. J Am Geriatr Soc 2003; 51: 737-744 [54]	Bibliographische Literaturrecherche	ja
Homma 2000	Homma A et al. Dement Geriatr Cogn Disord 2000; 11: 299-313 [55]	Bibliographische Literaturrecherche	ja
Kemp 2003	Kemp et al. J Neurol Neurosurg Psychiatry 2003; 74: 1567-1570 [56]	Bibliographische Literaturrecherche	nein ^(b)
Krishnan 2003	Krishnan KR et al. Am J Psychiatry 2003; 160: 2003-2011 [57]	Bibliographische Literaturrecherche	ja
Mohs 2001	Mohs RC et al. Neurology 2001; 57: 481-488 [58]	Bibliographische Literaturrecherche	ja
Moraes 2006	dos Santos Moraes et al. Sleep 2006; 199-205 [59]	Bibliographische Literaturrecherche	ja
Prasher 2002	Prasher VP et al. Int J Ger Psychiatry 2002; 17: 270-278 [60]	Bibliographische Literaturrecherche	ja
Rogers 1998	Rogers et al. Neurology 1998; 50: 136-145 [61]	Bibliographische Literaturrecherche	ja
Seltzer 2004	Seltzer B et al. Arch Neurology 2004; 61: 1852-1856 [62]	Bibliographische Literaturrecherche	ja
Tariot 2001	Tariot PN et al. J Am Geriatr Soc 2001; 49: 1590-1599 [63]	Bibliographische Literaturrecherche	ja
Tune 2003	Tune L et al. Am J Geriatr Psychiatry 2003; 11: 169-177 [64]	Bibliographische Literaturrecherche	ja
Winblad 2001	Winblad B et al. Neurology 2001; 57: 489-495 [65] Wimo A et al. Curr Med Res Opin 2004; 20: 1221-1225 [66] Wimo 2003 Dement Geriatr Cogn Disord 2003; 15: 44-54 [67]	Bibliographische Literaturrecherche	ja
Galantamin vs. Placebo			
Brodaty 2005	Brodaty H et al. Dement Geriatr Cogn Disord 2005; 20: 120-132 [68]	Bibliographische Literaturrecherche	ja
Erkinjuntti 2002	Erkinjuntti T et al. Lancet 2002; 359: 1283-1290 [69]	Bibliographische Literaturrecherche	ja

(Fortsetzung)

Tabelle 4 (Fortsetzung): Studienpool

Testsubstanz(en) Studie	Zugeordnete Volltextpublikation	Identifiziert durch	Einschluss in den Bericht
Raskind 2000	Raskind MA et al. Neurology 2000; 54: 2261-2268 [70]	Bibliographische Literaturrecherche	ja
Rockwood 2006	Rockwood K et al. CMAJ 2006; 174: 1099-1105 [71]	Bibliographische Literaturrecherche	ja
Tariot 2000	Tariot PN et al. Neurology 2000; 54: 2269-2276 [72] Cummins JL et al. Am J Psychiatry 2004; 161: 532-538 [73] Galasko D et al. J Am Geriatr Soc 2004; 52: 1070-1076 [74]	Bibliographische Literaturrecherche	ja
Wilcock 2000	Wilcock GK et al. BMJ 2000; 321: 1445-1449 [75]	Bibliographische Literaturrecherche	ja
Rivastigmin vs. Placebo			
B304 1998	Novartis Pharma AG. Study ENA B304 Final Study Report. 1998 [48]	Anfrage an Hersteller	ja
B351		Anfrage an Hersteller	nein ^(c)
Corey-Bloom 1998	Corey-Bloom J et al. Int J Ger Psychopharmacol 1998; 1: 55-65 [76] Kumar et al. Eur J Neurol 2000; 7: 159-169 [77]	Bibliographische Literaturrecherche	ja
Forette 1999	Forette F et al. Eur J Neurol 1999; 6: 423-429 [78]	Bibliographische Literaturrecherche	ja
Karaman 2005	Karaman et al. Dement Geriatr Cogn Disord 2005; 19: 51-56 [79]	Bibliographische Literaturrecherche	nein ^(a)
Rösler 1999	Rösler M et al. BMJ 1999; 318: 633-638 [80] Erkinjuntti et al. Int J Clin Pract 2002; 56: 791-796 [81]	Bibliographische Literaturrecherche	ja
Vergleich verschiedener Cholinesterasehemmer			
<i>Galantamin vs. Donepezil</i>			
Wilcock 2003	Wilcock G et al. Drugs Aging 2003; 20: 777-789 [82]	Bibliographische Literaturrecherche	ja
<i>Rivastigmin vs. Donepezil</i>			
Bullock 2005	Bullock R et al. Curr Med Res Opin 2005; 21: 1317-1123 [83] Bullock et al. Curr Med Res Opin 2006; 22: 483-494 [84] Touchon et al. Curr Med Res Opin 2006; 22: 49-59 [85]	Bibliographische Literaturrecherche	ja

(Fortsetzung)

Tabelle 4 (Fortsetzung): Studienpool

Testsubstanz(en) Studie	Zugeordnete Volltextpublikation	Identifiziert durch	Einschluss in den Bericht
Fuschillo 2001	Fuschillo C et al. Arch Gerontol Geriatr 2001; 7: 151-158 [86]	Bibliographische Literaturrecherche	ja
Kim 2002	Kim et al. Int Psychogeriatr 2002; 14: 187-195 [87]	Bibliographische Literaturrecherche	nein ^(d)
Thomas 2001	Thomas et al. Clin Neuropharmacol 2001; 24: 31-42 [88]	Bibliographische Literaturrecherche	nein ^(a)
Wang 2001	Wang Y et al. Chinese Journal of Neurology 2001; 34: 210-213 [89]	Bibliographische Literaturrecherche	ja
<i>Rivastigmin vs. Donepezil vs. Galantamin</i>			
Cumbo 2005	Cumbo E. Prim Care Comm Psych 2005; 10: 95-102 [90]	Bibliographische Literaturrecherche	ja
Publikationen zu mehreren Studien (gepoolte Analysen)			
Anand et al. Int J Ger Psychopharmacol 2000; 2: 68-72 [91]		Bibliographische Literaturrecherche	ja
Burns et al. Int J Geriatr Psychiatry 2004; 19: 243-249 [92]		Bibliographische Literaturrecherche	ja
Kurz et al. Alzheimer Dis Assoc Disord 2004; 18: 123-128 [93]		Bibliographische Literaturrecherche	ja
Marcusson et al. Alzheimer Dis Assoc Disord 2003; 17 Suppl 3: S86-S91 [94]		Bibliographische Literaturrecherche	ja
Orgogozo et al. Curr Med Res Opin 2004; 20: 1815-1820 [95]		Bibliographische Literaturrecherche	ja
Potkin et al. Prog Neuropsychopharmacol Biol Psychiatry 2002; 26: 713-720 [96]		Bibliographische Literaturrecherche	ja
Pratt et al. Int J Clin Pract 2002; 56: 710-717 [97]		Bibliographische Literaturrecherche	ja
Sano M et al. Int J Ger Psychiatry 2003; 18: 942-950 [98]		Bibliographische Literaturrecherche	ja
Schneider et al. Int J Ger Psychopharmacol 1998; 1: S26-S34 [99]		Bibliographische Literaturrecherche	ja
Whitehead et al. Int J Geriatr Psychiatry 2004; 19: 624-633 [100]		Bibliographische Literaturrecherche	ja
Wilkinson et al. Int J Clin Pract 2002; 56: 509-514 [101]		Bibliographische Literaturrecherche	ja
a: Studie konnte nicht eingeschlossen werden, da wesentliche Fragen zur Interpretation bis zur Fertigstellung des Vorberichts nicht geklärt werden konnten; eine entsprechende Anfrage an die Autoren blieb unbeantwortet.			
b: Studie konnte nicht eingeschlossen werden, da keine für die vorliegende Fragestellung relevanten Datenauswertungen durchgeführt wurden bzw. zur Verfügung gestellt werden konnten (Auskunft P. Kemp am 31.07.2006).			
c: Diese unveröffentlichte Studie konnte nicht eingeschlossen werden, da der Studienbericht bis Ende Juli 2006 nicht von Novartis zur Verfügung gestellt wurde.			
d: Studie konnte nicht eingeschlossen werden, da die Publikation keine für die vorliegende Fragestellung relevanten Datenauswertungen darstellt; eine entsprechende Anfrage an die Autoren blieb unbeantwortet.			

In die Bewertung eingeschlossen wurden die durch die Suche identifizierten, relevanten und vollständig publizierten Studien. Ausnahmen stellen folgende Studien dar, die (bisher) nicht in die Nutzenbewertung einbezogen werden konnten:

AD2000:

In der AD2000-Studie sind 565 Patienten mit leichter bis mittelschwerer Alzheimer-Demenz zum Vergleich von Donepezil gegenüber Placebo eingeschlossen worden. Die an 22 Zentren in England durchgeführte Studie betrachtete die Patienten über einen Zeitraum von maximal etwa 4 Jahren bis zu einer Heimaufnahme bzw. bis zu einem definierten Grad des Verlustes von alltagspraktischen Fähigkeiten (primäre Zielkriterien). Diese Studie wurde ausschließlich mit Mitteln des englischen Gesundheitssystems (National Health Service) durchgeführt. Sie ist als randomisiert und doppelblind beschrieben und die Ergebnisse zu den primären Endpunkten (Zeitspanne bis zum Eintritt des Ereignisses) sind mit Hilfe von Überlebenszeitanalysen ausgewertet worden. Im Gegensatz zu den üblichen randomisierten Zweigruppenvergleichen wurden die Patienten in der AD2000-Studie zum Großteil zweimal randomisiert. Zu Studienbeginn erfolgte eine Randomisierung aller eingeschlossenen Patienten in die beiden Gruppen Donepezil 5 mg und Placebo. Nach zwölf Wochen wurden die Patienten erneut randomisiert; diesmal in die drei Gruppen Donepezil 5mg, Donepezil 10mg und Placebo. Die Patienten sind entsprechend dieser Aufteilung in drei Therapiearme bis zu vier Phasen à 48 Wochen lang beobachtet worden, wobei zwischen diesen Phasen jeweils eine vier bis sechswöchige Washout-Phase stattfand. Die ungewöhnliche Rerandomisierung nach zwölf Wochen führte dazu, dass etwa 43 % der Patienten von Donepezil nach Placebo bzw. umgekehrt wechselten, etwa 11 % eine Dosissteigerung von 5mg auf 10mg Donepezil erfuhren und nur bei etwa 32 % der Patienten die ursprünglich begonnene Therapie beibehalten wurde (etwa 14 % sind innerhalb der ersten zwölf Wochen gestorben oder ausgeschieden, die genannten Prozentzahlen entstammen einer eigenen Berechnung anhand der Angaben aus der Publikation).

Es gibt verschiedene Aspekte bezüglich des Studiendesigns und der Darstellung der Ergebnisse in der Publikation [50], die die Aussagekraft der Resultate einschränken:

- **Therapiezuordnung:** Durch das zweimalige Randomisieren entstand eine Mischung der Studienpopulation hinsichtlich der Behandlung, bestehend aus Therapiewechseln, Patienten mit Dosiserhöhung und Patienten, die die Therapie nicht wechselten. Es kann daher nicht ausgeschlossen werden, dass die Ergebnisse durch Überhangeffekte verzerrt sind, insbesondere, da vor der zweiten Randomisierung keine Washout-Phase eingefügt wurde. Auch wenn die Autoren im Methodenteil ihrer Publikation angeben, dass das verwendete Analyseverfahren („standard mixed model techniques“) diese Effekte berücksichtigt, ist dies allenfalls für die sekundären Endpunkte (diverse Merkmale auf einem kontinuierlichen Zahlenniveau) möglich, allerdings auch hierfür in der Publikation nicht nachvollziehbar dargestellt. Bezüglich der Auswertung der primären Endpunkte

muss festgestellt werden, dass die Verwendung von Kaplan-Meier-Schätzern und des Log-Rank-Tests dem Studiendesign nicht gerecht wird. In der Publikation sind sämtliche Ergebnisse lediglich als Zweigruppenvergleich Donepezil vs. Placebo dargestellt, wobei der Beginn der Beobachtung der Zeitpunkt der ersten Randomisierung ist. Es finden sich keine getrennten Ergebnisse für die beiden Dosisstufen unter Donepezil. Auch Analysen, bei denen der Zeitpunkt der zweiten Randomisierung als Basis dient, werden nicht berichtet.

- **Washout-Phasen:** Während eine Washout-Phase nach den ersten 12 Wochen zur Verringerung möglicher Überhangeffekte sinnvoll gewesen wäre, liefern die Autoren für die vier- bis sechswöchigen Washout-Phasen zwischen den Behandlungsphasen keine Begründung.

Zur Klärung dieser Aspekte wurde eine Anfrage per E-Mail an die AD2000 Collaborative Group verschickt mit der Bitte um weitere Details zu Studiendesign und -durchführung sowie zu den Analysen. Diese Bitte wurde sowohl per E-Mail als auch schriftlich wiederholt, jedoch blieben die Anfragen unbeantwortet. Aufgrund der beschriebenen Aspekte ist die Interpretierbarkeit der Ergebnisse so stark eingeschränkt, dass es geboten schien, in diesem Bericht auf die Darstellung der Resultate und die Einbeziehung dieser in die Bewertung zu verzichten.

B351:

Ebenfalls nicht in die Nutzenbewertung eingeschlossen werden konnte die Studie B351, die (neben der Studie B304) in den Unterlagen der Herstellerfirma Novartis als potenziell relevante Studie identifiziert worden war. Zudem ging aus anderen Veröffentlichungen zu Rivastigmin [49,99] hervor, dass die Studie B351 hinsichtlich des angewendeten Designs und der untersuchten Zielkriterien für die vorliegende Nutzenbewertung relevante Ergebnisse beinhaltet. Zu beiden Studien wurden am 27.03.2006 die Studienberichte bei der Firma Novartis angefordert. Zu der Studie B304 wurde der Studienbericht von der Firma Novartis zur Verfügung gestellt und konnte somit in die Nutzenbewertung eingehen. Zu der Studie B351 hatte die Firma bis Ende Juli 2006 keine Unterlagen zur Verfügung gestellt.

Karaman 2005:

Die Publikation lässt grundlegende Fragen zum Design offen, u. a. hinsichtlich der Zuteilung der Patienten zu den Gruppen, der Anzahl der randomisierten Patienten sowie hinsichtlich der Behandlung sowohl für die Placebo- als auch für die Verum-Gruppe nach den ersten 8 Wochen. Eine Anfrage an die Autoren blieb unbeantwortet.

Kemp 2003:

In der Publikation wird keine für die vorliegende Fragestellung geeignete Auswertung der Daten dargestellt. Nach Auskunft der Autoren sind die Datensätze nicht mehr verfügbar, so dass keine entsprechenden Auswertungen zur Verfügung gestellt werden konnten.

Kim 2002:

In der Publikation wird keine für die vorliegende Fragestellung geeignete Auswertung der Daten dargestellt. Eine Anfrage an die Autoren blieb unbeantwortet.

Thomas 2001:

Die Publikation lässt grundlegende Fragen zum Design offen. So wird u. a. nicht klar beschrieben, ob die Behandlung in der Rivastigmin- und der Donepezil-Gruppe parallel erfolgte. Eine Anfrage an die Autoren blieb unbeantwortet.

Alle eingeschlossenen Studien stellen einen Vergleich der verschiedenen Cholinesterasehemmer zu Placebo oder untereinander an. Es wurden keine randomisierten Studien identifiziert, die den Nutzen eines Cholinesterasehemmers mit einer anderen medikamentösen Therapie oder einer nichtmedikamentösen Therapie bei der Alzheimer Demenz vergleichen.

Im folgenden Bericht werden die Ergebnisse der Studien getrennt für die verschiedenen Cholinesterasehemmer dargestellt.

5.1.5 Informationen aus der Anhörung

Aus der Anhörung ergaben sich folgende für die Nutzenbewertung relevante Informationen:

- Zusatzinformationen zur Methodik einiger Studien, die in die Bewertung der Studien- und Publikationsqualität eingingen, sowie
- Zusatzinformationen zu den Ergebnissen einzelner Studien, die in die Darstellung der Ergebnisse zu den jeweiligen Therapiezielen aufgenommen wurden.

Weitere im Rahmen der Stellungnahmen und der wissenschaftlichen Erörterung vorgebrachten Aspekte werden im Abschnitt 6 „Diskussion“ aufgegriffen.

In den Stellungnahmen wurden keine den Ein- und Ausschlusskriterien genügenden Studien genannt, die nicht im Vorbericht berücksichtigt wurden.

5.2 Charakteristika der in die Bewertung eingeflossenen Studien

5.2.1 Donepezil

5.2.1.1 Studiendesign, Dosierungen und Studienpopulationen

Insgesamt fanden sich 14 relevante Studien zu Donepezil im Vergleich mit Placebo, von denen 12 in die Nutzenbewertung eingeschlossen werden konnten (Tabelle 4). Einzelheiten zum Design und zu den grundsätzlichen Charakteristika der eingeschlossenen Studien finden sich in Tabelle 5. Nachfolgend werden die wesentlichen Aspekte zusammenfassend beschrieben.

Bis auf vier kleinere Studien waren alle Studien multizentrisch angelegt. Die Studiendauer betrug zumeist 24 Wochen. Es fanden sich jedoch auch zwei Studien mit einer kontrollierten und verblindeten Medikamentengabe über einen Zeitraum von einem Jahr (Winblad 2001) bzw. 13,5 Monaten (Mohs 2001). Eingesetzt wurden Dosierungen von 5 und 10 mg Donepezil täglich. Drei Studien, zumeist mit schwerer Erkrankten, erlaubten entweder eine fakultative Höherdosierung auf 10 mg oder eine Dosisreduktion auf 5 mg nach Ermessen des Studienarztes (Winblad 2001) oder eine Dosisreduktion auf 5 mg bei Verträglichkeitsproblemen nach zuvor geforderter Steigerung auf 10 mg (Gauthier 2002, Tariot 2001). Bei den Studien mit höherer Dosierung erfolgte die Dosissteigerung von 5 mg auf 10 mg nach frühestens einer Woche (Burns 1999), zumeist innerhalb von 6 Wochen.

Bei Gauthier 2002 handelt es sich um eine Subgruppenanalyse der Studie von Feldman et al. 2001, in die Patienten mit mittelschwerer bis schwerer Alzheimer Demenz (MMSE von 5–17) eingeschlossen worden waren. Für diesen Bericht wurden gemäß des Indikationsbereiches von Cholinesterasehemmern lediglich die Daten der Subpopulation mit mittelschwerer Demenz (MMSE von 10–17) herangezogen, die etwa 71 % der Gesamtpopulation ausmacht.

Die Zahl der in die Studien eingeschlossenen Patienten betrug – bis auf vier Fälle mit deutlich kleineren Patientenzahlen – zwischen 153 (Seltzer 2004) und 818 (Burns 1999, dreiarmlig). Insgesamt wurden ca. 1700 Patienten mit Donepezil und 1300 Patienten mit Placebo behandelt.

Die Studien unterschieden sich hinsichtlich ihrer Einschluss- und Ausschlusskriterien (Tabelle 6) z. T. wesentlich; diese Unterschiede spiegeln sich auch in den tatsächlich betrachteten Studienpopulationen wider (Tabelle 7) und sind deshalb im Folgenden ausführlich dargestellt.

Die meisten Studien schlossen Patienten mit einer wahrscheinlichen Alzheimer Demenz nach den NINCDS-ADRDA-Kriterien ein, wobei der MMSE meist zwischen 10 und 26 Punkten lag.

In der Studie von Burns 1999 wiesen über 80 % der Patienten eine leichte Demenz (CDR 1) auf, der durchschnittliche MMSE lag bei 20 Punkten. Vergleichbar mit dieser Studie waren die Patientenpopulationen der Studien von Rogers 1998 und Winblad 2001 sowie der kleineren Studien von Krishnan 2003 und Tune 2003. Die Erkrankungsschwere auf der CDR-Skala (80 % der Teilnehmer im Stadium CDR 1) lag auch für die Patienten bei Mohs 2001 in diesem Bereich, trotz insgesamt schlechterem durchschnittlichem MMSE (17 Punkte). Dies lässt sich aus den höheren Anforderungen an noch zu bewältigende Aktivitäten des täglichen Lebens gemäß den Einschlusskriterien erklären. Bei Homma 2000 wies ein höherer Anteil der Patienten eine bereits mittelschwere Erkrankung auf (ca. 35 % CDR 2, durchschnittlicher MMSE ca. 17). In dieser Studie fand sich eine deutlich geringere kognitive Beeinträchtigung in der Verum-Gruppe im Vergleich mit der Placebo-Gruppe zu Studienbeginn (Unterschied der Gruppen im MMSE +1,2 Punkte, $p = 0,035$; im ADAS-cog -3,9 Punkte, $p = 0,001$). Da jedoch die individuelle Progression (Baseline versus Studienende) betrachtet und in einer (zusätzlichen) Kovarianzanalyse (ANCOVA) auch der Baselinewert als Faktor berücksichtigt wurde (und laut Autoren keinen Einfluss auf das Ergebnis hatte), bleiben die Studienergebnisse interpretierbar.

Die Studie von Seltzer 2004 schloss ausschließlich sehr leicht beeinträchtigte Teilnehmer (MMSE 21–26) ein. Bei Patienten am oberen Rand dieses Bereichs (25–26) kann zum Teil vermutlich auch von einem Mild Cognitive Impairment gesprochen werden, wobei die Abgrenzung MCI/leichte AD allerdings nur unscharf ist. Immerhin 32 % der Patienten befanden sich in dem mit einem MCI zu vereinbarenden CDR-Stadium 0,5, die restlichen Patienten hatten eine leichte Demenz (CDR 1). In dieser Studie wurde auch in der Placebo-Gruppe ein lediglich minimales Fortschreiten der kognitiven Beeinträchtigung beobachtet, wie es in Gruppen mit zu Beginn nur leichten kognitiven Beeinträchtigungen häufig ist.

Gauthier 2002 berichtete über die Auswertung der Subgruppe von mittelschwer beeinträchtigten Patienten (MMSE 10–17) im Rahmen einer Studie [53], die auch schwer beeinträchtigte Patienten einschloss. Im Folgenden werden, soweit möglich, lediglich die Daten zu dieser Subgruppe dargestellt. Die jeweiligen Ergebnisse von Sub- und Gesamtgruppe waren im Therapieverlauf vergleichbar. Ähnliches gilt für die Studie Tariot 2001, in der Patienten mit leichter bis schwerer AD (MMSE 5–26) betrachtet wurden, die protokollgemäß alle Pflegeheimbewohner waren. Da 75 % der Patienten eine leichte bis mittelschwere Beeinträchtigung (MMSE 10–26) aufwiesen, ist die Studie hinreichend geeignet, für diese Population Aussagen zu treffen. Soweit separate Auswertungen zur Subgruppe der leicht bis mittelschwer beeinträchtigten Patienten durchgeführt worden sind, werden nur diese im Folgenden dargestellt. Auch hier zeigten sich vergleichbare Ergebnisse von Sub- und Gesamtgruppe. In diesen beiden Studien war das Vorliegen einer Bildgebung zu Studienbeginn nicht verpflichtend vorgesehen und es ist deshalb mit einer stärkeren zusätzlichen vaskulären Begleitpathologie zu rechnen.

Eine besondere Studienpopulation stellen die Patienten mit einer Trisomie 21 und einer möglichen Alzheimer Demenz bei Prasher 2002 dar. Da laut Fachinformation Donepezil für die Anwendung bei Menschen mit Trisomie 21 nicht ausdrücklich ausgeschlossen ist (vgl. [102]), wurde die Studie in die Nutzenbewertung mit aufgenommen, auch wenn diese Patientengruppe naturgemäß nur eine kleine Minderheit der Patienten mit Alzheimer Demenz repräsentiert. Zur Diagnose einer Demenz vom Alzheimer Typ wurden in dieser Gruppe gesonderte Kriterien angewandt (Diagnostic Research Criteria der WHO), deren Validität unklar ist. Ob bei Trisomie 21 pharmakogenetische Besonderheiten für die Wirkung von Cholinesterasehemmern vorliegen könnten, ist ebenfalls unklar. Angesichts der deutlichen Ungleichverteilung von prognostisch möglicherweise relevanten Merkmalen zwischen den randomisierten Gruppen bleibt die Bedeutung der Studienergebnisse jedoch auch für Menschen mit Trisomie 21 per se unsicher.

Die wichtigsten Ausschlusskriterien in den einzelnen Studien waren andere mögliche Ursachen für die Demenz, andere neurologische oder psychiatrische Erkrankungen sowie weitere bedeutende, nicht eingestellte Erkrankungen wie z. B. insulinpflichtiger Diabetes, Atemwegserkrankungen, hämatologische oder onkologische Erkrankungen.

In fünf Studien stellte ADAS-cog und/oder CIBIC-plus das primäre Zielkriterium dar (Tabelle 5). Zwei Studien (Krishnan 2003 und Tune 2003 mit 67 bzw. 28 Teilnehmern) untersuchten jeweils die Veränderungen neurobiologischer Marker mittels apparativer Verfahren als primäre Zielgröße (Magnetresonanztomographie und -spektroskopie, Positronen-Emissions-Tomographie) bzw. die Veränderung des REM-Schlafs unter Donepezilgabe (Moraes 2006 mit 35 Teilnehmern). Allerdings wurden auch Ergebnisse zu vorab definierten patientenrelevanten Therapiezielen dargestellt. In Tabelle 5 sind für alle Studien die jeweils primären Zielkriterien dargestellt. Darüber hinaus sind dort die gemäß Abschnitt 4.1.3 patientenrelevanten Zielkriterien aufgelistet. Bei der überwiegenden Zahl der Studien wurden die verwendeten Skalen primär kontinuierlich ausgewertet, häufig ohne Definition eines Responsekriteriums.

In der kleinen Studie von Tune 2003 findet sich eine (statistisch signifikante) Ungleichverteilung bei einem Baselinecharakteristikum, und zwar beim NPI. Für die Auswertung zu diesem Zielkriterium wurde allerdings die Differenz zu Behandlungsbeginn betrachtet und für die (unterschiedlichen) Baselinewerte adjustiert (ANCOVA). In der ebenfalls kleinen Studie von Moraes 2006 wurde ein Unterschied zwischen den beiden Gruppen im ADAS-cog von etwa 3 Punkten in der Analyse nicht berücksichtigt. Auch die Gruppenunterschiede im SIB zur Baseline bei den Patienten in der Studie von Prasher 2002 gingen offensichtlich nicht in die Analyse ein.

Ein besonderes Studiendesign wies die Studie von Mohs 2001 auf: Hier wurden die Teilnehmer nicht über einen fixen Zeitraum verfolgt, sondern bis zum Erreichen eines vorher definierten (primären) Endpunktes, der auch als Studienende für die jeweils Betroffenen

herangezogen wurde. Die Ergebnisse wurden im Sinne einer „Überlebenszeitanalyse“ ausgewertet. Dieses Design ließ Ergebnisse außerhalb dieser Überlebenszeitanalysen nicht mehr sicher interpretieren, da entweder sukzessive Patienten für die entsprechenden Auswertungen verloren gingen oder die Werte zum Zeitpunkt des Erreichens des Endpunktes verwendet wurden, womit das Prinzip des „zufälligen Fehlens“ („missing at random“) verletzt ist.

Bei Mohs 2001 wurde das Erreichen einer Verschlechterung der basalen Aktivitäten bzw. von mehreren instrumentellen Aktivitäten (mehr als 20 % der bei Studienbeginn durchführbaren iADL) des täglichen Lebens als klinischer Endpunkt definiert. Eine alternative Formulierung (Zunahme des CDR um einen Punkt) war nur für eine kleine, in beiden Armen ca. gleich große Gruppe ausschlaggebend für das Erreichen des Endpunkts.

Tabelle 5: Eingeschlossene Studien zu Donepezil

Studie (Weitere Studien- bezeichnungen)	Studiendesign	Studiendauer (kontrollierte Dauer der Verumgabe)	Interventionsgruppen und Zahl der randomisierten Patienten	Ort der Durchführung und Zahl der Zentren	Relevante Zielkriterien ^(a)
Burns 1999 (E044-304)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Donepezil 5 mg: n=271 2. Donepezil 10 mg: n=273 3. Placebo: n=274	AUS, B, CAN, F, D, GB, IRL, NZ, ZA (Gesamt:82 Zentren)	ADAS-cog, CIBIC-plus IDDD (modifiziert), QoL, unerwünschte Ereignisse
Gauthier 2002 ^(b) (MSAD; Don-NY-96- 002-324)	RCT, parallel, doppelblind, multizentrisch	6 Monate	Subgruppe mit mittelschwerer AD: 1. Donepezil (5-) 10 mg: n=102 2. Placebo: n=105	CAN (22), AUS (6), F (4) (Gesamt: 32 Zentren) ^(c)	CIBIC-plus sMMSE, DAD, IADL+, PSMS+, NPI, CSS ^(d) , unerwünschte Ereignisse
Homma 2000	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Donepezil 5 mg: n=136 2. Placebo: n=132 ^(e)	Japan (54 Zentren)	ADAS-J-cog^(f), J-CGIG CMCS, unerwünschte Ereignisse
Krishnan 2003	RCT, parallel, doppelblind, monozentrisch	6 Monate	1. Donepezil 10 mg: n=34 2. Placebo: n=33	USA (1 Zentrum)	N-Acetyl-Aspartat- Konzentration ADAS-cog, unerwünschte Ereignisse
Mohs 2001 (A001-312; The Functional Survival Study)	RCT, parallel, doppelblind, multizentrisch, <i>Überlebenszeitanalyse</i>	13,5 Monate	1. Donepezil 10 mg: n=214 2. Placebo: n=217	USA (31 Zentren)	Zeit bis zu einer klinisch feststellbaren funktionellen Verschlechterung^(g) ADFACS, CDR-SB, MMSE, unerwünschte Ereignisse
Moraes 2006	RCT, parallel, doppelblind, bizenrisch	6 Monate	1. Donepezil 10 mg: n=17 2. Placebo: n=18	Brasilien (2 Zentren)	REM-Schlaf ADAS-cog

(Fortsetzung)

Tabelle 5 (Fortsetzung): Eingeschlossene Studien zu Donepezil

Studie (Weitere Studien- bezeichnungen)	Studiendesign	Studiendauer (kontrollierte Dauer der Verumgabe)	Interventionsgruppen und Zahl der randomisierten Patienten	Ort der Durchführung und Zahl der Zentren	Relevante Zielkriterien ^(a)
Prasher 2002 (UK Down Syndrome Ageing Study)	RCT, parallel, doppelblind, monozentrisch	6 Monate	1. Donepezil 10 mg: n=16 2. Placebo: n=15	GB (1 Zentrum)	DMR SIB, NPI, unerwünschte Ereignisse
Rogers 1998 (A001-302)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Donepezil 5 mg: n=154 2. Donepezil 10 mg: n=157 3. Placebo: n=162	USA (20 Zentren)	ADAS-cog, CIBIC-plus QoL, unerwünschte Ereignisse
Seltzer 2004 (E2020-A001-402)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Donepezil 10 mg: n=96 2. Placebo: n=57	USA (17 Zentren)	ADAS-cog13 CDR-SB
Tariot 2001 ^(b)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Donepezil (5-) 10 mg: n=103 2. Placebo: n=105 Subgruppe mit leichter bis mittelschwerer AD 1. Donepezil (5-) 10 mg: n=76 2. Placebo: n=80	USA (27 Pflegeheime)	NPI-NH CDR-SB, MMSE, PSMS, unerwünschte Ereignisse
Tune 2003	RCT, parallel, doppelblind, monozentrisch	6 Monate	1. Donepezil 10 mg: n=14 2. Placebo: n=14	USA (1 Zentrum)	Hirnstoffwechsel (PET) ADAS-cog, NPI
Winblad 2001 (DON-NY-96-001)	RCT, parallel, doppelblind, multizentrisch	12 Monate	1. Donepezil (5-) 10 mg: n=142 2. Placebo: n=144	DK, FIN, N, S, NL (Gesamt: 28 Zentren)	GBS PDS, MMSE, NPI
<p>a: primäre sowie relevante Zielkriterien gemäß Abschnitt 4.1.3 (falls in der jeweiligen Studie als „primär“ deklariert Fettdruck).</p> <p>b: Alle Angaben beziehen sich – soweit nicht anders gekennzeichnet – auf die Publikation Gauthier 2002 [52], die über die Subgruppe der mittelschwer beeinträchtigten Patienten (MMSE 10–17) der ursprünglichen Studie von Feldman 2001 [53] berichtet.</p> <p>c: Angaben aus Feldman 2001 [53]</p> <p>d: Daten zu CSS sind veröffentlicht in Feldman 2003 [54] für die Gesamtpopulation der ursprünglichen Studie von Feldman 2001 (MMSE 5–17)</p> <p>e: Anzahl der Patienten in den Gruppen erschlossen aus Safety-Angaben auf S. 307 der Publikation.</p> <p>f: japanische Version des ADAS-cog</p> <p>g: „clinically evident decline“: Kriterium aus ADFACS bzw. CDR</p>					

Tabelle 6: Wesentliche Ein- und Ausschlusskriterien der Studien zu Donepezil

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Burns 1999	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26; CDR: 1-2 - CT oder MRT - insgesamt guter Gesundheitszustand 	<ul style="list-style-type: none"> - andere neurologische oder psychiatrische Erkrankungen - signifikante, nicht eingestellte gastrointestinale, renale, hepatische, endokrine oder onkologische Erkrankungen; Asthma
Gauthier 2002 ^(b)	<ul style="list-style-type: none"> - mittelschwere AD (wahrscheinlich oder möglich) - MMSE: 10-17 - FAST: ≤ 6 - CT oder MRT 	<ul style="list-style-type: none"> - andere mögliche Demenzursachen - andere primäre neurologische oder psychiatrische Diagnose - Delir, Depression oder andere Diagnose, die die Studienteilnahme beeinträchtigen könnte - klinische signifikante obstruktive Atemwegserkrankungen oder Asthma, hämatologische oder onkologische Erkrankungen, Vitamin-B12- oder Folatmangel, gastrointestinale, renale, hepatische, endokrine oder kardiovaskuläre Erkrankungen - Substanzmissbrauch in den letzten 10 Jahren - vollständige Pflegebedürftigkeit <p>Andere Komorbiditäten wie z. B. eingestellter Diabetes, Bluthochdruck oder Schilddrüsenerkrankung waren zulässig.</p>
Homma 2000	<ul style="list-style-type: none"> - leichte bis mittelschwere AD - MMSE: 10-26; CDR: 1-2 - ADAS-J-cog: ≥ 15 - Hachinski Score^(c): ≤ 4 - CT und MRT - Betreuungsperson 	<ul style="list-style-type: none"> - neurologische Auffälligkeiten (z. B. Parkinson), frühere Kopfraumata mit Bewusstseinsstörungen - Depression - schwere Komplikationen - Magen-/Darmgeschwür
Krishnan 2003	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26; CDR: 1-2 - Hachinski Score^(c): ≤ 4 - insgesamt guter Gesundheitszustand - Eignung für MRT (z. B. kein Herzschrittmacher) 	<ul style="list-style-type: none"> - zerebrovaskuläre Erkrankung - andere psychiatrische Primärdiagnose - jegliche instabile medizinische Erkrankung

(Fortsetzung)

Tabelle 6 (Fortsetzung): Wesentliche Ein- und Ausschlusskriterien der Studien zu Donepezil

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien^(a)
Mohs 2001	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE 12-20 bzw. 21; CDR: 1-2 - Hachinski Score^(c): ≤ 4 - Mindestanzahl erhaltener Aktivitäten des täglichen Lebens - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurologische oder psychiatrische Erkrankungen - Komplikationen der Demenzerkrankungen (organische Erkrankungen, Delirium, Depression, Wahnvorstellungen) - frühere Abhängigkeitserkrankung
Moraes 2006	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - CDR: 1-2 	<ul style="list-style-type: none"> - andere Demenzursachen - andere schwerwiegende medizinische oder psychiatrische Erkrankungen - mittlere bis schwere Schlafstörungen - Apnoe-Hypopnoe-Index (AHI) > 10 und Beinbewegungs-Index (PLMI) > 5 pro Stunde
Prasher 2002	<ul style="list-style-type: none"> - Down-Syndrom - leichte bis mittelschwere AD nach DRC-10 - bei einer Betreuungsperson wohnend 	<ul style="list-style-type: none"> - signifikante Erkrankungen (z. B. insulinpflichtiger Diabetes oder andere unbehandelte Stoffwechselerkrankung, Asthma, obstruktive Atemwegserkrankung, relevante unkontrollierte neurologische, gastrointestinale, hepatische oder kardiovaskuläre Erkrankung, Vitamin-B12- oder Folatmangel) - Eingestellte Epilepsie, behandelte Schilddrüsenerkrankung, stabile psychotrope Medikation waren zulässig.
Rogers 1998	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26; CDR: 1-2 - Betreuungsperson 	<ul style="list-style-type: none"> - insulinpflichtiger Diabetes oder andere endokrine Störung - Asthma oder obstruktive Atemwegserkrankung - klinisch signifikante, nicht eingestellte gastrointestinale, hepatische oder kardiovaskuläre Erkrankungen
Seltzer 2004	<ul style="list-style-type: none"> - leichte AD (wahrscheinlich) - MMSE: 21-26; CDR: 0,5-1 - Hachinski Score^(c): ≤ 4 - nur geringe Einschränkung in den Aktivitäten des täglichen Lebens 	<ul style="list-style-type: none"> - Gedächtnisstörung möglicherweise aufgrund anderer psychiatrischer oder neurologischer Störung oder nach Schädeltrauma

(Fortsetzung)

Tabelle 6 (Fortsetzung): Wesentliche Ein- und Ausschlusskriterien der Studien zu Donepezil

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Tariot 2001	<ul style="list-style-type: none"> - leichte bis schwere^(d) AD (wahrscheinlich oder möglich) mit oder ohne vaskuläre Komponente - MMSE: 5-26 - NPI-NH: mindestens 1 Symptom mehrmals wöchentlich - seit mindestens 1 Monat in Pflegeheim wohnend - medizinisch stabil (Vorerkrankungen zulässig) 	<ul style="list-style-type: none"> - andere neurologische Erkrankungen, Parkinson, vaskuläre Demenz - nicht eingestellte Erkrankungen sowie klinisch signifikante obstruktive Atemwegserkrankung, Asthma, Vitamin-B12-Mangel, hämatologische oder onkologische Erkrankungen, Hemiparese oder Aphasie - Krankenhausaufenthalt in den letzten 3 Monaten - Substanzmissbrauch innerhalb des letzten Jahres, sekundäre Demenz nach Alkoholmissbrauch
Tune 2003	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26; CDR: 1-2 - Hachinski Score^(c): ≤ 4 - CT oder MRT 	keine Angaben
Winblad 2001	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich oder möglich) - MMSE: 10-26 - insgesamt guter Gesundheitszustand - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurologische oder psychiatrische Primärdiagnose (insb. Depression oder vaskuläre Demenz) - klinisch signifikante und nicht eingestellte, aktive gastrointestinale, renale, hepatische, endokrine oder kardiovaskuläre Erkrankungen - Schilddrüsenunterfunktion - Neubildungen, insulinpflichtiger Diabetes, obstruktive Atemwegserkrankung, Asthma, hämatologische/onkologische Erkrankungen, perniziöse Anämie, Vitamin-B12- oder Folatmangel - Substanzmissbrauch innerhalb der letzten 10 Jahre
<p>a: über die üblichen Kontraindikationen (z. B. Wirkstoffunverträglichkeit) bzw. Ausschlusskriterien (z. B. konkurrierende medikamentöse Begleittherapien) hinausgehend</p> <p>b: Angaben aus Feldman 2001 [53]</p> <p>c: Die Hachinski-Skala wird zur Differentialdiagnose zwischen einer Alzheimer Demenz und einer vaskulären Demenz verwendet.</p> <p>d: Der größte Teil der eingeschlossenen Patienten (etwa 75 %) weist eine leichte bis mittelschwere AD auf (MMSE 10–26). Sofern für diese Subgruppe in der Publikation gesonderte Auswertungen berichtet wurden, werden diese dargestellt.</p>		

Tabelle 7: Charakterisierung der Studienpopulationen Donepezil

Studie	Alter (Jahre) ^(a)	Geschlecht w (%)	MMSE ^(a)	NPI ^(a)	Kommentar (zur Vergleichbarkeit etc.)	Studienabbrecher N (%)
Burns 1999						
DON 5 mg	72 (8,2) ^(b)	61	20 (4,9) ^(b)	-		60 (22)
DON 10 mg	72 (8,3) ^(b)	57	20 (3,3) ^(b)	-		72 (26)
Placebo	71 (8,3) ^(b)	55	20 (5,0) ^(b)	-		55 (20)
Gauthier 2002						
DON (5-) 10mg	74 (52-92)	69	14 (2,9) ^(b)	18 (17,1) ^(b)		19 ^(c) (19)
Placebo	74 (48-90)	57	14 (2,7) ^(b)	17 (16,4) ^(b)		12 ^(c) (11)
Homma 2000						
DON 5 mg	70 (7,6) ^(d)	68 ^(d)	18 (3,9) ^(d)	-	nur Basisdaten der PP-Gruppe gegeben, Einfluss der Unterschiede in Baseline MMSE und ADAS-cog in ANCOVA untersucht	17 (13) ^(c)
Placebo	69 (8,8) ^(d)	66 ^(d)	17 (3,9) ^(d)	-		22 (17) ^(c)
Krishnan 2003						
DON 10 mg	74 (7,0)	74	20 (4,8)	-		6 (18)
Placebo	72 (10,1)	70	19 (4,6)	-		10 (30)
Mohs 2001						
DON 10 mg	75 (8,8)	61	17 (2,9)	-		60 (28)
Placebo	75 (8,8)	65	17 (3,0)	-		56 (26)

(Fortsetzung)

Tabelle 7 (Fortsetzung): Charakterisierung der Studienpopulationen Donepezil

Studie	Alter (Jahre) ^(a)	Geschlecht w (%)	MMSE ^(a)	NPI ^(a)	Kommentar (zur Vergleichbarkeit etc.)	Studienabbrecher N (%)
Moraes 2006						
DON 10 mg	77 (6,6)	76	-	-	Donepezil-Gruppe etwas weniger stark beeinträchtigt ^(g)	k.A.
Placebo	75 (9,8)	61	-	-		k.A.
Prasher 2002						
DON 10 mg	53 (8,0)	37	-	8 (5,8)	deutliche Unterschiede hinsichtlich Geschlecht und SIB (Kognition)	2 (13) ^(f)
Placebo	55 (4,6)	64	-	8 (7,6)		2 (13) ^(f)
Rogers 1998						
DON 5 mg	73 (7,5) ^(b)	63	19 (5,0) ^(b)	-		23 ^(e) (15)
DON 10 mg	75 (7,5) ^(b)	62	19 (5,0) ^(b)	-		50 ^(e) (32)
Placebo	73 (7,6) ^(b)	61	19 (5,1) ^(b)	-		32 ^(e) (20)
Seltzer 2004						
DON 10 mg	73 (9,6)	50	24 (1,7)	-		26 (27)
Placebo	75 (8,8)	60	24 (1,3)	-		11 (19)
Tariot 2001						
DON (5-) 10mg	85 (64-98)	83	14 (5,4)	21 (14,5) ^(h)		19 (18)
Placebo	86 (65-102)	82	14 (5,8)	21 (14,7) ^(h)		27 (26)
Tune 2003						
DON 10 mg	74 (62-83)	79	21 (3,7)	18 (12,4)	Unterschiede zwischen den Gruppen im NPI; in ANCOVA für Auswertung zum NPI berücksichtigt	0 (0)
Placebo	72 (53-92)	71	21 (4,1)	9 (9,8)		2 (14)

(Fortsetzung)

Tabelle 7 (Fortsetzung): Charakterisierung der Studienpopulationen Donepezil

Studie	Alter (Jahre)^(a)	Geschlecht w (%)	MMSE^(a)	NPI^(a)	Kommentar (zur Vergleichbarkeit etc.)	Studienabbrecher N (%)
Winblad 2001						
DON (5-) 10mg	72 (8,6)	70	19 (4,4)	13 (13,8)		47 (33)
Placebo	73 (8,0)	59	19 (4,5)	12 (12,2)		47 (33)
<p>Erläuterungen: DON = Donepezil, k.A. = keine Angaben, N = Anzahl, w = weiblich Da lediglich für eine Studie (Gauthier 2002) Baselinedaten zu ADL-Skalen vorliegen, werden diese hier nicht dargestellt.</p> <p>a: Mittelwert mit SD (= Standardabweichung) oder Spannweite b: eigene Berechnung der SD aus Standardfehler c: Angaben der Herstellerfirma Eisai im Rahmen des Stellungnahmeverfahrens (vgl. SN Eisai) d: Angaben liegen lediglich für die Per-Protokoll-Population vor. e: N berechnet aus Prozentangaben f: %-Angabe aus N berechnet g: Donepezil-Gruppe zur Baseline mit ADAS-cog = 36, Placebo-Gruppe zur Baseline mit ADAS-cog = 39 h: NPI-NH</p>						

5.2.1.2 Studien- bzw. Berichtsqualität

Insgesamt war die Studien- bzw. Berichtsqualität mäßig. Einzelheiten hierzu finden sich in Tabelle 8. Die daraus resultierende Bewertung der biometrischen Qualität ergab in acht Fällen „leichte Mängel“, in vier Fällen „grobe Mängel“ (s. a. Abschnitt 4.3). Nachfolgend werden die wesentlichen Aspekte zusammenfassend dargestellt.

Alle Studien wurden als randomisiert, doppelblind und Placebo-kontrolliert beschrieben. Das genaue Vorgehen bezüglich der Randomisierung und der Verdeckung der Gruppenzuteilung (Concealment of Allocation) wurde in den Publikationen häufig nicht dargestellt. Für die meisten der Studien, wo dies der Fall war, lieferte die Firma Eisai im Rahmen des Stellungnahmeverfahrens weitere Informationen (vgl. SN Eisai^b). Insgesamt wurden so Randomisierung und Concealment fast immer als adäquat bewertet.

Aus den Publikationen war meist nicht ersichtlich, ob und inwiefern die Erheber der Zielkriterien im Hinblick auf andere Daten der Patienten verblindet waren. Dies ist insbesondere deshalb relevant, da aufgrund der charakteristischen Nebenwirkungen der Cholinesterasehemmer ansonsten eine Postrandomisierungsentblindung möglich ist. Die Hersteller gaben im Stellungnahmeverfahren an, dass die Erhebung des CIBIC-plus jeweils durch einen unabhängigen Rater durchgeführt worden sei, der gegenüber allen anderen Studienergebnissen verblindet gewesen sei (SN Pfizer, S. 3). Bezüglich des ADAS-cog wurde beschrieben, dass dieser üblicherweise nicht durch den Studienleiter, sondern durch ein anderes Mitglied des Studienteams erhoben wurde, meist bevor andere Messungen durchgeführt wurden. Eine formale Verblindung der ADAS-cog-Rater sei deshalb nicht notwendig, insbesondere da diese Personen im Allgemeinen keinen Zugriff auf andere Untersuchungsergebnisse hätten (SN Eisai, S. 9, sowie Protokoll). Damit ist die Verblindung gegenüber unerwünschten Ereignissen sowie anderen Testergebnissen sicher nur für die Erhebung des CIBIC-plus gegeben. Für die Studien Moraes 2006 und Prasher 2002 sind weitere Einschränkungen entsprechend vermerkt.

Keine Angaben zur Fallzahlplanung fanden sich in den Publikationen zu den beiden größeren Studien von Seltzer 2004 und Homma 2000, für die allerdings im Rahmen des Stellungnahmeverfahrens entsprechende Informationen von der Firma Eisai nachgeliefert wurden (vgl. SN Eisai), sowie in den beiden kleineren Studien zur Bildgebung (Krishnan 2003, Tune 2003), die jedoch nur Pilotstudiencharakter hatten. In manchen Punkten unklar blieben Angaben zur Fallzahlplanung bei Moraes 2006, Rogers 1998 und Tariot 2001. Die weiteren Studien enthielten eine Fallzahlberechnung für eine primäre Zielgröße, zumeist den

^b Bei Verweisen auf Stellungnahmen wird in den Fällen, in dem diese von Herstellerfirmen eingereicht wurden, jeweils der Firmenname genannt. Bei Verweisen auf Stellungnahmen von Fachgesellschaften, Verbänden, Arbeitsgruppen o. Ä. wird jeweils die erstunterzeichnende Person genannt (vgl. Anhang I).

CIBIC-plus. In der Studie von Burns 1999 wurde aufgrund der hohen Varianz der Baselinedaten die Fallzahlschätzung während der laufenden Studie nach oben korrigiert. In keiner der Studien fanden sich Angaben zum Umgang mit dem Problem des multiplen Testens bei Betrachtung zahlreicher Zielgrößen (z. B. eine statistische α -Adjustierung oder eine Begründung für die Nichtnotwendigkeit einer solchen), auch dann nicht, wenn mehrere primäre Zielkriterien formuliert worden waren. Insbesondere für die Patientenpopulation, die der Studie von Gauthier 2002 zugrunde liegt, wurden in verschiedenen Publikationen Ergebnisse von 10 Erhebungsinstrumenten zumeist sowohl in einer kontinuierlichen als auch kategorialen Auswertungsform mit verschiedenen Responderdefinitionen ohne eine Korrektur für multiples Testen berichtet.

Angaben zu Studien- bzw. Therapieabbruchern, Abbruchgründen sowie ein Diagramm zum Patientenfluss waren bei fast allen Studien vorhanden.

In fast allen Studien wurde eine von den Autoren als ITT-Analyse bezeichnete LOCF-Auswertung durchgeführt. Dabei wurde meist mindestens eine Postbaseline-Messung zur Aufnahme in die ITT-Population gefordert, was aber größtenteils zum Ausschluss von nur wenigen (< 5 %) der randomisierten Patienten führte. In den meisten Studien wurde somit – unter der Annahme, dass die LOCF-Strategie hier eine angemessene Ersetzungsstrategie für fehlende Werte darstellt – eine „adäquate“ Intention-to-treat-Analyse durchgeführt. Eine als relevant zu bezeichnende Abweichung vom ITT-Prinzip^c war letztlich nur in zwei Studien erkennbar (Homma 2000, Prasher 2002). Bei Burns 1999 blieb es unklar, da lediglich im Methodikteil der Publikation berichtet wurde, dass eine Auswertung gemäß ITT-LOCF erfolgen sollte, sonst aber keine Angaben zu den tatsächlich ausgewerteten Patienten zu finden waren. Im Rahmen des Stellungnahmeverfahrens wurden Angaben zum Anteil nichtberücksichtigter Patienten bezüglich des ADAS-cog angegeben (vgl. SN Eisai), nicht jedoch für den CIBIC-plus, obwohl auch dieser einen primären Endpunkt darstellte. Zu der Studie Mohs 2001, bei der die Zeit bis zum Erreichen eines bestimmten Endpunktes betrachtet wurde, wurden von der Firma Eisai Angaben zur Zahl der Patienten unter Risiko nach Behandlungsbeginn geliefert (vgl. SN Eisai).

Die Studien wiesen kaum relevante Dateninkonsistenzen auf. Die Studie von Prasher 2002 enthielt widersprüchliche Angaben zur Anzahl von schwerwiegenden Nebenwirkungen.

^c Dies kann in aller Regel nur für die als primär bezeichneten Zielkriterien mit einiger Sicherheit beurteilt werden, und diese Beurteilung impliziert, dass das Vorgehen nach der LOCF-Methode als adäquate Ersetzungsmethode für die hier zu betrachtende klinische Situation und die verwendeten Zielkriterien angesehen wird.

Tabelle 8: Studien- und Publikationsqualität Donepezil

Studie	Randomisierung/ Concealment adäquat	Verblindete Erhebung der Zielgrößen	Fallzahl- planung adäquat	Studienabbrecher genannt / Abbruchgründe genannt	Relevante Abweichung vom ITT- Prinzip	Relevante Daten- inkonsistenz innerhalb der Publikation	Studien- und Publikationsqualität
Burns 1999	ja ^(a) /ja ^(a)	(ja) ^(b)	ja ^(c)	ja/ja	unklar ^(d)	nein	grobe Mängel
Gauthier 2002	ja ^(c) /ja ^(a)	(ja) ^(b)	ja ^(e,f)	ja/ja ^(g)	nein	nein	leichte Mängel
Homma 2000	ja ^(a) /ja ^(a)	(ja) ^(b)	ja ^(a)	ja ^(a) /ja ^(a)	ja ^(h)	nein	grobe Mängel
Krishnan 2003	ja/ja ^(a)	(ja) ^(b)	nein	ja/ja	nein	nein	leichte Mängel
Mohs 2001	ja ^(a) /ja	ja	ja	ja ^(a) /ja	(nein) ⁽ⁱ⁾	nein	leichte Mängel
Moraes 2006	unklar/unklar	(ja) ^(j)	unklar	unklar	nein	nein	grobe Mängel
Prasher 2002	ja/unklar	(ja) ^(k)	(ja) ^(l)	ja/ja	ja ^(m)	ja ⁽ⁿ⁾	grobe Mängel
Rogers 1998	ja/ja ^(a)	(ja) ^(o)	unklar	ja/(ja) ^(p)	nein	nein	leichte Mängel
Seltzer 2004	ja ^(a) /ja ^(a)	(ja) ^(b)	ja ^(a)	ja/ja	nein	nein	leichte Mängel
Tariot 2001	ja/ja ^(a)	(ja) ^(b)	unklar	ja/ja ^(a)	nein ^(q)	nein	leichte Mängel
Tune 2003	ja ^(a) /ja ^(a)	(ja) ^(b)	nein	ja/ja	nein ^(r)	nein	leichte Mängel

(Fortsetzung)

Tabelle 8 (Fortsetzung): Studien- und Publikationsqualität Donepezil

Studie	Randomisierung/ Concealment adäquat	Verblindete Erhebung der Zielgrößen	Fallzahl- planung adäquat	Studienabbrecher genannt / Abbruchgründe genannt	Relevante Abweichung vom ITT- Prinzip	Relevante Daten- inkonsistenz innerhalb der Publikation	Studien- und Publikationsqualität
Winblad 2001	ja/ja ^(a)	(ja) ^(b)	ja ^(a)	ja/ja	nein	nein	leichte Mängel
<p>a: Der Bewertung liegen Angaben der Firmen Eisai und Pfizer im Rahmen des Stellungnahmeverfahrens zugrunde (vgl. SN Eisai, Pfizer).</p> <p>b: Studie wird als doppelblind beschrieben, keine Angaben zur Art der Verblindung der Endpunkterfasser in der Publikation. Die Herstellerfirmen gaben im Stellungnahmeverfahren an, dass die Verblindung der CIBIC-Rater gegenüber anderen Studienergebnissen jeweils sichergestellt worden sei (vgl. SN Pfizer). Der ADAS-cog werde üblicherweise nicht durch den Studienleiter, sondern durch ein anderes Mitglied des Studienteams erhoben, meist bevor andere Messungen durchgeführt würden. Eine formale Verblindung der ADAS-cog-Rater sei deshalb nicht notwendig, insbesondere da diese Personen im Allgemeinen keinen Zugriff auf andere Untersuchungsergebnisse hätten. Es wurden keine Angaben zur Erhebung der weiteren Endpunkte gemacht (vgl. SN Eisai).</p> <p>c: Fallzahlplanung erfolgte für CIBIC-plus, nicht für 2. primären Parameter ADAS-cog; Nachrekrutierung aufgrund hoher Varianz der Baselinewerte im ADAS-cog um ca. 500 Patienten.</p> <p>d: In der Publikation wurde nicht angegeben, wie viele Patienten in die Auswertung der primären Endpunkte eingingen. Im Rahmen des Stellungnahmeverfahrens wurden von Eisai diese Angaben zum primären Zielkriterium ADAS-cog geliefert, jedoch nicht zum CIBIC-plus, dem zweiten primären Endpunkt (vgl. SN Eisai).</p> <p>e: Angaben hierzu liegen in der Publikation Feldman 2001 [53] vor.</p> <p>f: Fallzahlplanung erfolgte nur für CIBIC-plus und nicht für betrachtete Subgruppe (MMSE 10-17), sondern für die der Studie zugrunde liegende Gesamtpopulation (MMSE 5-17).</p> <p>g: Abbruchgründe für die betrachtete Subgruppe von der Firma Eisai zur Verfügung gestellt (vgl. SN Eisai).</p> <p>h: Insgesamt gingen 11 % der randomisierten Patienten nicht in die ITT-LOCF-Auswertung des primären Zielkriteriums ADAS-cog ein (Donepezil-Gruppe 7 %; Placebo-Gruppe 14 %).</p> <p>i: Angaben zu den Patienten unter Risiko im Studienverlauf wurden von der Firma Eisai zur Verfügung gestellt (vgl. SN Eisai). Die Angaben zu den verwendeten statistischen Analyseverfahren (Kaplan-Meier-Analyse, Life-Table-Methode) sind nicht vollständig nachzuvollziehen.</p> <p>j: Studie wird als doppelblind beschrieben, zur Verblindung der Endpunkterfasser ist lediglich die Verblindung bei der Erfassung der Schlafparameter erwähnt.</p>							

(Fortsetzung)

Tabelle 8 (Fortsetzung): Studien- und Publikationsqualität Donepezil

- k: In der Studie wird beschrieben, dass alle an der Studie Beteiligten während der ganzen Studienlaufzeit gegenüber der Gruppenzuteilung verblindet waren, jedoch wird nicht dargestellt, auf welche Art die Verblindung aufrechterhalten wurde.
- l: Die geplante Fallzahl wird in der Studie nicht erreicht.
- m: Insgesamt gingen 13 % der randomisierten Patienten nicht in die ITT-LOCF-Auswertung des primären Zielkriteriums DMR ein (Donepezil-Gruppe 13 %; Placebo-Gruppe 13 %).
- n: Dateninkonsistenzen bezüglich unerwünschter Ereignisse
- o: Studie wird als doppelblind beschrieben, zur Verblindung der Endpunkterfasser ist lediglich die Verblindung des CIBIC-plus-Raters gegenüber den Ergebnissen der psychometrischen Tests, der Laborwerte und der unerwünschten Ereignisse erwähnt.
- p: Abbruchgründe nur teilweise genannt
- q: Auf Anfrage berichtete die Herstellerfirma Eisai, dass in die Auswertung des primären Zielkriteriums NPI-NH in der Donepezil-Gruppe und in der Placebo-Gruppe jeweils 103 Patienten eingingen.
- r: Auch wenn formal das Kriterium der ITT-Verletzung erfüllt war (Nichtberücksichtigungsrate in der Verumgruppe $0/14 = 0\%$ und in der Placebogruppe $1/14 = 7\%$), wurde dieser Unterschied zwischen den Gruppen nicht als relevante Abweichung gewertet, da er auf die Nichtauswertung eines einzigen Patienten in der gesamten Studie zurückzuführen ist.

5.2.2 Galantamin

5.2.2.1 Studiendesign, Dosierungen und Studienpopulationen

Sechs Studien zum Vergleich von Galantamin mit Placebo erfüllten die Einschlusskriterien für die vorliegende systematische Übersicht (Brodaty 2005, Erkinjuntti 2002, Raskind 2000, Rockwood 2006, Tariot 2000, Wilcock 2000; vgl. Tabelle 4). Design und untersuchte Patientenpopulationen der einbezogenen Studien sind in Tabelle 9 und Tabelle 11 dargestellt. Zusätzlich zu den Publikationen wurden vom Hersteller Janssen-Cilag Daten zur Verfügung gestellt. Diese nicht publizierten Daten werden jeweils gesondert gekennzeichnet. Im Folgenden werden wesentliche Aspekte der Studien zusammenfassend beschrieben.

Alle Studien waren multizentrisch angelegt. Die kontrollierte Beobachtungszeit in den eingeschlossenen Studien (randomisierte Phasen) variierte zwischen 4 und 6 Monaten. Bei einigen Studien schloss sich eine Nachbeobachtungsphase mit unverblindeter Verummedikation für alle interessierten Teilnehmer der verblindeten Studienphase an (open label follow-up studies). Daten dieser unkontrollierten Studienphase wurden für die Nutzenbewertung nicht herangezogen. In den Studien wurden Dosierungen von 8 mg, 16 mg, 24 mg und 32 mg Galantamin untersucht, wobei die Dosis der unretardierten Darreichungsform immer auf zwei Tabletten verteilt morgens und abends gegeben wurde. Da die Dosierung von 32 mg Galantamin in Deutschland für die Indikation Alzheimer Demenz nicht zugelassen ist, werden die entsprechenden Ergebnisse hier jeweils nicht dargestellt. Die am häufigsten verwendete Dosierung war 24 mg/Tag. Die Medikation wurde unterschiedlich rasch gesteigert, mit Erhöhungen um 8 mg pro Woche (Wilcock 2000, Raskind 2000) bzw. 4 mg pro Woche (Erkinjuntti 2002) bis 8 mg alle 4 Wochen (Brodaty 2005, Tariot 2000, Rockwood 2006). In zwei Studien war eine flexible Dosierung erlaubt (16 oder 24 mg Galantamin) (Brodaty 2005, Rockwood 2006). Bei Brodaty 2005 wurde sowohl die zweimal tägliche Gabe einer Tablette mit konventioneller Galenik als auch ein Galantamin-Präparat mit verzögerter Freisetzung untersucht, das eine morgendliche Einmalgabe ermöglichte (Galantamine prolonged release, GAL-PR).

Die Teilnehmerzahlen schwankten zwischen 130 und 978 pro Studie. Insgesamt wurden ca. 2000 Patienten mit Galantamin und 1200 Patienten mit Placebo behandelt.

Fünf Studien schlossen Patienten mit einer wahrscheinlichen Alzheimer Demenz entsprechend den NINCDS-ADRDA-Kriterien ein (Tabelle 10). Eine Ausnahme stellt die Studie von Erkinjuntti 2002 dar, in der sowohl Patienten mit einer wahrscheinlich vaskulär bedingten Demenz (VaD) als auch solche mit einer Alzheimer Demenz und zusätzlicher zerebrovaskulärer Erkrankung (AD-CVD) eingeschlossen wurden. Für diese letztgenannte Gruppe mit „Mischdemenz“ (AD-CVD) lag sowohl die Diagnose einer möglichen Alzheimer Demenz nach den NINCDS-ADRDA-Kriterien als auch eine zerebrale Bildgebung mit

signifikanter zerebrovaskulärer Schädigung vor. Diese AD-CVD-Gruppe überschneidet sich nach Angaben der Autoren nicht mit AD-Patienten aus anderen Studien zu Galantamin, in denen aufgrund der Bildgebung ein Ausschluss bei signifikanter vaskulärer Schädigung erfolgte. Einige Auswertungen zu der AD-CVD-Subgruppe wurden gesondert in der Publikation ausgewiesen und hier für die Übersicht herangezogen. Weitere, nicht veröffentlichte Daten dieser Studie wurden von der Firma Janssen-Cilag auf Anfrage sowie im Rahmen des Stellungnahmeverfahrens übermittelt (vgl. SN Janssen-Cilag). Bei der Interpretation dieser Ergebnisse ist zu berücksichtigen, dass möglicherweise Unterschiede im Krankheitsverlauf zwischen Patienten mit zusätzlicher zerebrovaskulärer Demenz und reiner Alzheimer Demenz bestehen, z. B. erhöhte Mortalität und Morbidität sowie raschere Progression der Erkrankung [103].

In alle Studien wurden Patienten mit leichter bis mittelschwerer Alzheimer Demenz eingeschlossen. Der Schweregrad der Demenz wurde mit der MMSE-Skala gemessen, der Einschlussbereich umfasste einen Skalenbereich von 10-25 Punkten. Gleichzeitig wurde eine Mindestfehlerzahl im ADAS-cog (≥ 12 bzw. ≥ 18) gefordert. Dadurch wurde sichergestellt, dass auch Patienten in den guten Leistungsbereichen (MMSE 24/25) bereits ein Mindestmaß an kognitiven Einschränkungen aufwiesen.

Die sonstigen Ein- und Ausschlusskriterien variierten leicht, im Allgemeinen wurden auch Patienten mit Begleiterkrankungen wie Hypertonie, Herzinsuffizienz, Typ-2-Diabetes eingeschlossen, wenn die jeweilige Erkrankung gut kontrolliert war. Patienten mit Magengeschwüren, Harnabflussstörungen sowie Herzerkrankungen, die zu einer Synkope führen konnten, wurden in einem Teil der Studien explizit ausgeschlossen. Patienten mit anderen bedeutenden psychiatrischen oder neurologischen Erkrankungen wurden generell nicht eingeschlossen. Eine bestehende psychiatrische Komedikation war zumeist erlaubt.

Entsprechend der Spannweite des Einschluss-MMSE-Scores in den einzelnen Studien unterschieden sich der mittlere MMSE-Score und das mittlere Alter zwischen den Studien nur leicht (Tabelle 11). In der Studie Erkinjuntti 2002 waren die Patienten in der Placebo-Gruppe hinsichtlich der Aktivitäten des täglichen Lebens (DAD) zur Baseline etwas schwerer beeinträchtigt (66 Punkte vs. 71 Punkte in der Galantamin-Gruppe). Es war unklar, ob dieser Unterschied in der Auswertung berücksichtigt wurde. Auch bei Rockwood 2006 wiesen die Patienten der Placebo-Gruppe eine schwerere Beeinträchtigung im DAD und ADAS-cog (27,9 in der Placebo-Gruppe vs. 24,2 in der GAL-Gruppe) auf. Bei Berücksichtigung dieser Unterschiede veränderte sich das Ergebnis nicht.

In fünf Studien waren ADAS-cog und CIBIC-plus primäre Zielkriterien (Tabelle 9), lediglich bei Rockwood 2006 stellte der Grad der Erreichung von individuell festgelegten Zielen, gemessen mit der Goal Attainment Scale (GAS), das primäre Zielkriterium dar. Weitere, sekundäre Endpunkte waren die alltagspraktischen Fähigkeiten, für deren Messung zwei unterschiedliche Skalen verwendet wurden (zwei Studien ADCS-ADL, vier Studien DAD).

Die begleitende Psychopathologie wurde in drei Studien unter Verwendung des NPI erfasst. Aus zwei Studien wurden zusammengefasste Daten zu Pflege- und Betreuungszeiten nachträglich veröffentlicht [98]. In einer Studie wurde die Belastung der betreuenden Personen anhand der Caregiving Burden Scale (CBS) gemessen (Rockwood 2006).

Die sechs Studien sind insgesamt in Bezug auf die Methodik und die Merkmale der Studienteilnehmer vergleichbar.

Tabelle 9: Eingeschlossene Studien zu Galantamin

Studie (Weitere Studien- bezeichnungen)	Studiendesign	Studiendauer (kontrollierte Dauer der Verumgabe)	Interventionsgruppen und Zahl der randomisierten Patienten	Ort der Durchführung und Zahl der Zentren	Relevante Zielkriterien^(a)
Brodaty 2005 (GAL-INT-10)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Galantamin Prolonged Release (PRC) 16-24 mg: n=320 2. Galantamin nichtretadiert (GAL) 16- 24 mg: n=327 3. Placebo: n=324	USA (66 Zentren), CAN, AUS, ZA, NZ (Gesamt: 93 Zentren)	ADAS-cog, CIBIC-plus ADCS-ADL, NPI, NPI-D, unerwünschte Ereignisse
Erkinjuntti 2002 ^(b) (GAL-INT-6)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Galantamin 24 mg: n=188 2. Placebo: n=97	CAN, DK, FIN, F, D, IRL, NL, Israel, PL, GB (Gesamt: 62 Zentren)	ADAS-cog, CIBIC-plus NPI, NPI-D, DAD, unerwünschte Ereignisse
Raskind 2000 (GAL-USA-1)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Galantamin 24 mg ^(c) : n=212 2. Placebo: n=213	USA (33 Zentren)	ADAS-cog, CIBIC-plus DAD, unerwünschte Ereignisse
Rockwood 2006	RCT, parallel, doppelblind, multizentrisch	4 Monate	1. Galantamin 16-24 mg: n=64 2. Placebo: n=66	CAN (14 Zentren)	GAS ADAS-cog, CIBIC-plus, DAD, CBS
Tariot 2000 (GAL-USA-10)	RCT, parallel, doppelblind, multizentrisch	5 Monate	1. Galantamin 8 mg: n=140 2. Galantamin 16 mg: n=279 3. Galantamin 24 mg: n=273 4. Placebo: n=286	USA (keine Angaben zur Zahl der Zentren)	ADAS-cog, CIBIC-plus ADCS-ADL, NPI, NPI-D, unerwünschte Ereignisse
Wilcock 2000 (GAL-INT-1)	RCT, parallel, doppelblind, multizentrisch	6 Monate	1. Galantamin 24 mg ^(c) : n=220 2. Placebo: n=215	CAN, NOR, FIN, F, D, S, NL, GB (Gesamt: 86 Zentren)	ADAS-cog, CIBIC-plus DAD, unerwünschte Ereignisse
<p>a: primäre sowie relevante Zielkriterien gemäß Abschnitt 4.1.3 (falls in der jeweiligen Studie als „primär“ deklariert Fettdruck)</p> <p>b: Subpopulation der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung (AD-CVD), Daten direkt vom Hersteller (Janssen-Cilag)</p> <p>c: Die Studie beinhaltete auch eine Interventionsgruppe mit 32 mg, die jedoch wegen fehlender Zulassung hier nicht dargestellt wird.</p>					

Tabelle 10: Wesentliche Ein- und Ausschlusskriterien der Studien zu Galantamin

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Brodaty 2005	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-24 - ADAS-cog: ≥ 18 - progressiver Abbau von kognitiven Fähigkeiten seit ≥ 6 Monaten - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankungen, kognitive Einschränkung anderer Ursache, vaskuläre Demenz - klinisch signifikante psychiatrische, hepatische, renale, pulmonale, metabolische, endokrine oder kardiovaskuläre Erkrankungen - Harnabflussstörung, Magen-/Darmgeschwür - Epilepsie
Erkinjuntti 2002	<p>Subpopulation AD+CVD:</p> <ul style="list-style-type: none"> - leichte bis mittelschwere AD (möglich) - CVD (wahrscheinlich; neuroradiologische Bildgebung) - MMSE: 10-25 - ADAS-cog: ≥ 12 - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankungen, kognitive Einschränkung anderer Ursache - kardiovaskuläre Erkrankung, die Studienabschluss infrage stellt - klinisch signifikante psychiatrische, hepatische, renale, pulmonale metabolische, endokrine oder kardiovaskuläre Erkrankungen - aktives Magen-/Darmgeschwür - Epilepsie oder früherer Substanzmissbrauch
Raskind 2000	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 11-24 - ADAS-cog: ≥ 12 - progressiver Abbau von kognitiven Fähigkeiten seit ≥ 6 Monaten - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankungen, zerebrovaskuläre Erkrankung - kardiovaskuläre Erkrankung, die Studienabschluss infrage stellt - klinisch signifikante psychiatrische, hepatische, renale, pulmonale, metabolische oder endokrine Erkrankungen - Harnabflussstörung, Magen-/Darmgeschwür - früherer Substanzmissbrauch oder Epilepsie <p>Stabile und gut eingestellte Begleiterkrankungen wie z. B. Bluthochdruck, nicht insulinpflichtiger Diabetes oder Schilddrüsenunterfunktion waren zulässig.</p>
Rockwood 2006	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - progressiver Abbau von kognitiven Fähigkeiten seit ≥ 6 Monaten - MMSE: 10-25 - ADAS-cog: ≥ 18 - Betreuungsperson 	<ul style="list-style-type: none"> - andere mögliche Demenzursachen - in Pflegeheim wohnend - relevante Kommunikationsschwierigkeiten (Sprach-, Hör-, Sehvermögen) - andere Erkrankungen

(Fortsetzung)

Tabelle 10 (Fortsetzung): Wesentliche Ein- und Ausschlusskriterien der Studien zu Galantamin

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Tariot 2000	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-22 - ADAS-cog: ≥ 18 - CT oder MRT - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankungen, zerebrovaskuläre Erkrankung - kardiovaskuläre Erkrankung, die Studienabschluss infrage stellt - klinisch signifikante psychiatrische, hepatische, renale, pulmonale, metabolische oder endokrine Erkrankungen - Harnabflussstörung, Magen-/Darmgeschwür - früherer Substanzmissbrauch oder Epilepsie <p>Stabile und gut eingestellte Begleiterkrankungen wie z. B. Bluthochdruck, nicht insulinpflichtiger Diabetes oder Schilddrüsenunterfunktion waren zulässig.</p>
Wilcock 2000	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - progressiver Abbau von kognitiven Fähigkeiten seit ≥ 6 Monaten - MMSE: 11-24 - ADAS-cog: ≥ 12 - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankungen, Multiinfarkt-Demenz oder zerebrovaskuläre Erkrankung - kardiovaskuläre Erkrankung, die Studienabschluss infrage stellt - klinisch signifikante psychiatrische, hepatische, renale, pulmonale, metabolische oder endokrine Erkrankungen - Harnabflussstörung, Magen-/Darmgeschwür - früherer Substanzmissbrauch oder Epilepsie <p>Stabile und gut eingestellte Begleiterkrankungen wie z. B. Bluthochdruck, nicht insulinpflichtiger Diabetes oder Schilddrüsenunterfunktion waren zulässig.</p>
<p>a: über die üblichen Kontraindikationen (z. B. Wirkstoffunverträglichkeit) bzw. Ausschlusskriterien (z. B. konkurrierende medikamentöse Begleittherapien) hinausgehend</p>		

Tabelle 11: Charakterisierung der Studienpopulationen Galantamin

Studie	Alter (Jahre) ^(a)	Geschlecht w (%)	MMSE ^(a)	NPI ^(a)	ADL ^(a)	Kommentar (zur Vergleichbarkeit etc.)	Studien-abbrecher N (%)
Brodaty 2005 ^(b)							
GAL PRC 16-24 mg	77 (7,6)	205 (64)	18 (4,0)	11 (13,8) ^(c,d)	53 (15,9) ^(c,d,e,f)		69 (22)
GAL 16-24 mg	77 (7,8)	208 (64)	18 (4,1)	13 (13,6) ^(c,d)	52 (16,4) ^(c,d,e,f)		76 (23)
Placebo	76 (8,0)	205 (64)	18 (4,1)	10 (11,7) ^(c,d)	55 (15,6) ^(c,d,e,f)		58 (18)
Erkinjuntti 2002 ^(c,g)							
GAL 24 mg	76 (6,8)	99 (53)	21 (4,0)	11 (11,6)	71 (23,6) ^(d,h)	<i>Patienten der Placebo-Gruppe hins. ADL etwas schwerer beeinträchtigt</i>	34 (18) ⁽ⁱ⁾
Placebo	78 (5,9)	49 (51)	20 (3,6)	11 (11,8)	66 (24,3) ^(d,h)		10 (10) ⁽ⁱ⁾
Raskind 2000							
GAL 24 mg	76 (7,3)	139 (66)	20 (4,4)	-	71 (21,8) ^(h)		68 (32)
Placebo	75 (8,8)	131 (62)	19 (4,4)	-	70 (23,4) ^(h)		41 (19)
Rockwood 2006							
GAL 16-24 mg	77 (8)	41 (64)	21 (3,3)	-	76 (19,7) ^(h)	<i>Patienten der Placebo-Gruppe schwerer beeinträchtigt⁽ⁱ⁾</i>	11 (17)
Placebo	78 (8)	41 (62)	20 (4,2)	-	71 (21,4) ^(h)		10 (15)
Tariot 2000							
GAL 8 mg	76 (7,1)	90 (64)	18 (3,5)	13 (14,2)	54 (14,2) ^(e)		32 (23)
GAL 16 mg	76 (8,4)	174 (62)	18 (3,3)	12 (13,4)	52 (15,0) ^(e)		60 (22)
GAL 24 mg	78 (6,6)	183 (67)	18 (3,3)	12 (13,2)	52 (16,5) ^(e)		61 (22)
Placebo	77 (8,5)	178 (62)	18 (3,4)	11 (11,8)	52 (15,2) ^(e)		46 (16)
Wilcock 2000							
GAL 24 mg	72 (8,3)	139 (63)	20 (3,4)	-	70 (21,4) ^(h)		44 (20)
Placebo	73 (7,6)	132 (61)	19 (3,5)	-	67 (22,5) ^(h)		29 (13)

(Fortsetzung)

Tabelle 11 (Fortsetzung): Charakterisierung der Studienpopulationen Galantamin

Erläuterungen:

GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, N = Anzahl, w = weiblich

a: Mittelwert mit SD (= Standardabweichung) oder Spannweite

b: Die Angaben zur Studienpopulation beziehen sich auf die Gruppe der randomisierten und behandelten Patienten (GAL-PRC: n=319, GAL: n=326, Placebo: n=320)

c: Daten direkt von Herstellerfirma Janssen-Cilag

d: Diese Angaben liegen nur für die ITT-LOCF-Population vor.

e: ADCS-ADL

f: Standardabweichung aus Standardfehler errechnet mit Anzahl der randomisierten Patienten, da keine Angaben zur Anzahl der ITT-LOCF-Population vorlagen.

g: Subpopulation der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung

h: DAD

i: Angaben der Herstellerfirma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens (vgl. SN Janssen-Cilag)

j: Unterschied wird in der Auswertung berücksichtigt und verändert das Ergebnis nicht

5.2.2.2 Studien- bzw. Berichtsqualität

Es gab geringfügige Unterschiede in der methodischen Qualität und der Berichtsqualität zwischen den Studien, bis auf eine Ausnahme. Insgesamt kann die Qualität als befriedigend bezeichnet werden. Einzelheiten hierzu sind in Tabelle 12 dargestellt. In fünf Fällen wurde die biometrische Qualität mit „leichte Mängel“ bewertet, in einem Fall mit „grobe Mängel“. Im Folgenden werden die wesentlichen Aspekte zur Studien- bzw. Berichtsqualität zusammenfassend dargestellt.

Alle Studien wurden als randomisiert beschrieben. Die meisten Publikationen berichteten von einer computergenerierten Randomisierung. Inwieweit die Gruppenzuteilung verdeckt wurde (Concealment of Allocation), blieb in der Hälfte der Studien anhand der Veröffentlichungen unklar. Aus zusätzlichen Unterlagen der Herstellerfirma ging für diese Studien hervor, dass das Concealment jeweils gewährleistet war (vgl. SN Janssen-Cilag). Bei Rockwood 2006 war die verdeckte Gruppenzuteilung formal durch eine zentrale Randomisierungsprozedur gewährleistet – allerdings ist hier denkbar, dass die ungewöhnlich kleinen 2er-Randomisierungsblöcke in Zusammenhang mit den typischen cholinergen Nebenwirkungen zu einer teilweisen Aufdeckung der Gruppenzuteilung geführt haben könnten.

Alle Studien wurden in den Publikationen als doppelblind bezeichnet. Die Verblindung der Patienten sowie der Behandler erschien in allen Studien adäquat. Eine Verblindung der CIBIC-Endpunkterfasser wurde bei Wilcock 2001 und Erkinjuntti 2002 explizit erwähnt. Für die Studie Tariot 2000 ging aus zusätzlichen Unterlagen der Herstellerfirma hervor, dass der CIBIC durch eine unabhängige Person beurteilt wurde; somit wurde die Verblindung für diesen Endpunkt als gesichert angesehen. Die ADAS-cog-Rater sollten ebenfalls möglichst wenig in die Behandlung involviert sein – dies war allerdings offenbar nicht immer möglich, so dass in diesen Fällen die Erhebung vor der Abfrage der unerwünschten Ereignisse durchgeführt wurde (vgl. SN Janssen-Cilag). Bei Rockwood 2006 wurde darauf hingewiesen, dass die Personen, die sowohl den GAS als auch den CIBIC bewerteten, gegenüber anderen Studienergebnissen und Nebenwirkungen verblindet waren. In den anderen Publikationen wurde das Risiko, dass Endpunkterfasser durch Kenntnisse der typischen cholinergen Nebenwirkungen im Studienverlauf entblindet werden können, nicht problematisiert bzw. diskutiert.

Alle Studien enthielten eine Fallzahlberechnung auf der Basis von früheren Wirksamkeitsstudien. Die Angaben hierzu waren für die meisten Studien vollständig. Die geplanten Fallzahlen wurden zumeist erreicht, nur bei Rockwood 2006 wurden statt der geplanten 152 lediglich 130 Patienten in die Studie eingeschlossen. Eine Begründung dafür findet sich in der Publikation nicht.

In allen Studien wurden eine Last-Observation-Carried-Forward-Auswertung vorgenommen, die als Intention-To-Treat-Analyse bezeichnet wurde (ITT-LOCF), wobei üblicherweise eine Post-Baseline-Erfassung gefordert wurde. Lediglich bei Wilcock 2000 genügte eine Messung während der Baseline-Erhebung, um in die ITT-Analyse aufgenommen zu werden. Alle Studien führten darüber hinaus eine Observed-Cases-Auswertung (OC) durch. In der Veröffentlichung zur AD-CVD-Subgruppe der Erkinjuntti 2002-Studie finden sich keine Angaben zur ITT-Population, sondern ausschließlich zu den „Observed Cases“. Für den vorliegenden Bericht wurden daher bislang unveröffentlichte Daten der Firma Janssen-Cilag verwendet, in denen auch die ITT-LOCF-Analysen für die relevante Subpopulation vorlagen. In dieser Studie betrug der Unterschied in der Nichtberücksichtigungsrate im Rahmen der ITT-LOCF-Auswertung des primären Zielkriteriums ADAS-cog zwischen den Gruppen 5 Prozentpunkte (Galantamin-Gruppe 6 %; Placebo-Gruppe 1 %). Im Rahmen des Stellungnahmeverfahrens wurden von der Herstellerfirma zusätzliche Angaben zur ITT-Analyse, bei der auch Patienten ohne Post-Baseline-Messung ausgewertet wurden, zur Verfügung gestellt. Hier lagen die Nichtberücksichtigungsraten für den ADAS-cog lediglich bei 3 % (Galantamin-Gruppe) bzw. 1 % (Placebo-Gruppe) und für den CIBIC-plus bei 9 % bzw. 5 % (vgl. SN Janssen-Cilag)^d. In der Studie Rockwood 2006 gingen in die Auswertung des GAS-Scores, des primären Endpunktes, alle Placebopatienten, jedoch nur 61 der 64 Galantamin-Patienten ein. Diese Differenz in der Berücksichtigungsrate von 5 Prozentpunkten wurde als relevante Abweichung vom ITT-Prinzip betrachtet.

Für fünf der sechs Studien waren in den Publikationen ausreichend Angaben zu der Anzahl der Studienabbrecher und den Gründen für den Studienabbruch zu finden (Brodaty 2005; Raskind 2000; Rockwood 2006; Tariot 2000; Wilcock 2000). Für die Erkinjuntti 2002-Studie lagen in der Veröffentlichung diese Informationen zur Gesamtpopulation vor, nicht jedoch für die betrachtete Subgruppe. Für diese stellte die Firma Janssen-Cilag die entsprechenden Angaben zur Verfügung (vgl. SN Janssen-Cilag).

Die Studien wiesen keine relevanten Dateninkonsistenzen auf. Widersprüchliche Angaben zu den Effekten auf den ADAS-cog durch die Vertauschung eines Vorzeichens bei Raskind 2000 konnten durch die Unterlagen der Herstellerfirma geklärt werden.

^d Im Folgenden werden zu der Studie Erkinjuntti 2002 zur besseren Vergleichbarkeit mit den übrigen Galantamin-Studien jeweils die Ergebnisse der ITT-LOCF-Analysen dargestellt. Die Ergebnisse waren mit denen der ITT-Analyse vergleichbar (vgl. SN Janssen-Cilag).

Tabelle 12: Studien- und Publikationsqualität Galantamin

Studie	Randomisierung/ Concealment adäquat	Verblindete Erhebung der Zielgrößen	Fallzahl- planung adäquat	Studienabbrucher genannt/Abbruch gründe genannt	Relevante Abweichung vom ITT- Prinzip	Relevante Daten- inkonsistenz innerhalb der Publikation	Studien- und Publikationsqualität
Brodaty 2005	ja/ja	(ja) ^(a)	ja ^(b)	ja/ja	nein	nein	leichte Mängel
Erkinjuntti 2002	ja/ja ^(b)	(ja) ^(c)	ja ^(b,d)	ja ^(b) /ja ^(b)	nein ^(b,e)	nein	leichte Mängel
Raskind 2000	ja/ja ^(b)	(ja) ^(f)	ja ^(g)	ja/ja	nein	nein	leichte Mängel
Rockwood 2006	unklar/ja	(ja) ^(h)	(ja) ⁽ⁱ⁾	ja/ja	ja ⁽ⁱ⁾	nein	grobe Mängel
Tariot 2000	ja/ja ^(b)	(ja) ^(k)	ja ^(g)	ja/ja	nein ^(l)	nein	leichte Mängel
Wilcock 2000	ja/ja	(ja) ^(c)	ja ^(g)	ja/ja	nein	nein	leichte Mängel

(Fortsetzung)

Tabelle 12 (Fortsetzung): Studien- und Publikationsqualität Galantamin

- a: In der Studie wird beschrieben, dass alle an der Studie Beteiligten während der ganzen Studienlaufzeit gegenüber der Gruppenzuteilung verblindet waren, jedoch wird nicht dargestellt, auf welche Art die Verblindung aufrechterhalten wurde.
- b: Der Bewertung liegen Angaben der Firma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens zugrunde (vgl. SN Janssen-Cilag).
- c: Studie wird als doppelblind beschrieben; zur Verblindung der Endpunkterheber ist lediglich beschrieben, dass die Bewertung im CIBIC-plus durch einen Rater vorgenommen wurde, der gegenüber anderen Aspekten der Studie verblindet war.
- d: Fallzahlplanung erfolgte für die der Studie zugrunde liegenden Gesamtpopulation, nicht für die Subgruppe AD+CVD.
- e: Bewertung bezieht sich auf die Daten zur „klassischen“ ITT-Analyse, deren Ergebnisse mit der ITT-LOCF-Analyse vergleichbar waren.
- f: In der Studie wird beschrieben, dass die Untersucher während der ganzen Studienlaufzeit gegenüber der Gruppenzuteilung verblindet waren, jedoch wird nicht dargestellt, auf welche Art die Verblindung aufrechterhalten wurde.
- g: Fallzahlplanung erfolgte ausschließlich für ADAS-cog, obwohl auch CIBIC-plus ein primäres Zielkriterium war. Es fehlen Angaben zur Variabilität.
- h: Die Person, die das Interview mit Patienten und Angehörigen zur Erfassung der Zielerreichung führte, war gegenüber unerwünschten Ereignissen (UEs) und allen anderen Studienergebnissen (außer CIBIC-plus, den sie ebenfalls erhob) verblindet; der behandelnde Arzt war gegenüber den UEs meist nicht verblindet.
- i: Die geplante Fallzahl wird in der Studie nicht erreicht.
- j: Insgesamt gingen 2 % der randomisierten Patienten nicht in die ITT-LOCF-Auswertung des primären Zielkriteriums GAS ein; der Gruppenunterschied betrug 5 Prozentpunkte (Galantamin-Gruppe 5 %; Placebo-Gruppe 0 %).
- k: Studie wird als doppelblind beschrieben; keine Angaben zur Verblindung der Endpunkterheber in der Publikation. Die Herstellerfirma gab im Stellungnahmeverfahren an, dass der CIBIC jeweils durch einen unabhängigen Rater erhoben worden sei. Der ADAS-cog-Rater habe nicht in die Behandlung eingebunden sein und keinen Zugang zu Angaben bzgl. der unerwünschten Ereignisse haben sollen. Falls dies nicht möglich war, führte der Rater den ADAS-Test durch, bevor er die unerwünschten Ereignisse erfragte. Es wurden keine Angaben zur Erhebung der weiteren Endpunkte gemacht (vgl. SN Janssen-Cilag).
- l: 40 Patienten eines Studienzentrums wurden aus der Wirksamkeitsanalyse wegen Verstößen des dortigen Studienleiters gegen das Studienprotokoll ausgeschlossen.

5.2.3 Rivastigmin

5.2.3.1 Studiendesign, Dosierungen und Studienpopulationen

Zum Vergleich von Rivastigmin mit Placebo wurden sechs relevante Studien identifiziert, von denen drei veröffentlichte (Corey-Bloom 1998, Forette 1999 und Rösler 1999) sowie eine unveröffentlichte Studie (B304) in die Nutzenbewertung einbezogen werden konnten. Die Daten zu der nicht publizierten Studie wurden vom Hersteller Novartis zur Verfügung gestellt und werden jeweils gesondert gekennzeichnet.

Einzelheiten zum Design und zu den allgemeinen Charakteristika der vorliegenden vier Studien sind in Tabelle 13 dargestellt. Die wesentlichen Aspekte der Studien werden im Folgenden zusammenfassend beschrieben.

Alle Studien waren multizentrisch angelegt. Die Studiendauer betrug 18 Wochen bei Forette 1999, in den anderen Studien 26 Wochen. Alle Studien umfassten jeweils drei Studienarme. Die Studien Corey-Bloom 1998 und Rösler 1999 untersuchten jeweils eine Niedrigdosis-Gruppe mit 1 bis 4 mg Rivastigmin und eine Hochdosis-Gruppe mit 6 bis 12 mg täglich im Vergleich zu Placebo. Die Dosis wurde jeweils wöchentlich in den ersten 7 (Corey-Bloom 1998) bzw. 12 (Rösler 1999) Wochen bis zum Zieldosisbereich erhöht; in dieser Zeit war keine Dosisreduktion vorgesehen. Anschließend konnte eine weitere Dosiserhöhung bis zur oberen Grenze oder Dosiserniedrigung bis zur unteren Grenze der jeweiligen Gruppe erfolgen. Falls in der Aufdosierungsphase die untere Zieldosis der jeweiligen Gruppe nicht erreicht wurde, erfolgte bei Corey-Bloom 1998 der Ausschluss der betroffenen Patienten aus der Studie. Im Falle von Nebenwirkungen konnte bei Rösler 1999 eine Dosis ausgelassen oder ein Antiemetikum hinzugegeben werden; für die Studie Corey-Bloom 1998 gibt es diesbezüglich keine Angaben. In der Studie B304 betrug die tägliche Dosis Rivastigmin je nach Verträglichkeit für den einzelnen Patienten zwischen 2 und 12 mg, wobei jeweils eine möglichst hohe Dosierung angestrebt wurde. In einer Gruppe wurde Rivastigmin in zwei täglichen Dosen, in der anderen Gruppe in drei täglichen Dosen verabreicht. Die Titrationsphase dauerte zwischen 10 Tage und 12 Wochen, abhängig davon, wie hoch die vom Patienten vertragene Dosierung war. Auch bei Forette 1999 wurde Rivastigmin in zwei oder in drei täglichen Dosen gegeben. Die tägliche Mindestdosis lag hier jedoch bei 6 mg, die maximale Dosis bei 12 mg täglich. Rivastigmin wurde stets bis zur höchstverträglichen Dosis aufdosiert. Bei Nebenwirkungen wurde ein Antiemetikum gegeben. Patienten, die nicht mindestens 6 mg tolerieren konnten, wurden im Verlauf der Studie ausgeschlossen.

In den Studien B304, Corey-Bloom 1998 und Rösler 1999 wurden jeweils um die 700 Patienten eingeschlossen. Die Studie Forette 1999, die auch als Phase-II-Studie bezeichnet wurde, war mit 114 Patienten deutlich kleiner. Insgesamt wurden etwa 1500 Patienten mit Rivastigmin und 700 Patienten mit Placebo behandelt.

In allen Studien wurden ausschließlich Patienten mit einer wahrscheinlichen Alzheimer Demenz nach den NINCDS-ADRDA-Kriterien eingeschlossen, mit einem leichten bis mittleren Schweregrad. Der Schweregrad wurde mit der MMSE-Skala gemessen. In den Studien B304, Corey-Bloom 1998 und Rösler 1999 wurden Patienten mit einem MMSE-Wert von 10 bis 26 eingeschlossen, bei Forette 1999 lag der MMSE-Wert zwischen 12 und 26. Ausschlussgründe waren insbesondere schwerwiegende bzw. nicht eingestellte Erkrankungen (Tabelle 14).

Die in den vier Studien untersuchten Populationen waren homogen bezüglich des Krankheitsbildes. Einzelheiten zu den Merkmalen der in die Studien eingeschlossenen Patienten sind in Tabelle 15 dargestellt. Innerhalb der Studien waren die Therapiegruppen bezogen auf die Basischarakteristika der Patienten – soweit beurteilbar – vergleichbar, auch wenn bei Corey-Bloom 1998 der Anteil der weiblichen Patienten in der Hochdosis-Gruppe etwas größer war und bei Forette 1999 die Patienten in der Placebo-Gruppe etwas älter waren als die in den Verum-Gruppen. Bei Rösler 1999 lagen Angaben zu den Basischarakteristika lediglich für die gesamte Studienpopulation, nicht jedoch für die einzelnen Gruppen vor.

Die Abbrecherraten betragen in den Placebo-Gruppen zwischen 8 und 15 %, in den Rivastigmin-Gruppen zwischen 14 % (Rösler 1999, Niedrigdosis-Gruppe) und 36 % (Forette 1999, Verabreichung 2x täglich).

In allen drei größeren Studien (B304, Corey-Bloom 1998, Rösler 1999) wurden der ADAS-cog sowie der CIBIC-plus untersucht und zusätzlich die PDS, der MMSE sowie die GDS erhoben. Bei Forette 1999 wurde neben dem ADAS-cog und dem CIBIC-plus die NOSGER eingesetzt. In allen vier Studien wurde der CIBIC-plus dahingehend ausgewertet, wie groß der Anteil der Patienten mit einer „mindestens minimalen Verbesserung“ ($\text{Score} \leq 3$) ist, statt, wie in den Studien zu anderen Cholinesterasehemmern zumeist üblich, nach dem Anteil der Patienten mit „keiner Verschlechterung oder Verbesserung“ ($\text{Score} \leq 4$).

Die drei größeren Studien waren in Bezug auf die Methodik und die Merkmale der Studienteilnehmer vergleichbar.

Tabelle 13: Eingeschlossene Studien zu Rivastigmin

Studie (Weitere Studien- bezeichnungen)	Studiendesign	Studiendauer (kontrollierte Dauer der Verumgabe)	Interventionsgruppen und Zahl der randomisierten Patienten	Ort der Durchführung und Zahl der Zentren	Relevante Zielkriterien ^(a)
B304 1998 ^(b)	<i>RCT, parallel, doppelblind, multizentrisch</i>	26 Wochen	1. Rivastigmin 2x/d (2-12 mg): n=229 2. Rivastigmin 3x/d (2-12 mg): n=227 3. Placebo: n=222	GB, IRL, AUS, CAN, ZA, I (Gesamt: 37 Zentren)	ADAS-cog, CIBIC-plus PDS, CAS, unerwünschte Ereignisse
Corey-Bloom 1998 (B352)	RCT, parallel, doppelblind, multizentrisch	26 Wochen	1. Rivastigmin 1-4 mg: n=233 2. Rivastigmin 6-12 mg: n=231 3. Placebo: n=235	USA (22 Zentren)	ADAS-cog, CIBIC-plus, PDS unerwünschte Ereignisse
Forette 1999 (B104)	RCT, parallel, doppelblind, multizentrisch <i>Phase II- Studie</i>	18 Wochen	1. Rivastigmin 2x/d (6-12 mg/d): n=45 2. Rivastigmin 3x/d (6-12 mg/d): n=45 3. Placebo: n=24	B, CAN, F, GB, NOR (Gesamt: 11 Zentren)	ADAS-cog, CIBIC-plus, NOSGER, unerwünschte Ereignisse ^(c)
Rösler 1999 (B303)	RCT, parallel, doppelblind, multizentrisch	26 Wochen	1. Rivastigmin 1-4 mg: n=243 2. Rivastigmin 6-12 mg: n=243 3. Placebo: n=239	Europa und USA (Gesamt: 45 Zentren)	ADAS-cog, CIBIC-plus, PDS unerwünschte Ereignisse
<p>a: primäre sowie relevante Zielkriterien gemäß Abschnitt 4.1.3 (falls in der jeweiligen Studie als „primär“ deklariert Fettdruck)</p> <p>b: Angaben aus dem Studienbericht des Herstellers</p> <p>c: Bei Forette 1999 waren keine primären Zielkriterien definiert.</p>					

Tabelle 14: Wesentliche Ein- und Ausschlusskriterien der Studien zu Rivastigmin

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien^(a)
B304 1998 ^(b)	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26 - Hachinski Score^(c): ≤ 4 - Betreuungsperson - CT oder MRT 	<ul style="list-style-type: none"> - andere mögliche Ursachen für die Demenz, zerebrovaskuläre Erkrankung - fortgeschrittene, schwerwiegende oder nicht eingestellte kardiovaskuläre Erkrankung; gastrointestinale, hepatische oder renale Erkrankungen, Atemwegserkrankungen - Magen-/Darmgeschwür - Epilepsie oder Substanzmissbrauch
Corey-Bloom 1998	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26 - CT oder MRT - Betreuungsperson 	<ul style="list-style-type: none"> - schwerwiegende und nicht eingestellte Erkrankungen <p>Die meisten Patienten mit Begleiterkrankungen wurden eingeschlossen.</p>
Forette 1999	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 12-26 	<ul style="list-style-type: none"> - signifikante medizinische, neurologische oder psychiatrische Erkrankungen
Rösler 1999	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-26 - Betreuungsperson 	<ul style="list-style-type: none"> - schwerwiegende und nicht eingestellte kardiovaskuläre Erkrankungen, schwere obstruktive Atemwegserkrankung oder andere lebensbedrohliche Erkrankungen wie z. B. rasch fortschreitende bösartige Neubildungen - Andere Begleiterkrankungen wie Bluthochdruck, nicht insulinpflichtiger Diabetes oder Arthritis waren zulässig.
<p>a: über die üblichen Kontraindikationen (z. B. Wirkstoffunverträglichkeit) bzw. Ausschlusskriterien (z. B. konkurrierende medikamentöse Begleittherapien) hinausgehend</p> <p>b: Angaben aus dem Studienbericht des Herstellers</p> <p>c: Die Hachinski-Skala wird zur Differentialdiagnose zwischen einer Alzheimer Demenz und einer vaskulären Demenz verwendet.</p>		

Tabelle 15: Charakterisierung der Studienpopulationen Rivastigmin

Studie	Alter (Jahre) ^(a)	Geschlecht w (%)	MMSE ^(a)	Kommentar (zur Vergleichbarkeit etc.)	Studienabbrecher N (%)
B304 1998 ^(b)					
RIV 2x/d (2-12 mg)	71 ± 8,2	57	19 (n.g.) ^(c)		54 (24)
RIV 3x/d (2-12 mg)	71 ± 7,9	60	18 (n.g.) ^(c)		38 (17)
Placebo	72 ± 8,7	60	19 (n.g.) ^(c)		33 (15)
Corey-Bloom 1998					
RIV 1-4 mg	75 (45-89)	57	20 (n.g.)		34 (15) ^(d)
RIV 6-12 mg	74 (50-89)	68	20 (n.g.)		82 (35) ^(d)
Placebo	75 (45-89)	58	20 (n.g.)		39 (17) ^(d,e)
Forette 1999					
RIV 2x/d (6-12 mg)	70 ± 9,9 ^(f)	n.g.	20 ± 4,2 ^(f)		16 (36) ^(d)
RIV 3x/d (6-12 mg)	72 ± 6,8 ^(f)	n.g.	19 ± 3,4 ^(f)		11 (24) ^(d)
Placebo	73 ± 4,8 ^(f)	n.g.	19 ± 3,8 ^(f)		2 (8) ^(d)
Rösler 1999					
RIV 1-4 mg	72 (45-95) ^(g)	59 ^(g)	20 (10-29) ^(g)	laut Text „vergleichbar“	34 (14) ^(d)
RIV 6-12 mg					79 (33) ^(d)
Placebo					31 (13) ^(d)
<p>Erläuterungen: RIV = Rivastigmin, N = Anzahl, n.g. = nicht genannt, w = weiblich Da lediglich für eine Studie (B304) Baselinedaten zu den Aktivitäten des täglichen Lebens und für keine Studie Baselinedaten zum NPI vorliegen, werden hier keine entsprechenden Daten dargestellt.</p> <p>a: Mittelwert mit SD (= Standardabweichung) oder Spannweite b: Daten aus dem Studienbericht des Herstellers c: Angaben liegen lediglich für die ITT-Population (RIV 2x/d: n=227; RIV 3x/d: n=227; Placebo: n=220) vor. d: %-Angabe aus N berechnet e: Für die Placebo-Gruppe sind in Abbildung 1 der Publikation 39 Abbrecher bei 235 randomisierten Teilnehmern genannt. Dies ergibt 196 Teilnehmer, die die Studie abschlossen, und nicht 197, wie in der Tabelle genannt. f: Angaben liegen lediglich für die Per-Protokoll-Population vor. g: Angaben nur für die gesamte Stichprobe vorhanden; nach den zum ADAS-cog vorliegenden Baselinedaten scheinen die Gruppen hinsichtlich der kognitiven Fähigkeiten vergleichbar.</p>					

5.2.3.2 Studien- und Berichtsqualität

Im Rahmen der biometrischen Qualitätsbewertung wurden zwei Studien mit „leichte Mängel“ und zwei mit „grobe Mängel“ beurteilt. Einzelheiten zur Studien- bzw. Berichtsqualität finden sich in Tabelle 16 und sind im Folgenden zusammenfassend dargestellt.

Alle Studien wurden als randomisiert, doppelblind und Placebo-kontrolliert beschrieben. Informationen bezüglich des genauen Vorgehens bei der Randomisierung fehlten lediglich bei Forette 1999. In drei Studien (B304, Forette 1999, Rösler 1999) blieb unklar, ob eine verdeckte Zuordnung der Patienten zu den Gruppen (Concealment of Allocation) gewährleistet war. Dies stellt jedoch bei tatsächlich doppelblinden Studien vermutlich keinen gravierenden Mangel dar, auch wenn genauere Angaben hierzu wünschenswert wären.

Alle Studien wurden in den Publikationen als doppelblind bezeichnet. Im Rahmen des Stellungnahmeverfahrens gab die Herstellerfirma Novartis bezüglich der Rater des CIBIC-plus an, dass diese keinen Zugriff auf Patientendaten oder irgendwelche anderen Messdaten, die in den Studien erhoben wurden, gehabt hätten (vgl. SN Novartis). Allerdings war aus den Publikationen sowie der Stellungnahme nicht ersichtlich, ob und inwieweit die Erheber der übrigen Zielkriterien im Hinblick auf jeweils andere Daten der Patienten verblindet waren. Angesichts der hohen Rate an Abbrüchen in den Studien von bis zu 36 %, zumeist aufgrund von unerwünschten Ereignissen, stellt sich die Frage, inwiefern die Verblindung von Patienten, Angehörigen und Behandlern während der Studiendurchführung aufrechterhalten werden konnte.

In den drei größeren Studien (B304, Corey-Bloom 1998, Rösler 1999) wurde jeweils eine adäquate Fallzahlplanung beschrieben. Hier wurden auch die Studienabbrecher sowie die jeweiligen Abbruchgründe angemessen dargestellt. Im Gegensatz dazu liegen bei der Studie Forette 1999 keine Angaben zur Fallzahlplanung vor. Die Anzahl der Studienabbrecher kann zwar nachvollzogen werden, jedoch wurden die Gründe nur unvollständig berichtet.

In der Studie B304 sowie Rösler 1999 wurden die Ergebnisse sowohl zur klassischen ITT-Analyse (jeweils definiert als die Population aller randomisierten Patienten) als auch zur ITT-LOCF-Population (randomisierte Patienten mit mindestens einer Messung während der medikamentösen Behandlung) und der Observed-Cases-Population dargestellt. Bei Corey-Bloom 1998 wurde zwar ein vergleichbares Vorgehen bei der statistischen Auswertung beschrieben, jedoch wurden nur die Ergebnisse der klassischen ITT- sowie der OC-Analyse berichtet. Im Folgenden werden deshalb jeweils die Daten zur klassischen ITT-Population dargestellt.

In der Studie Forette 1999 erfolgte die Auswertung nicht nach dem ITT-, sondern nach dem Per-Protokoll-Prinzip. In der Publikation wurde explizit darauf hingewiesen, dass dieses Auswertungsverfahren a priori festgelegt worden war, da die Studie in einer frühen Phase der

Entwicklung durchgeführt wurde. Die Aussagekraft der Studie ist dadurch jedoch erheblich eingeschränkt, da in den verschiedenen Gruppen bis zu etwa 50 % der ursprünglich randomisierten Patienten in der Wirksamkeitsanalyse nicht berücksichtigt wurden. Deshalb werden im Folgenden die Ergebnisse zu den Wirksamkeitsergebnissen nicht dargestellt, sondern lediglich zu den unerwünschten Ereignissen, da in diese Auswertung alle randomisierten Patienten eingingen. In der Studie Rösler 1999 betrug der Unterschied in der Nichtberücksichtigungsrate im Rahmen der ITT-LOCF-Auswertung des primären Zielkriteriums CIBIC-plus zwischen den Gruppen 6 Prozentpunkte (Rivastigmin Niedrigdosis-Gruppe 4 %; Rivastigmin Hochdosis-Gruppe 10 %; Placebo-Gruppe 4 %), was als relevante Abweichung vom ITT-Prinzip angesehen wurde.

Die Studien wiesen keine relevanten Dateninkonsistenzen auf, die die Aussage der Studien entscheidend beeinflussen könnten, sondern lediglich kleinere Diskrepanzen.

Tabelle 16: Studien- und Publikationsqualität Rivastigmin

Studie	Randomisierung/ Concealment adäquat	Verblindete Erhebung der Zielgrößen	Fallzahl- planung adäquat	Studienabbrecher genannt/ Abbruchgründe genannt	Relevante Abweichung vom ITT- Prinzip	Relevante Daten- inkonsistenz innerhalb der Publikation	Studien- und Publikationsqualität
B304 1998	ja/(ja)	(ja) ^(a)	ja	ja/ja	nein	nein	leichte Mängel
Corey-Bloom 1998	unklar/ja	(ja) ^(a)	ja ^(b)	ja/ja	nein	nein	leichte Mängel
Forette 1999	unklar/unklar	(ja) ^(c)	nein	ja/(ja) ^(d)	ja ^(e)	nein	grobe Mängel
Rösler 1999	ja/unklar	(ja) ^(a)	ja ^(b)	ja/ja	ja ^(f)	nein	grobe Mängel

a: Studie wird als doppelblind beschrieben; keine Angaben zur Verblindung der Endpunkterheber in der Publikation. Die Herstellerfirma gab im Stellungnahmeverfahren an, dass der CIBIC jeweils durch einen unabhängigen Rater erhoben worden sei, der keinen Zugriff auf Patientendaten oder irgendwelche anderen Messdaten gehabt habe, die in der Studie erhoben worden seien. Bezüglich der Erhebung des ADAS-cog sowie der weiteren Endpunkte wurden keine Angaben gemacht (vgl. SN Novartis).

b: Variabilitätsschätzer für ADAS-cog nicht angegeben, unklare Angabe für CIBIC-plus

c: Studie wird als doppelblind beschrieben, zur Verblindung der Endpunkterheber ist lediglich die Verblindung des CIBIC-plus-Raters gegenüber den Ergebnissen der psychometrischen Tests und der unerwünschten Ereignisse erwähnt.

d: Abbruchgründe nur teilweise genannt

e: Insgesamt gingen 39 % der randomisierten Patienten nicht in die Auswertung des ADAS-cog ein (Rivastigmin 2x/d-Gruppe 49 %; Rivastigmin 3x/d-Gruppe 38 %; Placebo-Gruppe 21 %).

f: Insgesamt gingen 6 % der randomisierten Patienten nicht in die ITT-Auswertung des primären Zielkriteriums CIBIC-plus ein (Rivastigmin Niedrigdosis-Gruppe 4 %; Rivastigmin Hochdosis-Gruppe 10 %; Placebo-Gruppe 4 %).

5.2.4 Vergleiche unterschiedlicher Cholinesterasehemmer

Es wurden sieben Studien identifiziert, die unterschiedliche Cholinesterasehemmer miteinander verglichen; fünf konnten in die Nutzenbewertung eingeschlossen werden. Davon verglich eine Studie Donepezil und Galantamin (Wilcock 2003), drei Studien verglichen Donepezil und Rivastigmin (Bullock 2005, Fuschillo 2001, Wang 2001) und eine Studie verglich alle drei Substanzen untereinander (Cumbo 2005). Einzelheiten zum Studiendesign und zu den Studienpopulationen sind in Tabelle 17 und Tabelle 19 dargestellt und werden im Folgenden zusammenfassend beschrieben.

5.2.4.1 Studiendesign, Dosierungen und Studienpopulationen

5.2.4.1.1 Galantamin vs. Donepezil

In der Studie Wilcock 2003 zum Vergleich von Galantamin und Donepezil wurden 188 Patienten mit MMSE-Werten zwischen 9 und 18 eingeschlossen. Von diesen wiesen etwa 14 % einen MMSE-Wert unter 12 auf. Die Studiendauer betrug 52 Wochen.

In der Galantamin-Gruppe erfolgte die Aufdosierung von Woche 1 bis 12 auf 16 mg pro Tag, danach konnte der Behandler die Dosis je nach Verträglichkeit auf 24 mg steigern (Verabreichung verteilt auf 2 Dosen täglich). In der Donepezil-Gruppe erhielten die Patienten in den ersten 4 Wochen 5 mg und ab der 5. Woche je nach Verträglichkeit 5 mg oder 10 mg (Verabreichung einmal täglich). 71 % der Galantamin-Patienten und 69 % der Donepezil-Patienten erhielten die maximale Dosierung. Die Medikation erfolgte offen, die Erhebung der Endpunkte wurde verblindet durchgeführt.

Als primärer Endpunkt diente die Bristol Activities of Daily Living Scale (BADLS) zur Beurteilung der alltagspraktischen Fertigkeiten. Des Weiteren wurden akzessorische Symptome (mittels NPI) und die kognitive Leistung mittels MMSE und ADAS-cog erhoben. Als weiterer Parameter wurde anhand des Screen for Caregiver Burden (SCGB) die objektive und subjektive Belastung der Pflegepersonen erfasst.

Die Patienten in beiden Gruppen erschienen weitgehend vergleichbar, lediglich der Frauenanteil war in der Donepezil-Gruppe etwas höher.

5.2.4.1.2 Rivastigmin vs. Donepezil

Zum Vergleich Rivastigmin vs. Donepezil fanden sich drei Studien, die den Einschlusskriterien entsprachen (Fuschillo 2001, Wang 2001, Bullock 2005).

Die Studie von Bullock 2005 zeichnete sich durch eine lange Beobachtungsdauer und einen großen Stichprobenumfang aus. In dieser doppelblinden Studie über insgesamt 2 Jahre wurden 998 Patienten an 94 europäischen und kanadischen Zentren randomisiert. Die Aufdosierung erfolgte über einen relativ langen Zeitraum von 16 Wochen. Die Medikation mit einer Dosis von maximal 12 mg Rivastigmin bzw. maximal 10 mg Donepezil wurde in identischer Form verkapselt verabreicht. Die Patienten erhielten in der Erhaltungsphase innerhalb des vorgegebenen Rahmens die bezüglich der Verträglichkeit höchstmögliche Dosis. Die durchschnittliche Dosis lag bei 9,4 mg in beiden Gruppen. Damit lagen die Dosierungsstärken für Donepezil und Rivastigmin – verglichen mit anderen Studien – im oberen Bereich. Als primärer Endpunkt diente die Severe Impairment Battery (SIB) als Maß für die kognitive Leistungsfähigkeit. Weiterhin wurden die globale Demenzschwere (mit der GDS), Aktivitäten des täglichen Lebens (mit der ADCS-ADL), akzessorische Symptome (mit dem NPI) und die kognitive Leistung mit dem MMSE erhoben.

Die Teilnehmer der Studie von Bullock 2005 setzten sich aus mittelschwer erkrankten (MMSE 10-20) ambulanten Patienten mit einer möglichen Alzheimer Demenz zusammen. Der Anteil an Patienten mit einer möglichen komorbiden Lewy-Body-Demenz betrug 3,6 % bzw. 4,4 %. Das Durchschnittsalter betrug 76 Jahre, 69 % waren Frauen, die mittlere Erkrankungsdauer lag bei 34 Monaten.

Die offene Studie von Fuschillo 2001 dauerte 30 Wochen. Rivastigmin wurde in der ersten Woche mit einer Einmaldosierung von 1,5 mg pro Tag verabreicht, dann erfolgte eine wöchentliche Erhöhung der Dosierung um jeweils 1,5 mg pro Tag bis zu 6 bis 9 mg (Verabreichung verteilt auf zwei Dosen täglich), wenn dies vertragen wurde. Donepezil wurde während der gesamten Studienlaufzeit einmal täglich in einer Dosierung von 5 mg gegeben. Die Wirksamkeit der beiden Präparate wurde anhand des ADAS-cog, der Physical Self-Maintenance Scale (PSMS) sowie des MMSE beurteilt.

Bei Fuschillo 2001 war die Anzahl der Teilnehmer mit 11 Personen in der Rivastigmin- und 16 in der Donepezil-Gruppe gering. Die Studie wurde in einer neuropsychogeriatrischen Station in Italien durchgeführt. Die Patienten wiesen eine leichte bis mittelschwere Alzheimer Demenz auf (MMSE von 10-21). Das Alter der Teilnehmer lag im Durchschnitt bei 68 (Donepezil-Gruppe) bzw. 66 Jahren (Rivastigmin-Gruppe), die mittlere Erkrankungsdauer bei 21 bzw. 22 Monaten. Der Frauenanteil betrug 56 % in der Donepezil- und 55 % in der Rivastigmin-Gruppe.

In der ebenfalls offenen, chinesischen Studie von Wang 2001 wurde Rivastigmin zunächst in einer Dosis von zweimal täglich 1,5 mg verabreicht und, wenn keine relevanten unerwünschten Wirkungen auftraten, nach vier Wochen auf zweimal täglich 3 mg erhöht. Donepezil wurde in einer Einmaldosis zunächst für 4 Wochen mit 5 mg gegeben und dann, wenn dies möglich war, auf 10 mg erhöht. Die Wirksamkeitsbeurteilung erfolgte anhand des MMSE, der Blessed-Roth Dementia Scale (die vor allem soziale Aktivitäten und ADLs erfasst) und der GDS (zur Schweregradbeurteilung der Demenz). Die primäre Wirksamkeitsbeurteilung erfolgte nach 16 Wochen.

In der Studie von Wang 2001 wurden in beiden Gruppen jeweils 62 Patienten randomisiert. Die Teilnehmer hatten eine leichte bis mittelschwere Demenz (MMSE von 10-26). Soziodemographische Daten zur Grundgesamtheit konnten der Publikation nicht entnommen werden, allerdings wurde vermerkt, dass sich die Gruppen bezüglich solcher Daten nicht statistisch signifikant unterschieden.

5.2.4.1.3 Rivastigmin vs. Donepezil vs. Galantamin

An der 18-monatigen offenen Studie von Cumbo 2005 nahmen insgesamt 101 Patienten teil. 37 Patienten in der Rivastigmin-Gruppe erhielten im Durchschnitt 9 mg Rivastigmin täglich, 31 Patienten in der Donepezil-Gruppe 10 mg täglich und die 33 Patienten in der Galantamin-Gruppe erhielten durchschnittlich 16 mg Galantamin. Über die Aufdosierungsschemata wurden keine Angaben gemacht. Das primäre Zielkriterium war die Zeit bis zum erstmaligen Auftreten von Verhaltensauffälligkeiten im Zusammenhang mit der Demenzerkrankung. Sekundäre Zielkriterien waren die Häufigkeit und Schwere der Symptome zum Zeitpunkt des Auftretens, gemessen anhand des NPI sowie der NPI-D (NPI Caregiver Distress Scale) und der Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD).

Es wurden Patienten mit einer leichten bis mittelschweren wahrscheinlichen Alzheimer Demenz (MMSE 10-27) eingeschlossen, die eine Erkrankungsdauer von mindestens drei Jahren aufwiesen. Zum Ort der Durchführung wurden keine näheren Informationen gegeben. Die Vergleichbarkeit der Patienten zwischen den drei Behandlungsgruppen war nicht beurteilbar, da hierzu keine separaten Informationen vorlagen.

Tabelle 17: Eingeschlossene direkte Vergleichsstudien zwischen Cholinesterasehemmern

Studie	Studiendesign	Studiendauer (kontrollierte Dauer der Verumgabe)	Interventionsgruppen und Zahl der randomisierten Patienten	Ort der Durchführung und Zahl der Zentren	Relevante Zielkriterien ^(a)
Galantamin vs. Donepezil					
Wilcock 2003	RCT, parallel, offen/raterverblindet, multizentrisch; Pilotstudie	12 Monate	1. Galantamin 16-24 mg: n=97 2. Donepezil 5-10 mg: n=91	GB (18)	BADLS ADAS-cog, NPI, SCGB, unerwünschte Ereignisse
Rivastigmin vs. Donepezil					
Bullock 2005	RCT, parallel, doppelblind, multizentrisch	24 Monate	1. Rivastigmin 3-12 mg: n=498 2. Donepezil 5-10 mg: n=500	AUS, CAN, F, D, I, E, GB Gesamt (94)	SIB ADCS-ADL, GDS, NPI, unerwünschte Ereignisse
Fuschillo 2001	RCT, parallel, offen, monozentrisch	7,5 Monate	1. Rivastigmin 6-9 mg: n=11 2. Donepezil 5 mg: n=16	Italien, 1 Zentrum	ADAS-cog, PSMS, unerwünschte Ereignisse ^(b)
Wang 2001	RCT, parallel, offen, multizentrisch	4 Monate	1. Rivastigmin 3-6 mg: n=62 2. Donepezil 5-10 mg: n=62	China, mehrere Zentren	MMSE, Blessed-Roth Dementia Scale, GDS, unerwünschte Ereignisse ^(b)

(Fortsetzung)

Tabelle 17 (Fortsetzung): Eingeschlossene direkte Vergleichsstudien zwischen Cholinesterasehemmern

Rivastigmin vs. Donepezil vs. Galantamin						
Cumbo 2005	RCT, parallel, offen	18 Monate	1. Rivastigmin 9 mg: n=37 2. Donepezil 10 mg: n=31 3. Galantamin 16 mg: n=33	keine Angaben		Zeit bis zum erstmaligen Auftreten von Verhaltensauffälligkeiten im Zusammenhang mit der Demenzerkrankung^(c) Häufigkeit und Schwere der Verhaltensauffälligkeiten, unerwünschte Ereignisse
a: primäre sowie relevante Zielkriterien gemäß Abschnitt 4.1.3 (falls in der jeweiligen Studie als „primär“ deklariert Fettdruck) b: In dieser Studie waren keine primären Zielkriterien definiert. c: erfasst anhand NPI, NPI-D, BEHAVE-AD						

Tabelle 18: Wesentliche Ein- und Ausschlusskriterien der direkten Vergleichsstudien zwischen Cholinesterasehemmern

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Galantamin vs. Donepezil		
Wilcock 2003	<ul style="list-style-type: none"> - mittelschwere AD (wahrscheinlich) - MMSE: 9-18 - progressiver Abbau von kognitiven Fähigkeiten seit ≥ 12 Monaten - CT oder MRT - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Erkrankung, Multiinfarktdemenz oder zerebrovaskuläre Erkrankung, kognitive Einschränkung anderer Ursache - andere medizinische Umstände, die den sicheren Studienabschluss des Patienten fraglich machen
Rivastigmin vs. Donepezil		
Bullock 2005	<ul style="list-style-type: none"> - mittelschwere^(b) AD (wahrscheinlich) - MMSE: 10-20 - Betreuungsperson 	<ul style="list-style-type: none"> - andere neurodegenerative Primärerkrankung (einschl. Parkinson) - jegliche fortgeschrittene, schwerwiegende oder nicht eingestellte Erkrankung oder Behinderung - Episode einer Major Depression, nicht eingestellte Epilepsie, Magen-/Darmgeschwür, schwere oder nicht eingestellte Asthmaerkrankung oder kardiovaskuläre Erkrankung, zerebrovaskuläre Erkrankung
Fuschillo 2001	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (möglich oder wahrscheinlich)^(c) - MMSE: 10-21 - CT oder MRT - Betreuungsperson 	<ul style="list-style-type: none"> - schwerwiegende und nicht eingestellte Erkrankungen
Wang 2001	<ul style="list-style-type: none"> - leichte bis mittelschwere AD - MMSE: 10-26 	<ul style="list-style-type: none"> - Hachinski Score^(d) ≥ 7 - andere mögliche Demenzursachen - aktive Epilepsie oder Magen-/Darmgeschwür - schwerwiegende kardiovaskuläre Erkrankung oder Asthma

(Fortsetzung)

Tabelle 18 (Fortsetzung): Wesentliche Ein- und Ausschlusskriterien der direkten Vergleichsstudien zwischen Cholinesterasehemmern

Studie	wesentliche Einschlusskriterien	wesentliche Ausschlusskriterien ^(a)
Rivastigmin vs. Donepezil vs. Galantamin		
Cumbo 2005	<ul style="list-style-type: none"> - leichte bis mittelschwere AD (wahrscheinlich) - MMSE: 10-27 - Krankheitsdauer \geq 3 Jahre - keine Verhaltensauffälligkeiten - Betreuungsperson 	<ul style="list-style-type: none"> - andere (frühere) neurologische oder psychiatrische Erkrankung - Substanzmissbrauch - klinisch signifikante medizinische oder chirurgische Störungen (unabhängig von deren Stabilität) - frühere Demenztherapien
<p>a: über die üblichen Kontraindikationen (z. B. Wirkstoffunverträglichkeit) bzw. Ausschlusskriterien (z. B. konkurrierende medikamentöse Begleittherapien) hinausgehend</p> <p>b: In der Publikation wird das Erkrankungsstadium als „moderate to moderately-severe“ bezeichnet.</p> <p>c: Angaben hierzu sind inkonsistent in der Publikation: im Abstract „wahrscheinlich oder möglich“, im Text lediglich „wahrscheinlich“.</p> <p>d: Die Hachinski-Skala wird zur Differentialdiagnose zwischen einer Alzheimer Demenz und einer vaskulären Demenz verwendet.</p>		

Tabelle 19: Charakterisierung der Studienpopulationen zu direkten Vergleichsstudien zwischen Cholinesterasehemmern

Studie	Alter (Jahre) ^(a)	Geschlecht w (%)	MMSE ^(a)	Kommentar zur Vergleichbarkeit	Studienabbrücker N (%)
Galantamin vs. Donepezil					
Wilcock 2003					
GAL 16-24 mg	74 (53-88)	56	15 (14,5; 15,7) ^(b)		19 (20)
DON 5-10 mg	73 (54-88)	68	15 (14,2; 15,3) ^(b)		20 (22)
Rivastigmin vs. Donepezil					
Bullock 2005					
RIV 3-12 mg	76 (6,6)	69	15 (3,0)		237 (48)
DON 5-10 mg	76 (6,8)	69	15 (2,9)		183 (37)
Fuschillo 2001					
RIV 6-9 mg	66 (9,2)	55	13 (3,3)		unklar ^(c)
DON 5 mg	68 (5,6)	56	14 (3,4)		
Wang 2001				nicht zu beurteilen, da wesentliche Angaben fehlen	
RIV 3-6 mg	n.g.	n.g.	17 (4,1)		2 (3)
DON 5-10 mg	n.g.	n.g.	18 (4,7)		1 (2)
Rivastigmin vs. Donepezil vs. Galantamin					
Cumbo 2005 ^(d)					
DON 10 mg		58		nicht zu beurteilen, da keine separaten Angaben vorliegen	
RIV 9 mg	76 (66-83)	57	17 (n.g.)		unklar ^(c)
GAL 16 mg		58			
Erläuterungen: DON = Donepezil, GAL = Galantamin, n.g. = nicht genannt, RIV = Rivastigmin, w = weiblich Zum NPI liegen Baselinedaten lediglich für eine Studie vor (Bullock 2005); die vorliegenden Baselinedaten zu den Aktivitäten des täglichen Lebens beziehen sich auf drei verschiedene Skalen, so dass sie für die Beurteilung der Vergleichbarkeit der Studienpopulationen nur bedingt hilfreich sind. Diese Daten werden deshalb nicht dargestellt.					
a: Mittelwert mit SD (= Standardabweichung) bzw. Spannweite					
b: Konfidenzintervall					
c: In der Publikation wurde nicht explizit berichtet, ob es Studienabbrüche gab.					
d: jeweils mittlere Dosierung angegeben					

5.2.4.2 Studien- und Berichtsqualität

Tabelle 20 fasst die Studien- und Publikationsqualität der Studien zum Vergleich verschiedener Cholinesterasehemmer untereinander zusammen.

5.2.4.2.1 Galantamin vs. Donepezil

In der Studie Wilcock 2003 wurde bei der Erstellung des Randomisierungsplans adäquat vorgegangen, allerdings liegen keine Angaben dazu vor, inwieweit die Gruppeneinteilung verdeckt vorgenommen wurde (Concealment of Allocation). Im Unterschied zu tatsächlich doppelblinden Studien stellt dies bei einer offenen Studie einen relevanten Qualitätsmangel dar. Auch die aus praktischen Gründen nicht angestrebte Verblindung der Patienten und Behandler senkt die methodische Qualität der Studie deutlich. Zwar wurde beschrieben, dass alle klinischen Messungen durch verblindete Rater durchgeführt wurden, allerdings wurde nicht dargestellt, wie die Raterverblindung tatsächlich aufrechterhalten wurde. Äußerungen der Patienten oder der pflegenden Personen zur Medikation waren vermutlich nicht immer zu vermeiden. Zudem ist nicht beschrieben, ob die Erfasser gegenüber den unerwünschten Ereignissen verblindet waren.

Aus der Beschreibung der Fallzahlplanung wurde nicht deutlich, ob die Studie eine angemessene statistische Teststärke aufwies. Sie war zudem nicht erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudie konzipiert. Hinsichtlich der Studienabbrüche ging zwar aus der Publikation hervor, wie viele Patienten die Studie abbrachen, allerdings lagen keine vollständigen Informationen zu den Gründen für die Studienabbrüche vor. Detailliertere Angaben zu den Abbruchgründen wurden von der Firma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens zur Verfügung gestellt. Zur statistischen Analyse wurden die Unterschiede Endpunkt – Baseline in der LOCF-Analyse (ITT-LOCF genannt) nach Kontrolle von Alter und Baseline-MMSE, jedoch ohne Kontrolle der jeweiligen Baselinewerte der Skalen herangezogen.

5.2.4.2.2 Rivastigmin vs. Donepezil

Alle drei Studien sind als randomisiert bezeichnet, eine Studie (Bullock 2005) wurde als doppelblind, die beiden anderen als offen beschrieben (Fuschillo 2001, Wang 201). In der Studie Bullock 2005 erfolgte die Randomisierung adäquat; die verdeckte Gruppenzuordnung war ebenfalls gewährleistet. Auch bei Wang 2001 wurde die Randomisierung angemessen durchgeführt, zu einer Verdeckung der Gruppenzuordnung fehlten allerdings Informationen. In der Studie Fuschillo 2001 blieben beide Aspekte unklar. Auch bei diesen beiden offenen Studien stellt die Unklarheit bezüglich der verdeckten Zuteilung einen wesentlichen Qualitätsmangel dar. Darüber hinaus wurde jeweils nicht berichtet, ob die Erhebung der Endpunkte verblindet erfolgte. In der Studie von Bullock 2005 dagegen wurde dargestellt,

dass alle an der Studie beteiligten Personen über die Studiendauer hinweg gegenüber der jeweiligen Behandlungsgruppe verblindet waren.

Bei Bullock 2005 wurde eine angemessene Fallzahlplanung für die Entdeckung eines Unterschiedes vorgenommen; insgesamt wurden 998 Patienten randomisiert. In den Studien Fuschillo 2001 und Wang 2001 dagegen wurden keine Angaben zu einer Fallzahlplanung gemacht – die betrachtete Patientengruppe war mit 27 randomisierten Patienten insbesondere bei Fuschillo 2001 zu klein, um selbst moderate Unterschiede zwischen den Gruppen aufdecken zu können. Alle drei Studien waren nicht erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudien konzipiert.

Sowohl die Anzahl der Studienabbrüche als auch die jeweiligen Gründe wurden bei Bullock 2005 und bei Wang 2001 dargestellt. In der Studie Fuschillo 2001 wurden hierzu keine Informationen gegeben.

Bei Bullock 2005 stellte die ITT-LOCF-Auswertung die primäre Analyse dar. Über die im Vergleich zu anderen Studien lange Studiendauer von zwei Jahren brach ein relativ großer Teil der Patienten die Studie ab (48 % in der Rivastigmin-Gruppe bzw. 37 % in der Donepezil-Gruppe). In der Studie Fuschillo 2001 blieb unklar, ob jeweils alle 27 Teilnehmer in alle Analysen eingegangen sind.

Relevante Dateninkonsistenzen waren in der Publikation zu Fuschillo 2001 bezüglich der unerwünschten Ereignisse vorhanden (vgl. Tabelle 20).

5.2.4.2.3 Rivastigmin vs. Donepezil vs. Galantamin

Zum direkten Vergleich von Rivastigmin, Donepezil und Galantamin liegt eine als randomisiert bezeichnete offene Studie von Cumbo 2005 vor. Zum Randomisierungsverfahren wurden in der Publikation keine näheren Angaben gemacht. Ebenso blieb unklar, ob die Gruppenzuteilung verdeckt erfolgte. Es wurde nicht berichtet, ob die Erhebung der Endpunkte verblindet erfolgte. Die Publikation enthielt keine expliziten Informationen darüber, ob eine Fallzahlplanung durchgeführt worden war. Auch diese Studie war offenbar nicht als Äquivalenz- bzw. Nichtunterlegenheitsstudie angelegt. Ebenfalls fehlen Angaben zu möglichen Studienabbrüchen. Es wurde jedoch berichtet, dass alle Patienten wie geplant an den Untersuchungen teilnahmen, was die Annahme zulässt, dass es keine Studienabbrüche gab und dass auch das ITT-Prinzip nicht verletzt wurde. Aufgrund der Unklarheit bezüglich der verdeckten Zuteilung und der nicht erfolgten Verblindung weist die Studie jedoch insgesamt grobe Mängel auf.

Tabelle 20: Studien- und Publikationsqualität Vergleichsstudien

Studie	Randomisierung/ Concealment adäquat	Verblindete Erhebung der Zielgrößen	Fallzahl- planung adäquat	Studienabbrecher genannt/ Abbruchgründe genannt	Relevante Abweichung vom ITT- Prinzip	Relevante Daten- inkonsistenz innerhalb der Publikation	Studien- und Publikationsqualität
Galantamin vs. Donepezil							
Wilcock 2003	ja/unklar	ja ^(a)	unklar	ja/ja ^(b)	nein	nein	grobe Mängel
Rivastigmin vs. Donepezil							
Bullock 2005	unklar/ja	ja ^(c)	ja	ja/ja	(nein) ^(d)	nein	leichte Mängel
Fuschillo 2001	unklar/unklar	nein ^(e)	nein	unklar/unklar	unklar	ja ^(f)	grobe Mängel
Wang 2001	ja/unklar	nein ^(e)	nein	ja/ja	nein	nein	grobe Mängel
Rivastigmin vs. Donepezil vs. Galantamin							
Cumbo 2005	unklar/unklar	nein ^(e)	nein	unklar/unklar	nein	nein	grobe Mängel
<p>a: Lediglich die Erheber der Endpunkte waren in der Studie verblindet.</p> <p>b: Der Bewertung liegen Angaben der Firma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens zugrunde (vgl. SN Janssen-Cilag).</p> <p>c: Die Studie wird als doppelblind bezeichnet. Es wird beschrieben, dass das an der Studie beteiligte Personal während der ganzen Studienlaufzeit gegenüber der Gruppenzuteilung verblindet war, jedoch wird nicht dargestellt, auf welche Art die Verblindung aufrechterhalten wurde.</p> <p>d: Auswertung erfolgte nach dem LOCF-Prinzip; allerdings gingen 5,4 % (Rivastigmin) bzw. 3,4 % der Patienten nicht in die Auswertung des Hauptzielkriteriums ein; problematisch ist darüber hinaus der relativ große und vor allem unterschiedlich große Anteil der Therapieabbrüche im Verlauf in den beiden Studiengruppen.</p> <p>e: offene Studie</p> <p>f: Berichtete Prävalenzen zu einzelnen unerwünschten Ereignissen sind nicht vereinbar mit der Zahl der Patienten.</p>							

5.3 Ergebnisse zu Therapiezielen

5.3.1 Donepezil

In den Studien zu Donepezil wurden zu einzelnen Zielkriterien, z. B. Aktivitäten des täglichen Lebens, oft verschiedene Skalen als Erhebungsinstrumente verwendet. Aufgrund fehlender Textangaben zu Mittelwerten und Streumaßen mussten Angaben zum Teil aus den Abbildungen entnommen werden, wodurch die Daten einer gewissen Unsicherheit unterliegen. Aus Abbildungen geschätzte Werte sind entsprechend gekennzeichnet.

5.3.1.1 Aktivitäten des täglichen Lebens

Die alltagspraktischen Fähigkeiten als Zielgröße wurden in sechs Studien mittels verschiedener Erhebungsinstrumente erfasst (Burns 1999, Gauthier 2002, Homma 2000, Mohs 2001, Tariot 2001, Winblad 2001), bei Mohs 2001 erfolgte darüber hinaus eine andere Art der Operationalisierung (Zeit bis zum Erreichen einer definierten Verschlechterung).

Bei Burns 1999 ließen sich zu Studienende unter der 10 mg-Dosis, nicht aber der 5 mg-Dosis, leichte (statistisch signifikante) Effekte im Vergleich zu Placebo in der IDDD-Subskala „complex tasks“ (instrumentelle Aktivitäten) nachweisen, wohingegen die basalen Aktivitäten (Subskala „self care“) im Verlauf nicht beeinflusst wurden (dazu wurden in der Publikation ansonsten keine genaueren Angaben gemacht). Über Ergebnisse zur Gesamtskala (instrumentelle + basale Aktivitäten) wurde nichts berichtet. Aufgrund der selektiv anmutenden Darstellungsweise sind diese Ergebnisse nur sehr eingeschränkt interpretierbar.

In der Studie von Gauthier 2002 wurden sowohl die basalen Aktivitäten des täglichen Lebens (PSMS-plus) als auch die instrumentellen Aktivitäten (IADL-plus) sowie eine kombinierte, von den beiden isolierten Maßen abhängige Skala eingesetzt (DAD). In dieser Studie hatte Donepezil einen statistisch signifikanten Effekt sowohl auf die instrumentellen Aktivitäten (IADL-plus) als auch (in etwas geringerem Ausmaß) auf die basalen Aktivitäten und das kombinierte Maß DAD.

Eine modifizierte Version der Crichton Scale (Caregiver-rated Modified Crichton Scale – CMCS), die verschiedene Aspekte der Aktivitäten des täglichen Lebens misst, wurde in der Studie Homma 2000 betrachtet. Es wurde eine signifikant geringere Verschlechterung in der Donepezil-Gruppe berichtet, gleichzeitig jedoch darauf hingewiesen, dass die psychometrische Güte dieser Skala bisher noch nicht untersucht worden ist.

Tabelle 21: Donepezil: Ergebnisse zu alltagspraktischen Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Gruppenunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Burns 1999 (24 Wochen)	DON 5 mg	IDDD Gesamt- skala	(271) ^(b)	n.g.	n.g.	? ^(c)	n.g.
	DON 10 mg		(273) ^(b)	n.g.	n.g.	(↗) ^(d)	n.g.
	Placebo		(274) ^(b)	n.g.			
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	DAD	92	0 ^(e) (13,4 ^(f))	- 9,25 (n.g.)	↗	<0,001
	Placebo		101	-9,25 ^(e) (14,3 ^(f))			
Homma 2000 (24 Wochen)	DON 5 mg	CMCS	103 ^(g)	1,03 (6,7)	-2,42	↗	0,010
	Placebo		99 ^(g)	3,45 (7,1)			
Tariot 2001 ^(h) (24 Wochen)	DON (5-) 10 mg	PSMS	(103) ^(b,h)	1	n.g.	? ^(c)	„n.s.“
	Placebo		(105) ^(b,h)	1			
Winblad 2001 (52 Wochen)	DON (5-) 10 mg	PDS	89 ⁽ⁱ⁾	-11 ^(e)	4 ^(f)	↗	„<0,05“
	Placebo		94 ⁽ⁱ⁾	-15 ^(e)			
<p>Erläuterung: DON = Donepezil, KI = Konfidenzintervall, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, N = Anzahl, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in Auswertung, sofern nicht anders angegeben b: Zahl der randomisierten Patienten, da keine genauen Angaben zur Anzahl der Patienten in Auswertung vorliegen c: Da in der Publikation keine Angaben zum beobachteten Gruppenunterschied gemacht werden, bleibt die Richtung der Veränderung unklar. d: Da für die Subskala „complex tasks“ ein statistisch signifikanter Vorteil zu Placebo berichtet wird, ist die Richtung des Effektes für die Gesamtskala vermutlich entsprechend, letztlich aber mit einer gewissen Unsicherheit behaftet. e: Werte aus Abbildung geschätzt f: eigene Berechnung g: Detaillierte Daten zur Auswertung liegen nur für die Per-Protokoll-Population vor; ITT-Analyse: p = 0,019. h: Dargestellte Patientenpopulation schließt auch Patienten mit schwerer AD ein (MMSE 5-26). i: Daten werden nur für die Observed Cases dargestellt; auch bei der LOCF-Auswertung ist das Ergebnis statistisch signifikant. Die Anzahl der Patienten in Auswertung wurde erschlossen aus den Angaben dazu, für wie viele Patienten Daten zu den Einzelitems der Skala vorlagen (89 bis 93 zu Donepezil, 94 bis 97 zu Placebo, ohne Berücksichtigung des Items zu „Safe Driving“).</p>							

Im Rahmen der Studie von Tariot 2001 fand sich kein statistisch signifikanter Effekt von Donepezil hinsichtlich einer Veränderung auf der Physical Self-Maintenance Scale (PSMS), einer stark an den basalen Fähigkeiten ausgerichteten Skala, zum Studienendpunkt. Genauere Daten wurden in der Studie nicht gegeben. Diese Auswertung bezog sich auf die gesamte Studienpopulation, die auch schwer beeinträchtigte Personen umfasste. Eine Subgruppenanalyse für die leicht und mittelschwer beeinträchtigten Patienten lag nicht vor.

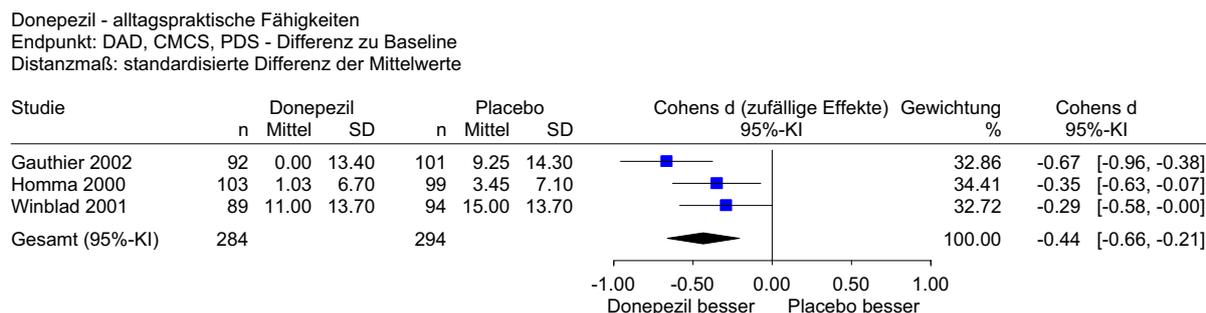
Bei Winblad 2001 wurde berichtet, dass bei den mit Donepezil behandelten Patienten zu allen Messzeitpunkten eine geringere Verschlechterung als bei den Patienten in der Placebo-Gruppe bezüglich der mittels der PDS erfassten Aktivitäten beobachtet werden konnte. Diese Unterschiede waren zum Endpunkt (LOCF) statistisch signifikant, allerdings wurden keine näheren Angaben hierzu gemacht.

Bei Mohs 2001 wurde die Zeit bis zum Erreichen einer klinisch auffälligen funktionellen Verschlechterung betrachtet, die wiederum durch ein bestimmtes Ausmaß der Verschlechterung in den basalen oder den instrumentellen ADL (jeweils gemessen anhand des ADFACS) oder durch die Zunahme des CDR-Wertes um mindestens einen Punkt definiert war. Die mediane Zeit bis zum Erreichen des Endpunktes „funktionelle Verschlechterung“ betrug 357 Tage (untere Grenze 95 %-KI: 280 Tage) für die Verum-Gruppe und 208 Tage (obere Grenze 95 %-KI: 252 Tage) für die Placebo-Gruppe (statistisch signifikanter Unterschied mit $p = 0,005^{\circ}$). Darüber hinaus enthielt die Studie Mohs 2001 weitergehende Auswertungen hinsichtlich der Unterschiede im ADFACS zwischen Placebo- und Verum-Gruppe beim jeweils letzten Messzeitpunkt sowie Gruppenunterschiede in den Veränderungen über die Studienlaufzeit hinweg. Das Design der Studie lässt Ergebnisse außerhalb der (primären) Überlebenszeitanalyse allerdings nicht mehr sicher interpretieren, da entweder sukzessive Patienten für die entsprechenden Auswertungen verloren gingen („observed cases“ [OC]-Analyse) oder die Werte der Patienten zum Zeitpunkt des Erreichens des Endpunktes verwendet wurden, womit das Prinzip des „zufälligen Fehlens“ (missing at random) verletzt ist.

Die Unterschiedlichkeit in den verwendeten Erhebungsinstrumenten und teilweise divergente Operationalisierung sowie die zum größten Teil unzureichende Darstellung der Ergebnisse (siehe Tabelle 21) ließen eine quantitative Zusammenfassung der Ergebnisse nur sehr eingeschränkt zu. In eine Meta-Analyse konnten nur drei der sechs Studien eingeschlossen werden, für zwei dieser drei Studien lagen darüber hinaus lediglich dezidierte Daten aus OC-Analysen vor, mit einem Datenverlust für 25 % (Homma 2000) bzw. 36 % (Winblad 2001) der randomisierten Patienten. In den Publikationen zu beiden Studien fanden sich jedoch andererseits Aussagen, dass sich die Ergebnisse aus diesen OC-Analysen nicht wesentlich von

[°] Dabei handelt es sich um den größeren von zwei angegebenen p-Werten (zu zwei unterschiedlichen statistischen Tests).

denen einer LOCF-Analyse unterschieden, so dass sie orientierend für diese Meta-Analyse Verwendung fanden. Die Meta-Analyse zeigt eine statistisch signifikante Überlegenheit von Donepezil gegenüber Placebo (Abbildung 2; standardisierte Mittelwertsdifferenz: -0,44 [95 %-KI: -0,21 – -0,66], Modell mit zufälligen Effekten) bei moderater Heterogenität ($I^2 = 46,3\%$). Der ermittelte Unterschied von nahezu einer halben Standardabweichung muss als Überschätzung des Effekts angesehen werden, da zum einen das eindeutig negative Ergebnis der Studie von Tariot 2001 hier nicht miteinbezogen werden konnte und zum anderen der in der Studie von Gauthier 2002 beobachtete Unterschied von mehr als 9 Punkten auf der DAD-Skala im Vergleich zu den anderen Cholinesterasehemmern (siehe Abschnitte 5.3.2.1 und 5.3.3.1) ungewöhnlich groß ist. Somit kann in der Gesamtschau und vor dem Hintergrund der diesbezüglich mangelhaften Berichterstattung allenfalls von einem Hinweis auf eine günstige Beeinflussung der alltagspraktischen Fähigkeiten durch Donepezil bei nicht ausreichend sicher schätzbarer Effektstärke ausgegangen werden. Diese Bewertung wird unterstützt durch die von der Herstellerfirma Eisai im Rahmen des Stellungnahmeverfahrens zusätzlich gelieferte Meta-Analyse. Die Analyse zeigt unter Einbezug der Studien Burns 1999 und Tariot 2001 einen etwas geringeren gemeinsamen Effekt ($d = -0,28$ [95 %-KI: -0,44 – -0,12]; vgl. SN Eisai).



Heterogenität: $Q=3.72$, $df=2$ ($p=0.156$), $I^2=46.3\%$
 Gesamteffekt: Z Score= -3.78 ($p=0.000$), $\tau^2=0.018$

Abbildung 2: Donepezil: Meta-Analyse zu Aktivitäten des täglichen Lebens

5.3.1.2 Begleitende Psychopathologie

Drei größere Studien nutzten den NPI (Winblad 2001, Gauthier 2002) bzw. den NPI-NH (Neuropsychiatrisches Interview–Nursing Home, Studie von Tariot 2001) zur Beurteilung der Beeinflussung einer begleitenden Psychopathologie. Während Gauthier 2002 einen statistisch signifikanten Effekt zugunsten von Donepezil beobachtete, unterschieden sich in den Studien von Winblad 2001 und Tariot 2001 die Gruppen nicht statistisch signifikant. Bei Winblad 2001 fand sich diesbezüglich nur eine kursorische Bemerkung, dass ab der 12. Studienwoche ein günstiger Effekt zugunsten Donepezil zu verzeichnen war; bei Tariot 2001 zeigte sich

dagegen ein Effekt in umgekehrter Richtung (Tabelle 22). Die Analyse von Tariot 2001 bezieht sich allerdings auf die Gesamtpopulation der betrachteten Patienten, die auch schwer beeinträchtigte Patienten einschloss (MMSE 5-26). Für die Subgruppe der Patienten mit leichter bis mittelschwerer Erkrankung wurde bzgl. des NPI-NH keine separate Auswertung berichtet.

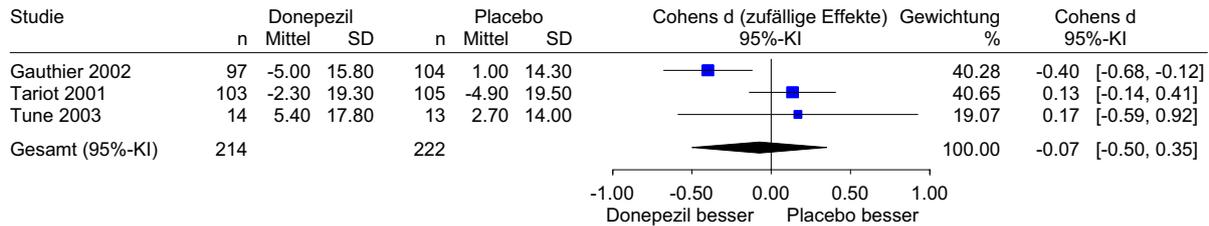
Auch in den beiden sehr kleinen Studien Prasher 2002 und Tune 2003 (jeweils ≤ 30 Patienten) wurde der NPI eingesetzt. Bei Prasher 2002 – hier nahmen ausschließlich Menschen mit Down-Syndrom teil – wurde unter Donepezil eine stärkere Verschlechterung beobachtet als unter Placebo ($p = 0,03$); dabei ist jedoch die Berechnungsgrundlage für den angegebenen p-Wert unklar. Bei Tune 2003 waren die Veränderungen im NPI in beiden Gruppen nicht statistisch signifikant unterschiedlich, allerdings ebenfalls mit einem Trend zuungunsten von Donepezil. Bezüglich der Subskalen des NPI wurden in den meisten Studien keine (Prasher 2002, Tune 2003, Winblad 2001) bzw. keine ausreichend interpretierbaren Daten (Gauthier 2002, Tariot 2001) berichtet.

Zusammenfassend geben die Studien keinen eindeutigen Hinweis auf eine günstige oder ungünstige Beeinflussung der Psychopathologie durch Donepezil. Eine Meta-Analyse (Abbildung 3) bestätigt diese Einschätzung. Auffällig ist eine stark ausgeprägte Heterogenität ($I^2 = 74,1 \%$), die durch den Unterschied in den Ergebnissen der Studien von Gauthier 2002 und Tariot 2001 hervorgerufen wird. Dieser Unterschied mag in dem besonderen Kollektiv der bei Tariot 2001 eingeschlossenen Patienten aus Pflegeheimen begründet sein; ansonsten fanden sich keine Unterschiede im Studiendesign oder den Patientencharakteristika, die diese Diskrepanz erklären könnten. Aufgrund der Heterogenität wurde ein Modell mit zufälligen Effekten gewählt. Bei Verwendung von festen Effekten zeigt sich ein etwas günstigerer Effekt zugunsten von Donepezil ohne statistische Signifikanz (Daten nicht dargestellt). In die Meta-Analyse konnte das Ergebnis der Studie von Prasher 2002 nicht aufgenommen werden, da für den relevanten Zeitpunkt 24 Wochen kein Variabilitätsmaß angegeben bzw. ein Variabilitätsmaß nicht eindeutig aus Abbildungen in der Publikation abschätzbar war.

Tabelle 22: Donepezil: Ergebnisse zur begleitenden Psychopathologie

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Gruppenunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	NPI	97	-5,0 (15,8 ^(b))	-5,9 (n.g.)	↗	0,002
	Placebo		104	1,0 (14,3 ^(b))			
Prasher 2002 ^(c) (24 Wochen)	DON 10 mg	NPI	27 ^(d)	-2,2 (n.g.)	+2,2 (n.g.)	↘	n.g. ^(e)
	Placebo			-4,4 (n.g.)			
Tariot 2001 ^(f) (24 Wochen)	DON (5-) 10 mg	NPI- NH	103 ^(g)	-2,3 (19,3 ^(b))	+2,6 (n.g.)	↘	„n.s.“
	Placebo		103 ^(g)	-4,9 (19,3 ^(b))			
Tune 2003 ^(c) (24 Wochen)	DON 10 mg	NPI	14	5,4 (17,8)	+2,8 (-9,9; 15,4)	↘	0,688
	Placebo		13	2,7 (14,0)			
Winblad 2001 (52 Wochen)	DON (5-) 10 mg	NPI	(142) ^(h)	n.g.	n.g.	↗	„n.s.“
	Placebo		(144) ^(h)	n.g.			
<p>Erläuterung: DON = Donepezil, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in der Auswertung, sofern nicht anders angegeben b: eigene Berechnungen c: unklar, ob berichtete Ergebnisse ITT-LOCF-Analyse darstellen d: Anzahl der Patienten in Auswertung aus Angaben zu Freiheitsgraden in der Publikation erschlossen, die Zuteilung zu den Gruppen ist jedoch unklar. e: In der Publikation wird lediglich ein p-Wert für den Unterschied über den gesamten Zeitraum angegeben. f: Daten gelten für die Gesamtpopulation der betrachteten Patienten (MMSE 5-26), da keine Subgruppenanalyse für die leicht bis mittelschwer beeinträchtigten Patienten vorliegt. g: Angaben direkt vom Hersteller h: Zahl der randomisierten Patienten, da keine genauen Angaben zur Anzahl der Patienten in Auswertung vorliegen.</p>							

Donepezil - begleitende Psychopathologie
 Endpunkt: NPI - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=7.71$, $df=2$ ($p=0.021$), $I^2=74.1\%$
 Gesamteffekt: Z Score=-0.34 ($p=0.731$), $\tau^2=0.095$

Abbildung 3: Donepezil: Meta-Analyse zur begleitenden Psychopathologie

5.3.1.3 Kognitive Leistungsfähigkeit

In sieben Studien (Burns 1999, Homma 2000, Krishnan 2003, Moraes 2006, Rogers 1998, Seltzer 2004, Tune 2003) wurde zur Untersuchung der Progression der kognitiven Defizite der ADAS-cog eingesetzt. Für drei Studien (Gauthier 2002, Tariot 2001, Winblad 2001) lagen Daten zum MMSE vor. Die Unterschiede zur Placebo-Gruppe in den ADAS-cog-Veränderungen zum jeweiligen Ausgangswert lagen am Studienendpunkt zwischen 1,5 (Donepezil 5 mg-Gruppe) und 3,1 Punkten (Donepezil 10 mg-Gruppe) zugunsten von Donepezil. Lediglich in der Studie Moraes 2006 war der beobachtete Effekt deutlich größer. Hier lagen auch zur Baseline die Werte in beiden Gruppen deutlich über denen in anderen Studien; darüber hinaus wurde die bereits zu Studienbeginn beobachtete Abweichung zwischen den Gruppen in der Auswertung nicht berücksichtigt. Es bleibt letztendlich unklar, worauf dieser Unterschied im Vergleich zu den übrigen Studien zurückzuführen ist. Die Wirkungen auf den MMSE bewegten sich zwischen 1,0 (Donepezil [5-] 10 mg-Gruppe) und 2,1 Punkten (Donepezil 10 mg-Gruppe) Unterschied zu Placebo und waren damit kongruent zum ADAS-cog.

In der Studie Prasher 2002 zur Alzheimer Demenz bei Menschen mit Down-Syndrom wurde ausschließlich der SIB zur Erfassung der kognitiven Funktionen verwendet. Der Unterschied zugunsten von Donepezil war statistisch nicht signifikant ($p = 0,06$), was jedoch an der geringen Stichprobengröße liegen mag.

Bei der Studie Mohs 2001, die die Zeit bis zum Erreichen eines definierten Endpunktes als Hauptzielgröße betrachtete, wurde ebenfalls über Veränderungen der kognitiven Funktionsfähigkeit (MMSE) berichtet. Hier wurde dargestellt, dass zum Endpunkt der Studie die Donepezil-Gruppe signifikant bessere MMSE-Werte aufwies. Die Interpretierbarkeit

dieses Ergebnisses ist jedoch aufgrund der Besonderheiten im Design (s. o.) stark eingeschränkt.

Zusammenfassend liegen konsistent positive Ergebnisse hinsichtlich der Verbesserung der kognitiven Funktionen in den verwendeten Skalen vor. Eine Meta-Analyse (Abbildung 4) zeigt keine Heterogenität in den Effekten und eine moderate, statistisch signifikante, Effektstärke in der Größenordnung einer halben Standardabweichung (Cohen's d = -0,51; 95 %-KI: -0,60 – -0,42). Für den ADAS-cog entspricht dies etwa 2,5 bis 3 Scorepunkten. Die Studien Moraes 2006 und Prasher 2002 konnten nicht in die Meta-Analyse eingeschlossen werden, da für den geschätzten Gruppenunterschied, gemessen an der Veränderung zwischen Baseline und Beobachtungsende, kein Variabilitätsmaß angegeben war und dieses sich aus den berichteten p-Werten auch nicht abschätzen ließ.

Tabelle 23: Donepezil: Ergebnisse zu kognitiven Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunter- schied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effektes		
Burns 1999 (24 Wochen)	DON 5 mg	ADAS- cog	(271) ^(b)	0,3 (4,9) ^(c)	-1,5 (n.g.)	↗	< 0,001
	DON 10 mg		(273) ^(b)	-1,3 (5,0) ^(c)	- 2,9 (n.g.)	↗	0,002
	Placebo	(274) ^(b)	1,8 (5,0) ^(c)				
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	sMMSE	91	1,5 (4,8) ^(c)	2,1 (n.g.)	↗	< 0,001
	Placebo		100	-0,5 (4,0) ^(c)			
Homma 2000 (24 Wochen)	DON 5 mg	ADAS- J-cog	126	-2,4 (5,1)	-2,5 (n.g.)	↗	< 0,001
	Placebo		113	0,1 (5,2)			
Krishnan 2003 (24 Wochen)	DON 10 mg	ADAS- cog	34	0,1 ^(c) (n.g.)	-3,1 ^(c) (n.g.)	↗	< 0,04
	Placebo		32	3,2 ^(c) (n.g.)			
Moraes 2006 (24 Wochen)	DON 10 mg	ADAS- cog	17	-7,4 ^(d) (n.g.)	-11,2 ^(d) (n.g.)	↗	<0,01 ^(e)
	Placebo		18	3,8 ^(d) (n.g.)			
Prasher 2002 ^(f) (24 Wochen)	DON 10 mg	SIB	27 ^(g)	-5,2 (n.g.)	10,8 (n.g.)	↗	n.g. ^(h)
	Placebo			-16,0 (n.g.)			

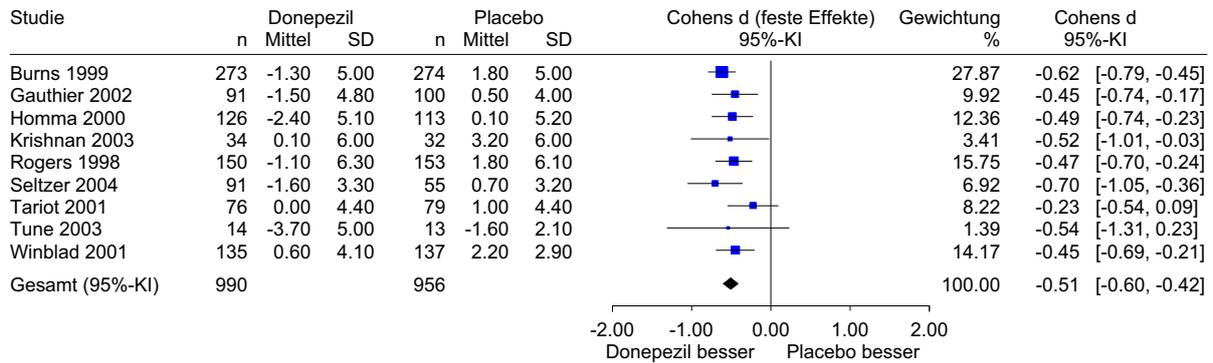
(Fortsetzung)

Tabelle 23 (Fortsetzung): Donepezil: Ergebnisse zu kognitiven Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunter- schied zu Placebo	p-Wert	
Rogers 1998 (24 Wochen)	DON 5 mg	ADAS- cog	152	-0,7 (6,3)	-2,5 (n.g.) ↗	< 0,001
	DON 10 mg		150	-1,1 (6,3)	-2,9 (n.g.) ↗	< 0,001
	Placebo		153	1,8 (6,1)		
Seltzer 2004 (24 Wochen)	DON 10 mg	ADAS- cog ¹³	91	-1,6 (3,3) ^(c)	-2,3 (n.g.) ↗	0,001
	Placebo		55	0,7 (3,2) ^(c)		
Tariot 2001 ⁽ⁱ⁾ (24 Wochen)	DON (5-) 10 mg	MMSE	76	0,0 (4,4) ^(c)	1,0 ^(c) (n.g.) ↗	„n.s.“
	Placebo		79	-1,0 (4,4) ^(c)		
Tune 2003 ^(f) (24 Wochen)	DON 10 mg	ADAS- cog	14	-3,7 (5,0)	-2,1 (-5,2; 1,0) ↗	0,186
	Placebo		13	-1,6 (2,1)		
Winblad 2001 (52 Wochen)	DON (5-) 10 mg	MMSE	135	-0,6 (4,1) ^(c)	1,6 ^(c) (n.g.) ↗	< 0,001
	Placebo		137	-2,2 (2,9) ^(c)		
<p>Erläuterung: DON = Donepezil, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in Auswertung, sofern nicht anders angegeben b: Da keine genauen Fallzahlen vorliegen, wird die ITT-Population angenommen. c: Werte aus Abbildung geschätzt, Standardabweichung jeweils eigene Berechnung d: eigene Berechnung e: p-Wert gilt für den Interaktionsfaktor Behandlung/Zeit in der zweifaktoriellen ANOVA f: unklar, ob berichtete Ergebnisse ITT-LOCF-Analyse darstellen g: Anzahl der Patienten in Auswertung aus Angaben zu Freiheitsgraden in der Publikation erschlossen, die Zuteilung zu den Gruppen ist jedoch unklar. h: In der Publikation wird lediglich ein p-Wert für den Unterschied über den gesamten Zeitraum zu mehreren Messzeitpunkten angegeben. i: Daten der Subgruppe MMSE 10-26; siehe auch Gesamtpopulation MMSE: 5-26 (n. sign.)</p>						

Eine getrennte Analyse der Studien nach der jeweiligen Dosierung (niedrige Dosis [5 mg], flexible Dosis [5-10 mg], hohe Dosis [10 mg]) zeigt, dass der beobachtete Effekt bei der höheren Dosierung größer ist (Cohen's d = -0,58; 95 %-KI: -0,70 – -0,46) als bei einer niedrigen oder flexiblen Dosierung (Abbildung 5 bis Abbildung 7).

Donepezil - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog, MMSE, SIB - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=6.43$, $df=8$ ($p=0.600$), $I^2=0\%$
 Gesamteffekt: $Z\text{ Score}=-10.99$ ($p=0.000$)

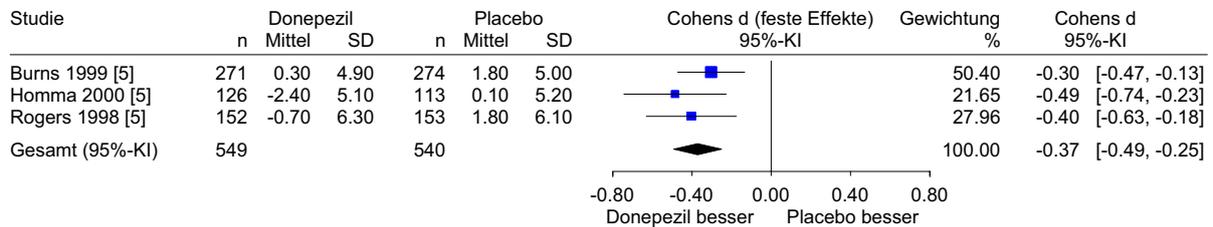
Bemerkung:

Vorzeichen bei MMSE und SIB geändert, d. h. kleine Werte entsprechen positiven Effekten

Krishnan 2003: Standardabweichungen aus $p\text{-Wert} = 0,04$ berechnet

Abbildung 4: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten

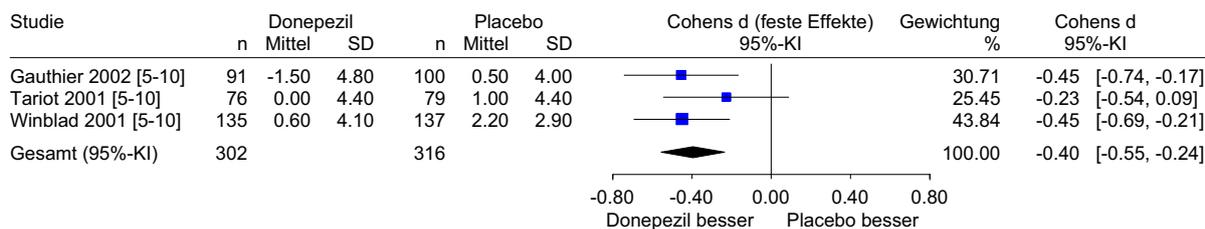
Donepezil - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog, MMSE, SIB - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=1.46$, $df=2$ ($p=0.482$), $I^2=0\%$
 Gesamteffekt: $Z\text{ Score}=-6.06$ ($p=0.000$)

Abbildung 5: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – niedrige Dosis

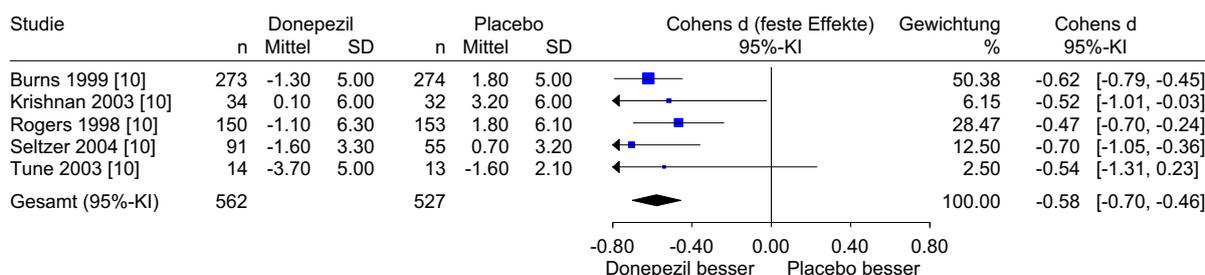
Donepezil - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog, MMSE, SIB - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=1.46$, $df=2$ ($p=0.483$), $I^2=0\%$
 Gesamteffekt: Z Score=-4.86 ($p=0.000$)

Abbildung 6: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – flexible Dosis

Donepezil - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog, MMSE, SIB - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=1.72$, $df=4$ ($p=0.788$), $I^2=0\%$
 Gesamteffekt: Z Score=-9.32 ($p=0.000$)

Abbildung 7: Donepezil: Meta-Analyse zu kognitiven Fähigkeiten – hohe Dosis

5.3.1.4 Krankheitsbezogene Lebensqualität

Instrumente zur Erfassung der Lebensqualität der Patienten wurden in den Studien von Rogers 1998 und Burns 1999 eingesetzt. In beiden Studien wurde die QoL-Skala von Blau 1977 [104] verwendet, die allerdings für die Verwendung bei Demenzpatienten nicht validiert ist [105]. In der Studie von Rogers 1998 zeigte sich im Unterschied zur Placebo-Gruppe in beiden Donepezil-Gruppen (5 mg und 10 mg) ein Trend zur Verbesserung der Lebensqualität. Allerdings war der Unterschied zur Placebo-Gruppe nur für die 5 mg-Gruppe und nur in der Woche 24 statistisch signifikant ($p = 0,05$). In der LOCF-Auswertung zum Studienendpunkt ergaben sich keine Unterschiede zwischen den Gruppen. In der Studie von Burns 1999 wurden keine Daten dargestellt, allerdings wurde auf fehlende statistisch signifikante

Unterschiede in der Lebensqualität zwischen den Gruppen mit 5 mg Donepezil, 10 mg Donepezil und Placebo bei hoher Antwortvariabilität hingewiesen.

Zusammenfassend lässt sich aus den Ergebnissen kein Nachweis für eine günstige Beeinflussung der krankheitsbezogenen Lebensqualität durch Donepezil ableiten.

5.3.1.5 Notwendigkeit einer vollstationären Pflege (Institutionalisierung)

Zu dem Zielkriterium „Notwendigkeit der vollstationären Pflege (Institutionalisierung)“ wurden lediglich in einer Veröffentlichung [67] zu der Studie Winblad 2001 Daten berichtet. Von 142 Patienten in der Donepezil-Gruppe und 144 Patienten in der Placebo-Gruppe lebten zu Studienbeginn 132 (93 %) bzw. 133 (92 %) zu Hause. Von diesen zogen während der Studienlaufzeit 9 bzw. 10 Patienten in eine andere, teurere Unterkunft (Wohneinrichtung mit Serviceleistungen, Altersheim, Wohngruppen, Pflegeheim oder andere Unterkunft). Es wird berichtet, dass in der Donepezil-Gruppe weniger Patienten in ein Pflegeheim übersiedelten als in der Placebo-Gruppe, allerdings war dieser Unterschied nicht statistisch signifikant (3 Patienten vs. 8 Patienten; $p = 0,13$). Zudem wurden zu den verschiedenen Arten der Unterbringung keine weiteren Angaben gemacht, so dass diese Auswertung nur sehr eingeschränkt interpretierbar ist.

5.3.1.6 Mortalität

Insgesamt war die Zahl der Todesfälle in den Studien gering und es zeigte sich kein auffälliger Unterschied zwischen den Gruppen. Angaben hierzu finden sich in Tabelle 24.

Einen eindeutigen Hinweis auf eine günstige oder ungünstige Beeinflussung der Mortalität geben die berichteten Daten nicht.

5.3.1.7 Unerwünschte Ereignisse

Abbrüche wegen unerwünschter Ereignisse wurden im Vergleich zur Placebo-Gruppe bei Patienten mit 10 mg-Dosierung in fünf von sechs Studien häufiger beobachtet (Odds Ratio 2,02 [95 %-KI: 1,44 – 2,84]; $p < 0,001$; $I^2 = 0$ %; Abbildung 10), bei Patienten mit flexibler Dosierung war dies in zwei von drei Studien der Fall (Odds Ratio 0,99 [95 %-KI: 0,46 – 2,12]; $p = 0,976$; $I^2 = 46$ %; Abbildung 9) und bei Patienten mit 5 mg-Dosierung in keiner Studie. Ähnlich verhält es sich mit dem Auftreten von schwerwiegenden unerwünschten Ereignissen, wobei hier die gepoolten Schätzer etwas niedriger ausfallen und keine statistische Signifikanz auch im Hochdosisbereich mehr vorliegt (Tabelle 24 und Abbildung 11 bis Abbildung 13). Auffällig ist eine moderate bis hohe Heterogenität bei den Studien mit flexibler Dosierung. Diese geht offensichtlich insbesondere auf die Studien Tariot 2001 und Winblad 2001 zurück; zu beachten ist hier das besondere Patientenkollektiv bei Tariot 2001 (Pflegeheimbewohner).

In drei Studien fanden sich keine detaillierten Angaben zu unerwünschten Ereignissen (Krishnan 2003, Moraes 2006 und Tune 2003), bei Prasher 2002 waren die Angaben inkonsistent, so dass sie nachfolgend nicht mit aufgeführt werden. Im Einklang mit dem Wirkprinzip der Acetylcholinesterasehemmer traten typische unerwünschte Wirkungen (Übelkeit, Erbrechen, Durchfall) ca. 3- bis 4-mal häufiger in den 10 mg-Verum-Gruppen als in der Placebo-Gruppe auf (Tabelle 25). In der 5 mg-Gruppe dagegen waren unerwünschte Ereignisse nur wenig häufiger als in der Placebo-Gruppe zu verzeichnen. Appetitlosigkeit wurde etwa 3-mal häufiger und Gewichtsverlust ca. doppelt so häufig in der 10 mg-Gruppe wie unter Placebo beobachtet. Weitere unerwünschte Ereignisse, die in den Verum-Gruppen häufiger auftraten, waren Schwindel, Kopfschmerzen und abdominelle Beschwerden.

Meta-Analysen bestätigen die Dosis-Wirkungs-Beziehung bei den o. g. unerwünschten Ereignissen mit statistisch signifikanter Risikoerhöhung für die Gruppen mit 10 mg-Gabe (Abbildung 8 bis Abbildung 27). Lediglich für Durchfall ist das Risiko auch in den 5 mg-Gruppen statistisch signifikant erhöht. Die Ergebnisse zeigen bis auf den oben angesprochenen Aspekt keine auffällige Heterogenität.

Tabelle 24: Donepezil: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Studie (Dauer)		Aufdosierung innerhalb des ersten Monats	N ^(a)	Studienabbruch N (%)	Todesfälle N	Schwerwiegende unerwünschte Ereignisse N (%)	Abbruch wegen unerwünschter Ereignisse N (%)	Gesamte unerwünschte Ereignisse N (%)
Burns 1999 (24 Wochen)	DON 5 mg	5 mg	271	60 (22)	1	19 (7)	24 (9)	213 (79)
	DON 10 mg	10 mg	273	72 (26)	2	29 (11)	50 (18)	234 (86)
	Placebo		274	55 (20)	2	25 (9)	27 (10)	207 (76)
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	10 mg	102	19 ^(b) (19)	1 ^(b)	14 ^(c) (14)	9 (9) ^(b)	84 (82)
	Placebo		105	12 ^(b) (11)	0 ^(b)	13 ^(c) (12)	5 (5) ^(b)	84 (80)
Homma 2000 (24 Wochen)	DON 5 mg	5 mg	136	17 (13) ^(b)	k.A.	k.A.	2 (1)	54 (40)
	Placebo		131 ^(d)	22 (17) ^(b)	k.A.	k.A.	6 (5)	33 (25)
Krishnan 2003 (24 Wochen)	DON 10 mg	5 mg	34	6 (18)	k.A.	k.A.	0	32 ^(e) (94)
	Placebo		33	10 (30)	k.A.	k.A.	1 (3 ^(e))	28 ^(e) (85)
Mohs 2001 (54 Wochen)	DON 10 mg	5 mg	214	60 (28)	3	29 ^(f) (14)	20 (9)	k.A.
	Placebo		217	56 (26)	4	23 ^(f) (11)	12 (6)	k.A.
Prasher 2002 (24 Wochen)	DON 10 mg	5 mg	16	2 (13)	k.A.	Inkonsistenzen ^(g)	k.A.	k.A.
	Placebo		15	2 (13)	k.A.	Inkonsistenzen ^(g)	k.A.	k.A.
Rogers 1998 (24 Wochen)	DON 5 mg	5mg	154	23 ^(c) (15)	0	7 (5)	9 ^(c) (6)	k.A.
	DON 10 mg	10 mg	157	50 ^(c) (32)	1	15 (10)	25 ^(c) (16)	k.A.
	Placebo		162	32 ^(c) (20)	1	9 (6)	11 ^(c) (7)	k.A.

(Fortsetzung)

Tabelle 24 (Fortsetzung): Donepezil: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Studie (Dauer)	Aufdosierung innerhalb des ersten Monats	N ^(a)	Studienabbruch N (%)	Todesfälle N	Schwerwiegende unerwünschte Ereignisse N (%)	Abbruch wegen unerwünschter Ereignisse N (%)	Gesamte unerwünschte Ereignisse N (%)
Seltzer 2004 (24 Wochen)	DON 10 mg	96	26 (27)	k.A.	5 (5)	15 (16)	67 (70)
	Placebo	57	11 (19)	k.A.	3 (5)	5 (9)	37 (65)
Tariot 2001 ^(h) (24 Wochen)	DON (5-) 10 mg	103	19 (18)	3	10 ^(c) (10)	11 (11)	99 (96)
	Placebo	105	27 (26)	7	17 ^(c) (16)	19 (18)	102 (97)
Tune 2003 (24 Wochen)	DON 10 mg	14	0	k.A.	k.A.	k.A.	k.A.
	Placebo	14	2 (14 ^(e))	k.A.	k.A.	k.A.	k.A.
Winblad 2001 (24 Wochen)	DON (5-) 10 mg	142	47 (33)	4	35 (25)	10 (7)	116 (82)
	Placebo	144	47 (33)	3	20 (14)	9 (6)	109 (76)
<p>Erläuterung: DON = Donepezil, k.A. = keine Angaben, N = Anzahl, Wo. = Woche</p> <p>a: Anzahl randomisierter Patienten b: Angaben der Herstellerfirma Eisai im Rahmen des Stellungnahmeverfahrens (vgl. SN Eisai) c: aus Prozentzahl berechnet d: Bei der Safety-Analyse wurde ein Patient in der Placebo-Gruppe ausgeschlossen, der nach Randomisierung zu keiner Untersuchung mehr erschien. e: aus N berechnet f: einschließlich Todesfälle (wurden in der Publikation gesondert berichtet) g: In der Publikation werden inkonsistente Angaben zur Anzahl der Patienten mit schwerwiegenden unerwünschten Ereignissen gemacht; Donepezil-Gruppe: 0, 8 bzw. 12 Patienten, Placebo-Gruppe: 0, 3 bzw. 7 Patienten. h: Daten gelten für die Gesamtpopulation der betrachteten Patienten (MMSE 5-26), da diese Angaben für die leicht bis mittelschwer beeinträchtigten Patienten nicht vollständig vorliegen.</p>							

Tabelle 25: Donepezil: unerwünschte Ereignisse

Studie (Dauer)	Auf- dosierung innerhalb des ersten Monats	N ^(a)	Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein	
														Anteil (%) mit UE
Burns 1999 (24 Wochen)	DON 5 mg	5 mg	271	7	4	10	5	4	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
	DON 10 mg	10 mg	273	24	16	16	9	8	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
	Placebo		274	7	4	4	5	1	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	10 mg	102	8	7	13	8	k.A.	8	k.A.	11	7	k.A.	k.A.
	Placebo		105	4	3	6	4	k.A.	4	k.A.	4	8	k.A.	k.A.
Homma 2000 (24 Wochen)	DON 5 mg	5 mg	136	4	1	4	k.A.	1	k.A.	0 ^(c)	3	1	k.A.	k.A.
	Placebo		131 ^(b)	1	2	3	k.A.	2	k.A.	2 ^(c)	1	2	k.A.	k.A.
Mohs 2001 (54 Wochen)	DON 10 mg	5 mg	214	9	k.A.	17	k.A.	6	4	13	9	6 ^(d)	k.A.	k.A.
	Placebo		217	4	k.A.	5	k.A.	2	6	10	3	1 ^(d)	k.A.	k.A.
Rogers 1998 (24 Wochen)	DON 5 mg	5 mg	154	4	3	9	10	2	k.A.	k.A.	k.A.	k.A.	5	k.A.
	DON 10 mg	10 mg	157	17	10	17	8	7	k.A.	k.A.	k.A.	k.A.	8	k.A.
	Placebo		162	4	2	7	4	2	k.A.	k.A.	k.A.	k.A.	2	k.A.

(Fortsetzung)

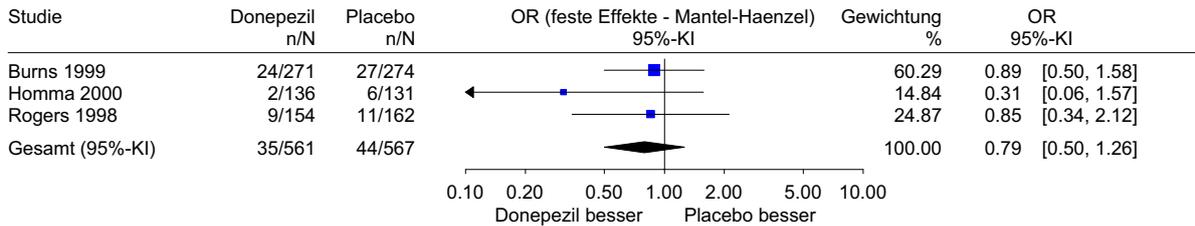
Tabelle 25 (Fortsetzung): Donepezil: unerwünschte Ereignisse

Studie (Dauer)	Auf- dosierung innerhalb des ersten Monats	N ^(a)	Anteil (%) mit UE										
			Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein
Seltzer 2004 (24 Wochen)	DON 10 mg	96	10	k.A.	20	8	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
	Placebo	57	4	k.A.	9	2	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
Tariot 2001 ^(c) (24 Wochen)	DON (5-) 10 mg	103	9	15	15	8	9	19	10	15	10	k.A.	k.A.
	Placebo	105	4	14	10	8	5	10	8	16	5	k.A.	k.A.
Winblad 2001 (52 Wochen)	DON (5-) 10 mg	142	11	k.A.	7	6	k.A.	k.A.	k.A.	8	2	k.A.	k.A.
	Placebo	144	9	k.A.	7	4	k.A.	k.A.	k.A.	6	6	k.A.	k.A.

Erläuterungen:
DON = Donepezil, k.A. = keine Angaben, N = Anzahl, UE = unerwünschte Ereignisse, Wo. = Woche

a: Anzahl randomisierter Patienten
b: Bei der Safety-Analyse wurde ein Patient in der Placebo-Gruppe ausgeschlossen, der nach Randomisierung zu keiner Untersuchung mehr erschien.
c: „restlessness“
d: „dyspepsia“
e: Daten gelten für die Gesamtpopulation der betrachteten Patienten (MMSE 5-26), da diese Angaben für die leicht bis mittelschwer beeinträchtigten Patienten nicht vollständig vorliegen.

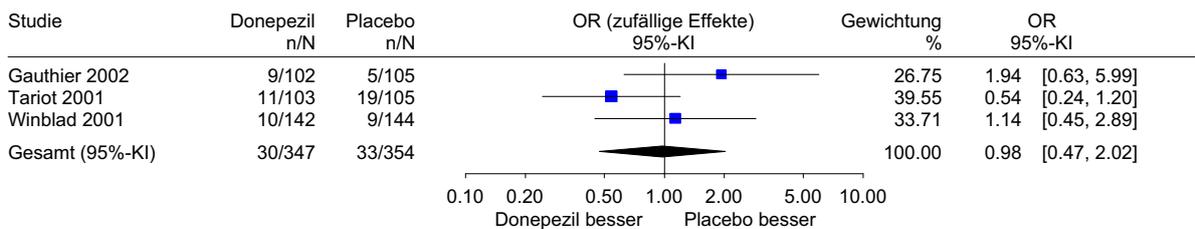
Donepezil - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



Heterogenität: $Q=1.46$, $df=2$ ($p=0.482$), $I^2=0\%$
 Gesamteffekt: Z Score=-0.98 ($p=0.328$)

Abbildung 8: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – niedrige Dosis

Donepezil - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



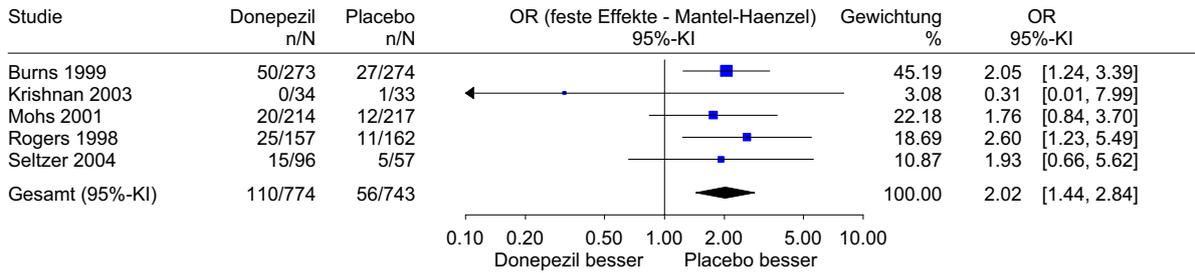
Heterogenität: $Q=3.56$, $df=2$ ($p=0.169$), $I^2=43.8\%$
 Gesamteffekt: Z Score=-0.06 ($p=0.950$), $\tau^2=0.181$

Bemerkung:

Aufgrund der Heterogenität wurde ein Modell mit zufälligen Effekten gewählt. Ein Modell mit festen Effekten ergibt ähnliche Ergebnisse.

Abbildung 9: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – flexible Dosis

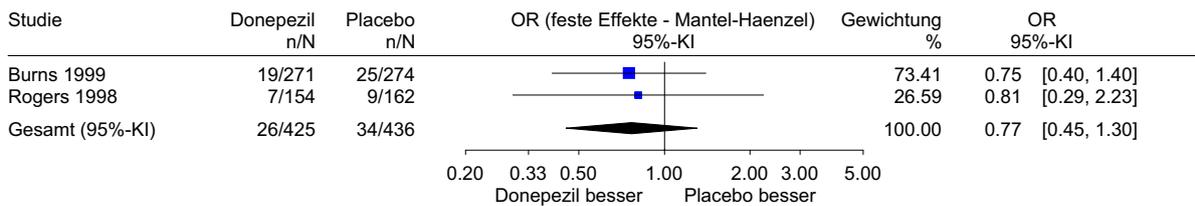
Donepezil - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



Heterogenität: $Q=1.85$, $df=4$ ($p=0.763$), $I^2=0\%$
 Gesamteffekt: $Z\ Score=4.06$ ($p=0.000$)

Abbildung 10: Donepezil: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse – hohe Dosis

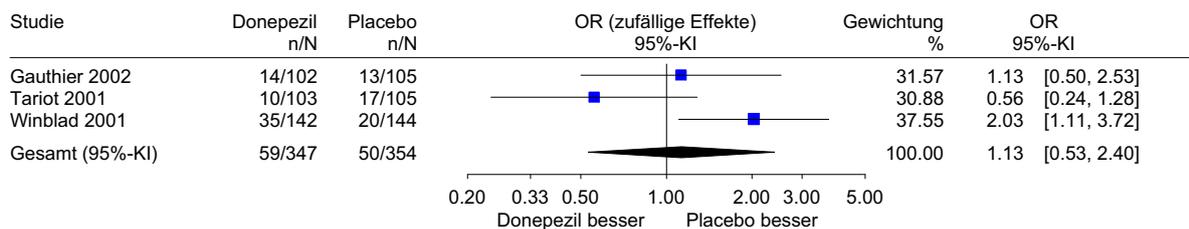
Donepezil - Schwerwiegende UEs
 Endpunkt: SUE erlitten (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit SUE



Heterogenität: $Q=0.02$, $df=1$ ($p=0.901$), $I^2=0\%$
 Gesamteffekt: $Z\ Score=-0.98$ ($p=0.325$)

Abbildung 11: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – niedrige Dosis

Donepezil - Schwerwiegende UEs
 Endpunkt: SUE erlitten (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit SUE



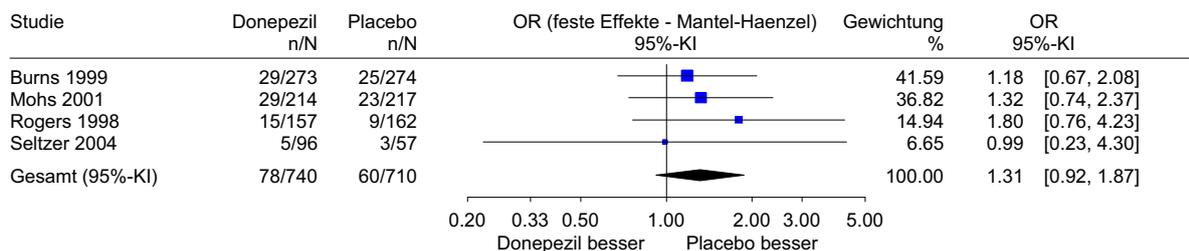
Heterogenität: $Q=6.12$, $df=2$ ($p=0.047$), $I^2=67.3\%$
 Gesamteffekt: Z Score=0.32 ($p=0.751$), $\tau^2=0.298$

Bemerkung:

Aufgrund der Heterogenität wurde ein Modell mit zufälligen Effekten gewählt. Ein Modell mit festen Effekten ergibt ähnliche Ergebnisse.

Abbildung 12: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – flexible Dosis

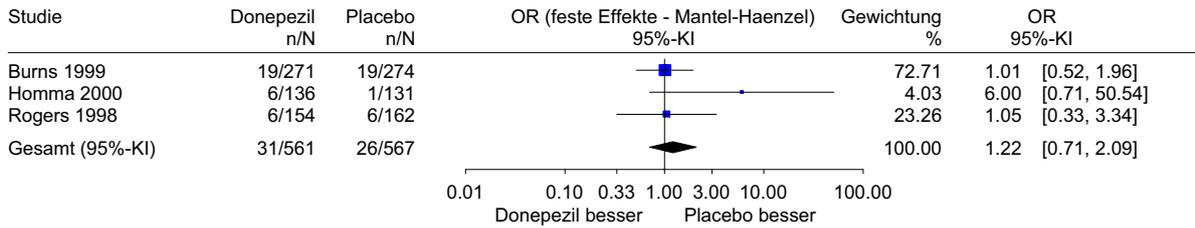
Donepezil - Schwerwiegende UEs
 Endpunkt: SUE erlitten (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit SUE



Heterogenität: $Q=0.79$, $df=3$ ($p=0.853$), $I^2=0\%$
 Gesamteffekt: Z Score=1.51 ($p=0.132$)

Abbildung 13: Donepezil: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen – hohe Dosis

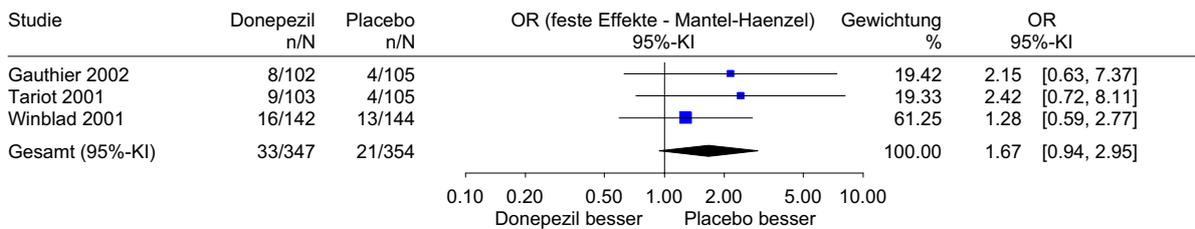
Donepezil - UEs
 Endpunkt: Übelkeit
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=2.52$, $df=2$ ($p=0.284$), $I^2=20.7\%$
 Gesamteffekt: Z Score=0.73 ($p=0.463$)

Abbildung 14: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – niedrige Dosis

Donepezil - UEs
 Endpunkt: Übelkeit
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



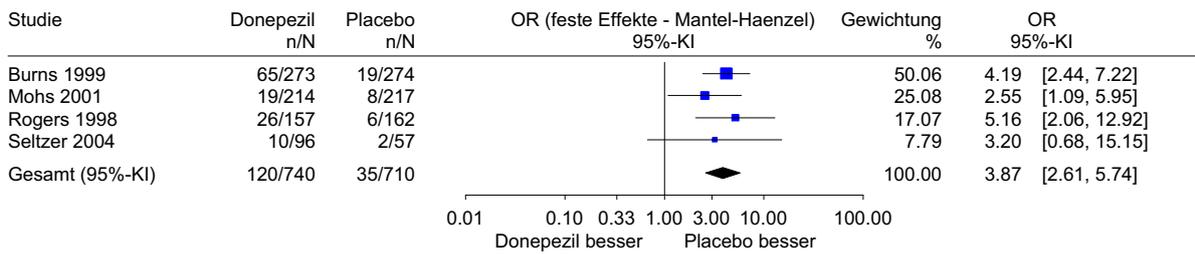
Heterogenität: $Q=0.98$, $df=2$ ($p=0.614$), $I^2=0\%$
 Gesamteffekt: Z Score=1.76 ($p=0.079$)

Abbildung 15: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – flexible Dosis

Donepezil - UEs

Endpunkt: Übelkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.46$, $df=3$ ($p=0.693$), $I^2=0\%$

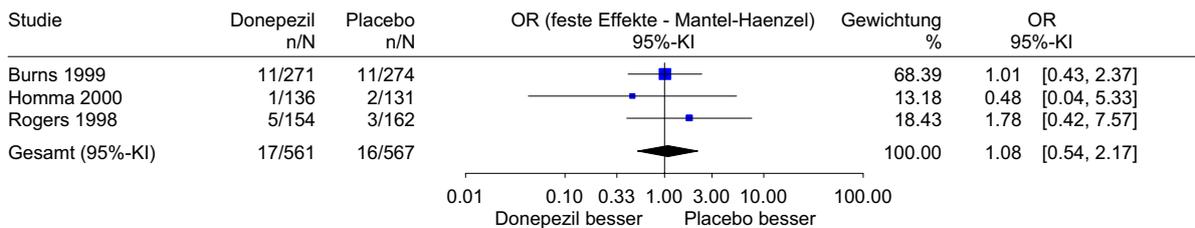
Gesamteffekt: Z Score=6.72 ($p=0.000$)

Abbildung 16: Donepezil: Meta-Analyse zum Endpunkt „Übelkeit“ – hohe Dosis

Donepezil - UEs

Endpunkt: Erbrechen

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf

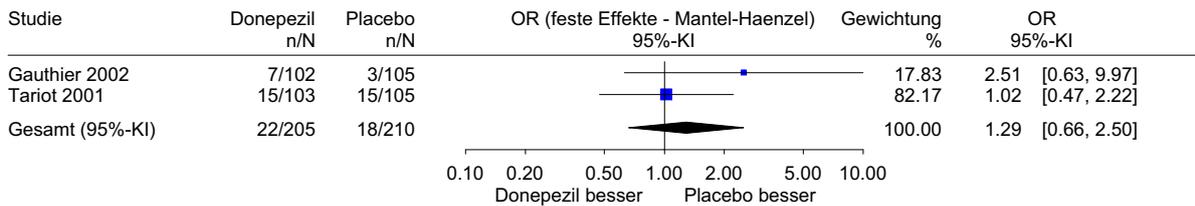


Heterogenität: $Q=0.92$, $df=2$ ($p=0.632$), $I^2=0\%$

Gesamteffekt: Z Score=0.22 ($p=0.823$)

Abbildung 17: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – niedrige Dosis

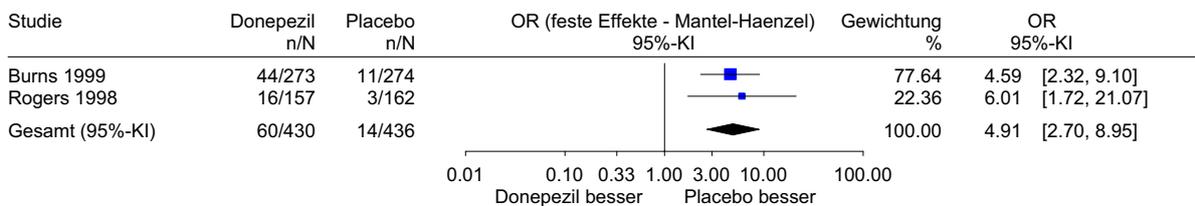
Donepezil - UEs
 Endpunkt: Erbrechen
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.23$, $df=1$ ($p=0.267$), $I^2=18.9\%$
 Gesamteffekt: $Z\text{ Score}=0.74$ ($p=0.456$)

Abbildung 18: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – flexible Dosis

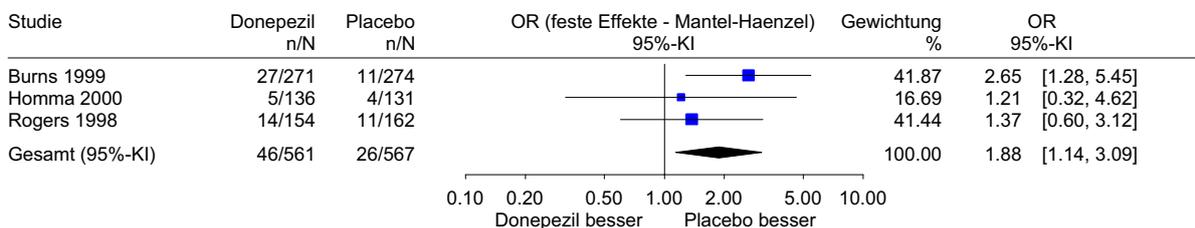
Donepezil - UEs
 Endpunkt: Erbrechen
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=0.14$, $df=1$ ($p=0.711$), $I^2=0\%$
 Gesamteffekt: $Z\text{ Score}=5.2$ ($p=0.000$)

Abbildung 19: Donepezil: Meta-Analyse zum Endpunkt „Erbrechen“ – hohe Dosis

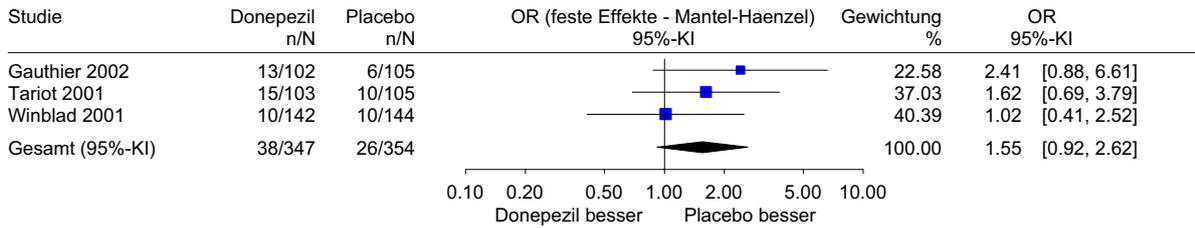
Donepezil - UEs
 Endpunkt: Diarrhoe
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.83$, $df=2$ ($p=0.400$), $I^2=0\%$
 Gesamteffekt: $Z\text{ Score}=2.49$ ($p=0.013$)

Abbildung 20: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – niedrige Dosis

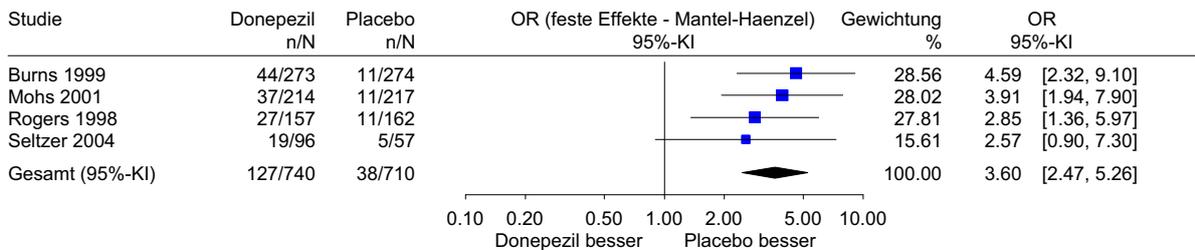
Donepezil - UEs
 Endpunkt: Diarrhoe
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.58$, $df=2$ ($p=0.454$), $I^2=0\%$
 Gesamteffekt: Z Score=1.65 ($p=0.099$)

Abbildung 21: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – flexible Dosis

Donepezil - UEs
 Endpunkt: Diarrhoe
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



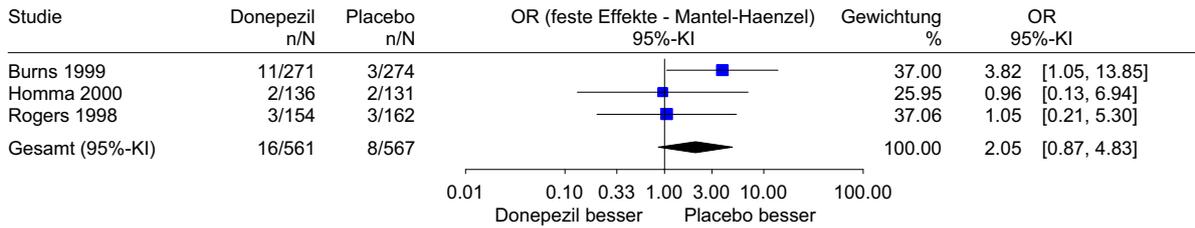
Heterogenität: $Q=1.33$, $df=3$ ($p=0.723$), $I^2=0\%$
 Gesamteffekt: Z Score=6.62 ($p=0.000$)

Abbildung 22: Donepezil: Meta-Analyse zum Endpunkt „Diarrhoe“ – hohe Dosis

Donepezil - UEs

Endpunkt: Appetitlosigkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=2.12$, $df=2$ ($p=0.347$), $I^2=5.5\%$

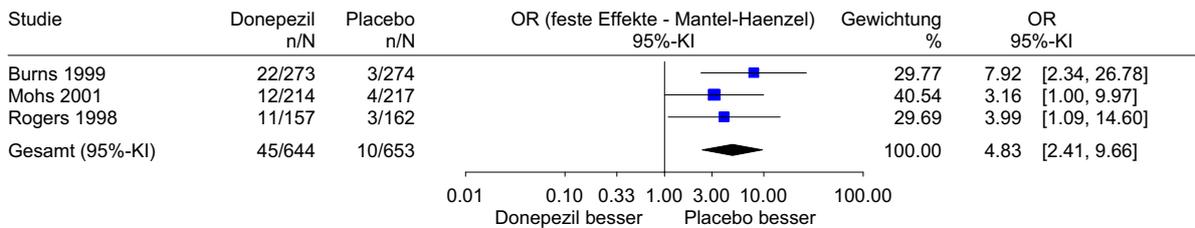
Gesamteffekt: Z Score=1.65 ($p=0.099$)

Abbildung 23: Donepezil: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ – niedrige Dosis

Donepezil - UEs

Endpunkt: Appetitlosigkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf

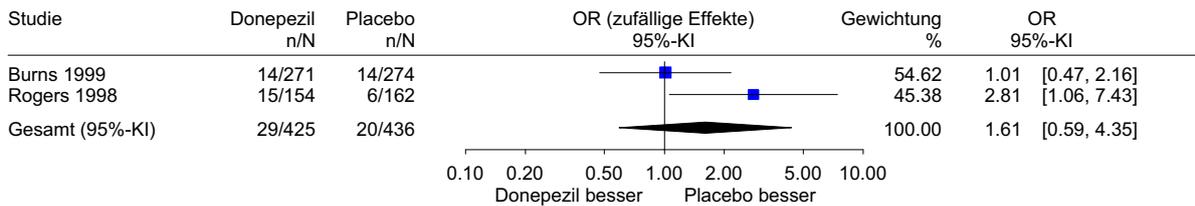


Heterogenität: $Q=1.24$, $df=2$ ($p=0.539$), $I^2=0\%$

Gesamteffekt: Z Score=4.45 ($p=0.000$)

Abbildung 24: Donepezil: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ – hohe Dosis

Donepezil - UEs
 Endpunkt: Schwindel
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



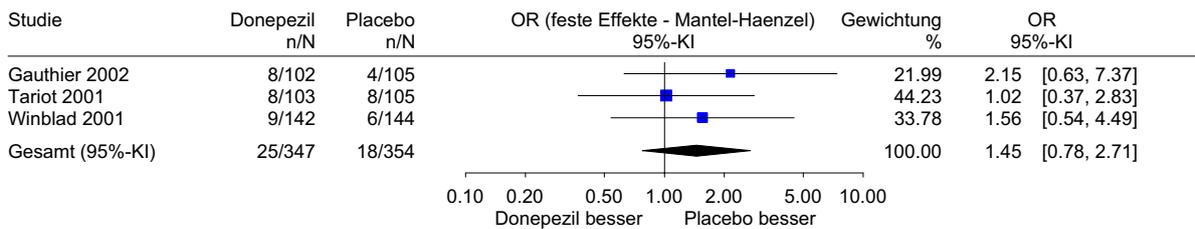
Heterogenität: $Q=2.62$, $df=1$ ($p=0.105$), $I^2=61.9\%$
 Gesamteffekt: Z Score=0.93 ($p=0.351$), $\tau^2=0.323$

Bemerkung:

Aufgrund der Heterogenität wurde ein Modell mit zufälligen Effekten gewählt. Ein Modell mit festen Effekten ergibt ähnliche Ergebnisse.

Abbildung 25: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – niedrige Dosis

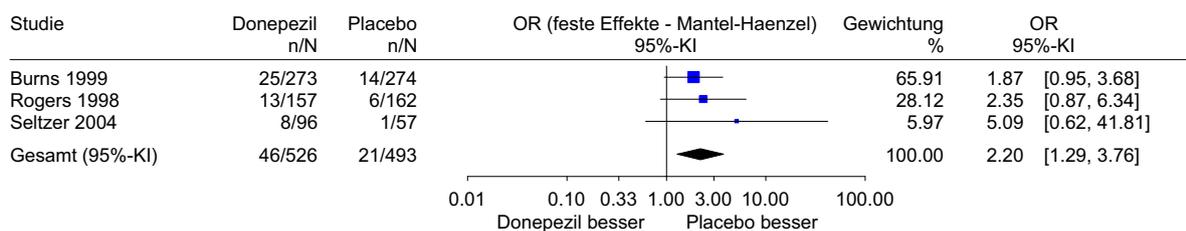
Donepezil - UEs
 Endpunkt: Schwindel
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=0.86$, $df=2$ ($p=0.650$), $I^2=0\%$
 Gesamteffekt: Z Score=1.16 ($p=0.244$)

Abbildung 26: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – flexible Dosis

Donepezil - UEs
 Endpunkt: Schwindel
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=0.84$, $df=2$ ($p=0.656$), $I^2=0\%$
 Gesamteffekt: Z Score=2.88 ($p=0.004$)

Abbildung 27: Donepezil: Meta-Analyse zum Endpunkt „Schwindel“ – hohe Dosis

5.3.1.8 Lebensqualität der (betreuenden) Angehörigen

Bei Gauthier 2002 wurde in einer Zusatzpublikation zu dieser Studie [54] der durch die Betreuung der Patienten erzeugte Stress für die Angehörigen („caregiver stress“) als Zielkriterium erfasst und dafür als Erhebungsinstrument eine für die Alzheimer Demenz modifizierte Version der Caregiver Stress Scale (CSS) eingesetzt. Einschlusskriterium für Patienten in die Studie war, dass ein betreuender Angehöriger („caregiver“) für mindestens 24 Stunden in der Woche mit dem Erkrankten zusammen sein musste. Auswertungen lagen lediglich für die Gesamtgruppe der Patienten mit mittelschwerer und schwerer Demenz (MMSE 5-17) vor. Zum Studienendpunkt (ITT-LOCF) war in der Donepezil-Gruppe praktisch keine Veränderung im Gesamtscore zu verzeichnen, während es in der Placebo-Gruppe zu einem leichten (Stress-) Anstieg kam; der Unterschied zwischen den Gruppen war statistisch nicht signifikant. Ähnliches spiegelte sich in der Auswertung der 11 Subdomänen des Scores wider: In sechs kam es zu einer günstigen Beeinflussung durch Donepezil (davon in einer statistisch signifikant), in zwei war kein größerer Unterschied erkennbar und in drei war die Placebo-Gruppe besser (ohne statistische Signifikanz).

Zusammenfassend lässt sich aus den Ergebnissen kein Nachweis für eine günstige Beeinflussung der Lebensqualität der betreuenden Angehörigen durch Donepezil ableiten.

5.3.1.9 Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en

Bei Gauthier 2002 bzw. der Zusatzpublikation [54] wurde auch die Betreuungszeit durch die pflegenden Angehörigen als Zielkriterium ermittelt. Hierzu gaben die Betreuer an, wie viele Minuten sie pro Tag bei den instrumentellen sowie den basalen Aktivitäten der Patienten (erfasst anhand des IADL-plus bzw. des PSMS-plus) assistierten. Auswertungen lagen wiederum nur für die Gesamtgruppe der Patienten mit mittelschwerer und schwerer Demenz

(MMSE 5-17) vor. Die Pflegezeiten für die Hilfe bei den instrumentellen und basalen Aktivitäten waren in der Verum-Gruppe zum Ausgangszeitpunkt zwar deutlich länger als in der Placebo-Gruppe (49 Minuten), allerdings wurde dies in der statistischen Analyse berücksichtigt. Zum Studienende (ITT-LOCF) betrug der Gruppenunterschied in der Veränderung der Pflegezeiten etwa 50 Minuten ($p = 0,004$). Während sich in der Donepezil-Gruppe der Pflegeaufwand von etwa 325 Minuten auf etwa 300 Minuten verringerte, nahm dieser in der Placebo-Gruppe von etwa 275 auf etwa 300 Minuten zu. Die absolute Dauer der Pflegezeit innerhalb der beiden Gruppen hatte sich somit offenbar zum Studienende angeglichen. Da diese Ergebnisse möglicherweise durch eine Regression zur Mitte erklärt werden könnten, ist die Validität der Daten unklar.

Zur Studie von Winblad 2001 wurden ebenfalls Daten zum zeitlichen Pflegeaufwand publiziert [66]. Als Zielgröße wurde die Pflegezeit mit einer Skala, der „Resource Utilization in Dementia“ (RUD), erfasst. Nach 52 Wochen wurde zwar ein statistisch signifikanter Unterschied in der Zunahme des Zeitbedarfs zugunsten von Donepezil (Differenz ca. 60 Minuten, $p = 0,03$) beobachtet, allerdings beruht diese Analyse auf nur noch etwa 50 % der ursprünglich randomisierten Patienten, so dass das Ergebnis nicht mit ausreichender Sicherheit zu interpretieren ist.

Die Interpretation der dargestellten Ergebnisse wird über die oben dargestellten spezifischen Aspekte hinaus weiter eingeschränkt: Wenn sich Patienten im Krankenhaus befanden, wurde bei Winblad 2001 der zeitliche Bedarf für den Pflegeaufwand für den entsprechenden Zeitraum offenbar auf den Wert 0 gesetzt. Dieses Vorgehen erscheint nicht adäquat, da so ein unerwünschtes Ereignis (Notwendigkeit einer Hospitalisierung) die Schätzung des Pflegeaufwands in eine günstige Richtung beeinflusst. In der Publikation wurde nicht berichtet, wie oft und in welchem Umfang ein solches Vorgehen erforderlich war. Von der Firma Pfizer wurde im Stellungnahmeverfahren dargelegt, dass es während der Studie bei 9 von 96 Donepezil- und 12 von 94 Placebo-Patienten zu einer Änderung der Wohnverhältnisse kam. 5 der 9 Donepezil- und 7 der 12 Placebo-Patienten brachen daraufhin die Studie ab, so dass damit in der Placebo-Gruppe der Betreuungsaufwand häufiger auf 0 gesetzt wurde (vgl. SN Pfizer). Allerdings klären diese Angaben nicht das Problem, dass beispielsweise bei einem Krankenhausaufenthalt die Betreuungszeit mit einem Nullwert in die Analyse eingeht.

Bei Gauthier 2002 wurde in der Publikation das zugrunde liegende Problem gar nicht erst dargestellt. Im Stellungnahmeverfahren wurde von der Firma Pfizer berichtet, dass für Patienten, für die keine Pflegezeit berichtet worden war (z. B. weil der Patient sich im Krankenhaus befand), der LOCF-Ansatz verwendet wurde (vgl. SN Pfizer). Hiervon unberührt bleibt die oben beschriebene Unsicherheit bezüglich der berichteten Ergebnisse.

Zusammenfassend lässt sich aus den Ergebnissen kein sicherer Hinweis auf eine günstige Beeinflussung des Betreuungsaufwandes durch Donepezil ableiten.

5.3.1.10 Ergänzende Informationen: klinisches Krankheitsstadium

Der klinische Gesamteindruck wurde in vier Studien unter Verwendung des CIBIC-plus bzw. des J-CGIC (Japanese-Clinical Global Impression of Change) ermittelt und war dort jeweils primäres Zielkriterium (Burns 1999, Gauthier 2002, Homma 2000 und Rogers 1998; Tabelle 26). Der Unterschied im Anteil der als unverändert bzw. gebessert beurteilten Patienten zwischen den Verum- und den Placebo-Gruppen betrug zwischen absolut 8 % (Donepezil 5 mg, Burns 1999) und 24 % (Donepezil 5 mg, Homma 2000) bzw. 23 % (Donepezil 10 mg, Gauthier 2002) zugunsten Donepezil. Die Studie von Rogers 1998 gab nur Veränderungen bezüglich der Anzahl der verbesserten Patienten (nicht unverändert und gebessert) an.

Bei Winblad 2001 wurde die Gottfries-Bråne-Steen-Skala (GBS) zur Einschätzung des Gesamteindrucks verwendet und stellte hier ebenfalls das primäre Zielkriterium dar. Während sich die mittlere Veränderung auf der Skala bei den beobachteten Patienten zwischen den Gruppen in den Wochen 24, 36 und 52 statistisch signifikant zugunsten der Donepezil-Gruppe unterschied ($p < 0,05$), war der Unterschied in der ITT-LOCF-Analyse zum Endpunkt nicht mehr statistisch signifikant ($p = 0,054$). Weiterhin wurde der Anteil von Patienten mit einer Verbesserung des Gesamteindrucks zu den Zeitpunkten 12, 24 und 52 Wochen berichtet. Der Unterschied zwischen den Gruppen betrug zwischen absolut 6 % (24 Wochen) und 10 % (52 Wochen).

In den Studien Seltzer 2004 und Tariot 2001 wurde ebenfalls nicht der CIBIC eingesetzt, sondern der CDR als globales Maß verwendet. In der Studie Seltzer 2004, die ausschließlich leicht beeinträchtigte Patienten (MMSE 21-26) einschloss, wurde lediglich berichtet, dass es hinsichtlich des CDR-SB keine Unterschiede zwischen den Behandlungsgruppen gab. Bei Tariot 2001 dagegen zeigte sich in der ITT-LOCF-Analyse nach 24 Wochen ein statistisch signifikanter Unterschied ($p < 0,05$) im CDR-SB (Nursing Home Version) sowohl für die betrachtete Gesamtpopulation als auch für die Subgruppe der Patienten mit leichter bis mittelschwerer Alzheimer Demenz (MMSE 10-26).

Mohs 2001 berichtete, dass die Donepezil-Gruppe zu mehreren Zeitpunkten, jedoch nicht zum Studienende, statistisch signifikant bessere Werte im CDR-SB aufwies. Das Design der Studie lässt Ergebnisse außerhalb der (primären) Überlebenszeitanalyse allerdings nicht mehr sicher interpretieren (siehe Abschnitt 5.2.2.1).

Zusammenfassend zeigt sich, dass Donepezil den klinischen Gesamteindruck im Vergleich zu Placebo verbessert. Eine Meta-Analyse erscheint aufgrund der nicht einheitlichen Instrumente und teilweise unterschiedlichen Operationalisierungen nicht sinnvoll. Darüber hinaus ist die Validität dieses Surrogats im Hinblick auf die Patientenrelevanz unklar (vgl. Abschnitt 4.5).

Tabelle 26: Donepezil: Ergebnisse zum klinischen Gesamteindruck

Studie (Dauer)	End punkt	N ^(a)	Anteil (%) mit Score ≤ 4 ^(b)	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Burns 1999 (24 Wochen)	DON 5 mg	CIBIC- plus	271 ^(c)	57	8% (n.g.)	↗	0,007 ^(d)
	DON 10 mg		273 ^(c)	63	14% (n.g.)	↗	< 0,001 ^(d)
	Placebo		274 ^(c)	49			
Gauthier 2002 (24 Wochen)	DON (5-) 10 mg	CIBIC- plus	98	70	23% (n.g.)	↗	< 0,001 ^(d)
	Placebo		105	47			
Homma 2000 (24 Wochen)	DON 5 mg	J-CGIC	133	81	23% (n.g.)	↗	< 0,001 ^(d)
	Placebo		128	58			
Rogers 1998 (24 Wochen)	DON 5 mg	CIBIC- plus	149	26 ^(e)	15% (n.g.)	↗	0,005 ^(d)
	DON 10 mg		149	25 ^(e)	14% (n.g.)	↗	< 0,001 ^(d)
	Placebo		152	11 ^(e)			
Mittlere Differenz zu Baseline (SD)							
Seltzer 2004 (24 Wochen)	DON 10 mg	CDR- SB	96 ^(c)	n.g.	n.g. (n.g.)		„n.s.“
	Placebo		57 ^(c)	n.g.			
Tariot 2001 (24 Wochen)	DON (5-) 10 mg	CDR- SB	76	-0,2 (2,8) ^(f)	-1,0 ^(f) (n.g.)	↗	< 0,05
	Placebo		79	0,8 (2,6) ^(f)			
Winblad 2001 (52 Wochen)	DON (5-) 10 mg	GBS	138	8 (18) ^(f)	-4 ^(f) (n.g.)	↗	0,054
	Placebo		144	12 (18) ^(f)			
<p>Erläuterung: DON = Donepezil, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes. ⁶ a: Anzahl ausgewerteter Patienten, sofern nicht anders angegeben b: Stabilität oder Verbesserung c: Anzahl randomisierter Patienten, Anzahl ausgewerteter Patienten in Publikation nicht genannt d: p-Werte beruhen nicht auf einer Auswertung der dichotomisierten Daten. e: Kriterium CIBIC ≤ 3 (nur Verbesserung) f: Werte aus Abbildung geschätzt</p>							

5.3.2 Galantamin

5.3.2.1 Aktivitäten des täglichen Lebens

In allen sechs Studien zu Galantamin im Vergleich gegen Placebo waren die Aktivitäten des täglichen Lebens sekundäre Zielkriterien (Tabelle 27). In vier Studien war die Veränderung des erreichten Wertes auf der Disability Assessment for Dementia Scale (DAD) das betrachtete Ergebnismaß (Erkinjuntti 2002, Raskind 2000, Rockwood 2006, Wilcock 2000). In zwei Studien (Brodaty 2005, Tariot 2000) kam die AD Cooperative Study Activities of Daily Living-Scale (ADCS-ADL) [106] zur Anwendung. Bei Erkinjuntti 2002 zeigte sich laut bislang unveröffentlichten Daten der Firma Janssen-Cilag auf der DAD-Skala nach 6 Monaten ein statistisch signifikanter Vorteil von Galantamin 24 mg gegenüber Placebo in der Subgruppe der Patienten mit Alzheimer Demenz und zerebrovaskulärer Erkrankung (AD-CVD-Subgruppe). Zur Studie Raskind 2000 lagen – wie auch zu Erkinjuntti 2002 – zusätzlich zur Publikation Daten des Herstellers Janssen-Cilag vor. Zwischen den Patienten der Placebo-Gruppe und der 24 mg Galantamin-Gruppe ergaben sich zum Studienende keine statistisch signifikanten Unterschiede in der Veränderung des DAD-Wertes. Auch bei Rockwood 2006 wurde berichtet, dass sich die betrachteten Gruppen (Placebo vs. 16–24 mg Galantamin) nach 4 Monaten im primären Auswertungsverfahren (ANOVA) nicht statistisch signifikant unterschieden ($p = 0,13$). In einer Mixed-Effects-Analyse, bei der die Differenzen im Schweregrad der Demenz zwischen den Gruppen zu Studienbeginn berücksichtigt wurden, lag der p-Wert an der Schwelle zur statistischen Signifikanz ($p = 0,051$). In der Studie Wilcock 2000 ergab sich zwischen der 24 mg Galantamin-Gruppe und der Placebo-Gruppe ebenfalls kein statistisch signifikanter Unterschied.

Auf der ADCS-ADL-Skala war bei Brodaty 2005 die Behandlung sowohl mit der retardierten (prolonged release) als auch mit der nichtretardierten Form von Galantamin (16–24 mg) in der LOCF-Analyse der Behandlung mit Placebo überlegen. In der Studie von Tariot 2000 waren die Ergebnisse hinsichtlich der ADCS-ADL in der 16 mg- und der 24 mg-Gruppe statistisch signifikant besser als in der Placebo-Gruppe, wobei die 16 mg-Gruppe eine etwas höhere Differenz zu Placebo aufwies. In der 8 mg-Gruppe zeigte sich kein statistisch signifikanter Effekt.

In der Studie Rockwood 2006 stellte der Grad der Erreichung von Behandlungszielen das primäre Ergebniskriterium dar. Hierzu wurden zum einen von Behandlern, zum anderen von den Patienten oder Angehörigen anhand der Goal Attainment Scale (GAS) Ziele in den Bereichen Kognition, Funktion, Verhalten, Freizeit und soziale Aktivitäten definiert und später die Veränderung im Vergleich zum Studienbeginn beurteilt. Damit stellt die GAS ein individualisiertes Ergebnismaß dar, das – je nachdem, welche Ziele definiert werden –,

unterschiedliche Lebensbereiche abdeckt.^f In der Studie setzten sowohl die Patienten/Angehörigen als auch die Behandler die meisten Ziele im Bereich Kognition und Funktion (67 % bzw. 60 %) und die wenigsten in den Bereichen Freizeit und soziale Aktivitäten (14 % bzw. 19 %). In der ITT-Analyse ergab sich nach vier Monaten für die Galantamin-Gruppe sowohl nach der Beurteilung der Patienten/Angehörigen als auch der der Behandler ein erhöhter GAS-Wert, was einer Verbesserung in den gesetzten Zielen im Vergleich zum Studienbeginn entspricht. Dabei lagen die GAS-Werte in der Galantamin-Gruppe für die von den Behandlern aufgestellten und beurteilten Ziele statistisch signifikant höher als in der Placebo-Gruppe (Cohen's d = 0,451; p = 0,02), nicht jedoch für die von den Patienten bzw. Angehörigen beurteilten Ziele (Cohen's d = 0,20; p = 0,27). Als mögliche Erklärung hierfür wird diskutiert, dass insbesondere hinsichtlich der Beurteilung der Zielerreichung durch die Patienten/Angehörigen sich auch die Placebo-Gruppe in diesem Zeitraum verbesserte. Auch bei Berücksichtigung der Unterschiede im Schweregrad der Demenz zu Studienbeginn zwischen den Gruppen ergab sich hinsichtlich der Patienten/Angehörigen-Ziele kein statistisch signifikanter Unterschied.

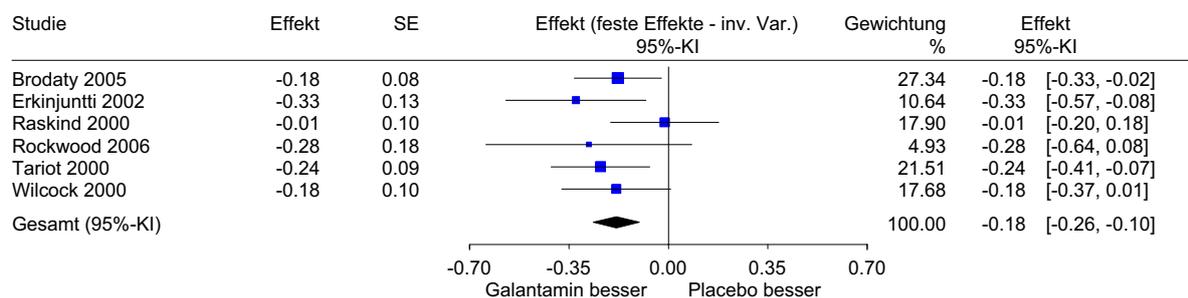
Zusammenfassend liefern die Studien überwiegend, jedoch nicht durchgehend, Hinweise darauf, dass Galantamin die Verschlechterung der Aktivitäten des täglichen Lebens im Vergleich zu Placebo verringern kann. Die Meta-Analyse aller Studien mit (16–)24 mg Galantamin als nicht retardierte Form (Abbildung 28) bestätigt die nur geringe Heterogenität in den Ergebnissen ($I^2 = 0,5 \%$) und zeigt einen kleinen, wenn auch statistisch signifikanten gepoolten Effekt in der Größenordnung von einem Fünftel Standardabweichung (Cohen's d = -0,18; 95 %-KI: -0,26 – -0,10). Dies entspricht auf der DAD-Skala etwa 3 Scorepunkten.

^f Da sich die Goal Attainment Scale von den anderen, psychometrischen Verfahren zur Erfassung der Aktivitäten des täglichen Lebens deutlich unterscheidet, werden hier für die Studie Rockwood 2006 sowohl die Ergebnisse aus dem DAD als auch aus der GAS dargestellt.

Tabelle 27: Galantamin: Ergebnisse zu alltagspraktischen Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Brodaty 2005 (6 Monate)	GAL-PRC 16- 24 mg	ADCS -ADL	(320) ^(b)	0,0 (8,6) ^(c)	2,7 (1,1; 3,9)	↗	< 0,001
	GAL 16-24 mg		(327) ^(b)	-1,0 (9,0) ^(c,d)	1,7 (0,2; 3,0)	↗	0,018
	Placebo		(324) ^(b)	-2,7 (10,1) ^(c)			
Erkinjuntti 2002 ^(e,f) (6 Monate)	GAL 24 mg	DAD	172	-1,0 (15,8) ^(c)	5,5 (n.g.) ^(g)	↗	< 0,01
	Placebo		93	-6,0 (14,5) ^(c)			
Raskind 2000 ^(f) (6 Monate)	GAL 24 mg	DAD	(212) ^(b)	-2,7 (14,9) ^(c)	0,2 (n.g.)	↗	„n.s.“
	Placebo		(213) ^(b)	-2,9 (15,8) ^(c)			
Rockwood 2006 (4 Monate)	GAL 16-24 mg	DAD	(64) ^(b)	n.g.	Cohen's d=0,28 (n.g.)	↗	0,13 ^(h)
	Placebo		(66) ^(b)	n.g.			
Tariot 2000 (5 Monate)	GAL 8 mg	ADCS -ADL	129	-3,2 (9,1) ^(c)	0,6 (n.g.)	↗	„n.s.“
	GAL 16 mg		255	-0,7 (8,0) ^(c)	3,1 (n.g.)	↗	< 0,001
	GAL 24 mg		253	-1,5 (9,5) ^(c)	2,3 (n.g.)	↗	< 0,01
	Placebo		262	-3,8 (9,7) ^(c)			
Wilcock 2000 (6 Monate)	GAL 24 mg	DAD	212	-3,2 (14,9) ^(c)	2,8 (-0,6; 6,1)	↗	0,1
	Placebo		210	-6,0 (15,7) ^(c)			
<p>Erläuterung: GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Anzahl ausgewerteter Patienten, sofern nicht anders angegeben b: Anzahl randomisierter Patienten, Anzahl ausgewerteter Patienten in Publikation nicht genannt c: eigene Berechnung d: Dateninkonsistenz bezüglich SD bzw. SE, in Tabelle 2 angegebener Wert von SE = 0,05 deckt sich nicht mit Wert in der Abbildung und müsste mit hoher Wahrscheinlichkeit SE = 0,5 lauten e: Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung f: Angaben direkt von der Herstellerfirma Janssen-Cilag g: Die Diskrepanz zwischen der berichteten Differenz und den angegebenen gruppeninternen Schätzern ist vermutlich auf Adjustierungen zurückzuführen. h: In einer zusätzlichen Analyse, bei der die Unterschiede im Demenzschweregrad sowie die Abbrecher berücksichtigt wurden, lag der p-Wert bei p = 0,051.</p>							

Galantamin - Aktivitäten des täglichen Lebens
 Endpunkt: ADCS-ADL, DAD - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=5.03$, $df=5$ ($p=0.413$), $I^2=0.5\%$
 Gesamteffekt: Z Score=-4.46 ($p=0.000$)

Bemerkung:

Rockwood 2006: SE aus p-Wert = 0,13 berechnet

Vorzeichen der Scores geändert, d. h. negative Werte entsprechen positiven Effekten

Darstellung des Gruppenunterschieds inklusive SE, da gruppeninterne Daten für Rockwood nicht publiziert

Abbildung 28: Galantamin: Meta-Analyse zu Aktivitäten des täglichen Lebens

5.3.2.2 Begleitende Psychopathologie

In drei von sechs Studien zu Galantamin wurde die begleitende Psychopathologie mittels des Neuropsychiatric Inventory (NPI) als Ergebnisparameter erhoben (Brodaty 2005, Erkinjuntti 2002, Tariot 2000). Die Durchschnittswerte der Baselinedaten zum NPI in den drei Studien lagen in dem engen Bereich von 10 bis 13 Punkten bei hoher Variabilität zwischen den einzelnen Patienten (Tabelle 11). Bei Brodaty 2005 zeigte sich kein statistisch signifikanter Effekt für Galantamin, ebenso wenig in der Subgruppe mit AD-CVD-Patienten der Studie von Erkinjuntti 2002. Tariot 2000 berichtete hingegen über statistisch signifikante Effekte zugunsten der Galantamin 16- und 24 mg-Gruppe, nicht aber in der 8 mg-Gruppe (Tabelle 28). Daten zu Subskalen des NPI lagen für die Studien Brodaty 2005 und Tariot 2000 sowie für die Subgruppe der AD-CVD-Patienten bei Erkinjuntti 2002 nicht vor.

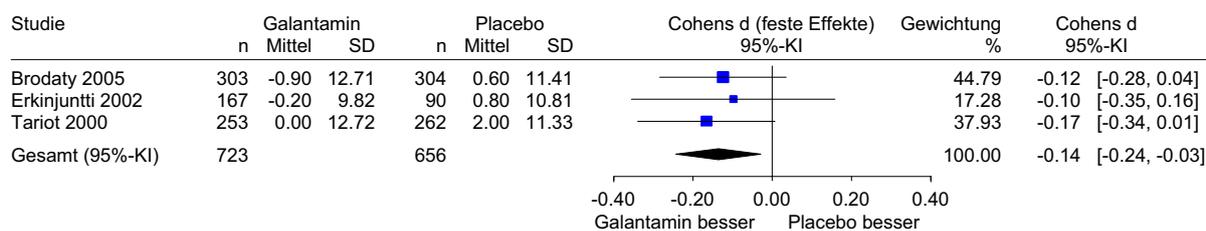
Zusammenfassend ergeben sich damit Hinweise auf einen Effekt von Galantamin bezüglich der Beeinflussung der neuropsychiatrischen Symptomatik. Die Meta-Analyse (

Abbildung 29) bestätigt dies und zeigt einen nur geringfügigen, wenn auch statistisch signifikanten, gepoolten Effekt in der Größenordnung von einem Siebtel Standardabweichung (Cohen's $d = -0,14$; 95 %-KI: $-0,24 - -0,03$). Dies entspricht auf die NPI-Skala umgerechnet etwa 1 bis 2 Scorepunkten.

Tabelle 28: Galantamin: Ergebnisse zur begleitenden Psychopathologie

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Brodaty 2005 (6 Monate)	GAL-PRC 16- 24 mg	NPI	299 ^(b)	0,6 ^(c) (11,4 ^(d))	0 ^(d) (-1,9; 1,8)	↔	0,941
	GAL 16-24 mg		303 ^(b)	-0,9 (12,7 ^(d))	-1,5 ^(d) (-3,4; 0,2)	↗	0,102
	Placebo		304 ^(b)	0,6 (11,4 ^(d))			
Erkinjuntti 2002 ^(b,e) (6 Monate)	GAL 24 mg	NPI	167	-0,2 (9,8 ^(d))	-1,0 (n.g.)	↗	„n.s.“
	Placebo		90	0,8 (10,8 ^(d))			
Tariot 2000 (5 Monate)	GAL 8 mg	NPI	129	2,3 (11,4) ^(f)	0,3 ^(d) (n.g.)	↘	„n.s.“
	GAL 16 mg		255	-0,1 (11,4) ^(f)	-2,1 ^(d) (n.g.)	↗	< 0,05
	GAL 24 mg		253	0,0 (11,1) ^(f)	-2,0 ^(d) (n.g.)	↗	< 0,05
	Placebo		262	2,0 (11,0) ^(f)			
<p>Erläuterung: GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in Auswertung b: Angaben direkt von der Herstellerfirma Janssen-Cilag c: In der Originalpublikation war hier ein Vorzeichen vertauscht. d: eigene Berechnung e: Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung f: Angaben aus Cummings et al. 2004 [73]</p>							

Galantamin - begleitende Psychopathologie
 Endpunkt: NPI - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=0.22$, $df=2$ ($p=0.895$), $I^2=0\%$
 Gesamteffekt: Z Score=-2.49 ($p=0.013$)

Abbildung 29: Galantamin: Meta-Analyse zur begleitenden Psychopathologie

5.3.2.3 Kognitive Leistungsfähigkeit

In allen sechs Studien wurde die Wirksamkeit von Galantamin bezüglich der kognitiven Funktionen mit der kognitiven Subskala der ADAS (ADAS-cog) gemessen (Brodaty 2005, Erkinjuntti 2002, Raskind 2000, Rockwood 2006, Tariot 2000, Wilcock 2000). In allen Studien ergaben sich für die Dosierungen 16 mg und 24 mg statistisch signifikante Vorteile für Galantamin gegenüber Placebo. Die 8 mg-Gruppe aus der Studie von Tariot 2000 dagegen zeigte keinen statistisch signifikanten Effekt (Tabelle 29).

Insgesamt zeigen die Studien konsistent einen Vorteil von Galantamin im Hinblick auf eine günstige Beeinflussung der kognitiven Funktionsfähigkeit, gemessen mit dem ADAS-cog. Zwischen den Studien zeigte sich keine Heterogenität der Ergebnisse. In der Meta-Analyse (Abbildung 30) über sechs Studien, in die jeweils die mit 24 mg bzw. 16–24 mg behandelten Patienten gingen, beträgt der durchschnittliche Unterschied in der Verbesserung des ADAS-cog etwa eine halbe Standardabweichung (Cohen's $d = -0,51$; 95 %-KI: $-0,59 - -0,43$) und damit etwa 3 Punkte auf der ADAS-cog-Skala.

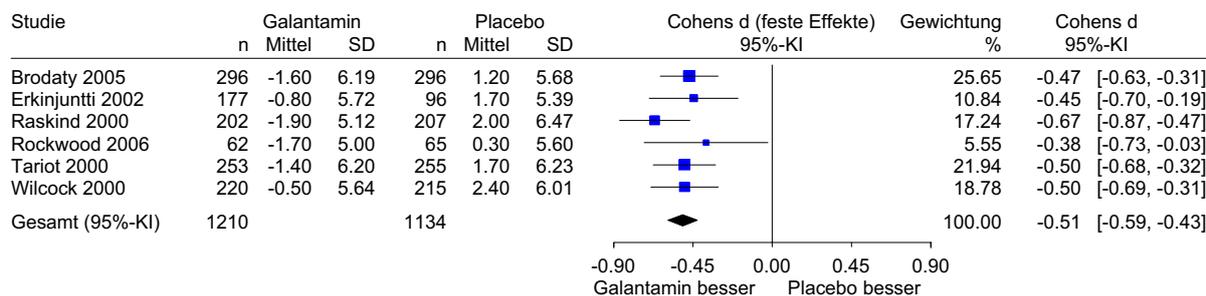
Tabelle 29: Galantamin: Ergebnisse zu kognitiven Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied zu Placebo		p-Wert
				Differenz (95 % KI)	Richtung des Effekts	
Brodaty 2005 (6 Monate)	GAL-PRC 16- 24 mg	ADAS-cog	291	-1,3 (5,3 ^(b))	-2,5 ^(b) (-3,3; -1,5)	↗ < 0,001
	GAL 16-24 mg		296	-1,6 (6,2 ^(b))	-2,8 ^(b) (-3,7; -1,9)	↗ < 0,01
	Placebo		296	1,2 (5,7 ^(b))		
Erkinjuntti 2002 ^(c,d) (6 Monate)	GAL 24 mg	ADAS-cog	177	-0,8 (5,7 ^(b))	-2,5 (n.g.)	↗ < 0,001
	Placebo		96	1,7 (5,4 ^(b))		
Raskind 2000 ^(d) (6 Monate)	GAL 24 mg	ADAS-cog	202	-1,9 ^(e) (5,1 ^(b))	-3,8 (n.g.)	↗ < 0,001
	Placebo		207	2,0 (6,5 ^(b))		
Rockwood 2006 (4 Monate)	GAL 16-24 mg	ADAS-cog	62	-1,7 (5,0) ^(f)	-2,0 ^(b) (n.g.)	↗ 0,04 ^(g)
	Placebo		65	0,3 (5,6) ^(f)		
Tariot 2000 (5 Monate)	GAL 8 mg	ADAS-cog	126	0,4 (5,8 ^(b))	-1,3 ^(b) (n.g.)	↗ „n.s.“
	GAL 16 mg		253	-1,4 (5,6 ^(b))	-3,1 ^(b) (n.g.)	↗ < 0,001
	GAL 24 mg		253	-1,4 (6,2 ^(b))	-3,1 ^(b) (n.g.)	↗ < 0,001
	Placebo		255	1,7 (6,2 ^(b))		
Wilcock 2000 (6 Monate)	GAL 24 mg	ADAS-cog	220	-0,5 (5,6 ^(b))	-2,9 ^(h) (-4,1; -1,6)	↗ < 0,001
	Placebo		215	2,4 (6,0 ^(b))		

Erläuterungen:
GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung
Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.

a: Zahl der Patienten in Auswertung
b: eigene Berechnung
c: Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung
d: Angaben direkt von der Herstellerfirma Janssen-Cilag
e: In der Publikation war (offenbar irrtümlicherweise) +1,9 angegeben, in den von der Herstellerfirma Janssen-Cilag bereitgestellten Daten jedoch -1,9.
f: Werte aus Abbildung geschätzt
g: Ergebnis auch bei Berücksichtigung der Gruppenunterschiede im Demenzschweregrad sowie der Abbrecher statistisch signifikant
h: Vorzeichen sind in Originalpublikation vertauscht

Galantamin - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog/11 - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=3.45$, $df=5$ ($p=0.630$), $I^2=0\%$
 Gesamteffekt: Z Score=-12.04 ($p=0.000$)

Bemerkung:

Rockwood 2006: Standardabweichungen aus Abbildung geschätzt

Abbildung 30: Galantamin: Meta-Analyse zu kognitiven Fähigkeiten

5.3.2.4 Krankheitsbezogene Lebensqualität

Zu dem Zielkriterium „Krankheitsbezogene Lebensqualität“ wurden in den Studien keine Daten berichtet.

5.3.2.5 Notwendigkeit der vollstationären Pflege (Institutionalisierung)

Zu dem Zielkriterium „Notwendigkeit der vollstationären Pflege (Institutionalisierung)“ wurden aus den kontrollierten Phasen der eingeschlossenen Studien keine Daten berichtet.

5.3.2.6 Mortalität

In den eingeschlossenen Studien wurde über nur wenige Todesfälle innerhalb des Berichtszeitraums berichtet. Dabei zeigte sich keine auffällige Häufung in einer der Behandlungsgruppen (Tabelle 30).

Insgesamt geben die Daten keinen Hinweis auf einen positiven oder schädlichen Effekt von Galantamin im Hinblick auf die Mortalität.

5.3.2.7 Unerwünschte Ereignisse

Bis auf Tariot 2000 wurden in allen Studien unter Galantamin deutlich höhere Abbruchraten wegen unerwünschter Ereignisse insbesondere in den Studien mit schnelleren Aufdosierungen (24 mg bzw. 16 mg innerhalb des ersten Monats) (Erkinjuntti 2002, Raskind 2000, Wilcock 2000) beobachtet. Angaben zu schwerwiegenden unerwünschten Ereignissen wurden nur bei Rockwood 2006 und Tariot 2000 in den Publikationen berichtet; für die übrigen Studien (bis

auf die relevante Subgruppe bei Erkinjuntti 2002) stellte die Herstellerfirma Janssen-Cilag entsprechende Angaben zur Verfügung. In der Gesamtschau mittels Meta-Analyse zeigten sich bezüglich der schwerwiegenden unerwünschten Ereignisse keine Unterschiede zwischen den Verum- und Placebo-Gruppen (Abbildung 32).

Typische cholinerge Nebenwirkungen traten in der Verum-Gruppe deutlich häufiger auf, besonders während der Aufdosierungsphase. In den Studien, in denen pro Woche um 8 mg aufdosiert wurde (Raskind 2000, Wilcock 2000), litt ca. ein Drittel der Patienten in der Verum-Gruppe unter Übelkeit und ein Fünftel unter Erbrechen (Tabelle 31). Insgesamt trat Übelkeit in allen Studien – bis auf die 8 mg-Gruppe bei Tariot 2000 – ca. 3- bis 4-mal häufiger in der Verum-Gruppe auf; die absoluten Differenzen lagen zwischen 8 % und 25 %, Erbrechen war (absolut) zwischen 5 % und 16 % häufiger (in der 8 mg-Gruppe bei Tariot 2000: 3 %). Eine Sensitivitätsanalyse hierzu hinsichtlich Studien mit Aufdosierung gemäß Fachinformation [107] und schnellerer Aufdosierung ergab weder konsistente noch deutliche Vorteile für die langsamere Aufdosierung. Unter Schwindel und Appetitlosigkeit litten etwa doppelt so häufig die Patienten in den Verum-Gruppen, in der 24 mg-Gruppe bei Raskind 2000 war Gewichtsverlust besonders ausgeprägt.

Zusammenfassend belegen die Studien ein gehäuftes Auftreten von unerwünschten Ereignissen durch die Einnahme von Galantamin gegenüber Placebo. Die im Folgenden dargestellten Meta-Analysen (Abbildung 31 bis Abbildung 37) zeigen keine auffällige Heterogenität zwischen den Ergebnissen.

Tabelle 30: Galantamin: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Studie (Dauer)		Aufdosierung innerhalb des ersten Monats	N ^(a)	Studien- abbruch N (%)	Todes- fälle N	Schwerwiegende unerwünschte Ereignisse N (%)	Abbruch wegen unerwünschter Ereignisse N (%)	Jegliche unerwünschte Ereignisse N (%)
Brodaty 2005 (6 Monate)	GAL-PRC 16-24 mg	8 mg	(319) ^(b)	68 (21)	3	35 ^(c) (11) ^(d)	28 (9)	253 (79)
	GAL 16-24 mg	8 mg	(326) ^(b)	75 (23)	1	39 ^(c) (12) ^(d)	24 (7)	235 (72)
	Placebo		(320) ^(b)	54 (17)	1	35 ^(c) (11) ^(d)	15 (5)	224 (70)
Erkinjuntti 2002 (6 Monate)	GAL 24 mg	16 mg	188 ^(d)	34 (18) ^(d)	2 ^(d)	k.A.	25 (13) ^(d)	330 ^(d) (83) ^(e)
	Placebo		97 ^(d)	10 (10) ^(d)	1 ^(d)	k.A.	5 (5) ^(d)	133 ^(d) (68) ^(e)
Raskind 2000 (6 Monate)	GAL 24 mg	24 mg	212	68 (32)	1	29 (14) ^(d)	49 (23)	195 (92)
	Placebo		213	41 (19)	1	27 (13) ^(d)	16 (8)	168 (79)
Rockwood 2006 (4 Monate)	GAL 16-24 mg	8 mg	64	11 (17)	0	5 (8) ^(f,g)	5 (8)	54 (84)
	Placebo		66	10 (15)	1	10 (15) ^(f,g)	2 (3)	41 (62)
Tariot 2000 (5 Monate)	GAL 8 mg	8 mg	140	32 (23)	1	14 (10)	9 (6)	106 (76)
	GAL 16 mg	8 mg	279	60 (22)	3	28 (10)	19 (7)	206 (74)
	GAL 24 mg	8 mg	273	61 (22)	3	35 (13)	27 (10)	219 (80)
	Placebo		286	46 (16)	4	31 (11)	20 (7)	206 (72)
Wilcock 2000 (6 Monate)	GAL 24 mg	24 mg	220	44 (20)	2 ^(d)	31 (14) ^(d,h)	31 (14)	182 (83)
	Placebo		215	29 (13)	2 ^(d)	27 (13) ^(d,h)	19 (9)	165 (77)

(Fortsetzung)

Tabelle 30 (Fortsetzung): Galantamin: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Erläuterung:

GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, k.A. = keine Angaben, N = Anzahl

a: Anzahl randomisierter Patienten

b: Anzahl der randomisierten und behandelten Patienten

c: aus %-Angabe berechnet

d: Angaben der Herstellerfirma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens (vgl. SN Janssen-Cilag)

e: Angaben beziehen sich auf die Gesamtpopulation der betrachteten Patienten (Galantamin: n=396, Placebo: n=196), einschließlich der mit ausschließlicher zerebrovaskulärer Demenz, da für die Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung keine separaten Angaben zu den unerwünschten Ereignissen vorliegen.

f: Prozentzahlen aus N berechnet

g: Zuordnung der Zahlen zu den beiden Gruppen geht aus der Publikation nicht ganz eindeutig hervor und ist hier so dargestellt, wie sie am wahrscheinlichsten zutrifft.

h: Todesfälle wurden in der Stellungnahme gesondert berichtet und hier inkludiert.

Tabelle 31: Galantamin: unerwünschte Ereignisse

Studie (Dauer)	Auf- dosierung innerhalb des ersten Monats	N ^(a)	Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein	
														Anteil (%) mit UE
Brodaty 2005 (6 Monate)	GAL PRC 16- 24 mg	8 mg	(319) ^(b)	17	7	5	10	6	4	7	8	2 ^(d)	4 ^(d)	k.A.
	GAL 16-24 mg	8 mg	(326) ^(b)	14	9	7	7	7	5	6	6	3 ^(d)	4 ^(d)	k.A.
	Placebo		(320) ^(b)	5	2	7	4	3	1	7	6	2 ^(d)	1 ^(d)	k.A.
Erkinjuntti 2002 ^(c, d) (6 Monate)	GAL 24 mg	16 mg	396 ^(c)	24 ^(d)	13 ^(d)	8 ^(d)	9 ^(d)	5 ^(d)	4 ^(d)	3 ^(d)	6 ^(d)	5 ^(d)	4 ^(d)	k.A.
	Placebo		196 ^(c)	7 ^(d)	6 ^(d)	5 ^(d)	5 ^(d)	2 ^(d)	2 ^(d)	4 ^(d)	6 ^(d)	6 ^(d)	5 ^(d)	k.A.
Raskind 2000 (6 Monate)	GAL 24 mg	24 mg	212	37	21	12	14	14	12	11 ^(d)	9 ^(d)	7	6 ^(d)	k.A.
	Placebo		213	13	8	10	11	6	5	16 ^(d)	8 ^(d)	4	4 ^(d)	k.A.
Rockwood 2006 (4 Monate)	GAL 16-24 mg	8 mg	64	23	17	k.A.	k.A.	11	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
	Placebo		66	6	3	k.A.	k.A.	2	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.

(Fortsetzung)

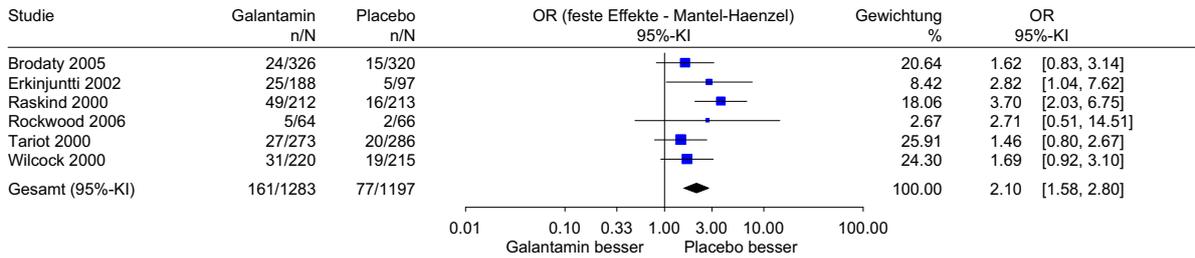
Tabelle 31 (Fortsetzung): Galantamin: unerwünschte Ereignisse

Studie (Dauer)	Auf- dosierung innerhalb des ersten Monats	N ^(a)	Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein	
														Anteil (%) mit UE
Tariot 2000 (5 Monate)	GAL 8 mg	8 mg	140	6	4	5	5 ^(c)	6	k.A.	15	4 ^(c)	k.A.	k.A.	k.A.
	GAL 16 mg	8 mg	279	13	6	12	5 ^(d)	7	5 ^(d)	10	7 ^(d)	4 ^(d)	4 ^(d)	k.A.
	GAL 24 mg	8 mg	273	17	10	6	7 ^(d)	9	5 ^(d)	8	5 ^(d)	3 ^(d)	5 ^(d)	k.A.
	Placebo		286	5	1	6	4 ^(d)	3	1 ^(d)	9	5 ^(d)	4 ^(d)	2 ^(d)	k.A.
Wilcock 2000 (6 Monate)	GAL 24 mg	24 mg	220	37	20	7	11	10	8	4 ^(d)	10	8 ^(d)	6 ^(d)	k.A.
	Placebo		215	12	4	7	5	0	1	8 ^(d)	3	5 ^(d)	5 ^(d)	k.A.

Erläuterung:
GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, k.A. = keine Angaben, N = Anzahl, UE = unerwünschte Ereignisse

a: Anzahl der randomisierten Patienten, sofern nicht anders angegeben
b: Anzahl der randomisierten und behandelten Patienten
c: Angaben beziehen sich auf die Gesamtpopulation der betrachteten Patienten, einschließlich der mit ausschließlicher zerebrovaskulärer Demenz, da für die Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung keine separaten Angaben zu den Nebenwirkungen vorliegen.
d: Angaben direkt von der Herstellerfirma Janssen-Cilag
e: Angaben der Herstellerfirma Janssen-Cilag im Rahmen des Stellungnahmeverfahrens (vgl. SN Janssen-Cilag)

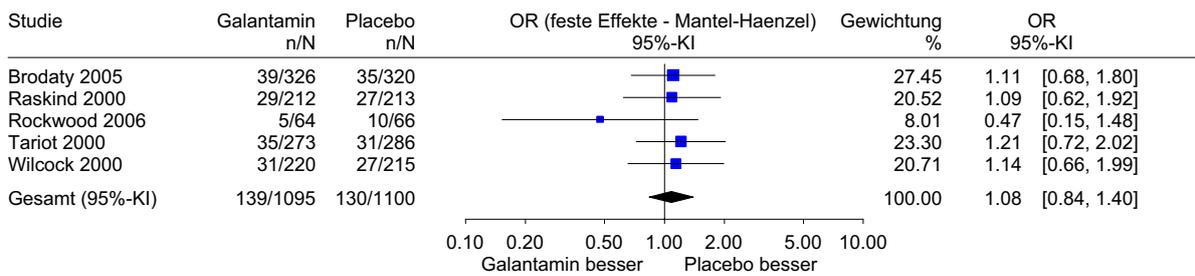
Galantamin - Abbruch wegen UEs (schnelle Aufdosierung)
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



Heterogenität: $Q=6.32$, $df=5$ ($p=0.276$), $I^2=20.9\%$
 Gesamteffekt: Z Score=5.08 ($p=0.000$)

Abbildung 31: Galantamin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse

Galantamin - Schwerwiegende UEs
 Endpunkt: SUE erlitten (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit SUE



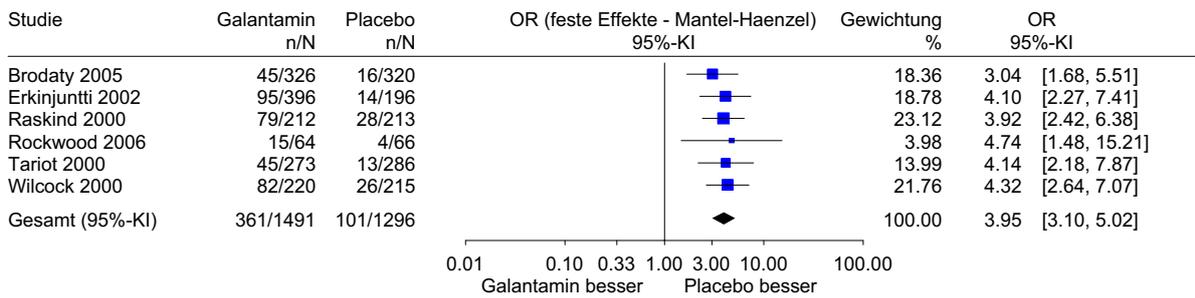
Heterogenität: $Q=2.25$, $df=4$ ($p=0.689$), $I^2=0\%$
 Gesamteffekt: Z Score=0.62 ($p=0.535$)

Abbildung 32: Galantamin: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen

Galantamin - UEs

Endpunkt: Übelkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1$, $df=5$ ($p=0.962$), $I^2=0\%$

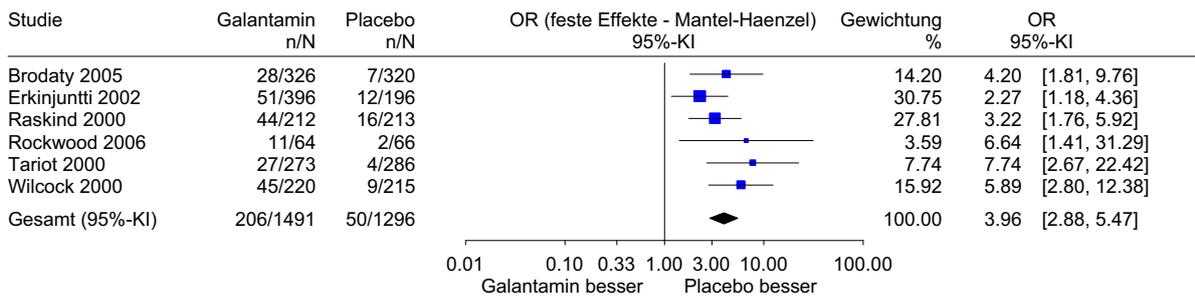
Gesamteffekt: Z Score=11.15 ($p=0.000$)

Abbildung 33: Galantamin: Meta-Analyse zum Endpunkt „Übelkeit“

Galantamin - UEs

Endpunkt: Erbrechen

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=6.3$, $df=5$ ($p=0.278$), $I^2=20.6\%$

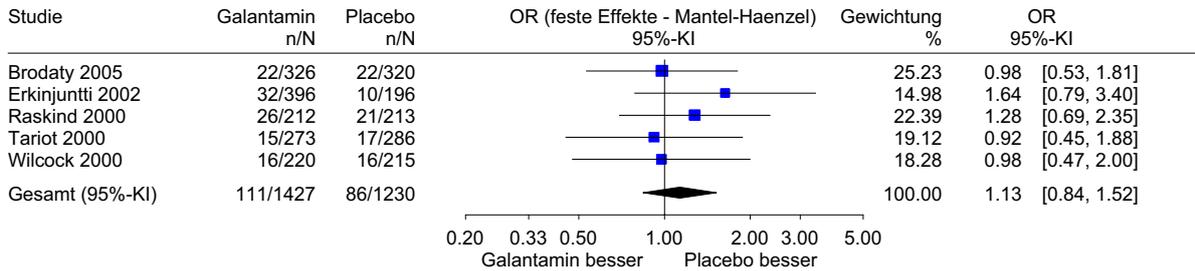
Gesamteffekt: Z Score=8.41 ($p=0.000$)

Abbildung 34: Galantamin: Meta-Analyse zum Endpunkt „Erbrechen“

Galantamin - UEs

Endpunkt: Diarrhoe

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.82$, $df=4$ ($p=0.768$), $I^2=0\%$

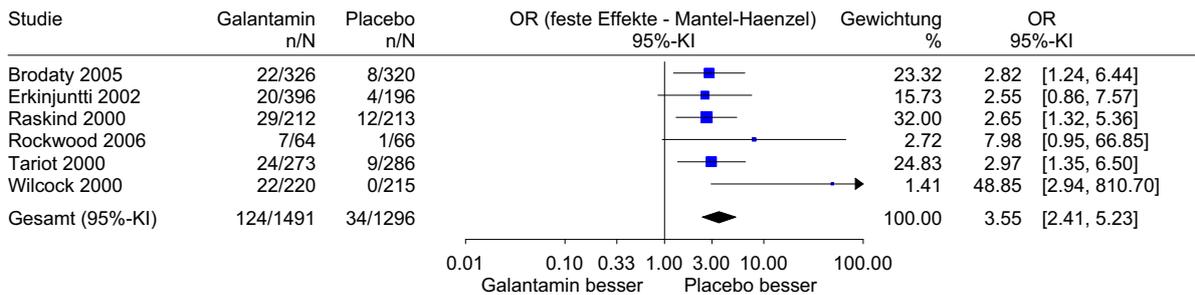
Gesamteffekt: Z Score=0.82 ($p=0.411$)

Abbildung 35: Galantamin: Meta-Analyse zum Endpunkt „Diarrhoe“

Galantamin - UEs

Endpunkt: Appetitlosigkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf

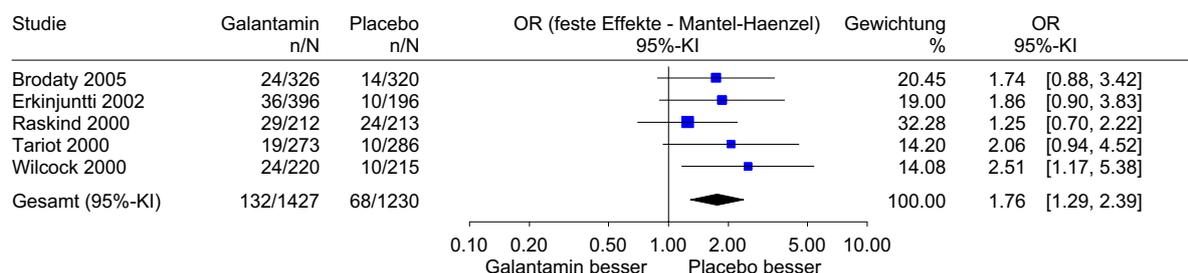


Heterogenität: $Q=5.41$, $df=5$ ($p=0.367$), $I^2=7.7\%$

Gesamteffekt: Z Score=6.4 ($p=0.000$)

Abbildung 36: Galantamin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“

Galantamin - UEs
 Endpunkt: Schwindel
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=2.37$, $df=4$ ($p=0.667$), $I^2=0\%$
 Gesamteffekt: Z Score= 3.59 ($p=0.000$)

Abbildung 37: Galantamin: Meta-Analyse zum Endpunkt „Schwindel“

5.3.2.8 Lebensqualität der (betreuenden) Angehörigen

In der Studie Rockwood 2006 wurde die Belastung der betreuenden Angehörigen anhand der Caregiving Burden Scale (CBS) untersucht. Hier zeigte sich nach vier Monaten kein statistisch signifikanter Unterschied zwischen der Galantamin- und der Placebo-Gruppe (standardisierte Mittelwertsdifferenz = -0,17; $p = 0,38$).

In der Publikation von Cummings 2004 [73] wurde über die Ergebnisse aus der Studie von Tariot 2000 zur Neuropsychiatric Inventory Caregiver Distress Scale (NPI-D) berichtet, die den emotionalen Stress der Angehörigen im Zusammenhang mit den psychopathologischen Symptomen der Alzheimer Erkrankung erfasst. Darüber hinaus lagen zu dieser Skala für die Studien Brodaty 2005, Erkinjuntti 2002 sowie für die beiden mit 16 mg bzw. 24 mg Galantamin behandelten Gruppen aus Tariot 2000 Daten der Herstellerfirma Janssen-Cilag vor. In keiner der Studien zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Die Zusammenfassung mittels einer Meta-Analyse (Abbildung 38) ergab mit einem sehr geringen, wenn auch statistisch signifikanten, gepoolten Effekt von etwa einem Zehntel Standardabweichung (Cohen's $d = -0,11$; 95 %-KI: -0,21 – -0,00) einen Hinweis auf eine günstige Beeinflussung der Lebensqualität der betreuenden Angehörigen.

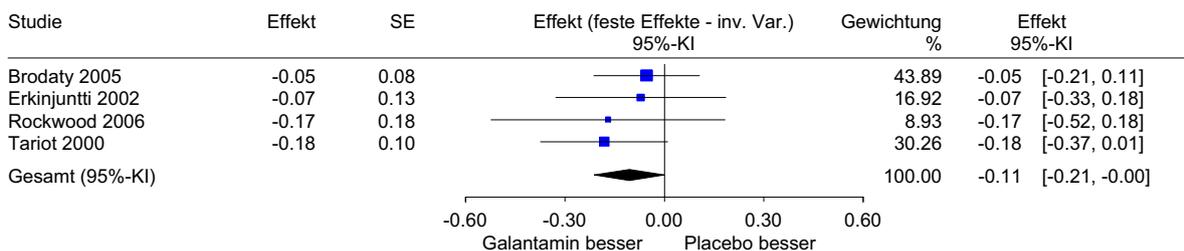
Tabelle 32: Galantamin: Ergebnisse zur Lebensqualität der (betreuenden) Angehörigen

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Brodaty 2005 ^(b) (6 Monate)	GAL-PRC 16-24 mg	NPI-D	299	0,1 (5,7 ^(c))	0,0 ^(d) (n.g.)	↔	„n.s.“
	GAL 16-24 mg		303	-0,4 (6,1 ^(c))	-0,3 ^(d) (n.g.)	↗	„n.s.“
	Placebo		304	0,3 (5,1 ^(c))			
Erkinjuntti 2002 ^(b,e) (6 Monate)	GAL 24 mg	NPI-D	167	-0,1 (5,4 ^(c))	-0,4 (n.g.)	↗	„n.s.“
	Placebo		90	0,3 (5,9 ^(c))			
Rockwood 2006 (4 Monate)	GAL 16-24 mg	CBS	(64) ^(f)	n.g.	Cohen's d = -0,17	↗	0,38
	Placebo		(66) ^(f)	n.g.			
Tariot 2000 ^(b) (5 Monate)	GAL 8 mg	NPI-D	116 ^(g)	1,2 (5,9) ^(g)	0,3 ^(c) (n.g.)	↘	„n.s.“ ^(g)
	GAL 16 mg		215	-0,1 (5,9 ^(c))	-1,1 (n.g.)	↗	„n.s.“
	GAL 24 mg		206	-0,2 (7,2 ^(c))	-1,2 (n.g.)	↗	„n.s.“
	Placebo		214	0,9 (5,9 ^(c))			

Erläuterung:
 GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung
 Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.

a: Zahl der Patienten in Auswertung, sofern nicht anders angegeben
 b: Angaben direkt von der Herstellerfirma Janssen-Cilag
 c: eigene Berechnung
 d: Diskrepanzen zwischen den berichteten Differenzen und den angegebenen gruppeninternen Schätzern sind vermutlich auf Adjustierungen zurückzuführen.
 e: Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung
 f: Anzahl randomisierter Patienten, Anzahl ausgewerteter Patienten in Publikation nicht genannt.
 g: Angaben aus Cummings et al. 2004 [73]

Galantamin - Lebensqualität der (betreuenden) Angehörigen
 Endpunkt: NPI-D, CBS - Differenz zu Baseline
 Distanzmaß: Standardisierte Differenz der Mittelwerte



Heterogenität: $Q=1.21$, $df=3$ ($p=0.750$), $I^2=0\%$
 Gesamteffekt: Z Score=-1.97 ($p=0.049$)

Bemerkung:

Rockwood 2006: SE aus p-Wert = 0,38 berechnet

Abbildung 38: Galantamin: Meta-Analyse zur Lebensqualität der (betreuenden) Angehörigen

5.3.2.9 Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en

Eine gepoolte Analyse zum zeitlichen Pflege- und Betreuungsaufwand auf der Basis der Daten der Studien von Wilcock 2000 und Raskind 2000 wurde von Sano 2003 [98] publiziert. Die Originaldaten wurden in den Ursprungspublikationen nicht berichtet. Die aufgewendete Pflegezeit wurde mittels des Allocation of Caregiver Time Survey (ACTS) erfasst; darüber hinaus wurde erfragt, wie viel Zeit die Patienten täglich ohne Betreuung verbringen konnten. Insgesamt wurden in der Analyse Daten von 411 Verum- und 414 Placebo-Patienten ausgewertet (97 % bzw. 98 % der randomisierten Patienten), wobei der Patientenfluss unklar bleibt. In der berichteten Population erscheinen die Verum- und die Placebo-Gruppe vergleichbar. Es wird berichtet, dass nach 6 Monaten die Betreuer der Patienten, die mit 24 mg Galantamin behandelt wurden, täglich im Mittel 32 Minuten weniger für die Unterstützung bei den alltagspraktischen Tätigkeiten aufwendeten als die Betreuer von Patienten in der Placebo-Gruppe ($p = 0,011$). Bei den Patienten mit mittelschwerer Alzheimer Demenz ($MMSE \leq 18$) lag die Differenz zwischen Verum- und Placebo-Gruppe mit 53 Minuten täglich noch deutlich höher ($p = 0,021$). Das Ergebnis für das Kriterium „Zeit ohne Betreuung“ war ähnlich. Verschiedenartige Auswertungen (kategoriale versus kontinuierliche Betrachtung der Zielkriterien, Adjustierung für Pflegeaufwand bei Baseline) im Sinne von Sensitivitätsanalysen ergaben konsistente Ergebnisse.

Auch in der Studie Rockwood 2006 wurde der ACTS eingesetzt, allerdings werden in der Publikation keine Ergebnisse hierzu berichtet.

Zusammenfassend liefern die vorliegenden Daten lediglich einen Hinweis auf eine günstige Beeinflussung des Betreuungsaufwands durch Galantamin.

5.3.2.10 Ergänzende Information: klinisches Krankheitsstadium

Alle Studien untersuchten Veränderungen im klinischen Gesamteindruck unter Verwendung des CIBIC-plus. In fünf Studien stellte er einen primären Endpunkt dar, und er wurde hier jeweils kategoriell unter Berücksichtigung aller sieben Stufen ausgewertet. In Tabelle 33 ist der Anteil der Responder dargestellt, das heißt der Anteil von Patienten mit einem Wert ≤ 4 (keine Verschlechterung oder Besserung). Der Responderanteil betrug in den Verum-Gruppen (mit Ausnahme der 8 mg-Gruppe) nach 5 bis 6 Monaten zwischen 62 % und 75 %. Der Anteil der Patienten in der Placebo-Gruppe, die als unverändert oder verbessert eingestuft wurden, lag zwischen 49 % und 58 %. In vier dieser fünf Studien war die Verteilung der CIBIC-plus-Werte zwischen den Galantamin- und den Placebo-Gruppen statistisch signifikant unterschiedlich.

Bei Rockwood 2006 stellte der CIBIC-plus einen sekundären Endpunkt dar und wurde lediglich im Rahmen eines Mittelwertsvergleichs ausgewertet. Nach vier Monaten war der CIBIC-plus in der LOCF-Auswertung in der Galantamin-Gruppe ebenfalls statistisch signifikant geringer als in der Placebo-Gruppe.

Insgesamt zeigen die Studien, dass Galantamin den klinischen Gesamteindruck im Vergleich zu Placebo verbessert.

Tabelle 33: Galantamin: Ergebnisse zu CIBIC-plus ≤ 4 (Stabilität oder Verbesserung)

Studie (Dauer)	End- punkt	N ^(a)	Anteil (%) mit Score ≤ 4	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Brodaty 2005 (6 Monate)	GAL-PRC 16- 24 mg	CIBIC -plus	291	62	+ 5% (n.g.)	↗	„n.s.“ ^(b)
	GAL 16-24 mg		302	63	+ 6% (n.g.)	↗	„n.s.“ ^(b)
	Placebo		301	57			
Erkinjuntti 2002 ^(c) (6 Monate)	GAL 24 mg	CIBIC -plus	172	75	+ 21% (n.g.)	↗	0,001 ^(b)
	Placebo		92	54			
Raskind 2000 (6 Monate)	GAL 24 mg	CIBIC -plus	186	73	+ 16% (n.g.)	↗	< 0,01 ^(b)
	Placebo		196	57			
Tariot 2000 (5 Monate)	GAL 8 mg	CIBIC -plus	128	53	+ 4% (n.g.)	↗	„n.s.“ ^(b)
	GAL 16 mg		256	66	+ 17% (n.g.)	↗	< 0,001 ^(b)
	GAL 24 mg		253	64	+ 15% (n.g.)	↗	< 0,001 ^(b)
	Placebo		261	49			
Wilcock 2000 (6 Monate)	GAL 24 mg	CIBIC -plus	206	62	+ 12% (n.g.)	↗	< 0,05 ^(b)
	Placebo		203	50			
Mittlere Differenz zu Baseline (SD)							
Rockwood 2006 (4 Monate)	GAL 16-24 mg	CIBIC -plus	61	n.g.	-1,0 ^(d) (n.g.)	↗	0,03
	Placebo		65	n.g.			
<p>Erläuterung: GAL = Galantamin, GAL-PRC = Galantamin Prolonged Release, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Anzahl Patienten in Auswertung b: p-Werte beruhen auf einem Vergleich der Verteilung der Patienten über 7 Kategorien (van Elteren-Test) c: Subgruppe der Patienten mit Alzheimer Demenz plus zerebrovaskulärer Erkrankung d: aus Abbildung geschätzt</p>							

5.3.3 Rivastigmin

5.3.3.1 Aktivitäten des täglichen Lebens

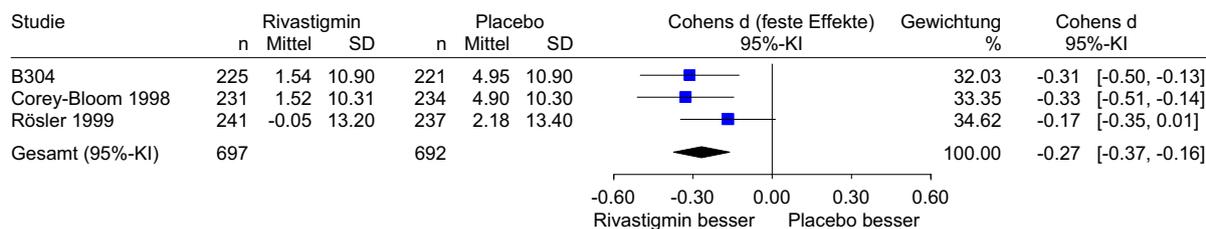
In drei Studien (B304, Corey-Bloom 1998 und Rösler 1999) wurden die Aktivitäten des täglichen Lebens anhand der Progressive Deterioration Scale (PDS) gemessen. Die Werte beruhen auf Angaben der Angehörigen oder Pflegenden. In der Studie B304, die eine flexible Dosierung bis 12 mg zuließ, wurden sowohl bei zwei- als auch bei dreimaliger Verabreichung pro Tag statistisch signifikante Verbesserungen im Vergleich zu Placebo beobachtet. Dabei war der beobachtete Unterschied zur Placebo-Gabe in der Gruppe mit dreimal täglicher Rivastigmingabe etwas größer. Bei Corey-Bloom 1998 zeigten sich in der PDS statistisch signifikante Effekte für die Hochdosis-Gruppe (6 bis 12 mg), nicht jedoch für die Niedrigdosis-Gruppe. Bei Rösler 1999 ergab sich lediglich für die Hochdosis-Gruppe ein nicht statistisch signifikanter Trend zugunsten Rivastigmin ($p = 0,07$), in der Niedrigdosis-Gruppe (1–4 mg) zeigte sich ein umgekehrter Trend.

Zusammenfassend liefern die Studien den Hinweis für eine günstige Beeinflussung der Aktivitäten des täglichen Lebens durch Rivastigmin in hoher Dosierung bzw. bei dreimal täglicher Gabe. Eine Meta-Analyse (Abbildung 39) zeigt keine Heterogenität in den Effekten und eine kleine, statistisch signifikante, Effektstärke in der Größenordnung einer Viertel Standardabweichung (Cohen's $d = -0,27$; 95 %-KI: $-0,37 - -0,16$). Für den PDS entspricht dies etwa drei Scorepunkten. Die Berücksichtigung der Ergebnisse aus der Studie B351, die bislang nicht veröffentlicht wurde und für die die Firma Novartis den Studienbericht nicht zur Verfügung gestellt hat, im Rahmen einer Sensitivitätsanalyse, führt zu statistisch auffälliger Heterogenität und verringert den Effektschätzer etwas. Allerdings bleibt die statistische Signifikanz sowohl für ein Modell mit festen ($p < 0,001$) als auch mit zufälligen Effekten ($p < 0,001$) erhalten (Daten in Anhang G dargestellt).

Tabelle 34: Rivastigmin: Ergebnisse zu alltagspraktischen Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (95 %-KI)	Gruppenunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
B304 1998 ^(b) (26 Wochen)	RIV 2-12 mg (2x/d)	PDS	227	-2,7 (n.g.)	2,3 ^(c) (n.g.)	↗	0,030
	RIV 2-12 mg (3x/d)		225	-1,5 (n.g.)	3,4 ^(c) (n.g.)	↗	0,001
	Placebo		221	-5,0 (n.g.)			
Corey- Bloom 1998 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	PDS	233	-5,2 [-6,5; -3,9]	-0,3 ^(c) (n.g.)	↘	n.g. ^(d)
	RIV Hoch-Dosis (6-12 mg)		231	-1,5 [-2,9; -0,2]	3,4 (1,5; 5,3)	↗	< 0,001
	Placebo		234	-4,9 [-6,2; -3,6]			
Rösler 1999 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	PDS	241	-3,4 [-5,0; -1,6]	-1,2 ^(c) (n.g.)	↘	„n.s.“
	RIV Hoch-Dosis (6-12 mg)		241	0,1 [-1,6; 1,8]	2,2 ^(c) (n.g.)	↗	0,07 ^(e)
	Placebo		237	-2,2 [-3,9; -0,5]			
<p>Erläuterung: KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, RIV = Rivastigmin, SD = Standardabweichung Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in Auswertung b: Daten aus dem Studienbericht des Herstellers c: eigene Berechnung d: Aus den Angaben im Text (S. 61: „This same effect was not seen for the low dose ENA 713 group“) kann auf Nichtsignifikanz geschlossen werden. e: Angabe aus Korrektur: BMJ 2001; 322:1456</p>							

Rivastigmin - Aktivitäten des täglichen Lebens
 Endpunkt: PDS - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=1.83$, $df=2$ ($p=0.400$), $I^2=0\%$
 Gesamteffekt: Z Score= -4.96 ($p=0.000$)

Bemerkung:

Vorzeichen geändert, d. h. negative Werte entsprechen positiven Effekten

B304: Standardabweichungen aus p -Wert = 0,001 berechnet

Corey-Bloom 1998, Rösler 1999: Standardabweichungen aus Konfidenzintervallen berechnet.

Abbildung 39: Rivastigmin: Meta-Analyse zu alltagspraktischen Fähigkeiten

5.3.3.2 Begleitende Psychopathologie

In der Studie Forette 1998 wurden das beobachtbare Verhalten und einzelne Aspekte der Psychopathologie mittels des NOSGER (Nurses' Observation Scale for Geriatric Patients) erfasst. Die Ergebnisse werden aufgrund des hohen Anteils an nicht berücksichtigten Patienten (20–50 % in den Gruppen) nicht dargestellt (siehe Abschnitt 5.2.3.2). In den Studien Corey-Bloom et al. 1998 und Rösler et al. 1999 wurde die Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD) eingesetzt [92], jedoch in den Originalpublikationen nicht berichtet. Die Firma Novartis teilte mit, dass der BEHAVE-AD in diesen Studien nicht ausgewertet, sondern nur zur Berechnung des CIBIC für die Domäne Verhalten herangezogen worden sei.

Damit liegen keine Hinweise auf eine günstige Beeinflussung der begleitenden Psychopathologie durch Rivastigmin vor.

5.3.3.3 Kognitive Leistungsfähigkeit

In allen vier vorliegenden Studien wurde der ADAS-cog zur Erfassung der kognitiven Defizite verwendet.

Für drei Studien (B304, Corey-Bloom 1998 und Rösler 1999) wurden ITT-Auswertungen berichtet. Für die beiden Behandlungsgruppen der Studie B304 (2–12 mg mit zwei- bzw. dreimal täglicher Verabreichung) zeigten sich statistisch signifikante Vorteile gegenüber der Placebo-Gruppe, ebenso wie in den Hochdosis-Gruppen (6–12 mg) bei Corey-Bloom 1998 und Rösler 1999. Für die Gruppe mit niedriger Dosierung (1–4 mg) bei Rösler 1999 dagegen

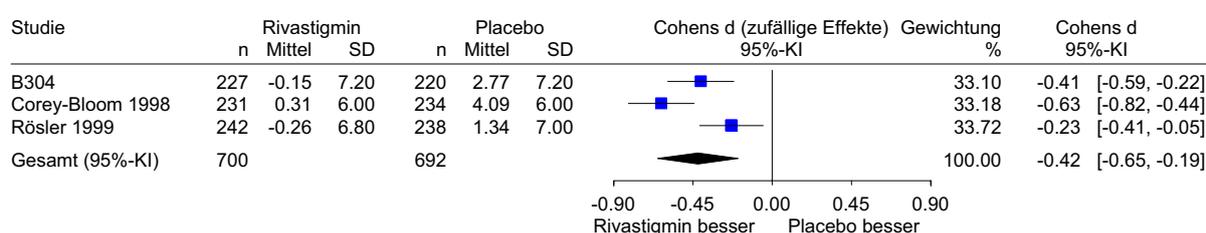
wurde kein statistisch signifikanter Unterschied zur Placebo-Gruppe beobachtet. Auch bei Corey-Bloom 1998 war in der Gruppe mit niedriger Dosierung der Vorteil gegenüber der Placebo-Gruppe deutlich geringer als in der Hochdosis-Gruppe. Hier wurde zwar kein p-Wert angegeben, allerdings deutet der berichtete Unterschied auf statistische Signifikanz hin.

Tabelle 35: Rivastigmin: Ergebnisse zu kognitiven Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (95 % KI)	Behandlungsunterschied zu Placebo		p-Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
B304 1998 ^(b) (26 Wochen)	RIV 2-12 mg (2x/d)	ADAS- cog	228	1,2 ^(c) (n.g.)	-1,61 ^(d) (n.g.)	↗	0,019
	RIV 2-12 mg (3x/d)		227	-0,2 ^(c) (n.g.)	-2,9 ^(d) (n.g.)	↗	< 0,001
	Placebo		220	2,8 ^(c) (n.g.)			
Corey-Bloom 1998 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	ADAS- cog	233	2,4 ^(c) (1,6; 3,1)	-1,7 ^(d) (n.g.)	↗	n.g.
	RIV Hoch-Dosis (6-12 mg)		231	0,3 ^(c) (-0,5; 1,1)	-3,8 (-4,9; -2,7)	↗	< 0,001
	Placebo		234	4,1 ^(c) (3,3; 4,9)			
Rösler 1999 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	ADAS- cog	242	1,4 ^(c) (0,5; 2,3)	0,1 ^(d) (n.g.)	↘	„n.s.“
	RIV Hoch-Dosis (6-12 mg)		242	-0,3 ^(c) (-1,1; 0,7)	-1,6 ^(d) (n.g.)	↗	0,011 ^(e)
	Placebo		238	1,3 ^(c) (0,4; 2,2)			
<p>Erläuterung: KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, n.s. = nicht statistisch signifikant, RIV = Rivastigmin Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Zahl der Patienten in Auswertung b: Daten aus dem Studienbericht des Herstellers c: Vorzeichen sind in Originalpublikation vertauscht d: eigene Berechnung e: Angabe aus Korrektur: BMJ 2001; 322:1456</p>							

Die vorliegenden Studien belegen den Nutzen einer Rivastigmin-Einnahme im Hinblick auf die kognitive Leistungsfähigkeit. Zudem legen Sie das Vorhandensein einer Dosis-Wirkungsbeziehung nahe. Die Meta-Analyse (Abbildung 40) für die jeweils höchste Dosis-Stufe bzw. die Gruppe mit häufigerer täglicher Einnahme zeigt eine starke Heterogenität in den Ergebnissen, die sich nicht eindeutig durch Unterschiede im Design oder den Patientencharakteristika zwischen den Studien erklären lässt. Die geschätzten Effekte bewegen sich im Rahmen von einem bis zwei Drittel der Standardabweichung. Dies entspricht etwa drei Scorepunkten auf der ADAS-cog-Skala. Die Hereinnahme der Ergebnisse aus der Studie B351 ändert am Gesamtergebnis nur wenig (Daten in Anhang G dargestellt).

Rivastigmin - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=9.12$, $df=2$ ($p=0.010$), $I^2=78.1\%$
 Gesamteffekt: Z Score=-3.63 ($p=0.000$), $\tau^2=0.032$

Bemerkung:

B304: Angabe zur Standardabweichung für die Placebo-Gruppe aus [49] für beide Gruppen eingesetzt
 Corey-Bloom 1998, Rösler 1999: Standardabweichungen aus Konfidenzintervallen berechnet

Abbildung 40: Rivastigmin: Meta-Analyse zu kognitiven Fähigkeiten

5.3.3.4 Krankheitsbezogene Lebensqualität

Zu dem Zielkriterium „Krankheitsbezogene Lebensqualität“ wurden in den Studien keine Daten berichtet.

5.3.3.5 Notwendigkeit der vollstationären Pflege (Institutionalisierung)

Zu dem Zielkriterium „Notwendigkeit der vollstationären Pflege (Institutionalisierung)“ wurden in den eingeschlossenen Studien keine Daten berichtet.

5.3.3.6 Mortalität

Todesfälle wurden in den Studien nur vereinzelt berichtet (siehe Tabelle 36).

Ein Hinweis auf eine günstige oder ungünstige Beeinflussung der Mortalität durch Rivastigmin lässt sich aus diesen Daten nicht entnehmen.

5.3.3.7 Unerwünschte Ereignisse

In den Studien war die Rate an unerwünschten Ereignissen und die Rate an hierauf zurückzuführenden Studienabbrüchen bei Rivastigmin-Patienten hoch, vor allem in der Hochdosis-Gruppe. Zum Beispiel erlitten in dieser Gruppe bei Corey-Bloom 1998 48 % der Patienten Übelkeit (Niedrigdosis: 14 %, Placebo: 11 %), 27 % Erbrechen (gegenüber 7 % bzw. 3 %), 20 % Appetitlosigkeit (gegenüber 8 % und 3 %) und 24 % Schwindel (gegenüber 15 % und 13 % in der Placebo-Gruppe). Bei Rösler 1999 verloren darüber hinaus 24 % aus der Hochdosis-Gruppe mehr als sieben Prozent ihres Körpergewichts (gegenüber 9 % und 7 %). In der B304-Studie war die Rate der unerwünschten Ereignisse unabhängig von der Einnahmefrequenz vergleichbar mit derjenigen aus den Hochdosis-Gruppen.

Zusammenfassend belegen die Studien ein gehäuftes Auftreten von unerwünschten Ereignissen durch die Einnahme von Rivastigmin in hoher Dosierung gegenüber Placebo. Die Meta-Analysen (Abbildung 41 bis Abbildung 47) für die jeweils höchste Dosis-Stufe bzw. die Gruppe mit häufigerer täglicher Einnahme zeigen im Hinblick auf Übelkeit, Erbrechen und Appetitlosigkeit ein Odds Ratio von jeweils etwa 8 und damit gegenüber Placebo eine erhebliche Risikoerhöhung. Die ausgeprägte Heterogenität in der Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse ist augenscheinlich auf die Studie B304 zurückzuführen, in der eine im Vergleich zu den anderen Studien niedrigere Mindestdosis bei flexibler Dosierung eingesetzt wurde. Die Effektrichtung aller vier Studien war jedoch identisch und das Gesamtergebnis statistisch signifikant. Die Berücksichtigung der in der Studie B351 beobachteten unerwünschten Ereignisse ändert an diesen Ergebnissen nur wenig (Daten in Anhang G dargestellt).

Im Niedrigdosis-Bereich deuten die Ergebnisse aus den beiden zur Verfügung stehenden Studien (Corey-Bloom 1998 und Rösler 1999) auf ein deutlich niedrigeres Schadenspotenzial hin. Jedoch besteht in diesem Dosisbereich auch kein nachgewiesener Nutzen.

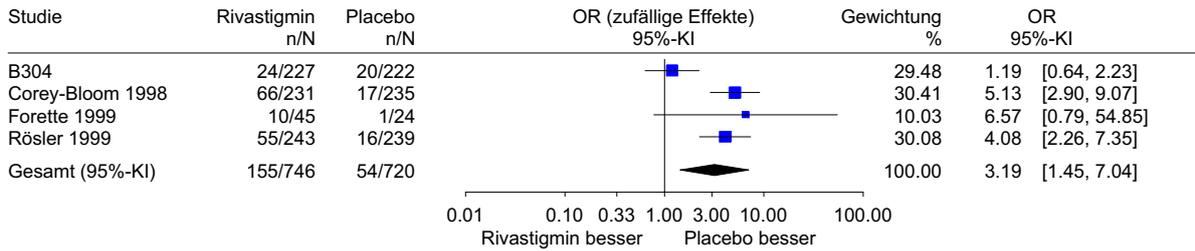
Tabelle 36: Rivastigmin: Studienabbrucher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Studie (Dauer)	Aufdosierung innerhalb des ersten Monats	N ^(a)	Studien- abbruch N (%)	Todes- fälle N	Schwerwiegende unerwünschte Ereignisse N (%)	Abbruch wegen unerwünschter Ereignisse N (%)	Jegliche unerwünschte Ereignisse N (%)	
B304 1998 ^(b) (26 Wochen)	RIV 2-12 mg (2x/d)	2 bis 6 mg	228 ^(c)	54 (24)	1	40 (18)	38 (17)	208 (91)
	RIV 2-12 mg (3x/d)	2 bis 6 mg	227	38 (17)	0	40 (18)	24 (11)	208 (92)
	Placebo		222	33 (15)	1	33 (15)	20 (9)	169 (76)
Corey-Bloom 1998 (26 Wochen)	RIV Niedrig-Dosis (1-4 mg)	k.A.	233	34 (15) ^(d)	0	k.A.	19 (8) ^(d)	„> 85 %“
	RIV Hoch-Dosis (6-12 mg)	k.A.	231	82 (35) ^(d)	1	k.A.	66 (29) ^(d)	„> 85 %“
	Placebo		235	39 (17) ^(d)	0	k.A.	17 (7) ^(d)	„> 85 %“
Forette 1999 (18 Wochen)	RIV 6-12 mg (2x/d)	6 mg ^(e)	45	16 (36) ^(d)	k.A.		14 (31) ^(d)	k.A.
	RIV 6-12 mg (3x/d)	6 mg ^(e)	45	11 (24) ^(d)	k.A.	gesamt: 13 (11) ^(d)	10 (22) ^(d)	k.A.
	Placebo		24	2 (8) ^(d)	k.A.		1 (4) ^(d)	k.A.
Rösler 1999 (26 Wochen)	RIV Niedrig-Dosis (1-4 mg)	bis zu 4 mg	243	34 (14)	0	„≈ 18 %“	18 (7)	172 (71)
	RIV Hoch-Dosis (6-12 mg)	bis zu 6 mg	243	79 (33)	1	„≈ 18 %“	55 (23)	220 (91)
	Placebo		239	31 (13)	0	„≈ 18 %“	16 (7)	172 (72)
Erläuterung: k.A. = keine Angaben, N = Anzahl, RIV = Rivastigmin								
a: Anzahl randomisierter Patienten, sofern nicht anders angegeben								
b: Daten aus dem Studienbericht des Herstellers								
c: Eine Person der 2x/d-Gruppe weigerte sich, die Studienmedikation einzunehmen und wurde deshalb aus allen Sicherheits-Analysen ausgeschlossen.								
d: %-Angabe aus N berechnet								
e: 2 mg initial, am 4. Tag + 1 mg, alle 4 Tage + 0,5 mg								

Tabelle 37: Rivastigmin: unerwünschte Ereignisse

Studie (Dauer)	Auf- dosierung innerhalb des ersten Monats	N ^(a)	Anteil (%) mit UE											
			Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein	
B304 ^(b) (26 Wochen)	RIV 2-12 mg (2x/d)	2 bis 6 mg	228 ^(c)	54	39	18	18	21	k.A.	9	18	15	5	6
	RIV 2-12 mg (3x/d)	2 bis 6 mg	227	48	30	17	17	19	k.A.	6	16	11	6	7
	Placebo		222	14	6	9	7	3	k.A.	12	10	5	5	3
Corey-Bloom 1998 (26 Wochen) Eindosierungs- phase	RIV Niedrigdosis (1-4 mg)	k.A.	233	14	7	k.A.	15	8	1	k.A.	k.A.	k.A.	5	1
	RIV Hochdosis (6-12 mg)	k.A.	231	48	27	k.A.	24	20	4	k.A.	k.A.	k.A.	10	3
	Placebo		235	11	3	k.A.	13	3	1	k.A.	k.A.	k.A.	4	1
Forette 1999 (18 Wochen)	RIV 6-12 mg (2x/d)	6 mg ^(d)	45	58	38	k.A.	27	18	k.A.	k.A.	16	k.A.	k.A.	k.A.
	RIV 6-12 mg (3x/d)	6 mg ^(d)	45	58	31	k.A.	9	16	k.A.	k.A.	20	k.A.	k.A.	k.A.
	Placebo		24	8	4	n.g.	0	0	k.A.	k.A.	4	k.A.	k.A.	k.A.
Rösler 1999 (26 Wochen)	RIV Niedrigdosis (1-4 mg)	bis zu 4 mg	242 ^(e)	17	8	10	10	3	9 ^(f)	k.A.	7	5	2	1
	RIV Hochdosis (6-12 mg)	bis zu 6 mg	242 ^(e)	50	34	17	20	14	24 ^(f)	k.A.	19	12	10	10
	Placebo		239	10	6	9	7	2	7 ^(f)	k.A.	8	3	3	2
Erläuterung: k.A. = keine Angaben, N = Anzahl, RIV = Rivastigmin, UE = unerwünschte Ereignisse														
a: Anzahl randomisierter Patienten, sofern nicht anders angegeben														
b: Daten aus dem Studienbericht des Herstellers														
c: Eine Person der 2x/d-Gruppe weigerte sich, die Studienmedikation einzunehmen und wurde deshalb aus allen Sicherheitsanalysen ausgeschlossen.														
d: 2 mg initial, am 4. Tag + 1 mg, alle 4 Tage + 0,5 mg														
e: Die Sicherheitsdaten wurden bei jeweils 242 (statt der randomisierten 243) Patienten in den beiden Verum-Gruppen erhoben.														
f: Gewichtsverlust > 7 % des Körpergewichtes														

Rivastigmin - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



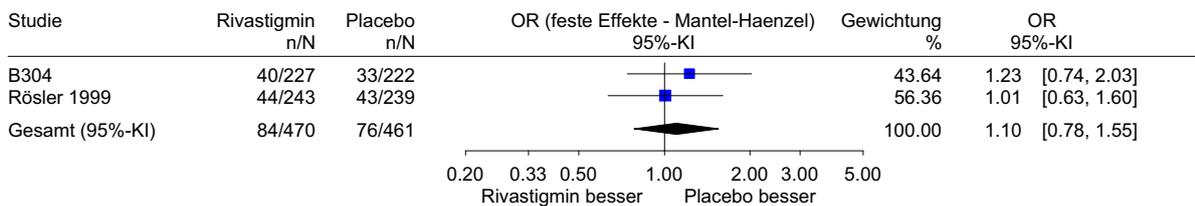
Heterogenität: $Q=13.31$, $df=3$ ($p=0.004$), $I^2=77.5\%$
 Gesamteffekt: Z Score= 2.88 ($p=0.004$), $\tau^2=0.451$

Bemerkung:

Aufgrund der Heterogenität wurde ein Modell mit zufälligen Effekten gewählt. Ein Modell mit festen Effekten ergibt ähnliche Ergebnisse.

Abbildung 41: Rivastigmin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse

Rivastigmin - Schwerwiegende UEs
 Endpunkt: SUE erlitten (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit SUE



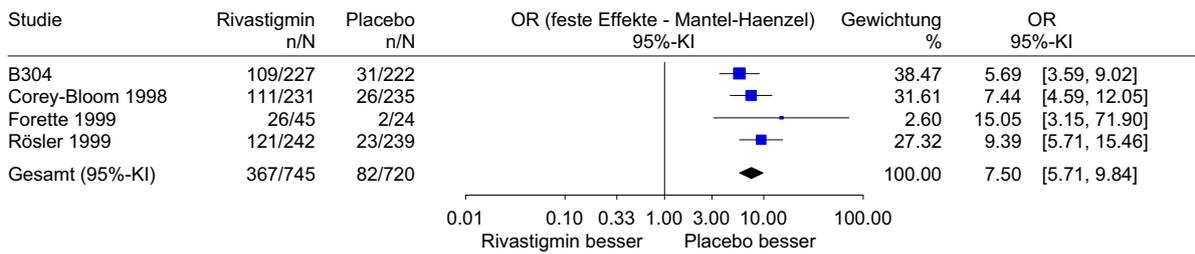
Heterogenität: $Q=0.31$, $df=1$ ($p=0.576$), $I^2=0\%$
 Gesamteffekt: Z Score= 0.56 ($p=0.574$)

Abbildung 42: Rivastigmin: Meta-Analyse zu schwerwiegenden unerwünschten Ereignissen

Rivastigmin - UEs

Endpunkt: Übelkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: Q=2.93, df=3 (p=0.403), I²=0%

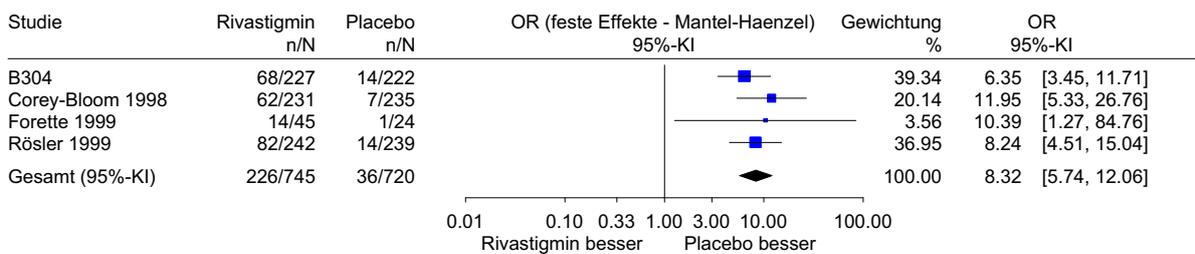
Gesamteffekt: Z Score=14.54 (p=0.000)

Abbildung 43: Rivastigmin: Meta-Analyse zum Endpunkt „Übelkeit“

Rivastigmin - UEs

Endpunkt: Erbrechen

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf

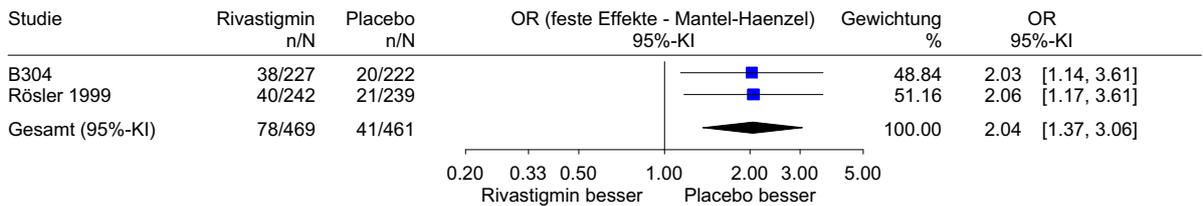


Heterogenität: Q=1.57, df=3 (p=0.667), I²=0%

Gesamteffekt: Z Score=11.18 (p=0.000)

Abbildung 44: Rivastigmin: Meta-Analyse zum Endpunkt „Erbrechen“

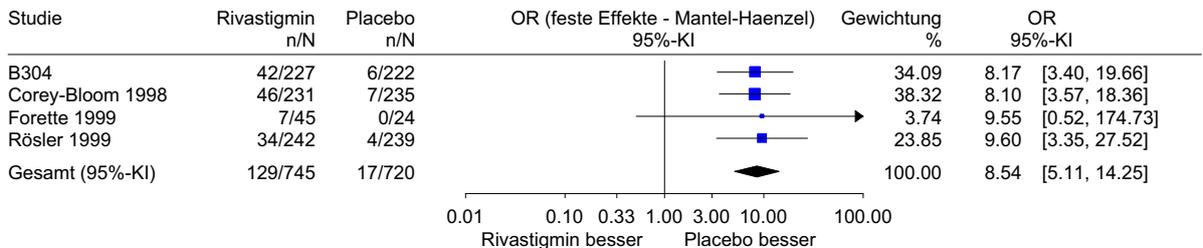
Rivastigmin - UEs
 Endpunkt: Diarrhoe
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=0$, $df=1$ ($p=0.976$), $I^2=0\%$
 Gesamteffekt: Z Score=3.48 ($p=0.000$)

Abbildung 45: Rivastigmin: Meta-Analyse zum Endpunkt „Diarrhoe“

Rivastigmin - UEs
 Endpunkt: Appetitlosigkeit
 Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf

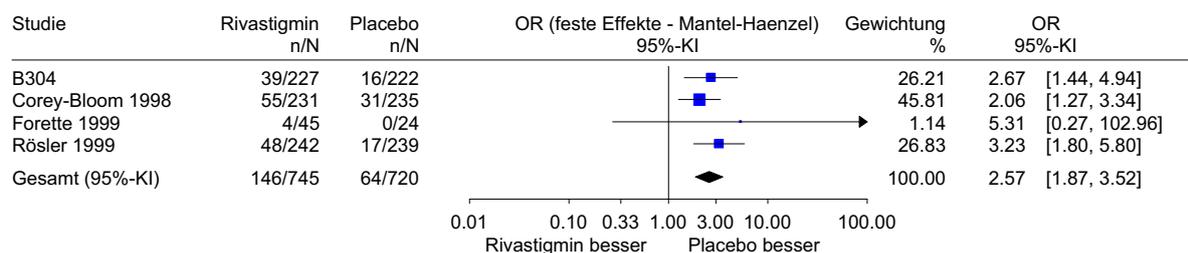


Heterogenität: $Q=0.08$, $df=3$ ($p=0.994$), $I^2=0\%$
 Gesamteffekt: Z Score=8.2 ($p=0.000$)

Abbildung 46: Rivastigmin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“

Rivastigmin - UEs
 Endpunkt: Schwindel

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=1.65$, $df=3$ ($p=0.649$), $I^2=0\%$
 Gesamteffekt: Z Score=5.86 ($p=0.000$)

Abbildung 47: Rivastigmin: Meta-Analyse zum Endpunkt „Schwindel“

5.3.3.8 Lebensqualität der (betreuenden) Angehörigen

Zu dem Zielkriterium „Lebensqualität der (betreuenden) Angehörigen“ wurden in den Studien keine Daten berichtet.

5.3.3.9 Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en

Im Studienbericht zur Studie B304 wurde berichtet, dass mit dem Amendment 6 des Studienprotokolls im Mai 1995 der Caregiver Activity Survey (CAS) als weiteres Effektivitätsmaß eingeführt wurde, um den Zusammenhang zwischen Rivastigmingabe und Betreuungszeiten zu erfassen. Im letzten Amendment des Studienprotokolls vom März 1997 wurde jedoch festgelegt, dass die Analyse des CAS nur im Rahmen einer Poolung aller vier Phase-III-Studien erfolgen sollte (S. 56: „Stated that the analysis of the CAS was only to be performed on data pooled from all four phase III studies.“). Es bleibt unklar, ob diese Entscheidung vor oder nach der Entblindung bzgl. der Patientendaten getroffen wurde – der letzte Patient vollendete die Studie im September 1996. Weiterhin lässt diese Anmerkung darauf schließen, dass auch in den drei anderen Phase-III-Studien (B351, Corey-Bloom 1998, Rösler 1999) der CAS eingesetzt worden ist. In den Veröffentlichungen zu Corey-Bloom 1998 sowie Rösler 1999 wurden hierzu jedoch keine Angaben gemacht. Es konnten auch keine Publikationen zu gepoolten Analysen dieser vier Studien zum CAS identifiziert werden. Die Firma Novartis teilte mit, dass es keine Auswertungen zu den CAS-Daten gebe.

Damit liegen keine Hinweise auf eine günstige Beeinflussung des Betreuungsaufwandes durch Rivastigmin vor. Die bislang nicht erfolgte Publikation der Auswertung offenbar diesbezüglich erhobener Daten legt den Schluss nahe, dass ein günstiger Effekt durch Rivastigmin auf dieses Zielkriterium sogar eher unwahrscheinlich ist.

5.3.3.10 Ergänzende Information: klinisches Krankheitsstadium

In allen vier Studien wurde der klinische Gesamteindruck unter Verwendung des CIBIC-plus ermittelt.

Der Unterschied im Anteil der als gebessert beurteilten Personen ($\text{Score} \leq 3$) wurde in zwei Studien mittels einer ITT-Auswertung überprüft (Tabelle 38). In der Studie B304 ergab sich ein statistisch signifikanter Vorteil nur für die Gruppe mit dreimal täglicher Rivastigmingabe (31 %), nicht jedoch für die Gruppe, in der Rivastigmin zweimal täglich verabreicht wurde (23 %; Placebo: 19 %). Bei Rösler 1999 zeigte sich nach 26 Wochen sowohl für die Hoch- als auch für die Niedrigdosis-Gruppe ein statistisch signifikant größerer Anteil an Patienten mit einer Besserung des klinischen Gesamteindrucks (37 % bzw. 30 %) als in der Placebo-Gruppe (20 %). Bei Corey-Bloom 1998 wurde als ITT-Analyse nicht der Anteil der als gebessert beurteilten Personen, sondern lediglich die kontinuierliche Auswertung der CIBIC-Skala berichtet. Es zeigte sich sowohl in der Hoch- als auch in der Niedrigdosis-Gruppe eine Verringerung des CIBIC-Scores, d. h. also eine Verbesserung, im Vergleich zu Placebo. Nur für die Hochdosis-Gruppe wurde ein p-Wert angegeben ($p < 0,01$).

Zusammenfassend liegen damit Hinweise vor, dass Rivastigmin den klinischen Gesamteindruck verbessern kann.

Tabelle 38: Rivastigmin: Ergebnisse zu CIBIC-plus ≤ 3 (Verbesserung)

Studie (Dauer)	Dosierung	End- punkt	N ^(a)	Anteil (%) mit Score ≤ 3	Behandlungsunterschied zu Placebo		p-Wert
					Differenz (95 %-KI)	Richtung des Effekts	
B304 ^(b) (26 Wochen)	RIV 2-12 mg (2x/d)	CIBIC- plus	222	23	4% ^(c) (n.g.)	↗	0,260
	RIV 2-12 mg (3x/d)		222	31	12% ^(c) (n.g.)	↗	0,002
	Placebo		216	19			
Rösler 1999 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	CIBIC- plus	233	30	10% ^(c) (n.g.)	↗	< 0,05
	RIV Hoch-Dosis (6-12 mg)		219	37	17% ^(c) (n.g.)	↗	< 0,001
	Placebo		230	20			
Mittlere Differenz zur Baseline (KI)							
Corey- Bloom 1998 (26 Wochen)	RIV Niedrig- Dosis (1-4 mg)	CIBIC- plus	233	0,2 (0,1; 0,4)	-0,3 ^(c) (n.g.)	↗	n.g.
	RIV Hoch-Dosis (6-12 mg)		231	0,2 (0,0; 0,4)	-0,3 (-0,5; -0,1)	↗	< 0,01
	Placebo		234	0,5 (0,3; 0,7)			
<p>Erläuterung: KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, RIV = Rivastigmin Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (↗) oder eine Verschlechterung (↘) darstellt; er enthält keine Informationen über die Größe des Effektes.</p> <p>a: Anzahl Patienten in Auswertung, sofern nicht anders angegeben b: Daten aus dem Studienbericht des Herstellers c: Werte berechnet</p>							

5.3.4 Vergleiche unterschiedlicher Cholinesterasehemmer

5.3.4.1 Galantamin vs. Donepezil

5.3.4.1.1 Aktivitäten des täglichen Lebens

Aktivitäten des täglichen Lebens, gemessen anhand der BADLS, stellten bei Wilcock 2003 das primäre Zielkriterium dar. Es wurde berichtet, dass die BADLS-Werte in beiden Gruppen bis zum 9. Monat konstant blieben und sich danach verschlechterten. In einer Kovarianzanalyse mit Alter und MMSE-Werten zum Zeitpunkt des Screenings als Kovariaten zeigte sich kein statistisch signifikanter Unterschied zwischen den Patienten unter Galantamin- bzw. Donepezilgabe ($p > 0,5$). Ein ähnliches Ergebnis zeigte sich in einer Subgruppenanalyse von Patienten mit einem MMSE-Wert zwischen 12 und 18.

5.3.4.1.2 Begleitende Psychopathologie

Im Hinblick auf Veränderungen im NPI zeigten sich bei Wilcock 2003 keine Unterschiede zwischen den Patienten in der Galantamin- und der Donepezil-Gruppe. Genaue Daten wurden in der Publikation nicht angegeben.

5.3.4.1.3 Kognitive Leistungsfähigkeit

In der Studie Wilcock 2003 wurden die Effekte von Galantamin bzw. Donepezil auf die kognitive Leistungsfähigkeit anhand von MMSE und ADAS-cog überprüft. Wie auch bei den anderen Vergleichen im vorliegenden Bericht werden an dieser Stelle in erster Linie die Ergebnisse zum ADAS-cog vorgestellt. Nach anfänglichen Verbesserungen im ADAS-cog in beiden Gruppen wiesen die Patienten zum Studienendpunkt nach 52 Wochen in beiden Gruppen eine Verschlechterung im Vergleich zum Ausgangswert auf, die in der Galantamin-Gruppe etwas geringer ausgeprägt war. Der Unterschied zwischen den beiden Gruppen war allerdings zum Studienendpunkt statistisch nicht signifikant.

Explorative Analysen legten hinsichtlich des ADAS-cog zwar einen stärkeren Effekt von Galantamin in der Subgruppe der mäßig dementen Patienten (MMSE: 12–18) nahe, allerdings kann der Publikation nicht entnommen werden, ob diese Subgruppenanalyse a priori geplant war. Zudem liegen keine Ergebnisse eines statistischen Interaktionstests vor. Die Ergebnisse zum MMSE entsprachen von der Tendenz her denen zum ADAS-cog.

5.3.4.1.4 Krankheitsbezogene Lebensqualität

Zu dem Zielkriterium „Krankheitsbezogene Lebensqualität“ wurden in der Studie keine Daten berichtet.

5.3.4.1.5 Notwendigkeit einer vollstationären Pflege (Institutionalisierung)

Zu dem Zielkriterium „Notwendigkeit einer vollstationären Pflege (Institutionalisierung)“ wurden in der Studie keine Daten berichtet.

5.3.4.1.6 Mortalität

In der Wilcock 2003-Studie starben unter Galantamin zwei Patienten (2,1 %) und unter Donepezil drei Patienten (3,3 %)

5.3.4.1.7 Unerwünschte Ereignisse

Unter Galantamin und Donepezil kam es bei Wilcock 2003 zu jeweils fast gleich hohen Abbrecherraten wegen unerwünschter Ereignisse (11 % vs. 12 %) und vergleichbaren Abbrecherraten insgesamt (20 % vs. 22 %). Auch der Anteil an Patienten mit jeglichen unerwünschten Ereignissen war in beiden Gruppen etwa gleich hoch (91 % in der Galantamin- vs. 93 % in der Donepezil-Gruppe). Übelkeit (Galantamin 20 %, Donepezil 18 %) und Erbrechen (Galantamin 18 %, Donepezil 14 %) traten in beiden Gruppen ungefähr gleich häufig auf. Geringe Unterschiede zuungunsten von Galantamin zeigten sich hinsichtlich Agitation (19 % vs. 12 %), Stürzen (17 % vs. 9 %) und Kopfschmerzen (17 % vs. 12 %).

Die Rate an schwerwiegenden unerwünschten Ereignissen lag bei 19 % unter Galantamin und 20 % unter Donepezil.

5.3.4.1.8 Lebensqualität der (betreuenden) Angehörigen

In der Studie Wilcock 2003 wurden mittels des Screen for Caregiver Burden (SCGB) die objektiven und subjektiven Belastungen der betreuenden Personen ermittelt. Über eine Verbesserung oder Aufrechterhaltung der objektiven Belastung berichteten in der Galantamin-Gruppe zum Studienendpunkt 67 % der Betreuer (von n = 57) bzw. hinsichtlich der subjektiven Belastung 68 % (von n = 85). In der Donepezil-Gruppe lagen die Werte bei 51 % (von n = 41) für die objektiven Belastungen und 49 % (von n = 79) für die subjektiven Belastungen. Über die statistische Signifikanz dieser Unterschiede wurde nichts berichtet. Diese Auswertung ist jedoch nur sehr eingeschränkt interpretierbar, da für bis zur Hälfte der betroffenen Betreuer keine Angaben vorlagen und es unklar bleibt, inwieweit die Angaben zu den subjektiven Belastungen auf (deutlich) größeren Fallzahlen beruhen als diejenigen zu den objektiven Belastungen, wenn doch die Werte für die subjektiven aus denen für die objektiven Belastungen errechnet wurden.

5.3.4.1.9 Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en

Zu dem angehörigerelevanten Zielkriterium „Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en“ wurden in der Studie keine Daten berichtet.

5.3.4.1.10 Ergänzende Information: klinisches Krankheitsstadium

Zu dem Zielkriterium „Klinisches Krankheitsstadium“ wurden in der Studie keine Daten berichtet.

5.3.4.2 Rivastigmin vs. Donepezil

5.3.4.2.1 Aktivitäten des täglichen Lebens

In der Bullock 2005-Studie ergab sich nach zwei Jahren in der ITT-LOCF-Auswertung ein statistisch signifikanter Vorteil in der ADCS-ADL-Skala zugunsten von Rivastigmin. So hatten sich nach zwei Jahren die Patienten in der Rivastigmin-Gruppe um 2,1 Punkte weniger verschlechtert als die Patienten in der Donepezil-Gruppe ($p = 0,047$). Die Ergebnisse sind allerdings von eingeschränkter Aussagekraft, da die Rate von Therapieabbrüchen in beiden Gruppen hoch und zwischen den beiden Gruppen unterschiedlich war. So hatten bis Studienende nahezu 50 % der Rivastigmin- und knapp 37 % der Donepezil-Patienten die Studie bereits vorzeitig beendet. In der Publikation waren lediglich die Least Squares (LS) Means einschließlich Standardfehler für die Veränderung pro Gruppe angegeben. Bei Verwendung des t-Tests ergibt sich ein mehr als doppelt so großer p-Wert ($p = 0,103$; eigene Berechnung). Dies ist auffällig, da es sich um eine randomisierte Studie mit vielen Patienten handelt, bei der ein großer Unterschied zwischen den beiden Auswertungen nicht zu erwarten ist. Eine Anfrage an die Firma Novartis lieferte die unadjustierten Mittelwerte und Standardabweichungen, aus denen sich für den t-Test ein p-Wert von 0,016 ergibt, was wiederum wesentlich kleiner ist als der aus der ANCOVA berechnete p-Wert. Es kann vermutet werden, dass diese Diskrepanzen aufgrund der Abbruchraten zustande kommen. Allerdings ließ sich dies nicht abschließend klären. Schließlich wird in der Publikation bemerkt, dass die Unterschiede in den Per-Protokoll-Auswertungen keine statistische Signifikanz mehr besaßen. Ob dies auf einer Änderung des Effektschätzers und/oder einer verminderten Power beruht, bleibt unklar.

Bei Fuschillo 2001 wurden zur PSMS-Skala lediglich die Ergebnisse von Intra-Gruppen-Vergleichen dargestellt. Sowohl in der Rivastigmin- als auch in der Donepezil-Gruppe zeigten sich nach 30 Wochen im Vergleich zum Ausgangswert praktisch keine Veränderungen auf der PSMS-Skala.

In der Studie Wang 2001 fanden sich nach 16 Wochen zwischen Rivastigmin und Donepezil keine Unterschiede in der Veränderung der Werte der Blessed Roth Dementia Scale ($p = 0,472$).

Zusammenfassend sind, bezogen auf alltagspraktische Fähigkeiten, die Daten nicht ausreichend für einen Beleg der Über- oder Unterlegenheit einer der beiden Wirkstoffe, sondern können allenfalls als diesbezüglicher Hinweis interpretiert werden.

Tabelle 39: Rivastigmin vs. Donepezil: Ergebnisse zu alltagspraktischen Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied (RIV vs. DON)		p- Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Bullock 2005 (24 Monate)	RIV 3-12 mg	ADCS /ADL	454	-12,8 (19,2 ^(b))	2,1 (n.g.)	RIV > DON	0,047
	DON 5-10 mg		475	-14,9 (19,6 ^(b))			
Fuschillo 2001 (7,5 Monate)	RIV 6-9 mg	PSMS	11	-0,5 ^(b)	-0,4 ^(b)	RIV < DON	n.g.
	DON 5 mg		16	-0,1 ^(b)			
Wang 2001 (4 Monate)	RIV 3-6 mg	BRDS	59/60 ^(c)	-1,8 (2,6)	-0,4 ^(b) (n.g.)	RIV > DON	0,472
	DON 5-10 mg		61	-1,4 (3,2)			

Erläuterung:
 DON = Donepezil, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, RIV = Rivastigmin, SD = Standardabweichung
 > = numerisch größerer Behandlungseffekt

a: Zahl der Patienten in Auswertung
 b: eigene Berechnung
 c: In der Publikation wurden widersprüchliche Angaben gemacht – es gingen 59 oder 60 Patienten in die Auswertung ein.

5.3.4.2.2 Begleitende Psychopathologie

Begleitende psychopathologische Symptome wurden nur in der Studie von Bullock 2005 untersucht. Hier zeigten sich im Rahmen der ITT-LOCF-Analyse im NPI-10 keine Behandlungsunterschiede zwischen den Patienten in der Rivastigmin- und der Donepezil-Gruppe. Die Veränderung nach zwei Jahren betrug im Vergleich zur Baselinemessung in der Rivastigmin-Gruppe 2,4 Punkte und in der Donepezil-Gruppe 2,9 Punkte ($p = 0,554$). Zu den Subskalen wurden keine Ergebnisse berichtet.

5.3.4.2.3 Kognitive Leistungsfähigkeit

In der Studie von Bullock 2005 ergab sich in dem primären Zielkriterium SIB (Severe Impairment Battery) nach zwei Jahren kein Unterschied zwischen den Patienten der Donepezil- und der Rivastigmin-Gruppe (ITT-LOCF). Die Auswertungen zum sekundären Ergebnisparameter MMSE bestätigten diese Ergebnisse.

Bei Fuschillo 2001 wurden zum ADAS-cog wiederum lediglich die Ergebnisse von Intra-Gruppen-Vergleichen dargestellt. In der Rivastigmin- und in der Donepezil-Gruppe zeigten sich nach 30 Wochen im Vergleich zum Ausgangswert Verbesserungen um durchschnittlich 3,8 bzw. 3,6 Scorepunkte (eigene Berechnung). Eine statistische Signifikanz dieses geringfügigen Unterschieds zwischen den Gruppen ist nicht anzunehmen.

In der Studie von Wang 2001 unterschied sich nach 16 Wochen das Ausmaß der Verbesserungen im MMSE zwischen den beiden Gruppen nicht statistisch signifikant ($p = 0,422$).

Somit liegen keine Belege für eine Überlegenheit von Donepezil oder Rivastigmin hinsichtlich einer positiven Beeinflussung der kognitiven Fähigkeiten vor.

Tabelle 40: Rivastigmin vs. Donepezil: Ergebnisse zu kognitiven Fähigkeiten

Studie (Dauer)	End- punkt	N ^(a)	Mittlere Differenz zu Baseline (SD)	Behandlungsunterschied (RIV vs. DON)		p- Wert	
				Differenz (95 %-KI)	Richtung des Effekts		
Bullock 2005 (24 Monate)	RIV 3-12 mg	SIB	471	-9,3 (23,9 ^(b))	0,6 ^(b) (n.g.)	RIV > DON	0,609
	DON 5-10 mg		483	-9,9 (24,2 ^(b))			
Fuschillo 2001 (7,5 Monate)	RIV 6-9 mg	ADAS -cog	11	-3,8 ^(b)	-0,2	RIV > DON	n.g.
	DON 5 mg		16	-3,6 ^(b)			
Wang 2001 (4 Monate)	RIV 3-6 mg	MMSE	59/60 ^(c)	2,5 (3,9)	0,6 ^(b) (n.g.)	RIV > DON	0,422
	DON 5-10 mg		61	1,9 (3,4)			
Erläuterung: DON = Donepezil, KI = Konfidenzintervall, N = Anzahl, n.g. = nicht genannt, SD = Standardabweichung > = numerisch größerer Behandlungseffekt a: Zahl der Patienten in Auswertung b: eigene Berechnung c: In der Publikation werden widersprüchliche Angaben gemacht – es gingen 59 oder 60 Patienten in die Auswertung ein.							

5.3.4.2.4 Krankheitsbezogene Lebensqualität

Zu dem Zielkriterium „Krankheitsbezogene Lebensqualität“ wurden in den Studien keine Daten berichtet.

5.3.4.2.5 Notwendigkeit einer vollstationären Pflege (Institutionalisierung)

Zu dem Zielkriterium „Notwendigkeit einer vollstationären Pflege (Institutionalisierung)“ wurden in den eingeschlossenen Studien keine Daten berichtet.

5.3.4.2.6 Mortalität

In der Studie von Bullock 2005 war in beiden Verum-Gruppen die Mortalität vergleichbar (Rivastigmin: 5,3 %, Donepezil: 6,8 %). Bei Fuschillo 2001 und Wang 2001 wurden keine Todesfälle berichtet.

5.3.4.2.7 Unerwünschte Ereignisse

Unter Rivastigmin kam es in der Studie von Bullock 2005 zu mehr Abbrechern wegen unerwünschter Ereignisse (26 % vs. 16 %) und zu mehr Abbrechern insgesamt (48 % vs. 37 %) als unter Donepezil. Die Autoren stellten die unerwünschten Ereignisse getrennt für die Aufdosierungsphase (16 Wochen) und Erhaltungsphase (Woche 17 bis 104) dar. Betrachtet man alle unerwünschten Ereignisse, so bestand in der Aufdosierungsphase ein deutlicher Unterschied zuungunsten von Rivastigmin. Der Anteil an Patienten mit mindestens einem unerwünschten Ereignis war in der Aufdosierungsphase unter Rivastigmin 17 Prozentpunkte höher als unter Donepezil (82 % vs. 65 %). Dabei unterschieden sich die beiden Gruppen vor allem bezüglich Übelkeit (33 % unter Rivastigmin vs. 15 % unter Donepezil), Erbrechen (28 % vs. 6 %), Appetitverlust (9 % vs. 4 %) und Gewichtsabnahme (6 % vs. 2 %).

Zwar war bei Bullock 2005 die Anzahl an Patienten mit unerwünschten Ereignissen insgesamt in der Erhaltungsphase zwischen den beiden Gruppen nicht unterschiedlich (79 % unter Rivastigmin vs. 77 % unter Donepezil), hinsichtlich Übelkeit, Erbrechen und Appetitverlust zeigten sich jedoch weiterhin Unterschiede (Übelkeit: 13 % vs. 5 %; Erbrechen: 15 % vs. 4 %, Appetitverlust 6 % vs. 3 %). Bei der Interpretation dieser Daten in der Erhaltungsphase ist zu beachten, dass Patienten, die die Studie wegen unerwünschter Ereignisse in der Aufdosierungsphase abgebrochen hatten, nicht mehr eingeschlossen waren. Die Rate an schwerwiegenden unerwünschten Ereignissen war insgesamt bei beiden Medikamenten gleich.

In der Studie von Fuschillo 2001 wurde berichtet, dass es hinsichtlich der Inzidenz unerwünschter Ereignisse keinen statistisch signifikanten Unterschied zwischen den Gruppen gab. Übelkeit und Erbrechen waren meist während der Titrationsphase von Rivastigmin zu beobachten; Schwindel und Kopfschmerzen traten bei höheren Rivastigmin-Dosen häufiger

auf und klangen ohne Behandlung ab. Schwerwiegende unerwünschte Ereignisse traten in der Studie nicht auf, ebenso brach kein Patient die Studie wegen unerwünschter Nebenwirkungen ab. Die in der Publikation zu den einzelnen unerwünschten Ereignissen dargestellten Prozentangaben sind allerdings nicht plausibel und können deshalb nicht weiter interpretiert werden.

In der Studie Wang 2001 brachen zwei Personen unter Rivastigmin und eine Person unter Donepezil die Studie wegen unerwünschter Ereignisse ab. Zur Häufigkeit des Auftretens der verschiedenen Nebenwirkungen lagen keine auswertbaren Daten vor.

Tabelle 41: Rivastigmin vs. Donepezil: Studienabbrecher, Todesfälle und Anzahl der Patienten mit unerwünschten Ereignissen

Studie (Dauer)		Aufdosierung innerhalb des ersten Monats	N ^(a)	Studien- abbruch N (%)	Todes- fälle N	Schwerwiegende unerwünschte Ereignisse N (%)	Abbruch wegen unerwünschter Ereignisse N (%)	Jegliche unerwünschte Ereignisse N (%)
Bullock 2005 (24 Monate)	RIV 3-12 mg	3 mg	498	237 (48)	26	157 (32)	128 (26)	406 (82) ^(b)
	DON 5-10 mg	5 mg	500	183 (37)	34	162 (32)	80 (16)	323 (65) ^(b)
Fuschillo 2001 (7,5 Monate)	RIV 6-9 mg	6 mg	11	k.A.	k.A.	0	0	k.A. ^(c)
	DON 5 mg	5 mg	16	k.A.	k.A.	0	0	k.A. ^(c)
Wang 2001 (4 Monate)	RIV 3-6 mg	3 mg	62	2 (3)	k.A.	k.A.	2 (3)	k.A.
	DON 5-10 mg	5 mg	62	1 (2)	k.A.	k.A.	1 (2)	k.A.
<p>Erläuterung: DON = Donepezil, k.A. = keine Angaben, N = Anzahl, RIV = Rivastigmin a: Anzahl randomisierter Patienten b: Angaben gelten für die Eindosierungsphase (Woche 1-16); Erhaltungsphase (Woche 17-104): Rivastigmin 79 %, Donepezil 77 %. c: Im Text wird lediglich berichtet, dass zwischen beiden Gruppen kein Unterschied in der Inzidenz unerwünschter Ereignisse bestand.</p>								

Tabelle 42: Rivastigmin vs. Donepezil: unerwünschte Ereignisse

Studie (Dauer)	Aufdosierung innerhalb des ersten Monats	N ^(a)	Übelkeit	Erbrechen	Diarrhoe	Schwindel	Appetit- losigkeit	Gewichts- verlust	Agitation	Kopf- schmerzen	Abdominale Beschwerden	Müdigkeit	Unwohlsein	
														Anteil (%) mit UE
Bullock 2005 (24 Monate) Eindosierungs- -phase	RIV 3-12 mg	3 mg	495 ^(b)	33	28	8	k.A.	9	6	7	6	k.A.	k.A.	k.A.
	DON 5-10 mg	5 mg	499 ^(b)	15	6	7	k.A.	4	2	10	5	k.A.	k.A.	k.A.
Fuschillo 2001 ^(c) (7,5 Monate)	RIV 6-9 mg	6 mg	11											
	DON 5 mg	5 mg	16											
Erläuterung: DON = Donepezil, k.A. = keine Angaben, N = Anzahl, RIV = Rivastigmin, UE = unerwünschte Ereignisse														
a: Anzahl randomisierter Patienten, sofern nicht anders angegeben														
b: In der Safety-Auswertung wurden nur die Personen berücksichtigt, die die Studienmedikation erhalten haben.														
c: Angesichts der Teilnehmerzahlen von 11 und 16 sind die berichteten %-Angaben nicht plausibel und werden deshalb hier nicht dargestellt.														

5.3.4.2.8 Lebensqualität der (betreuenden) Angehörigen

Zu dem angehörigerelevanten Zielkriterium „Lebensqualität der (betreuenden) Angehörigen“ wurde in der Studie Bullock 2005 der NPI-D erhoben. Ergebnisse für Subgruppen nach Alter werden in Bullock 2006 [84] berichtet. Die Analyse des NPI-D für die Gesamtstudienpopulation ist nicht publiziert. Auf Anfrage teilte die Firma Novartis mit, dass der NPI-D in der Gesamtstudienpopulation ausschließlich deskriptiv für die Observed Cases analysiert wurde. Da diese Auswertung eine begrenzte Aussagekraft hat, wurde sie nicht angefordert.

5.3.4.2.9 Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en

Zu dem angehörigerelevanten Zielkriterium „Betreuungsaufwand durch eine oder mehrere betreuende Person/en oder Institution/en“ wurden in den Studien keine Daten berichtet.

5.3.4.2.10 Ergänzende Information: klinisches Krankheitsstadium

Sowohl in der Studie Bullock 2005 als auch bei Wang 2001 wurde das klinische Krankheitsstadium mittels der GDS gemessen.

Bei Bullock 2005 ergab sich in der ITT-LOCF-Analyse ein Vorteil für Rivastigmin. So lag der Wert auf der Skala bei den Patienten in der Donepezil-Gruppe nach zwei Jahren durchschnittlich um 0,69 Punkte höher als zur Baselinemessung, in der Rivastigmin-Gruppe nur um 0,58 Punkte ($p = 0,049$ im Wilcoxon-Test). Die Ergebnisse sind allerdings aufgrund der hohen Rate der Therapieabbrüche, die zwischen den beiden Gruppen unterschiedlich war, nur eingeschränkt interpretierbar.

In der Studie von Wang 2001 zeigten sich zwischen Rivastigmin- und Donepezil-Bedingung keine Unterschiede in der Veränderung der GDS nach 16 Wochen ($p = 0,126$).

5.3.4.3 Rivastigmin vs. Donepezil vs. Galantamin

Lediglich die Studie von Cumbo 2005 untersuchte den direkten Vergleich zwischen den drei Cholinesterasehemmern Rivastigmin, Donepezil und Galantamin. In der Publikation zu dieser Studie wurden nur die Ergebnisse zum primären Zielkriterium in interpretierbarer Weise dargestellt. Deshalb wird im Folgenden auf eine Gliederung der Ergebnisdarstellung nach allen Zielkriterien gemäß Abschnitt 4.1.3 bzw. 4.4.2 verzichtet.

Das primäre Zielkriterium stellte die Zeit bis zum Auftreten von Verhaltensauffälligkeiten im Zusammenhang mit der Demenzerkrankung dar. Hierzu wurden die Skalen NPI, NPI-D und BEHAVE-AD verwendet. Die Wahrscheinlichkeit, nach 18 Monaten noch frei von entsprechenden Symptomen zu sein, betrug für die Patienten in der Rivastigmin-Gruppe

0,622 (SEM = 0,080), in der Donepezil-Gruppe 0,484 (SEM = 0,090) und in der Galantamin-Gruppe 0,546 (SEM = 0,087). In paarweisen Vergleichen fanden sich diesbezüglich keine statistisch signifikanten Unterschiede (Rivastigmin vs. Donepezil: $p = 0,055$ mit nominellem Vorteil für Rivastigmin; Rivastigmin vs. Galantamin: $p = 0,235$; Galantamin vs. Donepezil: $p = 0,365$). Allerdings erschweren die fehlenden Angaben zur Fallzahlplanung die Interpretation dieser nicht statistisch signifikanten Ergebnisse.

Insgesamt sieben der 101 Patienten litten unter Übelkeit (Rivastigmin: 8 %, Galantamin: 6 %, Donepezil: 6 %), die damit die häufigste Nebenwirkung darstellte. Andere unerwünschte Ereignisse (Erbrechen, Kopfschmerzen, Appetitverlust und Gewichtsverlust) wurden jeweils insgesamt von einer bis drei Personen berichtet. Schwerwiegende unerwünschte Ereignisse traten im Rahmen der Studie nicht auf.

5.3.4.4 Zusammenfassung der direkten Vergleichsstudien

Eine quantitative Zusammenfassung der vergleichenden Ergebnisse zu einzelnen Zielkriterien erscheint aufgrund der begrenzten Anzahl von Studien und der teils sehr unterschiedlichen Operationalisierungen nicht angemessen. Ein eindeutiger Beleg für einen Zusatznutzen einer der drei Substanzen gegenüber den anderen lässt sich aus den Daten nicht ableiten. Allerdings verfügten nur zwei der fünf Vergleichsstudien über eine für die Detektion von moderaten Unterschieden ausreichende Fallzahl.

Bei aller gebotenen Vorsicht der Interpretation liefert eine vergleichsweise große Studie bei bloßer Betrachtung des Ergebnisses einen Hinweis auf eine Überlegenheit von Rivastigmin gegenüber Donepezil bei der Beeinflussung der Aktivitäten des täglichen Lebens. Diese Beobachtung wird durch die Ergebnisse zweier sehr kleiner Studien jedoch nicht gestützt; zudem erschwert die hohe Abbruchrate insbesondere unter Rivastigmin die Interpretation der Ergebnisse. Unter Rivastigmin zeigen sich deutlich höhere Raten beim Auftreten von unerwünschten Ereignissen, insbesondere bezüglich Übelkeit, Erbrechen, Appetitverlust und Gewichtsabnahme. Im Hinblick auf die begleitende Psychopathologie deutet eine (kleine) Studie auf einen Vorteil von Rivastigmin gegenüber Donepezil; dieses Ergebnis wird wiederum durch die große Studie nicht gestützt. Auch wenn bei der Beeinflussung der kognitiven Fähigkeiten die drei eingeschlossenen Studien relativ konsistent keine Belege für eine Überlegenheit von Rivastigmin oder Donepezil liefern, so lässt sich daraus aber auch nicht die Vergleichbarkeit der beiden Substanzen diesbezüglich ableiten, da alle Studien nicht erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudien mit einer A-priori-Definition von irrelevanten Unterschieden geplant waren. Zur krankheitsbezogenen Lebensqualität der Patienten, zur Frage der Institutionalisierung und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden Daten berichtet.

Für den Vergleich von Galantamin und Donepezil liefern die beiden einbezogenen Studien keinen Hinweis auf eine Überlegenheit eine der beiden Substanzen bezogen auf die

Beeinflussung der Aktivitäten des täglichen Lebens, der begleitenden Psychopathologie und der Kognition. Zur krankheitsbezogenen Lebensqualität der Patienten, zur Frage der Institutionalisierung und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden bzw. keine eindeutig interpretierbaren Daten berichtet.

5.3.5 Sensitivitätsanalysen

Nach der Zusammenstellung der verfügbaren Daten zeigte sich, dass die Studien, die in die Meta-Analysen eingeschlossen werden konnten, zum ganz überwiegenden Teil mit leichten Mängeln behaftet waren. Bezüglich der Zielkriterien, zu denen ausreichend viele Studien in die Meta-Analysen eingingen, um eine Sensitivitätsanalyse nach biometrischer Qualitätsbewertung sinnvoll erscheinen zu lassen (dies traf insbesondere für den Bereich der Kognition zu), zeigten die Studien homogene Ergebnisse. Insofern wurde auf eine entsprechende Sensitivitätsanalyse verzichtet. Für einige Studien standen neben den ITT-Auswertungen auch Ergebnisse einer Per-Protokoll-Auswertung zur Verfügung. Die Ergebnisse der Meta-Analysen anhand der Per-Protokoll-Auswertung zeigten zumeist etwas positivere Effekte zugunsten der Interventionen. Da die Abweichungen von den Meta-Analysen für die ITT-Daten jedoch in allen Fällen unbedeutend waren, wurde von einer Darstellung in diesem Bericht abgesehen.

5.3.6 Subgruppenauswertungen

Die Durchführung von Subgruppenanalysen – so wie im Berichtsplan vorgesehen (siehe Abschnitt 4.4.5) – erwies sich angesichts der in den Einzelpublikationen berichteten Daten im Rahmen einer Meta-Analyse mit Ausnahme von Dosis-Wirkungs-Beziehungen als nicht möglich. Im Folgenden werden deshalb ergänzend die Ergebnisse aus publizierten Subgruppenanalysen dargestellt. Zu Donepezil lag eine Meta-Analyse individueller Patientendaten zu zehn Studien vor, von denen fünf eine Dauer von 24 Wochen aufwiesen und im Rahmen der vorliegenden Nutzenbewertung ausgewertet worden sind [100]. In einer weiteren Publikation wurden die Daten aus vier Donepezil-Studien gepoolt [97]. Zu Galantamin wurden in mehreren Publikationen Daten der Studien Raskind 2000, Tariot 2000, Wilcock 2000 und Rockwood 2001^g zusammenfassend ausgewertet [94,95,101] sowie Post-hoc-Analysen einzelner Studien vorgenommen [74]. Auch zu Rivastigmin lagen sowohl mehrere gepoolte Auswertungen, in denen die Daten der Studien Corey-Bloom 1998, Rösler 1999 und B351 bzw. B304 ausgewertet wurden [91,93,96,99], als auch Re-Analysen von Einzelstudien [77,81,84,85] vor. Alle nachfolgend aufgeführten Subgruppenanalysen waren nicht erkennbar a priori geplant, so dass von einem Post-hoc-Charakter ausgegangen werden muss. Die Aussagekraft ist weiterhin dadurch eingeschränkt, dass in keiner der Arbeiten ein

^g Die Studie Rockwood 2001 war wegen zu kurzer Beobachtungsdauer (3 Monate) nicht in die Nutzenbewertung aufgenommen worden.

statistischer Interaktionstest durchgeführt wurde, sondern lediglich die beobachteten Gruppenunterschiede zwischen den jeweiligen Subgruppen einander gegenübergestellt wurden. Damit können die Ergebnisse allenfalls als hypothesengenerierende Hinweise, nicht jedoch als abschließende Belege angesehen werden. Sie müssten dafür ggf. in gezielt geplanten Studien reproduziert werden.

5.3.6.1 Geschlecht

Nur zu Donepezil wurden nach Geschlecht getrennte Auswertungen berichtet [100], wobei sich keine Unterschiede zwischen Männern und Frauen zeigten. Damit reichte die Datengrundlage nicht aus, um alle Cholinesterasehemmer bezüglich dieser Fragestellung zu bewerten.

5.3.6.2 Alter

In den vorliegenden gepoolten Analysen bzw. Post-hoc-Auswertungen wurden die Altersgruppen unterschiedlich definiert. Für Donepezil [100] und Galantamin [94] wurden die untersuchten Patienten bei einer Altersgrenze von 80 Jahren dichotomisiert, bei Rivastigmin lag die Grenze bei 75 Jahren [99]. In den Publikationen fanden sich keine Hinweise auf deutliche Unterschiede in der Effektivität zwischen den Subgruppen. Die in einer Re-Analyse der Studie zum Vergleich von Rivastigmin und Donepezil [84] gefundenen Effekte zum NPI bei jüngeren Patienten ließen sich nicht von bereits vorhandenen Unterschieden der Behandlungsgruppen zur Baseline trennen.

5.3.6.3 Schweregrade der Demenz (leicht und mittelschwer)

In der gepoolten Analyse zu Donepezil [100] wurde berichtet, dass ein Vorteil im ADAS-cog für Donepezil (5 mg oder 10 mg) gegenüber Placebo unabhängig vom Schweregrad gefunden wurde. Allerdings wurden hierzu keine Daten präsentiert. In mehreren Auswertungen zu Galantamin [74,95,101] wurden – mit jeweils unterschiedlichen Operationalisierungen – Subgruppen nach dem Schweregrad der Demenz gebildet. Es ergaben sich Hinweise darauf, dass bei schwerer beeinträchtigten Personen der Behandlungseffekt deutlicher war als bei weniger schwer beeinträchtigten Personen. Auch für Rivastigmin zeigten sich ähnliche Hinweise [93,96].

5.3.6.4 Dosierungen der Cholinesterasehemmer

Bei Donepezil zeigte sich in der vorliegenden Nutzenbewertung eine stärkere Wirksamkeit bezüglich der Kognition bei einer täglichen Dosis von 10 mg als bei einer Dosis von 5 mg oder einer flexiblen Dosierung (5 bis 10 mg). Diese Dosis-Wirkungs-Beziehung wurde durch eine Meta-Analyse von individuellen Patientendaten bestätigt [100]. Auch für das Auftreten von unerwünschten Ereignissen zeigte sich ein Zusammenhang mit der Höhe der Dosierung [97] (vgl. auch Abschnitt 5.3.1.7). Für die anderen Therapieziele können keine Aussagen zu

dosisabhängigen Effektivitätsunterschieden abgeleitet werden, da hierzu keine aggregierbaren Daten aus Studien mit unterschiedlicher Dosierung vorliegen. Galantamin war in Dosierungen von 16 und 24 mg wie oben beschrieben vergleichbar wirksam, die Dosierung von 8 mg war nicht wirksam. Rivastigmin zeigte in der vorliegenden Nutzenbewertung in höherer Dosierung (6–12 mg) größere Effekte als in niedriger Dosierung (1–4 mg). Diese Dosis-Wirkungs-Beziehung bestätigte eine gepoolte Analyse [91]. Auch bei der Häufigkeit von unerwünschten Ereignissen zeigte sich ein Unterschied zwischen hoher und niedriger Dosierung.

5.3.6.5 Vorliegen unterschiedlicher Begleiterkrankungen

Patienten in allen Acetylcholinesterasehemmer-Studien sollten nach den Ein- und Ausschlusskriterien keine aktive Begleiterkrankung besitzen, die ein Verbleiben in der Studie gefährden könnte. Einzeln explizit aufgezählt wurden in einigen Studien als nicht erlaubte Begleiterkrankungen bzw. Therapien beispielsweise Asthma/COPD, aktive Magengeschwüre, Insulingebrauch, Miktionsstörungen und Erkrankungen, die eine kardiale Synkope begünstigen. Aufgrund des ähnlichen Wirkmechanismus und der fehlenden Vergleichsstudien bezüglich qualitativer Interaktionsunterschiede konnte eine Differentialindikation aufgrund der Komorbidität nicht abgeleitet werden.

In einigen publizierten Subgruppenanalysen wurden die Behandlungseffekte bei Vorliegen von Risikofaktoren betrachtet, die auch auf eine Mischform der Demenz hinweisen können. In einer Re-Analyse von Rivastigmin-Studien zeigten sich zwischen Patienten mit bzw. ohne vaskuläre Risikofaktoren (operationalisiert anhand der modifizierten Hachinski-Ischämie-Skala) [77] sowie zwischen Patienten mit und ohne Bluthochdruck [81] keine deutlichen Unterschiede hinsichtlich der Behandlungseffekte. Auch die Post-hoc-Analyse der Studie zum Vergleich von Rivastigmin und Donepezil hinsichtlich Anzeichen einer Erkrankung vom Lewy-Body-Typ [85] ergab keine deutlichen Nachweise für eine differentielle Wirksamkeit.

5.4 Zusammenfassung

In den vorliegenden Bericht wurden zum Vergleich der Cholinesterasehemmer Donepezil, Galantamin und Rivastigmin mit Placebo sowie untereinander 27 Studien mit insgesamt 9883 randomisierten Patienten eingeschlossen (Tabelle 43). Studien zum Vergleich mit anderen in Deutschland zugelassenen und verfügbaren medikamentösen und nichtmedikamentösen Interventionen konnten nicht identifiziert werden.

Tabelle 43: Anzahl eingeschlossener Studien und randomisierter Patienten

Substanz	Anzahl Studien	Randomisierte Patienten	
Donepezil	12	5 mg	561
		(5-) 10 mg	347
		10 mg	821
		Placebo	1276
		Σ	3005
Galantamin	6	8 mg	140
		16 mg	279
		16-24 mg	711
		24 mg	893
		Placebo	1201
Σ	3224		
Rivastigmin	4	1-4 mg	476
		2-12 mg	456
		6-12 mg	564
		Placebo	720
Σ	2216		
Direkte Vergleichsstudien			
Galantamin vs. Donepezil	1	188	
Rivastigmin vs. Donepezil	3	1149	
Rivastigmin vs. Donepezil vs. Galantamin	1	101	
Σ	27	9883	

Die Berichtsqualität der Studien muss überwiegend als mäßig, teils auch als mangelhaft angesehen werden. Zu dieser Einschätzung führen insbesondere wenig transparent dargestellte Patientenflüsse und eine kaum zu beobachtende konsequente Umsetzung des Intention-to-treat-Prinzips. In einigen Studien beträgt der Anteil der trotz Randomisierung nicht ausgewerteten Patienten für die primären Zielkriterien (gerundet) mindestens 11 %, was

sicher als relevante Abweichung vom ITT-Prinzip konstatiert werden muss. Für manche betrachteten Endpunkte, die in den Studien als sekundäre Zielgrößen mitgeführt wurden, ist dieser Anteil noch höher. Vier der fünf Studien zum Vergleich der verschiedenen Substanzen untereinander waren als nicht doppelblind angelegt, was per se in diesem Bereich als grober Designmangel angesehen werden muss.

Bis auf zwei Studien (beide zu Donepezil, Dauer jeweils etwa ein Jahr) waren alle Vergleiche mit Placebo lediglich auf einen Behandlungs- bzw. Beobachtungszeitraum von maximal 26 Wochen ausgerichtet. Auch wenn die länger währenden Studien keine grundsätzlich anderen Ergebnisse zeigten, lassen sich belastbare Aussagen im Wesentlichen nur für den 6-Monats-Zeitraum treffen. Immerhin drei der fünf Studien zum Vergleich von verschiedenen Substanzen untereinander hatten dem gegenüber eine Dauer von einem Jahr oder länger. Dafür litten diese Studien wiederum – bis auf eine zu Donepezil und Rivastigmin – neben der eingeschränkten Validität aufgrund eines unverblindeten Designs an zu geringen Fallzahlen, um Unterschiede entdecken, aber auch um die Vergleichbarkeit demonstrieren zu können.

Vergleiche mit Placebo

Die Tabelle 44 fasst die Ergebnisse der Studien zu den Vergleichen von Donepezil, Galantamin und Rivastigmin mit Placebo zusammen. Bei allen Studien ist ein dosisabhängiger Effekt zu beobachten, wobei im Niedrigdosisbereich bei Galantamin und Rivastigmin im Gegensatz zu Donepezil jeweils keine bzw. nur eine unsichere Wirksamkeit sichtbar ist. Bei Galantamin besteht kein erkennbarer Unterschied zwischen einer Dosierung von 16 mg und 24 mg. Bezüglich der in den Studien berichteten Häufigkeiten von unerwünschten Ereignissen wird diese Dosis-Wirkungs-Beziehung bestätigt.

Für das Therapieziel **Aktivitäten des täglichen Lebens** zeigen sich für alle drei Substanzen bei mittlerer und/oder hoher Dosis Hinweise auf eine günstige Beeinflussung. Die anhand von Meta-Analysen gefundenen durchschnittlichen Effekte betragen etwa drei Scorepunkte auf der DAD- bzw. PDS-Skala für Galantamin bzw. Rivastigmin. Entsprechende Schätzwerte für Donepezil lassen sich nicht ausreichend sicher ableiten, da bei der diesbezüglichen Meta-Analyse von einer Überschätzung des Effekts ausgegangen werden muss. Nichtsdestotrotz kann auch für Donepezil ein Hinweis auf eine günstige Beeinflussung angenommen werden.

Im Hinblick auf die **begleitende Psychopathologie** lassen sich keine Hinweise auf eine günstige oder ungünstige Beeinflussung durch Donepezil und Rivastigmin ableiten; bei Donepezil aufgrund nicht überzeugender Daten, bei Rivastigmin wegen praktisch nicht vorhandener Daten. Bei Galantamin hingegen zeigt sich ein Hinweis auf einen positiven Effekt, der allerdings mit etwa ein bis zwei Scorepunkten auf der NPI-Skala als äußerst gering einzustufen ist.

Tabelle 44: Zusammenfassung der Ergebnisse zu den Therapiezielen aus Placebo-kontrollierten Studien

Therapieziel	Donepezil	Galantamin	Rivastigmin
Patientenrelevante Therapieziele			
Aktivitäten des täglichen Lebens	↑	↑	↑
Psychopathologische Symptome	↔	↑	keine Daten
Kognitive Leistungsfähigkeit	↑↑	↑↑	↑↑
Krankheitsbezogene Lebensqualität	↔	keine Daten	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine Daten	keine Daten	keine Daten
Mortalität	(↔)	(↔)	(↔)
Unerwünschte Ereignisse	↓↓↓	↓↓↓	↓↓↓
Angehörigenrelevante Therapieziele			
Lebensqualität der (betreuenden) Angehörigen	↔	↑	keine bzw. unsichere Daten
Betreuungsaufwand	↔	↑	keine Daten
Ergänzende Information			
Klinisches Krankheitsstadium	↑↑	↑↑	↑↑
Dosis-Wirkungs-Beziehung			
	geringere Effektivität (Kognition) und weniger Nebenwirkungen für niedrige (5 mg) bzw. flexible Dosierung	kein günstiger Effekt und nicht konsistent mehr Nebenwirkungen bei 8 mg, sonst keine Unterschiede	unsicherer Effekt bei 1–4 mg
Erläuterung: ↑↑, ↓↓ = Beleg für günstigen bzw. ungünstigen Effekt. ↑, ↓ = Hinweis für einen günstigen bzw. ungünstigen Effekt ↔ = kein Hinweis auf Unterschied () = wenig Daten vorhanden			

Für die drei Substanzen kann der Nutzen bezüglich einer im Vergleich zu Placebo günstigen Beeinflussung der **Kognition** reklamiert werden. Dieser beträgt auf der ADAS-cog-Skala ca.

zwei (Donepezil 5 mg bzw. mit flexibler Dosierung) bis drei Scorepunkte (Donepezil 10 mg, Galantamin, Rivastigmin).

Zum Therapieziel **krankheitsbezogene Lebensqualität** liegen nur für Donepezil Daten aus zwei Studien vor, die keinen eindeutigen Hinweis auf einen günstigen oder ungünstigen Effekt liefern. Für Galantamin und Rivastigmin wurden dazu in den Studien überhaupt keine Daten berichtet.

Keinerlei (interpretierbare) Daten fanden sich auch zum Therapieziel **Notwendigkeit einer vollstationären Pflege** (Institutionalisierung).

Bei insgesamt in den Studien nur wenig berichteten Todesfällen lassen sich keine Hinweise auf eine günstige oder ungünstige Beeinflussung der **Mortalität** ableiten.

Bei allen Substanzen sind im höheren Dosisbereich größere Abbruchraten aufgrund **unerwünschter Ereignisse** zu verzeichnen. Es treten zudem häufiger solche unerwünschten Ereignisse auf, die im Einklang mit dem Wirkprinzip der Acetylcholinesterasehemmer stehen (z. B. Übelkeit, Erbrechen, Durchfall). Es findet sich kein Hinweis auf einen größeren Anteil von Patienten mit schwerwiegenden unerwünschten Ereignissen im Vergleich zu Placebo, wobei einschränkend bemerkt werden muss, dass die diesbezügliche Berichterstattung teilweise unzureichend war. Über eher selten bzw. erst nach längerer Einnahmedauer auftretende unerwünschte Ereignisse können aufgrund der Anlage der Studien und der Methodik der Berichterstattung kaum Aussagen getroffen werden.

Aus den vorliegenden Ergebnissen lassen sich für Donepezil keine Hinweise auf eine günstige oder ungünstige Beeinflussung der **Lebensqualität der betreuenden Angehörigen** erkennen. Bei Galantamin zeigt sich ein Hinweis auf einen – mit einer Größenordnung von einem Zehntel Standardabweichung allerdings als äußerst gering einzustufenden – positiven Effekt. Für Rivastigmin wurden dazu keinerlei Daten gefunden.

Es liegen Hinweise vor, dass zu Rivastigmin Daten bezüglich des Therapieziels **Betreuungsaufwand** in allen vier größeren Phase-III-Studien erhoben wurden. Diese Daten wurden allerdings bislang nicht publiziert, so dass hierzu keine Aussagen getroffen werden können. Die Daten für Donepezil sind aus vorwiegend methodischen Gründen nur wenig belastbar, so dass sich daraus ebenfalls kein Hinweis auf eine günstige Beeinflussung der Betreuungszeiten ergibt. Für Galantamin liegt aus einer Studie ein Hinweis auf einen diesbezüglich positiven Effekt vor.

Der **klinische Gesamteindruck** wird konsistent durch alle Substanzen gebessert.

Für Galantamin und Rivastigmin liegen Hinweise vor, dass bei schwerer beeinträchtigten Personen der Behandlungseffekt größer ist als bei weniger schwer beeinträchtigten Personen.

Bezüglich Alter, Geschlecht und begleitende Komorbidität lassen sich keine entsprechend differenzierten Aussagen treffen.

Vergleiche untereinander

Die Tabelle 45 fasst die Ergebnisse der Studien zu den Vergleichen der verschiedenen Cholinesterasehemmer untereinander zusammen.

Tabelle 45: Zusammenfassung der Ergebnisse zu den Therapiezielen aus Vergleichsstudien von Cholinesterasehemmern untereinander

Therapieziel	DON vs. GAL	DON vs. RIV	GAL vs. RIV
Patientenrelevante Therapieziele			
Aktivitäten des täglichen Lebens	(↔)	(↓) ^a	keine Daten
Psychopathologische Symptome	(↔)	↔	(↔)
Kognitive Leistungsfähigkeit	(↔)	↔	keine Daten
Krankheitsbezogene Lebensqualität	keine Daten	keine Daten	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine bzw. unsichere Daten	keine Daten	keine Daten
Mortalität	(↔)	↔	keine Daten
Unerwünschte Ereignisse	(↔)	↑↑	(↔)
Angehörigenrelevante Therapieziele			
Lebensqualität der (betreuenden) Angehörigen	keine Daten	keine Daten	keine Daten
Betreuungsaufwand	keine Daten	keine Daten	keine Daten
Ergänzende Information			
Klinisches Krankheitsstadium	keine Daten	keine Daten	keine Daten
Anmerkungen	in größerer Studie evtl. ungünstigere Dosierung für DON	evtl. ungünstigere Dosierung für DON	
Erläuterung: ↑↑, ↓↓ = Beleg für günstigen bzw. ungünstigen Effekt für DON vs. RIV ↑, ↓ = Hinweis für einen günstigen bzw. ungünstigen Effekt für DON vs. RIV ↔ = kein Hinweis auf Unterschied () = wenig Daten vorhanden DON = Donepezil, GAL = Galantamin, RIV = Rivastigmin a: Ergebnisse durch hohe Abbruchraten belastet			

Eine quantitative Zusammenfassung (Meta-Analyse) der vergleichenden Ergebnisse zu einzelnen Zielkriterien war aufgrund der begrenzten Anzahl von Studien und der teils sehr unterschiedlichen Operationalisierungen nicht angemessen. Nur zwei der fünf Studien verfügten über eine für die Detektion von moderaten Unterschieden ausreichende Fallzahl.

Für den Vergleich von Donepezil und Galantamin liefern die beiden einbezogenen Studien keinen eindeutigen Hinweis auf eine Überlegenheit einer der beiden Substanzen bezogen auf die Beeinflussung der Aktivitäten des täglichen Lebens, der begleitenden Psychopathologie, der Kognition und der therapiebezogenen unerwünschten Ereignisse. Zur krankheitsbezogenen Lebensqualität der Patienten, zur Frage der Institutionalisierung und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden bzw. keine eindeutig interpretierbaren Daten berichtet.

Beim Vergleich von Donepezil mit Rivastigmin weisen die Daten aus einer Studie bezüglich der Beeinflussung der Aktivitäten des täglichen Lebens zwar auf eine geringfügige Überlegenheit von Rivastigmin hin (Effektschätzer in der Größenordnung von einem Zehntel Standardabweichung), allerdings bestehen aus methodischen Gründen Zweifel an der Validität dieser Daten. Im Hinblick auf die begleitende Psychopathologie, die Kognition und die Mortalität gibt es keinen eindeutigen Hinweis auf einen Unterschied zwischen den beiden Substanzen. Unter Rivastigmin zeigen sich deutlich höhere Raten beim Auftreten von unerwünschten Ereignissen, insbesondere bezüglich Übelkeit, Erbrechen, Appetitverlust und Gewichtsabnahme. Zur krankheitsbezogenen Lebensqualität der Patienten und zu den angehörigerelevanten Zielkriterien wurden keine vergleichenden Daten berichtet.

Zum Vergleich von Galantamin und Rivastigmin liegen lediglich Ergebnisse aus einem Drei-Gruppen-Vergleich mit sehr niedrigen Fallzahlen vor. Hier sind keine Unterschiede bezüglich der Beeinflussung psychopathologischer Befunde und des Auftretens unerwünschter Ereignisse zu erkennen. Zu den anderen Therapiezielen liegen keine Daten vor.

Insgesamt lässt sich aus den nicht vorhandenen bzw. allenfalls geringfügigen (und mit einiger Unsicherheit behafteten) Unterschieden bei den Wirksamkeitsparametern in den Vergleichsstudien kein Nachweis einer Überlegenheit einer der Wirkstoffe den anderen gegenüber ableiten. Die Ergebnisse können allerdings auch nicht im Sinne einer Vergleichbarkeit der Substanzen interpretiert werden, da alle Studien nicht erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudien mit einer A-priori-Definition von irrelevanten Unterschieden geplant waren.

Anhand der vorliegenden Daten lassen sich kaum Aussagen zu einzelnen Subgruppen von Patienten machen. Auffällig war, dass bei fast allen Studien Patienten mit klinisch schwerwiegenden bzw. nicht gut eingestellten internistischen Erkrankungen ausgeschlossen wurden, so dass unklar ist, ob die Ergebnisse generell auf diese Patientenpopulation übertragbar sind.

6 DISKUSSION

Die drei zugelassenen Acetylcholinesterasehemmer Donepezil, Galantamin und Rivastigmin führen bei ihrem Einsatz dosisabhängig zu den in ihrem Wirkprinzip begründeten cholinergen gastrointestinalen Nebenwirkungen, die insbesondere, aber nicht nur ausschließlich in der Eindosierungsphase häufig sind. Davon sind je nach Substanz und Nebenwirkung 5 bis 10 % bei Donepezil, 5 bis 20 % bei Galantamin und 10 bis 40 % der Patienten bei Rivastigmin betroffen.

Dem gegenüber steht ein konsistenter Nutzen gegenüber Placebo im Bereich der Kognition mit einer Effektstärke von ca. 0,5 Standardabweichungen, was auf der ADAS-cog-Skala umgerechnet etwa drei Scorepunkten entspricht. Für weitere gut untersuchte Bereiche eines möglichen Nutzens zeigen sich geringe (Aktivitäten des täglichen Lebens) oder geringe und nicht für alle Cholinesterasehemmer konsistent nachweisbare Effekte (psychopathologische Begleitsymptome und Betreuungsaufwand für Galantamin). Darüber hinaus fanden sich für die übrigen im Berichtsplan definierten Bereiche eines patientenorientierten Nutzens keine oder keine verwertbaren Daten. Dass der klinische Gesamteindruck konsistent durch alle Substanzen gebessert wird, dürfte in der Gesamtschau folglich vor allem auf die Effekte bezüglich der Kognition zurückzuführen sein.

Das Ergebnis dieses Berichts wird im Wesentlichen durch die aktuellen systematischen Übersichten und HTA-Berichte auf diesem Gebiet [49,108-113] gestützt. Unterschiede bestehen allenfalls in der quantitativen Einordnung der gefundenen Hinweise auf eine günstige Beeinflussung der alltagspraktischen Fähigkeiten und der psychopathologischen Begleitsymptomatik.

Lediglich eine systematische Übersicht von Kaduszkiewicz et al. aus dem Jahr 2005 [114,115] kommt zu einer deutlich kritischeren Bewertung. Als Begründung dafür werden erhebliche methodische Mängel der Primärstudien angesichts nur geringer dort beobachteter Effekte angeführt. Wenngleich auch der vorliegende Bericht in einigen Fällen gravierende Mängel bei der Studien- und Berichtsqualität identifiziert und damit die Validität einzelner Studienergebnisse infrage gestellt hat, erscheint eine allein daraus resultierende Negativbewertung nicht angemessen, zumal die Wertigkeit einzelner methodischer Kritikpunkte von Kaduszkiewicz et al. bezweifelt werden kann [116]^h.

Problematisch ist die Einordnung der klinischen Relevanz der Ergebnisse zur Wirksamkeit, besonders in der Abwägung zu den vorhandenen, teilweise nicht unerheblichen Nebenwirkungen. Was bedeutet eine Mittelwertsdifferenz von drei Scorepunkten bei Veränderungen einer Skala? Ein häufig bemühtes Argument in diesem Zusammenhang

^h vergleiche hierzu auch eine Entgegnung der Autoren [117]

betrifft den Skalenbereich: Da der Skalenbereich des ADAS-cog beispielsweise 70 Punkte umfasse, sei eine Differenz von drei Punkten vermutlich nicht relevant. Eine solche Argumentation greift zu kurz, da in diesem Zusammenhang nicht die Spannweite – also die theoretisch denkbaren Ausprägungsmöglichkeiten –, sondern der Bereich, der tatsächlich ausgeschöpft wird – also letztlich die Variabilität (bei der zu betrachtenden Population) – von Interesse istⁱ.

Ebenfalls zumeist nicht hilfreich ist der Verweis auf einen so genannten „minimalen klinisch bedeutsamen Unterschied“ (minimal clinically important difference, MCID) nach Vorgaben durch Zulassungsbehörden oder aus der Literatur. Häufig werden solche MCID nämlich als individuelles Responsekriterium und nicht als Irrelevanzbereich für Gruppenunterschiede definiert. Für den ADAS-cog z. B. wurde von der FDA ein MCID von vier Scorepunkten aufgrund einer Expertenbefragung vorgeschlagen (zitiert nach [70,75,114]), so dass ggf. bei einer solchen individuellen Verbesserung auf dieser Skala von einem Nutzen für den einzelnen Patienten gesprochen werden kann. Unter der vereinfachenden Annahme, dass sich die Veränderungen im ADAS-cog durch eine Normalverteilung approximieren lassen, und unter Zugrundelegung eines Variabilitätsbereichs für diese Änderungen von etwa sechs Scorepunkten (einfache Standardabweichung, im Median etwa in den vorliegenden Studien beobachtet) sowie Annahmen über die mittleren Veränderungen in der Placebo-Gruppe lässt sich der Vergleich der Mittelwerte in einen Vergleich von (binären) Responderaten umrechnen (s. Tabelle 46). Vereinzelt wurden in den Publikationen zu den Studien Ergebnisse von Responderanalysen auf Grundlage der empirischen Daten berichtet (Donepezil: Rogers 1998; Galantamin: Erkinjuntti 2002, Raskind 2000, Tariot 2000, Wilcock 2000; Rivastigmin: B304, Rösler 1999). Die Ergebnisse dieser empirischen Analysen passen bis auf diejenigen der Studie von Rogers 1998 sehr gut mit den theoretisch abgeleiteten überein.

Bei einer derartigen Ergebnisdarstellung ergeben sich Differenzen von Response- (Erfolgs-) Raten zwischen ca. 13 und 16 Prozentpunkten (Verbesserung im ADAS-cog um vier Punkte). Die Unterschiede für die Beeinflussung der alltagspraktischen Fähigkeiten und – soweit überhaupt erkennbar – der psychopathologischen Symptome sind allerdings deutlich geringer. Bei einer optimistischen Schätzung einer standardisierten Mittelwertsdifferenz von etwa einem Viertel Standardabweichung bei den Aktivitäten des täglichen Lebens kann hier unter Zugrundelegung ähnlicher Responsekriterien wie für den ADAS-cog (zwei Drittel Standardabweichung) eine Differenz der Responderaten von ca. sieben bis acht Prozentpunkten angenommen werden.

ⁱ So wird man vermutlich auch nicht eine Mittelwertsdifferenz von 10 kg Körpergewicht beim Vergleich von zwei Interventionen zur Gewichtsreduktion vernachlässigen, nur weil das Gewicht von (erwachsenen) Menschen im Extremfall zwischen 35 und 350 kg schwanken kann.

Tabelle 46: Geschätzter und beobachteter Anteil Responder für den ADAS-cog bei Responsekriterium 4 Scorepunkte

Gruppe	Änderung	Anteil Responder (%)	Änderung	Anteil Responder (%)	Studie
	angenommen	geschätzt	beobachtet	beobachtet	
Verum	-1,5	34	-1,1	54	Rogers 1998 (DON, 10 mg)
Placebo	+1,5	18	+1,8	27	
			-0,8	35	Erkinjuntti 2002 (GAL, 24 mg)
			+1,7	22	
			-1,9	34	Raskind 2000 (GAL, 24 mg)
			+2,0	17	
			-1,4	37	Tariot 2000 (GAL, 24 mg)
			+1,7	20	
Verum	0,0	25	-0,5	29	Wilcock 2000 (GAL, 24 mg)
Placebo	+3,0	12	+2,4	15	
			-0,3	24	Rösler 1999 (RIV, 6-12 mg)
			+1,3	16	
			-0,2	23	B304 1998 (RIV, 2-12 mg)
			+2,8	13	
Geschätzter Anteil Responder über Normalverteilungsannahme und bei Standardabweichung von 6 Scorepunkten DON = Donepezil, GAL = Galantamin, RIV = Rivastigmin					

Auch wenn eine solche Darstellungsweise möglicherweise leichter interpretierbar erscheint, bleibt doch die Frage, was ein Response, also die Veränderung in einem abstrakten Score um einen bestimmten Betrag, nun unmittelbar und spürbar für den einzelnen Patienten bedeutet. Eine Antwort auf diese Frage steht letztlich noch aus und wird auf der „Score-Ebene“ allein vermutlich kaum zu erzielen sein. Insofern könnten Überlegungen hilfreich sein, sich der klinischen Relevanz der Effekte von Antidementiva anhand anderer Zielkriterien zu nähern.

Die im vorliegenden Bericht aktuellste Studie von Rockwood et al. [71] zu Galantamin verwendete den Grad der Erreichung von individuell festgelegten Zielen, gemessen mit der Goal Attainment Scale (GAS), als primäres Zielkriterium. Hierzu wurden zum einen von Behandlern, zum anderen von den Patienten oder Angehörigen anhand dieser Skala Ziele in den Bereichen Kognition, Funktion, Verhalten, Freizeit und soziale Aktivitäten definiert und später die Veränderung im Vergleich zum Studienbeginn beurteilt. Damit stellt die GAS ein individualisiertes Ergebnismaß dar, das – je nachdem, welche Ziele definiert werden – unterschiedliche Lebensbereiche abdeckt. In der ITT-Analyse ergab sich nach vier Monaten

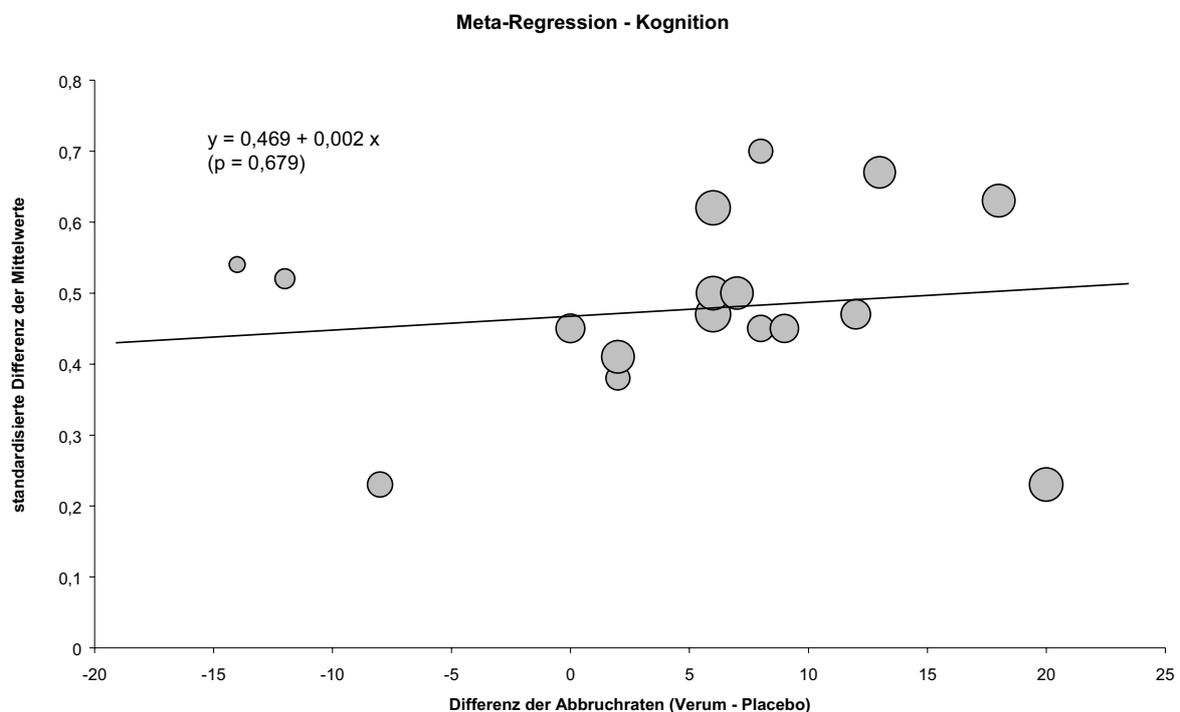
für die Galantamin-Gruppe sowohl nach der Beurteilung der Patienten/Angehörigen als auch der Behandler ein erhöhter GAS-Wert (Effektstärke [standardisierte Mittelwertsdifferenz]: 0,20 [Patienten/Angehörige] bzw. 0,45 [Behandler] Standardabweichungen), der Unterschied zu Placebo war jedoch nur für die von den Behandlern aufgestellten Ziele statistisch signifikant. Die Ergebnisse zur Veränderung bei der DAD- (Aktivitäten des täglichen Lebens) und ADAS-cog- Skala entsprachen in etwa denen der anderen Studien zu Galantamin.

Wenngleich auch hier eine andere Ergebnispräsentation wünschenswert gewesen wäre, z. B. mit Angabe des Anteils von Patienten, die (evtl. differenziert nach unterschiedlichen Bereichen) ihre Ziele erreicht hatten, so nähert sich ein solches Zielkriterium der Interpretierbarkeit im Hinblick auf die unmittelbaren und spürbaren Belange der Patienten. Vorteilhaft scheint weiterhin, dass auch negative Konsequenzen einer Intervention (Nebenwirkungen) darin einfließen können und somit eine Nutzen-Schaden-Abwägung erleichtert wird. Schließlich böte sich bei einer dichotomisierten Ergebnispräsentation (Anteil mit/ohne Zielerreichung) eine einfache Möglichkeit, fehlende Werte im Rahmen von Sensitivitätsanalysen interpretierbar zu modellieren.

Der letzte Aspekt (Umgang mit fehlenden Werten in der Auswertung) ist bei den Studien zu den Cholinesterasehemmern von besonderer Bedeutung, da zum einen kaum eine konsequente Umsetzung des ITT-Prinzips zu beobachten war und zum anderen die Ersetzungsstrategie anhand der LOCF-Methode (Fortschreiben des zuletzt erhobenen Wertes) immer wieder Anlass zu methodischer Kritik ist [114]. Diese stützt sich auf die Annahme, dass bei einer sich progredient verschlechternden Erkrankung ein frühzeitiges Ausscheiden der Patienten (z. B. aufgrund von Nebenwirkungen) durch das Fortschreiben der zu diesem Zeitpunkt noch günstigen Ergebnisse eine Verzerrung induzieren könnte. Wenngleich diese Annahme plausibel ist, so sollte sie doch – insbesondere bei dem vergleichsweise reichen Erkenntnismaterial – durch empirische Daten gestützt werden, bevor sie zu einer Abwertung der gefundenen Ergebnisse führt.

In den Placebo-kontrollierten Studien, die in den vorliegenden Bericht eingeschlossen werden konnten, wurden unter den Verum-Medikationen im Median Abbruchraten von 22 Prozentpunkten (Spannweite 0-35 %, ohne die Studie von Tune 2003: 14-35 %) beobachtet, in den Placebo-Gruppen im Median von 17 Prozentpunkten (11-33 %). Etwa die Hälfte der Abbrüche unter Verum ging auf unerwünschte Ereignisse zurück, unter Placebo betrug dieser Anteil im Median 44 %. Auch wenn nicht bei allen Patienten mit Abbruch der Medikation bei Studienende keine Werte für die Zielgrößen erhoben worden waren, stellt somit das Ersetzen fehlender Werte (anhand der LOCF-Methode) zweifelsohne ein Problem dar, zumal in den Studien unklar blieb, ob bei einem Abbruch trotz Vorliegen eines Wertes bei Studienende („retrieved drop outs“, RDO) nicht dennoch der Wert zum Zeitpunkt des Abbruchs fortgeschrieben wurde.

Die in den vorliegenden Studien beobachteten Therapieeffekte lassen aber beispielsweise bei der Zielgröße Kognition keinen Zusammenhang zu den Abbruchraten erkennen, auch nicht zu den Unterschieden in den Abbruchraten zwischen Verum und Placebo (Abbildung 48). In einer nachträglichen Auswertung von drei Rivastigmin-Studien (Corey-Bloom 1998, Rösler 1999 sowie die in den vorliegenden Bericht nicht aufgenommene Studie B351) untersuchten Farlow et al. [118] die Untergruppe von RDO-Patienten. Dabei zeigten Abbrecher im Verlauf eine Verschlechterung der Kognition (ADAS-cog, Differenz 26 Wochen minus Baseline). Die Verschlechterung war allerdings unter Placebo stärker ausgeprägt als unter Rivastigmin, so dass bei diesen Patienten die LOCF-Methode ggf. zu einer konservativen Verzerrung geführt hätte. Einschränkend ist hierbei anzumerken, dass nur knapp ein Drittel aller Abbrecher RDO-Patienten waren, so dass etwa zwei Drittel nicht in diese Auswertung eingehen konnten, und dass nicht der ADAS-cog-Wert zum Zeitpunkt des Abbruchs betrachtet wurde.



Aus mehrarmigen Studien wurde jeweils die Verum-Gruppe mit der höchsten Dosierung gewählt. Abbildung 48: Zusammenhang zwischen in den Placebo-kontrollierten Studien beobachtetem Effekt auf die Kognition und Differenz der Abbruchraten von Verum und Placebo.

Die theoretische Kritik an der Methodik zum Ersetzen fehlender Werte bei den in den vorliegenden Bericht eingeschlossenen Studien wird also durch empirische Daten nicht gestützt. Dennoch wären Sensitivitätsanalysen mittels anderer Ersetzungsstrategien [119] wünschenswert, um die Robustheit der Ergebnisse besser einschätzen zu können.

Im Übrigen können die bei einem größeren Anteil von Patienten zu beobachtenden Nebenwirkungen der Cholinesterasehemmer einer anderen Argumentation dienen: Sofern

solche Nebenwirkungen nicht zum Abbruch der Therapie führen, kann spekuliert werden, dass sie einen ungünstigen Einfluss insbesondere auf das Therapieziel der alltagspraktischen Fähigkeiten ausüben, so dass das Nutzenpotenzial bei weniger stark ausgeprägten Nebenwirkungen größer sein könnte. Auch diese Spekulation wird kaum durch die empirischen Daten aus dem vorliegenden Bericht gestützt, allerdings wiederum auch nicht in umgekehrter Richtung, dass also höhere Abbruchraten (generell oder unter Verum im Vergleich zu Placebo) mit größeren Therapieeffekten einhergehen.

Die aus den Mitteln des englischen NHS unterstützte AD2000-Studie zur Wertigkeit von Donepezil in einem möglichst anwendungsnahen Setting, deren primäres Zielkriterium die Zeit bis zu einer Heimaufnahme bzw. bis zu einem definierten Grad des Verlustes von alltagspraktischen Fähigkeiten war, konnte nicht in die vorliegende Nutzenbewertung aufgenommen werden. Die in der Publikation präsentierten Daten bzw. Auswertungen zu diesen beiden, im Prinzip hoch relevanten, primären Zielkriterien sind aus methodischen Gründen nicht interpretierbar. Diese Einschätzung wird von einigen anderen systematischen Übersichten geteilt [49,114] und ist kürzlich auch anderenorts ausführlich begründet worden [120,121]. Mehrfache Anfragen an die Autoren bezüglich dieser Vorbehalte blieben unbeantwortet.

In Gesamtschau der potenziell nützlichen und schädlichen Effekte der Cholinesterasehemmer wäre es für die Zukunft wünschenswert, gezielter solche Patienten zu identifizieren, die einen größeren Nutzen erfahren, und solche, die von einer Behandlung mit Cholinesterasehemmern nicht profitieren. Der aus nachträglichen Subgruppenanalysen abgeleitete Hinweis, dass ggf. schwerer an Demenz erkrankte Patienten mehr profitieren als weniger schwer erkrankte, sollte in zukünftigen Studien verifiziert werden. Möglicherweise lassen sich solche Ergebnisse durch einen Unterschied in der Veränderungssensitivität der eingesetzten Skalen in Abhängigkeit vom Schweregrad der Patienten erklären. Eine Begrenzung der Indikation für die Cholinesterasehemmer auf bestimmte Schweregrade lässt sich somit aus den vorliegenden Daten gegenwärtig nicht solide ableiten. Darüber hinaus wären Studien mit einer längeren Laufzeit zu begrüßen, und es sollten Zielkriterien eingesetzt werden, die den Nutzen für die Patienten unmittelbar und spürbar abbilden.

Schriftliche Stellungnahmen und wissenschaftliche Erörterung

Im Rahmen der schriftlichen Anhörung zum Vorbericht der vorliegenden Nutzenbewertung gingen insgesamt 17 substantielle Stellungnahmen ein (siehe Anhang I). Die Stellungnehmenden wurden zur Diskussion unklarer Aspekte in den Stellungnahmen zu einer mündlichen wissenschaftlichen Erörterung eingeladen, wobei für alle Stellungnahmen Vertreter anwesend waren (siehe Anhang H).

In den Stellungnahmen wurden 180 Referenzen zitiert. Dabei wurde auf zwei randomisierte, kontrollierte Studien sowie eine Meta-Analyse hingewiesen, die bisher im Bericht nicht

berücksichtigt gewesen seien (vgl. SN Gutzmann, SN Eisai, SN Pfizer, SN Wille). In der Studie Holmes 2004 [122] betrug die kontrollierte Behandlungsphase lediglich 12 Wochen. Die Studie Winblad 2006 [123] untersuchte ausschließlich Patienten mit schwerer Alzheimer Demenz (MMSE 1 bis 10). Damit erfüllten beide Arbeiten nicht die in dem zugrunde liegenden Berichtsplan definierten Ein- und Ausschlusskriterien für eine Berücksichtigung im Rahmen der vorliegenden Bewertung. Die Meta-Analyse Trinh 2003 [124] war bereits vor Veröffentlichung des Vorberichts auf weitere relevante Studien gescreent worden (vgl. Anhang C: Liste der gescreenten systematischen Übersichten).

In den Stellungnahmen kristallisierten sich sechs zentrale Diskussionspunkte heraus, die während der mündlichen wissenschaftlichen Erörterung ausführlich zur Sprache kamen. Diese Punkte werden im Folgenden dargestellt und diskutiert.

- Studienauswahl (Studiendauer, Studiendesign)

In mehreren Stellungnahmen wurde das gewählte Einschlusskriterium bezüglich der Studiendauer kritisiert. Von einigen Stellungnehmenden wurde angemerkt, dass auch kürzere Studien (≥ 12 Wochen) hätten eingeschlossen werden sollen (vgl. Anhang I, z. B. SN Eisai, SN Grass-Kapanke, SN Maier), von anderen, dass ausschließlich längere Studien (≥ 6 Monate) hätten berücksichtigt werden sollen (SN von Maxen, SN Wille) oder dass die Wahl des Einschlusskriteriums willkürlich sei. Das gewählte Vorgehen, Studien mit einer kontrollierten Beobachtungsphase von mindestens 16 Wochen einzuschließen, beruhte auf einer pragmatischen Entscheidung. Die EMA fordert eine Dauer von mindestens 6 Monaten [42]. Jedoch wird auf der anderen Seite hinsichtlich des individuellen Therapieverlaufs davon ausgegangen, dass sich eine Änderung zum Spontanverlauf gegebenenfalls bereits früher zeigt [43]. Für Donepezil und Rivastigmin ist dieser Sachverhalt im vorliegenden Bericht unerheblich, da alle zu Donepezil eingeschlossenen Studien mindestens 6 Monate dauerten und für Rivastigmin lediglich eine kürzere Studie vorlag (Forette 1999 mit 18 Wochen), die jedoch keine interpretierbaren Daten zu den Aspekten der Kognition und der Aktivitäten des täglichen Lebens lieferte. Bezüglich Galantamin hätte eine Vorgehensweise gemäß der EMA dazu geführt, dass zwei Studien (Rockwood 2006 mit 4 sowie Tariot 2000 mit 5 Monaten) nicht in die Nutzenbewertung eingegangen wären. Eine diesbezügliche Sensitivitätsanalyse ergab, dass sich auch bei Ausschluss dieser Studien nahezu identische Schätzer im Bereich der Kognition und der Aktivitäten des täglichen Lebens ergaben.

Im Rahmen des Stellungnahmeverfahrens wurde kritisiert, dass ausschließlich randomisierte kontrollierte Studien für die Nutzenbewertung herangezogen wurden (vgl. z. B. SN Gutzmann, SN Janssen-Cilag, SN Novartis, SN Pfizer, SN Wahler). Insbesondere wurde kommentiert, dass auch Daten aus unkontrollierten Open-label-follow-up-Phasen sowie Vergleiche mit historischen Kontrollen hätten berücksichtigt werden sollen.

Im vorliegenden Bericht wurde deshalb ausschließlich auf RCTs rekurriert, da diese mit der geringsten Ergebnisunsicherheit behaftet sind. Nur durch die Randomisierung kann eine strukturelle Gleichheit der Gruppen, auch hinsichtlich unbekannter Confounder, gewährleistet werden [125]. Bei der Auswertung von Studienphasen aus unkontrollierten Nacherhebungsphasen lässt sich der beobachtete Verlauf unter der Behandlung nicht gegenüber dem natürlichen Verlauf abgrenzen und so der Behandlungseffekt nicht erschließen. Das Heranziehen beispielsweise einer historischen Kontrollgruppe wirft wegen der unklaren Vergleichbarkeit der behandelten und unbehandelten Gruppe schwerwiegende methodische Probleme auf.

Das Vorgehen, sich auf randomisierte kontrollierte Studien zu beschränken, entspricht im Bereich der Forschung zu Cholinesterasehemmern bei Alzheimer Demenz dem internationalen Standard und wurde auch in jüngsten internationalen HTAs bzw. Systematischen Reviews entsprechend angewendet (z. B. [49,108,112]). Abweichend von der Darstellung in der Stellungnahme von Janssen-Cilag wurde auch bspw. in der Leitlinie des Scottish Intercollegiate Guidelines Network bei der Ableitung der Empfehlungen zu Cholinesterasehemmern ausschließlich auf RCTs Bezug genommen (vgl. SN Janssen-Cilag; [126]).

- Bewertung einzelner Aspekte des Studiendesigns

Im Rahmen des Stellungnahmeverfahrens wurde kritisiert, dass die Aspekte der Randomisierungsprozedur, der Verdeckung der Gruppenzuteilung und der Verblindung sich nicht in ausreichendem Maße in der Gesamtbewertung der Studienqualität widerspiegelt hätten (vgl. SN von Maxen, SN Wille).

Allerdings konnten die in der Anhörung von einigen Herstellern zur Verfügung gestellten Zusatzinformationen die im Vorbericht dargestellten Unklarheiten im Hinblick auf wesentliche Qualitätsaspekte zum größten Teil aufgelösen. Von den Herstellerfirmen Eisai und Janssen-Cilag wurde in den schriftlichen Stellungnahmen dargelegt, dass die Randomisierung in allen von ihnen durchgeführten Studien computergeneriert und zentral erfolgte; die Zuteilung der Patienten erfolgte über Randomisierungsnummern auf der Studienmedikation, die jeweils identisch in der äußeren Erscheinung und Verpackung gewesen sei. Auf Grundlage dieser Informationen wurde im vorliegenden Abschlussbericht das Concealment of Allocation für die allermeisten Studien als adäquat bewertet. Das Gleiche galt für die Angaben zur Randomisierungsprozedur.

In den Stellungnahmen wurde auf methodische Arbeiten verwiesen, in denen bspw. eine Effektüberschätzung von 41 % bei nicht gegebenem bzw. 30 % bei unklarem Concealment belegt worden sei (vgl. SN von Maxen, SN Wille). Angesichts der zusätzlichen Herstellerinformationen ist dieses Problem einer möglichen Überschätzung des Effektes durch ein nicht adäquates Concealment in den diesem Bericht zugrunde liegenden Studien

offensichtlich kaum relevant. Bei der Interpretation der in den Stellungnahmen herangezogenen Zahlen ist darüber hinaus zu berücksichtigen, dass bspw. die Arbeit von Schulz sich auf 33 Meta-Analysen aus dem geburtshilflichen Bereich mit 250 Studien aus den Jahren von 1955 bis 1992 bezieht, knapp 30 % der Studien stammten aus der Zeit vor 1980 [127]. Zwar ergab sich bei der Betrachtung von Studien mit adäquatem und unklarem Allocation Concealment unter Berücksichtigung weiterer Qualitätsaspekte ein um 30 % relativ größerer Therapieeffekt bei Studien mit unklarem Allocation Concealment, allerdings war dieser Zusammenhang in den einzelnen Meta-Analysen sehr unterschiedlich und sollte laut den Autoren nicht als genereller Wert interpretiert werden. Insbesondere das Alter der eingeschlossenen Studien und die Einschränkung auf den geburtshilflichen Bereich lassen vermuten, dass es sich – wenn überhaupt – um nur wenige im Sinne des Arzneimittelrechts zulassungsrelevante und damit GCP-konforme Studien handelte. Damit ist eine Übertragbarkeit auf die diesem Bericht zugrunde liegenden Studien fraglich.

Während – bis auf einige direkte Vergleichsstudien – alle Studien doppelblind durchgeführt wurden, war die Verblindung der Endpunkterheber in vielen Publikationen kaum beschrieben. Im Rahmen des Stellungnahmeverfahrens stellten die Hersteller auch zu diesem Aspekt weitere Informationen zur Verfügung bzw. beschrieben das übliche Vorgehen. Dabei wurde von den Firmen angegeben, dass der CIBIC-plus stets durch unabhängige Rater erfasst werde, die keinen Zugang zu jeglichen anderen Informationen bzgl. der Patienten hätten (vgl. SN Pfizer, SN Janssen-Cilag zur Studie Tariot 2001, SN Novartis). Hinsichtlich der Beurteilung des ADAS-cog waren die Anforderungen in den Studien offensichtlich weniger streng. Hierzu berichteten die Vertreter der Firma Novartis in der Erörterung, dass dieser üblicherweise jeweils durch Psychologen durchgeführt worden sei, die normalerweise nicht von der Station und häufig überhaupt nicht mit Patientenakten in Kontakt gekommen seien (vgl. Protokoll, Anhang H). In der Stellungnahme der Firma Eisai wurde dargestellt, dass der ADAS-cog üblicherweise nicht durch den Studienleiter, sondern durch ein anderes Mitglied des Studienteams erhoben werde, meist bevor andere Messungen durchgeführt würden (vgl. SN Eisai). Eine formale Verblindung der ADAS-cog-Rater sei deshalb nicht notwendig, insbesondere da diese Personen im Allgemeinen keinen Zugriff auf andere Untersuchungsergebnisse hätten (vgl. SN Eisai). Janssen-Cilag stellte in der Stellungnahme für die Studie Tariot 2001 dar, der ADAS-cog-Rater habe nicht in die Behandlung eingebunden sein und keinen Zugang zu Informationen bzgl. der unerwünschten Ereignisse haben sollen. Falls dies nicht möglich gewesen sei, habe der Rater den ADAS-Test durchgeführt, bevor er die unerwünschten Ereignisse erfragt habe (vgl. SN Janssen-Cilag). Zu der Frage der Verblindung bezüglich der Erhebung der Aktivitäten des täglichen Lebens sowie der neuropsychiatrischen Symptome wurden im Stellungnahmeverfahren von den Herstellern keine zusätzlichen Informationen geliefert.

Insofern war in den meisten Studien der Verblindungsstatus bei der Erhebung der für den vorliegenden Bericht als vornehmlich relevant definierten Endpunkte nicht vollständig

befriedigend und wurde dementsprechend im Rahmen der Qualitätsbewertung aufgrund dieser Angaben jeweils mit einem eingeschränkten „(ja)“ beschrieben.

- Umgang mit fehlenden Werten

Bewertung der Umsetzung des ITT-Prinzips

Das ITT-Prinzip, d. h. die vollständige Auswertung aller randomisierten Patienten in den ihnen laut Randomisierungscode zugeteilten Gruppen, stellt ein wichtiges Qualitätskriterium bei der Bewertung klinischer Studien dar. Nur so kann verhindert werden, dass die durch die Randomisierung hergestellte strukturelle Gleichheit der Gruppen durch die Nichtberücksichtigung von Patienten mit prognostisch bedeutsamen Merkmalsausprägungen wieder zerstört wird. Im Vorbericht wurde deshalb die Umsetzung des ITT-Prinzips als ein wesentliches Qualitätsmerkmal bewertet und wurden entsprechende Grenzwerte bezüglich der tolerierten Nichtberücksichtigungsrate festgesetzt. In den Stellungnahmen wurde kritisiert, die Prozentzahlen 10,5 (für den Gesamtanteil von fehlenden Patienten in der [Haupt-] Auswertung) und 4,5 (für die absolute Differenz des Anteils fehlender Patienten) erschienen willkürlich und seien nicht begründet; mit gleichem Recht hätte man eine Grenze bei 4, 12 oder 20 % ziehen können (vgl. SN Arznei-Telegramm, SN Kaduszkiewicz). Tatsächlich sind in der Literatur keine eindeutigen Grenzwerte festgelegt. Auch bspw. im Cochrane Handbook wird festgestellt, dass die Bewertung von Ergebnissen aus Studien mit mehr als minimalen Mengen von fehlenden Werten letztendlich eine Frage des Ermessens sei ([125], S. 113). Schulz und Grimes beschrieben 2002 eine „Faustregel“: Datenverluste von weniger als 5 % seien in aller Regel weniger kritisch, solche über 20% könnten die Validität ernsthaft infrage stellen [128]. Zu einem noch tolerablen Grenzwert beim differentiellen Datenverlust äußerten sie sich nicht, stellten aber fest, dass dieser (differentielle Datenverlust) noch problematischer sei. Insofern muss eine Grenzziehung in gewissem Sinne willkürlich sein. Für den vorliegenden Bericht wurde angenommen, dass eine Gesamtnichtberücksichtigungsrate von (gerundet) mindestens 11 % als kritisch anzusehen ist, genauso wie ein Unterschied in der Berücksichtigungsrate zwischen den Gruppen von (gerundet) mindestens 5 Prozentpunkten. Um die üblichen Rundungen (bei denen ab 10,5 % auf 11 % und ab 4,5 % auf 5 % aufgerundet wird) zu berücksichtigen, wurden im Vobericht die Zahlen mit Nachkommastelle angegeben. In Sensitivitätsanalysen zum ADAS-cog zeigte sich, dass aufgrund der großen Homogenität der Studienergebnisse auch eine moderate Verschiebung der Grenzen nach oben oder unten keine Änderung an der grundsätzlichen Aussage mit sich bringt: Von einer moderaten Verschiebung dieser Grenze der Nichtberücksichtigungsrate (vgl. Abschnitt 4.3) wären für Donepezil zunächst die Studien Gauthier 2002, Rogers 1998 und Krishnan 2003 betroffen. Eine Meta-Analyse ohne Studien, die dadurch als mit groben Mängeln behaftet eingestuft würden, führt zu keiner wesentlichen Änderung des gepoolten Schätzers (-0,45). Für Galantamin würde dies die Studien Brodaty 2005 und Tariot 2000 betreffen, ohne deren Berücksichtigung sich in der Meta-Analyse ein vergleichbarer gepoolter Schätzer (-0,55)

ergibt. Für Rivastigmin wären auch bei strengeren Grenzen keine weiteren Studien davon betroffen.

LOCF als Ersetzungsstrategie – Meta-Regression

Im Vorbericht wurde die Kritik an der LOCF-Methode als Ersetzungsstrategie im Rahmen einer sich progredient verschlechternden Erkrankung diskutiert. Dabei liegt die Annahme zugrunde, dass die Studienergebnisse bei einem frühzeitigen Ausscheiden der Patienten, insbesondere aufgrund von Nebenwirkungen, durch das Fortschreiben in einer für das Verum günstigen Weise verzerrt werden könnten (vgl. S. 172). Die Meta-Regression, die durchgeführt wurde, um den Einfluss der differentiellen Abbrüche im vorliegenden Bericht abzuschätzen, wurde im Rahmen des Stellungnahmeverfahrens kritisiert und dieser eine eigene Meta-Regression gegenübergestellt (vgl. SN Kaduszkiewicz). Dabei wurden die Studien, die im Vorbericht mit „groben Mängeln“ bewertet worden waren, nicht einbezogen. Die Begründung war, dass diese Beurteilung vor allem auf Unklarheiten bzw. deutlichen Abweichungen vom ITT-Prinzip beruht hätte, was die Ergebnisse möglicherweise deutlich verfälscht haben könnte (vgl. SN Kaduszkiewicz). Diese modifizierte Meta-Regression ergab nun einen statistisch signifikanten Zusammenhang zwischen der Differenz der Abbruchraten und der Effektgröße. Eigene Re-Analysen zeigten, dass dieser Unterschied im Ergebnis in erster Linie auf die Nichtberücksichtigung der Studie Rösler 1999 zurückzuführen ist. Angesichts der Tatsache, dass der Ein- oder Ausschluss einzelner Studien das Ergebnis so deutlich verändert, wird die Instabilität und damit begrenzte Aussagekraft dieser Analyse deutlich. Jedoch zeigen beide Meta-Regressionen bei gleichen Abbruchraten einen Behandlungseffekt von über 0,4 Standardabweichungen, der nur geringfügig unter dem global geschätzten Effekt von 0,5 Standardabweichungen liegt. Insofern widerlegt auch diese Analyse die Aussage des Berichtes nicht.

Davon unbenommen ist für zukünftige Studien eine vollständige Erfassung der Patienten zu fordern. Zwar werden verschiedene Ersetzungsstrategien als Alternative zu LOCF diskutiert (z. B. Multiple Imputation-Methoden [129], vgl. hierzu auch SN Eisai zur Re-Analyse der Studienergebnisse von Rogers 1998) – letztendlich bleibt aber festzustellen, dass die Nichtverfügbarkeit von Information durch nachträgliche Auswertungsmethoden nicht zu kompensieren ist. Um solchen Problemen zu begegnen und den potenziellen Datenverlust so niedrig wie möglich zu halten, wäre eine Möglichkeit, durch die Wahl pragmatischer und weniger komplexer Erfassungsinstrumente den Aufwand einer Nachbeobachtung für die Patienten zu verringern.

- Stellenwert des CIBIC-plus

Im Rahmen des schriftlichen Anhörungsverfahrens wurde von einigen Stellungnehmenden kritisch kommentiert, dass der CIBIC-plus herabgestuft worden sei. Dies sei klinisch nicht

korrekt, da gerade dieses Instrument dazu entwickelt worden sei, klinisch bedeutsame Veränderungen zu erfassen (vgl. z. B. SN Burns, SN Eisai, SN Maier, SN Pfizer).

Im Berichtsplan zur vorliegenden Nutzenbewertung war das Zielkriterium „klinisches Krankheitsstadium gemäß dem klinischen Eindruck“ nicht als patientenrelevanter Endpunkt definiert. Gemäß einem Amendment wurden die Ergebnisse hierzu als ergänzende Information dennoch dargestellt. Hiermit wurde unter anderem dem Umstand Rechnung getragen, dass „globales Ansprechen“ als eine der drei relevanten Domänen im Zulassungsverfahren definiert ist [42]. Wegen grundlegender Fragen hinsichtlich der Patientenrelevanz der Instrumente zur Erfassung des klinischen Gesamteindrucks – in den allermeisten Studien wurde der CIBIC-plus eingesetzt – gingen die Ergebnisse jedoch nicht primär in die Nutzenbewertung ein. Diese grundsätzlichen Fragen wurden auch vom Canadian Coordinating Office for Health Technology Assessment ausführlich diskutiert [130], das ebenfalls zu dem Schluss kam, dass derartige Maße als Zielgrößen nicht geeignet seien. Kritikpunkte bestehen unter anderem darin, dass je nachdem, wer die globale Verbesserung beurteilt (Ärzte, Pflegepersonal oder Angehörige), auf unterschiedliche Bereiche fokussiert werde (kognitive Fähigkeiten, notwendiger Pflegeaufwand oder Verhalten bzw. Alltagsaktivitäten). Damit spiegelten derartige Instrumente weniger den individuellen Grad an globaler Verbesserung wider, sondern vielmehr eine Verbesserung in demjenigen Funktionsbereich, der vom jeweiligen Rater als wichtig erachtet werde. Hinzu kommen speziell beim CIBIC-plus weitere problematische Aspekte. So variiert bspw. die Beurteilung in Abhängigkeit von der Reihenfolge, in der Patient und Betreuer befragt werden – mit ungünstigeren Einschätzungen, wenn zunächst die betreuende Person interviewt wird [130].

- Stellenwert der angehörigerelevanten Therapieziele

Einige Stellungnehmende (vgl. z. B. SN von Lützu-Hohlbein, SN Möller, SN Maier) kritisierten, dass im Amendment zum Berichtsplan angehörigerelevante Therapieziele herabgestuft worden seien. Tatsächlich wurde im vorliegenden Bericht unterschieden zwischen solchen Zielen, die unmittelbar aus Sicht des Patienten selber relevant sind, und solchen, bei denen die Perspektive der Angehörigen im Fokus steht – dies betraf die Ziele „Lebensqualität der (betreuenden) Angehörigen“ sowie „Höhe des Betreuungsaufwands durch eine oder mehrere betreuende Person(en) oder Institution(en)“. Ergebnisse zu diesen Therapiezielen wurden im Bericht aufgelistet und dargestellt, standen aber für die Nutzenbewertung an zweiter Stelle. Auch wenn eine solche Abgrenzung zwangsläufig nicht ganz trennscharf sein kann, wurde diese Unterscheidung mit der Absicht getroffen, darzustellen, dass es bei einer Nutzenbewertung um die Patientenperspektive und somit in erster Linie um den direkten Nutzen für die Patienten geht. Dies schien u. a. deshalb geboten, da – zumindest theoretisch – Konstellationen denkbar sind, in denen die Bedürfnisse der Patienten und die der Angehörigen nicht unbedingt gleichläufig sind (wenn bspw. eine Sedierung der Patienten die Belastung der Betreuer verringert). Dass eine Verbesserung des

Zustandes des Patienten zu einer geringeren Belastung und zu einer Erhöhung der Lebensqualität der Betreuungsperson führt, ist wünschenswert, und dementsprechend wurden auch im Rahmen des vorliegenden Berichts solche Zielkriterien abgebildet. Damit ist jedoch die Verbesserung des Patienten die notwendige Voraussetzung und steht somit an erster Stelle der Bewertung. Ohnehin ist die für die angehörigenrelevanten Zielkriterien vorhandene Evidenz quantitativ und qualitativ deutlich schwächer als bspw. für die Beeinflussung der kognitiven oder alltagspraktischen Fähigkeiten der Patienten. Eine stärkere Berücksichtigung dieser Aspekte hätte somit nicht zu einer anderen Bewertung der Cholinesterasehemmer geführt.

- ADAS-cog-Skala

In einigen Stellungnahmen wurde kritisiert, dass die Scores, die in den Studien untersucht wurden, lediglich Surrogate darstellten (SN von Maxen, SN Wille). Auch im Rahmen der mündlichen Erörterung wurde die ADAS-cog-Skala ausführlich diskutiert.

Die Relevanz der ADAS-cog-Skala leitet sich aus dem Krankheitsbild der Alzheimer Demenz ab. So ist die Alzheimer Demenz vornehmlich durch das Auftreten beobachtbarer kognitiver Einschränkungen gekennzeichnet, die mit einer Beeinträchtigung der sozialen oder beruflichen Leistungsfähigkeit verbunden sind [18]. Leitsymptom der Erkrankung ist die Entwicklung multipler kognitiver Defizite, die sich sowohl in einer Gedächtnisstörung als auch mindestens einer weiteren kognitiven Störung (Aphasie, Apraxie, Agnosie oder Störung der Exekutivfunktionen) zeigen [18]. Die ADAS-cog-Skala erfasst in 11 Items bzw. Subskalen die Symptombereiche Gedächtnis und Orientierung, Sprache und Praxie [130]. Damit bildet sie wichtige Symptombereiche der Alzheimer Erkrankung ab, auch wenn einzelne Bereiche (z. B. Störungen der Exekutivfunktion) nicht erfasst werden [130]. Die von der Skala abgebildeten Symptome ermöglichen es, zwischen kognitiv nicht beeinträchtigten Personen und Alzheimer Erkrankten zu diskriminieren [28,130]. Da diese Symptome den Kern der Erkrankung darstellen, erscheint es nicht plausibel, die Skala lediglich als Surrogat zu betrachten. In diesem Zusammenhang wurde darüber hinaus in einigen Stellungnahmen kritisiert, dass nicht a priori eine als klinisch relevant betrachtete Punktwertdifferenz definiert worden sei (vgl. SN von Maxen, SN Wille).

In der einschlägigen Literatur wird bezüglich der klinischen Relevanz üblicherweise auf die FDA rekurriert, die eine Umkehrung des natürlichen Verlaufs um mindestens 6 Monate als klinisch relevant ansehe (z. B. [131,132]), auch wenn diese Feststellung der herangezogenen Referenz nicht eindeutig zu entnehmen ist^j. Diese Zeitspanne wird häufiger (z. B. [130,132])

^j Hierbei handelt es sich um ein Gesprächsprotokoll des Peripheral and Central Nervous System Drugs Advisory Committee. Einer solchen Aussage am nächsten kommt der folgende Redebeitrag: „In polling a lot of physicians who treat Alzheimer’s patients – at least this came out at our symposium several weeks ago – most physicians

mit einer Veränderung von 4 Punkten auf der ADAS-cog-Skala gleichgesetzt, unter Bezug auf eine Arbeit aus dem Jahr 1988 [134]. Allerdings beziehen sich in der betreffenden Studie die Veränderungen bei den untersuchten Personen tatsächlich auf die ADAS-Gesamtskala, so dass eine derartige Differenz nicht zwangsläufig mit einer entsprechenden Veränderung der Subskala ADAS-cog gleichzusetzen ist. In einer weiteren Längsschnittstudie zum Verlauf der kognitiven Subskala ergab sich ein umgekehrt U-förmiger Zusammenhang zwischen der Krankheitsschwere und dem Ausmaß der Verschlechterung, mit einer mittleren Veränderung über 12 Monate von 9,6 Punkten (SD = 8,2) bei Alzheimer Patienten und praktisch keiner Veränderung (Verbesserung von 0,2 Punkten, SD = 2,0) bei gesunden Kontrollpersonen [135]. Über Patienten in den Placebo-Gruppen klinischer Studien wurde berichtet, dass der Verlauf mit etwa 5–6 Punkten Verschlechterung innerhalb eines Jahres [136] offensichtlich etwas günstiger sei als in anderen Stichproben [132]. Damit läge (bei Annahme eines linearen Verlaufs) die natürliche Verschlechterung über 6 Monate bei etwa 2,5 bis 3 Scorepunkten im ADAS-cog.

Darüber hinaus besteht derzeit kein Konsens darüber, wie groß eine Differenz im ADAS-cog sein muss, um als klinisch relevant zu gelten (vgl. SN Arznei-Telegramm) – weder auf individueller noch auf Gruppenebene. Deshalb scheint es plausibel, neben der biometrischen Einordnung des gefundenen Behandlungseffekts (eine Mittelwertsdifferenz von 0,5 SD wird im Allgemeinen als mittlere Effektgröße angesehen [137]) das Kriterium der 6-monatigen Krankheitsprogressionsverzögerung heranzuziehen.

Wird ein solcher Effekt im Sinne einer Responseudefinition auf die individuelle (Patienten-) Ebene übertragen, entspricht dies einer (absoluten) Differenz von etwa 15 Prozentpunkten Respondern – gemessen an der kognitiven Funktion – zwischen Verum und Placebo (siehe Tabelle 46^k) oder einer Erhöhung der Chance für einen Response um den Faktor 2,3 bis 2,5.

Im Rahmen des Stellungnahmeverfahrens wurden allerdings auch Responderanalysen kommentiert. Es wurde kritisiert, dass Responderanalysen zu einer Überschätzung des Therapieerfolgs führen können, wenn sich die Ergebnisse der verschiedenen Behandlungsgruppen um den Cut-off-Wert scharen (vgl. SN Arznei-Telegramm). Ein solches Argument kann immer ins Feld geführt werden, wenn quantitative Daten kategorisiert bzw. dichotomisiert werden. Es ist Ausdruck des Verhältnisses von auf der individuellen Ebene als relevant erachtetem Effekt (individuelle Veränderung auf der Skala, Responsekriterium) und der Variabilität dieser Veränderung. Damit führt es wieder auf die Frage, ob der als relevant erachtete Effekt tatsächlich relevant ist, und ist somit rekursiv. Darüber hinaus gilt es jedoch beim Zwischen-Gruppen-Vergleich für beide Gruppen gleichermaßen (es sei denn, in den

polled felt that a three- or four-month improvement was worth taking a drug for, and that something much less than that was probably not worth taking a drug for.” (S. 227) [133]

^k Dort wurde ein etwas höheres Responsekriterium verwendet. Eine weitere Responderdefinition wurde bspw. in Analysen der Firma Janssen-Cilag angewendet (vgl. SN Janssen-Cilag sowie [46]).

Gruppen existiert eine unterschiedliche Variabilität) und kann sowohl in die eine Richtung („Überschätzung“) wie auch in die andere („Unterschätzung“) wirken.

Weiterhin wurde in einer Stellungnahme angemerkt, dass die Annahme einer Normalverteilung – so wie im vorliegenden Bericht für die theoretisch abgeleiteten Responderanalysen erfolgt – bei einer eigentlich schiefen Verteilung zu einer Vergrößerung des Effektes führen könne (vgl. SN Kaduszkiewicz). Tatsächlich kann eine Verletzung der Normalverteilungsannahme sowohl zu einer Über- als auch einer Unterschätzung des Effektes führen. Allerdings führt die durchgeführte Differenzbildung (Differenz aus dem Wert am Ende und dem Wert zu Beginn) zu einer Symmetrisierung der Verteilung. Zudem stimmten – wie in Tabelle 46 beschrieben – die in einigen Studien berichteten Daten zu Respondern mit den theoretisch erwarteten recht gut überein.

Es bleibt einschränkend anzumerken, dass aus ursprünglich quantitativen Daten abgeleitete Responderanalysen hier einen Versuch darstellen, sich der Bedeutung der beobachteten Veränderungen und Mittelwertsunterschiede zu nähern; jedoch sind sie in aller Regel in ihrer Aussagekraft nicht vergleichbar mit entsprechenden Analysen bei tatsächlich dichotomen Kriterien. Insgesamt ist jedoch der konsistent nachgewiesene Effekt auf einer Skala, die einen patientenrelevanten Krankheitsaspekt abbildet, als direkter patientenrelevanter Nutzen zu bewerten [138]. Bestärkt wird diese Einschätzung durch den ebenfalls über alle Substanzen hinweg darzustellenden – wenn auch deutlich geringeren – Effekt bei den Skalen, die die Einschränkungen bei den Aktivitäten des täglichen Lebens abbilden sollen.

Davon unabhängig ist, wie bereits diskutiert (vgl. S. 174), für die Zukunft die Entwicklung von Instrumenten wünschenswert, die größeren Aufschluss darüber geben, inwiefern die erzielten Veränderungen für die Patienten im täglichen Leben eine spürbare Verbesserung mit sich bringen.

Auch die in einigen Stellungnahmen als alternativ einzusetzenden Zielkriterien wie Institutionalisierung und Betreuungsaufwand stellen letztlich auch nur Surrogate für den direkten patientenrelevanten Nutzen dar, da Konstellationen vorstellbar sind, in denen auf dieser Ebene als ungünstig einzustufenden Ergebnissen (Aufnahme in ein Pflegeheim, größerer Betreuungsaufwand) durchaus positive Ergebnisse auf Patientenseite gegenüberstehen (bewusster patientenseitiger Entschluss, größere Teilhabemöglichkeit). Darüber hinaus müssen Daten zur Institutionalisierung je nach Definition von „Institution“ differenziert betrachtet werden und sind aus Studien eines anderen Versorgungskontextes – wenn überhaupt – nur sehr eingeschränkt auf Deutschland übertragbar.

- Schlussfolgerung

Letztlich führen die in den Stellungnahmen vorgebrachten Kritikpunkte, die sowohl eine zu positive als auch eine zu negative Nutzenbewertung der Cholinesterasehemmer im

vorliegenden Bericht implizieren, zu keiner wesentlichen Änderung des bereits im Vorbericht getroffenen Fazits.

Ausblick

In der notwendigen Nutzen-Schaden-Abwägung stellt sich natürlich die Frage nach Therapiealternativen. Es wurden im Rahmen des vorliegenden Berichts keine Studien (gemäß Ein- und Ausschlusskriterien) identifiziert, die Cholinesterasehemmer mit einer medikamentösen oder nichtmedikamentösen Alternative verglichen hätten. Eine unlängst publizierte Studie, in der u .a. Patienten mit Donepezil oder Ginkgo biloba behandelt wurden, ist methodisch und in ihrer Darstellung zweifelhaft und darüber hinaus viel zu klein, um angesichts der nur begrenzt großen Effektstärke im Vergleich zu Placebo valide Aussagen zu einem Vergleich zwischen Donepezil und Ginkgo zu ermöglichen. Die Studie wurde im Vorbericht A05-19B ausführlich diskutiert [139,140].

Weiterhin bemerkenswert ist eine ebenfalls kürzlich veröffentlichte Arbeit, in der bei Patienten mit einer milden bis moderaten Demenz die Effekte einer Ergotherapie mit Fokus sowohl auf die Patienten als auch auf die betreuenden Personen im Vergleich zu einer Kontrollgruppe ohne entsprechende Kontextbedingungen (Wartelistenkontrolle) evaluiert wurden [141]. Auch wenn hier keine Differenzierung nach Demenztypen und kein Vergleich mit Cholinesterasehemmern erfolgte und auch die Studiendauer nur 3 Monate betrug, so übertrafen doch die beobachteten Effektstärken mit einem Gruppenunterschied von mehr als 2 Standardabweichungen beispielsweise auf der IDDD-Skala (Aktivitäten des täglichen Lebens) die der Studien aus dem vorliegenden Bericht um ein Vielfaches. Ein solcher Effekt wäre nahezu als dramatisch zu bezeichnen und bedarf sicherlich der Reproduktion und ausführlichen Diskussion der dazugehörigen Rahmen- und Kontextbedingungen. Insbesondere der Stellenwert einer Wartelistenkontrolle sollte kritisch diskutiert werden. Darüber hinaus muss beachtet werden, dass alle Patienten in dieser Studie unter einer Medikation mit Cholinesterasehemmern oder Memantine standen. Nichtsdestotrotz stellen diese Studienergebnisse die oft behauptete (Therapie-)Alternativlosigkeit bei Patienten mit einer Demenz infrage.

Um die Relevanz der durch Cholinesterasehemmer im Vergleich zu Placebo induzierten Effekte besser einordnen zu können, wären Studien mit deutlich längeren Laufzeiten unter kontrollierten Bedingungen zu fordern. Placebo-kontrollierte Langzeitstudien stoßen inzwischen auf ethische Bedenken (vgl. z. B. SN Gutzmann, SN Janssen-Cilag, SN Novartis, SN Pfizer). Angesichts der im vorigen Absatz kurz skizzierten Ergebnisse einer nichtmedikamentösen Intervention stellt womöglich ein entsprechender Vergleich mit der Gabe von Cholinesterasehemmern über einen Zeitraum von mindestens 1 bis 2 Jahren einen Ausweg aus dem (ethischen) Dilemma dar.

7 FAZIT

Die Cholinesterasehemmer Donepezil, Galantamin und Rivastigmin haben bei Patienten mit einer Alzheimer Demenz leichten bis mittleren Schweregrades einen Nutzen bezüglich des Therapieziels der kognitiven Leistungsfähigkeit. Für Donepezil gilt dies über alle eingesetzten Dosen hinweg, für Galantamin und Rivastigmin nur bei mittleren und hohen Dosen.

Weiterhin gibt es für alle drei Substanzen Hinweise auf einen Nutzen im Hinblick auf das Therapieziel der Aktivitäten des täglichen Lebens.

Für Galantamin gibt es darüber hinaus Hinweise auf einen Nutzen bezüglich der begleitenden psychopathologischen Symptome. Für Donepezil deuten die vorliegenden Daten auf keinen entsprechenden Nutzen, für Rivastigmin gibt es diesbezüglich keine Daten.

Für das Therapieziel der krankheitsbezogenen Lebensqualität liegen entweder ebenfalls keine Daten vor (Galantamin und Rivastigmin) oder sie liefern keinen Hinweis auf einen Nutzen (Donepezil).

Zur Vermeidung der Notwendigkeit einer vollstationären Pflege (Institutionalisierung) liegen keine interpretierbaren Daten vor.

Dem gegenüber steht für alle drei Wirkstoffe ein dosisabhängiger Schaden durch das Auslösen therapieassoziierter unerwünschter Ereignisse. Eine Beeinflussung der Mortalität lässt sich aus den vorliegenden Daten nicht ableiten, allerdings waren die Studien auch nicht angelegt, dazu Aussagen zu treffen.

Während es im direkten Vergleich Hinweise auf einen Zusatznutzen durch Rivastigmin gegenüber Donepezil bezüglich der Aktivitäten des täglichen Lebens gibt, hat Rivastigmin andererseits auch ein größeres Schadenspotenzial. Zu den anderen beiden möglichen Vergleichen (Galantamin versus Donepezil bzw. Galantamin versus Rivastigmin) lassen sich keine Aussagen treffen. Insgesamt lässt sich kein eindeutiger Vorteil für eine der drei untersuchten Substanzen aus den vorliegenden Daten ableiten.

Die oben getroffenen Aussagen beziehen sich im Wesentlichen auf den den meisten Studien zugrunde liegenden Zeitraum von bis zu 6 Monaten. Für eine weitergehende Nutzen-Schaden-Abwägung wären direkte Vergleichsstudien zu anderen Behandlungsoptionen (andere Medikamente oder nichtmedikamentöse Behandlungsstrategien) wünschenswert.

Der Stellenwert der Cholinesterasehemmer gegenüber anderen medikamentösen und nichtmedikamentösen Behandlungsoptionen ist mangels Daten unklar.

8 LISTE DER EINGESCHLOSSENEN STUDIEN

Donepezil vs. Placebo

Burns 1999 [51]

Burns A, Rossor M, Hecker J, Gauthier S, Petit H, Moller HJ, Rogers SL et al. The effects of donepezil in Alzheimer's disease - results from a multinational trial. *Dement Geriatr Cogn Disord* 1999; 10: 237-244.

Gauthier 2002 [52-54]

Gauthier S, Feldman H, Hecker J, Vellas B, Emir B, Subbiah P, Donepezil MSAD Study Investigators' Group. Functional, cognitive and behavioral effects of donepezil in patients with moderate Alzheimer's disease. *Curr Med Res Opin* 2002; 18: 347-354.

Feldman H, Gauthier S, Hecker J, Vellas B, Subbiah P, Whalen E, Donepezil MSAD Study Investigators Group. A 24-week, randomized, double-blind study of donepezil in moderate to severe Alzheimer's disease. *Neurology* 2001; 57: 613-620.

Feldman H, Gauthier S, Hecker J, Vellas B, Emir B, Mastey V, Subbiah P et al. Efficacy of donepezil on maintenance of activities of daily living in patients with moderate to severe Alzheimer's disease and the effect on caregiver burden. *J Am Geriatr Soc* 2003; 51: 737-744.

Homma 2000 [55]

Homma A, Takeda M, Imai Y, Udaka F, Hasegawa K, Kameyama M, Nishimura T. Clinical efficacy and safety of donepezil on cognitive and global function in patients with Alzheimer's disease. A 24-week, multicenter, double-blind, placebo-controlled study in Japan. E2020 Study Group. *Dement Geriatr Cogn Disord* 2000; 11: 299-313.

Krishnan 2003 [57]

Krishnan KR, Charles HC, Doraiswamy PM, Mintzer J, Weisler R, Yu X, Perdomo C et al. Randomized, placebo-controlled trial of the effects of donepezil on neuronal markers and hippocampal volumes in Alzheimer's disease. *Am J Psychiatry* 2003; 160: 2003-2011.

Mohs 2001 [58]

Mohs RC, Doody RS, Morris JC, Ieni JR, Rogers SL, Perdomo CA, Pratt RD et al. A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology* 2001; 57: 481-488.

Moraes 2006 [59]

dos Santos Moraes WA, Rollemberg Poyares D, Guilleminault C, Ramos LR, Bertolucci PHF, Tufik S. The effect of donepezil on sleep and REM sleep EEG in patients with Alzheimer Disease: a double-blind placebo-controlled study. *Sleep* 2006; 29: 199-205.

Prasher 2002 [60]

Prasher VP, Huxley A, Haque MS, Down syndrome Ageing Study Group. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Down syndrome and Alzheimer's disease-pilot study. *Int J Geriatr Psychiatry* 2002; 17: 270-278.

Rogers 1998 [61]

Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology* 1998; 50: 136-145.

Seltzer 2004 [62]

Seltzer B, Zolnouni P, Nunez M, Goldman R, Kumar D, Ieni J, Richardson S et al. Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial. *Arch Neurol* 2004; 61: 1852-1856.

Tariot 2001 [63]

Tariot PN, Cummings JL, Katz IR, Mintzer J, Perdomo CA, Schwam EM, Whalen E. A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting. *J Am Geriatr Soc* 2001; 49: 1590-1599.

Tune 2003 [64]

Tune L, Tiseo PJ, Ieni J, Perdomo C, Pratt RD, Votaw JR, Jewart RD et al. Donepezil HCl (E2020) maintains functional brain activity in patients with Alzheimer disease: results of a 24-week, double-blind, placebo-controlled study. *Am J Geriatr Psychiatry* 2003; 11: 169-177.

Winblad 2001 [65-67]

Winblad B, Engedal K, Soininen H, Verhey F, Waldemar G, Wimo A, Wetterholm AL et al. A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology* 2001; 57: 489-495.

Wimo A, Winblad B, Shah SN, Chin W, Zhang R, McRae T. Impact of donepezil treatment for Alzheimer's disease on caregiver time. *Curr Med Res Opin* 2004; 20: 1221-1225.

Wimo A, Winblad B, Engedal K, Soininen H, Verhey F, Waldemar G, Wetterholm AL et al. An economic evaluation of donepezil in mild to moderate Alzheimer's disease: results of a 1-year, double-blind, randomized trial. *Dement Geriatr Cogn Disord* 2003; 15: 44-54.

Galantamin vs. Placebo

Brody 2005 [68]¹

Brody H, Corey-Bloom J, Potocnik FC, Truyen L, Gold M, Damaraju CR. Galantamine Prolonged-Release Formulation in the Treatment of Mild to Moderate Alzheimer's Disease. *Dement Geriatr Cogn Disord* 2005; 20: 120-132.

Erkinjuntti 2002 [69]^f

Erkinjuntti T, Kurz A, Gauthier S, Bullock R, Lilienfeld S, Damaraju CV. Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a randomised trial. *Lancet* 2002; 359: 1283-1290.

Raskind 2000 [70]^f

Raskind MA, Peskind ER, Wessel T, Yuan W. Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology* 2000; 54: 2261-2268.

Rockwood 2006 [71]

Rockwood K, Fay S, Song X, MacKnight C, Gorman M, BotV-ISO-TADVI. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *CMAJ* 2006; 174: 1099-1105.

Tariot 2000 [72-74]^f

Tariot PN, Solomon PR, Morris JC, Kershaw P, Lilienfeld S, Ding C. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology* 2000; 54: 2269-2276.

Cummings JL, Schneider L, Tariot PN, Kershaw PR, Yuan W. Reduction of behavioral disturbances and caregiver distress by galantamine in patients with Alzheimer's disease. *Am J Psychiatry* 2004; 161: 532-538.

Galasko D, Kershaw PR, Schneider L, Zhu Y, Tariot PN. Galantamine maintains ability to perform activities of daily living in patients with Alzheimer's disease. *J Am Geriatr Soc* 2004; 52: 1070-1076.

Wilcock 2000 [75]^f

Wilcock GK, Lilienfeld S, Gaens E. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. Galantamine International-1 Study Group. *BMJ* 2000; 321: 1445-1449.

¹ Zu diesen Studien wurden zusätzlich Daten aus folgendem Dokument herangezogen: Shire Pharmaceuticals and Johnson & Johnson. Drugs for the treatment of Alzheimer's Disease. Submission to the National Institute of Clinical Excellence ("NICE-Dossier"). Übermittelt durch die Firma Janssen-Cilag GmbH. 2004.

Rivastigmin vs. Placebo

B304 1998 [48]

Novartis Pharma AG. Final Study Report. Prospective, randomised, multicentre, double-blind, parallel-group, placebo-controlled study to evaluate the efficacy and safety of individual highest well tolerated doses (range 2-12 mg/day) of SDZ ENA 713 given b.i.d. or t.i.d. in patients with mild to moderate probable Alzheimer's disease (Study ENA B304-E-00). 28.01.1998. Unveröffentlicht.

Corey-Bloom 1998 [76,77]

Corey-Bloom J, Anand R, Veach J. A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor in patients with mild to moderate severe Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998; 1: 55-65.

Kumar V, Anand R, Messina J, Hartman R, Veach J. An efficacy and safety analysis of Exelon in Alzheimer's disease patients with concurrent vascular risk factors. *Eur J Neurol* 2000; 7: 159-169.

Forette 1999 [78]

Forette F, Anand R, Gharabawi G. A phase II study in patients with Alzheimer's disease to assess the preliminary efficacy and maximum tolerated dose of rivastigmine (Exelon). *Eur J Neurol* 1999; 6: 423-429.

Rösler 1999 [80,81]

Rösler M, Anand R, Cicin SA, Gauthier S, Agid Y, Dal BP, Stahelin HB et al. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial. *BMJ* 1999; 318: 633-638.

Erkinjuntti T, Skoog I, Lane R, Andrews C. Rivastigmine in patients with Alzheimer's disease and concurrent hypertension. *Int J Clin Pract* 2002; 56: 791-796.

Vergleich verschiedener Cholinesterasehemmer

Galantamin vs. Donepezil

Wilcock 2003 [82]

Wilcock G, Howe I, Coles H, Lilienfeld S, Truyen L, Zhu Y, Bullock R et al. A long-term comparison of galantamine and donepezil in the treatment of Alzheimer's disease. *Drugs Aging* 2003; 20: 777-789.

Rivastigmin vs. Donepezil

Bullock 2005 [83-85]

Bullock R, Touchon J, Bergman H, Gambina G, He Y, Rapatz G, Nagel J et al. Rivastigmine and donepezil treatment in moderate to moderately-severe Alzheimer's disease over a 2-year period. *Curr Med Res Opin* 2005; 21: 1317-1327.

Bullock R, Bergman H, Touchon J, Gambina G, He Y, Nagel J, Lane R. Effect of age on response to rivastigmine or donepezil in patients with Alzheimer's disease. *Curr Med Res Opin* 2006; 22: 483-494.

Touchon J, Bergman H, Bullock R, Rapatz G, Nagel J, Lane R. Response to rivastigmine or donepezil in Alzheimer's patients with symptoms suggestive of concomitant Lewy body pathology. *Curr Med Res Opin* 2006; 22: 49-59.

Fuschillo 2001 [86]

Fuschillo C, La PS, Campana F, Pinto A, De SL. Cognitive deficits in Alzheimer's disease: treatment with acetylcholinesterase inhibitor agents. *Arch Gerontol Geriatr Suppl* 2001; 7: 151-158.

Wang 2001 [89]

Wang Y, Chen Q, Zhang Z, et a. The treatment by using rivastigmine for patients with Alzheimer disease: Results of a multicenter, randomized, open-labeled, controlled clinical trial. *Chin J Neurol* 2001; 34: 210-213.

Rivastigmin vs. Donepezil vs. Galantamin

Cumbo 2005 [90]

Cumbo E. Differential effects of rivastigmine, galantamine and donepezil on behavioral and psychological symptoms in patients with Alzheimer's disease: 18-month, randomized, open-label trial. *Prim Care Comm Psych* 2005; 10: 95-102.

Publikationen zu mehreren Studien (gepoolte Analysen)

Anand R, Messina J, Hartman R. Dose-response effect of rivastigmine in the treatment of Alzheimer's disease. *Int J Geriatr Psychopharmacol* 2000; 2: 68-72. [91]

Burns A, Spiegel R, Quarg P. Efficacy of rivastigmine in subjects with moderately severe Alzheimer's disease. *Int J Geriatr Psychiatry* 2004; 19: 243-249. [92]

Kurz A, Farlow M, Quarg P, Spiegel R. Disease stage in Alzheimer disease and treatment effects of rivastigmine. *Alzheimer Dis Assoc Disord* 2004; 18: 123-128. [93]

Marcusson J, Bullock R, Gauthier S, Kurz A, Schwalen S. Galantamine demonstrates efficacy and safety in elderly patients with Alzheimer disease. *Alzheimer Dis Assoc Disord* 2003; 17 Suppl 3: S86-S91. [94]

Orgogozo JM, Small GW, Hammond G, Van Baelen B, Schwalen S. Effects of galantamine in patients with mild Alzheimer's disease. *Curr Med Res Opin* 2004; 20: 1815-1820. [95]

Potkin SG, Anand R, Hartman R, Veach J, Grossberg G. Impact of Alzheimer's disease and rivastigmine treatment on activities of daily living over the course of mild to moderately severe disease. *Prog Neuropsychopharmacol Biol Psychiatry* 2002; 26: 713-720. [96]

Pratt RD, Perdomo CA, Surick IW, Ieni JR. Donepezil: tolerability and safety in Alzheimer's disease. *Int J Clin Pract* 2002; 56: 710-717. [97]

Sano M, Wilcock GK, Van BB, Kavanagh S. The effects of galantamine treatment on caregiver time in Alzheimer's disease. *Int J Geriatr Psychiatry* 2003; 18: 942-950. [98]

Schneider LS, Anand R, Farlow MR. Systematic review of the efficacy of rivastigmine for patients with Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998; 1: S26-S34. [99]

Whitehead A, Perdomo C, Pratt RD, Birks J, Wilcock GK, Evans JG. Donepezil for the symptomatic treatment of patients with mild to moderate Alzheimer's disease: a meta-analysis of individual patient data from randomised controlled trials. *Int J Geriatr Psychiatry* 2004; 19: 624-633. [100]

Wilkinson DG, Hock C, Farlow M, Van Baelen B, Schwalen S. Galantamine provides broad benefits in patients with 'advanced moderate' Alzheimer's disease (MMSE < or = 12) for up to six months. *Int J Clin Pract* 2002; 56: 509-514. [101]

9 LITERATUR

1. Fratiglioni L, De RD, guero-Torres H. Worldwide prevalence and incidence of dementia. *Drugs Aging* 1999; 15: 365-375.
2. Rocca WA, Hofman A, Brayne C, Breteler MM, Clarke M, Copeland JR et al. The prevalence of vascular dementia in Europe: facts and fragments from 1980-1990 studies. *Ann Neurol* 1991; 30: 817-824.
3. Rocca WA, Hofman A, Brayne C, Breteler MM, Clarke M, Copeland JR et al. Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980-1990 prevalence findings. *Ann Neurol* 1991; 30: 381-390.
4. Petersen RC, Doody RS, Kurz A, Mohs RC, Morris JC, Rabins PV. Current concepts in mild cognitive impairment. *Arch Neurol* 2001; 58: 1985-1992.
5. Erkinjuntti T, Bowler JV, DeCarli CS, Fazekas F, Inzitari D, O'Brien JT et al. Imaging of static brain lesions in vascular dementia: implications for clinical trials. *Alzheimer Dis Assoc Disord* 1999; 13 Suppl 3: S81-S90.
6. Desmond DW, Erkinjuntti T, Sano M, Cummings JL, Bowler JV, Pasquier F et al. The cognitive syndrome of vascular dementia: implications for clinical trials. *Alzheimer Dis Assoc Disord* 1999; 13 Suppl 3: S21-S29.
7. Kalaria R. Similarities between Alzheimer's disease and vascular dementia. *J Neurol Sci* 2002; 203-204: 29-34.
8. Andersen K, Launer LJ, Dewey ME, Letenneur L, Ott A, Copeland JR et al. Gender differences in the incidence of AD and vascular dementia: The EURODEM Studies. *Neurology* 1999; 53: 1992-1997.
9. Hofman A, Rocca WA, Brayne C, Breteler MM, Clarke M, Cooper B et al. The prevalence of dementia in Europe: a collaborative study of 1980-1990 findings. Eurodem Prevalence Research Group. *Int J Epidemiol* 1991; 20: 736-748.
10. Ganguli M, Dodge HH, Shen C, Pandav RS, DeKosky ST. Alzheimer disease and mortality: a 15-year epidemiological study. *Arch Neurol* 2005; 62: 779-784.
11. Waldorff FB, Rishoj S, Waldemar G. Identification and diagnostic evaluation of possible dementia in general practice. A prospective study. *Scand J Prim Health Care* 2005; 23: 221-226.
12. Barker WW, Luis C, Harwood D, Loewenstein D, Bravo M, Ownby R et al. The effect of a memory screening program on the early diagnosis of Alzheimer disease. *Alzheimer Dis Assoc Disord* 2005; 19: 1-7.
13. Butterfield DA, Castegna A, Lauderback CM, Drake J. Evidence that amyloid beta-peptide-induced lipid peroxidation and its sequelae in Alzheimer's disease brain contribute to neuronal death. *Neurobiol Aging* 2002; 23: 655-664.

14. Carter DB, Dunn E, McKinley DD, Stratman NC, Boyle TP, Kuiper SL et al. Human apolipoprotein E4 accelerates beta-amyloid deposition in APPsw transgenic mouse brain. *Ann Neurol* 2001; 50: 468-475.
15. Hsiao K, Chapman P, Nilsen S, Eckman C, Harigaya Y, Younkin S et al. Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science* 1996; 274: 99-102.
16. Löppönen M, Räihä I, Isoaho R, Vahlberg T, Kivelä S.L. Diagnosing cognitive impairment and dementia in primary health care - a more active approach is needed. *Age Ageing* 2003; 32: 606-612.
17. Weltgesundheitsorganisation. Internationale Klassifikation psychischer Störungen - ICD-10 Kapitel V (F): Diagnostische Kriterien für Forschung und Praxis. Bern: Verlag Hans Huber; 2004.
18. American Psychiatric Association. Diagnostisches und Statistisches Manual Psychischer Störungen. DSM-IV. Göttingen: Hogrefe; 1998.
19. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34: 939-944.
20. Cummings JL. Alzheimer's disease. *N Engl J Med* 2004; 351: 56-67.
21. Davies P, Maloney AJ. Selective loss of central cholinergic neurons in Alzheimer's disease. *Lancet* 1976; 2: 1403
22. Perry EK, Perry RH, Blessed G, Tomlinson BE. Necropsy evidence of central cholinergic deficits in senile dementia. *Lancet* 1977; 1: 189
23. Minger SL, Esiri MM, McDonald B, Keene J, Carter J, Hope T et al. Cholinergic deficits contribute to behavioral disturbance in patients with dementia. *Neurology* 2000; 55: 1460-1467.
24. Smith SC, Lamping DL, Banerjee S, Harwoo R, Foley B, Smith P et al. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technol Assess* 2005; 9(10): 1-93, iii-iv.
25. Sands LP, Ferreira P, Stewart AL, Brod M, Yaffe K. What explains differences between dementia patients' and their caregivers' ratings of patients' quality of life? *Am J Geriatr Psychiatry* 2004; 12: 272-280.
26. Mani RB. The evaluation of disease modifying therapies in Alzheimer's disease: A regulatory viewpoint. *Stat Med* 2004; 23: 305-314.

27. Whitehouse PJ. Harmonization of Dementia Drug Guidelines (United States and Europe): a report of the International Working Group for the Harmonization for Dementia Drug Guidelines. *Alzheimer Dis Assoc Disord* 2000; 14 Suppl 1: S119-S122.
28. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984; 141: 1356-1364.
29. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-198.
30. Guy W. ECDU assessment manual for psychopharmacology. Rockville, Md: National Institute of Mental Health; 1976. S. 218-222.
31. Reisberg B, Ferris SH. CIBIC-Plus Interview Guide. East Hannover, N.J.: Sandoz Pharmaceutical Corporation; 1994.
32. Rogers SL, Friedhoff LT, The Donepezil Study Group. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US Multicentre, Randomized, Double-Blind, Placebo-Controlled Trial. *Dementia* 1996; 7: 293-303.
33. Knapp MJ, Knopman DS, Solomon PR, Pendlebury WW, Davis CS, Gracon SI et al. A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. *JAMA* 1994; 271: 985-991.
34. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994; 44: 2308-2314.
35. Oremus M, Perrault A, Demers L, Wolfson C. Review of outcome measurement instruments in Alzheimer's disease drug trials: psychometric properties of global scales. *J Geriatr Psychiatry Neurol* 2000; 13: 197-205.
36. Perrault A, Oremus M, Demers L, Vida S, Wolfson C. Review of outcome measurement instruments in Alzheimer's disease drug trials: psychometric properties of behavior and mood scales. *J Geriatr Psychiatry Neurol* 2000; 13: 181-196.
37. Demers L, Oremus M, Perrault A, Champoux N, Wolfson C. Review of outcome measurement instruments in Alzheimer's disease drug trials: psychometric properties of functional and quality of life scales. *J Geriatr Psychiatry Neurol* 2000; 13: 170-180.
38. Demers L, Oremus M, Perrault A, Wolfson C. Review of outcome measurement instruments in Alzheimer's disease drug trials: introduction. *J Geriatr Psychiatry Neurol* 2000; 13: 161-169.
39. Albert SM, Jacobs DM, Sano M, Marder K, Bell K, Devanand D et al. Longitudinal study of quality of life in people with advanced Alzheimer's disease. *Am J Geriatr Psychiatry* 2001; 9: 160-168.

40. Cotrell V, Schulz R. The perspective of the patient with Alzheimer's disease: a neglected dimension of dementia research. *Gerontologist* 1993; 33: 205-211.
41. Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol* 1988; 41: 1065-1074.
42. The European Agency for the Evaluation of Medicinal Products (EMA). Note for Guidance in Medicinal Products in the Treatment of Alzheimer's Disease. London: EMA. Gelesen unter: <http://www.emea.eu.int/pdfs/human/ewp/055395en.pdf>. 1997.
43. Arzneimittelkommission der deutschen Ärzteschaft (AKdÄ). Demenz: Therapieempfehlungen der Arzneimittelkommission der deutschen Ärzteschaft. Berlin: AKdÄ; 2004.
44. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne Deal. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001; 134: 663-694.
45. Shire Pharmaceuticals and Johnson & Johnson. Drugs for the treatment of Alzheimer's Disease. Submission to the National Institute of Clinical Excellence ("NICE-Dossier"). Übermittelt durch die Firma Janssen Cilag GmbH. 2004.
46. Shire Pharmaceuticals and Johnson & Johnson. Response to the National Institute for Health and Clinical Excellence (NICE) - Request for additional analyses of July 2005. Übermittelt durch die Firma Janssen Cilag GmbH. 2005.
47. Tai CT, Liu CK, Sung SM, Pai MC, Hsu CY. The safety and efficacy of exelon in Alzheimer's patients: a multicentre, randomized, 26-week study in taiwan. *Int J Neuropsychopharmacol* 2000; 3: S356
48. Novartis Pharma AG. Final Study Report. Prospective, randomised, multicentre, double-blind, parallel-group, placebo-controlled study to evaluate the efficacy and safety of individual highest well tolerated doses (range 2-12 mg/day) of SDZ ENA 713 given b.i.d. or t.i.d. in patients with mild to moderate probable Alzheimer's disease (Study ENA B304-E-00). 1998.
49. Birks J. Cholinesterase inhibitors for Alzheimer's disease (Cochrane Review). *Cochrane Database Syst Rev* 2006; Chichester: John Wiley & Sons Ltd.
50. Courtney C, Farrell D, Gray R, Hills R, Lynch L, Sellwood E et al. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet* 2004; 363: 2105-2115.
51. Burns A, Rossor M, Hecker J, Gauthier S, Petit H, Moller HJ et al. The effects of donepezil in Alzheimer's disease - results from a multinational trial. *Dement Geriatr Cogn Disord* 1999; 10: 237-244.

52. Gauthier S, Feldman H, Hecker J, Vellas B, Emir B, Subbiah P et al. Functional, cognitive and behavioral effects of donepezil in patients with moderate Alzheimer's disease. *Curr Med Res Opin* 2002; 18: 347-354.
53. Feldman H, Gauthier S, Hecker J, Vellas B, Subbiah P, Whalen E et al. A 24-week, randomized, double-blind study of donepezil in moderate to severe Alzheimer's disease. *Neurology* 2001; 57: 613-620.
54. Feldman H, Gauthier S, Hecker J, Vellas B, Emir B, Mastey V et al. Efficacy of donepezil on maintenance of activities of daily living in patients with moderate to severe Alzheimer's disease and the effect on caregiver burden. *J Am Geriatr Soc* 2003; 51: 737-744.
55. Homma A, Takeda M, Imai Y, Udaka F, Hasegawa K, Kameyama M et al. Clinical efficacy and safety of donepezil on cognitive and global function in patients with Alzheimer's disease. A 24-week, multicenter, double-blind, placebo-controlled study in Japan. E2020 Study Group. *Dement Geriatr Cogn Disord* 2000; 11: 299-313.
56. Kemp PM, Holmes C, Hoffmann S, Wilkinson S, Zivanovic M, Thom J et al. A randomised placebo controlled study to assess the effects of cholinergic treatment on muscarinic receptors in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2003; 74: 1567-1570.
57. Krishnan KR, Charles HC, Doraiswamy PM, Mintzer J, Weisler R, Yu X et al. Randomized, placebo-controlled trial of the effects of donepezil on neuronal markers and hippocampal volumes in Alzheimer's disease. *Am J Psychiatry* 2003; 160: 2003-2011.
58. Mohs RC, Doody RS, Morris JC, Ieni JR, Rogers SL, Perdomo CA et al. A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology* 2001; 57: 481-488.
59. dos Santos Moraes WA, Rollemberg Poyares D, Guilleminault C, Ramos LR, Bertolucci PHF, Tufik S. The effect of donepezil on sleep and REM sleep EEG in patients with Alzheimer Disease: a double-blind placebo-controlled study. *Sleep* 2006; 29: 199-205.
60. Prasher VP, Huxley A, Haque MS, Down syndrome Ageing Study Group. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Down syndrome and Alzheimer's disease-pilot study. *Int J Geriatr Psychiatry* 2002; 17: 270-278.
61. Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology* 1998; 50: 136-145.
62. Seltzer B, Zolnouri P, Nunez M, Goldman R, Kumar D, Ieni J et al. Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial. *Arch Neurol* 2004; 61: 1852-1856.

63. Tariot PN, Cummings JL, Katz IR, Mintzer J, Perdomo CA, Schwam EM et al. A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting. *J Am Geriatr Soc* 2001; 49: 1590-1599.
64. Tune L, Tiseo PJ, Ieni J, Perdomo C, Pratt RD, Votaw JR et al. Donepezil HCl (E2020) maintains functional brain activity in patients with Alzheimer disease: results of a 24-week, double-blind, placebo-controlled study. *Am J Geriatr Psychiatry* 2003; 11: 169-177.
65. Winblad B, Engedal K, Soininen H, Verhey F, Waldemar G, Wimo A et al. A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology* 2001; 57: 489-495.
66. Wimo A, Winblad B, Shah SN, Chin W, Zhang R, McRae T. Impact of donepezil treatment for Alzheimer's disease on caregiver time. *Curr Med Res Opin* 2004; 20: 1221-1225.
67. Wimo A, Winblad B, Engedal K, Soininen H, Verhey F, Waldemar G et al. An economic evaluation of donepezil in mild to moderate Alzheimer's disease: results of a 1-year, double-blind, randomized trial. *Dement Geriatr Cogn Disord* 2003; 15: 44-54.
68. Brodaty H, Corey-Bloom J, Potocnik FC, Truyen L, Gold M, Damaraju CR. Galantamine Prolonged-Release Formulation in the Treatment of Mild to Moderate Alzheimer's Disease. *Dement Geriatr Cogn Disord* 2005; 20: 120-132.
69. Erkinjuntti T, Kurz A, Gauthier S, Bullock R, Lilienfeld S, Damaraju CV. Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a randomised trial. *Lancet* 2002; 359: 1283-1290.
70. Raskind MA, Peskind ER, Wessel T, Yuan W, The Galantamine USA-1 Study Group. Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. *Neurology* 2000; 54: 2261-2268.
71. Rockwood K, Fay S, Song X, MacKnight C, Gorman M, on behalf of the Video-Imaging Synthesis of Treating Alzheimer's Disease (VISTA) Investigators. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *CMAJ* 2006; 174: 1099-1105.
72. Tariot PN, Solomon PR, Morris JC, Kershaw P, Lilienfeld S, Ding C. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology* 2000; 54: 2269-2276.
73. Cummings JL, Schneider L, Tariot PN, Kershaw PR, Yuan W. Reduction of behavioral disturbances and caregiver distress by galantamine in patients with Alzheimer's disease. *Am J Psychiatry* 2004; 161: 532-538.

74. Galasko D, Kershaw PR, Schneider L, Zhu Y, Tariot PN. Galantamine maintains ability to perform activities of daily living in patients with Alzheimer's disease. *J Am Geriatr Soc* 2004; 52: 1070-1076.
75. Wilcock GK, Lilienfeld S, Gaens E. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. Galantamine International-1 Study Group. *BMJ* 2000; 321: 1445-1449.
76. Corey-Bloom J, Anand R, Veach J. A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor in patients with mild to moderate severe Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998; 1: 55-65.
77. Kumar V, Anand R, Messina J, Hartman R, Veach J. An efficacy and safety analysis of Exelon in Alzheimer's disease patients with concurrent vascular risk factors. *Eur J Neurol* 2000; 7: 159-169.
78. Forette F, Anand R, Gharabawi G. A phase II study in patients with Alzheimer's disease to assess the preliminary efficacy and maximum tolerated dose of rivastigmine (Exelon). *Eur J Neurol* 1999; 6: 423-429.
79. Karaman Y, Erdogan F, Koseoglu E, Turan T, Ersoy AO. A 12-month study of the efficacy of rivastigmine in patients with advanced moderate Alzheimer's disease. *Dement Geriatr Cogn Disord* 2005; 19: 51-56.
80. Rösler M, Anand R, Cicin SA, Gauthier S, Agid Y, Dal BP et al. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial. *BMJ* 1999; 318: 633-638.
81. Erkinjuntti T, Skoog I, Lane R, Andrews C. Rivastigmine in patients with Alzheimer's disease and concurrent hypertension. *Int J Clin Pract* 2002; 56: 791-796.
82. Wilcock G, Howe I, Coles H, Lilienfeld S, Truyen L, Zhu Y et al. A long-term comparison of galantamine and donepezil in the treatment of Alzheimer's disease. *Drugs Aging* 2003; 20: 777-789.
83. Bullock R, Touchon J, Bergman H, Gambina G, He Y, Rapatz G et al. Rivastigmine and donepezil treatment in moderate to moderately-severe Alzheimer's disease over a 2-year period. *Curr Med Res Opin* 2005; 21: 1317-1327.
84. Bullock R, Bergman H, Touchon J, Gambina G, He Y, Nagel J et al. Effect of age on response to rivastigmine or donepezil in patients with Alzheimer's disease. *Curr Med Res Opin* 2006; 22: 483-494.
85. Touchon J, Bergman H, Bullock R, Rapatz G, Nagel J, Lane R. Response to rivastigmine or donepezil in Alzheimer's patients with symptoms suggestive of concomitant Lewy body pathology. *Curr Med Res Opin* 2006; 22: 49-59.

86. Fuschillo C, La PS, Campana F, Pinto A, De SL. Cognitive deficits in Alzheimer's disease: treatment with acetylcholinesterase inhibitor agents. *Arch Gerontol Geriatr Suppl* 2001; 7: 151-158.
87. Kim JM, Shin IS, Yoon JS. Correlates of dropout, efficacy, and adverse events in treatment with acetylcholinesterase inhibitors in Korean patients with Alzheimer's disease. *Int Psychogeriatr* 2002; 14: 187-195.
88. Thomas A, Iacono D, Bonanni L, D'Andreamatteo G, Onofri M. Donepezil, rivastigmine, and vitamin E in Alzheimer disease: a combined P300 event-related potentials/neuropsychologic evaluation over 6 months. *Clin Neuropharmacol* 2001; 24: 31-42.
89. Wang Y, Chen Q, Zhang Z, et al. The treatment by using rivastigmine for patients with Alzheimer disease: Results of a multicenter, randomized, open-labeled, controlled clinical trial. *Chin J Neurol* 2001; 34: 210-213.
90. Cumbo E. Differential effects of rivastigmine, galantamine and donepezil on behavioral and psychological symptoms in patients with Alzheimer's disease: 18-month, randomized, open-label trial. *Prim Care Comm Psych* 2005; 10: 95-102.
91. Anand R, Messina J, Hartman R. Dose-response effect of rivastigmine in the treatment of Alzheimer's disease. *Int J Geriatr Psychopharmacol* 2000; 2: 68-72.
92. Burns A, Spiegel R, Quarg P. Efficacy of rivastigmine in subjects with moderately severe Alzheimer's disease. *Int J Geriatr Psychiatry* 2004; 19: 243-249.
93. Kurz A, Farlow M, Quarg P, Spiegel R. Disease stage in Alzheimer disease and treatment effects of rivastigmine. *Alzheimer Dis Assoc Disord* 2004; 18: 123-128.
94. Marcusson J, Bullock R, Gauthier S, Kurz A, Schwalen S. Galantamine demonstrates efficacy and safety in elderly patients with Alzheimer disease. *Alzheimer Dis Assoc Disord* 2003; 17 Suppl 3: S86-S91.
95. Orgogozo JM, Small GW, Hammond G, Van Baelen B, Schwalen S. Effects of galantamine in patients with mild Alzheimer's disease. *Curr Med Res Opin* 2004; 20: 1815-1820.
96. Potkin SG, Anand R, Hartman R, Veach J, Grossberg G. Impact of Alzheimer's disease and rivastigmine treatment on activities of daily living over the course of mild to moderately severe disease. *Prog Neuropsychopharmacol Biol Psychiatry* 2002; 26: 713-720.
97. Pratt RD, Perdomo CA, Surick IW, Ieni JR. Donepezil: tolerability and safety in Alzheimer's disease. *Int J Clin Pract* 2002; 56: 710-717.
98. Sano M, Wilcock GK, Van BB, Kavanagh S. The effects of galantamine treatment on caregiver time in Alzheimer's disease. *Int J Geriatr Psychiatry* 2003; 18: 942-950.

99. Schneider LS, Anand R, Farlow MR. Systematic review of the efficacy of rivastigmine for patients with Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998; 1: S26-S34.
100. Whitehead A, Perdomo C, Pratt RD, Birks J, Wilcock GK, Evans JG. Donepezil for the symptomatic treatment of patients with mild to moderate Alzheimer's disease: a meta-analysis of individual patient data from randomised controlled trials. *Int J Geriatr Psychiatry* 2004; 19: 624-633.
101. Wilkinson DG, Hock C, Farlow M, Van Baelen B, Schwalen S. Galantamine provides broad benefits in patients with 'advanced moderate' Alzheimer's disease (MMSE \leq 12) for up to six months. *Int J Clin Pract* 2002; 56: 509-514.
102. Eisai GmbH, Pfizer Pharma GmbH. Aricept 5 mg/10 mg: Fachinformation. Berlin: Rote Liste Service GmbH; 2005.
103. Bullock R, Erkinjuntti T, Lilienfeld S, GAL INT. Management of patients with Alzheimer's disease plus cerebrovascular disease: 12-month treatment with galantamine. *Dement Geriatr Cogn Disord* 2004; 17: 29-34.
104. Blau TH. Quality of life, social indicators and criteria of change. *Prof Psychol* 1977; 464-473.
105. Rogers SL, Doody RS, Mohs R, Friedhoff LT, and the Donepezil Study Group. Donepezil improves cognition and global function in Alzheimer Disease: A 15-week, double-blind, placebo-controlled study. *Arch Intern Med* 1998; 158: 1021-1031.
106. Galasko D, Bennett D, Sano M, Ernesto C, Thomas R, Grundman M et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord* 1997; 11 Suppl 2: S33-S39.
107. Janssen-Cilag GmbH. Reminyl® 4 mg/-8 mg/-12 mg Filmtabletten: Fachinformation. Berlin: Rote Liste Service GmbH; 2005.
108. Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E et al. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. *Health Technol Assess* 2006; 10: 1-378.
109. Pankiewicz O, Pelkowska A, Landa K. Efficacy and safety of rivastigmine. HTA Consulting; ohne Jahr.
110. Birks J, Harvey RJ. Donepezil for dementia due to Alzheimer's disease (Cochrane Review). *Cochrane Database Syst Rev* 2006; Chichester: John Wiley & Sons Ltd.
111. Sink KM, Holden KF, Yaffe K. Pharmacological treatment of neuropsychiatric symptoms of dementia. *JAMA* 2005; 293: 596-608.
112. Perras C, Shukla VK, Lessard CL, Skidmore B, Bergman H, Gauthier S. Cholinesterase inhibitors for Alzheimer's Disease: a systematic review of randomized controlled trials. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2005.

113. Hansen RA, Gartlehner G, Kaufer DJ, Lohr KN, Randolph LC, Carey T. Drug class review on Alzheimer's drugs. Chapel Hill: Oregon Health & Science Unit; 2005.
114. Kaduszkiewicz H, Zimmermann T, Beck-Bornholdt HP, van den Bussche H. Cholinesterase inhibitors for patients with Alzheimer's disease: systematic review of randomised clinical trials. *BMJ* 2005; 331: 321-327.
115. Kaduszkiewicz H, Beck-Bornholdt HP, van den Bussche H, Zimmermann T. Fragliche Evidenz für den Einsatz des Cholinesterasehemmers Donepezil bei Alzheimer-Demenz - eine systematische Übersichtsarbeit. *Fortschr Neurol Psychiatr* 2004; 72: 557-563.
116. Windeler J. Stellungnahme zur Publikation "Fragliche Evidenz für den Einsatz des Cholinesterasehemmers Donepezil bei der Alzheimer-Demenz - eine systematische Übersichtsarbeit (Fortschr Neurol Psychiatr 72 [2004] 557-563)". Essen: MDS; 2005.
117. Kaduszkiewicz H, Beck-Bornholdt HP, van den Bussche H, Zimmermann T. Fragliche Evidenz für den Einsatz des Cholinesterasehemmers Donepezil bei Alzheimer-Demenz: eine systematische Übersichtsarbeit. *Fortschr Neurol Psychiatr* 2004; 72(10): 557-63. Antwort der Autoren zur Stellungnahme von Windeler J vom 29. Dez. 2004. 18. Aug. 2005. Gelesen unter: <http://www.uke.uni-hamburg.de/institute/allgemeinmedizin/downloads/institut-allgemeinmedizin/StellungnahmezuWindeler.pdf>. 2005.
118. Farlow M, Potkin S, Koumaras B, Veach J, Mirski D. Analysis of outcome in retrieved dropout patients in a rivastigmine vs placebo, 26-week, Alzheimer disease trial. *Arch Neurol* 2003; 60: 843-848.
119. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001; 20: 3931-3946.
120. Kaiser T, Florack C, Franz H, Sawicki P. Donepezil bei Patienten mit Alzheimer-Demenz - die AD2000-Studie. *Med Klin* 2005; 100: 157-160.
121. Thomas S, Windeler J. Methodische Kritikpunkte zur Studie "Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double blind trial (Lancet 363 [2004] 2105-2115)". Essen: Medizinischer Dienst der Spitzenverbände der Krankenkassen; 2005.
122. Holmes C, Wilkinson D, Dean C, Vethanayagam S, Olivieri S, Langley A et al. The efficacy of donepezil in the treatment of neuropsychiatric symptoms in Alzheimer disease. *Neurology* 2004; 63: 214-219.
123. Winblad B, Kilander L, Eriksson S, Minthon L, Batsman S, Wetterholm AL et al. Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled study. *Lancet* 2006; (367): 1057-1065.
124. Trinh NH, Hoblyn J, Mohanty S, Yaffe K. Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: a meta-analysis. *JAMA* 2003; 289: 210-216.

125. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 [updated September 2006]. In: The Cochrane Library, Issue 4, 2006. Chichester, UK: John Wiley & Sons, Ltd: 2006.
126. Scottish Intercollegiate Guidelines Network Grading Review Group. Management of patients with dementia: A national clinical guideline. Edinburgh: SIGN; 2006.
127. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995; 273: 408-412.
128. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet 2002; 359: 781-785.
129. The European Agency for the Evaluation of Medicinal Products (EMA). Points to consider on missing data; Gelesen unter: <http://www.emea.eu.int/pdfs/human/ewp/177699EN.pdf>. 2001.
130. Wolfson C, Moride Y, Perrault A, Momoli F, Demers L, Oremus M. Drug treatments for Alzheimer's disease. II. A review of outcome measures in clinical trials. Ottawa; 2000.
131. Standish TIM, Molloy DW, Bédard M, Layne EC, Murray EA, Strang D. Improved reliability of the Standardized Alzheimer's Disease Assessment Scale (SADAS) compared with the Alzheimer's Disease Assessment Scale. J Am Geriatr Soc 1996; 44: 712-716.
132. Nadkarni NK, Black SE. Cognitive Outcomes. In: Rockwood K, Gauthier S (Ed). Trial designs and outcomes in dementia therapeutic research. London: Taylor & Francis; 2006.
133. FDA. Peripheral and Central Nervous System Drugs Advisory Committee Meeting. Rockville, MD: Department of Health and Human Services; Public Health Service; 1989.
134. Kramer-Ginsberg E, Mohs RC, Aryan M, Lobel D, Silverman J, Davidson M et al. Clinical predictors of course for Alzheimer Patients in a longitudinal study: a preliminary report. Psychopharmacol Bull 1988; 24: 458-462.
135. Stern RG, Mohs RC, Davidson M, Schmeidler J, Silverman J, Kramer-Ginsberg E et al. A longitudinal study of Alzheimer's disease: measurement, rate, and predictors of cognitive deterioration. Am J Psychiatry 1994; 151: 390-396.
136. Torfs K, Feldman H. 12-month decline in cognitive and daily function in patients with mild-to-moderate Alzheimer's disease: two randomized, placebo-controlled studies. Neurobiol Aging 2000; 21 (Suppl 1): 242-243.
137. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977.

138. Sawicki P. Nutzenbewertung in der Medizin - Stand Oktober 2006. *Gesundh ökon Qual manag* 2006; 11: 365-367.
139. Mazza M, Capuano A, Bria P, Mazza S. Ginkgo biloba and donepezil: a comparison in the treatment of Alzheimer's dementia in a randomized placebo-controlled double-blind study. *Eur J Neurol* 2006; 13: 981-985.
140. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Ginkgohaltige Präparate bei Alzheimer Demenz. Vorbericht A05-19B. Köln: IQWiG; 2006.
141. Graff MJL, Vernooij-Dassen MJM, Thijssen M, Dekker J, Hoefnagels WHL, Olde Rikkert MGM. Community based occupational therapy for patients with dementia and their care givers: randomised controlled trial. *BMJ* 2006; 333(7580): 1196-1299.

ANHANG A: SUCHSTRATEGIE

Primärrecherche

Datenbanken MEDLINE 66 und Pre-MEDLINE (Suchdatum: 13.04.2005, Suchmaske: Ovid)

#	Abfrage	Treffer
1	exp Alzheimer Disease	34415
2	dement\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	49046
3	alzheimer\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	44711
4	((cognit\$ or memory\$ or mental\$) and (decline\$ or impair\$ or los\$ or deteriorate\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	52641
5	1 or 2 or 3 or 4	119445
6	randomi?ed-controlled-trial.pt.	197725
7	controlled-clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	3893
8	randomi?ed-controlled-trials.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	39655
9	random-allocation.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	52978
10	double-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	80556
11	single-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	8724
12	clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	32003
13	clin\$ near trial\$.pt.	0
14	clin\$ trial\$.pt.	0
15	clin\$-trial\$.pt.	0
16	(clin\$ adj trial\$).pt.	0
17	(clin\$ adj trial\$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	186734
18	(clin\$ adj trial\$).pt.	0
19	(clin\$ adj trial\$).ti,ab.	83731
20	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	106636
21	placebo\$.ti,ab.	87216
22	random\$.ti,ab.	302049
23	research-design.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	46392
24	6 or 7 or 8 or 9 or 10 or 11 or 12.mp. or 17 or 19 or 20 or 21 or 22 or 23 [mp=title, original title, abstract, name of substance word, subject heading word]	602212
25	5 and 24	9744
26	(HIV\$ or AIDS or stroke or diabet# or heart or epilep# or schizophre#).mp. or child#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	869555
27	(normal control# or healthy control# or healthy volunteer#).mp. or normal volunteer#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	78924

	heading word]	
28	26 or 27	940842
29	25 not 28	8426
30	(TG=animal not (TG=human and TG=animal)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	0
31	donepezil.rn.	715
32	aricept.ti.ab.	53
33	exp galantamine/	522
34	Galantamin#.ti.ab.	242
35	galanthamin#.ti.ab.	227
36	nivalin#.ti.ab.	23
37	Lycoremin#.ti.ab.	0
38	reminyl.ti.ab.	21
39	rivastigmin#ti.ab.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	0
40	rivastigmin#.ti.ab.	313
41	exelon#.ti.ab.	0
42	31 or 32 or 33 or 34 or 35 or 36 or 38 Or 40.mp. or 41 [mp=title, original title, abstract, name of substance word, subject heading word]	1296
43	42 and 29	324

Datenbank EMBASE 88 (Suchdatum: 13.04.2005, Suchmaske: Ovid)^m

#	Abfrage	Treffer
1	exp Alzheimer Disease	
2	dement\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
3	alzheimer\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
4	((cognit\$ or memory\$ or mental\$) and (decline\$ or impair\$ or los\$ or deteriorate\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
5	1 or 2 or 3 or 4	
6	randomi?ed-controlled-trial.pt.	
7	controlled-clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
8	randomi?ed-controlled-trials.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
9	random-allocation.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
10	double-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
11	single-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
12	clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	

^m Zu dieser Recherche wurden die Treffer zu den einzelnen Suchschritten nicht dokumentiert.

13	clin\$ near trial\$.pt.	
14	clin\$ trial\$.pt.	
15	clin\$-trial\$.pt.	
16	(clin\$ adj trial\$).pt.	
17	(clin\$ adj trial\$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
18	(clin\$ adj trial\$).pt.	
19	(clin\$ adj trial\$).ti,ab.	
20	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
21	placebo\$.ti,ab.	
22	random\$.ti,ab.	
23	research-design.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
24	6 or 7 or 8 or 9 or 10 or 11 or 12.mp. or 17 or 19 or 20 or 21 or 22 or 23 [mp=title, original title, abstract, name of substance word, subject heading word]	
25	5 and 24	
26	(HIV\$ or AIDS or stroke or diabet# or heart or epilep# or schizophre#).mp. or child#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	
27	(normal control# or healthy control# or healthy volunteer#).mp. or normal volunteer#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	
28	26 or 27	
29	25 not 28	
30	(TG=animal not (TG=human and TG=animal)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
31	donepezil.rn.	
32	aricept.ti,ab.	
33	exp galantamine/	
34	Galantamin#.ti,ab.	
35	galanthamin#.ti,ab.	
36	nivalin#.ti,ab.	
37	Lycoremin#.ti,ab.	
38	reminyl.ti,ab.	
39	rivastigmin#.ti,ab.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
40	rivastigmin#.ti,ab.	
41	exelon#.ti,ab.	
42	31 or 32 or 33 or 34 or 35 or 36 or 38 or 40.mp. or 41 [mp=title, original title, abstract, name of substance word, subject heading word]	
43	42 and 29	255

Datenbank Cochrane Library (CLIB) (Suchdatum: 14.04.2005)

#	Abfrage	Treffer
1	<u>MeSH descriptor Alzheimer Disease explode all trees in MeSH products</u>	1140
2	<u>alzheimer* in All Fields, from 1800 to 2005 in CENTRAL</u>	2392
3	<u>alzheimer* in All Fields, from 1800 to 2005 in CENTRAL</u>	2392
4	<u>dement* in All Fields, from 1800 to 2005 in CENTRAL</u>	2471
5	<u>(cognit* or memory* or mental*) and (decline* or impair* or los* or deteriorate*) in All Fields, from 1800 to 2005 in CENTRAL</u>	3353
6	<u>(#1 OR #2 OR #4 OR #5)</u>	7919
7	<u>HIV*:ti or aids:ti or stroke:ti or diabet*:ti or heart:ti or epilep*:ti or schizophre*:ti or child*:ti OR Parkinson* :TI in All Fields, from 1800 to 2005 in CENTRAL</u>	2
8	<u>HIV* or aids or stroke or diabet* or heart or epilep* or schizophre* or child* OR Parkinson* in Record Title, from 1800 to 2005 in CENTRAL</u>	44002
9	<u>normal next control* OR healthy next control* OR healthy next volunteer* OR normal next volunteer* in Record Title, from 1800 to 2005 in CENTRAL</u>	3734
10	<u>(8 OR 9)</u>	51139
11	<u>(6 AND NOT 10)</u>	7058
12	<u>donepezil in All Fields, from 1800 to 2005 in CENTRAL</u>	270
13	<u>aricept in All Fields, from 1800 to 2005 in CENTRAL</u>	12
14	<u>galantamin* in All Fields, from 1800 to 2005 in CENTRAL</u>	110
15	<u>galanthamin* in All Fields, from 1800 to 2005 in CENTRAL</u>	21
16	<u>nivalin* in All Fields, from 1800 to 2005 in CENTRAL</u>	2
17	<u>lycoremin* in All Fields, from 1800 to 2005 in CENTRAL</u>	0
18	<u>reminyl in All Fields, from 1800 to 2005 in CENTRAL</u>	11
19	<u>Rivastigmin* in All Fields, from 1800 to 2005 in CENTRAL</u>	33
20	<u>exelon* in All Fields, from 1800 to 2005 in CENTRAL</u>	27
21	<u>(12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20)</u>	503
22	<u>(11 AND 21)</u>	347

Datenbank CHID (Suchdatum: 25.04.2005)

Suche anhand der Substanznamen gemäß der Benutzeroberfläche. Treffer: **54**

Nachrecherche 1

Datenbanken: MEDLINE 66 und Pre-MEDLINE (Suchmaske: Ovid, Suchdatum: 03.11.2005)

#	Abfrage	Treffer
1	exp Alzheimer Disease	36708
2	dement\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	52206
3	alzheimer\$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	49608
4	((cognit\$ or memory\$ or mental\$) and (decline\$ or impair\$ or los\$ or deteriorate\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	58390
5	1 or 2 or 3 or 4	130648
6	randomi?ed-controlled-trial.pt.	207743
7	controlled-clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	4265
8	randomi?ed-controlled-trials.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	44316
9	random-allocation.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	54519
10	double-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	83751
11	single-blind-method.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	9397
12	clinical-trial.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	35248
13	clin\$ near trial\$.pt.	200265
14	clin\$ trial\$.pt.	93852
15	clin\$-trial\$.pt.	112804
16	(clin\$ adj trial\$.pt.	0
17	(clin\$ adj trial\$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	200265
18	(clin\$ adj trial\$.pt.	0
19	(clin\$ adj trial\$.ti,ab.	93852
20	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	112804
21	placebo\$.ti,ab.	93570
22	random\$.ti,ab.	335625
23	research-design.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	49373
24	6 or 7 or 8 or 9 or 10 or 11 or 12.mp. or 17 or 19 or 20 or 21 or 22 or 23 [mp=title, original title, abstract, name of substance word, subject heading word]	654081
25	5 and 24	10792
26	(HIV\$ or AIDS or stroke or diabet# or heart or epilep# or schizophre#).mp. or child#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	913602
27	(normal control# or healthy control# or healthy volunteer#).mp. or normal volunteer#.ti. [mp=title, original title, abstract, name of substance word, subject heading word]	85025
28	26 or 27	990487

29	25 not 28	9303
30	(TG=animal not (TG=human and TG=animal)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	0
31	donepezil.rn.	794
32	aricept.ti,ab.	62
33	exp galantamine/	567
34	Galantamin#.ti,ab.	303
35	galanthamin#.ti,ab.	249
36	nivalin#.ti,ab.	23
37	Lycoremin#.ti,ab.	0
38	reminyl.ti,ab.	26
39	rivastigmin#.ti,ab.mp. [mp=title, original title, abstract, name of substance word, subject heading word]	0
40	rivastigmin#.ti,ab.	390
41	exelon#.ti,ab.	0
42	31 or 32 or 33 or 34 or 35 or 36 or 38 Or 40.mp. or 41 [mp=title, original title, abstract, name of substance word, subject heading word]	1458
43	42 and 29	369
44	limit 43 to yr="2005 - 2006"	34
45	"2005".ed,yr.	513026
46	43 and 45	34

Datenbank EMBASE 88 (Suchdatum: 03.11.2005, Suchmaske: Ovid)

#	Abfrage	Treffer
1	exp Alzheimer Disease/	45187
2	dement\$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	41257
3	alzheimer\$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	50468
4	((cognit\$ or memory\$ or mental\$) and (decline\$ or impair\$ or los\$ or deteriorate\$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	56430
5	1 or 2 or 3 or 4	119027
6	randomi?ed-controlled-trial.pt.	0
7	controlled-clinical-trial.mp.	3797
8	randomi?ed-controlled-trials.mp.	7098
9	random-allocation.mp.	509
10	double-blind-method.mp.	181
11	single-blind-method.mp.	26
12	clinical-trial.mp.	374849
13	(clin\$ adj trial\$.mp.	405792
14	(clin\$ adj trial\$.ti,ab.	84184
15	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).mp.	96186
16	(clin\$ adj trial\$.pt.	0
17	(clin\$ adj trial\$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	405792
18	(clin\$ adj trial\$.pt.	0
19	(clin\$ adj trial\$.ti,ab.	84184

20	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	96186
21	placebo\$.ti,ab.	88076
22	random\$.ti,ab.	286857
23	research-design.mp.	5708
24	6 or 7 or 8 or 9 or 10 or 11 or 12.mp. or 17 or 19 or 20 or 21 or 22 or 23	626965
25	5 and 24	13615
26	(HIV\$ or AIDS or stroke or diabet# or heart or epilep# or schizophre#).mp.	823909
27	(normal control# or healthy control# or healthy volunteer#).mp. or normal volunteer#.ti.	77408
28	26 or 27	893027
29	25 not 28	11009
30	(TG=animal not (TG=human and TG=animal)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]	0
31	donepezil.rn.	0
32	aricept.ti,ab.	69
33	exp GALANTAMINE/	1560
34	galantamin#.ti,ab.	343
35	galanthamin#.ti,ab.	207
36	nivalin#.ti,ab.	1
37	lycoremin#.ti,ab.	0
38	reminyl.ti,ab.	36
39	rivastigmin#.ti,ab.mp.	0
40	rivastigmin#.ti,ab.	436
41	exelon#.ti,ab.	0
42	31 or 32 or 33 or 34 or 35 or 36 or 38 or 40.mp. or 41	1649
43	42 and 29	587
44	limit 43 to yr="2005 - 2006"	78
45	"2005".em,yr.	382835
46	43 and 45	78

Datenbank Cochrane CENTRAL (CCTR) (Suchdatum: 03.11.2005)

#	Abfrage	Treffer
#1	<u>MeSH descriptor Alzheimer Disease explode all trees in MeSH products</u>	1231
#2	<u>alzheimer* in All Fields, from 2004 to 2005 in CENTRAL</u>	264
#3	<u>dement* in All Fields, from 2004 to 2005 in CENTRAL</u>	235
#4	<u>(cognit* or memory* or mental*) and (decline* or impair* or los* or deteriorate*) in All Fields, from 2004 to 2005 in CENTRAL</u>	498
#5	<u>(#1 OR #2 OR #3 OR #4)</u>	2600
#6	<u>HIV* or aids or stroke or diabet* or heart or epilep* or schizophre* or child* or parkinson* in Record Title, from 2004 to 2005 in CENTRAL</u>	3628
#7	<u>normal next control* or healthy next control* or healthy next volunteer* or normal next volunteer* in Record Title, from 2004 to 2005 in CENTRAL</u>	210
#8	<u>(#6 OR #7)</u>	4805
#9	<u>(#5 AND NOT #8)</u>	2373
#10	<u>donepezil in All Fields, from 2004 to 2005 in CENTRAL</u>	65
#11	<u>aricept in All Fields, from 2004 to 2005 in CENTRAL</u>	2

#12	<u>galantamin* in All Fields, from 2004 to 2005 in CENTRAL</u>	38
#13	<u>galanthamin* in All Fields, from 2004 to 2005 in CENTRAL</u>	1
#14	<u>nivalin* in All Fields, from 2004 to 2005 in CENTRAL</u>	0
#15	<u>lycoremin* in All Fields, from 2004 to 2005 in CENTRAL</u>	0
#16	<u>reminyl in All Fields, from 2004 to 2005 in CENTRAL</u>	2
#17	<u>Rivastigmin* in All Fields, from 2004 to 2005 in CENTRAL</u>	9
#18	<u>exelon* in All Fields, from 2004 to 2005 in CENTRAL</u>	1
#19	<u>(#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18)</u>	140
#20	<u>(#9 AND #19)</u>	103

Cochrane Reviews [9] | [DARE](#) [5] | [CENTRAL](#) [76] | Methodology Reviews [0] | CMR [0] | [HTA](#) [3] | [NHS EED](#) [9] | [About](#) [1]

Datenbank CHID (Suchdatum: 03.11.2005)

Suche anhand der Substanznamen gemäß der Benutzeroberfläche. Treffer: **1**

Nachrecherche 2

Datenbanken: Ovid MEDLINE(R) In-Process, Other Non-Indexed Citations, Ovid MEDLINE(R), Ovid MEDLINE(R) Daily Update (Suchmaske: Ovid, Suchdatum: 12.06.2006)

#	Abfrage	Treffer
1	exp ALZHEIMER DISEASE/	39086
2	alzheimer\$.ti,ab,ot.	46136
3	*DEMENTIA/	16991
4	(dementia or dement or demenz or demenc\$).ti,ot.	17072
5	or/1-4	66764
6	donepezil.ti,ab,ot,nm.	1175
7	aricept.ti,ab,ot,nm.	72
8	exp GALANTAMINE/	634
9	galantamin\$.ti,ab,ot,nm,rn.	739
10	galanthamin\$.ti,ab,ot,nm,rn.	273
11	nivalin\$.ti,ab,ot,nm,rn.	51
12	lycoramin\$.ti,ab,ot,nm,rn.	5
13	reminyl\$.ti,ab,ot,nm,rn.	32
14	rivastigmin\$.ti,ab,ot,nm,rn.	547
15	exelon\$.ti,ab,ot,nm,rn.	29
16	or/6-15	2080
17	exp RANDOM ALLOCATION/	57438
18	random\$.ti,ot.	54898
19	prospectiv\$.ti,ot.	42256
20	exp RANDOMIZED CONTROLLED TRIALS/	45681
21	randomized controlled trial.pt.	226483
22	exp CONTROLLED CLINICAL TRIALS/	48478
23	controlled clinical trial.pt.	73383
24	exp DOUBLE-BLIND METHOD/	88351
25	((single or double or triple) adj5 (mask\$ or blind\$) adj5 (study or trial or method)).ti,ab,ot,sh.	61082
26	or/17-25	426956
27	5 and 16 and 26	316

Datenbank: EMBASE 88 (Suchdatum: 12.06.2006, Suchmaske: Ovid)

#	Abfrage	Treffer
1	exp ALZHEIMER DISEASE/	48042
2	alzheimer\$.ti,ab,ot.	42924
3	*DEMENTIA/ or exp SENILE DEMENTIA/	18743
4	(dementia or dement or demenz or demenc\$).ti,ot.	15480
5	or/1-4	66663
6	exp DONEPEZIL/	3068
7	donepezil.ti,ab,ot,tn.	1030
8	aricept.ti,ab,ot,tn.	745
9	exp GALANTAMINE/	1811
10	galantamin\$.ti,ab,ot,tn.	410
11	galanthamin\$.ti,ab,ot,tn.	230
12	nivalin\$.ti,ab,ot,tn.	61
13	lycoramin\$.ti,ab,ot,tn.	14
14	reminyl\$.ti,ab,ot,tn.	390
15	exp RIVASTIGMINE/	1731
16	rivastigmin\$.ti,ab,ot,tn.	496
17	exelon\$.ti,ab,ot,tn.	481
18	or/6-17	4238
19	exp RANDOMIZATION/	19225
20	exp RANDOMIZED CONTROLLED TRIAL/	106163
21	random\$.ti,ot.	46724
22	prospectiv\$.ti,ot.	35371
23	exp DOUBLE-BLIND PROCEDURE/	59969
24	((single or double or triple) adj5 (mask\$ or blind\$) adj5 (study or trial or method)).ti,ab,ot,sh.	58088
25	or/19-24	208444
26	5 and 18 and 25	340

Datenbank Cochrane CENTRAL (CCTR) (Suchdatum: 12.06.2006)

#	Abfrage	Treffer
#1	MeSH descriptor Alzheimer Disease explode all trees in MeSH products	1309
#2	alzheimer* in Title, Abstract or Keywords in all products	2866
#3	MeSH descriptor Dementia , this term only in MeSH products	677
#4	demen* in Record Title in all products	1794
#5	(#1 OR #2 OR #3 OR #4)	4254
#6	donepezil in All Fields in all products	426
#7	aricept in All Fields in all products	23
#8	galantamin* in All Fields in all products	206
#9	galanthamin* in All Fields in all products	32
#10	nivalin* in All Fields in all products	2
#11	lycoramin* in All Fields in all products	0
#12	reminyl in All Fields in all products	21
#13	rivastigmin* in All Fields in all products	181
#14	exelon* in All Fields in all products	40
#15	(#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14)	722
#16	(#5 AND #15)	564

Cochrane Reviews [24] | Other Reviews [13] | **Clinical Trials [477]** | Methods Reviews [0] | Methods Studies [1] | Technology Assessments [15] | Economic Evaluations [33] | Cochrane Groups [1]

Datenbank CHID: Die Datenbank war zum Zeitpunkt der zweiten Nachrecherche nicht mehr verfügbar.

ANHANG B: LISTE DER IM VOLLTEXT ÜBERPRÜFTEN, ABER AUSGESCHLOSSENEN STUDIEN MIT AUSSCHLUSSGRÜNDEN

Nicht E1: vorgegebene Indikation nicht erfüllt (5)

Ballard C, Margallo-Lana M, Juszczak E, Douglas S, Swann A, Thomas A, et al. Quetiapine and rivastigmine and cognitive decline in Alzheimer's disease: randomised double blind placebo controlled trial. *BMJ* 2005; 330: 857-858.

Lopez-Pousa S. Pilot, multicenter, randomized, double-blind, controlled, parallel efficacy and safety study of rivastigmine vs placebo in the treatment of cognitive and non-cognitive symptoms in patients with moderate-to-severe Alzheimer's disease. *IFPMA Register* 2005.

Passmore AP, Bayer AJ, Steinhagen-Thiessen E. Cognitive, global, and functional benefits of donepezil in Alzheimer's disease and vascular dementia: results from large-scale clinical trials. *J Neurol Sci* 2005; 229-230: 141-146.

Salloway S, Ferris S, Kluger A, Goldman R, Griesing T, Kumar D, et al. Efficacy of donepezil in mild cognitive impairment: a randomized placebo-controlled trial. *Neurology* 2004; 63: 651-657.

Salloway SP, Pratt RD, Perdomo CA. Donepezil is well tolerated in patients with vascular dementia: a comparison of tolerability in vascular dementia patients and Alzheimer's disease patients. *Eur J Neurol* 2002; 9: 165-224.

Nicht E2: vorgegebene Prüf- bzw. Vergleichsintervention nicht erfüllt (8)

Finkel SI, Mintzer JE, Dysken M, Krishnan KR, Burt T, McRae T. A randomized, placebo-controlled study of the efficacy and safety of sertraline in the treatment of the behavioral manifestations of Alzheimer's disease in outpatients treated with donepezil. *Int J Geriatr Psychiatry* 2004; 19: 9-18.

Kristensen M, Richardson A, Van Osselaer N, Vangeneugden T. European multicentre placebo-controlled trial to determine the safety and efficacy of galantamine hydrobromide 40 mg/d in patients diagnosed with Alzheimer type dementia (GAL-95-05). 1997.

Onofri M, Thomas A, Luciano AL, Iacono D, Di RA, D'Andreamatteo G, et al. Donepezil versus vitamin E in Alzheimer's disease: Part 2: mild versus moderate-severe Alzheimer's disease. *Clin Neuropharmacol* 2002; 25: 207-215.

Patterson CE, Passmore AP, Crawford VL. A 6-month open-label study of the effectiveness and tolerability of galantamine in patients with Alzheimer's disease. *Int J Clin Pract* 2004; 58: 144-148.

Santens P, Ventura M. Donepezil in the treatment of mild to moderate Alzheimer's disease: report of a Belgian multicenter study. *Acta Neurol Belg* 2003; 103: 159-163.

Suh GH, Yeon JH, Uk LC, Hoon OB, Nam BJ, Jung HY, et al. A prospective, double-blind, community-controlled comparison of three doses of galantamine in the treatment of mild to moderate Alzheimer's disease in a Korean population. *Clin Ther* 2004; 26: 1608-1618.

Suh GH, Jung HY, Lee CU, Oh BH, Lee SK, Lee N, Kim J et al. Effect of the apolipoprotein E epsilon4 allele on the efficacy and tolerability of galantamine in the treatment of Alzheimer's disease. *Dement Geriatr Cogn Disord* 2006; 21: 33-39.

Weiser M, Rotmensch HH, Korczyn AD, Hartman R, Cicin SA, Anand R, et al. A pilot, randomized, open-label trial assessing safety and pharmacokinetic parameters of co-administration of rivastigmine with risperidone in dementia patients with behavioral disturbances. *Int J Geriatr Psychiatry* 2002; 17: 343-346.

Nicht E3: keine relevanten Zielgrößen (1)

Parnetti L, Amici S, Lanari A, Romani C, Antognelli C, Andreasen N, Minthon L et al. Cerebrospinal fluid levels of biomarkers and activity of acetylcholinesterase (AChE) and butyrylcholinesterase in AD patients before and after treatment with different AChE inhibitors. *Neurol Sci* 2002; 23 Suppl 2: S95-S96.

Nicht E4: keine RCT (33)

Aguglia E, Onor ML, Saina M, Maso E. An open-label, comparative study of rivastigmine, donepezil and galantamine in a real-world setting. *Curr Med Res Opin* 2004; 20: 1747-1752.

Albuquerque EX, Santos MD, Alkondon M, Pereira EF, Maelicke A. Modulation of nicotinic receptor activity in the central nervous system: a novel approach to the treatment of Alzheimer disease. *Alzheimer Dis Assoc Disord* 2001; 15 Suppl 1: S19-S25.

Almkvist O, Darreh ST, Stefanova E, Spiegel R, Nordberg A. Preserved cognitive function after 12 months of treatment with rivastigmine in mild Alzheimer's disease in comparison with untreated AD and MCI patients. *Eur J Neurol* 2004; 11: 253-261.

Baladi J-F, Bailey PAB, Black S, Bouchard RW, Farcnik KD, Gauthier S, Kertesz A et al. Rivastigmine for Alzheimer's disease: Canadian interpretation of intermediate outcome measures and cost implications. *Clin Ther* 2000; 22: 1549-1561.

Blesa R. Galantamine: therapeutic effects beyond cognition. *Dement Geriatr Cogn Disord* 2000; 11 Suppl 1: 28-34.

Böttcher BE. Therapie der Alzheimer-Demenz mit Donepezil: gut verträglich, wirksam und kostengünstig. *Neurol Rehabil* 2000; 6: 332-333.

Bullock R, Truyen L. Not all head-to-head trials are created equal: results from an open-label, short-term study seem inconsistent with previous donepezil literature. *Int J Geriatr Psychiatry* 2005; 20: 85-87.

Caro J, Ward A, Ishak K, Migliaccio-Walle K, Getsios D, Papadopoulos G, et al. To what degree does cognitive impairment in Alzheimer's disease predict dependence of patients on caregivers? *BMC Neurol* 2002; 2: 6.

Caro JJ, Getsios D. Long-Term Effects of Second-Generation Cholinesterase Inhibitors on Clinical Outcomes and Costs of Alzheimer's Disease. *Dis Manage Health Outcomes* 2003; 11.

Doody RS, Dunn JK, Clark CM, Farlow M, Foster NL, Liao T, et al. Chronic donepezil treatment is associated with slowed cognitive decline in Alzheimer's disease. *Dement Geriatr Cogn Disord* 2001; 12: 295-300.

Doody RS, Geldmacher DS, Gordon B, Perdomo CA, Pratt RD, Donepezil Study Group. Open-label, multicenter, phase 3 extension study of the safety and efficacy of donepezil in patients with Alzheimer disease. *Arch Neurol* 2001; 58: 427-433.

Desai AK, Grossberg GT. Rivastigmine for Alzheimer's disease. *Expert rev neurotherapeutics* 2005; 5: 563-580.

Ebell M. Does donepezil help patients with moderate Alzheimer's dementia preserve their ability to function independently? *Evidence Based Practice* 2002.

Farlow MR, Hake A, Messina J, Hartman R, Veach J, Anand R. Response of patients with Alzheimer disease to rivastigmine treatment is predicted by the rate of disease progression. *Arch Neurol* 2001; 58: 417-422.

Farlow MR. Update on rivastigmine. *Neurology* 2003; 9: 230-234.

Froelich L, Gertz HJ, Heun R, Heuser I, Jendroska K, Kornhuber J, et al. Donepezil for Alzheimer's disease in clinical practice--The DONALD Study. A multicenter 24-week clinical trial in Germany. *Dement Geriatr Cogn Disord* 2004; 18: 37-43.

Gasser US, Gasser T. Vergleich der Wirksamkeit von Cholinesterasehemmern und Ginkgo. *MMW Fortschr Med* 2001; 143: 40-41.

- Gu P, Jing QW, Ding XS, et al. Observation of effect on Alzheimer disease treated by Galanthamine hydrobromide capules. *Practical Geriatrics* 2000; 14: 307-308.
- Ieni J, Warner JP. Donepezil improved cognitive and global function in mild-to-moderate Alzheimer disease. *Evid-based Med* 1998; 3: 155.
- Joffres C, Bucks RS, Haworth J, Wilcock GK, Rockwood K. Patterns of clinically detectable treatment effects with galantamine: a qualitative analysis. *Dement Geriatr Cogn Disord* 2003; 15: 26-33.
- Kurz A, Van BB. Ginkgo biloba compared with cholinesterase inhibitors in the treatment of dementia: a review based on meta-analyses by the cochrane collaboration. *Dement Geriatr Cogn Disord* 2004; 18: 217-226.
- Lilienfeld S, Parys W. Galantamine: Additional benefits to patients with Alzheimer's disease. *Dement Geriatr Cogn Disord* 2000; 11: 19-27.
- Lilienfeld S. Galantamine - A novel cholinergic drug with a unique dual mode of action for the treatment of patients with Alzheimer's disease. *CNS Drug Rev* 2002; 8: 159-176.
- Luckmann R. Donepezil improved the clinical state and quality of life in moderate-to-severe Alzheimer disease. *ACP J Club* 2002; 136: 59.
- Mani RB. The evaluation of disease modifying therapies in Alzheimer's disease: A regulatory viewpoint. *Stat Med* 2004; 23: 305-314.
- Mossello E, Tonon E, Caleri V, Tilli S, Cantini C, Cavallini MC, et al. Effectiveness and safety of cholinesterase inhibitors in elderly subjects with Alzheimer's disease: a "real world" study. *Arch Gerontol Geriatr Suppl* 2004; 297-307.
- Nobili F, Vitali P, Canfora M, Girtler N, De LC, Mariani G, et al. Effects of long-term Donepezil therapy on rCBF of Alzheimer's patients. *Clin Neurophysiol* 2002; 113: 1241-1248.
- Peters O, Frolich L, Heuser I. Antidementive Kombinationstherapie. *Nervenheilkunde* 2005; 24: 476-482.
- Relkin NR, Reichman WE, Orazem J, McRae T. A large, community-based, open-label trial of donepezil in the treatment of Alzheimer's disease. *Dement Geriatr Cogn Disord* 2003; 16: 15-24.
- Schreiter GU, Gasser T. Ein Vergleich von Cholinesterasehemmern und Ginkgo-Extrakt in der Behandlung der Alzheimer-Demenz. *Fortschr Med* 2001; 135-138.
- Vennerica A, Shanks MF, Staff RT, Pestell SJ, Forbes KE, Gemmill HG, Murray AD. Cerebral blood flow and cognitive responses to rivastigmine treatment in Alzheimer's disease. *Neuroreport* 2002; 13: 83-87.
- Vetter C. Stabilization of cognition and everyday performance with galantamine therapy in dementia patients. *Tw Neurol Psychiatr* 2001; 24-25.
- Ward A, Caro JJ, Getsios D, Ishak K, O'Brien J, Bullock R, AHEAD Study Group. Assessment of health economics in Alzheimer's disease (AHEAD): treatment with galantamine in the UK. *Int J Geriatr Psychiatry* 2003; 18: 740-747.

A1: Beobachtungsdauer weniger als 16 Wochen (15)

- Agid YDBA. Efficacy and tolerability of rivastigmine in patients with dementia of the Alzheimer type. *Curr Ther Res Clin Exp* 1998; 59: 837-845.
- Boada-Rovira M, Brodaty H, Cras P, Baloyannis S, Emre M, Zhang R, et al. Efficacy and safety of donepezil in patients with Alzheimer's disease: results of a global, multinational, clinical experience study. *Drugs Aging* 2004; 21: 43-53.
- Greenberg SM, Tennis MK, Brown LB, Gomez-Isla T, Hayden DL, Schoenfeld DA, et al. Donepezil therapy in clinical practice: a randomized crossover study. *Arch Neurol* 2000; 57: 94-99.

Hegerl U, Mergl R, Henkel V, Gallinat J, Kotter G, Muller-Siecheneder F, Pogarell O et al. Kinematic analysis of the effects of donepezil hydrochloride on hand motor function in patients with Alzheimer dementia. *J Clin Psychopharmacol* 2003; 23: 214-216.

Holmes C, Wilkinson D, Dean C, Vethanayagam S, Olivieri S, Langley A, et al. The efficacy of donepezil in the treatment of neuropsychiatric symptoms in Alzheimer disease. *Neurology* 2004; 63: 214-219.

Homma A, Imai N, Hagiuchi S, Hasegawa H, Kameyama M. Late phase II clinical study of acetyl choline esterase E 2020 in patients with Alzheimer type dementia. *Clin Eval* 1998; 26: 251-284.

Jones RW, Soininen H, Hager K, Aarsland D, Passmore P, Murthy A, et al. A multinational, randomised, 12-week study comparing the effects of donepezil and galantamine in patients with mild to moderate Alzheimer's disease. *Int J Geriatr Psychiatry* 2004; 19: 58-67.

Markowitz JS, Gutterman EM, Lilienfeld S, Papadopoulos G. Sleep-related outcomes in persons with mild to moderate Alzheimer disease in a placebo-controlled trial of galantamine. *Sleep* 2003; 26: 602-606.

Nikolova G, Traykov L. Efficacy of donepezil in patients with Alzheimer's disease - Results of 12-week open clinical trial. *Acta Med Bulg* 2001; 28: -70.

Peng D, Xu X, Hou Q. The safety and efficacy of aricept in patients with Alzheimer disease. *Zhonghua Shen Jing Ge Za Zhi* 2002; 35: 19-21.

Rockwood K, Mintzer J, Truyen L, Wessel T, Wilkinson D. Effects of a flexible galantamine dose in Alzheimer's disease: a randomised, controlled trial. *J Neurol Neurosurg Psychiatry* 2001; 71: 589-595.

Rogers SL, Friedhoff LT. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US Multicentre, Randomized, Double-Blind, Placebo-Controlled Trial. The Donepezil Study Group. *Dementia* 1996; 7: 293-303.

Rogers SL, Doody RS, Mohs RC, Friedhoff LT. Donepezil improves cognition and global function in Alzheimer disease: a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. *Arch Intern Med* 1998; 158: 1021-1031.

Wilkinson D, Murray J. Galantamine: a randomized, double-blind, dose comparison in patients with Alzheimer's disease. *Int J Geriatr Psychiatry* 2001; 16: 852-857.

Wilkinson DG, Passmore AP, Bullock R, Hopker SW, Smith R, Potocnik FC, et al. A multinational, randomised, 12-week, comparative study of donepezil and rivastigmine in patients with mild to moderate Alzheimer's disease. *Int J Clin Pract* 2002; 56: 441-446.

A2: andere Demenzerkrankung (3)

Hull RP, Collins M, Frey J, McIntosh W, Chalmers A, Sauer R, Thein S. A randomized 26 week double blind placebo controlled trial to evaluate the safety and efficacy of galantamine in the treatment of vascular dementia. *Clinical Trials gov* 2002; 1-14.

Kurz A. Non-cognitive benefits of galantamine (Reminyl) treatment in vascular dementia. *Acta Neurol Scand Suppl* 2002; 178: 19-24.

Willan AR, Goeree R, Pullenayegum EM, McBurney C, Blackhouse G. Economic evaluation of rivastigmine in patients with Parkinson's disease dementia. *Pharmacoeconomics* 2006; 24: 93-106.

A3: Daten aus unkontrollierten „open label“ follow-up Phasen (14)

Bullock R, Erkinjuntti T, Lilienfeld S, GAL INT. Management of patients with Alzheimer's disease plus cerebrovascular disease: 12-month treatment with galantamine. *Dement Geriatr Cogn Disord.* 2004; 17: 29-34.

Doraiswamy PM, Krishnan KRR, Anand R, Sohn H, Danyluk J, Hartman RD, et al. Long-term effects of rivastigmine in moderately severe Alzheimer's disease: Does early initiation of therapy offer sustained benefits? *Prog Neuropsychopharmacol Biol Psychiatry* 2002; 26: 705-712.

Erkinjuntti T, Kurz A, Small GW, Bullock R, Lilienfeld S, Damaraju CV, et al. An open-label extension trial of galantamine in patients with probable vascular dementia and mixed dementia. *Clin Ther* 2003; 25: 1765-1782.

Farlow M, Anand R, Messina J, Jr., Hartman R, Veach J. A 52-week study of the efficacy of rivastigmine in patients with mild to moderately severe Alzheimer's disease. *Eur Neurol* 2000; 44: 236-241.

Geldmacher DS, Provenzano G, McRae T, Mastey V, Ieni JR. Donepezil is associated with delayed nursing home placement in patients with Alzheimer's disease. *J Am Geriatr Soc* 2003; 51: 937-944.

Grossberg G, Irwin P, Satlin A, Mesenbrink P, Spiegel R. Rivastigmine in Alzheimer disease: efficacy over two years. *Am J Geriatr Psychiatry* 2004; 12: 420-431.

Kurz AF, Erkinjuntti T, Small GW, Lilienfeld S, Damaraju CR. Long-term safety and cognitive effects of galantamine in the treatment of probable vascular dementia or Alzheimer's disease with cerebrovascular disease. *Eur J Neurol* 2003; 10: 633-640.

Lyketsos CG, Reichman WE, Kershaw P, Zhu Y. Long-term outcomes of galantamine treatment in patients with Alzheimer disease. *Am J Geriatr Psychiatry* 2004; 12: 473-482.

Pirttila T, Wilcock G, Truyen L, Damaraju CV. Long-term efficacy and safety of galantamine in patients with mild-to-moderate Alzheimer's disease: multicenter trial. *Eur J Neurol* 2004; 11: 734-741.

Prasher VP, Adams C, Holder R, Roy A, Cassidy G, Roy M, Mlele T et al. Long term safety and efficacy of donepezil in the treatment of dementia in Alzheimer's disease in adults with Down syndrome: Open label study. *Int J Geriatr Psychiatry* 2003; 18: 549-551.

Raskind MA, Peskind ER, Truyen L, Kershaw P, Damaraju CV. The cognitive benefits of galantamine are sustained for at least 36 months: a long-term extension trial. *Arch Neurol* 2004; 61: 252-256.

Rogers SL, Friedhoff LT. Long-term efficacy and safety of donepezil in the treatment of Alzheimer's disease: an interim analysis of the results of a US multicentre open label extension study. *Eur Neuropsychopharmacol* 1998; 8: 67-75.

Rogers SL, Doody RS, Pratt RD, Ieni JR. Long-term efficacy and safety of donepezil in the treatment of Alzheimer's disease: final analysis of a US multicentre open-label study. *Eur Neuropsychopharmacol* 2000; 10: 195-203.

Rösler M, Retz W, Retz JP, Dennler HJ. Effects of two-year treatment with the cholinesterase inhibitor rivastigmine on behavioural symptoms in Alzheimer's disease. *Behav Neurol* 1998; 11: 211-216.

A4: Doppelpublikationⁿ (19)

Blesa R, Davidson M, Kurz A, Reichman W, Van BB, Schwalen S. Galantamine provides sustained benefits in patients with 'advanced moderate' Alzheimer's disease for at least 12 months. *Dement Geriatr Cogn Disord* 2003; 15: 79-87.

Bullock R. The needs of the caregiver in the long-term treatment of Alzheimer disease. *Alzheimer Dis Assoc Disord* 2004; 18 Suppl 1: S17-S23.

Coyle J, Kershaw P. Galantamine, a cholinesterase inhibitor that allosterically modulates nicotinic receptors: effects on the course of Alzheimer's disease. *Biol Psychiatry* 2001; 49: 289-299.

ⁿ Dieser Kategorie wurden auch gepoolte Auswertungen oder Post-Hoc-Analysen zugeordnet, die keine zusätzlichen relevanten und als ausreichend valide angesehenen Informationen für die Nutzenbewertung enthielten.

- Deleu D. A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology* 2002; 58: 835-836.
- Dengiz AN, Kershaw P. The clinical efficacy and safety of galantamine in the treatment of Alzheimer's disease. *CNS Spectr* 2004; 9: 377-392.
- Farlow M, Lane R, Kudravalli S, He Y. Differential qualitative responses to rivastigmine in APOE epsilon 4 carriers and noncarriers. *Pharmacogenomics J* 2004; 4: 332-335.
- Farlow M, Potkin S, Koumaras B, Veach J, Mirski D. Analysis of outcome in retrieved dropout patients in a rivastigmine vs placebo, 26-week, Alzheimer disease trial. *Arch Neurol* 2003; 60: 843-848.
- Feldman H, Gauthier S, Hecker J, Vellas B, Hux M, Xu Y, Schwam EM et al. Economic evaluation of donepezil in moderate to severe Alzheimer disease. *Neurology* 2004; 63: 644-650.
- Feldman HH, Van BB, Kavanagh S. Effects of galantamine on activities of daily living in Alzheimer's disease: Evidence from six randomised double-blind placebo-controlled trials. *Res Pract Alzheimer's Dis* 2005; 10: 234-238.
- Finucane TE, Tariot PN, Cummings JL, Katz IR, Mintzer J, Perdomo CA, Schwam EM et al. Getting donepezil into the nursing home. *J Am Geriatr Soc* 2003; 51(1): 133-134.
- Dunbar F, Zhu Y, Brashear HR. Post hoc comparison of daily rates of nausea and vomiting with once- and twice-daily galantamine from a double-blind, placebo-controlled, parallel-group, 6-month study. *Clin Ther* 2006; 28: 365-372.
- Gauthier S. Efficacy of donepezil on maintenance of activities of daily living in patients with moderate-to-severe alzheimer's disease, and impact on caregiver burden. *Geriatrics and Aging* 2004; 7: 34-36.
- Grossberg G. Comparative study of donepezil and rivastigmine. *Int J Clin Pract* 2003; 57: 70-71.
- Herrmann N, Rabheru K, Wang J, Binder C. Galantamine treatment of problematic behavior in Alzheimer disease: post-hoc analysis of pooled data from three large trials. *Am J Geriatr Psychiatry* 2005; 13: 527-534.
- Kertesz A. Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a randomized trial. *Curr Neurol Neurosci Rep* 2002; 2: 503-504.
- Marder K. Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a randomized trial. *Curr Neurol Neurosci Rep* 2002; 2: 389-390.
- Morganroth J, Graham S, Hartman R, Anand R. Electrocardiographic effects of rivastigmine. *J Clin Pharmacol* 2002; 42: 558-568.
- Potkin SG, Anand R, Fleming K, Alva G, Keator D, Carreon D, et al. Brain metabolic and clinical effects of rivastigmine in Alzheimer's disease. *Int J Neuropsychopharmacol* 2001; 4: 223-230.
- Steinman MA, Covinsky KE, Tariot PN, Cummings JL, Katz IR, Mintzer J, Perdomo CA et al. Donepezil for nursing home patients with dementia: a reinterpretation of the evidence. *J Am Geriatr Soc* 2003; 51: 132-133.

A5: keine Volltextpublikation (49)

- Anand R, Messina J, Veach J, Hartman R. Effects of Rivastigmine in patients with moderately severe Alzheimer's disease. *Proceedings of the 6th International Stockholm/Springfield Symposium on Advances in Alzheimer Therapy*; 2000 Apr 5-8, Stockholm, Sweden 2000; 199.
- Bullock R, Lillienfeld S. Galantamine shows promising results in alzheimer's disease with cerebrovascular components and probable vascular dementia (preliminary results). *JNS* 2001; 187.
- Bullock R, Wilkinson DG, Passmore P, Hopker SW, Smith R, Potocnik FC, et al. Caregiver and physician determination of satisfaction with and ease of use of donepezil and rivastigmine treatment in alzheimer's disease patients. *Proceedings of the 17th Alzheimer's Disease International Conference*; 2001 Oct 25-27, Christchurch, New Zealand 2001; 39.

- Clary C, McRae T, Griesing T, Whalen E. The safety of donepezil and sertraline for the management of behavioral symptoms in patients with Alzheimer's disease. *Int J Neuropsychopharmacol* 2000; 3: S267.
- Doody R, Clark C, Farlow MM. Chronic Donepezil Treatment is Associated with Slowed Cognitive Decline in Alzheimer's disease patients. *Neurobiol Aging* 1998; 19: S303.
- Erkinjuntti T. Broad therapeutic benefits in patients with probable vascular dementia or Alzheimer's disease with cerebrovascular disease after treatment with galantamine. *Eur J Neurol* 2002; 9: 545.
- Engedal K, Soininen H, Verhey F, Waldemar G, Winblad B, Wimo A, et al. Donepezil improved or stabilized cognition over one year in patients with mild and moderate alzheimer's disease. *Eur Neuropsychopharmacol* 2000; 10: S368.
- Feldman H, Gauthier S, Hecker J, Vellas B, Subbiah P, Whalen E. Donepezil provides benefits in global function in moderate to severe Alzheimer's Disease. *Proceedings of the World Alzheimer Congress; 2000 Jul 9-13, Washington DC, USA 2000.*
- Feldman H, Gauthier S, Hecker J, Vellas B, Subbiah P, Whalen E, et al. Benefits of Donepezil on global function, behavior, cognition and ADLs in patients with moderate to severe Alzheimer's disease. *Neurology* 2000; 54: A469.
- Froelich L, Gertz HJ, Heun R, Heuser I, Jendroska K, Kornhuber J, et al. Donepezil for alzheimer's disease - the donald study - a multicenter 24 weeks clinical trial in germany. *Eur Neuropsychopharmacol* 2000; 10: S360.
- Gauthier S, Lussier I. Donepezil treatment is associated with alleviation of caregiver burden in alzheimer's disease: interim results from an open-label multicentre clinical trial. *Eur J Neurol* 1999; 6: 116.
- Gauthier S, Feldman H, Vellas B, Subbiah P. Efficacy of donepezil on functional, behavioural and cognitive symptoms in patients with moderate to severe alzheimer's disease. *J Am Geriatr Soc* 2000; 48: S2.
- Gauthier S, Feldman H, Hecker J, Vellas B, Subbiah P, Whalen E. Effects of donepezil on behaviour and other domains in moderate to severe alzheimer's disease. *Eur Neuropsychopharmacol* 2000; 10: S359.
- Grundman M. MCI trial with vitamin E and donepezil. *Proceedings of the 6th International Stockholm /Springfield Symposium on Advances in Alzheimer Therapy; 2000 Apr 5-8, Stockholm, Sweden 2000.*
- Hecker J, Foti D, Gauthier S, Vellas B, Subbiah P, Whalen E. Benefits of Donepezil in the treatment of behavioural problems in moderate to severe Alzheimer's Disease. *Proceedings of the World Alzheimer Congress; 2000 Jul 9-13, Washington DC, USA 2000.*
- Kewitz H, Berzewski H, Rainer M, Dal BP, Friedl E, Deisenhammer J, et al. Galantamine, a selective non-toxic acetylcholinesterase inhibitor is significantly superior over placebo in treatment of SDAT. *Neuropsychopharmacology* 1994; 10: 130S.
- Kurz A, Lilienfeld S, Damaraju CV. Galantamine is safe and effective for the long-term treatment of cognitive decline in patients with alzheimer disease with cerebrovascular components or probable vascular dementia. *Proceedings of the 7th International Geneva/Springfield Symposium on Advances in Alzheimer therapy, 2002 Apr 3-6, Geneva, Switzerland 2002; 206.*
- Lilienfeld S, Papadopoulos G. Galantamine alleviates caregiver burden in Alzheimer's disease. *14th Annual Meeting of the American Association for Geriatric Psychiatry; 2001 23-26 Feb, San Francisco, USA 2001.*
- McRae T, Griesing T, Whalen E. Donepezil and sertraline for the management of behavioral symptoms in patients with alzheimer's disease. *Neurology* 2000; 54: A416-A417.
- McRae T, Griesing T, Whalen E. Effectiveness of Donepezil on behavioural disturbances in mild to moderate Alzheimer's Disease patients. *Proceedings of the World Alzheimer Congress; 2000 Jul 9-13, Washington, DC, USA 2000.*
- McRae T, Knopman D, Duttagupta S, Ieni J, Provenzano G. Donepezil delays time to nursing home placement in patients with Alzheimer's disease. *14th Annual Meeting of the American Association for Geriatric Psychiatry; 2001 Feb 23-26, San Francisco, USA 2001.*
- Nunez M, Hasselbalch S, Heun R, Kalisvaart CJ, Kozubski W, Sakka P, et al. Donepezil-treated Alzheimer's disease patients with apparent initial cognitive decline demonstrate significant benefits when therapy is

continued: results from a randomised placebo-controlled trial. Second Annual Dementia Congress, 2003 Sep 12-14, Washington, DC, USA 2003.

O'Brien A. A pilot study comparing the effect of galantamine (Reminyl) with donepezil in patients with Alzheimer's Disease. National Research Register 2000. Studienregistrierung.

Parsa M, Poggi E, Barte L. Treatment of dementia patients with psychotic and behavioural symptoms with quetiapine and donepezil. *Eur Neuropsychopharmacol* 2000; 10: S302.

Parys W, Pontecorvo MJ. Treatment of Alzheimer's disease with galantamine, a compound with a dual mechanism of action. *Neurobiol Aging* 1998; 19 Suppl 4: S304.

Passmore P, Wetterberg P, Adler G, Bullock R, Soininen H, Aarsland D, et al. First head to head study comparing the tolerability, ease of use, and efficacy of donepezil and galantamine in Alzheimer's disease. *Proceedings of the 8th International Conference on Alzheimer's Disease and Related Disorders*; 2002 July 20-25, Stockholm, Sweden 2002.

Potocnik FCV, Smith R, Passmore P, Hock C, Wilkinson D, Maud CM, et al. Tolerability, ease of use, and efficacy of donepezil and rivastigmine in alzheimer's disease patients. 2001 Annual Meeting of the American Psychiatric Association; 2001 May 5-10, New Orleans, USA 2001.

Pratt RD, Gauthier S, Burns A, Perdomo CA. Donepezil provides long-term clinical benefits for patients with Alzheimers Disease. *Proceedings of the World Alzheimer Congress*; 2000 Jul 9-13, Washington, DC, USA 2000.

Pratt R, Perdomo C, The D. Donepezil is well tolerated in patients with vascular dementia: a comparison of safety and tolerability results from randomized, placebo-controlled clinical trials in vascular dementia patients and Alzheimer's disease patients. *Proceedings of the 8th International Conference on Alzheimer's Disease and Related Disorders*; 2002 July 20-25, Stockholm, Sweden 2002.

Pratt RD, Perdomo CA, The D. Donepezil improves cognition in patients with vascular dementia: results from study 307, a 24-week, randomized, double-blind, placebo-controlled trial. *Proceedings of the 7th International Geneva/Springfield Symposium on Advances in Alzheimer therapy*, 2002 Apr 3-6, Geneva, Switzerland 2002; 233.

Pratt RD, Perdomo CA, The D. Population characteristics and pattern of cognitive decline in patients with vascular dementia enrolled in two 24-week, randomized, double-blind, placebo-controlled trials. *Proceedings of the 7th International Geneva /Springfield Symposium on Advances in Alzheimer therapy*, 2002 Apr 3-6, Geneva, Switzerland 2002; 234.

Rainer M, Janoch P, Reiss A, et a. Galanthamine treatment in Alzheimer's disease: a preliminary evaluation of forty patients. *Can J Neurol Sci* 1993; 20 Suppl 4.

Robert P, Lebert F, Goni S, Touchon J, Vincent S, ARIAL-Study-Investigators-Collaborative-Group. The impact on caregiver distress of donepezil treatment of patients with mild Alzheimer's disease. *Proceedings of the Quality Research in Dementia Conference*; 2000 Nov 19-22, London, UK 2000.

Robert PH, Lebert F, Goni S, Touchon J. The impact of caregiver distress of donepezil treatment of patients with mild alzheimer's disease. 152nd Annual Meeting of the American Psychiatric Association. 1999 May 15-20 Washington, DC, USA 1999.

Rogers SL, Friedhoff LT. Donepezil provides long-term clinical benefits for patients with alzheimer's Disease (AD). *J Neurol Sci* 1997; 150: S296.

Rogers SL, Mohs RC, Friedhoff LT, Nightingale SL. Donepezil (E2020) Improves Cognition and Function in Patients with Mild to Moderately Severe Alzheimer's Disease Results from Phase III Trials: Donepezil approved for treatment of Alzheimer disease. *J Am Med Assoc* 1997; 277: 10.

Rogers SL, Friedhoff LT. Donepezil improves cognition in patients with mild to moderate AD: Results of ADAS-COG analysis in a 30-week phase III study. 10th European College of Neuropsychopharmacology Congress 1997 Sept 13-17, Vienna, Austria 1997.

Rogers SL, Friedhoff LT. Donepezil is well tolerated at clinically effective doses for the treatment of Alzheimer's disease (AD). 10th European College of Neuropsychopharmacology Congress 1997 Sept 13-17, Vienna, Austria 1997.

Rozzini L, Bargnani C, Bosio A, Chia F, Franzoni S, Leonardi R, et al. Comparison of efficacy and safety of rivastigmine and donepezil in patients with mild to moderate alzheimer disease: results from a multicentre randomised trial. Proceedings of the 7th International Geneva/Springfield Symposium on Advances in Alzheimer therapy, 2002 Apr 3-6, Geneva, Switzerland 2002; 240.

Salloway S. A double blind randomized pilot study to evaluate the effects of galantamine and donepezil on sleep and attention and gastrointestinal GI tolerance in patients with mild to moderate Alzheimer's disease AD. Clinical Trials Journal 2002; 1-2. Studienregistrierung.

Salloway S, Pratt RD, Perdomo CA, The D. Donepezil treated patients with vascular dementia demonstrated cognitive and global benefits: results from Study 308, a 24-week randomized double-blind, placebo-controlled trial. Proceedings of the 8th International Conference on Alzheimer's Disease and Related Disorders; 2002 July 20-25, Stockholm, Sweden 2002.

Sharma T. A prospective, randomized, multi-centre, double-blind, placebo-controlled, parallel-group study of the effect of Exelon on the time to clinical diagnosis of Alzheimer's disease in subjects with mild cognitive impairment. National Research Register 2000. Studienregistrierung.

Shua HJ, Smith J, Potel S. A head to head study of donepezil (aricept), rivastigmine (exelon) and galantamine (reminyl) for the treatment of Alzheimer's disease: safety, tolerability, clinical and caregiver impression after 4-5 months of treatment (a prospective study). Proceedings of the 8th International Conference on Alzheimer's Disease and Related Disorders; 2002 July 20-25, Stockholm, Sweden 2002.

Small GW, Lilienfeld S. Galantamine improves cognition in alzheimer's cerebrovascular disease. 155th Annual Meeting of the American Psychiatric Association; 2002 May 18-23, Philadelphia, USA 2002.

Tai CT, Liu CK, Sung SM, Pai MC, Hsu CY. The safety and efficacy of exelon in Alzheimer's patients: a multicentre, randomized, 26-week study in taiwan. Int J Neuropsychopharmacol 2000; 3: S356.

Thal L. A Randomized, Double-Blind, Placebo-Controlled Trial of Vitamin E and Donepezil HCL (Aricept) to Delay Clinical progression From Mild Cognitive Impairment. National Institute of Health Clinical Center Trials Database 1999. Studienregistrierung.

Veach KRRK, SM, Doraiswamy PM. Rivastigmine slows stage-specific global deterioration in alzheimer's disease. 152nd Annual Meeting of the American Psychiatric Association. 1999 May 15-20, Washington DC, USA 1999.

Vincent S, Andrews C, Lane R. Rivastigmine shows particular efficacy in Alzheimer patients with concomitant hypertension. Proceedings of the 7th International Geneva/Springfield Symposium on Advances in Alzheimer therapy, 2002 Apr 3-6, Geneva, Switzerland 2002; 253.

Winblad B, Engedal K, Soininen H, Verhey F, Waldemar G, Wimo A, et al. Donepezil enhances global function and activities of daily living compared with placebo in a one-year, double-blind trial in patients with mild to moderate alzheimer's disease. Proceedings of the Quality Research in Dementia Conference; 2000 Nov 19-22, London, UK 2000.

ANHANG C: LISTE DER GESCREENTEN SYSTEMATISCHEN ÜBERSICHTEN

1. Barner EL, Gray SL. Donepezil use in Alzheimer disease. *Ann Pharmacother* 1998; 32: 70-77.
2. Birks J. Cholinesterase inhibitors for Alzheimer's disease (Cochrane Review). *Cochrane Database Syst Rev* 2006; Chichester: John Wiley & Sons Ltd.
3. Birks J, Evans JG, Iakovidou V, Tsolaki M. Rivastigmine for Alzheimer's Disease (Cochrane Review). *Cochrane Database Syst Rev* 2000; Chichester: John Wiley & Sons Ltd.
4. Birks J, Harvey RJ. Donepezil for dementia due to Alzheimer's disease (Cochrane Review). *Cochrane Database Syst Rev* 2006; Chichester: John Wiley & Sons Ltd.
5. Clegg A, Bryant J, Nicholson T, McIntyre L, De BS, Gerard K, et al. Clinical and cost-effectiveness of donepezil, rivastigmine and galantamine for Alzheimer's disease: a rapid and systematic review. *Health Technol Assess* 2001; 5: 1-137.
6. Hansen RA, Gartlehner G, Kaufer DJ, Lohr KN, Randolph LC, Carey T. Drug class review on Alzheimer's drugs. Chapel Hill: Oregon Health & Science Unit; 2005.
7. Health Council of the Netherlands. Dementia. The Hague: Health Council of the Netherlands, 2002; publication no. 2002/04E
8. Herrmann N. Cognitive pharmacotherapy of Alzheimer's disease and other dementias. *Can J Psychiatry* 2002; 47: 715-722.
9. Ihl R. The impact of drugs against dementia on cognition in aging and mild cognitive impairment. *Pharmacopsychiatry* 2003; 36 Suppl 1: S38-S43.
10. Kaduszkiewicz H, Beck-Bornholdt HP, van den Bussche H, Zimmermann T. Fragliche Evidenz für den Einsatz des Cholinesterasehemmers Donepezil bei Alzheimer-Demenz - eine systematische Übersichtsarbeit. *Fortschr Neurol Psychiat* 2004; 72: 557-563.
11. Kaduszkiewicz H, Zimmermann T, Beck-Bornholdt HP, van den Bussche H. Cholinesterase inhibitors for patients with Alzheimer's disease: systematic review of randomised clinical trials. *BMJ* 2005; 331: 321-327.
12. Krall WJ, Sramek JJ, Cutler NR. Cholinesterase inhibitors: a therapeutic strategy for Alzheimer disease. *Ann Pharmacother* 1999; 33: 441-450.
13. Lanctot KL, Herrmann N, Yau KK, Khan LR, Liu BA, LouLou MM, et al. Efficacy and safety of cholinesterase inhibitors in Alzheimer's disease: a meta-analysis. *CMAJ* 2003; 169: 557-564.
14. Livingston G, Katona C. How useful are cholinesterase inhibitors in the treatment of Alzheimer's disease: a number needed to treat analysis (Structured abstract). *Int J Geriatr Psychiatry* 2000; 15(3): 203-207.
15. Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E et al. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. *Health Technol Assess* 2006; 10: 1-378.
16. Loy C, Schneider L. Galantamine for Alzheimer's disease (Cochrane Review). *Cochrane Database Syst Rev* 2004; Chichester: John Wiley & Sons Ltd.
17. Pankiewicz O, Pelkowska A, Landa K. Efficacy and safety of rivastigmine. *HTA Consulting*; ohne Jahr.

18. Pearson VE. Galantamine: a new alzheimer drug with a past life. *Ann Pharmacother* 2001; 35: 1406-1413.
19. Perras C, Shukla VK, Lessard CL, Skidmore B, Bergman H, Gauthier S. Cholinesterase inhibitors for Alzheimer's Disease: A systematic review of randomized controlled trials. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2005. Technology Report 58.
20. Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). *Neurology* 2001, 56: 1133-1142.
21. Ritchie CW, Ames D, Clayton T, Lai R. Metaanalysis of randomized trials of the efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer disease. *Am J Geriatr Psychiatry* 2004; 12: 358-369.
22. Santaguida P, Raina P, Booker L, Patterson C, Baldassarre F, Cowan D, et al. Pharmacological treatment of dementia. Agency for Healthcare Research and Quality (AHRQ); 2004. Evidence Report/Technology Assessment No. 97.
23. Sink KM, Holden KF, Yaffe K. Pharmacological treatment of neuropsychiatric symptoms of dementia. *JAMA* 2005; 293: 596-608.
24. Slattum PW, Johnson MA. Caregiver Burden in Alzheimer's Disease. *Consult Pharm* 2004; 19: 352-362.
25. Standridge JB. Pharmacotherapeutic approaches to the treatment of Alzheimer's disease. *Clin Ther* 2004; 26: 615-630.
26. Stein K. Donepezil in the treatment of mild to moderate senile dementia of the Alzheimer type (SDAT). Bristol: NHS Executive South and West, 1997. Development and Evaluation Committee Report No.69.
27. Stein K. Rivastigmine (Exelon TM) in the treatment of senile dementia of the Alzheimer type (SDAT). Bristol: NHS Executive South and West, 1998. Development and Evaluation Committee Report No. 89.
28. Trinh N-H, Hoblyn J, Mohanty S, Yaffe K. Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: A meta-analysis. *JAMA* 2003; 289: 210-216.
29. Whitehead A, Perdomo C, Pratt RD, Birks J, Wilcock GK, Evans JG. Donepezil for the symptomatic treatment of patients with mild to moderate Alzheimer's disease: a meta-analysis of individual patient data from randomised controlled trials. *Int J Geriatr Psychiatry*; 19: 624-633.
30. Wolfson C, Moride Y, Perrault A, Momoli F, Demers L, Oremus M. Drug treatments for Alzheimer's disease: I. A comparative analysis of clinical trials. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2000. Technology Report 8.

ANHANG D: ÜBERBLICK ÜBER DIE IN DEN STUDIEN ERHOBENEN ERGEBNISMAßE

Die folgende Tabelle gibt einen Überblick über die Ergebnismaße, die in den eingeschlossenen Studien erhoben worden sind und sich den Zielkriterien gemäß Abschnitt 4.1.3 zuordnen lassen.

Wenn in den Studien zu einem Therapieziel (z. B. kognitive Fähigkeiten) mehrere Skalen berichtet wurden, wird in der Nutzenbewertung im Allgemeinen jeweils nur eine (vorzugsweise diejenige, die am häufigsten verwendet wurde) berichtet (vgl. Abschnitt 4.4.2). Die jeweils nicht ausgewerteten Zielgrößen sind jeweils mit einer Klammer versehen. Wo in Einzelfällen Zielgrößen aus anderen Gründen als der Redundanz ausgeschlossen wurden oder aus besonderen Gründen mehr als eine Zielgröße zu einem Therapieziel ausgewertet worden ist, ist dies jeweils gekennzeichnet.

Tabelle 47: Gegenüberstellung von Zielkriterien des Berichtsplan und in den Studien verwendeten Ergebnismaßen

	Aktivitäten des täglichen Lebens ^(a)	psychopathologische Symptome ^(a)	kognitive Leistungsfähigkeit ^(a)	krankheitsbezogene Lebensqualität ^(a)	Notwendigkeit vollstationärer Pflege ^(a)	Lebensqualität der Angehörigen ^(b)	Betreuungsaufwand ^(b)	klinisches Krankheitsstadium ^(c)
Donepezil vs. Placebo								
Burns 1999	IDDD		ADAS-cog	QoL				CIBIC-plus (CDR-SB)
Gauthier 2002	DAD PSMS-plus ^(d) IADL-plus ^(d)	NPI	sMMSE (SIB)			CSS (SF-36) ^(e)	Zeitaufwand für IADL und PSMS	CIBIC-plus (FRS)
Homma 2000	CMCS		ADAS-J cog					J-CIGIG (MENFIS) (CDR-SB)
Krishnan 2003			ADAS-cog					
Mohs 2001	ADFACS		MMSE					CDR-SB
Moraes 2006			ADAS-cog					
Prasher 2002		NPI (ABS)	SIB					(DMR) ^(f) (allg. Eindruck der Betreuer) ^(g)
Rogers 1998			ADAS-cog (MMSE)	QoL				CIBIC-plus (CDR-SB)
Seltzer 2004			ADAS-cog13 (MMSE) (CMBT)					CDR-SB (PGAS)
Tariot 2001	PSMS	NPI-NH	MMSE					(CDR-SB)
Tune 2003		NPI	ADAS-cog					

(Fortsetzung)

Tabelle 47 (Fortsetzung): Gegenüberstellung von Therapiezielen des Berichtsplans und in den Studien verwendeten Zielgrößen

	Aktivitäten des täglichen Lebens ^(a)	psychopathologische Symptome ^(a)	kognitive Leistungsfähigkeit ^(a)	krankheitsbezogene Lebensqualität ^(a)	Notwendigkeit vollstationärer Pflege ^(a)	Lebensqualität der Angehörigen ^(b)	Betreuungsaufwand ^(b)	klinisches Krankheitsstadium ^(c)
Winblad 2001	PDS	NPI	MMSE					GBS (GDS)
Galantamin vs. Placebo								
Brody 2005	ADCS-ADL	NPI	ADAS-cog11 (ADAS-cog13) (memory/non-memory ADAS-cog)			NPI-D		CIBIC-plus
Erkinjuntti 2002	DAD	NPI	ADAS-cog11 (ADAS-cog13)			NPI-D		CIBIC-plus
Raskind 2000	DAD		ADAS-cog11 (ADAS-cog13)					CIBIC-plus
Rockwood 2006	GAS DAD ^(h)		ADAS-cog11 (Examination of Memory and Temporality) ^(e) (Red Pen Task) ^(e)			CBS	(ACTS) ^(e)	CIBIC-plus
Tariot 2000	ACDS-ADL	NPI	ADAS-cog11 (ADAS-cog13)			NPI-D		CIBIC-plus
Wilcock 2000	DAD		ADAS-cog11					CIBIC-plus
Rivastigmin vs. Placebo								
B304 1998	PDS		ADAS-cog (ADAS-cogA) (MMSE)				CAS ^(e)	CIBIC-plus (GDS)
Corey-Bloom 1998	PDS		ADAS-cog (MMSE)					CIBIC-plus (GDS)

(Fortsetzung)

Tabelle 47 (Fortsetzung): Gegenüberstellung von Therapiezielen des Berichtsplans und in den Studien verwendeten Zielgrößen

	Aktivitäten des täglichen Lebens ^(a)	psychopathologische Symptome ^(a)	kognitive Leistungsfähigkeit ^(a)	krankheitsbezogene Lebensqualität ^(a)	Notwendigkeit vollstationärer Pflege ^(a)	Lebensqualität der Angehörigen ^(b)	Betreuungsaufwand ^(b)	klinisches Krankheitsstadium ^(c)
Forette 1999		NOSGER	ADAS-cog (Wechsler Tests) ^(g)					CIBIC-plus
Rösler 1999	PDS		ADAS-cog (MMSE)					CIBIC-plus (GDS)
Galantamin vs. Donepezil								
Wilcock 2003	BADLS	NPI	ADAS-cog (MMSE)			SCGB		
Rivastigmin vs. Donepezil								
Bullock 2005	ADCS-ADL	NPI	SIB (MMSE)					GDS
Fuschillo	PSMS		ADAS-cog (MMSE)					
Wang 2001	BRDS		MMSE					GDS
Rivastigmin vs. Donepezil vs. Galantamin								
Cumbo 2005		NPI BEHAVE-AD				NPI-D		
<p>Erläuterung: () = im Rahmen des vorliegenden Berichtes nicht ausgewertet a: patientenrelevante Therapieziele b: angehörigerelevante Therapieziele c: ergänzende Information d: sich ergänzende Skalen, die unterschiedliche Aspekte abbilden und deshalb ebenfalls ausgewertet wurden e: Hierzu konnte keine Publikation identifiziert werden. f: Nicht ausgewertet, da keine gute Verlaufserfassung (Screeninginstrument) g: Nicht ausgewertet, da unklare Validität. h: Da sich die Goal Attainment Scale von den anderen, psychometrischen Verfahren zur Erfassung der Aktivitäten des täglichen Lebens deutlich unterscheidet, werden hier für die Studie Rockwood 2006 sowohl die Ergebnisse aus dem DAD als auch aus der GAS dargestellt. g: Wechsler logical memory, digit span, word fluency Test</p>								

ANHANG E: KURZBESCHREIBUNG DER ERGEBNISMASSE

Kurzdarstellung der Ergebnisparameter und Messinstrumente bei der Alzheimer Demenz, die im Rahmen der Nutzenbewertung ausgewertet wurden.

Instrument	Erläuterung / Kommentar
<i>Globale Ergebnisparameter</i>	
Clinical Global Impression of Change Scale (CGIC) Auch als CIBIC mit Input des Betreuenden: CIBIC-M oder CIBIC-plus	CIBIC ist eine häufig verwendete Skala. Es wird die Veränderung relativ zu einem Ausgangszustand unter Verwendung aller verfügbaren Informationen eingeschätzt. Mittlere bis gute Test-Retest- und Interrater-Reliabilität und Übereinstimmungsvalidität. Scores auf dem CGIC/CIBIC spiegeln nicht das Ausmaß der individuellen globalen Verbesserung wider. Ärzte verwenden jedoch häufig die klinische Psychopathologie als Basis zur Bestimmung der globalen Verbesserung, das Krankenpflegepersonal eher den pflegerischen Aufwand. Bei der Version mit Input des Betreuenden kann das Ergebnis davon abhängen, ob man zuerst den Betroffenen oder den Betreuenden interviewt.
Clinical Dementia Rating Scale (CDR) Auch: Clinical Dementia Rating Scale Sum of Boxes (CDR-SB)	Häufig verwendetes Maß zur Bestimmung der Demenzschwere mit guter Interrater-Reliabilität und mittlerer bis guter Übereinstimmungsvalidität. Misst eher die kognitiven Aspekte der Demenz, nicht den globalen Gesundheitsstatus.
Global Deterioration Scale (GDS)	Häufig zur Stadieneinteilung verwendetes Instrument. Kann allerdings den Schweregrad der Erkrankung falsch wiedergeben und sollte deshalb nicht zur Stadieneinteilung in Arzneimittelstudien verwendet werden.
Gottfries-Bråne-Steen-Skala (GBS)	Eher selten eingesetztes Verfahren mit befriedigenden bis sehr guten psychometrischen Eigenschaften, das zur Quantifizierung von Demenz in Arzneimittelstudien geeignet ist. Erfasst kognitive, funktionale und behaviorale Aspekte und kann damit als globales Instrument betrachtet werden.
<i>Kognitive Ergebnisparameter</i>	
Alzheimer's Disease Assessment Scale-cognitive (ADAS-cog)	Sehr häufig als primärer Parameter eingesetzte Skala mit hoher Reliabilität. Die meisten der ADAS-cog-Subskalen haben ihre Begrenzungen in der Fähigkeit, eine Veränderung an beiden Enden des Schweregrad-Spektrums zu erkennen.
Mini Mental State Examination (MMSE)	Sehr häufig als Screening für den Schweregrad eingesetzte Skala mit hoher Retest-Reliabilität. Der Nutzen der MMSE als Veränderungsmaß bei individuellen Patienten ist begrenzt, die Skala kann leichtere Veränderungen der Kognition nicht erfassen. Die Veränderungssensitivität ist besser bei leichter bis mittelschwerer Demenz, unzureichend bei schwerer Demenz.

Instrument	Erläuterung / Kommentar
Severe Impairment Battery (SIB)	Der SIB wurde entwickelt, um kognitive Funktionen bei Personen zu messen, die für andere neuropsychologische Tests zu schwer beeinträchtigt sind. Die Subskalen umfassen u. a. Aufmerksamkeit, Orientierung, Sprache, Gedächtnis und räumlich-visuelle Fähigkeiten.
<i>Funktionale Ergebnisparameter und Lebensqualität des Patienten</i>	
Alzheimer's Disease Cooperative Study-Activities of Daily Living (ADCS-ADL)	Gebräuchlicher strukturierter Fragebogen zur Erfassung der instrumentellen und basalen Aktivitäten des täglichen Lebens über ein breites Schweregrad-Spektrum. Die Sensitivität und Reliabilität dieser Skala sind etabliert.
Alzheimer's Disease Functional Assessment and Change Scale (ADFACS)	Erfasst instrumentelle Aktivitäten des täglichen Lebens (10 Items). Eine ausführliche Erhebung der psychometrischen Eigenschaften ist allerdings nicht erfolgt. Die Items erscheinen veränderungssensitiv und haben eine gute Test-Retest-Reliabilität.
Blessed-Roth Dementia Scale (BRDS)	Es existieren wenige Daten zur Validität sowie zur Reliabilität. Die Veränderungssensitivität ist unklar. Die Verwendung verschiedener Modifikationen erschwert den Vergleich über Studien hinweg.
Bristol Activities of Daily Living Scale (BADLS)	Erhebung von 20 Aktivitäten des täglichen Lebens durch einen Betreuer, Pflegenden. Die Skala ist spezifisch für Demenz-Patienten, hat eine gute Inhaltsvalidität und gute Test-Retest-Validität. Korreliert gut mit kognitiven Funktionstests.
Caregiver-rated modified Crichton Scale (CMCS)	Modifikation der Crichton Geriatric Rating Scale. 7-Item-Skala mit Fragen zum Verständnis von Zeit und Ort, Führen eines Gespräches, Kooperation, Ruhelosigkeit, Ankleiden, soziale Aktivitäten und Freizeit. Reliabilität nachgewiesen, Validität unklar.
Disability Assessment for Dementia (DAD)	46 Item-strukturiertes Interview oder Fragebogen für Betreuende. Die erfassten Beeinträchtigungen entsprechen der WHO-Definition. Gebräuchliches Verfahren mit hohem Ausmaß an interner Konsistenz, Interrater- und Test-Retest-Reliabilität. Erfasst instrumentelle und basale Aktivitäten des täglichen Lebens.
Goal Attainment Scale (GAS)	Es werden individuelle Problembereiche definiert, darauf abgestimmte Ziele definiert und später die Veränderung im Vergleich zum Studienbeginn beurteilt. Damit stellt die GAS ein individualisiertes Ergebnismaß dar, das je nachdem, welche Ziele definiert werden, unterschiedliche Lebensbereiche abdeckt. Höhere Änderungssensitivität als bspw. PSMS oder IADL (Rockwood et al. 2003).
Instrumental Activities of Daily Living (IADL, iADL)	Ein häufig verwendetes und oft zitiertes Instrument. Die Skala ist theoretisch gut fundiert und die Aktivitäten, die erfasst werden, sind mit hoher Wahrscheinlichkeit in den frühen Demenzstadien beeinträchtigt. Es fehlen allerdings gute Reliabilitätstestungen.
Interview for Deterioration in Daily Living in Dementia (IDDD)	Erscheint angemessen zur Beurteilung der Alltagsfähigkeiten bei leichter und mittelschwerer Demenz, misst die funktionale Beeinträchtigung in der Selbstpflege und bei komplexen Aktivitäten. Psychometrische Daten fehlen.

Instrument	Erläuterung / Kommentar
Physical Self-Maintenance Scale (PSMS)	Erfassung von 6 basalen Aktivitäten. Theoretisch gut fundiert und geeignet bei institutionalisierten Patienten. Bei Patienten, die nicht im Pflegeheim leben, starker Deckeneffekt möglich. Die Testung der psychometrischen Eigenschaften ist unvollständig.
Progressive Deterioration Scale (PDS)	Die Skala ist stadiensensitiv und hat gute Reliabilität und Validität, nicht geeignet bei mittelschwer bis schwer Erkrankten, da eine Reihe basaler Tätigkeiten nicht erfasst wird.
Quality of Life Scale (QoL)	Skala bildet das Konzept der Lebensqualität entsprechend der WHO-Definition ab. Es fehlen jedoch Daten zur Reliabilität, Validität und Veränderungssensitivität. Angesichts anderer verfügbarer Instrumente zur Lebensqualität ist der Einsatz dieser Skala eher fragwürdig.
<i>Lebensqualität bzw. Belastung der Angehörigen</i>	
Caregiver Activity Survey (CAS)	Misst die Zeit, die Angehörige dafür aufwenden, Patienten bei den Aktivitäten des täglichen Lebens zu unterstützen. 6-Item-Version umfasst folgende Aspekte: Kommunikation, Transport, Essen, Ankleiden, das Kümern um das äußere Erscheinungsbild, Beaufsichtigung. Angemessene Retest-Reliabilität und konvergente Validität (Davis et al. 1997).
Caregiver Stress Scale (CSS)	Wenig verwendetes Instrument, erfasst aufbauend auf einem umfassenden Stress-Modell verschiedene Aspekte der Belastung durch die Pflege von Angehörigen (z. B. problematisches Verhalten, Konflikte in der Familie, Vereinbarkeit mit dem Beruf, ökonomische Belastung, Ausmaß der erfahrenen Unterstützung) (Pearlin 1990). Bei Gauthier 2002 verwendet (Angaben in [53]) und für die Anwendung im Rahmen von Alzheimer-Studien angepasst.
Screen for Caregiver Burden (SCGB)	Skala zur raschen Erfassung der Belastung pflegender Angehöriger. Interne Konsistenz, Test-Retest-Reliabilität und Validität sind zufriedenstellend.
Neuropsychiatric Inventory Caregiver Distress Scale (NPI-D)	Erfasst die mit den psychiatrischen Symptomen der Alzheimer Demenz verbundene Belastung der betreuenden Angehörigen. Zu den Symptomdomänen des NPI wird die dadurch jeweils erfahrene emotionale oder psychische Belastung auf einer 6-stufigen Skala angegeben. Angemessene Retest- und Interrater-Reliabilität. (Kaufert 1998)
<i>Psychopathologische Symptome</i>	
Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-D)	Fremdbeurteilungsskala zur Einschätzung des globalen Schweregrads sowie zur differenzierten Beschreibung von Verhaltensstörungen und psychopathologischen Symptomen bei Alzheimer Patienten. Stadiensensitiv, gute Validität, Reliabilität nur wenig untersucht.
Nurses' Observation Scale for Geriatric Patients (NOSGER)	Das NOSGER-Instrument erfasst beobachtbare Verhaltensweisen in verschiedenen Domänen (Gedächtnis, instrumentelle Aktivitäten des täglichen Lebens, Körperpflege, Stimmung, soziales Verhalten, störendes Verhalten). Selten eingesetzt, validiert, hat hohe Interrater- und Test-Retest-Reliabilität.
Neuropsychiatric Inventory (NPI)	Gängiges Instrument zur Erfassung des Verhaltens bzw. akzessorischer neuropsychiatrischer Symptome mit zufriedenstellender Reliabilität und Validität. Für Bewohner eines Pflegeheimes steht die modifizierte Version NPI-Nursing Home (NPI-NH) zur Verfügung.

Die Angaben beruhen, soweit nicht anders gekennzeichnet, auf folgenden Quellen:

Collegium Internationale Psychiatriae Salarum (Hrsg.). Internationale Skalen für Psychiatrie. Göttingen: Belz Test GmbH; 2000.

Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E and Clegg A. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. Health Technol Assess 2006; Vol. 10: No. 1.

Wolfson C, Moride Y, Perrault A, Momoli F, Demers L, Oremus M. Drug treatments for Alzheimer's disease. II. A review of outcome measures in clinical trials. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2000.

Zusätzliche Publikationen wurden zu einzelnen Skalen herangezogen:

Davis KL, Marin DB, Kane R, Patrick D, Peskind ER, Raskind MA et al. The Caregiver Activity Survey (CAS): development and validation of a new measure for caregivers of persons with Alzheimer's disease. Int J Geriatr Psychiatry, 1997; 12: 978-988.

Kaufert DI, Cummings JL, Christine D, Bray T, Castellon S, Masterman D et al. Assessing the impact of neuropsychiatric symptoms in Alzheimer's Disease: the Neuropsychiatric Inventory Caregiver Distress Scale. J Am Geriatr Soc, 1998; 46: 210-215.

Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist, 1990; 30: 583-594.

Rockwood K, Howlett S, Stadnyk K, Carver D, Powell C, Stolee P. Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. J Clin Epidemiol, 2003; 56: 736-743.

ANHANG F: BISLANG UNVERÖFFENTLICHTE INFORMATIONEN PHARMAZEUTISCHER UNTERNEHMEN

Kurzbeschreibung der Studie B304 (1998)

Die Studie B304 war eine 26-wöchige, multizentrische, randomisierte, doppelblinde, parallele, dreiarmlige Phase-II/III-Studie, in der bei Patienten mit leichter bis mittelschwerer Alzheimer Demenz die jeweils höchste vertragene Dosis Rivastigmin (2–12 mg/Tag) in zwei- oder dreimal täglicher Verabreichung mit Placebo verglichen wurde.

Primäres Ziel war es, bei Patienten mit wahrscheinlicher Alzheimer Demenz die Wirksamkeit und Verträglichkeit der höchsten individuell gut tolerierten Dosis Rivastigmin in zwei- oder dreimal täglicher Verabreichung gegen Placebo zu vergleichen. Das sekundäre Ziel bestand darin, die verschiedenen Dosierungsschemata (zwei- bzw. dreimal tägliche Rivastigmingabe) hinsichtlich der Wirksamkeit und Sicherheit untereinander zu vergleichen und die Veränderungen bei den Aktivitäten des täglichen Lebens zu erfassen.

Primäre Zielkriterien der Studie waren die Veränderungen nach 26 Wochen im Vergleich zu Baseline bzgl. ADAS-cog und CIBIC-plus, sekundäre Zielkriterien waren u.a. PDS sowie ADAS-cogA; CAS war tertiäres Zielkriterium.

Die Patienten wurden den Behandlungsgruppen gemäß einer computergenerierten Randomisierungsliste zugeteilt. Versiegelte Umschläge enthielten die Information über die Gruppenzuteilung und wurden an die Zentren gegeben. Die Kapseln, die Rivastigmin bzw. Placebo enthielten, waren in ihrem Äußeren identisch. Die Anzahl der eingenommenen Kapseln war in allen Gruppen jeweils gleich.

Die Studie bestand aus einer etwa 42-tägigen Screeningphase, während der die Patienten untersucht, die Einschlusskriterien geprüft und nicht gestattete Medikationen (z. B. Psychostimulanzien, anticholinerge Substanzen etc.) beendet wurden. Anschließend wurde die Baselineerhebung der verschiedenen Ergebnismaße vorgenommen; danach erfolgte die Randomisierung.

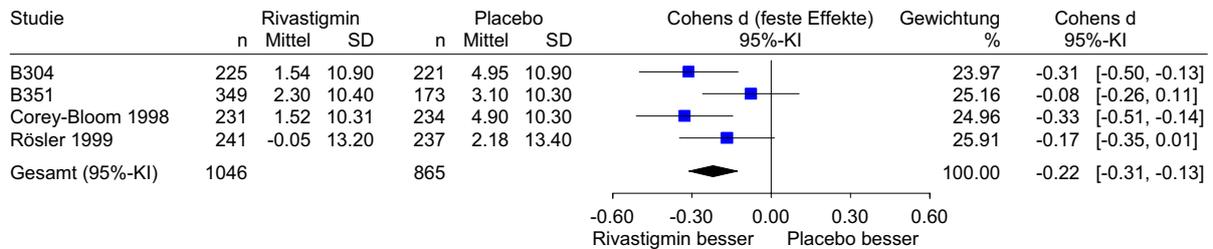
Die Ergebnismaße CIBIC-plus, ADAS-cog und PDS wurden in den Wochen 12, 18 und 26 erhoben, der MMSE sowie die GDS lediglich in Woche 26. Falls ein Patient die Studie frühzeitig beendete, sollten diese Ergebnismaße jeweils zum Zeitpunkt der Beendigung erhoben werden. Die Skala CAS wurde in den Wochen 6, 12, 18 und 26 erhoben. Unerwünschte Ereignisse wurden jeweils im Verlauf der Studie dokumentiert.

Daten und Studieninformationen, die in die vorliegende Nutzenbewertung eingegangen sind, finden sich in den Tabellen und im Text des Berichts zur Nutzenbewertung.

ANHANG G: ZUSÄTZLICHE META-ANALYSEN

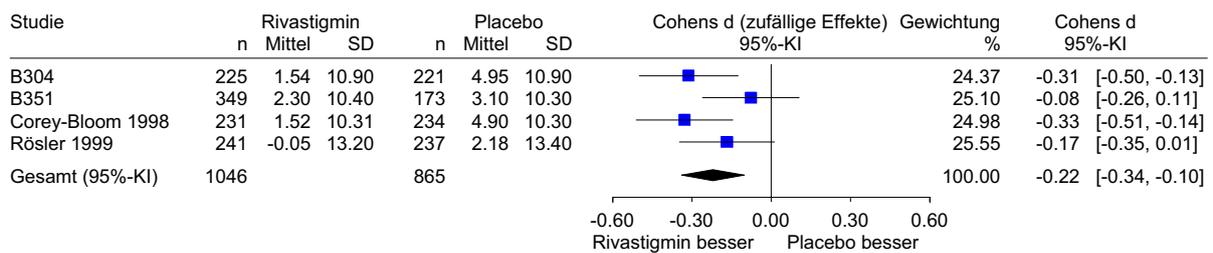
Im Folgenden werden Meta-Analysen zu Rivastigmin präsentiert, in denen Ergebnisse zur Studie B351 (jeweils aus [49]) mit einbezogen wurden.

Rivastigmin - Aktivitäten des täglichen Lebens
 Endpunkt: PDS - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: Q=4.97, df=3 (p=0.174), I²=39.7%
 Gesamteffekt: Z Score=-4.71 (p=0.000)

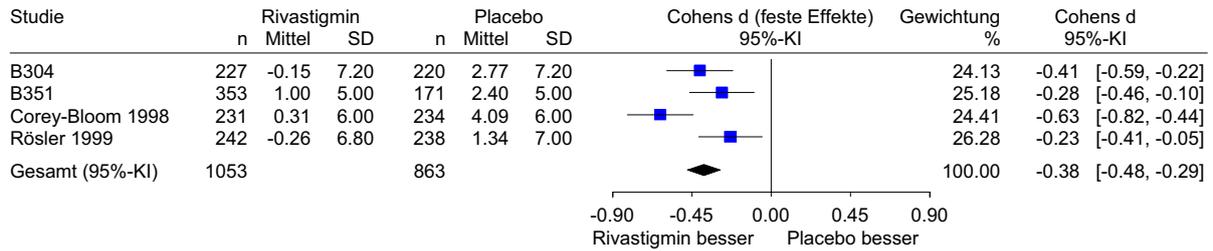
Rivastigmin - Aktivitäten des täglichen Lebens
 Endpunkt: PDS - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: Q=4.97, df=3 (p=0.174), I²=39.7%
 Gesamteffekt: Z Score=-3.67 (p=0.000), tau²=0.006

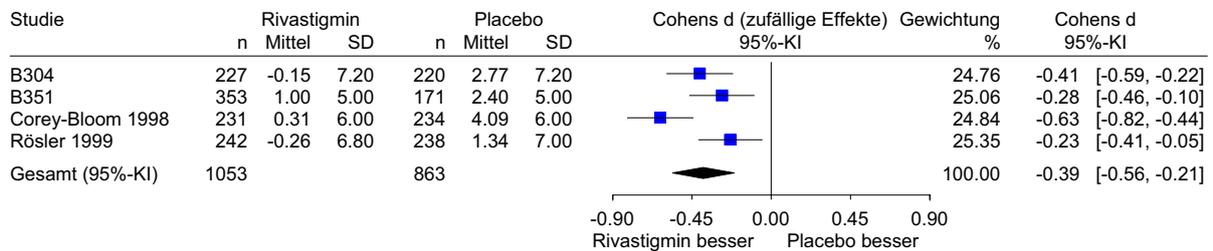
Abbildung 49: Rivastigmin: Meta-Analyse zu alltagspraktischen Fähigkeiten (mit B351, feste und zufällige Effekte)

Rivastigmin - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=10.74$, $df=3$ ($p=0.013$), $I^2=72.1\%$
 Gesamteffekt: Z Score=-8.16 ($p=0.000$)

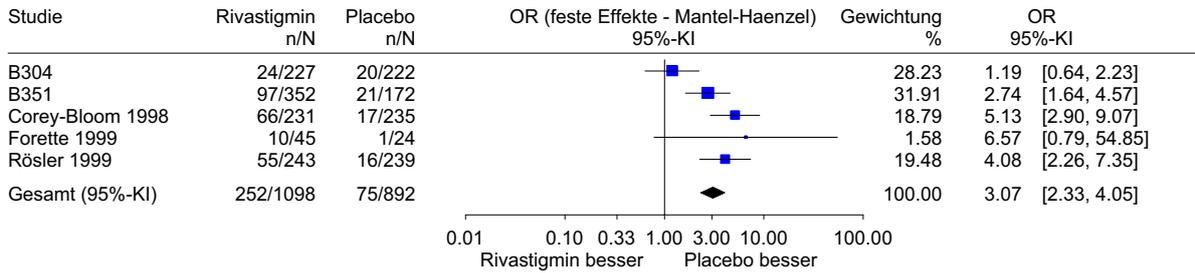
Rivastigmin - kognitive Leistungsfähigkeit
 Endpunkt: ADAS-cog - Differenz zu Baseline
 Distanzmaß: standardisierte Differenz der Mittelwerte



Heterogenität: $Q=10.74$, $df=3$ ($p=0.013$), $I^2=72.1\%$
 Gesamteffekt: Z Score=-4.34 ($p=0.000$), $\tau^2=0.023$

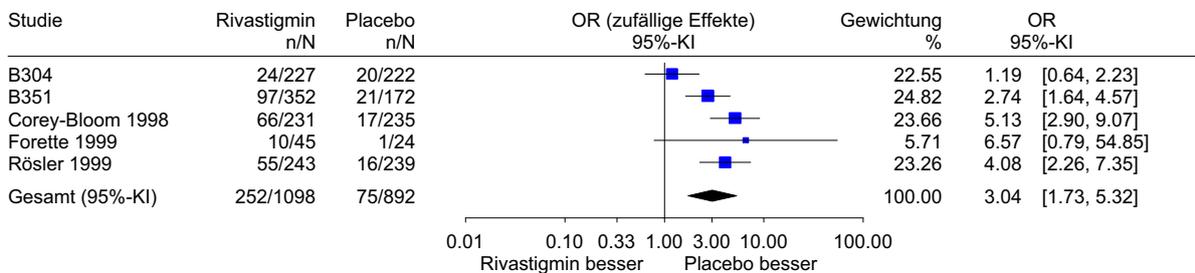
Abbildung 50: Rivastigmin: Meta-Analyse zu kognitiven Fähigkeiten (mit B351, feste und zufällige Effekte)

Rivastigmin - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



Heterogenität: $Q=13.48$, $df=4$ ($p=0.009$), $I^2=70.3\%$
 Gesamteffekt: Z Score=7.95 ($p=0.000$)

Rivastigmin - Abbruch wegen UEs
 Endpunkt: Studie aufgrund UE vorzeitig abgebrochen (ja, nein)
 Distanzmaß: Odds Ratio der Anteile von Patienten mit Abbruch



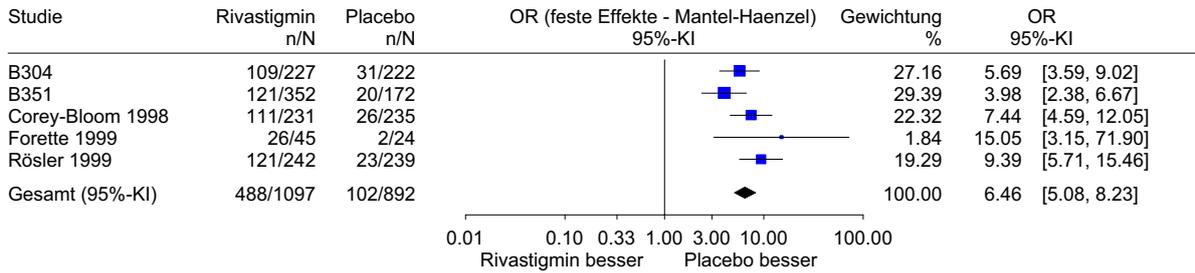
Heterogenität: $Q=13.48$, $df=4$ ($p=0.009$), $I^2=70.3\%$
 Gesamteffekt: Z Score=3.88 ($p=0.000$), $\tau^2=0.261$

Abbildung 51: Rivastigmin: Meta-Analyse zum Studienabbruch wegen unerwünschter Ereignisse (mit B351, feste und zufällige Effekte)

Rivastigmin - UEs

Endpunkt: Übelkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



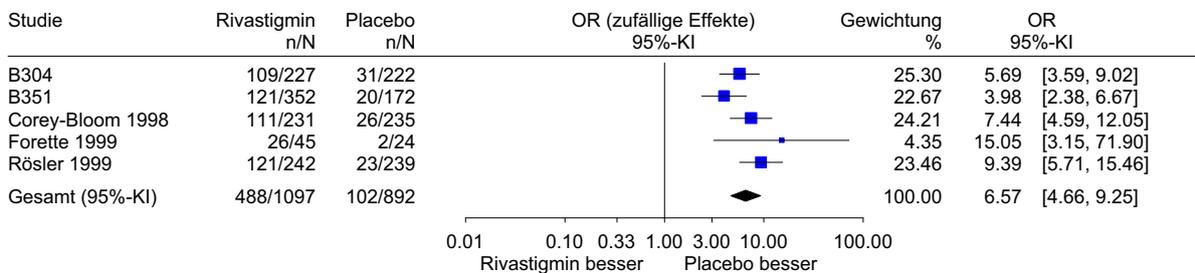
Heterogenität: $Q=7.3$, $df=4$ ($p=0.121$), $I^2=45.2\%$

Gesamteffekt: Z Score=15.17 ($p=0.000$)

Rivastigmin - UEs

Endpunkt: Übelkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: $Q=7.3$, $df=4$ ($p=0.121$), $I^2=45.2\%$

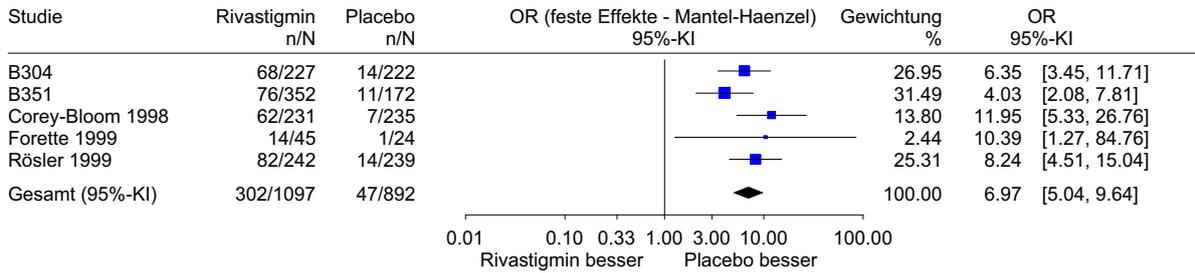
Gesamteffekt: Z Score=10.77 ($p=0.000$), $\tau^2=0.066$

Abbildung 52: Rivastigmin: Meta-Analyse zum Endpunkt „Übelkeit“ (mit B351, feste und zufällige Effekte)

Rivastigmin - UEs

Endpunkt: Erbrechen

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



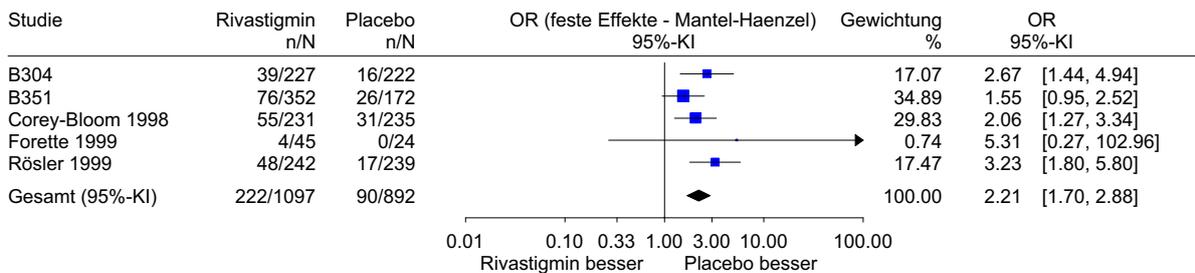
Heterogenität: $Q=4.87$, $df=4$ ($p=0.301$), $I^2=17.9\%$
Gesamteffekt: Z Score=11.74 ($p=0.000$)

Abbildung 53: Rivastigmin: Meta-Analyse zum Endpunkt „Erbrechen“ (mit B351)

Rivastigmin - UEs

Endpunkt: Schwindel

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



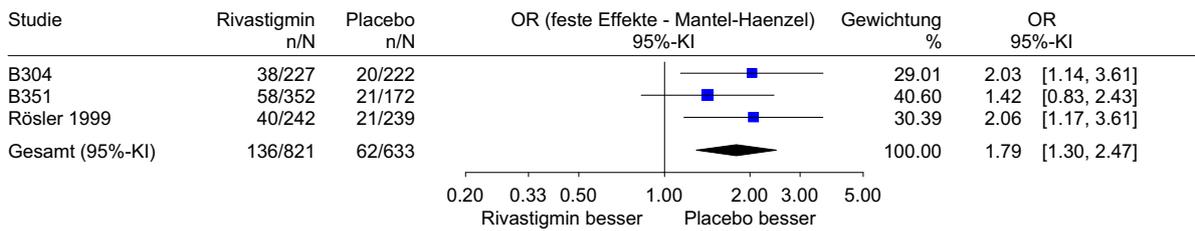
Heterogenität: $Q=4.46$, $df=4$ ($p=0.348$), $I^2=10.2\%$
Gesamteffekt: Z Score=5.87 ($p=0.000$)

Abbildung 54: Rivastigmin: Meta-Analyse zum Endpunkt „Schwindel“ (mit B351)

Rivastigmin - UEs

Endpunkt: Diarrhoe

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



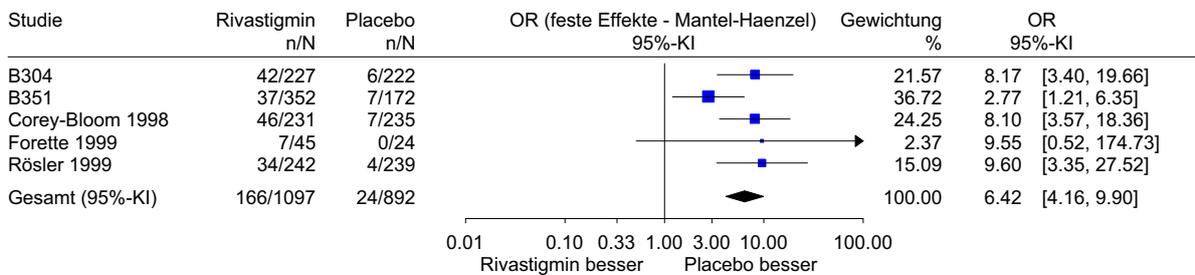
Heterogenität: Q=1.14, df=2 (p=0.566), I²=0%
Gesamteffekt: Z Score=3.54 (p=0.000)

Abbildung 55: Rivastigmin: Meta-Analyse zum Endpunkt „Diarrhoe“ (mit B351)

Rivastigmin - UEs

Endpunkt: Appetitlosigkeit

Distanzmaß: Odds Ratio der Anteile von Patienten mit UE im Studienverlauf



Heterogenität: Q=5.18, df=4 (p=0.269), I²=22.8%
Gesamteffekt: Z Score=8.42 (p=0.000)

Abbildung 56: Rivastigmin: Meta-Analyse zum Endpunkt „Appetitlosigkeit“ (mit B351)

ANHANG H: PROTOKOLL DER WISSENSCHAFTLICHEN ERÖRTERUNG

Protokoll der wissenschaftlichen Erörterung

Zu Stellungnahmen zum Vorbericht A05-19A:

„Cholinesterasehemmer bei Alzheimer Demenz“

Am 14.11.2006 im IQWiG

Teilnehmerliste:

Name	Institution
Dr. Latif Akintade	Eisai GmbH
Dr. Monika Baier	Novartis Pharma GmbH
Dr. Matthias Binek	VFA e. V.
Prof. Dr. Alistair Burns	University of Manchester
Renate Demski	ALZheimer-ETHik e. V.
Dr. Dipl.-Psych. Tobias Drieling	Universitätsklinikum Freiburg
Dr. Anna-Sabine Ernst	IQWiG
Dr. Manfred Gogol	Deutsche Gesellschaft für Gerontologie und Geriatrie
Dr. Dipl.-Psych. Brigitte Grass-Kapanke	privat, Mettmann
Jutta Halbekath	Arznei-Telegramm
PD Dr. Martin Haupt	DGGPP e. V.
Dr. Tatjana Heinen-Kammerer	VFA e. V.
Dipl.-Psych. Julia Hommerich	IQWiG
Prof. Dr. Michael Hüll	Universitätsklinikum Freiburg
Sabine Jansen	Deutsche Alzheimer Gesellschaft e. V.
Dr. Hanna Kaduszkiewicz	Universitätsklinikum Hamburg-Eppendorf
Dr. Thomas Kaiser	IQWiG
Dipl.-Psych. Julia Kreis	IQWiG
PD Dr. Stefan Lange	IQWiG
Friedhelm Leverkus	Pfizer GmbH
Prof. Dr. Wolfgang Maier	Kompetenznetz Demenzen
Dr. Andreas von Maxen	Arznei-Telegramm

Name	Institution
Dr. Angelika Mehnert	Janssen-Cilag GmbH
Prof. Dr. Hans-Jürgen Möller	Hirnliga e. V.
Prof. Dr. Peter Sawicki	IQWiG
Prof. Dr. Klaus Schmidtke	Universitätsklinikum Freiburg
Dr. Helga Schmitz	Eisai GmbH
Dr. Guido Skipka	IQWiG
Dr. Albrecht Stöffler	Merz Pharmaceuticals GmbH
Koen Torfs, MSc	Janssen-Cilag GmbH
Dr. Ferenc Tracik	Novartis Pharma GmbH
Dr. Dr. Stefan Weinmann	Bezirkskrankenhaus Günzburg
Dr. Beate Wieseler	IQWiG
Dr. Hans Wille	Institut für Klinische Pharmakologie, Klinikum Bremen-Mitte
Dr. Yvonne Wirth	Merz Pharmaceuticals GmbH
Dr. Rainer Zerfaß	Pfizer GmbH
Dr. Dipl.-Psych. Thomas Zimmermann	Universitätsklinikum Hamburg-Eppendorf

Tagesordnung

- TOP 1** **Einführung, Vorstellung des Auftrags und der Vorgehensweise des IQWiG**
- TOP 2** **Studienauswahl**
 a) Dauer
 b) Studiendesign
- TOP 3** **Patientenrelevanz der erhobenen Endpunkte**
 a) Angehörige
 b) CIBIC
 c) Klinisch relevante Unterschiede
 d) Responderanalysen
- TOP 4** **Bewertung der Studienqualität**
 a) Concealment of Allocation
 b) Verblindung
 c) Umgang mit fehlenden Werten: Umsetzung des ITT-Prinzips, LOCF als Ersetzungsstrategie
- TOP 5** **Verschiedenes**

Beginn: 10:00 Uhr, Ende: 15:10 Uhr

<i>Redner</i>	<i>Inhalte</i>
TOP 1	Vorstellung des Auftrags und der Vorgehensweise des IQWiG
Sawicki	<p>Begrüßung der Teilnehmer. Erläuterung der Tagesordnung und der Vorgehensweise in der Erörterung. Es würden jeweils die aus Sicht des IQWiG unklaren Aspekte benannt werden mit der Bitte um Beantwortung durch die Teilnehmer sowie der Bitte darum, eine divergente oder auch unterstützende Meinung jeweils darzustellen. Information, dass eine Tonbandaufnahme angefertigt werde, die der Erstellung eines Protokolls diene. Hinweis, dass die allgemeinen Methoden des IQWiG im Rahmen dieser Erörterung nicht zur Diskussion stünden.</p> <p>Auf Rückfragen der Teilnehmer zum Prozedere hinsichtlich des Protokolls Klarstellung, dass ein Inhaltsprotokoll erstellt werde, welches nicht zur Abstimmung gestellt werde und als Produkt des Institutes gekennzeichnet sei. Es werde an den Gemeinsamen Bundesausschuss weitergegeben und im Abschlussbericht veröffentlicht. Eigene Tonbandaufzeichnungen der Teilnehmer seien nicht möglich. Selbstverständlich sei es aber jedem freigestellt, eigene Protokolle anzufertigen. Die Interessenkonflikte würden nicht publiziert, stünden aber während der Sitzung zur Einsicht zur Verfügung, allerdings dürften die Informationen nicht nach außen weitergegeben werden.</p>
Möller	Hinweis, dass es neben Interessenkonflikten im Sinne von Bezahlungen durch die pharmazeutische Industrie auch noch andere Konflikte gebe, wie bspw. die Herausgeberschaft einer Zeitschrift.
Sawicki	<p>Die Interessenkonflikte bezögen sich nicht nur auf die Verbindungen zur pharmazeutischen Industrie, sondern auf alle potenziellen Interessenkonflikte, die vorliegen könnten. Es würden in 8 Punkten die bei internationalen Zeitschriften üblichen Interessenkonflikte abgefragt.</p> <p>Verzicht auf Vorstellungsrunde angesichts knapper Zeit.</p>
Möller	Hinweis auf Schreibfehler bei den Tischschildern.
Demski	Frage nach Übersicht über Interessenkonflikte der Teilnehmer der Erörterung.
Sawicki	Eine solche Übersichtsliste sei nicht vorhanden, der Vorschlag werde aufgenommen.
TOP 2	Studienauswahl
2 a)	Studiendauer
Kreis	Im Vorbericht sei die Dauer von 16 Wochen als Einschlusskriterium festgelegt worden. In den Stellungnahmen sei einerseits geäußert worden, dass man ausschließlich längere (d. h. mindestens 6-monatige) Studien, sowie andererseits, dass man auch noch kürzere (d. h. 3-monatige Studien) hätte berücksichtigen sollen.
Sawicki	Bitte um Wortmeldungen dazu.
Haupt	Je nach betrachteten Symptomen sei eine unterschiedliche Studiendauer angemessen. Stimme dem Vorbericht zu, dass zur Betrachtung der Wirkung auf kognitive Leistungen eine Dauer von 2 oder 3 Monaten zu kurz sei, ab 16 Wochen sei hier angemessen. Bei nicht kognitiven Symptomen seien auch kürzere Zeiträume ausreichend, da es keine Hinweise darauf gebe, dass nicht kognitive Symptome bei einer längeren Behandlung noch stärker verbessert werden könnten, als bereits nach 2 bis 3 Monaten sichtbar sei.
von Maxen	16 Wochen schienen ein Kompromiss zu sein zwischen Zahlen der AkdÄ und der EMEA. Problematisch sei, dass keine sachliche Begründung gegeben worden sei für das gewählte Kriterium. Aus ihrer Sicht sei insbesondere der Langzeitverlauf unter der Behandlung wichtig. Deshalb sei die Vorgabe der EMEA geeignet, nämlich eine Mindestdauer von 6 Monaten.
Wille	Da eine Dauertherapie mit den Präparaten angestrebt werde, seien aus seiner Sicht auch 6 Monate zu kurz. Eine Studiendauer von mindestens einem Jahr sei deshalb angemessen,

<i>Redner</i>	<i>Inhalte</i>
	insbesondere zur Berücksichtigung weiterer patientenrelevanter Endpunkte.
Torfs	Bei gut angelegten Studien mit sinnvollen Endpunkten solle eine Studie auch berücksichtigt werden, egal ob sie nun über 2 Wochen oder 2 Jahre gehe.
Sawicki	Man müsse sich allerdings immer fragen, welche Studiendauer aus klinischer Sicht Sinn mache.
Maier	Prinzip der EBM sei es, alle verfügbare Evidenz zu berücksichtigen, auch wenn Begrenzungen bzgl. der Studiendauer berechtigt sein mögen. Entsprechend einer prinzipientreuen Umsetzung der EBM-Methodik müsse deshalb jede Studie berücksichtigt werden. Außerdem spreche für die Berücksichtigung von 12-Wochen-Studien, dass viele Studien mit methodischen Mängeln behaftet seien und diese natürlich mit zunehmender Studiendauer bei einer solchen schwerwiegenden Erkrankung immer größer würden. Plädiert deshalb für die Berücksichtigung von 12-Wochen-Studien.
Sawicki	Frage, auf welche Referenz er sich bei dieser Aussage bzgl. der Methodik der EBM (Forderung des Einbezugs aller Studienergebnisse bzw. Studiendesigns) beziehe?
Maier	Könne das entsprechende Zitat gerne nachreichen.
Demski	Komme von einer Betroffenen-Organisation. Aus dieser Sicht sei die Diskussion akademisch. Problem sei, dass langfristig Nebenwirkungen aufträten, z. B. Schlaganfälle oder epileptische Anfälle nach 4 oder 5 Jahren Einnahme, und diese Patienten nicht wahrgenommen würden. Es sei überhaupt nicht erforscht, zu welchem Zeitpunkt noch gravierende Nebenwirkungen auftreten würden. Eine sehr frühzeitige Einnahme von Cholinesterasehemmern könne sehr schlimm erkrankte Menschen produzieren. Vorschlag: computergestütztes System zur zentralen Erfassung von Diagnosen, Verschreibungen und Nebenwirkungen.
Gogol	Ausschluss kürzerer Studien stelle eine erhebliche Fehlerquelle im Vorbericht dar. Angesichts der vielfältigen Symptomatologie werde dadurch ein Teil nicht erfasst. Auch kürzere Studien sollten deshalb in gewisser Weise berücksichtigt werden, auch wenn nur zur Erfassung der Responderraten und Komorbidität. Allerdings sei die Datenlage dazu, wie lange Patienten behandelt werden sollten, außerordentlich unbefriedigend, weil nur ein geringer Zeitraum übersehen werden könne. Zum Beispiel sei unklar, was nach einem halben oder einem Jahr geschehe. Dafür würden weitere (längere) Arbeiten benötigt.
Sawicki	Das sei immer das Problem der IQWiG-Berichte, dass man sich mehr Evidenz wünschen würde.
Tracik	Placebo-kontrollierte Langzeitstudien seien ethisch problematisch. Beim Versuch einer länger angelegten Placebo-kontrollierten Studie habe es mit einigen Ethikkommissionen Probleme gegeben, da diese eine Placebo-Behandlung über 16 bzw. 24 Wochen als unethisch angesehen hätten. Verweis darauf, dass es nicht randomisierte Studien zum Langzeitverlauf gebe.
Kaduszkiewicz	Zustimmung zu Herrn von Maxen. Problematisch sei, dass das 16-Wochen-Kriterium willkürlich erscheine. Allerdings kenne sie die 12-Wochen-Studien und sei der Überzeugung, dass deren Einbeziehung die Gesamtaussage des Berichtes nicht verändert hätte.
Möller	Es sei willkürlich, aus dem Spektrum aller RCTs nur solche mit bestimmter Dauer auszuwählen. Plädiert für Einbezug auch von 3-Monats-Studien. Eventuell würde dies sogar die Wahrscheinlichkeit eines positiven Befundes für die Antidementiva reduzieren. Aus klinischen Gründen seien längere Studien tatsächlich wünschenswert, allerdings seien bei einer Dauer von mehr als einem Jahr randomisierte kontrollierte Studien kaum möglich.

<i>Redner</i>	<i>Inhalte</i>
	Problematisch seien bei derartig langen Studien die Drop-outs und die Durchführungspraktikabilität sowie ethische Vorbehalte bei Placebo-kontrollierten Studien.
Grass-Kapanke	<p>Ausschluss von 3-Monats-Studien erscheine ihr nicht sonderlich sinnvoll. Es komme darauf an, was man eigentlich messen wolle.</p> <p>3-Monats-Studien seien eigentlich strenger in der Bewertung eines Effektes, da hier die natürliche Krankheitsprogression nicht so weit fortgeschritten sei, so dass ein Effekunterschied deutlich kritischer beurteilt werde. Gerade zur Betrachtung von Responderanalysen seien 3-Monats-Studien sinnvoll.</p> <p>Studien mit einer Dauer von mehr als einem Jahr seien für Patienten oft nicht akzeptabel. Verweist auf Probleme mit Ethikkommission.</p>
Wille	Gerade aus ethischen Gründen müssten Langzeitstudien gegen Placebo – sowohl aus Wirksamkeitsbeurteilungskriterien als auch aus Sicherheitskriterien – gefordert werden. Wenn sie jetzt nicht eingefordert würden, werde man sie nie mehr bekommen. Man wisse einfach nicht, was nach 2 bis 3 Jahren Behandlung passiere. In den ersten Monaten einer Behandlung zeigten sich lediglich Effekte auf die Symptomatologie, erst bei längerer Dauer könne man den Effekt auf die Krankheitsprogression beurteilen.
Binek	Wenn kürzere Studien dieselben Ergebnisse lieferten, dann hätten sie ja auch aufgenommen werden können – insbesondere vor dem Hintergrund, dass der Kompromiss hinsichtlich der Studiendauer nicht richtig begründet worden sei und nur 27 Studien ausgewertet worden seien.
Demski	Hinweis zur Entscheidung der Ethikkommission, Studien als unethisch abzulehnen: Im Saarland werde derzeit eine Kupfer-Studie von Dr. Pajonk durchgeführt. Die Landesregierung im Saarland habe mit Eisai und Pfizer einen Vertrag „Netzwerk Demenz“ geschlossen. Nun habe die Ethikkommission des Landes beschlossen, dass Kupfer nur in Verbindung mit Aricept gegeben werden dürfe, obwohl es aus ihrer Sicht auch ohne möglich wäre. Ihre Frage sei nun, wie und wo Pharmafirmen Ethikkommissionen beeinflussen.
Sawicki	Diskussion dieser Frage sei hier nicht möglich, da Beschränkung auf spezifische Aspekte des Berichtes notwendig.
Maier	Die Repräsentativität einer Stichprobe nehme mit der Dauer der Placebo-kontrollierten Studie erheblich ab. Auch dies spreche dafür, keine artifizielle untere Grenze zur minimalen Studiendauer zu setzen, sondern alle Studien gleichermaßen zu berücksichtigen. Er werde versuchen, für den Zusammenhang zwischen Studiendauer und Repräsentativität eine Referenz zu liefern.
Möller	Äußert Zustimmung zu Maier. Bei längeren Placebo-kontrollierten Antidementiva-Studien gingen die Drop-out-Quoten linear weiter. Votiert ebenfalls dafür, keine Studien wegen zu kurzer Dauer auszuschließen. Als Wissenschaftler stimme er Herrn Wille voll zu, dass es wichtig sei, Langzeitstudien durchzuführen. Allerdings gehe er als jemand, der Studien durchführe, davon aus, dass Ethikkommissionen entsprechende längere Studiendauern nicht bewilligen würden. Verweis auf entsprechende Diskussion im Bereich der Akuttherapie von Depressionen und auf das Buch von Helmchen und Lauter zu Fragen der Ethik. Eine konkrete Ablehnung einer Ethikkommission in Deutschland sei ihm nicht bekannt.
Wille	Behauptung, dass Ethikkommissionen solche Studien nicht zulassen würden, treffe nicht zu. Ihm sei nicht bekannt, dass jemals eine Placebo-kontrollierte Langzeitstudie zu Antidementiva abgelehnt worden sei.
Tracik	Ihnen sei von einer Ethikkommission in Vorgesprächen zu einer Langzeitstudie für die Firma Novartis ein negatives Votum avisiert worden, dies könne von ihm auch so belegt werden. Daraufhin hätten sie die Studie auf ganz wenige Zentren in Deutschland beschränken müssen. Allerdings habe das vor der AMG-Novelle stattgefunden, ob sich das

<i>Redner</i>	<i>Inhalte</i>
	jetzt geändert habe ...
Lange	<p>Kurze Zusammenfassung aus der Sicht des IQWiG: Bezüglich der Studiendauer sei eine Festlegung notwendig, gleichzeitig habe die Diskussion gezeigt, dass es sowohl Argumente für die Berücksichtigung kürzerer als auch längerer Studien gebe. Eine Begründung sei immer artifiziell, irgendwann müsse man sich einfach festlegen. Verweis auf Hinweis von Kaduszkiewicz, die bereits festgestellt habe, dass sich im Prinzip vermutlich auch bei einer anderen Festlegung sowohl in die eine als auch in die andere Richtung nicht viel geändert hätte.</p> <p>Kommentar zu zwei methodischen Punkten: Längere Studien müssten nicht zwangsläufig größere Mängel aufweisen. Es gebe in der Literatur genügend Langzeitstudien mit exzellentem Design. Auch in Langzeitstudien gebe es vernünftige Wege, mit Therapieabbruchern umzugehen. Die Frage nach der Repräsentativität von RCTs (bzw. zum Vergleich von Langzeit- vs. Kurzzeit-RCTs) sei ein alter Streit – es gebe auch Evidenz, dass sich RCTs von nicht randomisierten Studien gar nicht so sehr unterscheiden würden.</p>
2 b)	Studiendesign
Kreis	<p>Auch zu der Frage nach dem angemessenen Studiendesign seien in den Stellungnahmen unterschiedliche Ansichten deutlich geworden. In der Nutzenbewertung habe man sich auf die Auswertung von RCTs beschränkt, weil diese mit der geringsten Ergebnisunsicherheit behaftet seien.</p> <p>In einigen Stellungnahmen sei darauf hingewiesen worden, dass insbesondere für Endpunkte wie bspw. Pflegeheimweisung Daten aus anderen Studiendesigns als RCTs vorlägen. In den Stellungnahmen sei angemerkt worden, dass deshalb auch Daten aus unkontrollierten Open-label-follow-up-Phasen sowie Vergleichen mit historischen Kontrollen berücksichtigt werden sollten.</p>
Möller	<p>Aus wissenschaftlicher Sicht liege das IQWiG mit der Beschränkung auf RCTs zunächst sehr richtig – RCTs seien der Königsweg, um klare Aussagen zu machen.</p> <p>Allerdings sei dies ja bereits durch die Zulassungsbehörden ausgewertet worden und die Funktion des IQWiG gehe ja über die der Zulassungsbehörden hinaus. Gerade für patientenrelevante Kriterien wie Institutionalisierung solle deshalb auch andere Evidenz berücksichtigt werden, da hierzu lange Placebo-kontrollierte Studien schwierig durchzuführen seien – u. a. weil die Patienten und Angehörigen sich bei wahrgenommener Verschlechterung fragten, ob sie das Medikament oder ein Placebo erhielten, und die Teilnahme an der Studie infrage stellten. Deshalb das Votum, auch andere Designs, die wissenschaftlich nicht so ergiebig seien, sekundär mit einzubeziehen (z. B. naturalistische Studien zu Institutionalisierung), da Placebo-kontrollierte RCTs über 2 oder 5 Jahre nicht machbar seien.</p>
Sawicki	Frage, ob es Daten dazu gebe, wie viele Patienten in doppelblinden, Placebo-kontrollierten Studien richtig errieten, in welcher Gruppe sie seien.
Möller	Dieser Aspekt werde eher nicht erfragt, um die Rater nicht für diesen Aspekt zu sensibilisieren. Allerdings lägen die Rater seiner Erfahrung nach häufig daneben, weil die Placebo-Effekte unterschätzt würden und auch unter Placebo entsprechend der Patientenaufklärung die typischen Nebenwirkungen angegeben würden.
Sawicki	Deshalb habe er nachgefragt – wenn man postuliere, dass Langzeitstudien nicht durchführbar seien, weil Patienten unter Placebo das Medikament wegen Unwirksamkeit häufig absetzten, dann impliziere dies, dass die Patienten die Gruppenzugehörigkeit richtig erraten könnten.
Möller	Das sei so nicht richtig. Sowohl Patienten unter Verum als auch unter Placebo verschlechterten sich im Laufe der Zeit in ihrem Zustand objektiv bzw. subjektiv und setzten wegen der Unsicherheit darüber, ob sie eine wirksame Behandlung erhielten, die

<i>Redner</i>	<i>Inhalte</i>
	Medikation ab.
Halbekath	<p>Frage, was gegen praxisorientierte Placebo-kontrollierte RCTs spreche. Verweis auf CATIE-Studie, die zwar nicht Placebo-kontrolliert sei, aber ein sehr praxisorientiertes Setting gehabt habe. Solange noch Unsicherheit über den patientenrelevanten Nutzen eines Antidementivums auch in der Langzeittherapie bestehe, hänge die Machbarkeit einer Studie auch davon ab, wie die Patienten aufgeklärt würden. Ohne derartige Studien werde man nie eine valide Antwort auf die Frage erhalten, was Antidementiva wirklich bringen.</p> <p>Frage, was gegen eine Nachbeobachtung von Patienten spreche, die die Therapie abbrechen.</p>
Grass-Kapanke	<p>Problematisch sei bei Placebo-kontrollierten Studien die Unsicherheit der Patienten darüber, dass sie möglicherweise kein wirksames Medikament erhielten. Deshalb werde – unabhängig von der Gruppe – bei einer Verschlechterung häufig auf eine garantierte Behandlung gedrängt.</p> <p>Eine entsprechende Nachbeobachtung von Therapieabbrechern werde, soweit möglich, bereits durchgeführt, allerdings sei sie schwierig und auch für die Patienten und Angehörigen sehr aufwändig. So sinke die Bereitschaft der Patienten und Angehörigen, wenn kein direkter Nutzen mehr für sie bestehe.</p>
Torfs	<p>Unsicherheit der Patienten sei in blinden Studien immer ein Problem und führe dazu, dass die Patienten nicht so lange in einer Studie bleiben wollten. Zu Langzeitstudien: Die mediane Zeit bis zur Heimeinweisung betrage 49 Monate, eine entsprechend lange randomisierte, Placebo-kontrollierte Studie sei nicht möglich. Deshalb seien Effekte zur Heimeinweisung nur in Case-Control-Studien nachweisbar.</p>
Sawicki	<p>Frage, ob es Studien gebe, in denen eine entsprechend lange Laufzeit angestrebt worden sei, die aber aus den angeführten Gründen hätten abgebrochen werden müssen.</p>
Torfs	<p>AD2000-Studie. Eine eigene entsprechende Studie gebe es nicht. In den 90er Jahren seien mit anderen Produkten Placebo-kontrollierte 1-Jahres-Studien durchgeführt worden, aber heute sei das nicht mehr möglich.</p>
Maier	<p>Man müsse unterscheiden zwischen dem, was mit methodischem Purismus erreichbar sei und was damit nicht erreichbar sei – dies müsse sich in einem anderen Setting der Real World als Evidenz erweisen. Placebo-kontrollierte Mehrjahresstudien ließen sich, obwohl wünschenswert, nicht erreichen. Die progrediente Natur der Alzheimer Demenz führe zu Verschlechterungen und in der Folge dann zu Studienabbrüchen. Aufgrund der Wirksamkeitsunterschiede zwischen Verum und Placebo seien die Abbruchraten differentiell zwischen den Gruppen mit entsprechendem Auswertungsdilemma. Die vorhandene Evidenz solle gewürdigt werden. Ein Langzeit-Outcome wie Pflegeheimweisung lasse sich nicht im Rahmen einer Placebo-kontrollierten Studie untersuchen (Verweis auf AD2000-Studie), bestenfalls in Vergleichsstudien zwischen 2 Substanzen. CATIE sei ein Beispiel für eine solche Real-World-Studie, für die sie plädierten.</p>
Gogol	<p>RCTs seien der Goldstandard, allerdings auch mit erheblichen Biasfaktoren behaftet. Bei bestimmten Fragestellungen, wie auch Demenzen, seien RCTs möglicherweise nicht möglich oder sehr schwierig. Seiner Meinung nach seien – auch längere – RCTs bei Demenz nicht generell unmöglich, aber sehr aufwändig. Wenn keine RCTs vorhanden seien, müsse man sich die vorhandenen Studien anschauen, aber dabei die vielfältigen Biasfaktoren genau berücksichtigen.</p>
Sawicki	<p>Klarstellung wegen wiederholten Verweises auf Methodik der EBM: Sackett habe sich auf die Situation des einzelnen Arztes bezogen, der bei der Behandlung individueller Patienten die für die spezifische Frage vorhandene Evidenz heranziehen müsse – wenn es keine RCTs dazu gebe, müsse er bis hin zu Expertenmeinungen alles berücksichtigen, um dem Patienten individuell zu helfen. Das IQWiG führe allerdings nicht EBM, sondern Evidence based</p>

<i>Redner</i>	<i>Inhalte</i>
	Health Care durch. Dabei müsse man nicht zwangsläufig auf die beste zur Verfügung stehende Evidenz zurückgreifen, wenn dies bspw. nur eine einzelne Expertenmeinung sei, und auf dieser Grundlage dann bspw. eine Therapie erstatten. Bitte um Unterscheidung dieser verschiedenen Ansätze.
Zerfass	Hinweis, dass Alzheimer Demenz eine fortschreitende Erkrankung sei, die zum Tode führe. Es gebe keine Behandlungsalternative. Bei Placebo-kontrollierten Langzeitstudien enthalte man Patienten eine wirksame Therapie vor. Die Wirksamkeit hätten Zulassungsbehörden in über 80 Ländern festgestellt.
Jansen	Aus Sicht der Angehörigen seien Placebo-kontrollierte Langzeitstudien – auch wenn Erkenntnisse über die Spätfolgen wünschenswert wären – nach fast 10 Jahren, die die Medikamente auf dem Markt seien, und angesichts der Tatsache, dass sie in den Leitlinien aller Fachgesellschaften stünden, nicht mehr vermittelbar. Könne sich nicht vorstellen, dass man überhaupt noch genügend Angehörige und Patienten zur Studienteilnahme rekrutieren könne.
Sawicki	Bestätigung, dass nach langer Verfügbarkeit von Präparaten auf dem Markt eine adäquate Evaluation sehr schwierig sei.
Kaiser	Hinweis, dass RCTs im Sinne von Real World auch ohne besonders strenge Ein- und Ausschlusskriterien möglich seien. Das Argument der Placebo-Kontrolle werde dadurch entkräftet, dass in den Bericht auch offene Studien aufgenommen worden seien. Darüber hinaus wären auch z. B. Studien zu Cholinesterasehemmern im Vergleich zu Best Supportive Care mit aufgenommen worden. Mit entsprechender Aufklärung über die noch offenen Fragen hinsichtlich Langzeitnutzen und -sicherheit, die ja im Unterschied zur Wirksamkeit von den Zulassungsbehörden nicht geprüft worden seien, seien dann auch längere Studien möglich.
Möller	Hinweis auf Mitgliedschaft in Gremien des BfArM und der EMEA. Bitte an Kaiser, Nutzen (in Abgrenzung zur von den Zulassungsbehörden festgestellten Wirksamkeit) genauer zu definieren.
Kaiser	Beispielsweise Institutionalisierung sei kein Kriterium, das die Zulassungsbehörden forderten.
Möller	So genannte Real-World-Studien seien nicht wirklich erfolgreich gewesen, Verweis auf Kritikpunkte bei der AD2000- und der CATIE-Studie. Solange keine anderen verfügbar seien, sollten deshalb naturalistische Daten verwendet werden, wo in Ex-post-Analysen versucht worden sei, den medikamentösen Effekt auf die Pflegeheimweisung darzustellen. Eine Placebo-kontrollierte 6-Monats-Studie zu Donepezil habe hinsichtlich Pflegeheimweisung keine signifikanten Unterschiede gezeigt, es gebe aber zu Donepezil auch einige naturalistische Studien.
Wille	Placebo-kontrollierte Studien oder Studien gegen Best Supportive Care seien durchführbar. Bisherige Argumentation, dass man den Patienten gegenüber dies nicht mehr vertreten könne, weil eine wirksame Therapie vorliege, sei etwas fahrlässig: Wirksamkeit sei nach wie vor unklar und nur durch entsprechende Studien herauszubekommen. Angesprochene Probleme könnten bspw. durch pragmatische Studien mit ITT-Analyse berücksichtigt werden und die möglicherweise auftretenden Verdünnungseffekte müssten dann bei der Interpretation entsprechend berücksichtigt werden. Plädiere insbesondere hinsichtlich der Endpunkte Pflegeheimweisung, Komorbidität und Mortalität für kontrollierte Studien gegen Placebo oder gegen Best Supportive Care.
Leverkus	Gehe heute nicht darum, eine Studie zu planen, sondern um die Frage, inwiefern dem IQWiG Evidenz aus Nicht-RCTs zur Entscheidungsfindung dienen könne. Verweis auf Zitat von Sawicki beim VFA-Symposium. In dieser Situation liege ja Evidenz aus RCTs bereits vor, deshalb könne dann zu anderen Aspekten, die in RCTs schwierig abzubilden seien, auch andere Evidenz berücksichtigt werden, wenn diese auch einen geringeren

<i>Redner</i>	<i>Inhalte</i>
	Stellenwert habe.
Sawicki	Klarstellung zum von Herrn Leverkus genannten Zitat: Er habe damals darauf verwiesen, dass, wenn man nicht randomisierte Designs zur Beurteilung des Nutzens heranziehe, man sich dann über die möglichen Fehler und die sich daraus ergebenden Implikationen klar sein müsse.
Möller	Hinweis, dass es auch bei naturalistischen Studien Probleme durch Drop-outs gebe.
Tracik	Frage an Herrn Kaiser, ob er tatsächlich offene randomisierte Langzeitstudien anerkennen würde (Zustimmung durch Herrn Kaiser). Anmerkung zu Herrn Wille, dass Wirksamkeit belegt sei, schließlich seien die Präparate zugelassen.
Wille	Zulassung bedeute nicht Nutznachweis, sondern Verkehrsfähigkeit und ausreichende Sicherheit.
Maier	Anregung, zu diskutieren, ob randomisiertes Design, das entwickelt sei für den Wirksamkeitsnachweis, auch das optimale Design sei für die Nutzenbewertung.
Lange	Wirksamkeit sei eine notwendige, jedoch nicht hinreichende Bedingung für Nutzen. RCTs seien entwickelt worden, um interpretierbare Daten zu erhalten, unabhängig ob für Wirksamkeit, Nutzen oder Schaden. Es sei in jedem Setting möglich zu randomisieren. In solchen Diskussionen zu RCTs werde immer auf den Bereich der Kardiologie verwiesen, jedoch stürben auch dort Patienten und trotzdem seien solche Studien möglich. Entscheidend sei, dass durch ein vernünftiges Design eine möglichst große Ergebnissicherheit erreicht werde – insbesondere, wenn man es mit relativ moderaten Effekten zu tun habe.
Demski	Trotz langer Markteinführung seien bei entsprechender Aufklärung der Patienten auch jetzt noch entsprechende Studien möglich.
Sawicki	Alle wesentlichen Aspekte der Dauer und des Studiendesigns seien nun angesprochen worden. Anmerkung, dass man bei unvollständiger Evidenz immer den möglichen Schaden aufgrund von Tun und den aufgrund von Nichts-Tun abwägen müsse.
Möller	Verweis auf Nikolaus-Urteil.
PAUSE	
Tracik	<i>Einreichung eines Schreibens der Ethikkommission der Bayerischen Landesärztekammer an die Novartis Pharma GmbH, in dem bzgl. einer geplanten Studie die Auffassung vertreten wird, dass eine Placebo-Gabe über 8 Monate nicht vertretbar sei. Entgegennahme durch Herrn Sawicki.</i>
Sawicki	Bedankt sich, ggf. werde man nachfragen, ob auch die Koordinationsstelle der Ethikkommissionen diese Auffassung vertrete.
TOP 3	Patientenrelevante Endpunkte
3 a)	Angehörige
Kreis	Im Amendment zum Berichtsplan sei unterschieden worden zwischen angehörigen- und patientenrelevanten Zielen. Dabei seien die Endpunkte „Lebensqualität von (betreuenden) Angehörigen“ sowie „Betreuungszeiten“ als angehörigenrelevant eingestuft worden. Natürlich sei dem IQWiG bewusst, dass bei einer Alzheimer Erkrankung die Belastung der Angehörigen sehr groß sei, und auch, dass eine solche Unterscheidung nicht immer ganz trennscharf sein könne. Dennoch stehe bei der Nutzenbewertung die Perspektive der Patienten im Fokus.
Jansen	Aus Sicht der Deutschen Alzheimer Gesellschaft hätten diese angehörigenrelevanten Endpunkte auch für die Patienten einen entscheidenden Nutzen. Gehe davon aus, dass das, was den Angehörigen entlaste und dessen Lebensqualität steigere, letztendlich auch den

<i>Redner</i>	<i>Inhalte</i>
	Patienten zugute komme, u. a. wenn so ein möglichst langer Verbleib zu Hause ermöglicht werde. Deshalb würden sie diese Punkte gerne stärker berücksichtigt sehen.
Grass-Kapanke	Es bestehe eine starke Korrelation zwischen der Lebensqualität der pflegenden Angehörigen (Caregiver) mit den – insbesondere nicht kognitiven – Symptomen der Alzheimer Demenz. Die Höhe des Betreuungsaufwandes sei eng verbunden mit den Fähigkeiten im Bereich der Alltagsaktivitäten. Insgesamt bestehe also eine enge Verknüpfung zwischen diesen verschiedenen Aspekten. Die AD betreffe nicht nur die Patienten, sondern in erheblichem Ausmaß auch die Caregiver.
Maier	Angesichts der methodischen Schwierigkeiten, die Pflegeheimweisung in randomisierten Designs zu betrachten, könne der Pflegeaufwand der Betreuer als Proxy hierfür dienen und solle deshalb besonderes Gewicht erhalten. Er könne nachvollziehen, dass das Gesundheitssystem zunächst patientenorientiert vorgehe und nicht das Wohlergehen der Angehörigen finanziere. Betonung, dass er aus diesen methodischen Schwierigkeiten bzgl. der Pflegeheimweisung auf den Aspekt des Pflegeaufwands fokussieren würde, weniger auf Aspekte wie Lebensqualität der Angehörigen.
Gogol	Demenz betreffe auch das soziale Umfeld. Deshalb müsse in einer Nutzenbewertung auch betrachtet werden, wie durch Interventionen die Angehörigen entlastet werden könnten, auch möglicherweise im Hinblick auf einen volkswirtschaftlichen Nutzen.
Haupt	Die Gesundheit des Pflegenden und des Patienten seien sehr eng verknüpft. Über eine Änderung des Verhaltens des Pflegenden könne sich auch eine Veränderung beim Demenzkranken ergeben (Mediatoreneffekt). Hinweis auf das Dilemma beim Endpunkt Heimeinweisung. Heimeinweisung als Endpunkt sei ein inhomogenes Konstrukt, da zum Teil Angehörige zu diesem Schritt auch ermuntert würden und es auch sehr gute Heime mit modellhaft guter Pflege gebe.
Schmidtke	Gibt zu bedenken, dass die Belastung der Pflegenden ja ausschließlich eine Folge der Krankheit des Kranken sei. Deshalb sei es nicht ganz nachvollziehbar, warum hier eine zusätzliche Varianz eingebracht werde, denn schließlich müssten die psychopathologischen Symptome und die Krankheitsschwere die Varianz der Angehörigenbelastung größtenteils abdecken. Auch bei anderen Erkrankungen mit schwerer Angehörigenbelastung werde diese nicht üblicherweise in die Kriterien einbezogen, weil man sich auf den Kranken fokussiere.
Kaduszkiewicz	Auch wenn solche Endpunkte relevant seien, seien sie doch nur in einer Minderheit der Studien erhoben worden und dabei meist in sehr unterschiedlicher Weise. Somit sei die Evidenz hinsichtlich der angehörigenrelevanten Endpunkte sehr viel schlechter als die Evidenz bezüglich der Kognition der Patienten. Damit könne man zwar über die Berücksichtigung dieser Endpunkte diskutieren, aber es führe sowieso nicht zu mehr positiver Evidenz.
Maier	Zu Herrn Schmidtke: Stattdessen könne man den Sachverhalt auch so betrachten, dass der Angehörige als derjenige, der dies am besten beurteilen könne, den durch den Patienten verursachten Pflegebedarf einstufe. In diesem Sinne würde dann keine neue Varianzquelle im Sinne einer zusätzlich betrachteten Person eingeführt, sondern der Angehörige fungiere lediglich als Rater.
Lange	Klarstellung, dass auch die als angehörigenrelevant eingestuften Endpunkte im Bericht aufgegriffen und dargestellt worden seien, dass sie aber für die Nutzenbewertung eben – nach dem Patienten – an zweiter Stelle stünden. Dieses Vorgehen sei durch die Diskussion bestätigt worden. Bei einer Korrelation zwischen Lebensqualität der Angehörigen und Merkmalen des Patienten könne man sich fragen, warum dies nicht direkt bei den Patienten erhoben werde. Frage nach der Validität des Betreuungsaufwandes als Surrogat für Pflegeheimweisung. Bestätigung, dass Pflegeheimweisung möglicherweise nicht unbedingt immer ein Negativendpunkt sein müsse. Hinweis, dass das IQWiG in dem

<i>Redner</i>	<i>Inhalte</i>
	Bericht nicht die GKV-Perspektive, sondern die Patientenperspektive vertrete.
Weinmann	Betonung, dass angehörigerelevante Ziele in dem Bericht berücksichtigt worden seien. Allerdings seien ihm Studien, die eine gute Korrelation zwischen Lebensqualität der Angehörigen und Symptomen der Patienten zeigten, nicht bekannt. Die Studie von Smith und anderen habe gezeigt, dass die Beurteilung der Lebensqualität der Betroffenen durch die Angehörigen nicht unbedingt immer der eigenen Einschätzung durch die Patienten entspreche.
Demski	Es sei richtig, zunächst die Patienten in den Fokus zu stellen. Es gebe auch immer mehr Menschen bei ihnen in der Beratung, bei denen kein Angehöriger greifbar sei. Zur Validität des Kriteriums Heimeinweisung – viele Faktoren spielten hierbei eine Rolle, z. B. ob jemand alleine lebe oder auch die Einstellung des Angehörigen.
Jansen	Hinweis, dass die Eignung der Messinstrumente zur Erfassung patientenrelevanter Endpunkte generell überdacht werden sollte. Heimeinweisung bspw. sei leicht messbar, die Erfassung von Lebensqualität aber sehr schwierig. Verweis auf eigene Erfahrungen bei der Evaluation eines Schulungsprogramms mit dem SF-36.
Möller	Die Belastung der Angehörigen sei eine wichtige Größe. Trotzdem könne er die Position des IQWiG verstehen, dass Ziele aus Sicht der Angehörigen hintangestellt worden seien, da es tatsächlich eine Reihe von messtechnischen Problemen gebe. Hinweis darauf, dass Antidementiva schon im Zulassungsverfahren eine hohe Zulassungsschranke hätten (2- bzw. 3-Schwellen-Mechanismus). Vergleich der NNTs bei Antidementiva und ASS-Therapie.
Gogol	Psychopathologie der Patienten sei sehr unterschiedlich und mache jeweils für die Angehörigen mehr oder weniger Probleme. Je nachdem könne Heimeinweisung manchmal ein guter Endpunkt sein, manchmal aber auch eine sehr gute Lösung darstellen, um die häusliche Situation zu erleichtern.
3 b)	<i>CIBIC</i>
Kreis	Möglicherweise sei der Eindruck entstanden, der CIBIC als Instrument zur globalen Beurteilung des klinischen Gesamteindrucks sei nachträglich herabgestuft worden. Richtigstellung, dass der CIBIC vielmehr nachträglich als Endpunkt aufgenommen und dargestellt worden sei, allerdings habe er in der Nutzenbewertung weniger Gewicht erhalten.
Weinmann	Zunächst habe man entschieden, den CIBIC wegen der unklaren psychometrischen Eigenschaften und wegen der unklaren Patientenrelevanz nicht zu berücksichtigen. Später habe man in der Diskussion entschieden, ihn als zusätzlichen Ergebnisparameter mit aufzunehmen, allerdings stehe er ihrer Ansicht nach nicht im Vordergrund.
Maier	Bei einer sehr globalen, komplexen Erkrankung wie der AD sei ein Globalmaß als Ergänzung der spezifischen Parameter unerlässlich. Tatsächlich sei der CIBIC psychometrisch nicht wie eine Skala analysierbar. Dies sei allerdings kein Argument gegen die Verwendung, vielmehr müsse die psychometrische Theorie hier weiterentwickelt werden.
Zerfass	Er sei sehr befremdet gewesen über die Einschätzung bezüglich des CIBIC. Verweis auf Forderung des CIBIC durch FDA und EMEA und auf die Entstehung des CIBIC in der Zusammenarbeit zwischen FDA und der Firma Parke-Davis. Er sei entwickelt worden mit dem Zweck, die klinische Relevanz von Veränderungen auf spezifischen Skalen zu messen. In den verschiedenen Skalen würden immer einzelne Symptome abgefragt. Im CIBIC könne man dagegen scores, wenn ein Patient berichte, dass er sich insgesamt besser fühle und wieder stärker am Leben teilnehmen könne. Verweis auf ein mitgebrachtes Exemplar des CIBIC. Das Instrument sei wenig bekannt, weil es im klinischen Alltag keine Rolle spiele. Erläuterung der verschiedenen vom CIBIC erfassten Aspekte.

<i>Redner</i>	<i>Inhalte</i>
Möller	Zustimmung zu Vorrednern. Der CIBIC spiegele die Sichtweise des Arztes in der Klinik. Er habe eine hohe Validität und bilde sozusagen „arztbeurteilte Patientenrelevanz“ ab. Der CIBIC sei auch wichtig als Brücke zur Interpretation des ADAS-cog und werde deshalb ja von FDA und EMEA gefordert. Das IQWiG müsse die Methodiken der Zulassungsbehörden primär akzeptieren.
Burns (auf Englisch)	<i>CIBIC bilde das ab, was man im klinischen Alltag jeden Tag mache, und erfasse die verschiedenen Symptombereiche. Er sei reliabel und valide. Es sei eine Stärke des CIBIC, dass er den klinischen Eindruck abbilde. Er sei komplementär zur Erfassung der verschiedenen Symptombereiche und deshalb enorm wichtig.</i> <i>Bemerkungen zur AD2000-Studie: Sie habe positive Ergebnisse gezeigt für Kognition und Aktivitäten des täglichen Lebens, allerdings sei sie unterpowerd gewesen für die anderen Endpunkte (Pflegeheimweisung und Erreichen einer vorher definierten Verschlechterung). Die Patienten seien aufgrund der Ein- und Ausschlusskriterien nicht repräsentativ gewesen.</i>
Grass-Kapanke	CIBIC werde im Alltag nicht angewendet, weil er sehr aufwändig in der Durchführung sei. Aufgrund der fortschreitenden Degeneration werde nur die Hälfte des Punkteranges ausgeschöpft (Erhalt bzw. Verschlechterung). Beschreibung verschiedener Maßnahmen zur Erhöhung der Interrater-Reliabilität (Protokollierungen, Schulungen, Einsatz externer Rater). Insgesamt gebe es sicherlich noch Entwicklungsbedarf hinsichtlich der psychometrischen Skalen, allerdings gebe es keine Forschungsförderung zur Entwicklung neuer Skalen. Daher seien die zur Verfügung stehenden Skalen mit der größtmöglichen Sorgfalt zu verwenden und die zurzeit vorhandenen Verfahren seien durchaus aussagekräftig.
Sawicki	Hinweis auf wachsende Bereitschaft, unabhängige Gelder der Solidargemeinschaft für solche Zwecke zur Verfügung zu stellen.
Zimmermann	Nachfrage an Befürworter des CIBIC, warum in den Studien immer die 1- bis 7er-Skala mit dem entsprechend undifferenzierten Ergebnis dargestellt werde. Dies sei für die psychometrische Weiterverarbeitung und Bewertung der Ergebnisse sehr problematisch.
Wille	CIBIC sei ein wichtiges Instrument, weise aber verschiedene Probleme auf: Dass er in der Praxis nicht angewendet werde, sei ein großer Nachteil, außerdem sei er sehr anfällig für Verzerrungen bei nicht ausreichender Verblindung. Hinweis auf Studie Rockwood 2006, bei der sich eine Diskrepanz gezeigt habe zwischen der positiven Beurteilung der Zielerreichung durch die Ärzte auf der einen und der negativen Beurteilung durch Patienten und Angehörige auf der anderen Seite. Beurteilungen durch den Arzt und durch die Patienten bzw. Angehörigen könnten also durchaus differierende Ergebnisse bringen.
Lange	Am CIBIC störe ihn ein gewisser paternalistischer Ansatz, da dieser davon ausgehe, dass die Ärzte immer wüssten, was gut für den Patienten sei. Dies sei der wichtigste Vorbehalt gegenüber dem CIBIC gewesen, dass er nicht direkt patientenzentriert sei. Auch angesichts der Anforderungen der Zulassungsbehörden wolle sich das IQWiG doch eine gewisse Autonomie erlauben. Diskussion sei interessant und spannend gewesen, habe allerdings aus seiner Sicht den direkten Patientenbezug nicht aufgezeigt.
Maier	Zur Frage von Herrn Zimmermann: CIBIC sei keine Skala, sondern eine Bewertung mit Mittelwert und Standardabweichung. Er diene sozusagen als Validierungskriterium, welches den klinischen Ablauf mimt. Durch vorgegebene Fragen etc. werde der Bias eingeschränkt. Er sei ein Arzturteil und habe natürlich etwas Paternalistisches, da der Arzt sage, was relevant sei und dies entsprechend bewerte. Dabei werde er in der Regel den Patienten mitberücksichtigen.
Burns (auf Englisch)	<i>Hinweis auf die Unterschiede zwischen der Goal Attainment Scale, die Perspektiven der Patienten und Angehörigen erfasse, und dem CIBIC, der die globale Veränderung beim Patienten reflektiere.</i>

<i>Redner</i>	<i>Inhalte</i>
Möller	Unterschied, der sich zwischen Behandlungsgruppen zeige, hänge auch ab von der Variabilität der Messwerte. Zur Gewährleistung der Rater-Konstanz im CIBIC seien spezifische CIBIC-Rater eingeführt worden.
Binek	Frage, ob die vielen konstruktiven Beiträge überhaupt noch berücksichtigt werden könnten oder ob es bereits zu spät sei, da der Vorbericht ja bereits veröffentlicht sei. Schlägt vor, zukünftig so wichtige Fragen bereits im Vorfeld zu klären.
Sawicki	Dem Institut seien das Messinstrument und die angesprochenen Probleme und Vorteile natürlich bekannt gewesen. Das Institut verfare so, dass, wenn sich wichtige Aspekte aus einer Anhörung ergäben, diese so weit berücksichtigt würden, wie dies erforderlich sei – methodisch und auch zur Verbesserung der Qualität des Berichtes. Im Extremfall könne dies dazu führen, dass ein völlig neuer Vorbericht erstellt würde mit einem neuen Stellungnahmeprozess, auch wenn das in diesem Projekt nicht der Fall sein werde. Die Anhörungen und Erörterungen seien durchaus sehr ernst zu nehmen. Auch nach dem Abschlussbericht sei es nicht zu spät. Wenn bspw. in einigen Jahren eine sehr gute Langzeitstudie veröffentlicht würde, die die Ergebnisse des Abschlussberichtes widerläge, dann werde dem Gemeinsamen Bundesausschuss mitgeteilt, dass der Abschlussbericht neu evaluiert und ggf. um die neuen Daten ergänzt werden müsse. Es sei nicht zu spät für die Aufnahme relevanter Aspekte in den Bericht.
3 c)	<i>Klinisch relevante Unterschiede</i>
Kreis	In den Stellungnahmen sei eine Vorabdefinition klinisch relevanter Unterschiede gefordert worden, obwohl gleichzeitig darauf hingewiesen worden sei, dass es derzeit keinen Konsens über eine klinisch relevante Punktwertdifferenz gebe. Frage an die Stellungnehmenden, wie ein Vorschlag aus ihrer Sicht ausgesehen hätte.
Maier	Aus rein statistischer bzw. biometrischer Sicht sei die Frage klar – klinisch relevant sei jeder Effekt, der größer als eine halbe Standardabweichung sei. Dies sei der Fall beim ADAS-cog mit einer Effektgröße von einer halben Standardabweichung; ebenfalls von Relevanz, wenn auch von schwacher, sei die Besserung bei den Alltagsaktivitäten mit einer etwas geringeren Effektgröße von etwa 0,3. Etwas störend im Vorbericht sei, dass dort häufiger Mittelwertsunterschiede genannt würden, ohne jedoch die Standardabweichung anzugeben.
Grass-Kapanke	Hinweis, dass bei der Interpretation die Eigenschaften der jeweiligen Skala berücksichtigt werden müssten. So sei der Punkterange 0–70 im ADAS-cog sehr theoretisch. Patienten in Studien hätten meistens Werte zwischen 15 und 20 Punkten; bei Punktwerten über 30 seien die Patienten bereits sehr massiv beeinträchtigt. Damit sei das relevante Punktespektrum deutlich reduziert. Außerdem sei der ADAS kein lineares Instrument. Beim NPI habe der Summenscore gar keine Aussagekraft; hier sei es wichtiger, die einzelnen Symptombereiche zu betrachten. Auch bei ADL-Skalen, die üblicherweise von den Angehörigen ausgefüllt würden, gebe es verschiedene Probleme. So falle den Angehörigen die Bewertung der Fragen insbesondere anfangs recht schwer, außerdem verändere sich das Bewertungsverhalten über die Dauer einer Studie. Ein Interview durch einen geschulten Rater sei hier wünschenswert. Die zur Verfügung stehenden Skalen seien prinzipiell in Ordnung, allerdings müsse man sie genau kennen und richtig bewerten können.
Torfs	Diskussion im Vorbericht zum Effekt auf dem ADAS-cog sei sehr balanciert und Fußnote auf S. 166 setze das sehr gut in den Kontext. Die 3 bis 4 Punkte Änderung müsse man in Zusammenhang mit dem Ausgangswert der Patienten sehen, dann zeige sich deren Bedeutung (Größenordnung vergleichbar mit dem Effekt auf der PANSS-Skala für Schizophrenie). Verweis auf die Rockwood-Studie, in der die meisten Patienten ihre Therapieziele im Bereich der Kognition und der Alltagsaktivitäten formuliert hätten. Somit sei Kognition für die Patienten relevant und die Veränderung auf der Skala auch.

<i>Redner</i>	<i>Inhalte</i>
Jansen	Die Punkte sagten nichts darüber aus, was die Patienten, die eine gute Response hätten, bei dieser Verbesserung tatsächlich empfänden. Im Vorbericht sei auch dargelegt, dass wenig bekannt sei über die Menschen, die besser, und diejenigen, die schlechter profitierten. Man müsse alles dafür tun, dass diejenigen, die davon profitierten, dies auch weiterhin könnten. Verweis auf persönliche Erfahrung mit einer 41-jährigen Patientin, die von Verbesserungen (erhöhte Konzentrationsfähigkeit, geringere Aggressivität) nach der Einnahme von Medikamenten berichtet habe – Frage, auf welchen Instrumenten sich dies überhaupt abbilde und wie viele Punkte dies jeweils ausmache.
Zerfass	Betonung, dass ADAS-cog kein reiner Gedächtnistest sei. Kognition umfasse Auffassung, Konzentration, Sprache, Rechnen, Praxie. Es könne durchaus sein, dass ein Punkt auf der Skala hoch relevant sei für einen Patienten. Verweis auf Demenzpatienten, der in einem Krankenhaus wegen Desorientierung zu Tode gekommen sei – ggf. hätte hier ein Punkt auf der Skala einen Unterschied gemacht. Ein Punkt könne sein, ob ein Patient noch Auto fahren könne, für einen anderen Patienten könne es sein, dass ein Punkt gar nichts bedeute. Bei manchen Patienten könne ein Punkt über die weitere Existenz in seinem sozialen und beruflichen Leben entscheiden.
Möller	Patientenurteil sei in späteren Stadien der Erkrankung sehr schwierig, da die Patienten kaum mehr dazu in der Lage seien. Verweis auf Vorgehensweise der Zulassungsbehörden, die klinische Relevanz durch Doppelung oder Verdreifachung der Outcome-Ebenen sicherzustellen. Betrachtung der Brücke zwischen den 3 Instrumenten (Korrelationen) und der Responderanalysen sei relevant. Verweis auf die Situation hinsichtlich der Patientenrelevanz in der internen Medizin (Beispiel ASS).
Akintade (auf Englisch)	<i>Die Domänen Kognition, Funktion und Verhalten sollten nicht als unabhängig voneinander betrachtet werden, vielmehr überlappten sie. Kognition sei das Wichtigste, da Funktion und Verhalten davon abhängen.</i>
Schmidtke	Effektstärke sei der einzige universal anwendbare Effektmaßstab, deshalb sei keine Diskussion über die Bedeutung von 2 oder 3 oder 4 Punkten notwendig. Kritik daran, dass in Pharmastudien Effektstärken häufig nicht ausreichend beschrieben seien. Die Beobachtung besonders starker Verbesserung einzelner Patienten solle angesichts sehr großer Variabilität und niedriger Reliabilität mit großer Zurückhaltung aufgenommen werden. Die mittleren Veränderungen seien kleiner als der Unterschied zwischen 2 Messzeitpunkten bei einer Person.
Haupt	Erläuterung zu Charakteristika der Skalen in Abhängigkeit vom Schweregrad. Beim neuropsychiatrischen Inventar seien je nach Schweregrad der Erkrankung unterschiedliche Ausprägungen in den verschiedenen Subskalen zu erwarten. Deshalb sei es wichtig, nicht den Globalscore, sondern die Subskalen zu betrachten.
Halbekath	Es bestehe eine Inkonsistenz zwischen dem Bericht und dem Fazit. Zitat aus Vorbericht S. 8, wo dargestellt werde, dass es umstritten sei, ab welchem Ausmaß einer Beeinflussung (bezogen auf eine Skala) von einem Nutzen gesprochen werden könne. Auch am Ende des Berichtes werde infrage gestellt, ob die Frage nach dem patientenrelevanten Nutzen der Antidementiva überhaupt anhand der Skalen beantwortet werden könne. Es komme einem Salto gleich, dass dann im Fazit von einem Nutzen der Antidementiva in Hinsicht auf die Kognition gesprochen werde. Die vorher geäußerten Zweifel kämen im Fazit gar nicht mehr vor. Der Begriff „Nutzen“ habe auch im Alltagsgebrauch eine eindeutig positive Konnotation und sei nicht zu vereinbaren mit so erheblichen Zweifeln. Im Fazit des Abschlussberichtes solle sich widerspiegeln, dass die Frage, ob die beobachteten Effekte tatsächlich patientenrelevant seien, nicht habe beantwortet werden können.
Sawicki	Einwand sei aus wissenschaftlicher Sicht berechtigt. In den ersten Berichten habe man jeweils die Zweifel und das Für und Wider dargestellt. Schließlich gebe es nie die Wahrheit, sondern immer einen Grad des Irrtums. Allerdings fordere das Ministerium eine klare, unmissverständliche Empfehlung – das bedeute, das IQWiG müsse sich eindeutig festlegen,

<i>Redner</i>	<i>Inhalte</i>
	ob man einen Nutzen sehe oder nicht. Hinweis darauf, dass das IQWiG die Wissenschaft in eindeutiges gesundheitspolitisches Handeln übersetzen müsse.
3 d)	Responderanalysen
Kreis	In den Stellungnahmen sei angemerkt worden, dass Responderanalysen eventuell zu einer Überschätzung des Therapieeffektes führen könnten. Bitte an die Stellungnehmenden des Arznei-Telegramm und der Hamburger Arbeitsgruppe um eine Erläuterung ihrer Bedenken.
Lange	Die Beschreibung der Responderanalysen sei ein Versuch gewesen, Mittelwertsdifferenzen zwischen Gruppen eine andere Interpretationsmöglichkeit auf der individuellen Ebene zu geben. Dabei sei es möglich, als Kriterium eine Veränderung um 4 Punkte oder gar keine Verschlechterung oder eine Verbesserung um 7 Punkte heranzuziehen, allerdings liefen diese jeweils auf Ähnliches hinaus. Man könne sich natürlich methodisch darüber streiten, inwiefern dies jeweils sinnvoll sei.
Kaduszkiewicz	Es sei völlig richtig, Responderanalysen zu versuchen. Problematisch sei allerdings die dafür notwendige Annahme einer Normalverteilung. Falls die Verteilung schief sei, würde das Ergebnis anders aussehen. Deshalb hätten sie diese Annahme kritisiert, da man nicht davon ausgehen könne, dass es tatsächlich eine Normalverteilung geben müsse. In den Studien seien oft nicht genügend Daten für entsprechende Analysen angegeben (Beispiel CIBIC-plus). Deshalb der Appell, dass man, statt Annahmen zu machen, lieber die Originaldaten einfordern solle.
Lange	Nachfrage, ob es nun um ADAS-cog oder CIBIC gehe.
Kaduszkiewicz	Responderanalysen könne man für beide Kriterien machen.
Lange	Dies sei doch ein relevanter Unterschied. Beim CIBIC als eher ordinalen Skala sei die Auswertung nach Kategorien schon sinnvoller als die Betrachtung von Mittelwertsunterschieden. Beim ADAS-cog stelle sich dies noch etwas anders dar. Hinweis darauf, dass häufig fälschlicherweise Mittelwertsunterschiede zwischen Gruppen verglichen würden mit individuellen Erfolgskriterien. Beispielsweise könne es bei einem Mittelwertsunterschied von 3 Punkten durchaus einen mehr oder weniger großen Anteil von Patienten geben, die sich verschlechterten, aber auch solche, die sich relativ stark verbesserten. Aus diesen Responderanalysen könne man dann NNTs berechnen, ein Vergleich mit NNTs bei ASS sei allerdings aus seiner Sicht nicht adäquat.
Möller	Responderanalysen dienten hauptsächlich dazu, mehr Einblick in die Daten zu bekommen und zu sehen, wie viele Patienten wirklich profitierten.
von Maxen	Habe mehrere Einwände gegen Responderraten: Bei geringem Mittelwertsunterschied könne die Setzung eines Cut-off-Wertes, der genau zwischen diesen beiden Responderraten liege, zu einer Überschätzung des Therapieeffektes führen. Verweis auf Mühlhauser-Arbeit zur unterschiedlichen Beurteilung des Therapieerfolgs anhand verschiedener Ergebnisdarstellungen.
Lange	Das hänge aber auch von der Variabilität ab.
von Maxen	Ja, das hänge auch von der Variabilität ab. Das weitere Problem bestehe aber darin, dass man auch bei der Responserate nicht wisse, was sie eigentlich bedeute. Schließlich lege man hier etwas zugrunde (2, 3 oder 4 Punkte), bei dem nicht klar sei, wie klinisch relevant das eigentlich sei. Demnach sei auch bei daraus abgeleiteten NNTs nicht klar, was diese bedeuteten, weil nicht klar sei, was die 4 Punkte bedeuteten – möglicherweise bedeuteten diese auch bei jedem Patienten etwas anderes. Somit seien solche NNT-Berechnungen auch nicht vergleichbar mit denen aus anderen Bereichen. Selbst wenn man NNTs berechnen wolle, müsse man dann auch NNHs betrachten, also wie vielen Patienten es mit der Behandlung schlecht gehe. Dabei komme man auf Grundlage der vorhandenen Daten zu NNHs von 5 hinsichtlich Übelkeit bei Donepezil unter flexibler Dosierung.

<i>Redner</i>	<i>Inhalte</i>
Skipka	Anmerkung zur Argumentation von Kaduszkiewicz. Eine schiefe Verteilung könne unter einer Normalverteilungsannahme nicht nur zu einer Überschätzung des Effektes, sondern auch zu einer Unterschätzung führen. Wie auch im Vorbericht beschrieben, hätten sie die zu erwartenden und berichteten Responseraten verglichen und dies habe – bis auf eine Studie – gut gepasst.
Lange	Auch wenn die Verteilung des ADAS-cog schief sei, betrachte man hier ja doch Differenzen – und die seien nach dem zentralen Grenzwertsatz immer normal verteilt. Die Responderanalysen seien letztlich ein Versuch gewesen, sich dem Ganzen zu nähern. Als wichtig erachte er immer noch die Frage, die ja auch vorhin zur Diskussion gestellt worden sei, auf die es aber nur eine teilweise Antwort gegeben habe, was 1, 3 oder 7 Punkte im Einzelfall bedeuten könnten.
Maier	Warne davor, in diesem Zusammenhang Responderanalysen überzubewerten. In anderen Bereichen der Medizin leiteten sie sich, ebenso wie NNTs, von diskreten Endpunkten ab, die es aber hier nicht gebe – der einzige diskrete Endpunkt sei Heimeinweisung und die Problematik hierzu sei bereits diskutiert worden. Bei allen anderen Endpunkten sei das Responsekriterium willkürlich als Cut-off-Wert auf einer Skala definiert. Der Begriff suggeriere aber, dass er so etwas abbilde, wie Frau Jansen vorher gemeint habe – dass es Patienten gebe, die davon profitierten, und solche, die nicht davon profitierten. Er plädiere deshalb dafür, die Responderanalysen in diesem Feld, in dem es keine diskreten Endpunkte gebe, bei einer solchen Nutzenbewertung auf jeden Fall niedrig zu hängen.
Akintade (auf Englisch)	<i>Da AD eine neurodegenerative Erkrankung sei, würden in einigen Ländern bei Responderanalysen bereits kleinere als zu erwartende Verschlechterungen als relevanter Nutzen für die Patienten betrachtet. Eisai habe für NICE eigene Responderanalysen mit sehr stringenter Responsedefinition durchgeführt (Patienten mit keiner Verschlechterung oder mit Verbesserung im Bereich Kognition und funktioneller oder behavioraler Verbesserung). Analysen hätten einen Nutzen für Cholinesterasehemmer in allen 3 Domänen gezeigt. Verweis auf zusätzliche eigene Sensitivitätsanalysen zu verschiedenen Skalen im Bereich Kognition, Funktion und Verhalten, die konsistente Ergebnisse gezeigt hätten und so den Effekt nochmals belegten – genauso wie die Tatsache, dass die Ergebnisse für die verschiedenen Cholinesterasehemmer recht ähnlich seien.</i> <i>Daten seien vom Medical Research Council unabhängig ausgewertet worden und auf der NICE-Website verfügbar. Deren Fazit sei gewesen, dass die Analysen robust seien und dass Cholinesterasehemmer in allen 3 Domänen einen Benefit zeigten.</i>
Burns (auf Englisch)	<i>Verweis auf die Schwierigkeiten der Übersetzung der Responderdefinition in die klinische Praxis. Patienten in der eigenen Klinik wünschten sich häufig einfach, dass der Zustand stabil bleibe und sich nicht weiter verschlechtere.</i>
Gogol	Wie in anderen Bereichen der Medizin sei es auf jeden Fall notwendig, abzuschätzen, was man mit einem therapeutischen Prinzip erreiche und welche Patienten auf die Therapie ansprächen und welche nicht. Man könne sich über die methodischen Schwierigkeiten bei der Erfassung bzw. Quantifizierung unterhalten, aber benötigt würden Responderanalysen auf jeden Fall.
Stöffler	Betonung, dass bei dieser Erkrankung individuell bereits die Stabilisierung bzw. die Verlangsamung der Verschlechterung ein sinnvolles Therapieziel darstellen könne.
Schmidtke	Allgemein würde ein Effekt von einer halben Standardabweichung als relevant angesehen. Damit sei der entsprechende Punktwert ein sinnvolles Kriterium für eine Response. Alles andere sei dann eben weniger als eine halbe Standardabweichung; damit sei die Diskussion bei einer normalverteilten Größe artifiziell. Anders sei es beim CIBIC-plus, bei dem es den Endpunkt „Verbesserung oder keine Verschlechterung“ gebe. Bei dieser ordinalskalierten Größe sei eine Mittelwertbildung nicht sinnvoll. Häufig würden in Publikationen sowohl Response- als auch Mittelwertanalysen berichtet.

<i>Redner</i>	<i>Inhalte</i>
	Wenn zunächst festgelegt worden sei, den Mittelwert auszuwerten, sei die Betrachtung der Responderanalysen wenig aussagekräftig.
	<i>Pause</i>
TOP 4	Bewertung der Studienqualität
4 a)	<i>Concealment of Allocation</i>
Kreis	In einigen Stellungnahmen sei auf die Bewertung der Studienqualität eingegangen worden, insbesondere auch auf die Bewertung des Concealment of Allocation. In den Publikationen sei das Concealment of Allocation häufig nicht genau beschrieben worden, so dass dieses mit „unklar“ bewertet worden sei. Man sei davon ausgegangen, dass in tatsächlich doppelblinden Studien die Beschreibung des Concealment eine weniger zentrale Rolle spiele als in offenen Studien. Hinweis auf die Möglichkeit, im Rahmen dieser Erörterung zu einem besseren Verständnis diesbezüglich zu kommen. Deshalb Bitte an die Hersteller, das jeweilige Vorgehen genauer zu beschreiben.
Schmitz	Der Aspekt des Concealment sei von der Firma aufgegriffen und durch eine genaue Analyse der Clinical Study Reports untersucht worden. Sei überzeugt, dass die Mängel ausgeräumt werden konnten. Im Anhang der Stellungnahme sei das Concealment für alle Studien, bei denen Unklarheiten bestanden, genau beschrieben worden.
Lange	Sie hätten für alle Studien beschrieben, dass die Randomisierung computergeneriert erfolgte und dass die Medikation dadurch verblindet sei, dass die aktiven Substanzen und Placebo identisch erschienen und verpackt gewesen seien. Nachfrage, ob also die Zentren mit Packungen versehen worden seien, die nur eine Nummer hatten?
Schmitz	Das sei für jede Studie einzeln beschrieben. Bei z. B. zulassungsrelevanten Studien habe dies natürlich noch keine Handelsware sein können, das sei dann mit Nummern versehen gewesen.
Lange	Seiner Ansicht nach sei eine solche Beschreibung ausreichend.
Wille	Das seien nun Informationen, die ihnen nicht vorgelegen hätten.
Kreis	Hinweis, dass diese Informationen in den Stellungnahmen enthalten seien.
Wille	Verweis auf systematische Übersichtsarbeiten, die zeigten, dass auch ein unklares Concealment zu einer Effektüberschätzung von 30 % führe. Wenn das für alle Studien zufrieden stellend geklärt sein solle, sei das ein Punkt, den er nun nicht mehr weiter beurteilen könne.
Tracik	Könne sich Frau Schmitz nur anschließen, bei ihnen sei es genauso gelaufen. Hinweis darauf, dass viele der Studien von den Zulassungsbehörden auditiert und auch als GCP-konform bewertet worden seien. Sehe als Kritikpunkt, dass in den Publikationen zum Teil nicht ausreichend darauf eingegangen werde; in neueren Publikationen würden auch Aspekte des Concealment eindeutig dargestellt.
Wille	Bitte an Herrn Tracik, die genauen Maßnahmen zur Sicherstellung des Concealment zu konkretisieren.
Tracik	Indem sie die ganze Studie GCP-gerecht durchführten, indem sie eine Quality Assurance hätten, also eine eigene Abteilung, die genau das Concealment überwache, indem sie eine zentrale Computerrandomisierung durchführten. Auch wenn das in den Publikationen zum Teil nicht gut abgebildet sei, könne er versichern, dass die Studien GCP-gerecht gemacht worden seien.
Torfs	Er könne das für jede einzelne Studie vorlesen, das sei alles in den Studienprotokollen sehr gründlich beschrieben und auch so durchgeführt. Bei Johnson & Johnson gebe es eine eigene professionelle Gruppe mit hunderten von

<i>Redner</i>	<i>Inhalte</i>
	Mitarbeiten, die ausschließlich mit der Durchführung von klinischen Studien beschäftigt seien. Das laufe alles GCP-konform und werde auditiert; wenn jemand einen Fehler mache, werde er entlassen. Randomisierungs- oder Concealmentfehler seien nicht akzeptabel. Wenn man Zentren finde, die die Studie nicht GCP-gemäß durchführten, dann würden diese getrennt ausgewertet und in der nächsten Studie nicht eingeschlossen. Fragen zur Studienqualität ließen sich zu 99 % aus den Protokollen und Berichten klären, auch wenn nicht alles einfließe in die Publikationen.
Lange	Nachfrage, worum es sich bei den Briefumschlägen handele, von denen in der Stellungnahme von Janssen berichtet werde.
Leverkus	In den Firmen werde es häufig so gemacht, dass es Notfallumschläge gebe, die nur im Notfall geöffnet würden.
Mehnert	Bestätigt, dass es sich um die Umschläge handele, die geöffnet würden, wenn eine Entblindung stattfinden müsse.
Torfs	Ergänzung: aus Sicherheitsgründen.
Lange	Es sei allerdings nicht beschrieben, dass alles identisch verpackt sei.
Mehnert	Doch, da stehe: „Similarity in appearance of dispensed medication.“
Torfs	„Identical in appearance, taste and smell.“
Kaduszkiewicz	Hinweis, dass sich dies auf die Medikamente, aber nicht die Briefumschläge beziehe.
Grass-Kapanke	Auf den Umschlägen stehe üblicherweise jeweils nur die Randomnummer, also die Nummer, die auch auf den Medikamentenpackungen stehe.
Sawicki	Frage an Wille, ob damit die Frage nach dem Concealment aus seiner Sicht beantwortet sei.
Wille	Diese Angaben könne er jetzt nur so hinnehmen.
Lange	Klarstellung, dass bisher im Bericht auch nicht davon ausgegangen worden sei, dass das Concealment nicht gegeben sei, sondern dass es bisher unklar gewesen sei. Er kenne die Arbeit von Schulz, nach der auch unklares Concealment ein Problem sei. Allerdings habe Schulz in seiner Arbeit eher Studien aus anderen Bereichen und auch eher nicht zulassungsrelevante Studien betrachtet.
Wille	Verweis auf weitere Arbeiten neben Schulz, z. B. von Jüni et al. Man solle ein adäquates Concealment doch auch für tatsächlich doppelblinde Studien fordern. Appell, dies auch in den Publikationen entsprechend zu beschreiben.
Sawicki	Fazit, dass das IQWiG die für den Bericht notwendigen Informationen erhalte.
4 b)	Verblindung
Kreis	Auch der Aspekt der Verblindung sei in den Stellungnahmen immer wieder angesprochen worden. Die Verblindung der Endpunkterheber sei in den Publikationen häufig nicht genau beschrieben, wenn überhaupt, dann meist lediglich in Bezug auf den CIBIC-plus-Rater. In den Stellungnahmen seien von den Herstellern zusätzliche Informationen hierzu geliefert worden. Auch hier wolle man gerne noch einige Aspekte bei den Herstellern nachfragen. Z. B. sei von Novartis festgestellt worden, dass der CIBIC-plus-Rater verblindet gewesen sei gegenüber anderen Messdaten, allerdings seien keine Angaben gemacht worden zur Verblindung der Rater des ADAS-cog und der Aktivitäten des täglichen Lebens, bspw. gegenüber Nebenwirkungen.
Tracik	Die ADAS-cog-Daten seien in den meisten Studien durch Psychologen durchgeführt worden, die nicht auf der Station gearbeitet hätten und somit häufig überhaupt nicht mit den Patientenakten in Kontakt gekommen seien. Sie hätten auch nicht nachfragen sollen, ob Nebenwirkungen bestünden oder nicht. Beim CIBIC-plus sei dies eindeutig geregelt, da

<i>Redner</i>	<i>Inhalte</i>
	habe der Rater gar nicht von der Abteilung sein dürfen.
Sawicki	Frage, ob weitere Unklarheiten bestünden.
Kreis	Bei der Stellungnahme der Firma Eisai sei gesagt worden, dass nur für die Burns-Studie aus den internen Dokumenten hervorgehe, dass der CIBIC-Rater verblindet sei. Deshalb die Nachfrage, wie dies in den anderen Studien ausgesehen habe.
Zerfass	Es sei Voraussetzung für die Durchführung des CIBIC, dass der Durchführende von allem anderen keine Ahnung haben dürfe. Deswegen gebe es spezielle Mannschaften, die extrem geschult würden und durch das Land reisten.
Sawicki	Nachfrage, ob sichergestellt werde, dass dies auch so passiere, bzw. ob es eine Überprüfung gebe.
Tracik	Dies werde durch das Monitoring sichergestellt. Es gebe Monitore, die das in entsprechenden Monitoring-Berichten festhielten; Verstöße müssten entsprechend GCP angegeben werden, dies werde auch von den Behörden überprüft.
Kreis	Nachfrage zu S. 9 der Eisai-Stellungnahme. Dort stehe: „The blinding of endpoint raters is only described for the CIBIC-plus in the Burns-Study.“ Deswegen liege es nahe, für die anderen Studien nochmals nachzufragen. Zitat aus der Stellungnahme zur Verblindung der ADAS-cog-Rater.
Schmitz	Dies sei der gleiche Fall wie von Herrn Tracik beschrieben. ADAS-cog werde von Personen erhoben, die normalerweise keinen Zugang zu den anderen Daten hätten. Zu den anderen Studien habe man keine genauere Beschreibung im CSR gefunden, auch diese Studien seien aber nach dem üblichen Vorgehen und den normalen klinischen Gegebenheiten durchgeführt worden, wie von Herrn Zerfass beschrieben.
Sawicki	Frage, ob diese Angaben genügten.
Kreis	Man nehme das so auf.
Kaduszkiewicz	Hinweis darauf, dass die Verblindung der Erheber auch gegenüber den Nebenwirkungen extrem wichtig sei. Dies gehe immer etwas unter, wenn man davon spreche, dass etwas „normalerweise“ so sei. Allerdings enthalte der CIBIC die Frage nach anderen Krankheiten des Patienten, so dass mit einer hohen Wahrscheinlichkeit hier auch auftretende Nebenwirkungen genannt würden.
Wille	Frage, ob mal untersucht worden sei, wie hoch die Rate des richtigen Erkennens der Gruppenzuteilung gewesen sei.
Leverkus	Das von Wille angesprochene Konzept sei relativ neu, seit 2004 gebe es die ersten Aufsätze dazu. Deshalb habe es das damals nicht gegeben und sei nicht durchgeführt worden. Außerdem sei eine solche Vorgehensweise auch gar nicht so allgemein anerkannt.
Möller	Ein solches Vorgehen setze auch Artefakte, da es die Sensibilität der Rater fördere, eventuell direkt nach solchen distinguierenden Nebenwirkungen zu suchen.
Wille	Bemerkung, dass man eine solche Erfassung natürlich zum Abschluss und nicht während einer Studie machen solle.
Möller	In Antidementiva-Studien werde bereits ausgesprochen viel verblindet – schließlich sei ja der Patient bereits verblindet, der Arzt durch das Placebo-Kontrollverfahren per se auch. Weitere Forderungen bedeuteten einen Exzess von Methodik, wie es ihn sonst im Bereich der klinischen Psychopharmakologie nicht gebe.
Torfs	Wenn man am Ende nach der Gruppenzugehörigkeit fragen würde, dann stünde dies natürlich auch im Protokoll. Entsprechend werde während der ganzen klinischen Studie versucht zu raten und die Studie somit verzerrt. Zitiert die Beschreibung der Verblindung der ADAS-cog-Rater in der Studie Tariot 2001

<i>Redner</i>	<i>Inhalte</i>
	von Seite 4 der Anlage zur Stellungnahme. Man könne natürlich nie ausschließen, dass es mal nicht so laufe wie vorgesehen. Aber so stehe es im Protokoll und so würden auch die mitmachenden Ärzte trainiert sowie die Monitors gebrieft.
Leverkus	Verweist auf das Dilemma: Wenn man am Ende der Behandlung den Patienten frage, dann tippe er bei einem wahrgenommenen Nutzen auf das Verum und bei keinem wahrgenommenen Nutzen auf das Placebo. Wenn man eine Korrelation feststelle, könne dies also auch darauf zurückzuführen sein, dass der Patient durch das Medikament einen Nutzen erfahre. Günstiger sei es deshalb eigentlich, während der Behandlung zu fragen. Dies führe allerdings wiederum dazu, dass versucht werde, die Gruppenzugehörigkeit zu erraten und man damit die Verblindungsbemühungen untergrabe. Insgesamt sei dies eine methodisch wichtige Frage, die aber nicht einfach zu lösen sei.
Wille	Hinweis, dass ein solches Vorgehen natürlich nicht die Lösung des Problems sei, dass aufgrund der spezifischen Nebenwirkungen eine Postrandomisierungsentblindung zu erwarten sei. Trotzdem könne ein solches Verfahren aber sinnvoll sein, um ein Gefühl für dieses Problem zu bekommen.
Demski	Frage, warum von den Herstellern nachträglich noch Daten angefordert worden seien. Schließlich könnten die Hersteller sich dadurch schützen, dass sie bestimmte Daten nicht weitergäben. Kritisiert Verfahren, nachträglich Daten einzufordern, die die Hersteller in irgendeiner Form verändern könnten.
Sawicki	Die Firmen hätten eine moralische Verpflichtung gegenüber den Patienten, die an der Studie teilgenommen hätten, alle Daten offenzulegen; schließlich nähmen die Patienten an einer Studie teil, um zur Verbesserung der Behandlung beizutragen. Ohne eine Publikation der Ergebnisse entfalle unter Umständen die Grundlage für die Zustimmung der Patienten zu der Studie.
Demski	Frage, warum es nicht vorab sehr gute, vollständige Publikationen gebe. Bei Nachforderungen von Informationen, die nur bei den Pharmafirmen lagerten, stelle sich die nahe liegende Frage, ob diese nicht eventuell nur ein Teil oder leicht veränderte Informationen geben würden.
Sawicki	Man wisse tatsächlich von Firmen, die sich nach Abschluss von Studien weigerten, diese Daten zur Verfügung zu stellen. Dadurch könne eine Bewertung des IQWiG nur in eine positive Richtung verzerrt werden, da die Firmen positive Ergebnisse nicht zurückhalten würden.
Demski	Frage, ob IQWiG solche Studien, von denen keine Daten vorlägen, quantitativ auch in den Bericht habe einfließen lassen.
Kreis	Im Rahmen dieses Projektes sei eine Studie, B351, von Novartis nicht zur Verfügung gestellt worden, dies sei auch dargestellt worden im Bericht.
Tracik	Daten einer Übersichtsarbeit negativer Art seien auch mit eingerechnet.
4 c)	<i>Umgang mit fehlenden Werten</i>
Lange	Es gebe in dieser Hinsicht 2 relevante Aspekte – zum einen die Tatsache, dass fehlende Werte aufträten, die nicht in die Analyse eingingen, und damit verbunden die Frage, wie dies in der Qualitätsbewertung berücksichtigt werde, und zum anderen das angewendete LOCF-Verfahren zur Ersetzung fehlender Werte. Hinsichtlich der Qualitätsbewertung könne man sich darüber streiten, welcher Anteil an Patienten, die randomisiert worden, aber (trotz Ersetzungsstrategie) nicht in die Auswertung eingegangen seien, noch akzeptabel sei. Im Vorbericht seien 2 Grenzen festgelegt worden, die beschrieben, ab welchem Anteil an fehlenden Patientendaten dies als problematisch betrachtet worden sei. Im Vorbericht sei dies mit 10,5 % bzw. 4,5 Prozentpunkten Unterschied zwischen den Gruppen beschrieben. Man sei von den Werten 10 % und 5 %

<i>Redner</i>	<i>Inhalte</i>
	ausgegangen – lediglich durch die Berücksichtigung von Rundungen seien nun diese krummen Zahlen entstanden, da man bspw. ab 10,5 % aufrunde. Eigentlich habe es 10 % und 5 % heißen sollen. Aber auch da seien Einwände gekommen, z. B. von der Deutschen Gesellschaft für Allgemeinmedizin, warum man nicht 11,7 % genommen habe. Auch das werde betrachtet, ob sich dadurch etwas an den Ergebnissen ändere.
Skipka	Sensitivitätsanalysen hätten gezeigt, dass auch entsprechende Verschiebungen dieser Grenzen nach oben oder unten keine Änderung an der grundsätzlichen Aussage mit sich brächten.
Lange	Zu Begründung könne man also sagen, dass die Bestimmung einer Grenze notwendig gewesen sei und es in der Literatur keine eindeutige Festlegung hierzu gebe. Beschreibung des Dilemmas bei der LOCF-Methodik: Auf der einen Seite gebe es fehlende Werte, auf der anderen Seite könne es bei einer chronisch progredienten Erkrankung durch differentielle Abbruchraten zu einer künstlichen Verzerrung kommen. Frage, ob es andere Lösungsvorschläge zum Umgang damit gebe.
Schmidtke	Vorschlag, statt LOCF mittlere Verläufe der Placebo-Patienten einzusetzen.
Maier	Hinsichtlich der Bewertung des ITT-Prinzips gebe es keine perfekte Lösung – man müsse es so oder so ähnlich machen, wie das IQWiG es gemacht habe. Problematisch bei LOCF seien insbesondere differentielle Drop-out-Raten zwischen Verum und Placebo, wobei die Verluste in der Verumgruppe wegen der initialen Nebenwirkungen meist größer seien. Dies führe zu Verzerrungen, so dass LOCF unbefriedigend sei. Auch wenn die Zulassungsbehörden LOCF forderten, müsse sich das IQWiG dem ja nicht anschließen. Regressionsmethoden, die die Progression der Krankheit berücksichtigten würden, seien vielleicht vielversprechender.
Kaduszkiewicz	Zum Hinweis von Herrn Skipka bezüglich der festgelegten Grenzen: Selbst wenn sich dadurch im IQWiG-Bericht nichts verändere, bleibe doch die Frage, was Drop-out-Raten in den Studien selber bewirkten, was sich dann im Bericht gar nicht niederschlage. Hinweis auf Empfehlung, statt LOCF den Mittelwert der Placebo-Gruppe zu nehmen. Dies werde auch bei Winblad 2006 gemacht, so dass sich diese Diskussion allmählich auch in den Studien niederschlage. Generell müsse man sich aber fragen, inwiefern diese hohen Drop-out-Raten von 20 % und mehr überhaupt noch eine Ersatzstrategie möglich machten. Ohne Originaldaten könne man aber letztendlich nie korrekt einschätzen, ob die Ergebnisse dadurch verfälscht würden – diese erhalte man aber von den Herstellern nicht.
Leverkus	LOCF sei sicherlich problematisch. Verweis darauf, dass LOCF auch von den Zulassungsbehörden gefordert werde. Hinweis auf beispielhafte Durchführung eines anderen Ersetzungsverfahrens, das ebenfalls die gleichen Ergebnisse erbracht habe wie die LOCF-Analyse. Deshalb glaube er, dass LOCF durchaus valide sei. Die Meta-Regression des IQWiG im Bericht habe gezeigt, dass dadurch keine relevante Verzerrung zu erwarten sei. Frau Kaduszkiewicz habe in ihrer eigenen Meta-Regression die Studien mit Qualitätsmängeln nicht berücksichtigt. Diese Mängel seien ja nun geklärt, so dass diese Studien nun wieder eingeschlossen werden könnten. Davon unabhängig bestätige deren eigene Analyse das LOCF-Prinzip: Sie argumentiere ja, dass eine Verzerrung durch die höhere Abbruchrate in der Verumgruppe zustande komme. Jedoch werde an der Stelle, wo die Regressionsgerade das Interzept schneide (die Abbruchrate in beiden Gruppen also gleich sei), ein Effekt von Cohen's $d = 0,4$ geschätzt, der mit dem Effekt von $d = 0,5$ sehr vergleichbar sei.
Kaduszkiewicz	Zur Aussage von Herrn Leverkus, dass die Bewertung „grober Mangel“ sich durch die Diskussion erledigt habe – dies sehe sie nicht so. Diese Bewertung sei im Wesentlichen durch eine relevante Abweichung vom ITT-Prinzip begründet gewesen und dies sei bisher nicht aus der Welt geschafft.

<i>Redner</i>	<i>Inhalte</i>
Leverkus	In den Stellungnahmen hätten sie doch die kritisierten Punkte aus den CSR berichtet und hätten gezeigt, dass dort keine Abweichungen seien.
Lange	Wie auch Frau Kaduszkiewicz bereits gesagt habe, hätten Aspekte wie das Concealment in der Bewertung ja weniger große Bedeutung erhalten. Insofern könne eine Bewertung mit „grob mangelhaft“ nicht durch zusätzliche Informationen zur Frage des Concealment nun deutlich verbessert werden. Einige Studien seien mit „grobem Mängeln“ beurteilt worden, weil es eine relevante Abweichung vom ITT-Prinzip gegeben habe. Diese Bewertung werde sich nicht durch zusätzliche Informationen zum Concealment ändern.
Leverkus	Bei den Studien, die im Verdacht gestanden hätten, dass das ITT-Prinzip verletzt sei, hätten sie in der Stellungnahme Daten geliefert und damit hoffentlich das IQWiG überzeugt, dass dies nicht der Fall sei. Sonst müsse man darüber noch mal reden.
Lange	Dies betreffe teilweise Daten von Eisai, aber man habe nicht zu allen Studien solche Informationen erhalten.
Kaduszkiewicz	Eine beispielhafte Analyse mit einem anderen Ersetzungsverfahren sei sehr gut, allerdings stelle sich die Frage, ob dies auch in anderen Studien zu demselben Ergebnis geführt hätte. Beim Versuch, den Einfluss der Drop-outs in Simulationsrechnungen zu analysieren, habe sich gezeigt, dass sich dies letztlich nur mit Originaldaten für jede einzelne Studie wirklich beurteilen lasse.
Hüll	Eine ITT-Verletzung lasse sich kaum heilen. Es sei sehr schwierig zu beurteilen, inwiefern es tatsächlich systematische Verzerrungen gebe. Die Annahme, dass die LOCF-Analyse bessere Werte bringe, werde durch die Daten gar nicht immer bestätigt. Teilweise seien die LOCF-Werte schlechter als die der Observed Cases am Ende der Studie. Dies liege daran, dass man sich in den Studien unterschiedlich bemüht habe, den letzten Wert festzustellen. Auch die Länge der Studien habe einen Einfluss darauf, inwiefern es tatsächlich zu Verzerrungen kommen könne. Hilfreich für die Zukunft wäre es, wenn Patienten zum Studienende noch mal einbestellt würden.
Möller	Hinweis, dass LOCF-Analyse die Standardanalyse der Zulassungsbehörden sei, gleichzeitig werde aber auch die Per-Protokoll-Analyse verlangt. Die jeweiligen möglichen Fehler würden bei der Interpretation beachtet. Es gebe neuere Analysemethoden, die von den Zulassungsbehörden etwas skeptisch betrachtet würden (z.B. MMRM). Es sei sinnvoll, auch solche Verfahren einzusetzen, allerdings würde man daraus nicht unbedingt klüger. Darstellung von Frau Kaduszkiewicz, dass Studien mit mehr als 20 % Drop-out in beiden Studienarmen schlecht seien, widerspreche seiner Auffassung. Es gebe viele Studien mit – im Sinne der Studiendurchführung – schwierigen Patientengruppen (z. B. Schizophrenie- oder Demenzerkrankte), die solche Drop-out-Raten aufwiesen – in manchen Bereichen komme man an der Realität nicht vorbei. Dies bringe zwar methodische Probleme mit sich, die man dann aber lösen müsse.
Zimmermann	Es gehe nicht darum, dass es Drop-outs in den Studien gebe, sondern darum, dass über diese Patienten nicht weiter berichtet werde. ITT verlange nicht, dass es keine Drop-out-Raten gebe, sondern dass die eingeschlossenen Patienten auch am Ende ausgewertet würden.
Möller	Hinweis darauf, dass es unterschiedliche ITT-Definitionen gebe. (<i>Verabschiedet sich.</i>)
Baier	Zustimmung, dass LOCF nicht optimal sei. Zu einer Studie gebe es Daten zu Retrieved Drop Outs, die allerdings angesichts des geringen zur Verfügung stehenden Platzes in Publikationen nicht veröffentlicht worden seien. Es handele sich um die Studie, bei der sie im IQWiG-Bericht wegen einer Verletzung des ITT-Prinzips im CIBIC kritisiert worden seien. Die Analyse der Retrieved Drop Outs zeige, dass diese Ergebnisse mit der LOCF-Auswertung übereinstimmten.

<i>Redner</i>	<i>Inhalte</i>
Sawicki	Hinweis darauf, dass, wenn Daten in einer Publikation nicht untergebracht werden könnten, diese ja auch immer im Internet zur Verfügung gestellt werden könnten.
Baier	Der Studienreport könne auch nachgereicht werden.
Sawicki	Genau. Es sei immer schön, wenn man so etwas in Vollständigkeit vorliegen habe.
Stöffler	Plädoyer, die Wirklichkeit der Studiendurchführung nicht aus dem Blick zu verlieren. Insbesondere bei – wie heute morgen gefordert – längeren Studien werde eine Obergrenze von bspw. 20 % Drop-outs sofort als absurd deutlich. Zu Retrieved-Drop-out-Analysen: Nach seiner Erfahrung sei es sehr schwierig, Patienten zu einer Nachuntersuchung zu bewegen – in einer eigenen Studie seien weit unter der Hälfte der Patienten dazu bereit gewesen.
Sawicki	Erläuterung der Begriffsdefinition von Drop-outs zur Abgrenzung von Studienabbruchern. Entscheidend sei die Studienbegleitung, um die Patienten zur Nachuntersuchung zu motivieren.
Stöffler	In Demenzstudien sei es kaum möglich, alle Patienten zu einer Nachuntersuchung zu bewegen, da die Untersuchungen sehr zeitaufwändig (z. B. 30 Minuten für ADAS-cog) und auch belastend für den Patienten seien.
Maier	Ein Problem der Nachuntersuchungen bei Demenzkranken sei auch, dass es bei Patienten im fortgeschrittenen Stadium schwierig sei, einen Informed Consent einzuholen.
Sawicki	Dies müsse doch beim Einschluss in die Studie eingeholt worden sein und gelte dann für die gesamte Studie.
Wille	Eine abschließende Untersuchung sei in Studien immer vorgesehen und dem werde also auch zu Beginn zugestimmt.
Grass-Kapanke	Informed Consent könne bei einer Nachuntersuchung nach einem oder anderthalb Jahren nicht einfach vorausgesetzt werden. Für die Patienten seien die Untersuchungen, zu denen sie meist von ihren Angehörigen gebracht werden müssten, sehr aufwändig und stellten eine erhebliche Belastung dar.
Demski	Die erhobenen wissenschaftlichen Bedenken sollten öffentlich gemacht werden. Die Frage, wie es dem Patienten – bspw. nach 10 Jahren – auf verschiedensten Ebenen gehe, solle als Endpunkt aufgenommen werden.
Maier	Frage, warum die Nutzenbewertung auf der Grundlage von Publikationen erfolge und nicht – wie bei den Zulassungsbehörden – auf der Grundlage von Originalstudien. Damit ließen sich auch Fragen wie bspw. die zur Relevanz der Verletzung des ITT-Prinzips eher abschließend klären.
Lange	Verweis auf die Rahmenvereinbarung mit den Herstellern über die Überlassung von vertraulichen Daten. Eine Auswertung von Originaldaten sei angesichts der zur Verfügung stehenden Ressourcen kaum leistbar. In Publikationen sollten die wichtigen Informationen enthalten sein. Dass es den Patienten zu schlecht gehe, um die Tests durchzuführen, könne kein Argument sein, warum Patienten nicht nachbeobachtet würden – schließlich sei dies eine relevante Information. Möglicherweise seien andere, pragmatischere Endpunkte sinnvoll, die mit weniger Aufwand erhoben werden könnten. Auch wenn man die Grenzen bezüglich der Drop-out-Raten willkürlich ziehen könne, so seien doch Raten von 30 bis 40 % wohl schon sehr, sehr problematisch.
Burns (auf Englisch)	<i>Gründe für das Drop-out seien vielfältig und es sei sicherlich sinnvoll zu versuchen, möglichst viele Patienten zu erfassen. Es gebe allerdings auch eine Gruppe von Personen, die tatsächlich ihre Einwilligung zurückziehe. In solchen Fällen sei eine Nachverfolgung auch aus Sicht der Ethikkommissionen sehr problematisch.</i>

<i>Redner</i>	<i>Inhalte</i>
Kaduszkiewicz	Wenn es so schwierig sei, Patienten zu einem Besuch zu bewegen, sei es womöglich auch sinnvoll, sie zu Hause zu besuchen.
Leverkus	Man möge bei der Bewertung beachten, dass die Studien aus der Mitte der 90er Jahre stammten.
TOP 5	Verschiedenes
Sawicki	Danke für die Diskussion der vielen wichtigen Aspekte und die Klärung von möglichen Missverständnissen. Bitte um Diskussion der Interpretation der Ergebnisse.
Maier	Diskussion sei hervorragend gewesen. Nutzen definiere sich aus seiner Sicht als eine erweiterte Wirksamkeit. Damit gebe es einen klaren Nutzen für die Cholinesterasehemmer, der auch biometrisch abgesichert sei als Effektgröße von einer halben Standardabweichung für den ADAS-cog und auch als Effekt bei den Alltagsfunktionen. Die aus Sicht des IQWiG vorhandenen Defizite in der Studienlage sollten beschrieben werden, damit auch eine Verantwortung der öffentlichen Hand zur Finanzierung von Studien, die für das IQWiG hinlänglich informativ seien, bestehe.
Sawicki	Es gebe 3 mögliche Ergebnisse einer Bewertung: 1. Der Nutzen sei belegt. 2. Es sei belegt, dass es keinen Nutzen gebe. 3. Es sei unklar, ob ein Nutzen bestehe. Zurzeit noch sei die dritte Alternative mit einer Einschränkung der Erstattung verbunden. Allerdings sei aus seiner Sicht stattdessen wünschenswert, dass es in diesen Fällen die Möglichkeit für den G-BA gebe, die Erstattung unter Auflagen (z. B. im Rahmen von versorgungsrelevanten Studien) zu gewähren.
Maier	Betonung, dass auf der gegenwärtigen Grundlage der Nutzen nachgewiesen sei. Allerdings habe niemand in den bisherigen Studien den Nutzen richtig nachweisen können, weil niemand danach gefragt worden sei. Die Industrie habe bisher die Wirksamkeit nachweisen müssen und habe dies konsequent getan. Weil vor dem IQWiG niemand die entsprechenden Fragen gestellt habe, gebe es auch solche Studien nicht. Deshalb solle man in der Übergangphase etwas liberaler sein in der Nutzenbewertung, auch im Interesse der Patienten.
Sawicki	Betonung, dass es nicht möglich sei zu sagen, dass der Nutzen aufgrund wie auch immer gearteter Erfahrung da, aber noch nicht in Studien belegt sei. Denn dies impliziere, dass es einen Erkenntnisweg außerhalb eines kontrollierten Vergleiches gebe, der den Nutzen belege.
Maier	Zurzeit gebe es noch ein relatives Unwissen, was einen eigenständigen Nutzenbegriff angehe. Um dieses Unwissen zu überbrücken, könne man sich an einem wirksamkeitsnahen Nutzenbegriff orientieren. In diesem Sinne sei der Nutzen belegt. (<i>Verabschiedet sich.</i>)
von Maxen	Man solle sich nicht zurückziehen auf die Anforderungen der Zulassungsbehörden. Die Substanzen seien bereits lange auf dem Markt und es habe – unabhängig von den Anforderungen der Zulassungsbehörden – Gelegenheit gegeben, patientenrelevante Endpunkte zu definieren und zu untersuchen. Er sehe im Bericht immer noch einen Bruch zwischen der Studienerläuterung bzw. der Diskussion und dem Fazit. Der Nutzen werde hauptsächlich in der Verbesserung in einem kognitiven Score gesehen. Allerdings sei nach wie vor unklar geblieben, was die ggf. valide belegte Verbesserung von 3 Punkten bedeute. Dies werde auch so im Vorbericht beschrieben, allerdings werde dann im Fazit auf einmal von einem Nutzen gesprochen. Stattdessen solle dem G-BA weitergegeben werden, dass nicht klar sei, was diese Verbesserung in der Praxis bedeute.
Burns (auf Englisch)	<i>Aus seiner Sicht, auch aus der Erfahrung in der klinischen Praxis, hätten die Medikamente einen enormen Effekt. Die Evidenz für die Wirksamkeit sei überwältigend. Auch von NICE</i>

<i>Redner</i>	<i>Inhalte</i>
	<i>und Cochrane seien die Cholinesterasehemmer positiv bewertet worden. Bei NICE gehe es nun nicht um die Frage der Wirksamkeit oder Sicherheit, sondern um die Kosteneffektivität. Die Medikamente sollten allerdings nur einen Teil der Behandlung darstellen. Wichtig sei auch, dass sowohl den Patienten als auch den Betreuern zusätzliche Unterstützung geboten werde. Nichtmedikamentöse Therapieansätze seien sehr wichtig. Seine Empfehlung sei, dass die Medikamente zur Verfügung gestellt werden sollten, allerdings unter der Auflage, dass die Diagnosestellung und Therapiebegleitung von einem Facharzt durchgeführt würden.</i>
<i>Akintade (auf Englisch)</i>	<i>Von über 80 Zulassungsbehörden weltweit seien die Ergebnisse bestätigt worden. Es gebe keine Behandlungsalternativen. Nicht für alle, aber für einige Patienten bringe es einen Benefit. Verweis auf NICE und Cochrane. Empfiehlt positives Fazit.</i>
<i>Wille</i>	Der Nutzen sei aus Sicht der Patienten nicht hinreichend nachgewiesen. Es gebe Wirksamkeit auf bestimmten Instrumenten bzw. Skalen, die offenbar gesichert seien. Deren Bedeutung für den Patienten sei aber unklar. Bei einer uneingeschränkten Empfehlung werde die Frage, was dieser Effekt eigentlich für die Patienten bedeute, nie mehr gelöst. Sinnvoll seien möglicherweise bestimmte Auflagen, z. B. die Lieferung weiterer Daten im Rahmen von Versorgungsforschung oder eine Beschränkung der Gabe insofern, als dass im Sinne einer Therapiekontrolle sichergestellt werde, dass die Patienten darauf reagierten und die Präparate auch nur so lange verordnungsfähig sein sollten.
<i>Gogol</i>	Datenlage sei nicht so, dass man den Nutzen eindeutig bejahen oder eindeutig verneinen könne. IQWiG müsse den Ärzten die Möglichkeit geben, zu agieren – unter Auflagen, z. B. im Rahmen von Versorgungsforschung –, aber nicht nur, wie von NICE vorgegeben, bei einem MMSE von 10 bis 20 Punkten.
<i>Torfs</i>	Eine ähnliche Diskussion habe man vor 5 Jahren schon bei NICE geführt. Er sei sicher, dass man auch in 5 Jahren noch die Diskussion führen könne, weil auch die besten Studien den gleichen Effekt zeigen würden. Und auch wenn dann einige methodische Probleme ausgeräumt würden (ITT, LOCF etc.), dann würde man doch immer noch darüber diskutieren, was dieser Effekt bedeute. Es gebe schon viele gute Studien, die die Effekte konsistent zeigten. Daraus müsse man auch endlich die Konsequenzen ziehen.
<i>Kaiser</i>	Frage, ob damals NICE nicht gefordert habe, dass man – neben dem Effekt auf den ADAS-cog – andere Dinge untersuchen müsse.
<i>Torfs</i>	NICE habe naturalistische Daten gefordert, keine Langzeit-RCTs. Verweis auf vorhandene nicht randomisierte Evidenz aus Langzeitstudien, epidemiologischen Daten und Case-Control-Studien.
<i>Kaduszkiewicz</i>	Ihrer Ansicht nach zeigten die Studien einen Nutzen, dessen klinische Relevanz unklar sei. Darüber hinaus werde der Nutzen durch die methodischen Mängel in den Studien geschwächt. Es gebe weder einen Beweis für die Wirksamkeit noch für die Unwirksamkeit. Einschränkungen seien notwendig, zum Beispiel wie in Belgien (Register aller Patienten, die Antidementiva erhalten) oder England (Indikation zur Therapie nur gemeinsam durch Gerontopsychiater und Hausarzt sowie Vorgaben zur Kontrolle und Dokumentation der Therapie). Dadurch entstehe ein Datenschatz zur Wirkung der Medikamente in der klinischen Praxis. Wünsche sich eine Nutzenbewertung nichtmedikamentöser Verfahren durch IQWiG. Hinweis darauf, dass es neben der medikamentösen Behandlung auch noch andere Ansätze für die Patienten gebe.
<i>Hüll</i>	Es sei kritisiert worden, dass der Bericht eine Welle schlage in dem Sinne, dass man von einer geringen Wirksamkeit zu einem gewissen Nutzen komme. Er gehe davon aus, dass Wirksamkeit eine Voraussetzung für Nutzen sei. Die Frage sei, ob der Überstieg von Wirksamkeit zu Nutzen die Zusammenschau verschiedener Ebenen sei oder ein quantitatives Element beinhalte. Die schwierige Frage sei, ob das Instrument des ADAS-cog nicht in sich, weil es die Kernsymptomatik abbilde, auch einen gewissen Nutzen

<i>Redner</i>	<i>Inhalte</i>
	<p>bestimme und nicht nur Wirksamkeit.</p> <p>Wenn man sage, dass dieses Instrument etwas abbilde, was für den Patienten von Nutzen sei, dann habe man dort eine sehr geringe Wirksamkeit und einen sehr geringen Nutzen, aber diesen geringen Nutzen könne man nicht zurückverwandeln in gar keinen Nutzen. Dies sei höchstens möglich, wenn es einen Vergleich mit anderen Interventionen gebe, gegenüber denen der hier erzielte Nutzen absolut vernachlässigbar sei. Aber ansonsten sei diese „Rolle“ von einer geringen Wirksamkeit zu einem geringen Nutzen nicht anders aufzulösen.</p>
Sawicki	<p>Eine Kosten-Nutzen-Analyse könne dieses Problem auflösen. Dann fiel das Ergebnis wahrscheinlich so aus wie bei NICE, dass der geringe Nutzen die Kosten nicht wert sei. Dies sei aber nicht Bestandteil des vorliegenden Projektes und das Gesetz lasse auch Kosten-Nutzen-Analysen nicht zu. Deshalb werde hier nur der Nutzen betrachtet.</p>
Schmidtke	<p>Eine halbe Standardabweichung sei kein geringer, sondern ein moderater Nutzen. Es sei keine Frage, dass Nutzen ein statistischer Effekt auf einer Skala sei, die einen patientenrelevanten Endpunkt abbilde. Darüber hinaus könnten nur verschiedene medizinische Interventionen gegeneinander abgewogen werden.</p> <p>Zu Herrn von Maxen und Herrn Wille: Das Rekurren auf andere Messebenen wie Pflegeheimweisung führe in die falsche Richtung. Der Nutzen bemesse sich am Patienten. Das Kernsymptom der Demenz sei der Verlust der kognitiven Fähigkeiten, deswegen müsse man auch hier den Nutzen festmachen.</p>
Haupt	<p>Begriff Nutzen sei aus seiner Sicht nicht medizinisch, sondern eher gesamtgesellschaftlich definierbar. In der klinischen Arbeit sehe er nicht wenige Patienten, die den Nutzen erfahren. Allerdings solle sehr viel mehr im Bereich der nichtmedikamentösen Ansätze geforscht werden.</p>
Tracik	<p>Die Diskussion sei sehr nützlich gewesen. Bei der Behandlung der Alzheimer Demenz gehe es primär um die Effekte in der Kognition und im klinischen Gesamtbild. Der klinisch erfahrbare Nutzen bilde sich in den Studien zum großen Teil ab.</p>
Halbekath	<p>Zitiert aus Vorbericht S. 167: „... bleibt doch die Frage, was ein Response, also die Veränderung in einem abstrakten Score um einen bestimmten Betrag, nun unmittelbar und spürbar für den einzelnen Patienten bedeutet. Eine Antwort auf diese Frage steht letztlich noch aus und wird auf der Score-Ebene allein vermutlich kaum zu erzielen sein.“ Wenn man das so schreibe, wie komme man dann zu dem Fazit?</p>
Demski	<p>Analyse zeige, dass nur Surrogatparameter vorlägen. Sie könne nicht nachvollziehen, wie man auf dieser Grundlage die Medikamente empfehlen könne.</p> <p>Vorschlag an alle, ob es nicht möglich sei zu akzeptieren, dass das IQWiG diese Medikamente ablehne. Stattdessen könne Geld von den Pharmaunternehmen zusammengelegt werden, um damit Forschung zu finanzieren.</p>
Jansen	<p>Hinweis darauf, dass es keine anderen medikamentösen Therapieoptionen gebe. Es werde noch keine Kosten-Nutzen-Diskussion geführt. Solange es keine besseren Therapieoptionen gebe, solle die Verschreibung dieser Medikamente auch möglich sein. Einschränkungen gebe es ja bereits, z. B. empfehle die Arzneimittelkommission der deutschen Ärzteschaft bereits eine Therapiekontrolle. Eine derartige Kontrolle der Nebenwirkungen und Wirksamkeit erwarte sie auch vom Arzt, das sei für sie eine Selbstverständlichkeit.</p>
Sawicki	<p>Dank an alle Teilnehmer für das Kommen und die auf hohem Niveau geführte Diskussion. Wünscht allen eine gute Heimreise.</p>
<i>Ende der Sitzung</i>	

Nach Ende der Erörterung wurden keine weiteren Unterlagen durch die Teilnehmer eingereicht.

ANHANG I: STELLUNGNAHMEN

	Name	Institution	Seite
1	Renate Demski	ALZheimer-ETHik e.V.	270
2	Dr. Andreas von Maxen Dr. Heide Gieck Jutta Halbekath Wolfgang Becker-Brüser	Arznei-Telegramm	276
3	Heike von Lützu-Hohlbein Sabine Jansen	Deutsche Alzheimer Gesellschaft e.V.	284
4	Dr. Manfred Gogol Dr. Dieter Lüttje Dr. Hansjörg Werner	Deutsche Gesellschaft für Gerontologie und Geriatrie BAG Geriatrie Berufsverband Deutscher Internisten	289
5	Prof. Dr. Hans Gutzmann PD Dr. Martin Haupt	Deutsche Gesellschaft für Gerontopsychiatrie und - psychotherapie e.V.	296
6	Dr. Helga Schmitz Dr. Latif Akintade Friedhelm Leverkus (Pfizer)	Eisai GmbH	303
7	Prof. Dr. Hans-Jürgen Möller	Hirnliga e.V.	320
8	Dr. Angelika Mehnert Koen Torfs, MSc Dr. Klaus Suwelack	Janssen-Cilag GmbH	325
9	Dr. Hans Wille	Klinikum Bremen-Mitte gGmbH (Institut für Klinische Pharmakologie)	342
10	Prof. Dr. Wolfgang Maier	Kompetenznetz Demenzen	348
11	Dr. Yvonne Wirth Dr. Albrecht Stöffler	Merz Pharmaceuticals GmbH	352
12	Dr. Ferenc Tracik Dr. Monika Baier Dr. Dieter Götte	Novartis Pharma GmbH	359

	Name	Institution	Seite
13	Dr. Michael Warmbold Dr. Rainer Zerfaß Hilmar Mehlig (Eisai)	Pfizer GmbH	365
14	Dr. Dipl.-Psych. Brigitte Grass-Kapanke	privat, Mettmann	373
15	Dr. Hanna Kaduszkiewicz Dr. Dipl.-Psych. Thomas Zimmermann Prof. Dr. Hendrik van den Bussche	Universitätsklinikum Hamburg-Eppendorf (Institut für Allgemeinmedizin)	379
16	Prof. Dr. Alistair Burns	University of Manchester	394
17	Dr. Steffen Wahler Dr. Matthias Binek	Verband Forschender Arzneimittelhersteller e.V.	398

Stellungnahme ALZheimer-ETHik e. V., Demski



EINGEGANGEN: 05. Okt. 2006

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

Stellungnahme zum Vorbericht A05/19-A Cholinesterasehemmer bei Alzheimer Demenz

Name, Vorname, Titel und Funktion des/der Stellungnehmenden:

1. Demski, Renate
- 2.
- 3.

Kontakt für weitere Nachfragen:

Name: Renate Demski
Institution: Alzheimer - ETHik e.V.
Anschrift: Lappanbenedele 10 59063 Hamm
Tel./Fax-Nr.: 02381/ 510-15
E-Mail-Adresse: alzath@net.com

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme ALZheimer-ETHik e. V., Demski (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Hann, 4.10.06
Ort/Datum

Reinhold Demski
Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme ALZheimer-ETHik e. V., Demski (Fortsetzung)

Als unabhängige Selbsthilfeorganisation nimmt ALZheimer-ETHik e.V. zum Vorbericht des IQWiG A05/19-A: Cholinesterasehemmer bei Alzheimer Demenz (vom 8. 9. 2006) wie folgt Stellung:

1. Wertschätzung für alle Detail-Arbeit – Anfrage / Vorschlag

Für einen medizinisch-wissenschaftlichen Laien, wie es der Verfasser dieser Stellungnahme ist, erscheint die Arbeit sehr sorgfältig, sehr engagiert (wo Zahlen fehlten, wurde sogar geschätzt ...), allen Hinweisen der Wirksamkeit von Cholinesterasehemmern nachzugehen. Frage wurde jedoch beim Lesen des umfangreichen Vorberichts, ob ein solches Verfahren IQWiG nicht auf Dauer überfordern muss, angesichts der Fülle von Medikamenten, die gleichfalls überprüft werden müssten. Zudem: Neue wissenschaftliche Studien müssten die Arbeit an bereits bewerteten Medikamenten erneut in Gang setzen.

Vorschlag von Frau von Stösser (Am Ginsterhahn 16 53562 St. Katharinen Tel 02644/3686, Mail Stoesser-Standard@t-online.de):

ein computergestütztes System, in dem Ärzte ihre Diagnosen und Behandlungsvorschläge, auch Medikamente, die sie sowieso in den Computer eingeben, per Mouse-Klick an eine Zentraldatei weiterleiten müssen, die auch unerwünschte Ereignisse zentral registriert und umgekehrt direkt an den Arzt melden kann.

Dr. med. Volker Mohr, Bundesgeschäftsstelle Qualitätssicherung GGMBH Kanzlerstr. 4 40472 Düsseldorf Tel 0211/280729-0 hielt ein solches computergestütztes System für realisierbar. Auch der Datenschutz sei zu gewährleisten. (Diese GGMBH hat die Qualitätsberichte für Krankenhäuser bereits übernommen.)

Ein solches System wäre für sämtliche Medikamente geeignet und ständig aktuell.

2. Der Vorbericht enthält ein zu weit gespanntes Netz von Bedingungen.

a) Wir teilen die im Vorbericht durchgängig geäußerten Bedenken gegenüber verschiedenen Testverfahren, ob und inwieweit sie in der Lage sind, Demenz zu diagnostizieren. Nicht nachvollziehbar ist aus der Sicht von ALZheimer-ETHik jedoch, dass Studien, die Testergebnisse eben jener bedenklichen Testverfahren als Maßstab zu Grunde legen (feststellen, ob ein Gleichstand oder eine Veränderung der intellektuellen Fähigkeit durch eine Medikamentengabe erreicht wurde) trotzdem vom Vorbericht so ernst genommen werden, als gäbe es alle Bedenken gegen die Tests nicht mehr.

Nur solche Studienergebnisse hätten berücksichtigt werden dürfen, denen Testergebnisse aus eindeutig validen Testverfahren zu Grunde liegen.

Wahrnehmungen/ Überlegungen der Selbsthilfe:

Demenz, z.B. Alzheimer per Test diagnostizieren und den Verlauf verfolgen zu wollen, war unserer Kenntnis nach zu Beginn die Idee einer Pharmafirma. Ob und inwieweit jedoch die Bewältigung des Alltags - wobei auch Zufriedenheit und Freude eine ganz wichtige Rolle spielen - mit diesen Testergebnissen übereinstimmte, wurde nicht erfasst. Relevant für Betroffene und ihre Angehörigen sind jedoch die alltagsrelevanten Fähigkeiten und die individuelle Bewältigungsstrategien, nicht Punkte auf einer letztlich willkürlich gesetzten Skala. Alltagspraktische Fähigkeiten können von unterschiedlichen Angehörigen desselben Kranken unterschiedlich eingeschätzt werden.

Zudem ist die Verwendung dieser Tests im Alltag problematisch.

ArzthelferInnen werden angeleitet, diese durchzuführen. „Demenz-Busse“ mit Unterstützung von Pharmafirmen (Eisai-Pfizer und Janssen-Cilag) und Krankenkassen, in denen gleichfalls per Test diagnostiziert wird, sind in der Sommerzeit quer durch Deutschland unterwegs.

Selbsthilfe wird aufgefordert, diese Tests Interessierten vorzulegen.

Die Diagnose Demenz kann einen ernsthaften Einbruch in die Biografie bedeuten und sollte nur nach sorgfältiger Diagnose in einem Vertrauensverhältnis zum Arzt mitgeteilt und in ihren Konsequenzen bedacht werden.

Dadurch, dass der Vorbericht zwar Tests kritisch sieht, sie aber letztlich doch wieder mit ihren Ergebnissen akzeptiert, wird den Tests ein Gewicht gegeben, das ihm im Interesse Betroffener und potentiell Betroffener nicht zukommen dürfte.

b) Nur Studien mit weniger als 10 % Abbruchquote sollen laut Vorbericht berücksichtigt werden, auch ohne dass immer festgestellt wird (nachträglich festgestellt werden konnte) welche Gründe zu dem Abbruch der Einnahme des Medikaments geführt haben. Die Gründe

Stellungnahme von ALZheimer-ETHik e. V. zum Vorbericht des IQWiG A05/19-A: Cholinesterasehemmer bei Alzheimer Demenz]

Stellungnahme ALZheimer-ETHik e. V., Demski (Fortsetzung)

für den Abbruch können gravierend sein. Umgekehrt könnte gesehen werden, dass jeder 10. bereits in den wenigen Monaten der Einnahme Schaden davon getragen hat.

Wahrnehmung der Selbsthilfe:

Auf den ersten Blick banale unerwünschte Ereignisse wie „Durchfall“ werden in Patienteninformationen aufgeführt. Durchfall ist jedoch für Betroffene erheblich belastend, da er sozial isolieren kann, da in Folge intellektuelle, geistige Anregungen fehlen – nicht mehr wagen, das Haus zu verlassen, Verunreinigungen von Kleidung und Einrichtungsgegenständen. Zusätzliche Medikamente würden erforderlich.

Wichtiger noch:

Ernsthafte unerwünschte Ereignisse werden der Krankheit zugeordnet und nicht als Nebenwirkung des Medikaments? Unserer Wahrnehmung nach litten Patienten unter Reminyl verstärkt unter Depression. Patienten unter Aricept erlitten eine Art von epileptischen Anfällen, auch Schlaganfälle. Patienten unter Exelon (auch kombiniert mit atypischen Neuroleptika) Schwächezustände und Schlaganfälle.

BfARM forderte ALZheimer-ETHik auf, solche Fälle zu melden, bzw melden zu lassen. Eine solche Meldung kann nicht noch zur zusätzlichen Aufgabe der Selbsthilfe gemacht werden. Ärzte sind verpflichtet, Nebenwirkungen zu melden, was sie jedoch nur zu 10 % tun (laut Arzneimittelkommission).

Lösung aus der Sicht von ALZheimer-ETHik: siehe Vorschlag Punkt 1.

c) Einbeziehung von Ergebnissen, die nur dem Pharma-Unternehmen selbst zur Verfügung standen. Auch wenn diese Ergebnisse das Gesamtergebnis des Vorberichts nicht bestimmen (auch ohne sie gäbe es die beschränkte Empfehlung für Cholinesterasehemmer), so wirkt dieses Verfahren der Einbeziehung von Daten des Herstellers zumindest stützend auf das Gesamtergebnis.

Zudem:

Die Einbeziehung solcher Daten ist mehr als fragwürdig und wird im Vorbericht nicht reflektiert. Kein Hersteller wird Daten zur Verfügung stellen, die die Wirksamkeit seines Produkts fragwürdig machen? Auch wenn sie nachgereicht werden? IQWiG ist dafür angetreten, nur valide wissenschaftliche Studien zur Bewertung von Medikamenten und anderen Verfahren zu berücksichtigen – somit Maßstäbe für Übersichtsarbeiten zu setzen. Das Einbeziehen von Hersteller-Daten dürfte und sollte in zukünftigen Übersichtsarbeiten von welcher Person oder Institution auch immer nicht Praxis werden. Einer Verfälschung von Ergebnissen werden sonst Tür und Tor geöffnet.

Wahrnehmung der Selbsthilfe:

Gerade von Herstellern der Cholinesterasehemmer werden Patienteninformationen kostenlos verteilt, die nicht als objektive Information zu bezeichnen sind. Aussagen von Pharmafirmen jetzt auch noch auf wissenschaftlicher Ebene zuzulassen, stärkt das Monopol der Meinungsbildung dieser Firmen.

d) Zu kurzer Zeitraum der Studien. Überwiegend Studien von 3 Monaten bis maximal 24 Monaten Dauer der Medikamenteneinnahme werden berücksichtigt.

Wahrnehmung der Selbsthilfe:

Vor allem nach längerer Einnahme der Cholinesterasehemmer (3, 4, 5 Jahre) kommt es unserer Wahrnehmung nach verstärkt zu den schwerwiegenden unerwünschten Ereignissen, die in 2.b) dargestellt. Das computergestützte System wäre auch hier hilfreich (siehe Punkt 1). Notwendig wären entweder dieses System oder langwierige Langzeitstudien auch besonders deswegen, da mittlerweile allgemein eine möglichst frühe Diagnose der Demenz und in Folge eine möglichst frühe Medikamentengabe empfohlen werden und somit die Einnahme des Medikaments sich auch über einen jahrelangen Zeitraum erstrecken kann und damit mehr Menschen in größere Gefahr geraten, durch Nebenwirkungen zusätzlich zu erkranken.

3. Zur Einbeziehung von Experten und der Deutschen Alzheimer Gesellschaft (DALZ) (10)

Die Namen und Positionen derer, die jetzt im Nachhinein zum Vorbericht Stellung nehmen, werden bekannt, jedoch nicht diejenigen, die IQWiG von sich aus als „externe Sachverständige“ für die Erstellung des Vorberichts einbezog. Für Selbsthilfe wäre es sicherlich wünschenswert zu wissen, wer in welcher Form mitgearbeitet hat.

Der Vorbericht stellt dar, dass Angehörige der DALZ zur Definition von patientenrelevanten

Stellungnahme von ALZheimer-ETHik e. V. zum Vorbericht des IQWiG A05/19-A: Cholinesterasehemmer bei Alzheimer Demenz 2

Stellungnahme ALZheimer-ETHik e. V., Demski (Fortsetzung)

Endpunkten befragt wurden. Von Interesse wäre sicherlich zu wissen, was Angehörige nannten mit welcher Gewichtung dies von IQWiG Gewicht berücksichtigt wurde, z.B. folgende Punkte (?):

- keine zusätzlichen Beschwerden / Erkrankungen durch Cholinesterasehemmer auch nicht nach Jahren
- Zufriedenheit, Freude des Erkrankten (auf dieses Thema geht auch ein Artikel von Matthias Synofzik ein („Wirksam, indiziert – und dennoch ohne Nutzen? Die Ziele der medikamentösen Demenz-Behandlung und das Wohlergehen des Patienten“, in: Z Gerontol Geriat 4/2006, 301-307)
- Gesundheit, Zufriedenheit, Freude des /der Angehörigen in seiner/ ihrer Aufgabe der Begleitung
- keine Übersiedlung in ein Heim (oder Sonstiges)
- keine Klinikeinweisungen
- Verbesserung der kognitiven Fähigkeiten, wenn diese spürbar / erkennbar bewirkt, Notwendiges noch erledigen zu können, das intellektuelle Fähigkeiten erfordert. (Punkte in Tests nicht als Surrogatparameter)
- gesellschaftlich akzeptiert zu sein – Betroffene wie Angehörige

4. Keine genügende Berücksichtigung des Zusammenhangs von Qualität und Wirtschaftlichkeit?

IQWiG trägt in seinem Namen auch den Aspekt der Wirtschaftlichkeit, der in dem Vorbericht keine Rolle spielt?

Eindeutige Erfahrung von pflegenden Angehörigen ist, dass ihre eigene Einstellung und die Art und Weise des Umgangs mit ihrem dementen Angehörigen seine kognitiven wie auch alltagspraktischen Fähigkeiten erheblich fördern oder mindern können – dies in leichten bis hin zu schweren Stadien der Demenz.

Es gälte, Angehörige und Ehrenamtliche, denen förderndes Verhalten von sich aus gelingt, verstärkt anzuerkennen oder diese Einstellung und Art und Weise des Umgangs verstärkt zu lehren (auf Kosten von Krankenkassen bereits akzeptiert), insbesondere Alleinstehenden verstärkt „psychiatrische Begleitung“ oder „Alltagsbegleitung“ zukommen zu lassen (wie auch bereits von Ärzten verschreibbar, jedoch in der Realität wenig genutzt wird). Das Budget müsste wie die „stundenweise Verhinderungspflege“ frei genutzt werden können. Neu müsste die Leistung für psychosoziale Begleitung oder Alltagsbegleitung auch Bewohnern/ Patienten in Einrichtungen zu Gute kommen können – und sei es durch seriöse und konsequente Schulung des Personals (Festlegung von obligatorischen Lerninhalten, auch obligatorischer menschlicher und sozialer Kompetenz).

Die Gabe von Medikamenten statt der Realisierung von alternativer Begleitung beklagt auch der Artikel in British Medical Journal vom 11. September 2006;333:491-493 / doi:10.1136/bmj.38945.478160.94, Anthony J Pelosi et al. „Role of cholinesterase inhibitors in dementia needs rethinking“. Ich zitiere und übersetze frei die „Summary Points“ (S. 493, siehe Anlage): „Memory-Kliniken ziehen mit ihrer Verschreibung der umstrittenen Cholinesterasehemmern Geldmittel ab, die einer hochqualifizierten, integrierten Pflege zugute kommen sollten.“ In Deutschland stehen meiner Wahrnehmung nach nicht nur Memory-Kliniken in dieser Verantwortung, sondern auch Fach- und Hausärzte mit eigener Praxis, wie auch Beratungsstellen mit medizinischen Laien besetzt, die die Notwendigkeit dieser Medikamentengabe Ratsuchenden nahe legen. Hinzu kommen Internetseiten mit seriös klingenden Titeln, wie z.B. „altern-in-wuerde.de“, Schriften von Seniorenorganisationen, wie z.B. der Deutschen Seniorenliga, kostenlose Medizinbeilagen zu Tageszeitungen, fast wöchentlich mindestens eine Studienerfolgsmeldung über Cholinesterasehemmer in der „Ärztzeitung“ auf einer Pharma-Veranstaltung.

Die DALZ trägt als Fachgesellschaft eine besondere Verantwortung für die Empfehlung der Cholinesterasehemmer. Der Bewertung von IQWiG vorgehend, veröffentlicht sie 2006 wiederum in ihrer Broschüre (die kostenlos verteilt und versandt wird, siehe Anlage) „Das Wichtigste über die Alzheimer-Krankheit. Ein kompakter Ratgeber“ (neunte, aktualisierte Ausgabe, S. 17) „Donepezil, Galantamin und Rivastigmin (...) sind zugelassen für die Behandlung der leicht- bis mittelschweren Alzheimer-Krankheit. Sie halten über ungefähr ein

Stellungnahme von ALZheimer-Ethik e. V. zum Vorbericht des IQWiG A05/19-A: Cholinesterasehemmer bei Alzheimer Demenz 3

Stellungnahme ALZheimer-ETHik e. V., Demski (Fortsetzung)

Jahr die geistige Leistungsfähigkeit und die Alltagsbewältigung aufrecht.“

In Kontrast steht dies zudem zu Erfahrungen der DALZ in der Beratung, die mit denen von ALZheimer-ETHik übereinstimmen. In ihrer Beratung von Mensch zu Mensch macht die DALZ auch die Erfahrung, dass Patienten mit den Medikamenten nicht zurechtkommen und keine Wirksamkeit feststellen. (Frau Schneider-Schelte (DALZ) während des ev. Kirchentags in Hannover 2005 im Gespräch mit R. Demski (ALZheimer-ETHik).)

Zudem nicht nachvollziehbar ist mir die Entscheidung der DALZ, für die Wirksamkeit der Cholinesterasehemmer zu plädieren, da Pharmafirmen selbst in der Aussage über die Wirksamkeit ihrer Medikamente bereits vorsichtiger verfahren. So heißt es in der gleichfalls kostenlos verteilten und versandten Patienteninformation der „Alzheimer-Hilfe“, einer Initiative von Eisai-Pfizer (Hersteller von Aricept) („Schwierige Situationen im Alltag. Ein Leitfaden für Angehörige und Pflegekräfte“, S. 7, siehe Anlage): „Selbst wenn der Patient Medikamente erhält, die die Verschlechterung der Krankheitssymptome eine Zeit lang aufhalten können, wird sich das Krankheitsbild verschlechtern.“

Ob in einer Real- oder Wirklichkeitsform über die Wirksamkeit der Cholinesterasehemmer Aussagen getroffen werden – die Medikamente verzögern die Erkrankung oder sie können verzögern – halte ich für einen gravierenden Unterschied.

Die nicht leichte Aufgabe von IQWiG bestünde darin, der bereits seit Jahren in Gang gesetzten Flut von unzutreffenden Informationen über Cholinesterasehemmer Einhalt zu verschaffen. (?)

Folgende Summen stünden bei Wegfall der Verschreibung von Cholinesterasehemmern für psychosoziale Begleitung oder Alltagsbegleitung zur Verfügung:

<u>Aricept</u> (10 mg, 1 Tabl. täglich (4,7 €)).....	30 Tage: <u>120,1 €</u> jährlich: <u>1.440,7 €</u>
<u>Exelon</u> (4,5 oder 6 mg, 2 Tabl. täglich (1,70 €))....	30 Tage: <u>102,0 €</u> jährlich: <u>1.224,0 €</u>
<u>Reminyl</u> (24 mg, 1 Tabl. täglich (4,30 €))	30 Tage: <u>129,00 €</u> jährlich: <u>1.548,0 €</u>

Einzelne Ärzte sind dazu übergegangen, zwei der Cholinesterasehemmer gleichzeitig einem Patienten zu verschreiben oder auch in Kombination mit Memantine.

Hinzu kommen Kosten für die große Auswahl von ruhig stellenden Medikamenten, deren Wirkung, Neben- und Nachwirkungen IQWiG gesondert nach patientenrelevanten Eckpunkten bewerten müsste. (So z.B. allein mindestens 157 Medikamente aus einer der Gruppen der ruhig stellenden Medikamente, der Benzodiazepine, müsste IQWiG bewerten und nach Möglichkeit auch in Kombination mit anderen Medikamenten. Erneut erweist sich die Notwendigkeit des Vorschlags des computergestützten Systems in Punkt 1.)

Nochmals: Eindeutig effektiv für eine Verbesserung jeglicher Demenzsymptome sind die Einstellung zum Menschen mit Demenz und die Art und Weise des Umgangs ihm. Diese gilt es in jeder mit dementen Menschen befassten Personengruppe zu stärken und nicht Finanzen für zweifelhaft wirksame Medikamente bereit zu stellen.

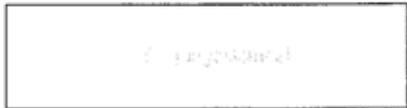
Ich hoffe, durch meine Stellungnahme, die auf verschiedenen Ebenen die Problematik der Cholinesterasehemmer angeht, Beiträge zu einem behutsamen Umgang mit dem Diagnostizieren von Demenz, zur Nicht-Verschreibung dieser Medikamente, zur Stärkung von psychosozialer Begleitung und Alltagsbegleitung von Menschen mit Demenz, zur Stärkung der Angehörigen und zum Bedenken der IQWiG auferlegten Bewertungspraxis geleistet zu haben.



Renate Demski, 4. 10. 2006

Stellungnahme von ALZheimer-ETHik e. V. zum Vorbericht des IQWiG A05/19.A: Cholinesterasehemmer bei Alzheimer Demenz 4

Stellungnahme Arznei-Telegramm, von Maxen



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

Vorbericht: Cholinesteraseinhibitoren bei Alzheimer Demenz, Auftragsnummer A05/19A

Name, Vorname; Titel und Funktion des/der Stellungnehmenden:

1. Dr. von Maxen, Andreas
wissenschaftlicher Redakteur der Arzneimittelinformation Berlin GmbH
2. Dr. Gieck, Heide
wissenschaftliche Redakteurin der Arzneimittelinformation Berlin GmbH
Halbekath, Jutta
3. wissenschaftliche Redakteurin der Arzneimittelinformation Berlin GmbH
4. Becker-Brüser, Wolfgang Herausgeber des arznei-telegramm,
Arzneimittelinformation Berlin GmbH

Kontakt für weitere Nachfragen:

Name: DR. Andreas von Maxen

Institution: Arzneimittelinformation Berlin GmbH

Anschrift: Bergstraße 38 A, 12169 Berlin

Tel./Fax-Nr.: Tel: 030 - 79490239, Fax: 030 - 79490220

E-Mail-Adresse: redaktion@arznei-telegramm.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Arzneitelegramm, von Maxen (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für **alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Böger, 10.10.2006
Ort/Datum

[Signature]
Unterschrift 1

[Signature]
Unterschrift 2

[Signature]
Unterschrift 3

[Signature]

Vielen Dank für Ihre Stellungnahme.

Stellungnahme Arzneitelegramm, von Maxen (Fortsetzung)

08/10 2006 18:23 FAX +49 30 79490219

arznei-telegramm

004

arznei-telegramm

A.T.I. Arzneimittelinformation Berlin GmbH

Bergstr. 38 A, Wasserturm, D-12169 Berlin

E-Mail: redaktion@arznei-telegramm.de
vertrieb@arznei-telegramm.de
Telefax: (0 30) 79 49 02 20 (Redaktion)
(0 30) 79 49 02 18 (Vertrieb)
<http://www.arznei-telegramm.de>
Geschäftsführer: Wolfgang BECKER-BRÜSER
Amtsgericht Berlin-Charlottenburg HRB 24207

**Schriftliche Stellungnahme zum Vorbericht:
Cholinesterasehemmer bei Alzheimer Demenz
(Auftrag A05/19-A), Version 1.0 vom 01.09.2006**

Der vorliegende Vorbericht zum Nutzen von Cholinesterasehemmern bei Alzheimer Demenz weist zahlreiche Inkonsistenzen, methodische Schwächen und Ungereimtheiten auf, die dringend einer Verbesserung bedürfen. Eine grundsätzliche Neubewertung der ganzen Substanzklasse ist unter Berücksichtigung der Mängel erforderlich.

Anmerkungen zur projektspezifischen Methodik

1. Einschlusskriterien

Unter 4.1.5. werden die Einschlusskriterien der zu bewertenden Studien beschrieben. Offenbar als Kompromiss zwischen Empfehlungen der europäischen Arzneimittelbehörde EMEA und Angaben aus einer Therapieempfehlung der Arzneimittelkommission der Deutschen Ärzteschaft wird eine Mindestbeobachtungszeit von 16 Wochen festgelegt. Diese Festlegung erscheint willkürlich und nicht begründbar. Die Europäische Arzneimittelbehörde (EMA) fordert in ihrer „Note for Guidance on Medicinal Products in the Treatment of Alzheimer's Disease“,¹ dass kurzzeitige Verbesserungen in kontrollierten Studien über mindestens sechs Monate belegt sein sollen. Studien über ein Jahr und länger seien jedoch wünschenswert, um den Erhalt des Effektes nachzuprüfen. Demgegenüber handelt es sich bei der Zeitdauer von zwölf Wochen, die einer Publikation der Arzneimittelkommission der Deutschen Ärzteschaft entnommen ist, um eine Angabe zur Therapiekontrolle einzelner Patienten unter laufender Behandlung mit Cholinesteraseinhibitoren, ist also einem völlig anderen Kontext zuzuordnen.

Vorschlag: In Anlehnung an die Empfehlungen der EMEA sollten nur Studien mit einer Mindestbeobachtungsdauer von sechs Monaten berücksichtigt werden.

2. Definition der patientenrelevanten Therapieziele

Unter 4.4.2 werden „patientenrelevante Therapieziele“ definiert, die in die Bewertung einfließen sollen. Unter den Zielgrößen werden jedoch lediglich Surrogat-Scores aufgezählt, die die patientenrelevanten Ziele repräsentieren sollen. Es fehlen Angaben und Diskussion zur Patientenrelevanz dieser Skalen. Das Ausmaß einer potenziellen Besserung, welches als relevant angesehen werden kann, wurde offensichtlich nicht prospektiv festgelegt. Dies erscheint bei Surrogatkriterien wie den hier angewandten jedoch erforderlich.

Im Gegensatz zu der Erfassung „harter“ klinischer Endpunkte wie Mortalität, Inzidenz von Herzinfarkten oder Insulten bei kardiovaskulären Erkrankungen gibt es bei der Alzheimer-Erkrankung keine idealen Messinstrumente, die patientenrelevante Veränderungen im Krank-

1

Stellungnahme Arznei-Telegramm, von Maxen (Fortsetzung)

06/10 2006 16:23 FAX +49 30 79490210

arznei-telegramm

005

heitsverlauf zuverlässig erfassen.¹ Nach Vorgaben der amerikanischen Arzneimittelbehörde FDA und der EMEA müssen Änderungen in mindestens zwei Bereichen nachgewiesen werden: zwingend in einem Test zur Kognition und zusätzlich entweder hinsichtlich des allgemeinen klinischen Eindrucks oder der alltäglichen Aktivitäten. Es gibt derzeit keinen Konsens darüber, wie groß eine Punktwertdifferenz beispielsweise im ADAS-COG sein muss, um als klinisch relevant gelten zu können. Seitens der FDA wurde 1989 eine Differenz von vier Punkten im ADAS-COG-Score mit der Abnahme kognitiver Fähigkeiten unbehandelter Erkrankter über drei bis sechs Monate gleichgesetzt.² Diese Punktzahl wird zum Teil als Grenzwert für eine klinisch relevante Verbesserung angesehen.

Die positive Bewertung der Cholinesterasehemmer im Vorbericht basiert im Wesentlichen auf einer stringenten Verbesserung der Patienten im ADAS-COG-Score um durchschnittlich etwa drei Punkte. Konkret bedeutet eine Differenz von drei Punkten im ADAS-COG beispielsweise, dass von zwölf zuvor gelesenen Wörtern, die mit zwölf unbekanntem Wörtern gemischt präsentiert werden, zwei mehr erkannt werden und von sechs Punkten zur zeitlichen Orientierung zusätzlich ein Punkt, zum Beispiel der Monat, erinnert wird.⁴ An der Relevanz dieser Verbesserung kann gezweifelt werden.

Aufgrund der erheblichen Fluktuation im Verlauf der Alzheimer-Erkrankung ist der z.B. im ADAS erfasste Unterschied zwischen den mit Cholinesteraseinhibitoren und mit Placebo behandelten geringer als intra-individuell erfasste Leistungsunterschiede zu verschiedenen Zeitpunkten. Trotz der statistisch signifikanten Unterschiede in den kognitiven Leistungen zwischen Placebo- und Verumgruppe sind Ärzte in der AD 2000 Studie nicht in der Lage, „Responder“ oder „Nicht-Responder“ zuverlässig den Behandlungsgruppen zuzuordnen (low signal [treatment effect] to noise [test-retest variability] ratio).³

Als Beleg für den klinisch relevanten Nutzen der Cholinesteraseinhibitoren werden in Studien und auch im Vorbericht Responderraten verglichen, bei denen häufig ein Cut-off bei vier bzw. sieben Punkten im ADAS-COG gewählt wird. Die Verwendung dieser Responderraten kann jedoch zu einer Überschätzung des Therapieerfolges führen, wenn sich die Ergebnisse der verschiedenen Behandlungsgruppen um den Cut-off herum scharen. Aufgrund der geringen Unterschiede der gemittelten Ergebnisse ist im vorliegenden Fall hiervon auszugehen.

Vorschlag: Die Patientenrelevanz der berücksichtigten Skalen muss begründet werden. Welche Punktwert-Unterschiede im Vergleich zur Placebogruppe als klinisch relevant anzusehen sind, muss definiert und begründet werden.

3. Übertragbarkeit der Studienergebnisse

Bei Patienten in Wirksamkeitsstudien mit Cholinesterasehemmern handelt es sich um eine hochselektive Population. Obwohl Patienten mit Alzheimer-Demenz häufig internistische Begleiterkrankungen und – abhängig vom Krankheitsstadium – in etwa 50% bis 90% klinisch relevante neuropsychiatrische Symptome aufweisen,^{6,7} sind diese – und damit ein Großteil der Erkrankten – in den ausgewerteten Studien nicht hinreichend erfasst. Mindestens 23 der insgesamt 27 im Vorbericht berücksichtigten Studien schließen Patienten mit relevanten internistischen und/oder neuropsychiatrischen Begleiterkrankungen aus.

Vorschlag: Es muss im Bericht zum Ausdruck kommen, auf welche eingeschränkten Patientenkollektive die Studienergebnisse übertragbar sind.

Stellungnahme Arznei-Telegramm, von Maxen (Fortsetzung)

06/10 2006 18:23 FAX +49 30 79490219

arznei-telegramm

006

4. Definition für „Abweichung vom ITT-Prinzip“

Im Vorbericht (4.3) wird definiert, dass eine relevante Verletzung des ITT-Prinzips konstatiert wurde, wenn die Rate der in den primären Auswertungen unberücksichtigt gebliebenen Patienten bei mindestens 10,5% oder der Unterschied in den Nichtberücksichtigungsraten zwischen den verschiedenen Behandlungsgruppen mindestens 4,5 Prozent-Punkte betrug. Beide Prozentzahlen erscheinen willkürlich gewählt und werden nicht begründet. Mit gleichem Recht hätte man eine Grenze bei 4%, 12% oder 20% ziehen können.

Vorschlag: Die Definition und Festlegung der oben genannten Grenzwerte muss begründet werden.

5. Beurteilung der Studienqualität

Die Qualität der eingeschlossenen Studien wird eingeteilt in „keine erkennbaren Mängel“, „leichte Mängel“, „grobe Mängel“ sowie „unklar“ (4.3). Warum Studien als mit leichten oder mit groben Mängeln behaftet eingestuft wurden, ist aus dem Methodenteil nicht ableitbar. Die tabellarisch dargestellte Qualitäts-einteilung erscheint willkürlich und nicht nachvollziehbar.

Eine verdeckte Zuteilung (Concealment of allocation) ist lediglich in einer von zwölf Studien mit Donepezil, drei von sechs Studien mit Galantamin und zwei von vier Studien mit Rivastigmin sicher gewährleistet. Im Vorbericht wird dies als unbedeutend beurteilt, da die Studien verblindet durchgeführt wurden (5.2.1.2; 5.2.2.2; 5.2.3.2). Dieser Argumentation können wir nicht folgen. Verzerrungen, die durch eine fehlerhafte Randomisierung entstehen, können nicht durch verblindete Durchführung der Studie ausgeglichen werden. Der Nutzen einer Behandlung in Studien mit unklarem Concealment of allocation wird um 30% überschätzt.⁸ Sind die Studien eindeutig ohne verdeckte Zuteilung durchgeführt worden, liegt die Überschätzung sogar bei 41% des Therapieerfolgs.^{8,9} Studien ohne adäquates Concealment of allocation kommen wesentlich häufiger zu signifikanten Therapieerfolgen als solche mit sauberer Randomisierung.¹⁰

Von gleicher Bedeutung wie ein ausreichendes Concealment of allocation ist unseres Erachtens eine verblindete Erfassung der Endpunkte. Jede Domäne soll nach Maßgabe der EMEA von einem verblindeten Untersucher erhoben werden.¹ Auch die amerikanische Zulassungsbehörde betont die Notwendigkeit der verblindeten Endpunkterhebung.¹¹

Im Vorbericht (5.2.1.2) werden Studien, bei denen sowohl die verdeckte Zuteilung als auch die verblindete Endpunkterhebung unklar ist, lediglich als mit „leichten Mängeln“ behaftet eingestuft, obwohl durch diese Mängel eine relevante Überschätzung des Therapieerfolgs zu befürchten ist. Bei einer nur als leicht mangelhaft bewerteten Donepezil-Studie ist über die oben genannten Mängel hinaus unklar, ob tatsächlich randomisiert wurde.

Unverständlicherweise werden ohne weiteren Kommentar auch als „grob mangelhaft“ deklarierten Studien in die im Vorbericht durchgeführten Metaanalysen eingeschlossen. Selbst wenn man den ungenügenden Kriterien der Autoren folgt, sind fünf von zwölf Donepezil-Studien, eine von sechs Galantamin-Studien und zwei von vier Rivastigmin-Studien grob mangelhaft. Allein unter den berücksichtigten Donepezil-Studien werden mehr als die Hälfte der Patienten (1.583 von 3.005, 53%) in methodisch grob mangelhafte Studien eingeschlossen. Studienmängel werden im Vorbericht zwar aufgeführt, sie wirken sich auf die Bewertung des Nutzens der Cholinesterasehemmer aber offensichtlich nicht aus.

3

Stellungnahme Arznei-Telegramm, von Maxen (Fortsetzung)

06/10 2006 18:24 FAX +49 30 79490219

arznei-telegramm

007

Es fehlt darüber hinaus eine formale Analyse auf Publikationsbias. Eine solche Analyse erscheint in dem Kontext erforderlich, da nahezu alle Studien mit Cholinesterasehemmern von den Herstellern durchgeführt wurden. Dies lässt auf die Möglichkeit unveröffentlichter Daten schließen.

Vorschlag: Die methodische Bewertung aller eingeschlossenen Studien muss vor diesem Hintergrund neu aufgerollt werden. Methodisch grob mangelhafte Studien sollten aus der Auswertung ausgeschlossen werden, da die Gefahr einer Überschätzung des Therapieerfolges zu groß und nicht korrigierbar ist. Eine formale Analyse des Publikationsbias ist durchzuführen.

6. Last observation carried forward (LOCF)

In den im Vorbericht ausgewerteten Studien fällt eine hohe Drop-out-Rate auf. Bei der Hälfte der plazebokontrollierten Studien brechen im Verumarm mindestens 22% der Patienten, im Plazeboarm mindestens 17% der Patienten die Teilnahme ab. Fehlende Nachbeobachtungsdaten werden bei den meisten Auswertungen durch die vor Studienabbruch erhobenen Werte ersetzt (last observation carried forward, LOCF). Das ist bei einer sich kontinuierlich verschlechternden chronischen Erkrankung ein problematisches Verfahren, das zur Überschätzung des Nutzens führen kann, insbesondere dann, wenn, wie in den hier betrachteten Studien, die Therapieeffekte klein sind und die Abbruchrate unter Verum insgesamt höher. Den Autoren des Vorberichtes reichen die grundsätzlichen Bedenken nicht aus: „Wenngleich diese Annahme (Verzerrung durch LOCF) plausibel erscheint, so sollte sie doch – insbesondere bei dem vergleichsweise reichen Erkenntnismaterial – durch empirische Daten gestützt werden, bevor sie zu einer Abwertung der gefundenen Ergebnisse führt (Seite 168).“

Hiermit wird kurzerhand eine inakzeptable Umkehr der Beweislast gefordert. In der Pflicht stehen vielmehr Hersteller bzw. Studienautoren, die eine weitgehend lückenlose Nachverfolgung von Studienabbrechern gewährleisten und dokumentieren müssen. Dies ist bei dem zur Rede stehenden Patientenkollektiv durchaus machbar.

Als Hinweis darauf, dass die LOCF-Methode zu keiner wesentlichen Überschätzung des Therapieeffektes beiträgt, wird die Arbeit von FARLOW et al. zitiert, in der die weiter nachbeobachteten Studienabbrecher (so genannte „retrieved drop out-population“, RDO) in Rivastigmin-Studien gesondert untersucht werden. Diese ist jedoch als Beleg aus mehreren Gründen ungeeignet:

1. Von 2.126 Patienten in Rivastigmin-Studien brechen 538 (25%) vorzeitig ab, von diesen bilden lediglich 166 (8%) die so genannte RDO-Population. Inwieweit diese Gruppe repräsentativ für alle Studienabbrecher ist, bleibt unklar und wird nicht diskutiert. Die Arbeit von FARLOW et al. zeigt jedoch, dass von 70% der Studienabbrecher keine vollständigen Verlaufsdaten erhoben wurden.
2. Gemessen wird in der Publikation der Unterschied zwischen Plazebo- und Verumgruppen in Bezug auf die Ausgangsdaten zu Beginn der Studie. Aus der AD2000-Studie ist ableitbar, dass die unter Einnahme von Cholinesteraseinhibitoren verbesserten Leistungen in kognitiven Tests nach Absetzen des Medikaments relativ rasch wieder abfallen. Interessant in diesem Zusammenhang wäre daher auch die Entwicklung der Ergebnisse vom Zeitpunkt des Abbruchs der Medikamenteneinnahme bis zum Ende der Studie.
3. Inwieweit sich die Ergebnisse der FARLOW-Studie auf Untersuchungen mit Donepezil und Galantamin übertragen lassen, bleibt unklar.

4

Stellungnahme Arznei-Telegramm, von Maxen (Fortsetzung)

06/10 2008 18:24 FAX +49 30 79490219

arznei-telegramm

008

Die im Vorbericht dargestellte Metaregression, die nachweisen soll, dass der ermittelte Therapieerfolg unabhängig von dem Unterschied der Drop-out-Rate zwischen Placebo- und Verumgruppe sei, ist nicht geeignet, den Verdacht auf Verzerrung durch LOCF-Auswertung zu entkräften, da die Berechnung auch hier auf einem Datenpool von zu einem erheblichen Teil methodisch mangelhaften Studien beruht und eine Verzerrung auch denkbar ist, wenn sich die Drop-out-Raten in Placebo- und Verumgruppen nicht wesentlich unterscheiden.

7. Interpretation in der Evidenz

Der Vorbericht nivelliert bzw. schönert in der Zusammenfassung unter 5.4 und im Fazit unter 7 die Ergebnisse zu den alltagspraktischen Fähigkeiten und widerspricht damit den vorangegangenen eigenen Evidenzbewertungen.

Zu Donepezil: Nach Aussage der Autoren (5.3.1.1) kann auf Grund unterschiedlicher, teilweise nicht validierter Skalen sowie unzureichender und möglicherweise selektiver Darstellung von Ergebnissen eine Effektstärke von Donepezil auf alltagspraktische Fähigkeiten nicht sicher abgeschätzt werden.

Zu Galantamin: Die Autoren stellen fest (5.3.2.), dass sich in drei von sechs Studien mit Galantamin in üblichen Skalen zu Aktivitäten des täglichen Lebens (DAD, ADCS-ADL) eine Überlegenheit für Verum im Vergleich zu Placebo errechnen lässt. Allerdings ist eine dieser Studien (Erkinjuntti 2002) klein und weist eine hohe Abbruchrate auf (54 % Verum, 34% Placebo, in Tabelle 11 im Vorbericht falsch angegeben mit 26% bzw. 17%), in einer weiteren Studie sind nur die zwei höheren Dosierungen Placebo überlegen. In drei Studien ist Galantamin nicht signifikant überlegen.

Zu Rivastigmin (5.3.3.): Für Rivastigmin wird konstatiert, dass sich für alltagspraktische Fähigkeiten eine signifikante Effektstärke in der Größenordnung einer Viertel Standardabweichung errechnen lässt. Dieser Effekt ist nicht als moderat – wie im Vorbericht angegeben – sondern gemäß üblicher Interpretation als klein einzustufen.¹³ Zudem wird diese Effektstärke ohne Berücksichtigung der unveröffentlichten Negativ-Studie B351 errechnet.

In der Zusammenfassung unter 5.4 und im Fazit unter 7 wird dann jedoch für alle drei Substanzen ohne Erwähnung der oben ausgeführten Einschränkungen konstatiert, dass „Hinweise auf eine günstige Beeinflussung“ vorliegen.

Vorschlag: Der Nutzen in Bezug auf alltagspraktische Fähigkeiten kann nach den im Vorbericht bereits aufgeführten erheblichen Einschränkungen nicht als gesichert gelten. Die „Hinweise“ reichen hierzu nicht aus. Dies muss in der Gesamtbewertung zum Ausdruck kommen. Es muss zudem klar gestellt werden, dass nach wie vor eine Nachweispflicht besteht, in diesem Bereich einen Nutzen zu belegen.

- 1 EMEA: Note for Guidance on Medicinal Products in the Treatment of Alzheimer's Disease; zu finden unter: <http://www.emea.eu.int>
- 2 LEBER, P. (FDA): Guidelines for the clinical evaluation of antedementia drugs, 1st draft. Rockville: US Food and Drug Administration 1990
- 3 BECKER, R.E.: Am. J. Alzheimer Dis. Other Demen. 2006; 21: 81-4
- 4 ROSSEN, W.G. et al.: Am. J. Psychiatry 1984; 141: 1856-64
- 5 ROCKWOOD, K.: J. Neurol. Neurosurg. Psychiatry 2004; 75: 677-85
- 6 LYSKETSOS, C.G. et al.: JAMA 2002; 288: 1475-83
- 7 SCHMITT, F.A., WICHEMS, C.H.: Prim. Care Compagnie J. Clin. Psychiatry 2006; 8: 158-69

5

Stellungnahme Arznei-Telegramm, von Maxen (Fortsetzung)

06/10 2006 18:24 FAX +49 30 79490219

arznei-telegramm

009

- 8 SCHULZ, K.F. et al.: JAMA 1995; 273: 408-12
- 9 JÜNI, P. et al.: BMJ 2001; 323: 42-6
- 10 HEWITT, C. et al.: BMJ 2005; 330: 1057-8
- 11 FDA: N. Engl. J. Med. 1991; 324: 349-52
- 12 ROCKWOOD, K., MacKNIGHT, C.: Neuroepidemiology 2001; 20: 51-6

6

Stellungnahme Deutsche Alzheimer Gesellschaft e. V., von Lützu-Hohlbein

Von : Deutsche Alzheimer Ges.

++49 30 259379529

06 Okt. 2006 15:50

S2



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:
Cholinesterasehemmer bei Alzheimer Demenz
Auftrag A05 19-A

Name, Vorname, Titel und Funktion des/der Stellungnehmenden:

1. Heike von Lützu-Hohlbein, Vorsitzende
2. Sabine Jansen, Geschäftsführerin
- 3.

Kontakt für weitere Nachfragen:

Name: Sabine Jansen
Institution: Deutsche Alzheimer Gesellschaft e.V. Selbsthilfe Demenz
Anschrift: Friedrichstr. 236, 10969 Berlin
Tel./Fax-Nr.: 030-2593795-11 Fax -29
E-Mail-Adresse: sabine.jansen@deutsche-alzheimer.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

1

Von : Deutsche Alzheimer Ges.

++49 30 259379529

06 Okt. 2006 15:48 51

 Deutsche Alzheimer Gesellschaft e.V. 10960 Berlin

Institut für Qualität und Wirtschaftlichkeit im
Gesundheitswesen



Deutsche
Alzheimer
Gesellschaft e.V.
Selbsthilfe Demenz

Berlin, den 6. Oktober 2006

**Stellungnahme der Deutschen Alzheimer Gesellschaft
zum Vorbericht des IQWiG, Auftrags A05/19-A**

Die Deutsche Alzheimer Gesellschaft begrüßt, dass das IQWiG in seinem Vorbericht den Nutzen der Acetylcholinesterasehemmer grundsätzlich bestätigt hat. Dies stimmt überein mit den Aussagen von Angehörigen, die nicht immer, aber doch in vielen Fällen durch diese Medikamente eine Verbesserung des Zustandes ihrer Kranken beobachten. Außerdem ist die Deutsche Alzheimer Gesellschaft erleichtert, dass damit die Leitlinien deutscher und internationaler Fachgesellschaften bestätigt werden. Für medizinische Laien ist es nicht nachzuvollziehen, wenn in der Presse Dispute von Wissenschaftlern zu Lasten der Kranken ausgetragen werden, zumal wenn Medikamente schon seit Jahren zugelassen sind.

Im Folgenden möchten wir einige inhaltliche Anmerkungen zu dem Bericht vornehmen:

1. Zielgrößen (patientenrelevante Endpunkte)

Eine wissenschaftliche fundierte Definition der Indikationsbezogenen patientenrelevanten Endpunkte steht aus. Bei der Auflistung der einbezogenen Endpunkte stellt sich die Frage, ob diese überhaupt genau gegeneinander abgegrenzt werden können. Sind nicht Effekte auf die Kognition auch relevant für die Aktivitäten des alltäglichen Lebens? Oder hängt nicht die Fähigkeit, die Aktivitäten des täglichen Lebens zu bewältigen, eng mit anderen mit der Krankheit verbundenen Symptomen, wie z.B. Agitiertheit, Depression usw. zusammen? In dem im April 2005 stattgefundenen Gespräch mit Vertretern des IQWiG haben wir z.B. besonders den Erhalt der Selbst-

Anschrift:

Deutsche Alzheimer Gesellschaft e.V.
Friedrichstraße 236
10960 Berlin
Tel.: 030/259 37 95-0
Fax: 030/259 37 95-29
Alzheimer-Telefon: 01803/17 10 17
www.deutsche-alzheimer.de
info@deutsche-alzheimer.de

Spendenkonto:

Bank für Sozialwirtschaft Berlin
BLZ 100 305 00
Kont. 337 78 00

Vorstand:

1. Vorsitzende:
Dipl.-Inf. Helke von Lützu-Hohlbein,
München

2. Vorsitzende:
Pia Maria Ulrike Kriebel, Darmstadt

Beisitzer:
Ingrid Fuhrmann, Berlin
Prof. Dr. med. Alexander Kurz, München
Heldegarde Harwit, Gera
Birbel Schöner, Bochum
Karin Sabker, Schölkern

Mitgliedschaften:

Alzheimer Europe
Alzheimer's Disease International
Bundesarbeitsgemeinschaft
Selbsthilfe e.V.

Von : Deutsche Alzheimer Ges.

++49 30 259379529

06 Okt. 2006 15:48 S2

ständigkeit betont. Diese ist sowohl durch den Erhalt der kognitiven Fähigkeiten als auch durch die Fähigkeit, noch Tätigkeiten des alltäglichen Lebens bewältigen können, gekennzeichnet.



Wir bedauern, dass wichtige Zielparameter durch das Amendment 1 des Berichtsplans im Juni 2006 in ihrer Bedeutung abgewertet wurden, nämlich die angehörigerelevanten Therapieziele „Besserung bzw. der Erhalt der Lebensqualität der Angehörigen“ und die „Reduktion der Höhe des Betreuungsaufwandes“ sowie der durch den behandelnden Arzt zu erhebende klinische Eindruck.

Aus dem Fehlen von Daten zu den definierten Endpunkten bzw. von validierten Messinstrumenten, z.B. zur Lebensqualität, kann nicht geschlossen werden, dass eine Wirksamkeit nicht vorliegt. Anstatt mit dem Satz „Nutzen nicht belegt“ zu schließen, wäre zu fordern, dass für die Zukunft Studien aufgesetzt würden, um diese Endpunkte, die von der EMEAS nicht für die Zulassung verlangt werden, dennoch auf ihre Wirksamkeit zu prüfen. Es darf aus unserer Sicht nicht sein, dass fehlende Daten, die nie eingefordert wurden, im Nachhinein zu einer Negativ-Bewertung führen.

2. Nutzen

Zur Nutzenbewertung allgemein stellt sich die grundsätzliche Frage, wer diesen festlegt. Leider lässt sich aus dem Methodenpapier des IQWiG eine wissenschaftlich fundierte Definition nicht herauslesen. Nur um die Wirksamkeit kann es nicht gehen. Diese wurde bereits im Zulassungsverfahren bewertet. Wenn ein Nutzen beschrieben wird, sollten dieser insbesondere aus Sicht der Patienten und Angehörigen formuliert werden.

3 Interaktion unterschiedlicher Strategien

Der Vorbericht stellt richtig fest, dass bei der Alzheimer-Krankheit verschiedene Interventionsbereiche (biologische, psychologische und soziale Eben) in hohem Maße interagieren. Alle Bereiche haben aus Sicht der Kranken und Angehörigen ihre Wertigkeit. Das gilt in gleichem Maße für die Pflege.

Keine der Strategien sollte zugunsten einer anderen abgewertet werden. In Hinblick auf den einzelnen Menschen sollten mit seinem sozialen Umfeld geeignete Interventionen gewählt werden können. Die medizinische Behandlung sollte nicht in Gegensatz zur Pflege gesetzt werden. Pflege im klassischen Sinn ist besonders im mittleren bzw. späteren Stadium wichtig. Deshalb kann es keine Option sein, Geld für die medizinische Behandlung, die besonders im frühen Stadium vorhandene Fähigkeiten erhalten will, zu sparen mit dem Ziel, damit mehr Personal in Pflegeheimen zu schaffen.

Stellungnahme Deutsche Alzheimer Gesellschaft e. V., Lützu-Hohlbein (Fortsetzung)

Von : Deutsche Alzheimer Ges.

++49 30 259379529

06 Okt. 2006 15:58 S1

4. Einbeziehung von Patienten

Im Vorbericht wird darauf hingewiesen, dass die Sicht von Angehörigen bei der Definition der patientenrelevanten Endpunkte durch die Konsultation von Vertretern der Deutschen Alzheimer Gesellschaft erfolgte (Seite 10). Der einzige Kontakt zur Deutschen Alzheimer Gesellschaft war ein einmaliges etwa zweistündiges Gespräch am 5. April 2005, d.h. direkt nach Erhalt des Auftrages. Weitere Kontakte fanden nicht statt. Dies ist aus Sicht der Deutschen Alzheimer Gesellschaft nicht als Konsultation zu sehen.



5. Zukünftige Studien

Zu begrüßen ist die Feststellung des Vorberichts, dass es für die Zukunft wünschenswert wäre, gezielter solche Patienten zu identifizieren, die einen größeren Nutzen erfahren, und solche, die von einer Behandlung nicht profitieren (S. 170).

6. Fazit

Aus der vom IQWiG im Vorbericht bestätigten Wirksamkeit muss eine Nutzen-diskussion allein aus Sicht der Patienten und Angehörigen erfolgen. Für Patienten und Angehörige ist auch eine Verhinderung der Verschlechterung ihrer Demenzerkrankung wichtig. Auch wenn nicht alle Kranken von den Medikamenten profitieren, darf aus Sicht der Deutschen Alzheimer Gesellschaft keine negative Empfehlung ausgesprochen werden. Alle Kranken sollten – genau wie bei anderen Indikationen – die Chance haben, die Wirksamkeit einer Behandlung zu erproben. Dass nach einer angemessenen Frist eine Abwägung für eine Weiterverschreibung erfolgt, halten wir für selbstverständlich.

Für die Deutsche Alzheimer Gesellschaft

Heike von Lützu-Hohlbein
Vorsitzende

Sabine Jansen
Geschäftsführerin

Stellungnahme DGGG/BAG/BDI, Gogol

06/10/2006 14:05 +49-5156-782287

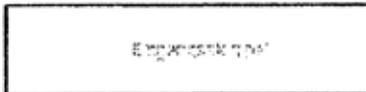
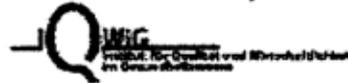
KH LINDENBRUNN INNER

S. 01/11

Dr. Gogol per Fax zur Freiblegung

Fax 0221-35685-851

- 11 Seiten -



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Antrag des IQWiG/Projektbeschreibung mit Auftragsnummer:

A05-19A „Cholinesterase-Hemmer bei Alzheimer Demenz“

Name, Vorname, Titel und Funktion des/der Stellungnahmebers:er:

1. Gogol, Manfred – Sektion II Deutsche Gesellschaft für Gerontologie und Geriatrie (DGGG)
2. Lüttje, Dieter – Bundesarbeitsgemeinschaft Klinisch-Geriatriischer Einrichtungen (BAG Geriatrie)
3. Werner, Hansjörg – Sektion Geriatrie im Berufsverband Deutscher Internisten (BDI)

Kontakt für weitere Fachfragen:

Name: Dr. med. Manfred Gogol
Institution: Klinik für Geriatrie, Krankenhaus Lindenzentrum
Anschrift: Lindenbrunn 1, 31863 Coppenbrügge
Tele-/Fax-Nr.: Tel. 05156-782-295, Fax. 05156-782-287
E-Mail-Adresse: gogol@krankenhaus-lindenbrunn.de

Bitte entnehmen Sie dem IQWiG-Formblatt die folgende:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

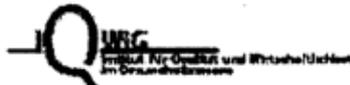
1

Stellungnahme DGGG/BAG/BDI, Gogol (Fortsetzung)

05/10/2006 14:05 +49-5156-782287

KH LINDENBRUNN INNER

S. 02/11



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal 5 DIN-A4-Seiten.

Ja

Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal 10 DIN-A4-Seiten.

Ja

Für alle Zitate sind Originalpublikationen (Volltext) beigelegt.

Entfällt

Die Stellungnahme ist von allen Stellungnehmenden unterschrieben.

z.T.*

Das Formblatt zur Darlegung potenzieller Interessenskonflikte liegt für alle Stellungnehmenden einzeln ausgefüllt vor.

z.T.*

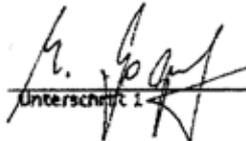
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

* Rest folgt

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Coppenbrügge, 06.10.2006

Ort/Datum


Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

2

06/10/2006 14:05 +49-5156-782287

KH LINDENBRUNN INNER

S. 87/11

Stellungnahme zum Vorbericht des IQWiG „Cholinesterasehemmer bei Alzheimer Demenz“

In Übereinstimmung mit den Ergebnissen von NICE (1-3) und Ihnen (4) ist u.E. nach festzustellen, dass Medikamente der Gruppe der Acetylcholinesterase-Hemmer klinisch oftmals hilfreich bei Patienten mit Morbus Alzheimer sind. Dies findet sich gleichfalls in Übereinstimmung mit unabhängigen Untersuchern der Cochrane Library (5-8). Insbesondere bei leichter bis mittelschwerer Alzheimer-Demenz ist die Evidenz hinsichtlich der Kognition eindeutig. Bezogen auf Alltagskompetenz und begleitende Psychopathologie gibt es insgesamt deutliche Hinweise auf eine – in der Tendenz – positive Beeinflussung.

Es muss allerdings deutlich darauf hingewiesen werden, dass hier weitere zahlreiche Einflussfaktoren seitens des Settings einer an Demenz erkrankten Person wirken bzw. wirken können, die nur schwer – auch in randomisierten Studien – abgeschätzt werden können. Bedenklich stimmt aber unserer Ansicht nach die Tendenz, die zu Grunde liegende zeitliche Limitierung der Studien – überwiegend 6 Monate – ggf. dafür heranzuziehen, den Einsatz dieser Medikamentengruppe generell auf diesen Zeitraum und / oder die oben genannte Schwere der Alzheimer-Demenz zu beschränken. Dies berücksichtigt noch nicht neuere Daten, die allerdings aus kleineren Untersuchungskollektiven resultieren. Gleichfalls ist das Problem der Beendigung der Medikation auf Grund des Krankheitsprogresses – und einer damit verbundenen überproportionalen Verschlechterung infolge der Therapieterminierung – noch ungenügend erforscht. Vielmehr kommt es auch in Zukunft darauf an bei ausgewählten Indikationen – einschließlich Mischformen von Demenzen – und Verläufen einen kontrollierten Behandlungsversuch beginnen bzw. fortführen zu dürfen einschließlich der Möglichkeit der Heranziehung weiterer Substanzgruppen (7, 9-19). Diese Patientenkollektive sind allerdings in Zukunft noch näher zu beschreiben. Vor allem bei Vorliegen eines Responder-Status ist die Indikation für eine weitere Behandlung wie bei anderen chronischen Erkrankungen generell zu bejahen. Vergleichbar dem progredienten Verlauf einer chronischen Herzinsuffizienz findet sich hier in den Leitlinien gleichfalls keine fachlich begründete Indikation für einen vollständigen oder partiellen Behandlungsabbruch auf Grund des Fortschreitens der Erkrankung (20). Die sorgfältige Diagnostik und Behandlung jeder Begleitmorbidität einschließlich der Frage der Früh- und weiterführenden Rehabilitation – z.B. nach Frakturen oder Schlaganfall – nach den Richtlinien evidenzbasierter Medizin muss Bestandteil der Behandlung Demenzkranker sein.

Unseres Erachtens ist es von essentieller Bedeutung die medikamentöse Therapie der Alzheimer-Demenz im Rahmen eines multidimensionalen und multiprofessionellen Behandlungsverfahrens einzusetzen, der insbesondere folgende Komponenten beinhalten sollte:

1. die antidepressive Therapie bei Vorliegen einer Depression (21,22),
2. die Therapie kardiovaskulärer Risikofaktoren, insbes. Herzrhythmusstörungen, arterielle Hypertonie, Herzinsuffizienz, Diabetes mellitus und evtl. Hyperlipidämie (20,23),
3. die präventive Behandlung von altersassoziierten Begleitmorbidität einschl. Impfungen u.a.,
4. die Unterstützung der Bezugspersonen (24-26),
5. Interventionen wie körperliches Training (25,27-30), Behandlung von Unter- (31-33) und Überernährung (23,35-38), Identifikation und Management des Frailty-Syndroms (39-41), soziale Aktivitäten (42-44), Alkoholkarenz (45,46), Schlafhygiene (47) u.a.,
6. der Einsatz aktivierender und mobilisierender Maßnahmen (v.a. Physio- (48) und Ergotherapie (49,50) sowie (Neuro)Psychologie (51)) und
7. weitere nicht-pharmakologische Maßnahmen, wie sie insbesondere in der stationären Altenpflege – z.B. die Validation – angewendet werden. Nur im Rahmen eines solchen

Stellungnahme DGGG/BAG/BDI, Gogol (Fortsetzung)

06/10/2006 14:05 +49-5156-782287

KH LINDENBRUNN INNER

S. 08/11

multimodalen Behandlungsansatzes kann auch in progredienten und schweren Verläufen die Sinnhaftigkeit einer medikamentösen Intervention mit Acetylcholinesterase-Hemmern optimiert und ggf. erfolgreich nachgewiesen werden.

Schlussendlich sei auf zwei weitere Problemfelder verwiesen: Zum einem das Delir mit dem Risiko der nachfolgenden Demenz (52,53), welches in Zukunft eine relevante Indikation für den Einsatz von Acetylcholinesterase-Hemmern darstellen könnte und zum anderen die Bereitschaft von Ärzten, allein auf Grund des Alters eines Patienten oder seines aktuellen Zustandes eine sinnvolle Behandlungsoption zu verweigern (54-56). Die große Zahl gesund alt werdender Menschen rechtfertigt einen Leistungsausschluss altersbedingt nicht (57,58).

Literatur:

1. NICE – Guidance on the use of Donepezil, Rivastigmine and Galantamine for the treatment of Alzheimer's Disease. 2001. <http://www.nice.org.uk/page.aspx?o=TA019guidance> (Zugriff 06.10.2006).
2. NICE – Clinical and costeffectiveness of Donepezil, Rivastigmine and Galantamine for Alzheimer's Disease. 2002. <http://www.nice.org.uk/page.aspx?o=14427> (Zugriff 06.10.2006).
3. NICE – Donepezil, Rivastigmine and Galantamine (review) and Memantine for the treatment of Alzheimer's Disease. 2006. <http://www.nice.org.uk/page.aspx?o=322952> (Zugriff 06.10.2006).
4. IQWiG – Cholinesterasehemmer bei Alzheimer Demenz. Vorbericht – vorläufige Nutzenbewertung – Version 1.0. Stand 01.09.2006. <http://www.iqwig.de/index.403.html> (Zugriff 06.10.2006).
5. Birks J. Cholinesterase inhibitors for Alzheimer's disease. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD0005593.
6. Birks J, Harvey RJ. Donepezil for dementia due to Alzheimer's disease. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD001190.
7. Loy C, Schneider L. Galantamine for Alzheimer's disease and mild cognitive impairment. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD001747.
8. Birks J, Grimley Evans J, Iakovidou V, Tsolaki M. Rivastigmine for Alzheimer's disease. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD001191.
9. McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD003154.
10. Craig D, Birks J. Rivastigmine for vascular cognitive impairment. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD004744.
11. Craig D, Birks J. Galantamine for vascular cognitive impairment. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD004746.
12. Malouf R, Birks J. Donepezil for vascular impairment. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD004395.
13. Wild R, Pettit T, Burns A. Cholinesterase inhibitors for dementia with Lewy bodies. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD003672.
14. Maidment I, Fox C, Boustani M. Cholinesterase inhibitors for Parkinson's disease dementia. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD 004747.
15. Birks J, Grimley Evans J. Ginkgo biloba for cognitive impairment and dementia. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD003120.
16. Gauthier S et al. Mild cognitive impairment. *Lancet* 2006;367:1262-70.
17. Birks J, Flicker L. Donepezil for mild cognitive impairment. The Cochrane Database of Systematic Reviews 2006 Issue 3. CD006104.
18. Petersen RC et al. Vitamin E and Donepezil for the treatment of mild cognitive impairment. *N Engl J Med* 2006;352:2379-88.

Stellungnahme DGGG/BAG/BDI, Gogol (Fortsetzung)

06/10/2006 14:05 +49-5156-782287

KH LINDENBRUNN INNER

S. 89/11

19. Winblad B et al. Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled Study. *Lancet* 2006;367:1057-65.
20. Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the the diagnosis and treatment of Chronic Heart Failure of the European Society of Cardiology. *Eur Heart J* 2005;26:1115-1140.
21. Van Gelder BM et al. – Marital status and living situation during a 5-year period are associated with a subsequent 10-year cognitive decline in older men: The FINE study. *J Gerontol* 2006;61B:P213-9.
22. Gatz JL et al. – Do depressive symptoms predict Alzheimer's Disease and dementia? *J Gerontol* 2005;60A:744-7.
23. Blennow K et al. Alzheimer's disease. *Lancet* 2006;368:387-403.
24. Emanuel EJ et al. Understanding economic and other burdens of terminal illness: the experience of patients and their caregivers. *Ann Intern Med* 2000;132:451-9.
25. Tori L et al. – Exercise plus behavioral management in patients with alzheimer disease. A randomized controlled trial. *JAMA* 2003;290:2015-22.
26. Amirkhanyan AA et al. Parent care and the stress process : findings from panel data. *J Gerontol* 2006;61B:S248-55.
27. Larson EB et al. Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Ann Intern Med* 2006;144:73-81.
28. Abbott RD et al. – Walking and dementia in physically capable elderly men. *JAMA* 2004;292:1447-53.
29. Wang L et al. – Performance-based physical function and future dementia in older people. *Arch Intern Med* 2006;166:1115-20.
30. Prohaska T et al. Physical activity, public health, and aging : critical issues and research priorities. *J Gerontol* 2006;61B:S267-73.
31. Karakelides H et al. Sarcopenia of aging and its metabolic impact. *Curr Top Dev Biol* 2005;68:123-48.
32. Steen B. Maximizing outcome of dementia treatment: the role of nutrition. *Arch Gerontol Geriatr Suppl* 2004;9:413-7.
33. Martin CT et al. Nutritional risk and low weight in community-living older adults : a review of the literature (1995-2005). *J Gerontol* 2006;61A:927-34.
34. Mamhidir AG et al. Underweight, weight loss and related risk factors among older adults in sheltered housing – a Swedish follow-up study. *J Nutr Health Aging* 2006;10:255-62.
35. Yaffe K et al. – The metabolic syndrome, inflammation and risk of cognitive decline. *JAMA* 2004;292:2237-42.
36. Gustafson D et al. – An 18-year follow-up of overweight and risk of alzheimer disease. *Arch Intern Med* 2003;163:1524-8.
37. Whitmer RA et al. – Obesity in middle age and future risk of dementia: a 27 year longitudinal population based Study. *BMJ* 2005;330:1360-4.
38. Scarmeas N et al. – Mediterranean diet and risk for Alzheimer's Disease. *Ann Neurol* 2006;59:912-21.
39. Paganelli R et al. Frailty of older age: the role of the endocrine-immune interaction. *Curr Pharm Des* 2006;12:3147-59.
40. Walston J et al. Research agenda for frailty in older adults : towards a better understanding of physiology and etiology: summary from the American Geriatrics Society / National Institute of Aging research conference on frailty in older adults. *J Am Geriatr Soc* 2006;54:991-1001.
41. von Kanel R et al. Effect of Alzheimer caregiving stress and age on frailty markers interleukin-6, C-reactive protein, and D-dimer. *J Gerontol* 2006;61A:963-9.
42. Verghese J et al. – Leisure activities and the risk of dementia in the elderly. *N Engl J Med* 2003;348:2508-16.

Stellungnahme DGGG/BAG/BDI, Gogol (Fortsetzung)

06/10/2006 14:05 +49-5156-782287

KH LINDENBRUNN INNER

S. 10/11

43. Coyle JT – Use it or lose it – do effortful mental activities protect against dementia? *N Engl J Med* 2003;348:2489-90.
44. Ghisletta P et al. Does activity engagement protect against cognitive decline in old age? Methodological and analytical considerations. *J Gerontol* 2006;61B:P253-61.
45. Mukamal KJ et al. – Prospective study of alcohol consumption and risk of dementia in older adults. *JAMA* 2003;289:1405-13.
46. Anttila T et al. Alcohol drinking in middle age and subsequent risk of mild cognitive impairment and dementia in old age: a prospective population based study. *BMJ* 2004;329:539-44.
47. Fetveit A et al. Sleep duration during the 24-hour day is associated with the severity of dementia in nursing home patients. *Int J Geriatr Psychiatry* 2006;21:945-50.
48. Koczy P et al. [Effectiveness of a multifactorial intervention to reduce physical restraints in nursing home residents with dementia]. *Z Gerontol Geriatr* 2005;38:33-9.
49. Gitlin LN et al. Maintenance of effects of the home environmental skill-building program for family caregivers and individuals with Alzheimer's disease and related disorders. *J Gerontol* 2005;60A:368-74.
50. Voigt-Radloff S et al. [Controlled trials on the efficacy of occupational therapy with elderly. Part II: Evidence for prioritized diseases and disabilities]. *Z Gerontol Geriatr* 2004;37:450-8.
51. Sitzer DI et al. Cognitive training in Alzheimer's disease: a meta-analysis of the literature. *Acta Psychiatr Scand* 2006;114:75-90.
52. Inouye SK. Delirium in older persons. *N Engl J Med* 2006;354:1157-65.
53. Lundström M et al. Dementia after delirium in patients with femoral neck fractures. *J Am Geriatr Soc* 2003;51:1002-6.
54. Hamel MB et al. Patient age and decisions to withhold life-sustaining treatments from serious ill, hospitalized adults. *Ann Intern Med* 1999;130:116-25.
55. Farber NJ et al. Physicians' decisions to withhold and withdraw life-sustaining treatment. *Arch Intern Med* 2006;166:560-4.
56. Raymond V et al. Prevalence of mental incapacity in medical inpatients and associated risk factors: cross-sectional study. *Lancet* 2004;364:1421-7.
57. Schoenhofen EA et al. Characteristics of 32 supercentenarians. *J Am Geriatr Soc* 2006;54:1237-40.
58. Calvert JF et al. Dementia-free survival among centenarians : an evidence-based review. *J Gerontol* 2006;61A:951-6.

Dr. med. Manfred Gogol (korrespondierender Autor)
Chefarzt
Klinik für Geriatrie
Krankenhaus Lindenbrunn
Lindenbrunn 1
31863 Coppenbrügge
Tel. 05156-782-295
Fax. 05156-782-287
Email: gogol@krankenhaus-lindenbrunn.de

Vorsitzender Sektion II
Deutsche Gesellschaft für Gerontologie und Geriatrie (DGGG)

Dr. med. Dieter Lütjje

Stellungnahme DGGG/BAG/BDI, Gogol (Fortsetzung)

05/18/2006 14:05 +49-5156-782267

KH LINDENBRUNN INNER

S. 11/11

Chefarzt
Klinik für Geriatrie
Klinikum Osnabrück
Sedanstr. 115
49090 Osnabrück

Vorsitzender Bundesarbeitsgemeinschaft Klinisch-Geriatricher Einrichtungen (BAG
Geriatrie)

Dr. med. Hansjörg Werner
Chefarzt
Klinik für Geriatrie
Evangelisches Krankenhaus Elisabethenstift
Landgraf-Georg-Str. 100
64287 Darmstadt

Vorsitzender Sektion Geriatrie im Berufsverband Deutscher Internisten (BDI)

Interessenkonflikte:

M. Gogol: keine

D. Lütjke: Honorare für Vortragstätigkeit Fa. Eisai und Novartis

H. Werner: Honorare für Vortragstätigkeit Fa. Eisai und Janssen-Cilag



Eingangstempel

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung und des Auftraggebers

Auftrag A05/19-A: Cholinesterasehemmer bei Alzheimerdemenz vom 07.09.06

Namen, Vorkämpfer und Funktionen der Stellungnahme

1. Gutzmann, Hans (Prof. Dr. med.) - Präsident der Deutschen Gesellschaft für Gerontopsychiatrie und -psychotherapie e.V. (DGGPP e.V.)
2. Haupt, Martin (PD Dr. med.) - Vizepräsident der Deutschen Gesellschaft für Gerontopsychiatrie und -psychotherapie e.V. (DGGPP e.V.)
- 3.

Kontakt

Name: Prof. Dr. med. Hans Gutzmann
Institution: Krankenhaus Hedwigshöhe, Klinik für Psychiatrie u. Psychotherapie
Anschrift: Höhensteig 1, 12526 Berlin
Tel./Fax-Nr.: 030/6747-3007 ; -3002
E-Mail-Adresse: h.gutzmann@alexius.de

Bittens / Aufforderungen zur Stellungnahme (VGL 10)

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme DGGPP e. V., Gutzmann (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für alle Stellungnehmenden einzeln **ausgefüllt** vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Berlin, d. 09.10.06
Ort/Datum

Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

2

Deutsche Gesellschaft für Gerontopsychiatrie und -psychotherapie e.V.

DGGPP, Prof. Dr. H. Gutzmann, KH Hedwigshöhe, Buntzelstr. 36, 12526 Berlin

EINGEGANGEN 06. Okt. 2006



Stellungnahme der Deutschen Gesellschaft für Gerontopsychiatrie und -psychotherapie e.V. zum Vorbericht (vorläufige Nutzenbewertung, Auftrag A05/19-A) Cholinesterasehemmer bei Alzheimerdemenz vom 01.09.2006

Mit dem 1. September 2006 publizierte das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen nach Beauftragung durch den Gemeinsamen Bundesausschuss im Februar 2005 den Vorbericht zu den Cholinesterase-Hemmern (ChE-H) Donepezil, Galantamin und Rivastigmin im Internet.

Im wesentlichen bestätigt das Institut einen Nutzen der Substanzen in ihren Indikationsbereichen der leichten und mittelschweren Alzheimer Demenz.

Zum Indikationsbereich von Rivastigmin bei der Demenz bei Parkinson Krankheit und der Wirksamkeit der Cholinesterasehemmer bei der Demenz mit Lewy-Körperchen lag keine Beauftragung des Instituts vor.

Der mehr als 200 Seiten umfassende Vorbericht ist gründlich recherchiert, sorgfältig ausgearbeitet und in den Schlussfolgerungen mehrheitlich nachvollziehbar. Es ergeben sich aber bei näherer Prüfung einige wichtige Anmerkungen und Korrekturerfordernisse.

Nach einer Darstellung der Alzheimer Krankheit, ihrer Epidemiologie, Diagnose und therapeutischen Möglichkeiten werden die Kernziele der Untersuchung benannt. Es handele sich um die Nutzenbewertung einer langfristigen Behandlung mit einem ChE-H bei der Alzheimer Demenz im Vergleich zu einer Placebo-Gabe hinsichtlich patientenrelevanter Therapieziele und um die Nutzenbewertung einer langfristigen Behandlung mit einem ChE-H bei der Alzheimer Demenz im Vergleich zu einer Behandlung mit einem anderen

Präsident

Prof. Dr. H. Gutzmann
Krankenhaus Hedwigshöhe
Buntzelstr. 36
12526 Berlin
Tel. 030-67 41 3000
h.gutzmann@alexus.de

Stv. Vorsitzender

PD Dr. M. Haupt
Hohenzollernstr. 5
40211 Düsseldorf
Tel. 0211-109 1040
m.haupt@alzheimer-
praxis-duesseldorf.de

Schatzmeister

R. Kortus
Zentrum für Psychiatrie
Schloßstr. 50
71364 Winnenden
Tel. 07195-900 2700
r.kortus@zfn.de

Schriftführer

Dr. D. K. Wolter
Westfälische Klinik Münster
Postfach 202252
48046 Münster
Tel. 0251-591 5208/69
d.wolter@wfkp-hf.org

Beisitzer

Prof. Dr. G. Adler
Mannheim

Dr. L. M. Drach
Schwerin

PD Dr. W. Hewer
Rotweil

Dr. J. Tenter
Ravensburg-Weissenau

Wissenschaftl. Beirat

Prof. Dr. H. Förstl
Prof. Dr. G. Heuft
Prof. Dr. S. Kanowski
Prof. Dr. H.-J. Möller
Prof. Dr. H. Radebold

Weiterbildung:

Dr. Chr. Kretschmar

Stellungnahme DGGPP e. V., Gutzmann (Fortsetzung)

medikamentösen oder nichtmedikamentösen Therapieoption hinsichtlich patientenrelevanter Therapieziele.

Zur Einbeziehung der Sicht der Angehörigen in die Nutzenbewertung seien auch Vertreter der Deutschen Alzheimer Gesellschaft gehört worden.

In der Darstellung der Methodik mit ihrer eingehenden Auflistung der für die Analyse als relevant betrachteten Arzneimittelstudien fällt auf, dass zur Beurteilung von medikamentös bedingten Therapieveränderungen ausschließlich randomisierte und doppelblinde klinische Studien, sogenannte RCT (randomized clinical trials)-Studien in die Prüfung einbezogen wurden. Zudem sind alle diejenigen Studien ausgeschlossen worden, die eine kürzere Laufzeit als 16 Wochen hatten, unabhängig davon, ob es sich um RCT-Studien handelte. Für die später folgende Auswertung hat dieses Vorgehen zwei nachteilige Folgen:

1. der Ausschluss von RCT mit kürzerer Laufzeit schließt Studien zur Wirksamkeit der ChE-H bei nichtkognitiven Störungen teilweise aus, gleichwohl zeigt die klinische Erfahrung, dass nichtkognitive im Unterschied zu kognitiven Störungen in weniger als 16-24 Wochen wirksam medikamentös behandelt werden können.
2. der Ausschluss von Studien, die nicht strikt dem erforderlichen Design einer RCT folgen, hat die Konsequenz, dass sämtliche vorliegende und unkontrollierte open-label follow-up-Studien mit mehr als 1-jähriger Laufzeit nicht in die Bewertung eingebracht wurden. Für die Beurteilung eines ChE-H-Effektes auf Zielgrößen, wie Institutionalisierung oder Mortalität, sind bei der Alzheimer Demenz mit Einschluss von leichten und mittelschweren Kranken jedoch langfristige Beobachtungen notwendig. Im ersten halben (24 Wochen) oder im ersten Jahr (52 Wochen) sind endpunktrelevante Veränderungen in den genannten beiden Variablen nur selten zu erwarten. Andererseits sind aber RCT mit längerer Studienzeit als 6-12 Monaten ethisch nicht vertretbar, da den einem Scheinmedikament zugeordneten Studienpatienten die Wirkung des Verum nicht länger vorenthalten werden darf.

Stellungnahme DGGPP e. V., Gutzmann (Fortsetzung)

- Für das methodische Vorgehen ist zu empfehlen, dass bei der Beurteilung des Therapieeffektes der ChE-H auf nichtkognitive Symptome auch kürzere, also 12-wöchige, RCT-Studien akzeptiert werden sollten (z.B. Holmes et al. 2004 für Donepezil) und für die Bewertung eines Therapieeffektes auf die Heimeinweisungsrate auch die sogenannten Extensionsstudien, die für alle im Vorbericht geprüften ChE-H vorliegen. Nicht akzeptabel ist beispielsweise die auf Seite 92 des Vorberichts gemachte Bemerkung, dass zu dem „Zielkriterium Notwendigkeit der vollstationären Pflege (Institutionalisierung) in den eingeschlossenen Studien keine Daten berichtet“ würden. Die methodisch restriktive Vorgehensweise des alleinigen Einschlusses von Studien über 16 Wochen schlägt sich im Vorbericht daher auch auf Seite 160 nieder, wo es heißt, dass bei Donepezil und Rivastigmin keine eindeutigen Belege für deren Wirkung auf nichtkognitive Symptome erkennbar wären bei Galantamin nur geringe Wirkungen. Hierzu ist ferner anzumerken, dass die Metaanalyse von Trinh et al. (2003), die der Wirkung der ChE-H auf nichtkognitive Symptome einen erkennbaren, wenngleich begrenzten positiven Effekt zuspricht, weder im Literaturverzeichnis noch im Anhang zu den ausgeschlossenen Studien Erwähnung findet. Möglicherweise ist sie nicht aufgenommen worden, weil auch Studien zu Metrifonat und Tacrin in die Metaanalyse eingingen.
-) In gleicher Weise trifft zudem die Kritik gegenüber dem Vorbericht auf die Ausführungen auf Seite 160 zu, dass insgesamt eine zu geringe Zahl von RCT-Studien mit einer Laufzeit über 1 Jahr zur Verfügung stünden. Die klinische Erfahrung lehrt seit etlichen Jahren, dass über einen derart langen Zeitraum Alzheimerkranken auch aus noch so überzeugenden wissenschaftlichen Gründen eine ChE-H-Gabe vorenthalten werden darf.

In der vorgelegten Untersuchung ist der Abschnitt der Vergleiche unterschiedlicher ChE-H konstruktiv kritisch bearbeitet. Es ist den Verfassern zuzustimmen, dass die gegenwärtig vorliegenden klinischen Studien zur vergleichenden Wirksamkeit verschiedener ChE-H keine Rückschlüsse auf eine relevante differentielle Wirksamkeit zulassen. Selbst der im Vorbericht erwähnte Wirksamkeitsvorteil von Rivastigmin gegenüber Donepezil bleibt eng begrenzt.

Stellungnahme DGGPP e. V., Gutzmann (Fortsetzung)

Die Dosisabhängigkeit der unerwünschten Wirkungen der ChE-H und ihr primärer Austragungsort im gastrointestinalen Bereich wird im Vorbericht nachvollziehbar herausgearbeitet und ist angemessen dargestellt.

Die darüber hinaus klar und kritisch vorgenommenen Auseinandersetzungen mit der in der wissenschaftliche Gemeinschaft kontrovers diskutierten AD-2000-Studie und mit den von Kaduszkiewicz und Mitarbeitern vom Lehrstuhl für Allgemeinmedizin in Hamburg vorgetragenen negativen Kritiken an der Effektstärke der ChE-H reihen sich in die überwiegend kompetente Analyse des Vorberichts angemessen ein.

Letztlich kommt daher der Vorbericht zu dem Schluss, dass die ChE-H einen konsistenten Nutzen gegenüber Placebo im Bereich der Kognition aufweisen und einen möglichen Nutzen bezogen auf die Aktivitäten des täglichen Lebens. Den Verfassern des Vorberichts ist bei der Erläuterung des Nutzens auf die kognitive Leistung zu folgen, wenn sie die 3 Punkte-Änderung auf dem kognitiven Teil des Instruments ADAS darlegen. Nicht die generell gegebene Punktespanne von 0-70 Punkten sei entscheidend für die Effektbeurteilung der 3 Punkte, sondern allein relevant sei die bei der Alzheimer Demenz tatsächlich abgedeckte Punktespanne, die sich auf eine deutlich geringere Spanne erstreckt (S. 165-166). Aus klinischer Erfahrung ist ein in diesem Sinne einzuordnender 3 Punkte-Unterschied auf der ADAS-cog-Skala sehr wohl klinisch bedeutsam und für die pflegenden Bezugspersonen im Alltag spürbar.

Der Beurteilung des Vorberichts bezogen auf die nichtkognitiven (psychopathologische Begleitsymptome) Symptome kann nicht gefolgt werden. Mögliche Effekte bleiben infolge des Ausschlusses von RCT-Studien mit einer kürzeren (< 16 Wochen) Laufzeit unberücksichtigt, zumindest ungeprüft. Der Maßstab einer mindestens 16 Wochen dauernden RCT mag für die Beurteilung eines Therapieeffektes bei kognitiven Störungen angemessen sein, für die nichtkognitiven Symptome ist eine kürzere Laufzeit einer RCT durchaus angemessen, um eine klinisch relevante Veränderung nachweisen zu können.

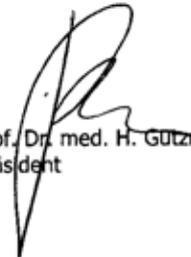
Stellungnahme DGGPP e. V., Gutzmann (Fortsetzung)

Langfristige Effekte der ChE-H, wie sie bei Variablen wie Heimeinweisung oder Mortalität untersucht werden könnten, lassen sich bei der Alzheimer Demenz nur in RCT-Studien über 2-3 Jahre identifizieren. Der gegenwärtige Kenntnisstand zu den Therapieeffekten der ChE-H lässt jedoch aus ethischen Gründen eine derart lange Laufzeit nicht mehr zu. Daher sollten auch die Ergebnisse aus open-label follow-up Studien in die Bewertung aufgenommen werden. Sie wurden bedauerlicherweise aus a priori gesetzten methodischen Kriterien ausgeschlossen (siehe S. 201-202).

Literatur:

Holmes C, Wilkinson D, Dean C, Vethanayagam S, et al. (2004) The efficacy of donepezil in the treatment of neuropsychiatric symptoms in Alzheimer disease. *Neurology* 63: 214-219

Trinh N-H, Hoblyn J, Mohanty S, Yaffe K (2003) Efficacy of cholinesterase-inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease. *JAMA* 289: 210-216



Prof. Dr. med. H. Gutzmann
Präsident



PD Dr. med. M. Haupt
Vizepräsident

Stellungnahme Eisai GmbH



Eingangsstempel

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

A05/19A Cholinesterasehemmer bei Alzheimer-Demenz

Name, Vorname, Titel und Funktion des/der Stellungnehmenden:

1. Schmitz, Helga, Dr.; Medizinische Direktorin, Eisai GmbH
2. Akintade, Latif, Dr.; Medical Director, Eisai Europe Limited
3. Friedhelm Leverkus, Leiter der Biometrie, Pfizer Pharma GmbH

Kontakt für weitere Nachfragen:

Name: Dr. Helga Schmitz
Institution: Eisai GmbH
Anschrift: Lyoner Straße 36, 60528 Frankfurt
Tel./Fax-Nr.: 069/66585-43 / 069/66585-45
E-Mail-Adresse: helga_schmitz@eisai.net

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Eisai GmbH (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal 5 DIN-A4-Seiten.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal 10 DIN-A4-Seiten.



Für alle Zitate sind Originalpublikationen (Volltext) beigelegt.



Die Stellungnahme ist von allen Stellungnehmenden unterschrieben.



Das Formblatt zur Darlegung potenzieller Interessenskonflikte liegt für alle Stellungnehmenden einzeln ausgefüllt vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Frankfurt, 05.10.06
Ort/Datum

A. Schulz
Unterschrift 1

[Handwritten Signature]
Unterschrift 2

[Handwritten Signature]
Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme Eisai GmbH (Fortsetzung)

Stellungnahme der Eisai GmbH zum IQWiG-Vorbericht A05/19-A: Cholinesterasehemmer bei Alzheimer-Demenz

Sehr geehrte Damen und Herren,

zu den von Ihnen am 8.9.2006 publizierten Vorbericht A05/19-A „Cholinesterasehemmer bei Alzheimer-Demenz“ möchten wir gerne wie folgt Stellung nehmen:

1. Im Vorbericht fehlende Originalstudien

Die Studie von Holmes⁷ ist eine 12-wöchige, randomisierte, doppelblinde Studie von Donepezil an Patienten mit leichter bis mittelschwerer Alzheimer-Demenz (NINCDS-ADRDA-Kriterien; MMSE 10-26) und neuropsychiatrischen Symptomen (Baseline-NPI-Wert > 11). Der Doppelblind-Phase ging eine 12-wöchige open-label-Behandlung mit Donepezil voraus. 134 Patienten erhielten 6 Wochen lang 5 mg Donepezil, anschließend 10 mg Donepezil für 6 weitere Wochen über einen Gesamtzeitraum von 24 Wochen. Diese Studie liefert wichtige Ergebnisse für die Therapieziele „begleitende Psychopathologie (Verhaltensstörungen)“ und „Betreuungsaufwand“ (im Sinne von „Caregiver Distress“).

Fazit: Die Studie von Holmes⁷ muss in die Bewertung eingehen, da sie die formalen, vom IQWiG festgelegten Kriterien einer 24-wöchigen Behandlungsdauer erfüllt.

2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht

A. Wirksamkeit von Donepezil bezüglich der einzelnen Therapieziele

A1. Aktivitäten des täglichen Lebens (Alltagsfähigkeiten; ADLs)

Es ist nicht gerechtfertigt, den in der Studie von Gauthier 2001 beobachteten Unterschied auf der DAD-Skala als „ungewöhnlich groß“ anzuzweifeln. Der Vergleich der Ergebnisse zum DAD der Gauthier-Studie mit denen der vier Galantaminstudien (für Rivastigmin liegen keine Ergebnisse zum DAD vor) ist fraglich, weil die Patientenpopulationen nicht vergleichbar sind. Während bei allen Galantamin-Studien der Baseline-DAD-Wert etwa 70 war und der Baseline-MMSE-Wert zwischen 19 und 21 lag, waren in der Gauthier-Studie die entsprechenden Werte deutlich niedriger (DAD: 58; MMSE: 14), d.h. die Erkrankung war bei diesen Patienten stärker vorangeschritten. Die Wahrscheinlichkeit, bestimmte Alltagsfähigkeiten zu verlieren, hängt vom Schweregrad der Alzheimer-Demenz, der mit Hilfe des MMSE bestimmt wurde, ab⁵. Ein MMSE-Wert von 16 stellt beim Verlust der Alltagsfähigkeiten einen wichtigen Übergangspunkt dar⁵. Die Patienten in den Galantamin-Studien befinden sich auf der einen Seite, die der Gauthier-Studie auf der anderen Seite dieses Übergangspunktes. Auch in einer retrospektiven Analyse einer Studie mit Rivastigmin¹⁰ konnte gezeigt werden, dass der therapeutische Effekt abhängig davon war, welcher Level der Erkrankung zu Beginn der Studie vorlag. Patienten mit einem niedrigeren Level von Störungen der Alltagsfunktionen zu Studienbeginn tendierten dazu dieses Level beizubehalten, während die Behandlung von stärker beeinträchtigten Patienten im Mittel zu einem verzögerten Verlust von Alltagsfähigkeiten führte. Eine solche Reduktion des Verlustes von Alltagsfähigkeiten oder Stabilisierung von ADLs auf dem Baselinelevel konnte in einer Reihe von Donepezil-Studien gezeigt werden (Feldman, Winblad, Burns). **Fazit:** Der in der Meta-Analyse ermittelte Effekt darf nicht aufgrund der Ergebnisse der Gauthier-Studie als Überschätzung des Effekts angesehen werden.

Darüber hinaus ist auch die auf 54 Wochen angelegte Studie von Mohs, gerade aufgrund ihres besonderen Designs geeignet, den positiven Effekt von Donepezil auf die Alltagsfähigkeiten zu demonstrieren. Die Studie von Mohs ist eine doppelblinde plazebo-kontrollierte Multizenterstudie. Primärer Endpunkt war die Zeit bis zu der Verschlechterung der Basis-ADL um mindestens einen Punkt oder der Verschlechterung der instrumentellen ADL um mindestens 20% oder der Verschlechterung (=Anstieg) des CDR-Wertes um mindestens einen Punkt. Patienten, die wegen anderer Gründe die Studie abbrachen oder vor Beobachtungsende aus der Studie ausschieden (lost to follow-up), wurden als zensiert betrachtet. Der Vorteil dieser Studie liegt darin, dass nicht ein Summenscore als primärer Endpunkt, sondern ein in besonderem Maße für Patienten relevantes Kriterium als primärer Endpunkt gewählt wurde. Alle drei Komponenten betreffen prospektiv definierte und klinisch evidente Endpunkte der Alltagsbewältigung. Insofern ist die Studie gezielt darauf ausgerichtet eine Beeinträchtigung der Alltagsfunktion zu ermitteln. Der Grund für die hohe Dropout-Rate ist das entsprechend angelegte Studiendesign. Vergleicht man aber Patienten, die aufgrund anderer als den als primäre Endpunkte definierten Gründen ausgeschieden sind, so ist die Anzahl vergleichbar (26 % Plazebo vs. 28% Donepezil). Da das Erreichen des Endpunktes „Verschlechterung der Funktion“ Zeit benötigt, wurden Patienten, die diese Kriterien nicht erfüllten, standardmäßig „zensiert“ in die primäre Analyse mit einbezogen. D.h., es wird die Information, dass bei einem Patienten bis zur Zensierung kein Ereignis eingetreten ist, berücksichtigt. In der Mohs Studie wurden die ADL-Funktionen bei Patienten unter Donepezil 72 % länger

Stellungnahme Eisai GmbH (Fortsetzung)

erhalten als in der Placebo-Gruppe, was eine Verzögerung der klinisch nachweisbaren funktionellen Verschlechterung um durchschnittlich 5 Monate bedeutete. **Fazit:** Auch in der Mohs-Studie wird der positive Effekt von Donepezil auf die Alltagsfähigkeiten deutlich.

Da einige Daten zu den Alltagsfähigkeiten nicht umfassend in die Publikationen eingegangen sind, haben wir die in Tabelle 21 des Vorberichts fehlenden Daten aus den Clinical Study Reports (CSR) ergänzt (siehe Anhang). Unter Berücksichtigung dieser Daten zeigt die Meta-Analyse weiterhin eine statistisch signifikante Überlegenheit von Donepezil (siehe Anhang: Abbildung 1). Aufgrund des Cohen's d von 0,28 kann von einem mittelstarken Effekt gesprochen werden. Der Heterogenitätstest ist grenzwertig, I^2 liegt mit 0,57 in einem als moderat gekennzeichneten Bereich. Selbst wenn die beiden Studien mit den größten und kleinsten Werten (Gauthier und Tariot) nicht berücksichtigt würden, blieben die Ergebnisse konstant (Cohen's $d=0,27$, $p<0,0001$; siehe Anhang: Abbildung 2).

Fazit: Es ist also nicht nur von einem Hinweis, sondern von einem Nachweis eines Effekts von Donepezil auf die ADL auszugehen.

A2. Begleitende Psychopathologie (Verhaltensstörungen)

Zwei Limitationen für die vom IQWiG ausgewählten Studien können dazu führen, dass der positive Effekt von Donepezil bzgl. der Verhaltensstörungen unterschätzt wird: 1. In den Studien von Winblad, Prasher und Tune war der Baseline-NPI-Wert niedrig (8-18), d.h., das Vorhandensein und die Ausprägung von Verhaltensstörungen zu Studienbeginn war relativ gering, da normalerweise Patienten mit schwereren Verhaltensstörungen aufgrund des Studienprotokolls von der Studie ausgeschlossen werden. 2. Die Verhaltensstörungen waren nicht so stark ausgeprägt, da diese schon vor Studienbeginn (und auch während der Studie) mit Psychopharmaka behandelt wurden. In der Studie von Tariot 2001 erfolgte bei mehr als der Hälfte der Patienten die nicht kontrollierte Gabe von Psychopharmaka. Dadurch könnten die uneindeutigen Ergebnisse dieser Studie und die Heterogenität zur Studie von Gauthier 2002 erklärt werden. Auch ein I^2 von 74% deutet darauf hin, dass die Studien sehr heterogen sind und eine Zusammenfassung in einer Metaanalyse fraglich ist.

Zwei Studien verdeutlichen die genannten Limitationen der Untersuchungen. In der in ihrer Art einmaligen, prospektiv geplanten, spezifisch auf die Erfassung der Wirksamkeit von Donepezil auf Verhaltensstörungen mit dem NPI als primärem Wirksamkeitsparameter angelegten Studie von Holmes⁷ lag der Baseline-NPI-Wert bei 25. Nach 12-wöchiger Donepezil-Behandlung verbesserten sich sowohl der NPI-Gesamtwert (um mehr als 10 Punkte) als auch neun der zehn Verhaltensstörungen, die mit dem NPI erfasst werden. Die Patienten erhielten für weitere 12 Wochen entweder Donepezil oder Placebo (doppelblinde, randomisierte Studienphase). Innerhalb dieses Zeitraums verschlechterten sich die Verhaltensstörungen in der Placebo-Gruppe signifikant (NPI-Gesamtwert: +3,3 Punkte), während sich die Donepezil-Patienten weiter verbesserten (-2,9 Punkte). In der zweiten Studie⁴ hatten die Patienten relativ schwere Verhaltensstörungen (NPI = 31) und alle Psychopharmaka wurden vor Studienbeginn abgesetzt. Durch die 20-wöchige Donepezil-Behandlung (8 Wochen Donepezil, anschließend 12 Wochen Donepezil plus Placebo (Kontrollgruppe: 12 Wochen Donepezil plus ein Antidepressivum)) wurde der NPI-Gesamtwert um mehr als 6 Punkte gesenkt (signifikante Verbesserung gegenüber Baseline). Eine Vorteil von knapp 2 NPI-Punkten gegenüber Placebo entspricht einer Verringerung der Häufigkeit oder Schwere eines bestimmten neuropsychiatrischen Symptoms¹⁶.

In dem im Auftrag des NICE erstellten HTA-Bericht⁸ wird zusammenfassend festgestellt, dass es Hinweise gibt, dass Donepezil im Vergleich zu Placebo einen Effekt auf die Verbesserung oder die Limitierung einer weiteren Verschlechterung der Verhaltensstörungen (gemessen mit dem NPI) hat.

Fazit: Die detailliertere Analyse der vorhandenen Evidenz zeigt, dass es sehr wohl deutliche Hinweise auf die günstige Beeinflussung der Verhaltensstörungen durch Donepezil gibt und dass der Effekt von Donepezil klinisch relevant ist.

B. Bewertung einzelner Studien

B1. Prasher/Tariot

Bei der Interpretation der Ergebnisse der Studien von Prasher (Down-Syndrom und AD) und Tariot (25 % der Patienten mit schwerer AD) muss berücksichtigt werden, dass in diesen Studien die Wirksamkeit von Donepezil in Indikationen untersucht wurde, für die Donepezil (noch) nicht zugelassen ist.

B2. Vergleichsstudien

Beim direkten Vergleich zwischen Donepezil und Rivastigmin bezogen auf die Alltagsfähigkeiten heißt es auf Seite 147 „Die Ergebnisse [der Studie Bullock 2005] sind allerdings von eingeschränkter Aussagekraft, da

Stellungnahme Eisai GmbH (Fortsetzung)

die Rate von Therapieabbrüchen in beiden Gruppen hoch und zwischen den beiden Gruppen unterschiedlich war". Unter Berücksichtigung dieser Einschränkung und der beiden anderen Vergleichsstudien wird zusammenfassend festgestellt, dass „die Daten nicht ausreichend für einen Nachweis der Über- oder Unterlegenheit einer der beiden Wirkstoffe“ (S. 148) sind. Bitte stellen Sie diese Ergebnisse in der Tabelle 45 richtig dar: (↔) statt (↓). Bitte korrigieren Sie auch das Fazit „...Hinweise auf einen Zusatznutzen durch Rivastigmin gegenüber Donepezil beim Therapieziel Besserung bzw. Verhinderung der Beeinträchtigung der Aktivitäten des täglichen Lebens gibt...“ (S. 171) entsprechend der Zusammenfassung von Seite 148.

Darüber hinaus ist die Aussagekraft der Bullock Studie fraglich, da etwa 50% mehr Rivastigmin-Patienten die Studie in der Titrationsphase aufgrund von unerwünschten Ereignissen abgebrochen haben. Die LOCF-Analyse führt so zu einer Überschätzung der Wirksamkeit von Rivastigmin.

Im Vorbericht heißt es „... ergab sich nach zwei Jahren in der ITT-LOCF-Auswertung ein statistisch signifikanter Vorteil in der ADCS-ADL-Skala zugunsten von Rivastigmin...Es kann vermutet werden, dass diese Diskrepanzen aufgrund der Abbruchraten zustande kommen...“. In der Publikation wird bemerkt, dass die Unterschiede in den Per-Protokoll-Auswertungen keine statistische Signifikanz mehr besaßen. Ob dieses auf einer Änderung des Effektschätzers und/oder einer verminderten Power beruht, bleibt unklar.

Außerdem wurde ein langsames und für Rivastigmin vorteilhaftes Titrationsschema verwendet, was zu einer Verzerrung der Ergebnisse bezüglich der Häufigkeit von unerwünschten Ereignissen führen könnte.

Weiterhin werden Aussagen aufgrund von Subgruppen-Analysen gemacht, deren Analysen auf der ITT-LOCF-Population beruhen und somit zu statistischer Verzerrung der Ergebnisse führen. Beispielhaft Lebensqualität der (betreuenden) Angehörigen: Es wurde nur die Analyse der Subgruppen nach Alter publiziert, nicht aber der Gesamtpopulation. Es ist unklar, ob Subgruppenanalysen, wie vorgelegt, a priori definiert wurden und entsprechend valide sind.

C. Studienqualität

Bei den Publikationen der von Eisai und Pfizer durchgeführten Studien wurden teils geringe, teils grobe „Mängel“ festgestellt, welche in dieser Form nicht bestehen. Im Anhang weisen wir nach, in welchen Punkten diese Kritik unbegründet ist, und bitten um entsprechende Korrektur.

Einige der vorliegenden Studien waren zulassungsrelevant. Die Zulassungsbehörden prüfen die Validität der Ergebnisse und die Qualität der Studien anhand der Clinical Study Reports und der Daten. Von der pharmazeutischen Industrie gesponserte klinische Prüfungen müssen nach GCP durchgeführt werden und folgen den ICH Guidelines. Die Zulassungsbehörden in mehr als 80 Ländern weltweit haben ihre Entscheidung zu Donepezil (Aricept®) nach Prüfen der Validität und Qualität der Studien positiv getroffen.

C1. Vergleich der Baseline-Gruppen

Wir stimmen dem IQWiG zu, dass es in den einzelnen Studien keine Probleme bezüglich der Ungleichheit der Gruppen gibt. Dennoch möchten wir darauf hinweisen, dass ein so genannter Homogenitäts-Test zum Vergleich der Baseline-Gruppen nicht überbewertet werden soll und nicht nötig ist (siehe auch¹³). Ein p-Wert < 0,05 bezüglich einzelner Patientenmerkmale bedeutet nicht, dass ein Problem mit der Homogenität der Baseline-Gruppen besteht.

C2. Multiples Testen

Wird wie in den EMEA Guidelines die Signifikanz für beide Hauptparameter gefordert, so ist eine Alpha-Adjustierung nicht notwendig. Die Wirksamkeit ist nur dann nachgewiesen, wenn beide Parameter Signifikanz zeigen; zeigt nur ein Parameter Signifikanz, ist der Nachweis der Wirksamkeit nicht erfolgt. Für die Beurteilung der Aussage einer einzelnen Studie, wie sie die Zulassungsbehörden und Reviewer einer Zeitschrift vornehmen, ist die Alpha-Adjustierung sicher relevant. Für die Synthese der Daten in einer Metaanalyse gehen die p-Werte nicht ein, erzeugen keinen Bias und sind damit irrelevant.

C3. Fallzahlplanung

Die Fallzahlplanung ist bei der ethischen Beurteilung einer einzelnen Studie von Relevanz. Auch kann eine zu niedrige Power darauf hinweisen, dass vorhandene Unterschiede nur mit einer geringen Wahrscheinlichkeit aufgedeckt werden können. Bei einer Metaanalyse werden Studien zusammengefasst, d.h. die Beschreibung einer prospektiven Fallzahlplanung in der Publikation ist sicherlich wünschenswert, aber sie ist für diese Fragestellung als Qualitätsmerkmal einer Studie nicht relevant. Eine post-hoc Poweranalyse kann jedoch zeigen, dass mögliche Unterschiede gar nicht aufgedeckt werden können.

Stellungnahme Eisai GmbH (Fortsetzung)

C4. LOCF

Das IQWiG stellt fest, dass die theoretische Kritik an der Methodik zum Ersetzen fehlender Werte mittels LOCF bei den in den Vorbericht eingeschlossenen Studien nicht durch empirische Daten gestützt wird. Allerdings hält das IQWiG Sensitivitätsanalysen mittels anderer Ersetzungsstrategien für wünschenswert, um die Robustheit der Ergebnis besser einschätzen zu können. Wir möchten den Standpunkt des IQWiG, dass die theoretische Kritik am LOCF-Ansatz durch empirische Daten nicht gestützt wird, untermauern. Dazu sind wir der Bitte des IQWiG, Sensitivitätsanalysen mittels anderer Ersetzungsstrategien durchzuführen, gerne nachgekommen. Wir haben beispielhaft eine Re-Analyse für die Studie von Rogers durchgeführt. Mit dem Multiple-Imputation-Verfahren^{11, 12} wurden die fehlenden Werte alternativ zum LOCF-Verfahren ersetzt. Es wird davon ausgegangen, dass der Ausfallmechanismus „MAR“ (missing at random) ist, d.h. durch die vorhandenen Informationen zum ADAS-Cog abgebildet werden kann. Die Informationen, die eingeflossen sind, sind die ADAS-Cog-Werte zu Baseline und Behandlung. Die Annahme ist die, dass mit Verschlechterung des ADAS-Cog die Wahrscheinlichkeit des Dropouts im Verlauf steigt. Wir schließen ein „informativ censoring“ aus. Es wurden drei verschiedene Verfahren (MCMC, Propensity Score, Regression) angewandt.

Da der Verlauf unter den verschiedenen Therapien unterschiedlich ist, wurden die Ersetzungsverfahren für jede Medikationsgruppe getrennt durchgeführt.

Die Ergebnisse der Multiple Imputation bestätigen die LOCF-Analysen (siehe Anhang). Alle Baseline-Final Veränderungen fallen ein wenig geringer aus (Tabelle 2). Da dieses in der Placebo-Gruppe aber überproportional geschieht, ist der Medikationseffekt sogar noch größer (Tabelle 3).

Fazit: Der Vergleich des LOCF-Ansatzes mit anderen Ersetzungsstrategien liefert keinen Hinweis, dass der LOCF-Ansatz zu einer Verzerrung zum Vorteil des Verums führt. Die theoretische Kritik an dem LOCF-Ansatz ist unbegründet.

C5. AD2000

Wir begrüßen die Entscheidung, die AD 2000 Studie nicht in die Nutzenbewertung mit einzubeziehen. Die im Vorbericht genannten, allesamt zutreffenden und validen Kritikpunkte möchten wir um eine Reihe wichtiger Aspekte ergänzen. Als einen der wichtigsten zusätzlichen Kritikpunkte erachten wir die Einschlusskriterien der Studie, die sich zum einen nicht ausschliesslich im Rahmen der zugelassenen Indikation bewegen, zum anderen eine korrekte Interpretation der Ergebnisse nicht zulassen. Die Einschlusskriterien erlauben eine Aufnahme von Patienten mit Vaskulärer Demenz und mit gemischten Demenz-Formen, so dass die Studie schon aus formalen Gründen ausgeschlossen werden muss. Eine Voraussetzung für die Einbeziehung von Studien zur Evaluation durch das IQWiG ist nämlich, dass die Studie sich im Rahmen der zugelassenen Indikation bewegt. In der AD 2000 waren jedoch ca. 20 % der Patienten mit Vaskulärer Demenz bzw. mit M. Parkinson eingeschlossen. Die Interpretation der Ergebnisse hinsichtlich dieser unterschiedlichen Patientengruppen ist äußerst problematisch. Darüber hinaus sollten laut Studienprotokoll gerade diejenigen Patienten, bei denen eine klare Indikation für Donepezil besteht, nicht in die Studie eingeschlossen werden. Stattdessen sollten nur Patienten, bei denen ein echter Zweifel (substantial uncertainty) besteht, ob sie von einer Therapie mit Donepezil profitieren könnten, eingeschlossen werden. Auf diese Patientengruppe kann dann auch nur verallgemeinert werden. Selbst wenn diese Studie keine methodischen Mängel gehabt hätte, hätten keine Aussagen über Patienten getätigt werden können, die nach Ansicht der Ärzte von Donepezil profitieren können. Dieses ist jedoch die interessierende Population, die im Fokus der Nutzenbewertung steht.

Bei der AD 2000 Studie sollte neben dem Effekt von Donepezil auch der Effekt von Aspirin untersucht werden. Die Begleitmedikation (bzw. zweite Studienmedikation) sowie die hier im Gegensatz zu der üblichen Studienpraxis zugelassenen schwerwiegenden Begleiterkrankungen führen zu weiteren Unwägbarkeiten hinsichtlich der Interpretation der Ergebnisse, insbesondere auch bzgl. der Zuordnung und Häufigkeit von unerwünschten Ereignissen und Drop outs.

Weiterhin ist nicht klar, ob die Studie auf Basis der seinerzeit geforderten Qualitätsstandards (good clinical practice) durchgeführt wurde. Hinsichtlich der Patientenzahl, der Behandlungsdauer und Endpunktauswahl sind ebenfalls eine Reihe kritischer Anmerkungen angebracht⁶, z.B. die für den Endpunkt „Zeit bis zur Pflegeheimübersiedlung“ unangemessene Auswahl der Studienpopulation und die Anfechtbarkeit der Definition des zweiten Hauptzielkriteriums „Progredienz der Behinderung“, definiert durch den Verlust wichtiger, für die Alltagstauglichkeit notwendiger Fähigkeiten.

Die Studie war für die aus ihr gezogenen Schlussfolgerungen unterpowert. Somit konnten mit der geringen Fallzahl, selbst wenn man alle oben genannten Bedenken außer Acht lassen würde, vorhandene Unterschiede gar nicht aufgedeckt werden.

Stellungnahme Eisai GmbH (Fortsetzung)

Fazit: Die AD2000 wurde aufgrund ihrer methodischen Schwachpunkte auch im aktuellsten Cochrane Review zu Donepezil³ nicht berücksichtigt. In Erwägung all dieser Kritikpunkte und Einschränkungen halten wir es auch für künftige Evaluationen für ausgeschlossen, die Studienergebnisse der AD 2000 in eine Nutzenbewertung mit einzubeziehen. Weitere Details zu den Kritikpunkten: siehe^{1, 6, 9, 14, 15}.

Gesamt-Fazit: Wo keine oder nur geringe Mängel konstatiert werden können, muss die Tabelle 8 des Vorberichts entsprechend der hier gegebenen Informationen positiv verändert werden. Die AD2000 darf weder in den Abschlussbericht noch in künftige Nutzenbewertungen einbezogen werden.

3. Anmerkungen zur projektspezifischen Methodik

A. Nicht berücksichtigte Evidenz

Aufgrund der vom IQWiG gewählten Ein-/Ausschlusskriterien sind wichtige Studienergebnisse nicht berücksichtigt (siehe z.B. Punkt 2: bei dem Therapieziel „Verhaltensstörungen“ weisen wir auf weitere, relevante Studien hin).

Die vom IQWiG im Berichtsplan gewählte Methodik, ausschließlich RCTs mit einer Studiendauer von mindestens 16 Wochen zur Bewertung zuzulassen, ist unseres Erachtens nicht nachvollziehbar. In Deutschland ist es allgemein akzeptierter Standard, dass eine Beobachtungsdauer von mindestens 12 Wochen erste Hinweise auf den therapeutischen Effekt liefern kann. Dies wird durch die Empfehlungen der Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ) nachhaltig bestätigt².

Auch sollten Studien niedrigeren Evidenzgrades mit einbezogen werden, wenn RCTs nicht vorhanden sind.

Fazit: Als Ausschlusskriterium muss "mindestens 12 Wochen" definiert werden und weitere, relevante Studien (siehe Punkt 2.) sollten berücksichtigt werden.

B. Klinischer Gesamteindruck

Die im Rahmen des Amendments vorgenommene Abwertung des klinischen Gesamteindrucks erscheint aus unserer Sicht nicht nachvollziehbar. Die Skalen zur Erfassung des klinischen Gesamteindrucks (z.B. der CIBIC-plus) sind bei der Bewertung des therapeutischen Effektes bezüglich der Alzheimer-Krankheit international anerkannte und validierte Verfahren. Andere Behörden, die sich der evidenz-basierten Nutzenbewertung widmen (z.B. NICE), lassen keinen Zweifel an der Bedeutung der Sinnhaftigkeit des klinischen Gesamteindrucks als relevante Messgröße.

Tatsächlich wurden die Skalen zur Erfassung des klinischen Gesamteindrucks zuerst von den Zulassungsbehörden gefordert um sicherzustellen, dass die Wirkung der Antidementiva klinisch relevant ist. Per Definitionem werden mit diesen Skalen die verschiedenen Stufen einer klinisch bedeutsamen Veränderung abgebildet, die nicht mit quantitativen, leistungsbezogenen Meßverfahren wie der ADAS-Cog oder dem MMSE erfasst werden.

Die Verbesserung des klinischen Gesamteindrucks durch Donepezil ist eindeutig nachgewiesen. Dadurch wird für ein weiteres patientenrelevantes Therapieziel der Nutzen von Donepezil deutlich.

Fazit: Der klinische Gesamteindruck muss in vollem Umfang in die Nutzenbewertung der Cholinesterase-Hemmer einfließen. Die Verbesserung des klinischen Gesamteindrucks durch Donepezil muss auch ins Fazit des IQWiG-Berichts als patienten-relevanter Nutzen aufgenommen werden.

C. Bewertung der Studienqualität

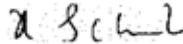
C1. ITT-Prinzip

Das ITT-Prinzip besagt, dass Patienten, die randomisiert werden, auch analysiert werden sollen. Es ist Standard, dass Patienten, die keine Werte unter Therapie haben, nicht in die Analyse eingehen. Dies wird auch in der ICH-E9 Guideline beschrieben (siehe Anhang). Diesem Prinzip folgen die vorliegenden Donepezil-Studien, wie im Anhang näher erläutert wird.

C2. Deskriptive Bewertung der Studienqualität

Im Sinne einer einheitlichen Abhandlung der in die Bewertung eingeflossenen Substanzen ist ein identischer Aufbau im Bereich "Studien- und Berichtsqualität" anzustreben. Derzeit werden lediglich 2 der 3 Substanzen mit einem Gesamturteil bewertet. **Fazit:** Für alle 3 Substanzen, also auch für Rivastigmin, muss ein Gesamturteil zur Studienqualität gefällt werden.

Darüber hinaus bitten wir das IQWiG, die im Anhang aufgeführten kleinen Mängel im Vorbericht zu korrigieren.


Dr. Helga Schmitz


Dr. Latif Akintade


Friedhelm Leverkus

Stellungnahme Eisai GmbH (Fortsetzung)

Anhang zur Stellungnahme der Eisai GmbH zum IQWiG-Vorbericht A05/19-A

Literaturverzeichnis

1. Akintade, L; Zaiac, M; Rleni, J; McRae, T: **AD2000: Design and conclusion**. In: Lancet 364 (2004) S.1214-1215.
2. Arzneimittelkommission der Deutschen Ärzteschaft: **Therapieempfehlungen der Arzneimittelkommission der Deutschen Ärzteschaft: Demenz**. Berlin 2004.
3. Birks, J; Harvey, RJ: **Donepezil for dementia due to Alzheimer's disease**. In: Cochrane Database.Syst Rev.(2006) Nr. 1, S.CD001190.
4. Cummings, JL; McRae, T; Zhang, R: **Effects of donepezil on neuropsychiatric symptoms in patients with dementia and severe behavioral disorders**. In: Am.J Geriatr Psychiatry. 14 (2006) Nr. 7, S.605-612.
5. Feldman, HH; Van Baelen, B; Kavanagh, SM; Torfs, KE: **Cognition, function, and caregiving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis**. In: Alzheimer.Dis Assoc.Disord 19 (2005) Nr. 1, S.29-36.
6. Frölich, L; Heuser, I; Wellek, S: **Zur klinischen Relevanz von Cholinesterase-Inhibitoren bei Alzheimer Krankheit - Eine klinische und biometrische Kritik der AD2000-Studie**. In: psychoneuro 32 (2006) Nr. 1, S.32-37.
7. Holmes, C; Wilkinson, D; Dean, C; Vethanayagam, S; Olivieri, S; Langley, A; Pandita-Gunawardena, ND; Hogg, F; Clare, C; Damms, J: **The efficacy of donepezil in the treatment of neuropsychiatric symptoms in Alzheimer disease**. In: Neurology 63 (2004) Nr. 2, S.214-219.
8. Loveman E; Green: **The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease**. In: Health Technology Assessment; VOL: 10 (1); p.iii-160 /2006/(2006) Nr. 1, S.p-160.
9. Maelicke, A: **Kritische Bewertung der Ergebnisse Langzeitbehandlung bei Demenz vom Alzheimer-Typ (AD2000-Studie)**. In: psychoneuro 30 (2004) Nr. 9, S.514-516.
10. Potkin, SG; Anand, R; Hartman, R; Veach, J; Grossberg, G: **Impact of Alzheimer's disease and rivastigmine treatment on activities of daily living over the course of mild to moderately severe disease**. In: Prog Neuropsychopharmacol Biol Psychiatry 26 (2002) S.713-720.
11. SAS Institute Inc.Cary NC USA: **The MI Procedure (Auszug); Sas Online Doc, Version 8**. 1999 <http://support.sas.com/rnd/app/da/new/802ce/stat/chap9/index.htm> (6.10.06).
12. Schafer, JL: **Multiple imputation: a primer**. In: Stat Methods Med Res 8 (1999) Nr. 1, S.3-15.
13. Senn, S: **Statistical Issues in Drug Development**. Chichester. John Wiley & Sons 1997.
14. Silver, G; Kay, R; Phillips, R: **Appendix A: Critical Analysis of the AD2000 study**. 2006 <http://www.nice.org.uk/download.aspx?o=342613> (20.9.06).
15. Thomas, S; Windeler, J: **Methodische Kritikpunkte zur Studie "Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double blind trial (Lancet 363 [2004] 2105-2115)"**. Essen: MDS 2005.
16. Trinh, NH; Hoblyn, J; Mohanty, S; Yaffe, K: **Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: a meta-analysis**. In: JAMA: the journal of the American Medical Association 289 (2003) Nr. 2, S.210-216.

Stellungnahme Eisai GmbH (Fortsetzung)

Meta-Analysen zu den Alltagsfähigkeiten

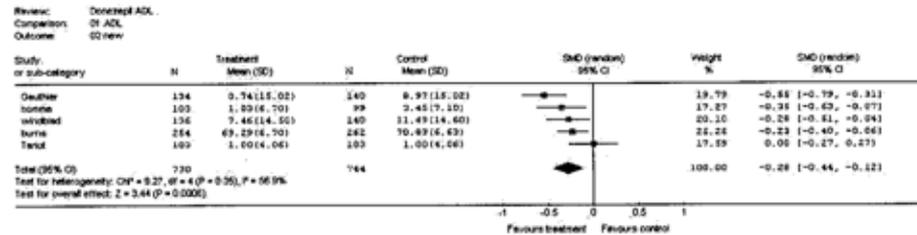


Abbildung 1: Meta-Analysen zu den Alltagsfähigkeiten unter Berücksichtigung aller Studien

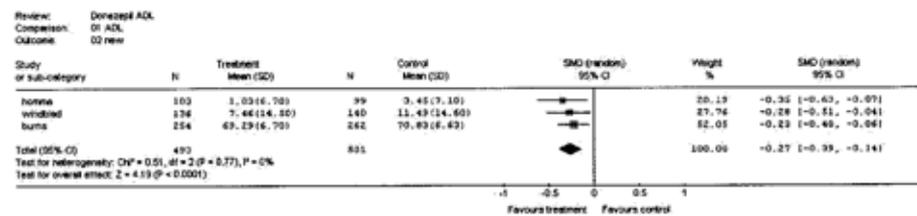


Abbildung 2: Meta-Analysen zu den Alltagsfähigkeiten ohne Berücksichtigung der Studien mit den größten und kleinsten Werten

LOCF vs. Multiple Imputation

Study 302- Rogers 1998

	N Baseline	N week 24	„Missings“
Placebo	153	131	14%
5mg	152	130	17%
10mg	150	104	31%

Table 0: Distribution of Ns

	LOCF	MCMC	Propensity Score	Regression
Placebo	1.73 (0.47)	2.08 (0.52)	2.02(0.54)	2.08 (0.52)
5mg	-0.79 (0.40)	-0.71 (0.42)	-0.76 (0.41)	-0.75 (0.42)
10mg	-1.14 (0.46)	-1.08 (0.55)	-1.04(0.59)	-1.11 (0.56)

Table 1: ADAS-Cog Baseline Final Change- Unadjusted Means with SEM

	LOCF	MCMC	Propensity Score	Regression
Placebo	1.72(0.44)	2.16 (0.60)	2.03(0.63)	2.19 (0.60)
5mg	-0.78 (0.44)	-0.62 (0.60)	-0.73(0.62)	-0.63 (0.64)
10mg	-1.15 (0.45)	-1.01 (0.65)	-1.03(0.68)	-1.02(0.66)

Table 2: ADAS-Cog Baseline Final Change- LS Means with SEM

	LOCF	MCMC	Propensity Score	Regression
Placebo vs 5mg	2.50 (p<0.001)	2.78 (p<0.001)	2.76 (p<0.001)	2.82 (p<0.001)
Placebo vs 10 mg	2.87 (p<0.001)	3.17 (p<0.001)	3.07 (p<0.001)	3.20 (p<0.001)

Table 3: ADAS-Cog Baseline Final Change – Contrasts with p-value

Textpassage aus der ICH-E9 Guideline zum ITT-Prinzip:

There are a limited number of circumstances that might lead to excluding randomised subjects from the full analysis set including the failure to satisfy major entry criteria (eligibility violations), the failure to take at least one dose of trial medication and the lack of any data post randomisation. Such exclusions should always be justified. Subjects who fail to satisfy

Stellungnahme Eisai GmbH (Fortsetzung)

Mängelliste

Wir bitten das IQWiG, folgende kleine Mängel im Vorbericht zu korrigieren:

Tab. 43: 12 statt 11 Studien für DON (siehe Tab. 4) und Patientenzahl entsprechend korrigieren

S. 162: Unerwünschte Ereignisse (UE): Es fehlt eine differenzierte Darstellung der Häufigkeit der UE wie auf S. 165 beschrieben. Bitte ergänzen.

S. 17: Randomisierungsprozess und Verdeckung der Gruppenzuweisung (allocation concealment): Was verstehen die Autoren unter einem "angemessenen Vorgehen"?

S. 93: Mortalität: „Einen eindeutigen Hinweis auf einen schädlichen Effekt geben die berichteten Daten nicht.“ bitte umformulieren zu „Ein Hinweis auf eine günstige oder ungünstige Beeinflussung der Mortalität durch Donepezil lässt sich aus diesen Daten nicht entnehmen.“ (vgl. Rivastigmin S. 136) Begründung: Bei Rivastigmin und Galantamin wurde auch beide möglichen Richtungen (günstig oder ungünstig) dargestellt.

Studypublication quality

In this document we will comment on the issues raised by IQWiG on study quality in general and secondly on each single study.

The topics on study quality have been focussing on the following points:

1. Randomization/Concealment
2. Blinding of endpoint raters
3. Sample size planning
4. Study dropouts named/reasons given
5. ITT analysis inadequate

1. Randomization/Concealment

Regarding the question of randomisation and adequate concealment we can confirm after checking of the clinical study reports for all Eisai and/or Pfizer sponsored studies in the table (Burns, Gauthier, Homma; Krishnan, Mohs, Rogers, Seltzer, Tariot, Tune und Winblad) that randomisation /concealment was adequate.

It has been in all cases a computer generated randomisation and blinding of medication by ensuring active and placebo treatments appear identical and are packed identically.

Although it is not explicitly stated in some publications, investigators and patients, as well as Pfizer/Eisai personnel were blinded to treatment allocation in these studies.

Following we have cited some relevant passages out of the CSRs:

Burns (unclear/unclear):

(CSR page 7, 8)...“patient randomisation schedule was computer generated.. Randomisation of study drug was performed in blocks of 6 (page 3)

each medication card included a two-part double-blind label (page 8)

CSR chapter. 4.1.2 and 4.1.3 ‘identical appearing...donepezil or placebo...Each dosage was supplied as two tablets within a sealed, blister-packaged card.’ Randomisation specified as ‘performed in blocks of six and each centre was allocated such blocks based on enrolment’ (3.1)

Gauthier: (yes d: Information on this is found in the publication of Feldman 2001 /unclear)

CSR page 31/32: ...doses were ‘individually packaged...indistinguishable from each other...within a sealed, blister-packaged card’ (9.4.6). Randomisation was undertaken by Pfizer by computer generation.

Homma (unclear/unclear):

....Each patient was assigned to a treatment group according to a computer generated schedule; randomized allocation schedule with a block size of 6 provided by an independent organization (controller committee) which allocates test drug and placebo and ensures the identity of appearance of packages and formulations between drugs tested.

Krishnan: (yes /unclear):

CSR page 26:... ‘Identical-appearing...tablets of donepezil or placebo were used...Study medication was supplied in sealed, blister-packaged cards...’ (9.4.2 and further details given in 9.4.6). Randomisation was computer generated (9.4.3).

Stellungnahme Eisai GmbH (Fortsetzung)

Mohs (unclear/yes):

CSR page 34/35:...according to a computerized schedule generated (chapter 5.4.3..identical appearance, tear off label5.4.6)

Rogers (yes/unclear):

(chapter 4.1.3, CSR page 5,6):...patient randomisation schedule was computer generated, identical appearance, tear off label etc

Seltzer (unclear/unclear):

CSR page 25,28 The appearance of the medications matched (9.4.1). A computer generated randomisation schedule was generated ...Page 38 randomized in blocks of four, changed to six to reflect the 2:1 randomisation scheme

Tariot (yes/unclear):

CSR page 22,23:...patient randomisation schedule was computer generated (page 23) 'Identical appearing...tablets of donepezil-HCl and placebo were used...within a sealed, blister-packaged card' (9.4.5).

We do not understand why IQWiG states "unclear" concerning concealment (see publication page 1592).

Tune (unclear/unclear):

CSR page 26-28:...randomisation and blinding: 'The study medications were...tablets of identical appearance...within a sealed, blister-packaged card' (5.4.6). The randomisation schedule was computer generated.....

Winblad (yes/unclear):

CSR page 29,30:: Study medication placebos were of identical appearance to Aricept tablets and packed in blister packs (4.1.2). Randomisation was performed using computer generated lists, by Pfizer, Germany, randomising in blocks of four.

2. Blinding of endpoint raters

All studies are double blind and blinding was sufficient. The blinding of endpoint raters is only described for the CIBIC plus in the Burns-study.

However, for the ADAS cog- a performance based test - it is usual practice, that not the investigator but another member of the study team (psychologist, physician, trained nurse) is performing the test , mostly before other measurements are done.. Therefore we believe a formal blinding of raters is not necessary in these cases, especially as in general these persons will not have access to other study results.

3. Sample size planning

Sample size planning is relevant in designing a study. Poor sizing of the study has two potential risks: 1) the study is too small to show a treatment difference when there is a real treatment difference; 2) the study is too large when there is a real treatment difference so that more patients than necessary receive an inferior treatment. Upon completion of the study, if a treatment difference is shown then, clearly, the study was large enough.

In confirmatory studies data from previous studies using the study treatments should be used in the calculation of sample size (ICH E9, section 3.5), however a number of the studies in this report were exploratory or pilot and there were no relevant previous studies and therefore sample sizes had to be based upon clinical judgement.

Burns: (yes b: Sample size planning for CIBIC-plus, not for 2nd primary parameter ADAS-cog; later recruiting of approx. 500 patients due to high variance of the baseline values in the ADAS-cog.)

It is a legitimate way for sample size planning not to include the second primary endpoint.

Gauthier: (yes e: Sample size planning only for CIBIC-plus and not for observed sub-group (MMSE 10-17), but for the total population (MMSE 5-17) underlying the study.)

CIBIC was identified as the primary efficacy variable and therefore it was appropriate to calculate the sample size on this measure. MMSE was one of the secondary measures. It must be noted that study sample size was selected to provide adequate statistical power for the overall population rather than within each subset. (CSR S 86)

Stellungnahme Eisai GmbH (Fortsetzung)

Homma (no):

(CSR page 3)

Final global clinical rating

The result of simulations relating to the power of test with a level of significance in the U test of two-tailed 5%, assuming the distribution shown in (3) (repeated 10,000 times), gave 114 cases per group for a power of test of 80%.

Table 4. Predicted distribution of final global clinical rating

Substance	Markedly improved	Improved	Slightly improved	No change	Slightly worse	Worse	Markedly worse
E2020	0%	25%	35%	35%	5%	0%	0%
Placebo	0%	15%	25%	55%	5%	0%	0%

Table 5. Predicted final changes in ADAS-Jcog scores

Substance	Average	SD
E2020	-0.90	5.1
Placebo	1.87	5.86

Krishnan (no):

(based on clinical judgement CSR page 43) This is adequate for a pilot project.

Seltzer (no):

Page 38 CSR: Determination of the overall study sample size was estimated based on determination of sample size estimates for the primary efficacy variable (ADAS-cog) at a significance (alpha) level of 0.05. Sample sizes were further evaluated based on a review of results of the Phase III clinical trials of donepezil HCl in patients with AD who had an MMSE score ≥ 21 . Table 2 below summarizes data that were used for the calculation of sample size.

Table 2. Sample Size Estimates for ADAS-cog

Variable Name	Improvement	σ	Nd:Np (2:1)
ADAS-cog	3	4.64	60:30

The power calculation presented in Table 2 was selected to support an overall study power of 80% to show a 3-point improvement on the ADAS-cog compared to placebo. An estimated sample size of 60 donepezil patients and 30 placebo patients was determined to be needed. To account for a potential 23% dropout rate, the sample size was increased to bring the planned number of patients to 100 donepezil and 50 placebo patients, for a total of 150 patients.

Tune (no):

This is adequate for a pilot project.

CSR page 40 "The sample size for this study was determined based on a review of the results of the phase II and phase III studies of donepezil-HCL and the available literature on PET studies in Alzheimer's disease. Because this was the first use of 18 FDG-PET to evaluate potential effects of donepezil on glucose metabolism in Alzheimer's disease there was insufficient data available to perform a power analysis of sample size requirements for detecting a net benefit associated with donepezil treatment. Therefore, the sample size was determined based primarily on clinical judgement and not on statistical justification."

Winblad (unclear):

The primary efficacy variable selected to estimate sample size was the GBS scale. Assuming a standard deviation (SD) of 25 and a mean difference between study drug groups of 10 points at study end, required sample size was 100 patients per group to achieve a power of 0.8 (alpha=0.05). It was considered desirable to obtain data on 200 patients completed as per protocol requirement, and therefore reasonable to include 300 patients, as it was assumed that 30% of patients would drop out before study end

Stellungnahme Eisai GmbH (Fortsetzung)

4. Study dropouts named/reasons given

Gauthier: (yes f: Complete listing of the reasons for dropping out only for the total population of the study in the publication Feldman 2001 [53])

CSR page 86 appendix 16.2, Table 2.11

TABLE 2.1.1

PAGE 1 OF 1

DONEPEZIL PROTOCOL DON-NY-96-002-324

PREMATURE DISCONTINUATIONS

SAFETY POPULATION (MMSE >= 10)

NUMBER (%) OF SUBJECTS	Donepezil 102	Placebo 105
DISCONTINUATIONS		
Related to study drug	5 (4.9)	0
Adverse event	5 (4.9)	0
Laboratory abnormality	0	0
Special safety test	0	0
Patient died	0	0
Not related to study drug	14 (13.7)	12 (11.4)
Adverse event	4 (3.9)	5 (4.8)
Laboratory abnormality	0	0
Special safety test	0	0
Patient died	1 (1.0)	0
Protocol violation	2 (2.0)	0
Lost to follow-up	0	0
Did not meet entrance criteria	0	0
Withdrew consent	3 (2.9)	5 (4.8)
Other*	4 (3.9)	2 (1.9)
TOTAL	19 (18.6)	12 (11.4)

* Patient 00180110 (Donepezil) had a related adverse event when he withdrew.

Note: Percentages are based on the total number of subjects in the safety population (MMSE >= 10) in each treatment group.

Stellungnahme Eisai GmbH (Fortsetzung)

Rogers: (yes h Reasons for dropping out only given in part)

Reasons for dropping out are summarised in the following table 1.1 from the report:

Table 1.1
Final Termination Summary

	Placebo	E2020 (10mg)	Total
Number of Subjects	217	214	431
Completed The Entire Study	43 (19.8%)	67 (31.3%)	110 (25.5%)
Left After Meeting Attrition Criteria (By Investigator)	118 (54.4%)	87 (40.7%)	205 (47.6%)
Discontinued Prematurely	56 (25.8%)	60 (28.0%)	116 (26.9%)
Adverse Events	16 (7.4%)	23 (10.7%)	39 (9.0%)
Intercurrent Illness	0 (0.0%)	0 (0.0%)	0 (0.0%)
Request of Subject/Inv.	11 (5.1%)	14 (6.5%)	25 (5.8%)
Medication non-compliance	1 (0.5%)	2 (0.9%)	3 (0.7%)
Protocol Violation	7 (3.2%)	3 (1.4%)	10 (2.3%)
Entered Nursing Facility	2 (0.9%)	2 (0.9%)	4 (0.9%)
Other	19 (8.8%)	14 (7.9%)	33 (8.1%)

Data Source: Data Listing 19

Tariot: (yes h: Reasons for dropping out only given in part)

Summary of patient disposition given in Table B, 10.1:

Table B: Summary of Patient Disposition

	Treatment Group	
	Placebo	Donepezil HCl
Total randomized	105	103
Total completed	78(74.3%)	84(81.6%)
Total number withdrawn	27(25.7%)	19(18.4%)
Reason for withdrawal		
Related to study Drug		
Adverse event	13(12.4%)	7(6.8%)
Not related to study drug	14(13.3%)	12(11.7%)
Adverse event	6(5.7%)	4(3.9%)
Patient's or Investigator's request	2(1.9%)	2(1.9%)
Protocol violation	3(2.9%)	2(1.9%)
Other	3(2.9%)	4(3.9%)
Total Discontinuations	27(25.7%)	19(18.4%)

Source: Tables 1.1 and 2.1

Stellungnahme Eisai GmbH (Fortsetzung)

Homma: (yes h: Reasons for dropping out are only given in part, number of study dropouts is not provided broken down by groups)

CSR:

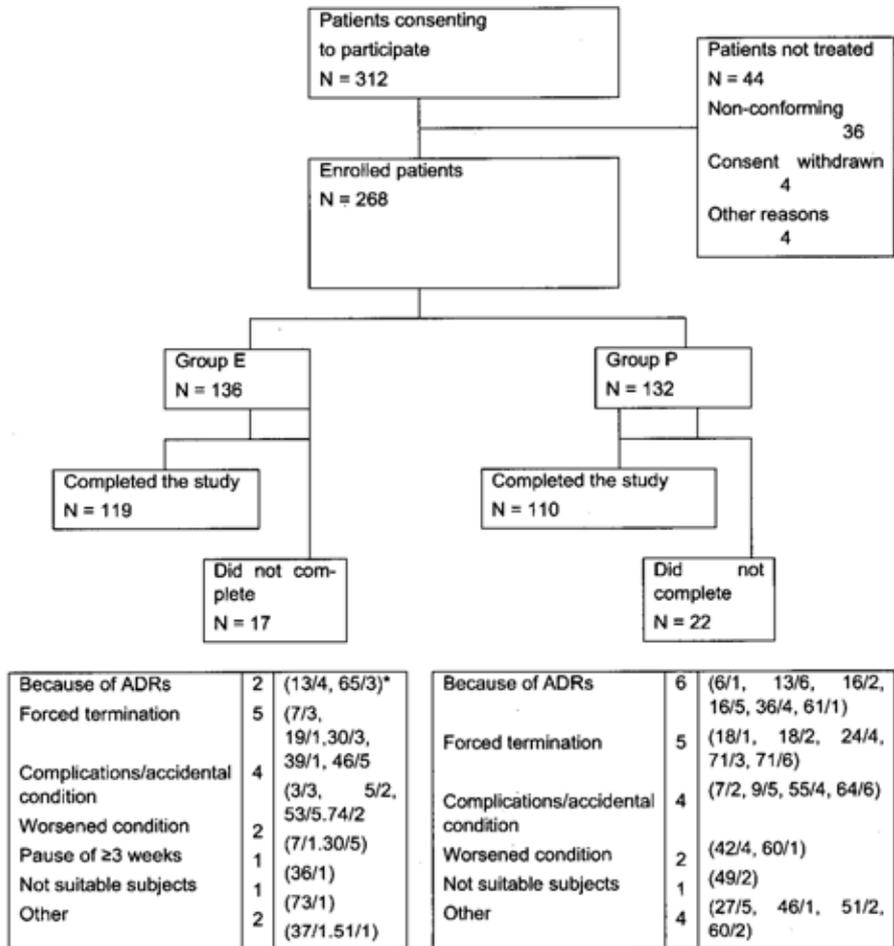


Fig. 2. Selection of patients

* Figures show set/number

Mohs: (yes j: Information on the number of patients at risk in the course of the study is missing for the main target criterion.)

CSR appendix 16.1a

App. 16.1a is shown below under chapter 5. ITT analysis

Stellungnahme Eisai GmbH (Fortsetzung)

5. ITT analysis inadequate

Burns: (unclear: Sample size missing from the ITT information. A "classic" ITT population is defined, however it is unclear how patients for whom there was no second measurement were handled)

(chapter 6.4, page 23 CSR)The ITT population is defined as '...all patients who were randomised to treatment, received at least one dose of study medications, provided complete Baseline and at least some post-Baseline assessment data.'

	Placebo	5mg	10	total
Patients enrolled	274	271	273	818
Patients baseline	272	271	272	815
Patients ADAS cog ITT endpoint analysis	264	262	254	780
Difference	10	9	19	38
%	3.65%	3.3%	7.0%	4,6%

The difference between groups is less than 4%; the total lies with 4.6 clear below 10.

Homma: (yes i: Overall, 11 % of the randomized patients were not used in the ITT LOCF analysis of the primary target criterion ADAS-cog (Donepezil group 7 %; placebo group 14 %).

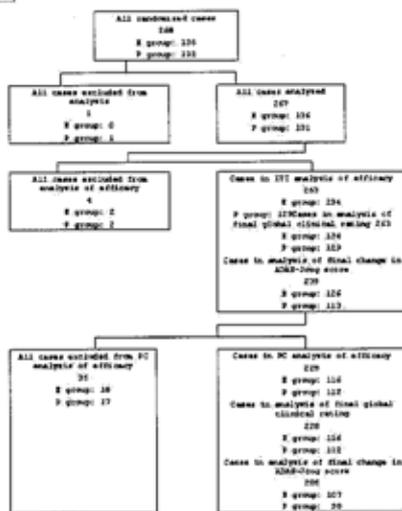


Fig. 2. Breakdown of cases for analysis

The ITT analysis was adequate. The ITT population consists of the patients who had at least one primary endpoint available. For the final global clinical rating endpoint we have a difference to randomised patients of 1.5% in the Donepezil group; and 2.3 % in the Placebo group. So the study fulfils the IQWiG criteria for an adequate ITT analysis. For the ADAS-Cog about 11% of patients were not included in the ITT population. As both primary endpoints have shown positive results, a bias concerning ADAS Cog is not reasonable to expect in this case.

Difference between groups is 7% (above the limit set by IQWiG)

Stellungnahme Eisai GmbH (Fortsetzung)

Mohs: (unclear, k: Number of patients at risk no longer stated after start of the study)

The number at risk at certain time-points is available in , App 16.1a (see below)

4% (9) of the placebo group and 3.2% (7) of the donepezil group could not be included in the ITT analysis. The analysis of the primary endpoint has been done with "time-to -event" methods (Kaplan-Meier, Logrank-test, Wilcoxon-Test, Cox-PH)

Mohs 312 App. 16.1a

Appendix 16.1a
Time to attrition by Investigator: Intent-to-treat population
Life-Table Method using protocol visits

The LIFETEST Procedure
Life Table Survival Estimates
TRT = H2020

Interval (Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	
0	43	11	7	203.5	0.0541	0.0159
43	85	20	12	183.0	0.1092	0.0231
85	127	18	9	152.5	0.1186	0.0261
127	169	9	12	124.0	0.0726	0.0213
169	211	7	7	105.5	0.0664	0.0242
211	253	9	5	82.5	0.0973	0.0302
253	295	5	2	80.0	0.0623	0.0271
295	337	5	1	73.5	0.0680	0.0284
337	379	0	49	43.5	0	0
379	.	0	19	9.5	0	0

Appendix 16.1a
Time to attrition by Investigator: Intent-to-treat population
Life-Table Method using protocol visits

The LIFETEST Procedure
Life Table Survival Estimates
TRT = Placebo

Interval (Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	
0	43	7	2	207.0	0.0338	0.0126
43	85	37	10	194.0	0.1907	0.0282
85	127	30	8	148.0	0.2027	0.0320
127	169	14	15	106.5	0.1315	0.0327
169	211	11	4	83.0	0.1325	0.0372
211	253	7	3	68.5	0.1022	0.0365
253	295	4	3	58.5	0.0684	0.0320
295	337	5	2	52.0	0.0962	0.0409
337	379	1	34	29.0	0.0345	0.0339
379	.	0	11	5.5	0	0

On request, we will make further listed sections of the individual CSRs regarding study quality available to the IQWiG in confidence.

Stellungnahme Hirnliga e. V., Möller



EINGEGANGEN 06. Okt. 2006

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

Cholinesterasehemmer bei Alzheimer Demenz
A05/19-a

Name, Vorname; Titel und Funktion des/der Stellungnehmenden:

1. Möller, Hans-Jürgen, Prof. Dr. med.
Vorsitzender der Hirnliga e.V.
- 2.
- 3.

Kontakt für weitere Nachfragen:

Name: Dr. Thomas Kunczik, Geschäftsführer
Institution: Hirnliga
Taubenweg 3
Anschrift: 51674 Wiehl
Tel./Fax-Nr.: 02262 999 99 17 / 02262 999 99 16
E-Mail-Adresse: gf@hirnliga.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik
(bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

1

Stellungnahme Hirnliga e. V., Möller (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln **ausgefüllt** vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Möller, 5.10.06
Ort/Datum

[Handwritten Signature]
Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme Hirnliga e. V., Möller (Fortsetzung)

Stellungnahme der Hirnliga e.V. zum Vorbericht (Vorläufige Nutzenbewertung, Auftrag A05/19-A) Cholinesterasehemmer bei Alzheimerdemenz v. 8.9.2006

Die Hirnliga e.V. begrüßt die Veröffentlichung des Vorberichts des IQWiG. Er leistet einen nicht unerheblichen Beitrag zur Beseitigung der in unserer ersten Stellungnahme zu den Berichtsplänen bemängelten fehlenden Transparenz.

Ergebnis und Qualität einer wissenschaftlichen Untersuchung und der daraus resultierenden Publikationen hängen ganz entscheidend von der Definition der Ausgangsbedingungen und den gewählten Untersuchungsverfahren ab. Das IQWiG definiert seine Ausgangsbedingungen und Untersuchungsmethoden in Berichtsplänen.

Das Verfahren zur Erstellung der Berichtspläne ist, im Gegensatz zu üblichem wissenschaftlichen Vorgehen, so gewählt, dass eine Diskussion der Methoden mit wissenschaftlichen Fachexperten bewusst ausgeschlossen wird.

Die so in den Berichtsplänen (A05-19A-C) festgeschriebenen Methoden spiegeln damit nicht den Stand der Wissenschaft wieder. Auch Aussagen zum Nutzen der Antidementiva leiden unter dieser methodischen Beschränkung.

Am Beispiel der Rückstufung der Bedeutung des Endpunktes „Besserung bzw. Erhalt der Lebensqualität der (betreuenden) Angehörigen“ im Amendment zu den Berichtsplänen sei dies nachfolgend dargestellt.

Nach moderner Auffassung ist die Demenz eine Erkrankung, die nicht nur die Betroffenen selbst, sondern auch ihre pflegenden Angehörigen betrifft. Die überwiegende Mehrzahl der Demenzkranken wird zu Hause von Angehörigen gepflegt. Der 4. Altenbericht beschreibt in 4.2.3.1 sehr ausführlich die Belastung der Pflegenden und deren durch die Betreuung auftretende Erkrankungen.

Die Lebensqualität der Angehörigen und deren Verbesserung ist damit ein Nutzenbeleg erster Güte. Die Rückstufung der Bedeutung des Endpunktes „Besserung bzw. Erhalt der Lebensqualität der (betreuenden) Angehörigen“ ist fehlerhaft. (s.o.). Bildet sich der Nutzen einer therapeutischen Intervention nicht nur beim Patienten selbst, sondern zugleich auch bei den pflegenden Angehörigen ab, so ist dies für die praktische Versorgungssituation bedeutsamer als der alleinige Nachweis durch Arzt-Fremdbeurteilungen beim Patienten. Es sei in diesem Kontext betont, dass bei der Arzneimittel-Zulassung die Angehörigen-Beurteilung der alltagsrelevanten Aktivitäten der Patienten eine wichtige Rolle spielt.

Vor dem Hintergrund, dass der vorgelegte Vorbericht wie am obigen Beispiel dargestellt, nicht den Stand der Wissenschaft widerspiegelt, sind aus Sicht der Hirnliga e.V. beispielhaft und nicht abschließend folgende Anmerkungen zu machen:

1. Es trifft nicht zu, dass die Alzheimerkrankheit nach MMSE und GDS in drei Schweregrade eingeteilt wird (Seite 3). Der MMSE eignet sich nur bedingt zur Schweregradeinstufung. Die GDS eignet sich zu Schweregradeinstufung, unterscheidet aber sieben Schweregrade. Die Schweregradfrage ist von

besonderer Bedeutung, wenn festgestellt werden soll, bei welchem Schweregrad von welcher Intervention etwas zu erwarten ist.

2. Die Aufteilung der Behandlungsbereiche nach Cummings (Seite 5) ist eine von vielen verschiedenen. Sie ist kein gängiger Standard der Demenzforschung. Eine Aufteilung nach den Zugangswegen medikamentös, psychotherapeutisch und umfeldstrukturierend könnte eher Klarheit schaffen.
3. Der Abschnitt „Therapieevaluation bei Alzheimer-Demenz“ (S. 6) wirkt oberflächlich, da die Messinstrumente nur oberflächlich betrachtet werden, die umfangreiche Literatur zu Stärken und Schwächen der Instrumente nicht berücksichtigt wird. Vollkommen außen vor bleibt die Frage, ob Instrumente z. B. zur Erfassung der „Quality of life“ tatsächlich erfassen, was sie zu erfassen vorgeben. Dieser Aspekt ist besonders relevant, da Forderungen nach Messung von bestimmten Kriterien damit verknüpft sind, ob überhaupt valide Instrumente zur Verfügung stehen (s.a. S. 12). In diesem Bereich sollte das IQWiG klar mehr Expertise erkennen lassen.
4. Mit der Frage, ob überhaupt valide Messinstrumente vorhanden sind, ist auch die Frage nach RCTs verknüpft (S. 13). Ohne valide Messinstrumente sind RCTs kaum vorstellbar.
5. Für mehrere Zielgrößen (S. 12) fehlt es an validen Messinstrumenten. Ohne Messinstrumente können auch keine Daten gewonnen werden, die bei einer Entscheidung zu Wirksamkeit und Nutzen hilfreich wären. Auch auf den ersten Blick sinnvolle Zielgrößen wie „Vermeidung der Notwendigkeit einer vollstationären Pflege“ relativieren sich, wenn bedacht wird, dass die Versorgungsdichte mit qualitativ hochwertigen Einrichtungen im niederschweligen und ambulanten Bereich ebenso wie mit stationären Versorgungseinrichtungen allein in einer Stadt sehr heterogen sein kann. Je nach Qualität und Verfügbarkeit kann so eine kürzere oder längere Zeit vor der vollstationären Pflege stehen. Zusätzlich wäre bei Kenntnis neuer Versorgungseinheiten wie ambulanten Hausgemeinschaften zu berücksichtigen, was eigentlich dieses scheinbar harte Kriterium definiert.
6. Das Studieneinschlusskriterium „Einschlussdauer“ ist weiterhin nicht hinreichend begründet (S. 13). Nach Auflisten zweier bereits genutzter Kriterien wird ohne Begründung ein drittes neues Kriterium eingeführt. Ein sachlich wissenschaftlicher Grund ist nicht angegeben. Es gibt in der wissenschaftlichen Literatur keinen Hinweis darauf, warum gerade 16 Wochen ein Kriterium sein sollten. Studien mit kürzerer Studiendauer sind sicherlich nicht ohne Grund durchgeführt worden. Es wird nicht deutlich, warum z. B. ein Medikament ausgeschlossen werden sollte, das nur über drei Monate geprüft worden ist.
7. Als „Vollpublikationen“ gewertete Berichte, sollten dann öffentlich gemacht werden, wenn keine Rechte anderer (z.B. Gefährdung von Patentansprüchen etc.) verletzt werden.
8. Für eine wirkliche Verfahrenstransparenz sollte der „Extraktionsbogens“ (S. 17) publiziert werden sowie die Datenexzerpte der einzelnen Studien öffentlich zugänglich sein.

Abgesehen von diesem Problem erfolgt die Analyse der eingeschlossenen Studien sachlich wissenschaftlich. Folgerichtig entspricht auch die Einschätzung der Wirksamkeit (S. 170) im Wesentlichen der Sichtweise der Zulassungsbehörden, der für die Analyse von Evidence in Rahmen der evidence based medicine bekannten

Stellungnahme Hirnliga e. V., Möller (Fortsetzung)

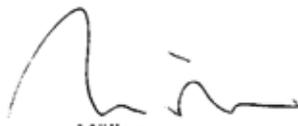
Cochrane Gruppe und der anderer wissenschaftlich und klinisch qualifizierter Fachkreise.

Wir gehen davon aus, dass das IQWiG die noch anstehenden Bewertungen therapeutischer Optionen bei Demenz unter gleichen Maßstäben bewertet.

Sehr begrüßenswert ist auch die mit gleicher methodischer Sorgfalt vorgenommene Diskussion publizierter kritischer Studien und sachlich wissenschaftlich nicht belegter Meinungsäußerungen.

Wie im Vorbericht beschrieben, führen die Cholinesterasehemmer bei einem nicht unerheblichen Teil der Behandelten zu einer Besserung von Krankheitssymptomen auch nach einem halben Jahr. Zur Bedeutung dieses Wirkungsumfanges für die Erkrankten gibt es klare Aussagen von behandelnden Ärzten, Angehörigen der Betroffenen, die ein zwar nicht heilendes, aber in dem gefundenen Umfange wirksames Medikament nicht vorenthalten bekommen wollen.

) Gezeichnet am 5.10.2006


Möller

Stellungnahme Janssen-Cilag GmbH



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:
Auftrag A05/19-A, Cholinesterasehemmer bei Alzheimer Demenz

Name, Vorname; Titel und Funktion des/der Stellungnehmenden:
1. Mehnert, Dr. Angelika; Director Health Economics & Reimbursement CNS
2. Torfs, Koen, MSc; Leiter Health Economics & Public Affairs
3. Suwelack, Dr. Klaus; Executive Business Director CNS

Kontakt für weitere Nachfragen:
Name: Dr. Angelika Mehnert
Institution: Janssen-Cilag GmbH
Anschrift: Raiffeisenstraße 8; 41470 Neuss
Tel./Fax-Nr.: Tel: 021,7/955 358; Fax: 02137/955 931
E-Mail-Adresse: amehner1@jacde.jnj.com

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:
1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

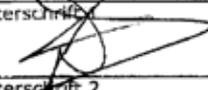
- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigelegter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Neuss, 6.10.2006



Unterschrift 1



Unterschrift 2



Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

2

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

**Stellungnahme des Unternehmens Janssen-Cilag GmbH
zum IQWiG-Vorbericht
"Cholinesterasehemmer bei Alzheimer Demenz" [Auftrag A05/19-A] vom 01.09.2006**

1. Im Vorbericht fehlende Originalstudien

1.1 Langzeitdaten (Siehe auch Punkt 3)

Der Ausschluss anderer als randomisierter kontrollierter Studien, insbesondere zu Behandlungsergebnissen, die einen Beobachtungszeitraum von 12 Monaten und länger erfordern, führt zu einer unvollständigen Bewertung und irreführenden Schlussfolgerungen, was die Evidenzbasis ("keine Daten") betrifft. Es wurde auch eine Reihe offener "Follow-up"-Studien mit Galantamin durchgeführt. Obwohl nicht randomisiert, können solche Studien bei entsprechend vorsichtiger Interpretation sehr wohl Aufschluss über die Langzeit-Effekte einer Galantamin-Behandlung geben. Mittlerweile liegen Langzeitdaten von 4 Jahren mit Galantamin vor. Janssen-Cilag verfügt auch über publizierte Ein-Jahres-Daten von mit Placebo behandelten Patienten.^{1,2} Diese Daten wurden in doppelblind-randomisierten Studien für Sabeluzole erhoben, dessen Entwicklung wegen fehlender Wirksamkeit aber eingestellt wurde. Die Studien hatten jedoch gleiche Einschlusskriterien wie die Galantamin-Studien, und die Patientenpopulationen sind daher vergleichbar. Aus diesen historischen Einjahres-Placebodaten und aus epidemiologischen Langzeit-Untersuchungen von Alzheimer-Patienten, die nicht mit Cholinesterasehemmern (ChE-H) behandelt wurden, lassen sich Rückschlüsse auf den natürlichen, unbehandelten Verlauf der Krankheit ziehen.

Der Vergleich von Placebo-Daten der Raskind 2000-Galantamin-Studie³, historischen Ein-Jahres-Placebodaten und Langzeit-Verlaufsdaten unbehandelter Patienten zeigt, dass der Verlauf in den Galantamin-Placebo-Armen über 6 Monate quasi identisch mit dem der historischen Ein-Jahres-Placebogruppe ist, und dass die 6- und 12-Monats-Verläufe der Placebo-Gruppen dem naturalistischen Einjahres-Verlauf unbehandelter Patienten entsprechen, was die Gültigkeit dieses indirekten Vergleiches unterstützt.⁴

1.1.1 Kognition

In Langzeit-Studien von 3 Jahren und länger zeigten mit Galantamin behandelte Patienten gegenüber historischen Verlaufsdaten sowohl einer unbehandelten Patientenkohorte als auch der oben genannten placebobehandelten Patientenkohorte, die beide hinsichtlich Krankheitsschwere, Alter und Geschlecht mit den Galantamin-Patienten vergleichbar waren, eine deutliche Verzögerung des Krankheitsverlaufs und einen längeren Erhalt kognitiver Fähigkeiten.^{4,5} Dabei sind die Effekte auf die Kognition nach 12, 18, 36 und 48 Monaten in allen Langzeit-Studien mit Galantamin konsistent von ähnlicher Größenordnung (vgl. auch NICE-Dossier, S. 23-26) Die Galantamin-Patienten wurden ursprünglich in Doppelblindstudien zu Galantamin randomisiert, was mögliche Verzerrungen durch eine Positivselektion einschränkt. Zusätzliche Analysen zeigten zudem, dass der Verlauf von Patienten, die die Galantamin-Behandlung vor Ablauf von 3 Jahren beendeten, bis zum Behandlungsabbruch dem der durchgängig 3 Jahre behandelten Patienten entsprach⁴, was darauf hindeutet, dass eine verzerrende "Responder-Selektion" auch im weiteren Behandlungsverlauf unwahrscheinlich ist.

Die Ergebnisse dieser vergleichenden Analysen weisen darauf hin, dass ein nach 6 Monaten erreichter Behandlungseffekt nicht nur bestehen bleibt, sondern dass der Behandlungseffekt mit der Dauer der Behandlung weiter zunimmt.

1.1.2 Alltagsaktivitäten

Ähnliche Langzeit-Ergebnisse zeigten sich auch bei den Alltagsaktivitäten: Während über ein Jahr mit Placebo behandelte Patienten deutliche Verluste von Alltagsaktivitäten aufwiesen (Differenz vs. Baseline auf der DAD-Skala: 12,4 Punkte¹), wurden solche Verluste unter Galantamin-Behandlung halbiert oder sogar ganz vermieden (Differenz vs. Baseline auf der DAD-Skala: 6,2 bzw. 1,7 Punkte³). Das entspricht dem Erhalt von 2-6 von 46 im DAD geprüfter Alltagsfähigkeiten, die unter Placebobehandlung verloren gingen.² Ergebnisse aller gepoolten 12-Monats-Daten des DAD kommen zu gleichen Ergebnissen (Differenz der gepoolten Studien: 6,1 Punkte; vgl. NICE-Dossier, S. 28)

1.1.3 Heimeinweisung

a) US-Studien: Unter Berücksichtigung von Unterschieden hinsichtlich Angebots- und Versorgungsumfeld können Daten zweier verschiedener US-Studien nützliche Informationen zu Langzeiteffekten hinsichtlich Institutionalisierung liefern. Lopez et al (2002)⁶ verglichen 135 mit ChE-H behandelte

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

Patienten mit einer gematchten Kontrollgruppe nicht mit ChE-H behandelten Patienten über bis zu 8 Jahren. Die Ergebnisse zeigen eine signifikante Reduktion des relativen Risikos einer Heimeinweisung (RR=0,33, 95% KI 0,57–0,70, p=0,004). Feldman et al. untersuchten den Effekt einer Galantamin-Behandlung auf das Risiko einer dauerhaften Heimeinweisung, indem sie Heimeinweisungsdaten aus einer Langzeit-Studie mit galantamin-behandelten Patienten den Daten einer vergleichbaren Kohorte unbehandelter Patienten gegenüberstellten. Cox-Regressionsanalysen zeigen, dass nach drei Jahren noch 86% der kontinuierlich mit Galantamin behandelten Patienten zuhause betreut werden konnten, während dies nur für 61% in der Vergleichskohorte galt.⁷ (vgl. auch "NICE-Dossier", S. 27 u. 29)

b) Internationale Studien (Europäische Länder und Kanada): Bei Patienten aus drei doppelblinden, randomisierten Studien wurde weiterhin eine retrospektive Untersuchung der Zeit bis zu einer dauerhaften Heimeinweisung durchgeführt. Hierzu wurde bei 596 Patienten, die bei Ende der Doppelblind-Studie nicht institutionalisiert waren, über einen Zeitraum bis zu 7 Jahren untersucht, ob und wann sie in ein Pflegeheim aufgenommen wurden und wie die Behandlungsgeschichte nach Ausscheiden aus den klinischen Studien verlief. Eine Cox-Regressionsanalyse, die den Effekt relevanter Baseline-Faktoren berücksichtigte, ergab, dass eine längere Behandlungsdauer das Risiko einer Einweisung signifikant senkte (hazard rate: 0,726;p<0,0001). Das bedeutet, dass ein zusätzliches Jahr mit Galantamin-Behandlung mit einer 27%-igen Senkung des relativen Risikos einer Heimeinweisung assoziiert ist.⁸ Weitere relevante Faktoren waren u.a. eine höhere Alltagskompetenz bei Baseline, die das Risiko einer Einweisung signifikant senkte (p<0,0001), während allein zu leben (d.h. keinen festen, gleichbleibenden Betreuer zu haben) das Institutionalisierungsrisiko signifikant erhöhte (p<0,0074).⁸ Wenn diese Befunde in den Kontext der AD2000-Studie (im Vorbericht zitiert) gestellt werden, zeigt sich, dass der Effekt einer kontinuierlichen Behandlung über ein Jahr (d.h. bis zur ersten Wash-out-Phase bei AD2000) auf das Risiko einer Heimeinweisung in beiden Studien ähnlich groß ist. Die größere Variabilität in der AD2000 Studie hinsichtlich (1) der Behandlungsdauer bzw. -kontinuität, (2) der geringeren Trennschärfe der Vergleichsgruppen (zweite Randomisierung nach 13 Wochen) und (3) der geringeren Patientenzahl (293 Patienten mit kontinuierlicher Behandlung über ein Jahr) führen in dieser Studie aber zu einem breitem Konfidenzintervall.

Auch wenn es sich bei allen diesen offenen Studien nicht um randomisierte kontrollierte Studien handelt, sollten sie in die Bewertung der Therapieziele mit aufgenommen werden, um nicht den Eindruck zu wecken, dass zu diesen Therapiezielen gar keine Evidenz vorläge.

1.2 Responder-Analysen (siehe auch Punkt 3.1)

Response ist ein wichtiges Beurteilungskriterium, das vor allem in der klinischen Praxis mitentscheidend ist bei der Frage der Weiterbehandlung über die Zeit und hätte mehr Platz im Vorbericht verdient. Dem IQWiG wurden im Dezember 2005 Responderanalysen von 5 doppelblinden randomisierten Studien mit Galantamin (GAL-INT-1; GAL-USA-1; GAL-USA-10; GAL-INT-10, GAL-INT-6 (AD-CVD)) zur Verfügung gestellt, die aber nicht in den Vorbericht eingingen. Die Responder-Diskussion in Kapitel 6 betrachtet nur *eine* Responderdefinition, noch dazu eine sehr konservative, die den formulierten Therapiezielen nicht gerecht wird. Für eine angemessene Betrachtung sollten die zur Verfügung gestellten Daten mit herangezogen werden.

1.3 Therapieassoziierte unerwünschte Ereignisse

In Tabelle 31 und Abbildungen 31-37 fehlen die Angaben zu Galantamin PRC (Einmal-Tägliche Gabe; 16-24 mg/Tag; N=320; Brodaty 2005) und Galantamin 16 mg/Tag (N=279; Tariot 2000). Diese relevanten und großen Behandlungsarme der beiden Studien sollten in die Analysen mit einbezogen werden, da sie mit dem zulassungsgemäßen Gebrauch von Galantamin - sowohl qua Dosis als auch qua Aufdosierungsschema - übereinstimmen und zur Vollständigkeit der Metaanalyse beitragen.

2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht

2.1 Vorbericht Kapitel 5.2.2.2: Studien- und Berichtsqualität und 5.2.4.2.1: Galantamin vs Donepezil
Ergänzende Angaben zu Studieneigenschaften, die in Tabelle 12 bzw.20 des Vorberichts als "unklar" bezeichnet werden, sowie detaillierte Auskunft zu Abbruchgründen und -zahlen für die Subgruppe AD-CVD aus Erkinjuntti (2002) (Tabellen 11 u. 12) befinden sich im Anhang.

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

Bei Wilcock 2000 werden ITT-Daten (nicht ITT-LOCF, wie auf S. 59 des Vorberichts beschrieben) berichtet: "We performed a 6 months intention to treat analysis that included all randomised patients who had any efficacy assessment, whether at baseline or during treatment." (Wilcock, 2000; S. 1446)

Bezüglich ergänzender Daten zur Subgruppe AD-CVD wurde Dr. Erkinjuntti nur nach OC und LOCF-Analysen gefragt (E-mail Dr. Lange an Dr. Erkinjuntti vom 11.08.2005). Ergänzende ITT-Daten zu dieser Studie für die AD-CVD Gruppe befinden sich ebenfalls im Anhang. Die Ergebnisse dieser "klassischen" ITT-Analysen stimmen für alle untersuchten Skalen (ADAS-cog; DAD, CIBIC-plus; NPI) mit denen der ITT-LOCF Analyse überein.

Die insgesamt geringe Anzahl von Studienabbruchern in der Subgruppe AD-CVD, die Konsistenz der ITT-Ergebnisse mit denen anderer Analysemethoden und die Auskunft über die Abbruchgründe informieren sowohl über die Qualität der Erkinjuntti-Studie (auch Subgruppe) als auch die Robustheit der Studienergebnisse für Galantamin bei Patienten mit AD-CVD.

2.2 Vorbericht Kapitel 5.3.6.4 Dosierungen der Cholinesterasehemmer und Kapitel 7 Fazit

Der Satz "Galantamin war in Dosierungen von 16 und 24 mg wie oben beschrieben vergleichbar wirksam, die Dosierung von 8 mg war nicht wirksam" in Punkt 5.3.6.4 sollte folgendermaßen modifiziert werden: "Galantamin war in den beiden zugelassenen Erhaltungsdosierungen von 16 und 24 mg/Tag wie oben beschrieben vergleichbar wirksam, die Anfangsdosis von 8 mg/Tag war nicht wirksam". Begründung: Gemäß Fachinformation⁹ ist die Tagesdosis von 8 mg ausdrücklich nur eine Anfangsdosis während der ersten 4 Behandlungswochen. Für die im Vorbericht formulierten Fragestellungen sind aber nur die Erhaltungsdosierungen, gemäß Fachinformation also nur 16 mg/Tag und 24 mg/Tag, relevant. Entsprechend sollte auch das Fazit (Kapitel 7), 1. Absatz, letzter Satz umformuliert werden: statt "... für Galantamin und Rivastigmin nur bei mittleren und hohen Dosen" also "...für Galantamin für die gemäß Fachinformation empfohlenen Erhaltungsdosierungen, ...". Auch die UEs werden im Vorbericht schließlich nur für Erhaltungsdosierungen, aber nicht für 8 mg/Tag untersucht. Ergänzend weisen wir darauf hin, dass in einer post-hoc Analyse von Patienten mit mittelschwerer Alzheimerdemenz (MMSE 10-18) Patienten unter 24 mg/Tag bessere Wirkung auf ADAS-cog zeigten als Patienten unter 16 mg/Tag Galantamin (Differenz: 1,83 Punkte; p=0,046), was auf eine Dosis-Wirkungs-Beziehung in dieser Population hindeutet.¹⁰

2.3 Vorbericht Kapitel 5.3.2.7 Therapieassoziierte unerwünschte Ereignisse

Die Analyse weist darauf hin, dass erhöhte Abbruchraten wegen unerwünschter Ereignisse hauptsächlich in Studien mit *schneller Aufdosierung* beobachtet wurden (Vorbericht S. 118, 1. Absatz). In Absatz 2 auf Seite 118 werden sodann die Raten für Übelkeit und Erbrechen gerade derjenigen Studien hervorgehoben, die ein Ein-Wochen-Aufdosierungsschema einsetzten, das *in der Fachinformation⁹ nicht empfohlen* wird. Für eine praxisrelevante Darstellung sollte aber das Augenmerk vor allem auf die Studien gelenkt werden, in denen die Aufdosierung *gemäß der Fachinformation* erfolgte. Eine schnelle Aufdosierung ist nicht in Übereinstimmung mit der Fachinformation, die *ausdrücklich eine langsame Aufdosierung* in 4-Wochen-Schritten empfiehlt. Es sollte daher der 1. Absatz durch die Aussage ergänzt werden, dass in den Studien, in denen die Aufdosierung gemäß der Fachinformation erfolgte, die Abbruchraten aufgrund unerwünschter Ereignisse nur bei 7-10% lagen (Placebo: 3-7%). Auch der Satz im 2. Absatz "In den Studien, in denen pro Woche um 8 mg aufdosiert wurde (Raskind 2000, Wilcock 2000), litt ca. ein Drittel der Patienten in der Verum-Gruppe unter Übelkeit und ein Fünftel unter Erbrechen (Tabelle 22)." sollte daher wie folgt modifiziert oder ergänzt werden: "In den Studien, die das in der Fachinformation empfohlene 4-Wochen-Aufdosierungsschema einsetzten (Tariot 2001, Brodaty 2005; Rockwood 2006), traten bei ca. einem Sechstel der Patienten Übelkeit und bei ca. einem Zehntel Erbrechen auf (Tabelle 31)." (Druckfehler bei Tabellennummer auf S. 118: dort irrtümlich Tabelle 22).

Weiterhin sollten in die Analysen zu unerwünschten Ereignissen (Abbildungen 31-37 und Tabelle 31) die Daten zu Galantamin PRC (Einmal-Tägliche-Gabe) und zu Tariot 2000, 16 mg/Tag, mit eingeschlossen werden (s.o., Punkt 1.4).

Tabelle 30: Ergänzende Informationen zu Tab. 30 befinden sich im Anhang. Die Prozentzahlen der Studienabbrucher bei Rockwood 2006 befinden sich in dessen Publikation auf S. 1102, linke Spalte, letzte Zeile, und brauchen nicht berechnet zu werden (vgl. Erläuterung i zu Tab. 30).

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

3. Anmerkungen zur projektspezifischen Methodik**3.1 Responder-Analysen**

In einer chronisch progredienten Krankheit wie Alzheimerdemenz stellt bereits der Erhalt des Gesundheits-Status quo einen Therapieerfolg dar. Dies spiegelt sich folgerichtig in den im Vorbericht definierten Therapiezielen wider: "Besserung bzw. Verhinderung der Beeinträchtigung der Aktivitäten des täglichen Lebens", "Besserung bzw. Erhalt der kognitiven Leistungsfähigkeit", "Besserung bzw. Erhalt des klinischen Krankheitsstadiums" u.ä. Eine Beschränkung der in Kap. 6 diskutierten Responderanalysen auf ein nur auf eine Verbesserung abzielendes Response-Kriterium "Verbesserung auf ADAS-cog um mindestens 4 Punkte" ist daher nicht geeignet, das im Vorbericht formulierte Therapieziel und die darauf bezogene Wirksamkeit der Präparate bzw. den Anteil der von der ChE-Hemmer-Behandlung profitierenden Patienten angemessen abzubilden. Unter Zugrundelegen einer Response-Definition, die die im Vorbericht definierten Therapieziele adäquat abbildet (s.u.), wurden Responderanalysen von 5 doppelblinden randomisierten Studien mit Galantamin durchgeführt und dem IQWiG am 2.12.2005 zur Verfügung gestellt.

*"The patient was stable or improved (change ≤ 0) on cognition as measured by the ADAS-Cog 11 scale vs. baseline assessment AND was stable /improved vs. baseline assessment on EITHER the global measure (CIBIC+ or CIGC) OR functional measure (ADCS-ADL or DAD) OR behavioural measure (NPI where collected)."*¹¹

Die so definierte Responder-Rate für Galantamin-Patienten betrug 59% (789/1335, einschließlich Patienten, die wegen mangelnder Wirksamkeit die Studie abgebrochen hatten). Die entsprechende Rate unter Placebo war 37% (334/898). Der mittlere Behandlungseffekt (Differenz der Mittelwerte der ADAS-cog-Änderung zwischen Galantamin-Respondern und Placebo-Patienten) nach 5-6 Monaten betrug 6,5 Punkte und war damit deutlich größer als eine Standardabweichung in der GAL-Responder-Gruppe (SD: 3,89) oder der Placebo-Gruppe (SD: 6,06).¹²

Langzeit-Daten (siehe dazu auch Punkt 3.2) zeigen darüber hinaus, dass der mittlere ADAS-cog-Score der GAL-Responder erst nach 18 Monaten wieder Baseline-Niveau erreichte. Weiterhin verlief die beobachtete ADAS-cog-Verschlechterung nach den ersten 18 Monaten langsamer als es dem geschätzten Verlauf dieser Patienten bei keiner Behandlung entsprochen hätte. Der Verlauf der als "non-Responder" klassifizierten Patienten entsprach eher der für sie geschätzten Verschlechterung ohne Behandlung, obwohl auch hier ein ansehnlicher Anteil den für sie erwartete ADAS-cog-Anstieg um eine Standardabweichung und mehr unterschritten (d.h. sich deutlich weniger verschlechterten als prognostiziert).¹³

3.2 Vorbericht Kapitel 4.1.4: Ausschluss nicht randomisierter Studien mit der Begründung "Für alle (...) genannten Interventionen und alle (...) genannten Therapieziele ist eine Evaluation im Rahmen von randomisierten, kontrollierten Studien möglich und praktisch durchführbar."

Evaluationen insbesondere von Endpunkten, die eine sehr lange Beobachtungsdauer erfordern (z.B. Zeit, bis Vollzeitpflege erforderlich ist; Langzeiteffekte), sind in einem randomisierten, kontrollierten Design versus Placebo nicht durchzuführen: das Vorenthalten einer wirksamen Therapie über längere Zeit ist schon aus ethischen Gründen nicht verantwortbar. Auch aus praktischen Gründen sind solche Studien nicht (mehr) durchführbar: 1) Die Verfügbarkeit wirksamer Medikamente erschwert nicht nur die Rekrutierung von Patienten in eine Placebo-kontrollierte Studie, sondern führt auch zu hohen Abbruchraten von Patienten, die eine sichere wirksame Behandlung einer (bei gleichmäßiger Randomisierung: 50%-) Wahrscheinlichkeit einer Scheinbehandlung vorziehen. Die um über 80% unter Plan liegende Anzahl rekrutierter Patienten (565 vs 3000 geplant) und die hohen Abbruchraten der einzigen Langzeitstudie (>12 Monate), in der dieses versucht wurde, AD 2000, bestätigen dies eindrucksvoll. 2) Die Aussagekraft von Wirksamkeitsparametern wie Institutionalisierung ist nicht nur kritisch zu bewerten, ein solcher Endpunkt erfordert auch extrem hohe Patientenzahlen für eine angemessen gepowerte Studie.¹⁴ Die praktische Durchführbarkeit solcher Studien ist daher zusätzlich zu den angeführten ethischen, Rekrutierungs- und Retentions-Gründen (s.o.) aufgrund der geforderten Patientenzahlen äußerst fraglich.

¹¹ Zur Adäquanz eines Vergleiches von Verum-Responder-Scores zu Gesamt-Placebo-Scores (und zur Unangemessenheit (Bias) anderer Vergleiche) siehe Report von Matthews F., Review of memantine submission, and detailed investigation of submissions for donepezil, rivastigmine and galantamine including new analyses, Final version 06.01.2006, Seite 6. Dieser Report wurde im Auftrag des NICE vom Medical Research Council, einer unabhängigen Forschungseinrichtung, erstellt und ist veröffentlicht auf der NICE-Webseite: <http://www.nice.org.uk/page.aspx?o=288664> Die Responderdefinition wurde ebenfalls von dieser unabhängigen Einrichtung überprüft: es zeigten sich keine Hinweise auf Heterogenität zwischen den Studien und kein verwässerter Effekt auf ADAS-cog gegenüber einer engeren Responderdefinition (ebd.).

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

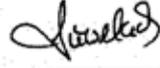
Eine umfassende Nutzenbewertung sollte daher die gesamte Datenlage umfassen und vorhandene Evidenz nicht auf Grund formalistischer Vorgaben von vornherein von der Betrachtung ausschließen.

Im Gegensatz zum vom IQWiG verfolgten Ansatz wird z.B. vom Scottish Intercollegiate Guidelines Network Grading Review Group, SIGN (deren Richtlinien zur Evidenzbewertung im IQWiG-Methodenpapier aufgegriffen werden) bei der Erstellung evidenzbasierter Guidelines zur Alzheimer-Therapie¹⁵ nicht-randomisierte Evidenz sehr wohl betrachtet: "At the case control study level, there is support for long term use of cholinesterase inhibitors to delay institutionalisation", mit Referenz zu der case-control-Studie von Lopez et al, 2002⁶.

Daneben ist die *theoretische* Durchführbarkeit einer Studienform (die aber praktisch nicht durchgeführt wurde) kein logisches Argument, andere Studienformen, die *tatsächlich* durchgeführt wurden, auszuschließen.

Insgesamt sind die im Vorbericht ausgewerteten Studien zu Galantamin sorgfältig und kenntnisreich recherchiert und analysiert worden. Die gewählte Methodik, die auf der ausschließlichen Analyse kontrollierter randomisierter Studien beruht, wird jedoch der Erkrankung und dem Reichtum der Datenlage nicht gerecht, da dadurch viel Evidenz unberücksichtigt und unerwähnt blieb und damit nicht im Sinne eines Patientennutzens gewürdigt wurde.

¹ Feldman HH, Van Baelen B, Kavanagh SM, Torfs KEL, Cognition, Function, and Caregiving Time Patterns in Patients With Mild-to-Moderate Alzheimer Disease. A 12-Month Analysis. *Alzheimer Dis Assoc Disord* 2005;19:29-36.
² Feldman H, Sauter A, Donald A, Gélinas I, Gauthier S, Torfs K, Parys W, Mehnert A. The Disability Assessment for Dementia scale: a 12-month study of functional ability in mild to moderate severity Alzheimer disease. *Alz Dis Assoc Disord* 2001;15:89-95.
³ Raskind MA, Peskind ER, Wessel T, Yuan W. Galantamine in AD: a 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology* 2000;54:2261-8.
⁴ Raskind MA, Peskind ER, Truyen L, Kershaw P, Damaraju CV. The cognitive benefits of galantamine are sustained for at least 36 months: a long-term extension trial. *Arch Neurol* 2004;61:252-6.
⁵ Pirttilä T, Wilcock G, Truyen L, Damaraju V. Long-term efficacy and safety of galantamine in patients with mild-to-moderate Alzheimer's disease: multicenter trial; *Eur J Neurol* 2004;11:734-741.
⁶ Lopez OL, Becker JT, Wisniewski S, Saxton J, Kaufer DI and DeKosky ST. Cholinesterase inhibitor treatment alters the natural history of Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2002;72:310-314.
⁷ Feldman H, Van Baelen B, Kavanagh S. (2004a) Admission to nursing home: evidence from US studies of galantamine. *Poster, 8th International Montreal/Springfield Symposium on Advances in Alzheimer Therapy*, Montreal, Canada, 2004.
⁸ Feldman H, Kavanagh S, Brashear HR, Van Baelen B. Effect of galantamine on time to residential or nursing home admission. *Poster, 8th European Federation of Neurological Societies*, Paris, France, Sept. 4-7, 2004.
⁹ Janssen-Cilag, Fachinformation REMINYL[®] 1 x täglich 8 mg/-16 mg/-24 mg Hartkapseln retardiert, Stand: Sept. 2005.
¹⁰ Schwalen S, Van Baelen B, Kavanagh S. Comparison of the effects of 24mg versus 16mg galantamine in patients with moderate Alzheimer's disease. *Poster, EFNS, Athens, Greece, 17-20th September 2005*.
¹¹ Shire Pharmaceuticals Limited and Johnson & Johnson Pharmaceutical Research & Development, L.L.C, Response to the National Institute for Health and Clinical Excellence (NICE) Request for Additional Analyses of July 2005: Additional Analyses using Studies of Galantamine in Subjects with Dementia of the Alzheimer's Type, 21 October 2005. Datei "1. Galantamine submission of additional evidence to NICE 21Oct05.zip" per E-Mail an IQWiG 2.12.2005.
¹² vgl. (11); Appendix 2, Tabelle 2.9.2. Enthalten in Datei "5. GAL NICE OCT 05 Appendix 2 RTF tables.doc" in "Appendices without SAS files.zip", per E-Mail an IQWiG 2.12.2005.
¹³ vgl. (11); Appendix 3, Figure 9 und 10. Enthalten in Datei "5. GAL NICE OCT 05 Appendix 3 LT graphs 1 to16_BW3.doc" in "Appendices without SAS files.zip", per E-Mail an IQWiG 2.12.2005.
¹⁴ Thomas S und Windeler J, Methodische Kritikpunkte zur Studie "Long-term donepezil treatment in 565 patients with Alzheimer's disease (AS2000): randomized double blind trial", 14.1.2005. http://www.mds-ev.org/download/AD2000_Methodenkritik-Endfassung%2014-01-05.pdf#search=%22MDS%20AD2000%22
¹⁵ Scottish Intercollegiate Guidelines Network Grading Review Group, Management of patients with dementia. A national clinical guideline; Edinburgh, Februar 2006. URL: <http://www.sign.ac.uk/pdf/sign86.pdf>

		
Dr. Angelika Mehnert	Koen Torfs	Dr. Klaus Suwelack

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

**Anlagen zur Stellungnahme des Unternehmens Janssen-Cilag GmbH
zum IQWiG-Vorbericht
"Cholinesterasehemmer bei Alzheimer Demenz" [Auftrag A05/19-A] vom 01.09.2006**

1. Ergänzende Daten zu Kapitel 5.2.2.2: Studien- und Berichtsqualität und Tabelle 12, und Kap. 5.3.2 Galantamin

1.1. Erkinjuntti 2002: "grobe Mängel"

Ergänzend zu den auf IQWiG-Anfrage gelieferten Daten für die Subgruppe AD-CVD der Studie GAL-INT-6 (Erkinjuntti 2002) stellen wir die folgenden Daten zur Verfügung, um etwaige Zweifel an der Studienqualität auszuräumen.

(a) Studienabbrecher genannt / Abbruchgründe genannt

Information hinsichtlich Abbruchgründen und -zahlen für die Subgruppe AD-CVD befinden sich in Tabelle 1 unten. Wie zu erwarten, bezog sich der größte Anteil der Abbrüche in beiden Studienarmen auf unerwünschte Ereignisse, überwiegend und für die Substanzklasse typisch im gastrointestinalen Bereich (vgl. auch Erkinjuntti 2002, Tabelle 4, S. 1088). Die höheren Abbruchraten spiegeln das schnelle (1-Wochen-) Aufdosierungsschema wider, das in den früheren Studien eingesetzt wurde, aber nicht dem in der Galantamin-Fachinformation empfohlenen 4-Wochen-Schema entspricht.

Tabelle 1: GAL-INT-6 Subgruppe AD-CVD: Abbruchgründe

Reason for discontinuation	Galantamine N (%)	Placebo N (%)	All patients N (%)
Adverse Event	25 (13)	5 (5)	30 (11)
Death	2 (1)	1 (1)	3 (1)
Insufficient response	1 (<1)	0 (0)	1 (<1)
Subject lost to follow up	0 (0)	1 (1)	1 (<1)
Subject non-compliant	2 (1)	1 (1)	3 (1)
Subject withdrew consent	2 (1)	1 (1)	3 (1)
Other	2 (1)	1 (1)	3 (1)
Completed	154 (82)	87 (90)	241 (85)

* Hand-searching of verbatims showed reasons recorded as follows "Caregiver refused to bring patient in for visits", "Patient went to nursing home" and "The patient refused"
Source: Janssen Pharmaceutica analyses in response to IQWiG data on file

(b) Abweichung vom ITT-Prinzip

(i) Die in der Publikation berichteten Daten beruhen auf ITT-LOCF Analyse, die voraussetzt, dass zumindest eine Nach-Baseline Wirksamkeitsbewertung zur Berechnung des Behandlungseffekts vorliegt. Es wurde aber alternativ auch eine "klassische" ITT- Analyse durchgeführt, in der für Patienten, die die Studie vor der ersten Nach-Baseline-Bewertung verlassen hatten, auch Baselinewerte fortgeschrieben wurden.

Die Ergebnisse dieser Analysen für Galantamin vs. Placebo für ADAS-Cog, DAD, NPI und CIBIC-plus stimmen mit denen der ITT-LOCF Analysen überein (Tabellen 2, 3, 4 und 5 unten). Die Analyse der Abbruchgründe, aus der hervorgeht, dass nur eine sehr geringe Anzahl von Patienten die Studie wegen mangelnder Wirksamkeit vorzeitig verlassen haben, unterstützt diese Übereinstimmung.

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

Tabelle 3: GAL-INT-6: DAD in Subgruppe AD-CVD

Table 6-20: Total DAD Score: Mean and Mean Change From Baseline at Month 6 by Diagnosis

Diagnosis (dementia) Analysis	Placebo			GAL 12 mg b.i.d.			GAL minus placebo		
	n	Mean ± SE	Mean change ± SE	n	Mean ± SE	Mean change ± SE	n	LS means (95% CI)	LS means (95% CI)
Mixed Observed	86	60.6 ± 2.78	-6.4 ± 1.58	153	70.2 ± 1.97	-0.6 ± 1.23	61	(2.05, 10.10)	6.1
LOCF	93	59.6 ± 2.71	-6.0 ± 1.47	172	69.5 ± 1.89	-1.0 ± 1.15	4.9	(1.17, 8.72)	4.9
CITT	96	59.8 ± 2.65	-5.8 ± 1.43	183	68.6 ± 1.86	-0.9 ± 1.08	4.9	(1.28, 8.45)	4.9

LOCF = last observation carried forward; CITT = classical intent to treat; LS = least square means based on an ANOVA model including treatment and country as factors.

Source: Display.EFF.DAD.SA and EFF.DAD.SB.

Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 91

Tabelle 2: GAL-INT-6: ADAS-cog in Subgruppe AD-CVD

Table 6-13: ADAS-cog/11 Mean Scores and Mean Changes From Baseline at Month 6 by Diagnosis

Diagnosis (dementia) Analysis	Placebo			GAL 12 mg b.i.d.			GAL minus placebo		
	n	Mean ± SE	Mean change ± SE	n	Mean ± SE	Mean change ± SE	n	LS means (95% CI)	p-value*
Mixed Observed	87	25.7 ± 1.3	1.8 ± 0.60	152	21.4 ± 0.86	-1.0 ± 0.46	2.7	(-4.16, -1.17)	p<0.001
LOCF	96	26.5 ± 1.26	1.7 ± 0.55	177	22.1 ± 0.79	-0.8 ± 0.43	2.3	(-3.75, -0.93)	p=0.001
CITT	96	26.5 ± 1.26	1.7 ± 0.55	182	22.0 ± 0.77	-0.7 ± 0.42	2.3	(-3.74, -0.96)	p<0.001

LOCF = last observation carried forward; CITT = classical intent to treat; LS = least square means based on an ANOVA model including treatment and country as factors.

* p-values for pairwise comparison with placebo based on a 2-way ANOVA model.

Source: Display.EFF.ADAS.32B and EFF.ADAS.32C.

Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 83

Tabelle 4: GAL-INT-6: CIBIC-plus in Subgruppe AD-CVD: ITT-Analyse

Table 6-14: Distribution of CIBIC-plus Scores in the Mixed Dementia Subgroup at Month 6: Observed Case

CIBIC-plus scores*	Placebo n=84			GAL 12 mg b.i.d. n=155		
	n (%)	Cumulative %	n (%)	Cumulative %	p-value*	
Markedly improved	1 (1.2%)	1.2%	4 (2.6%)	2.6%	p<0.001	
Moderately improved	2 (2.4%)	3.6%	12 (7.7%)	10.3%		
Minimally improved	13 (15.5%)	19.0%	33 (21.3%)	31.6%		
No change	29 (34.5%)	53.6%	67 (43.2%)	74.8%		
Minimally worsened	29 (34.5%)	88.1%	27 (17.4%)	92.3%		
Moderately worsened	10 (11.9%)	100%	10 (6.5%)	98.7%		
Markedly worsened	0 (0.0%)	100%	2 (1.3%)	100%		

* 7-point rating scale

* p-value for comparison with placebo.

Source: Display.EFF.CIB.16.

Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 85

Tabelle 5: GAL-INT-6: NPI in subgruppe AD-CVD

Table 6-21: Total NPI Score: Mean and Mean Change From Baseline at Month 6 by Diagnosis

Diagnosis (dementia) Analysis	Placebo			GAL 12 mg b.i.d.			GAL minus placebo		
	n	Mean ± SE	Mean change ± SE	n	Mean ± SE	Mean change ± SE	n	LS means (95% CI)	LS means (95% CI)
Mixed Observed	83	11.6 ± 1.53	0.7 ± 1.22	148	9.6 ± 1.02	-0.9 ± 0.76	-2.2	(-4.88, 0.56)	-2.2
LOCF	90	12.1 ± 1.48	0.8 ± 1.14	167	10.9 ± 1.11	-0.2 ± 0.76	-1.3	(-3.88, 1.33)	-1.3
CITT	94	12.0 ± 1.43	0.5 ± 1.11	179	10.8 ± 1.05	-0.2 ± 0.71	-1.0	(-3.45, 1.51)	-1.0

LOCF = last observation carried forward; CITT = classical intent to treat; LS = least square means based on an ANOVA model including treatment and country as factors.

Source: Display.EFF.NPI.SA and EFF.NPI.SB.

Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S.93

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

(c) Erkinjuntti 2002: Fallzahlplanung

3.6.1. Determination of Sample Size

The primary efficacy parameters were the change from baseline ADAS-cog/11 and the CIBIC-plus, at Month 6 (observed case) of the double-blind phase. Since both parameters needed to be significant to have a positive result, no adjustment for multiplicity was needed.

Previous studies conducted by the sponsor (GAL-INT-1) suggested that placebo-treated patients exhibit deterioration at Month 6, i.e., an increase of 2.4 (standard deviation [SD] 5.8) points on the ADAS-cog/11 and an improvement (decrease) of -0.7 (SD 6.0) points for galantamine-treated patients (12 mg b.i.d). This difference of 3.1 points in ADAS-cog/11 score between placebo- and galantamine-treated patients is clinically meaningful.

With 90% power and a 2-sided significance level of 0.05 the required number of observations for the ADAS-cog/11 change at Month 6 is 120 for galantamine 12 mg b.i.d. and 60 for the placebo group. Assuming a discontinuation rate of approximately 20%. 75 patients in the placebo group and 150 in the galantamine group were to be included.

The CIBIC-plus score at Month 6 was observed to have the following distribution (in GAL-INT-1): markedly improved 0 % (placebo) versus 0 % (galantamine), moderately improved 1% (placebo) versus 4 % (galantamine), minimally improved 17% (placebo) versus 17% (galantamine), no change 32 % (placebo) versus 47% (galantamine), minimally worse 33% (placebo) versus 27% (galantamine), moderately worse 16% (placebo) versus 4% (galantamine) and markedly worse 1 % (placebo) versus 2 % (galantamine). Simulation revealed that 140 patients treated with placebo and 280 patients treated with galantamine were required to detect a significant result at the 5% significance level with 90% power.

To account for a 20% discontinuation rate, 175 patients were to be recruited for placebo treatment and 350 patients for galantamine treatment. A total of 525 patients were to be recruited. With this number of patients, both comparisons could be performed at the 5% significance level with an overall power of at least 80%.

(Quelle: Janssen Research Foundation, Statistical Analysis plan GAL-INT-6, S. 49-50)

1.2. Brodaty: Fallzahlplanung

3.11.2.2. Sample Size Determination

The calculations of sample size and statistical power were based on the primary hypotheses to demonstrate treatment effects on change from baseline in ADAS-cog/11 and the CIBIC-plus scores at Week 26. The power and sample size calculations of this study were based on two-sided tests with a significance level of 5%. With an anticipated 230 subjects per group (a total of 690 subjects) completing the study, and with approximately 55% subjects in the placebo group with "no change" or "improved" CIBIC-plus score, the study was estimated to have approximately 90% power to detect a 15% difference between galantamine CR and placebo and >95% statistical power to detect a mean difference of 2.5 points (with SD = 6.2) in change from baseline at Week 26 in ADAS-cog/11 score between the placebo and galantamine CR groups.

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

This study had an estimated 90% and >95% power for the tests for CIBIC-plus and change in ADAS-cog/11 scores, respectively. Thus, the overall global power for testing the hypotheses across the dual endpoints was approximately 90%. With the assumption of a 22% dropout rate across groups during the double-blind treatment period, the study would need to randomize 295 subjects to each group (i.e., a total of 885 subjects in the study). Since the study needed to show statistical significance in both primary efficacy endpoints simultaneously, both endpoints were tested at the 5% significance level without any adjustment for the multiple primary endpoints. Since there was only 1 primary between-treatment comparison in this study, i.e., the comparison between galantamine CR and placebo groups for each primary efficacy measure, there was no need for the adjustment for multiple comparisons. Other comparisons were exploratory.

(Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-10, S. 49-50)

1.3 Tariot: Verblindung

3.2.3.1.1. Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog/11)

(...)

To eliminate the risk of breaking the blind, the ADAS rater was not to be involved in the patient's treatment, nor was he or she to have access to the record of reported adverse events. If this was not possible, then the ADAS rater performed the ADAS test before he or she asked the patient about adverse events.

3.2.3.1.2. Clinician's Interview-Based Impression of Change - Plus Caregiver Input (CIBIC-plus)

The score for the Clinician's Interview-Based Impression of Change - Plus Caregiver Input (CIBIC-plus) was a second primary dependent variable. An independent, experienced, and properly trained clinician provided a global impression of the patient's deterioration or improvement over the course of the trial, based on separate interviews with the patient and caregivers. For reference purposes, the CIBIC rater had the option of video or audiotaping the baseline interview. The CIBIC rater elicited from the caregiver only factual information about the patient's functioning and refrained from asking the caregiver's opinion about any possible change in the patient's condition.

The CIBIC-plus was performed at Visits 2, 3, 4, and 5 (baseline, Week 4, Week 13, and Month 5 or termination). To ensure that only trained CIBIC raters performed the testing, the name and signature of the CIBIC rater was documented at the time of each test. The scoring system for the CIBIC-plus assessment is provided in Annex 1.

(Quelle: Janssen Research Foundation, Clinical Research Report GAL-USA-10; S. 45 u. 46)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

1.4. Concealment in Studien GAL-INT-6, GAL-USA-1, GAL-USA-10
Tabelle 6: Concealment in Studien GAL-INT-6, GAL-USA-1, GAL-USA-10

Trial	Methods of allocation concealment			
	Central randomization scheme & sequential allocation of envelopes and medication	Similarity of appearance of dispensed medication	Handling of randomization allocation information	Other
GAL-INT-6 ¹	<p>Central randomization scheme & sequential allocation of envelopes and medication</p> <p>“Subjects admitted to the trial will be randomly allocated to one of the 2 treatment groups according to a randomization code generated [centrally] by the Janssen Research Foundation. In each centre 2 patients will be randomised to receive trial drug for each 1 patient randomised to receive placebo. In each centre, subject numbers will be assigned consecutively starting with the lowest number available and at least 6 subjects should be entered.”</p>	<p>Similarity of appearance of dispensed medication</p> <p>“The trial medication will be provided by the Janssen Research Foundation (JRF). The treatment will consist of tablets which will be identical in appearance, taste, and smell. The tablets will contain 4mg, 8mg, 12mg of galantamine or no active ingredient.”</p>	<p>Handling of randomization allocation information</p> <p>“The investigator will be provided with a blinded code for each subject, containing details of the treatment in the titration and double blind phase. The code should only be broken in case of an emergency where the further treatment of the subject is dependent on the trial medication he or she has been receiving. The date and time of the code breaking as well as the reason for it must be recorded on the Code Breaking Form in the CRF. In addition, the reason must be documented on the Adverse Event Form of the CRF. The trial monitor should be notified immediately. If the code is broken the subject must be withdrawn from the trial and must be followed up if appropriate. All envelopes, whether open or sealed will be collected at the end of the trial.”</p>	<p>Other</p> <p>“The investigator, his/her designee or a hospital pharmacist must maintain an adequate record of the receipt and distribution of all trial supplies using the Drug Accountability Form which must be available for inspection at any time.”</p>
GAL-USA-1 ²	<p>“Subjects admitted to the trial will be randomly allocated to one of the 3 treatment groups according to a randomization code generated [centrally] by the Janssen Research Foundation. Balancing will ensure that in each centre, subject numbers will be assigned consecutively starting with the lowest number available and at least 6 subjects should be entered.”</p>	<p>“Trial medication was provided by the Janssen Research Foundation (JRF). The medication was formulated in tablets that were identical in appearance, taste, and smell, and which contained either no active ingredient or 4mg, 8mg, 12mg, or 16mg galantamine.”</p>	<p>“The investigator will be provided with a blinded code for each subject, containing details of the treatment in the titration and double blind phase. The code should only be broken in case of an emergency where the further treatment of the subject is dependent on the trial medication he or she has been receiving. The reason must be documented on the Adverse Event Form of the CRF. The trial monitor should be notified immediately. If the code is broken the subject must be withdrawn from the trial and must be followed up if appropriate. All codes whether open or sealed will be collected at the end of the trial.”</p>	<p>“The investigator, or a hospital pharmacist must maintain an adequate record of the receipt and distribution of all trial supplies using the Drug Dispensing Form and the Drug Return Form. These forms must be available for inspection at any time.”</p>
GAL-USA-10 ³	<p>“Patients admitted to the trial were randomly allocated to one of the treatment groups according to a randomization code generated [centrally] by the JRF. The patient allocation ratios to the four treatment groups (Placebo, GAL 4mg bid, GAL 8mg bid and GAL 12mg bid) were 2:1:2:2... Balancing ensured that, at each site, equal numbers of patients entered each treatment group as required. At each site patient numbers were assigned consecutively starting with the lowest number available, and at least 6 subjects were to be entered.”</p>	<p>“The Janssen Research Foundation (JRF) provided all trial medication. The trial medication was formulated in tablets that were identical in appearance, taste, and smell and which contained no active ingredient or 4mg, 8mg or 12mg of galantamine.”</p>	<p>“For each patient, the investigator was provided with a blinded code containing details of the treatment in both the escalation and double blind phases of the trial. This code could be broken only in case of an emergency where further treatment of the patient depended on knowledge of the trial medication that the patient had received. The date and time of the code breaking as well as the reason had to be recorded on the Code Breaking Form in the CRF. Additionally, the reason for breaking the code had to be documented on the Adverse Event Form of the CRF. The trial monitor was to be notified immediately. The patient was to be withdrawn from trial medication and would continue to be followed.”</p>	<p>“The investigator acknowledges that the drug supplies are investigational and as such must be handled strictly in accordance with the protocol and container label...Supplies should be dispensed under the supervision of the investigator or the assistant investigator, or by a hospital pharmacist.”</p>

Sources: 1. Janssen Research Foundation Clinical Study Protocol GAL-INT-6 2. Janssen Research Foundation Clinical Study Protocol GAL-USA-1, Janssen Research Foundation Clinical Study Report GAL-USA-1, 3. Janssen Research Foundation Clinical Study Protocol GAL-USA-10, Janssen Research Foundation Clinical Study Report GAL-USA-10.

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

Die klinischen Studien GAL-INT-6, GAL-USA-1 und GAL-USA-10 - wie auch GAL-INT-1 und GAL-INT-10 - setzten strikte Concealment -Methoden ein, die die Standards der Good Clinical Practice entsprachen und mit den internen standard operating procedures JRF/PRI-GCP-SOP-350 übereinstimmten. Die eingesetzten Verfahren waren so gestaltet, dass eine Doppelverblindung von Patienten, Prüfern, Studienmonitoren und Datenanalysten gegenüber der Gruppenzuteilung gesichert war und ein Code-Bruch nur bei Eintreten eines Notfalls oder einer Schwangerschaft erlaubt war (Tabelle 6).

- Randomisierung wurde zentral von Janssen Research Foundation durchgeführt
- Patientennummern wurden konsekutiv vergeben, beginnend mit der kleinsten verfügbaren Nummer
- Studienmedikationsdarreichung: Tabletten mit identischem Aussehen, Geschmack und Geruch
- Prüfer (Investigators) waren verblindet gegenüber Gruppenzuweisung; Codes durften nur im Falle eines medizinischen Notfalles gebrochen werden

Study centres were monitored for adherence to study protocol including assignment codes and allocation of study medication.

2. Ergänzende Daten zu Kap. 5.2.4.2.1 u. Tabelle 20: Galantamin vs Donepezil: Abbruchgründe

2.1. GAL-GBR-2 - Abbruchgründe

Tabelle 7: GAL-GBR-2 - Abbruchgründe

Reason for discontinuation	Galantamine N (%)	Donepezil N (%)
Adverse Event	11 (11)	11 (12)
Death	2 (2)	3 (3)
Insufficient response	0 (0)	2 (2)
Subject withdrew consent	4 (4)	1 (1)
Other	2 (2)	3 (3)
Completed	78 (80)	71 (78)

Source: Johnson & Johnson Pharmaceutical Research & Development GAL-GBR-2 A Pilot Study Comparing the Effect of Galantamine (Reminyl) With Donepezil in Patients With Alzheimer's Disease: Clinical Study Synoptic Report, Raritan New Jersey, US.

3. Ergänzende Daten zu Kapitel 5.3.2.7: Therapieassoziierte unerwünschte Ereignisse

3.1 Ergänzende Angaben zu Brodaty 2005: schwere unerwünschte Ereignisse

6.2.2.2. Serious Adverse Events

6.2.2.2.1. By Randomized Treatment Group

In this study, 111 (12%) of 965 subjects each reported 1 or more serious adverse events during the double-blind treatment period or within 30 days after the last dose of study medication. The number of subjects reporting at least 1 serious adverse event was similar among the treatment groups (GAL-CR: 11%; GAL-IR: 12%; placebo: 11%). Most of the serious adverse events reported in this study were considered to be unrelated (doubtful or no relationship) to study treatment by the investigators.

(Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-10, S. 90)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

3.2. Ergänzende Angaben zu Erkinjuntti 2002: Unerwünschte Ereignisse und Todesfälle**3.2.1. Alle unerwünschten Ereignisse****Tabelle 8: GAL-INT-6: Studienabbrecher Gesamt****Table 7-1: Most Frequent Adverse Events Occurring in ≥5% of Patients in Any Group**

System organ class Preferred term	Placebo	GAL 12mg b.i.d.	All patients
	N = 196 n (%)	N = 396 n (%)	N = 592 n (%)
Total number of patients with Aes	133 (67.9)	330 (83.3)	463 (78.2)

(Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 97)

3.2.2. Todesfälle

Die in den Erläuterungen zu Vorbericht Tabelle 30 kommentierte "Inkonsistenz" der Zahlen bei den Todesfällen erklärt sich wie folgt: In der Publikation Erkinjuntti et al, 2006 gibt das Flussdiagramm auf S. 1286 korrekt die Anzahl der Patienten an, die die Studie vorzeitig wegen Todes *abgebrochen* haben, so wie auch in den *Trial Termination Forms* dokumentiert (s. Tabelle 9). Die Anzahl der Todesfälle *insgesamt* (GAL: 9, PLC: 7) basiert auf den *Adverse Event Reportings* und bezieht auch Todesfälle ein, die *nach Beendigung* der doppelblinden Studie auftraten. Insgesamt traten in der Galantamin-Gruppe 4 Todesfälle in der Doppelblind-Phase, 1 Todesfall innerhalb 30 Tagen nach Ende der Doppelblindphase und 4 Fälle (Plc: 2) später als 30 Tage nach Ende der Doppelblind-Phase ein. Von diesen letzten 4 Fällen der Galantamin-Gruppe wurden 2 Fälle aber unbeabsichtigt der Doppelblindphase zugordnet, da die Todesfälle bei database closure bereits in die Datenbank eingetragen waren, aber das Startdatum der open-label-Periode noch nicht. Tabelle 4.3 ("NICE Dossier", S. 34) berichtet daher 7 Fälle unter Galantamin (s. "7.1.3.1 Deaths" unten).

Tabelle 9: GAL-INT-6: Studienabbruchgründe, doppelblinde Studienphase

Display SUB.6: Trial Termination Reasons During Double-Blind

PAGE 1 (LAST PAGE)

02OCT2005 14:14
SYSTEM USED: PDSYD02/pdrun

JRF - TRIAL GAL-INT-6 - ANALYSIS OF DOUBLE-BLIND PHASE

DISPLAY SUB.6: TRIAL TERMINATION REASONS DURING DOUBLE-BLIND

POPULATION: ALL SUBJECTS

ANALYSIS PHASE: DOUBLE-BLIND

* COMPLETED 6 MONTHS OF DOUBLE-BLIND

TERM, TYPE TERM, REASON	PLACEBO NO. SUBJECTS (%)	GAL 12 MG BID NO. SUBJECTS (%)	ALL SUBJECTS NO. SUBJECTS (%)
TOTAL NUMBER OF SUBJECTS	196	396	592
COMPLETED*	143 (83.2)	294 (74.2)	437 (77.2)
DISCONTINUED	53 (36.8)	102 (25.8)	155 (22.8)
ADVERSE EVENT	16 (8.2)	78 (19.7)	94 (15.9)
DEATH	5 (2.6)	4 (1.0)	9 (1.5)
INSUFFICIENT RESPONSE	2 (1.0)	1 (0.3)	3 (0.5)
OTHER	6 (3.1)	9 (2.3)	15 (2.5)
SUBJECT LOST TO FOLLOW-UP	2 (1.0)	0 (0.0)	2 (0.3)
SUBJECT NON-COMPLIANT	1 (0.5)	4 (1.0)	5 (0.8)
SUBJECT WITHDREW CONSENT	1 (0.5)	6 (1.5)	7 (1.2)

(Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 136)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

7.1.3.1. Deaths

According to the analysis, 5 patients in the placebo group and 7 in the galantamine group died due to a serious adverse event with onset either during treatment or within 30 days of discontinuation of double-blind treatment. It should be noted that 2 galantamine-treated patients (A30125 and A30928) died during the open-label period of this trial and were inadvertently included as deaths in the double-blind period for analysis. At the time of double-blind database closure, the start dates of open-label medication for these patients were not yet confirmed. The other 10 deaths occurred during double-blind treatment or within 30 days after treatment discontinuation.

(Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 98-99)

3.3 Ergänzende Angaben zu *Raskind 2000*: Schwerwiegende unerwünschte Ereignisse

Tabelle 10: GAL-USA-1: Todesfälle, andere schwere unerwünschte Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse: doppelblinde Studienphase

Table 28: Number (%) of deaths, other serious adverse events, and discontinuations due to adverse events

Category	Placebo N =213	GAL 12 mg bid N =212
Deaths ^a	1 (0.5%)	1 (0.5%)
Discontinued due to adverse event	16 (7.5%)	49 (23.1%)
Discontinued due to serious adverse event	9 (4.2%)	11 (5.2%)
Serious adverse event ^a	27 (12.7%)	29 (13.7%)

^a: Events occurred during double-blind treatment phase and within 30 days after treatment discontinuation.

(Quelle: Janssen Research Foundation, Clinical Research Report GAL-USA-1, S. 94)

3.4 Ergänzende Angaben zu *Wilcock 2000*: Schwerwiegende unerwünschte Ereignisse

Tabelle 11: GAL-INT-1: Todesfälle, andere schwere unerwünschte Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse

Table 28: Incidence of deaths, discontinuations due to AE, and serious AEs other than death

	Placebo N=215	GAL 12 mg b.i.d. N=220
Deaths	2 (0.9%)	2 (0.9%)
Discontinued due to AE	19 (8.8%)	31 (14.1%)
Discontinued due to serious AE	11 (5.1%)	8 (3.6%)
Serious AEs other than death	25 (11.6%)	29 (13.2%)

(Quelle: Janssen Research Foundation, Clinical Research Report GAL-INT-1, S. 79)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

3.5 Ergänzende Angaben zu *Tarivot 2000*: Schwindel (Dizziness)

Vorbericht Figur 37 geht von 11 Fällen in der Placebo-Gruppe aus; es sind aber 10, siehe Tab. 10:

Tabelle 12: GAL-USA-10: Anzahl Patienten mit unerwünschtem Ereignis Schwindel (Dizziness)

**Table 24: Incidence of most frequent ($\geq 5\%$) all adverse events:
number (%) of patients**

WHO system-organ class WHO-preferred term	Placebo (N=286)	GAL 4 mg bid (N=140)	GAL 8 mg bid (N=279)	GAL 12 mg bid (N=273)
Central & peripheral nervous system disorders	42 (14.7%)	22 (15.7%)	55 (19.7%)	64 (23.4%)
Dizziness	10 (3.5%)	7 (5.0%)	15 (5.4%)	19 (7.0%)
Headache	13 (4.5%)	5 (3.6%)	19 (6.8%)	13 (4.8%)

(Quelle: Janssen Research Foundation, Clinical Research Report GAL-USA-10, S. 86)

Stellungnahme Janssen-Cilag GmbH (Fortsetzung)

CORRIGENDUM

zur Anlage zur Stellungnahme des Unternehmens Janssen-Cilag GmbH zum IQWiG-Vorbericht "Cholinesterasehemmer bei Alzheimer Demenz" [Auftrag A05/19-A] vom 01.09.2006

Table 4: GAL-INT-6: CIBIC-plus- in Subgruppe AD-CVD; ITT-Analyse

Table 6-16: Distribution of CIBIC-plus Scores in the Mixed Dementia Subgroup at Month 6: CITT Analysis

CIBIC-plus scores ^a	Placebo n=92		GAL 12 mg b.i.d. n=172		p-value ^b
	n (%)	Cumulative %	n (%)	Cumulative %	
Markedly improved	1 (1.1%)	1.1%	4 (2.3%)	2.3%	p=0.001
Moderately improved	3 (3.3%)	4.3%	13 (7.6%)	9.9%	
Minimally improved	14 (15.2%)	19.6%	34 (19.8%)	29.7%	
No change	32 (34.8%)	54.3%	76 (44.2%)	73.8%	
Minimally worsened	31 (33.7%)	88.0%	31 (18.0%)	91.9%	
Moderately worsened	11 (12.0%)	100.0%	12 (7.0%)	98.8%	
Markedly worsened	0 (0.0%)	100.0%	2 (1.2%)	100.0%	

^a 7-point rating scale
^b p-value for comparison with placebo.

Source: Display EFE.CIB.16.

Quelle: Janssen Research Foundation, Clinical Study Report GAL-INT-6, S. 86
 CITT = classical intent to treat.

Stellungnahme Klinikum Bremen-Mitte, Wille



Engelgasse 10000

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:
Cholinesterasehemmer bei Demenz
A 05 - 13a

Name, Vorname, Titel und Funktion des/der Stellungnehmenden:

1. Wille, Dr. Hans
2. Oberarzt
3. Institut für Klinische Pharmakologie, Bremen in 28177 Bremen

Kontakt für weitere Nachfragen:

Name: Dr. Hans Wille
Institution: Institut klinische Pharmakologie
Anschrift: Klinikum Bremen-Mitte, 28177 Bremen
Tel./Fax-Nr.: 0421-437-348 FAX - 3326
E-Mail-Adresse: h.wille@klinikpharm-bremen.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Klinikum Bremen-Mitte, Wille (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für **alle Stellungnehmenden** einzeln **ausgefüllt** vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Bremen, 6.10.06
Ort/Datum

Haus Wille
Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

2

Stellungnahme Klinikum Bremen-Mitte, Wille (Fortsetzung)

**Schriftliche Stellungnahme zum Vorbericht des IQWiG:
Cholinesterasehemmer bei Alzheimer Demenz
[A05-19A]**

Der veröffentlichte Vorbericht zum Nutzen von Cholinesterase(ChE)hemmern bei Demenz erscheint sehr intensiv und hervorragend recherchiert. Der Bewertung der einzelnen Studien kann aber nicht in jedem Fall zugestimmt werden. Zudem weist der Bericht bei der Interpretation der Daten methodische Inkonsistenzen und Intransparenzen auf, die zu nicht nachvollziehbaren Schlussfolgerungen führen.

1. Im Vorbericht fehlende Studien:

Die Literaturrecherche ist ausführlich und gründlich durchgeführt. Es erscheint allerdings nicht klar, weshalb eine im April 2006 veröffentlichte Studie⁹ nicht mit in die Auswertung einbezogen wurde. Ansonsten sind keine Studien bekannt, die nach den vorgegebenen Kriterien hätten erfasst werden müssen und nicht erfasst wurden.

Der Ausschluss der AD2000 auf Grund (noch nicht gelöster) methodischer Mängel ist vertretbar und im Grunde korrekt. Allerdings ist zu konstatieren, dass die AD2000 die bisher einzige Studie darstellt, die ohne Diskussion als primären Zielparameter einen wirklich relevanten klinischen und sicher messbaren Endpunkt untersucht. Vor dem Hintergrund, dass andere Studien, die als „mit groben Mängeln“ behaftet beurteilt werden (s.u.), eingeschlossen wurden, erscheint der Ausschluss der AD2000 jedoch fragwürdig und nur schwer nachvollziehbar.

2. Anmerkungen zur Bewertung von Originalstudien im Vorbericht

a) Differenzen zur Beurteilung einzelner Kriterien (Tab. 8, 12, 16)

(Nicht alle Originalarbeiten lagen vor)

Donepezil: Burns 1999: Abbrecher/Gründe je nicht genannt; Mohs 2001: Concealment unklar; ITT verletzt; Winblad 2001: ITT verletzt;

Rivastigmin: Corey-Bloom: keine verblindete Erhebung; Rösler: keine verblindete Erhebung

Galantamin: Brodaty: ITT verletzt; Erkinjuntti: Randomisierung unklar, kein Concealment, verblindete Erhebung unklar; Raskind: kein Concealment, verblindete Erhebung unklar; Tariot: kein Concealment.

Bei den übrigen Kriterien fand sich in den Studien eine weitgehende Übereinstimmung mit den im Bericht dargestellten Bewertungen.

Obwohl bei einzelnen Kriterien sicher Bewertungsspielräume bestehen, erschien im Bericht die Bewertung der Kriterien eher großzügig und zu Gunsten der Studien. Bei den benannten Abweichungen handelte es sich jeweils um schlechtere Bewertungen als im Bericht vorgenommen. Insgesamt erschien damit die Qualität der Studien zu positiv eingeschätzt.

b) Beurteilung der Studien- und Publikationsqualität (Tab. 8, 12, 16)

Die Kategorisierung der Studien in „leichte Mängel“ und „grobe Mängel“ ist nicht transparent und in vielen Fällen nicht nachvollziehbar. In den eigenen Vorgaben des IQWiG zur Beurteilung der biometrischen Qualität (Erläuterung zur Dokumentation und Extraktion von Daten bei RCT, Version 3.0) heißt es: „Bei *groben Mängeln* wäre die Gesamtaussage grundsätzlich (auch dann?) in Frage zu stellen, wenn die Mängel behoben würden“ und „*Leichte Mängel* liegen dann vor, wenn ... deren Behebung ... Gesamtaussage der Studie nicht wesentlich beeinflussen würde“.

Bei Verwendung dieser Kriterien scheint nicht nachvollziehbar, dass für Studien ohne gesichertes Concealment und zusätzlich ohne verblindete Endpunkterhebung nur „leichte

Stellungnahme Klinikum Bremen-Mitte, Wille (Fortsetzung)

Mängel“ konstatiert werden (Donepezil: Winblad 2001, Tune 2003, Tariot 2001, Seltzer 2004, Krishnan 2003, Gauthier 2002; Rivastigmin: Rösler 1999; Galantamin: Tariot 2000). Bei einigen Studien ist zudem trotz „leichter Mängel“ – teils zusätzlich – die Randomisierung (Donepezil: Seltzer 2004, Tune 2003; Rivastigmin: Corey-Bloom 1998; Galantamin: Rockwood 2006) oder eine adäquate Fallzahlplanung (Donepezil: Krishnan 2003, Rogers 1998, Setzer 2004, Tariot 2001, Tune 2003, Winblad 2001) nicht gesichert.

Im Vorbericht wird mehrfach festgestellt (z.B. S. 47, 59), dass ein **Concealment** bei verblindeten Studien kein entscheidender Faktor für die Qualität der Studie ist. Dem ist klar zu widersprechen. Bei Studien unklaren bzw. fehlendem Concealment wurden in systematischen Untersuchungen Effekt-Überschätzungen um 30% bzw. um 37-41% verglichen mit adäquaten Concealment gefunden^{1,2}. Der Fehler durch unklares Concealment zu Gunsten der zu prüfenden Intervention ist größer als der durch inadäquate Verblindung oder Randomisierung^{3,4,7}. Es ist deshalb davon auszugehen, dass durch Einbezug von Studien mit unklarem oder inadäquaten Concealment im Bericht eine generelle Überschätzung der Effekte zu Gunsten der ChE-Hemmer stattgefunden hat. Dies ist von besonderer Bedeutung, da bei ChE-Hemmern durch typische UAW von einer häufigen Entblindung während der Studie auszugehen ist (s.u.). Ein adäquates Concealment ist nur bei 11 der 12 Placebo-Studien mit Donepezil, 2 der 4 Rivastigmin-Studien und 3 der 6 Galantamin-Studien eingehalten.

Eine **verblindete Erhebung der Zielgrößen** (durch Untersucher, die an der Versorgung der Patienten unbeteiligt sind) ist von besonderer Bedeutung, wenn bei der Bewertung der Zielgrößen subjektive Momente eine Rolle spielen und/oder die Intervention wegen typischer Begleiteffekte durch Patienten und/oder Behandelnde erkennbar ist^{5,10}. Beides trifft auf die Studien für ChE-Hemmer zu: 1. Alle ausgewerteten Zielgrößen waren Scores auf Grundlage von Befragungen und Einschätzungen. 2. Das UAW-Profil der ChE-Hemmer ist typisch und bekannt; in den Studien des Berichts finden sich jeweils (!) bis zu 5-15% höhere Raten für Übelkeit, Erbrechen, Diarrhoe und Schwindel unter Verum als unter Placebo. Es ist somit von einer bedeutenden Post-Randomisierungs-Entblindung auszugehen (Tab. 25,31,37). Die FDA hat schon 1991 darauf hingewiesen, dass bei Studien mit Alzheimer-Mitteln die Auswerter gegenüber der Therapiezuweisung verblindet sein müssen⁶. Eine Überprüfung der Güte der Verblindung wäre ein wichtiges weiteres Qualitätskriterium gewesen, wurde aber für die eingeschlossenen Studien nicht durchgeführt. Insgesamt ist nur bei 8 der 12 Donepezil-Studien, 3 der 4 Rivastigmin-Studien und 1 der 6 Galantamin-Studien von einer adäquaten Verblindung auszugehen.

Zusammenfassend stellen inadäquates Concealment und unverblindete Endpunkterhebung wesentliche Qualitätskriterien dar, die bei 16 bzw. 12 der 22 eingeschlossenen Studien (bei 9 sogar gemeinsam) nicht erfüllt waren. Die zu erwartenden Effektüberschätzungen (bei allen bewerteten Effektivitätskriterien) sind nicht gebührend berücksichtigt. Sensitivitäts-Analysen für adäquates und inadäquates Concealment oder Verblinden der Untersucher wären sicher notwendig gewesen, sind aber nicht durchgeführt. Daten zum adäquaten Concealment und zur verblindeten Endpunkterhebung hätten auch bei den Autoren/Herstellern nachgefragt werden können.

c) Gemeinsame Auswertung von Studien mit „leichten“ und „groben“ Mängeln

Bei der gemeinsamen Auswertung der Daten wird die Studienqualität der Daten nicht weiter berücksichtigt. Es ist aber bekannt, dass in Studien minderer gegenüber solchen besserer Qualität Effekte um 34% überschätzt werden². 5 der Donepezil-Studien, 2 der Rivastigmin-Studien und 1 der Galantamin-Studien weisen nach Einschätzung der Autoren selbst „grobe Mängel“ auf; dies entspricht bei Galantamin zwar weniger als 10% der Patienten, aber bei Donepezil mehr als 50% und bei Rivastigmin mehr als 35% der Patienten. Die gemeinsame Auswertung von Studien unterschiedlicher Qualität ist zwar weiterhin ein methodisch nicht

Stellungnahme Klinikum Bremen-Mitte, Wille (Fortsetzung)

gelöstes Problem; auf jeden Fall wäre auch hier eine Sensitivitäts-Analyse für Studien mit guter und schlechter Qualität notwendig gewesen. Sie wurde nicht durchgeführt.

d) „Last-Observation carried forward“-Methode (LOCF)

Mehr als 20% der Studienteilnehmer brachen insgesamt die Studie vorzeitig ab; der Anteil war unter Placebo etwa 5% kleiner als unter den ChE-Hemmern. Es muss bis zum Beweis des Gegenteils davon ausgegangen werden, dass die für Studienabbrecher in den Studien dann angewandete LOCF-Methode zu einer systematischen Effektüberschätzung (gegenüber der angestrebten und zu fordernden ITT-Methode aller nicht verstorbenen oder sonstig zensierten Studienteilnehmer am Studienende) geführt hat, da die Demenz eine progrediente Erkrankung ist und bei früheren Messungen die Messwerten besser ausfallen.

Die Argumente in der Diskussion des Berichtes (S. 168, 169), dass von einer Überschätzung des Effektes nicht auszugehen ist, überzeugen nicht – als Beleg können sie in keiner Weise gelten. In der zitierten Studie von Farlow (gepoolte Nach-Analyse der Studienabbrecher aus 3 RCT) wurde nur ein der RDO-Patienten („retrieved drop outs“) erfasst. Dass hier Verum-Patienten besser abschnitten als Placebo-Patienten, kann allein durch Selektion von Patienten bedingt sein, die für eine nochmalige Untersuchung bereit waren; es beweist nicht, dass eine LOCF-Auswertung eher konservativ ist. Auch die Abb. 47 (S. 169) taugt nicht zum Beleg, dass LOCF-Auswertungen valide Ergebnisse erbringen: es ist hier unklar, wie die Auswahl der betrachteten Studien erfolgte und wie ihre methodische Qualität war. Bei Ignorierung der beiden Studien links oben und der rechts unten („Ausreißer“?) würde man bei der Meta-Regression sehr wahrscheinlich eine Gerade bekommen, die durch den Null-Punkt (der Effekt-Differenz) liefe und einen Zusammenhang zwischen Abbruchraten-Differenz (Verum – Placebo) und Effektstärke zeigen würde; größere Effekte unter Verum könnten dann durch die im Vergleich zu Placebo höheren Abbruchraten erklärt werden (ebenfalls spekulativ – andere Interpretationen wären auch denkbar).

Zusammenfassend liegen somit keine ausreichenden Belege dafür vor, dass durch die LOCF-Methode die ermittelten Effekte nicht überschätzt sind. Dieser Beweis müsste positiv für das Verum erbracht werden, bevor die Wahrscheinlichkeit einer Effektüberschätzung verneint wird.

e) Publikations-Bias

In dem Bericht fehlt ein formalisiertes Verfahren zur Überprüfung eines ev. Publikations-Bias (Funnel-Plot oder ähnliche, auch quantitative Instrumente).

f) Multiples Testen

In fast allen Studien wird die Wirksamkeit an Hand zweier oder mehr Kriterien geprüft; in keiner fand eine Adjustierung wegen multiplen Testens statt. Dieser Umstand wurde bei der Berichterstellung nicht berücksichtigt. Außerdem flossen in den Bericht häufig Daten von sekundär erhobenen Zielkriterien ein (an deren Validität generell gezweifelt werden kann), für die ebenfalls keine Adjustierungen des Signifikanzniveaus erfolgte.

g) Studiendauer

Für den Einschluss wird lediglich eine Mindestbeobachtung von 16 Wochen gefordert. Dies erscheint praxisfern und geht nicht konform mit Forderungen der EMEA, die einen Wirksamkeitsnachweis über 6 Monate als Mindestzeit fordert und über 12 und mehr Monate empfiehlt⁸.

3. Generelles zur projektspezifischen Methodik

Stellungnahme Klinikum Bremen-Mitte, Wille (Fortsetzung)

Ein zentraler Kritikpunkt an der Methodik für die Berichterstellung ist sicher die unter dem Aspekt der Patientenrelevanz unzureichend klare Definition der interessierenden Endpunkte. Veränderungen in den Scores, wie sie in den Studien verwendet wurden, sind als Surrogate (bestenfalls Intermediär-Endpunkte) zu sehen, deren klinische Relevanz unklar oder im besten Fall strittig ist und interindividuell erheblich variieren kann. Da – soweit bekannt – außer der AD2000 (s.o.) keine Demenz-Studien mit tatsächlich harten, patientenrelevanten Endpunkten (Fremdhilfe, Pflegebedarf, Heimunterbringung etc.) vorliegen, ist der Rückgriff auf Studien mit Erfassung von Score-Änderungen als Surrogate für Wirksamkeit vielleicht plausibel, kann aber das grundsätzliche Problem fehlender relevanter Daten nicht lösen. Es wäre zu wünschen und nötig gewesen, dass der Bericht klar und unmissverständlich festhält, dass wegen Fehlens solcher Studien die klinische Wirksamkeit der ChE-Hemmer hinsichtlich patientenrelevanter Verbesserungen der Symptomatik oder des Krankheitsverlaufs derzeit nicht beurteilt werden kann.

Wenn schon auf Scores zurückgegriffen wurde, hätten diejenigen Minimaländerungen, die als klinisch relevant angesehen werden sollten, im Bericht prädefiniert werden müssen. Welche Bedeutung die beobachteten mittleren Änderungen (z.B. im ADAS-cog, DAD oder CIBIC) unter den ChE-Hemmern für die Patienten haben, bleibt weitgehend unklar. Von der FDA werden beispielsweise seit Jahren Änderungen im ADAS-cog erst ab 4 Punkte als relevant eingeschätzt, die im Mittel unter keinem der ChE-Hemmer erreicht wurden. Der prädiktive Wert der Score-Instrumente bezüglich patientenrelevanter therapeutischer Effekte ist im Bericht weder thematisiert noch diskutiert worden, Gleiches gilt für die Validität und für die Reliabilität der Scores. Dies ist umso bedauerlicher, als die in den Studien eingesetzten Scores und Skalen nur schwer zugänglich und deshalb für Ärzte, Patienten und Betreuer weitgehend intransparent sind und in der Praxis auch kaum angewandt werden.

Bremen, 6.10.2006

Dr. Hans Wille
Institut für Klinische Pharmakologie
Klinikum Bremen-Mitte

Literatur:

1. Schulz, KF et al.: JAMA 1995; 273: 408-12
2. Moher, D et al.: Lancet 1998; 352: 609-13
3. Jüni, P et al.: BMJ 2001; 323: 42-46
4. Schulz, KF et al.: 2002; 359: 614-18
5. Day, SJ et al.: BMJ 2000; 321: 504-2000
6. FDA: NEJM 1991; 324: 349-52
7. Hewitt, C et al.: BMJ 2005; 330: 1057-58
8. EMEA: Note for Guidance ... (unter www.emea.europa.eu/)
9. Winblad, B et al.: Lancet 2006; 367: 1057-65
10. Schulz, KF et al.: Lancet 2002; 359: 696-700

Stellungnahme Kompetenznetz Demenzen, Maier



Titel: A05-19A

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:
Cholinesterasehemmer bei Alzheimer-Demenz
A05/19A

Name, Vorname; Titel und Funktion des/der Stellungnehmenden:

1. Maier, Wolfgang; Prof. Dr. med.
Sprecher des 'Kompetenznetz Demenzen'
- 2.
- 3.

Kontakt für weitere Nachfragen:

Name: Koordination "Kompetenznetz Demenzen"
Institution: c/o Zentralinstitut für Seelische Gesundheit
Anschrift: 75, 68072 Mannheim
Tel./Fax-Nr.: 0621 17203-2903 (Tel.) -2005 (Fax)
E-Mail-Adresse: demenz@zi-mannheim.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik
(bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Kompetenznetz Demenzen, Maier (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt. *beifügt*
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Bonn, 6.10.06
Ort/Datum

Unterschrift 1
Prof. Dr. med. W. Maier

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme Kompetenznetz Demenzen, Maier (Fortsetzung)

IQWiG: A05/19-A
(Vorbericht: Cholinesterasehemmer bei Alzheimer-Demenz)

Stellungnahme
(Prof. Dr. W. Maier für das "Kompetenznetz Demenzen"):

Die Bestätigung des Nutzens hinsichtlich der Zielparameter "kognitive Leistungsfähigkeit" und "ADL-Funktionen" entspricht dem aktuellen und von breitem Konsens getragenen nationalen und internationalen medizinischen Wissensstand zum Ausmaß und der Bedeutung der ChE-Inhibitoren in der Therapie der Alzheimer-Demenz.

Durch die ausschließliche Beschränkung der Analyse und Bewertung auf bestimmte randomisierte und kontrollierte Studien, die wesentlich über theoretisch-biometrische Kriterien begründet wird und die klinischen Besonderheiten der Demenzerkrankungen außer Acht lässt, werden weitere wichtige Wirksamkeitsaspekte nicht adäquat gewürdigt, dies betrifft insbesondere Daten zur Pflegeheimweisung.

Die Beschränkung der zu untersuchenden Studien auf eine Mindest-Studiendauer von 16 Wochen erscheint willkürlich und lässt eine Reihe von Studien mit 12 Wochen Untersuchungsdauer außer Acht.

Auch wurden angehörigerelevante Therapieziele als patientenrelevante Zielparameter wegen der klinischen Besonderheiten der Demenzen (zunehmender Verlust der Fähigkeit zur Selbstbestimmung bei Patienten) durch das Amendment des Berichtsplans im Juni 2006 in ihrer Bedeutung abgewertet, obwohl diese in der klinischen Beurteilung einer Demenz und für die Beurteilung von Therapieeffekten eine wesentliche Rolle spielen.

Weiterhin wurde die Bedeutung des durch den behandelnden Arzt (oder: Studienarzt bei RCT) zu erhebenden klinischen Eindrucks als patientenrelevanter Zielparameter zu einer "ergänzenden" Information herabgestuft, dies ist unlogisch und klinisch inkorrekt, weil gerade dieser Parameter per definitionem darauf angelegt ist, für einen geschulten Arzt/Beobachter klinisch bedeutsame Änderungen zu erfassen. Das klinische Arzturteil hat eine belegte hohe Validität; die standardisierte Beurteilung von Therapieeffekten kann das klinische Arzturteil nicht vollständig ersetzen, da diese stets nur Teilaspekte abbildet. Dieser Gesichtspunkt ist umso relevanter, je umfassender und schwerer die Erkrankung ist. Die Ergebnisse zum klinischen Gesamteindruck aus den analysierten Studien wurden aber lediglich als "ergänzende Information" in den Vorbericht einbezogen und blieben somit in der Gesamtbewertung weitgehend unberücksichtigt.

Stellungnahme Kompetenznetz Demenzen, Maier (Fortsetzung)

Der Vorbericht bezieht sich auf die publizierte Studienlage, für die erhebliche methodische Mängel konstatiert werden. Zweifellos ist die Studienqualität häufig unbefriedigend. In diesem Zusammenhang muss der Schlussfolgerung von IQWiG beigeprflichtet werden, dass rein aus diesen Mängeln eine "resultierende Negativbeurteilung nicht angemessen" ist. Die angemahnten Mängel resultieren nämlich einmal aus dem Umstand, dass die bewerteten randomisierten Studien nicht zur Nutzenbewertung, sondern lediglich zu Zwecken der dann auch jeweils erfolgten Zulassung durch Arzneimittelbehörden erstellt wurden; die dabei zugrunde liegenden Rohdaten der Studien bilden sich unter Umständen nur unzulänglich in den publizierten Studien ab, die der IQWiG-Bewertung zugrunde lagen. Zum anderen erzwingt die Erkrankungsschwere Unzulänglichkeiten (z.B. Protokollverletzungen, Drop outs).

Vor allem sollte auch das Fehlen von medikamentösen Behandlungsalternativen in die Gesamtbewertung des Nutzens von ChE-Inhibitoren für die Versorgung von Demenzpatienten einbezogen werden. Auch eine Aufrechnung von Pflegeleistungen gegen medikamentöse Behandlungsweisen ist ärztlich unsinnig und fällt wegen des sehr geringen Anteils von Medikamenten- an den Gesamtkosten auch kaum ins Gewicht. Hierauf hätte in einer Präambel o.ä. zur evidenzbasierten gutachterlichen Stellungnahme verwiesen werden sollen, um eine ethische und wertorientierte Einordnung der evidenzbasierten Analyse der Medikamenteneffekte zu ermöglichen.

Stellungnahme Merz Pharmaceuticals GmbH



Auftragsnummer.:	A05/19-A
-------------------------	----------

Eingangsstempel

Formblatt zur schriftlichen Stellungnahme zum Vorbericht

Name, Vorname; Titel und Funktion des/der Stellungnehmenden
1. Herr Dr. med. Albrecht Stöffler Leiter Klinische Forschung und Entwicklung ZNS
2. Frau Dr. med. Yvonne Wirth Gruppenleiterin Medizinische Statistik
3.

Kontaktdaten
Name: Herr Dr. Stöffler / Frau Dr. Wirth
Institution: Merz Pharmaceuticals GmbH
Anschrift: Eckenheimer Landstraße 100, 60318 Frankfurt am Main
Tel./Fax-Nr.: Tel. 069 1503-604 bzw. -507 / Fax 069 1503-759 bzw. -803
E-Mail-Adresse: albrecht.stoeffler@merz.de / yvonne.wirth@merz.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt
1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)



Auftragsnummer.:	A05/19-A

Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zum Stellungnahmeverfahren zu Vorberichten des Instituts“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für alle Stellungnehmenden einzeln ausgefüllt vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zum Stellungnahmeverfahren zu Vorberichten des Instituts“ habe ich / haben wir zur Kenntnis genommen.

Frankfurt, 24.10.06
Ort/Datum

Unterschrift 1 Dr. A. Stöffler

Unterschrift 2 Dr. Y. Wirth

Unterschrift 3

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)

**Stellungnahme zum Vorbericht des IQWiG vom 01. September 2006
Cholinesterasehemmer bei Alzheimer-Demenz**

Das IQWiG hat in seinem Vorbericht über die Cholinesterasehemmer bei Alzheimer Demenz auch die Effekte auf die begleitende Psychopathologie als einen der klinisch relevanten Endpunkte beurteilt. Es kommt dabei zu folgenden Schlußfolgerungen für die drei beurteilten Substanzen:

"Zusammenfassend geben die Studien keinen eindeutigen Hinweis auf eine günstige oder ungünstige Beeinflussung der Psychopathologie durch Donepezil."

"Zusammenfassend ergeben sich damit Hinweise auf einen Effekt von Galantamin bezüglich der Beeinflussung der neuropsychiatrischen Symptomatik. Die Meta-Analyse (Abbildung 29) bestätigt dies und zeigt einen nur geringfügigen, wenn auch statistisch signifikanten, gepoolten Effekt in der Größenordnung von einem Siebtel Standardabweichung (Cohen's $d = -0,14$; 95 %-KI: $-0,24 - -0,03$). Dies entspricht auf die NPI-Skala umgerechnet etwa 1 bis 2 Scorepunkten."

[Verwertbare Daten zu NOSGER und BEHAVE-AD lagen nicht vor]. "Damit liegen keine Hinweise auf eine günstige Beeinflussung des Betreuungsaufwandes durch Rivastigmin vor."

Diese Einschätzung sollte kritisch hinterfragt werden. Verhaltensstörungen sind sehr häufig bei der Alzheimer Demenz (Suh et al., 2004). Neuere Studien haben gezeigt, dass 92% der Demenzpatienten im Verlauf der Erkrankung eine Verhaltensstörung entwickeln (Hope et al., 1999; Wynn et al., 2004). Bei bis zu 80% von ihnen manifestieren sich sogar zwei oder mehr Verhaltensauffälligkeiten während der Erkrankung. Mit fortschreitender Demenzprogression nimmt die Häufigkeit der Verhaltensstörungen zu. Die am häufigsten bei der Alzheimer Demenz auftretenden Symptome sind Wahnvorstellungen, Halluzinationen, Agitation/Aggression, Depression und Angst (Craig et al., 2005). Diese Symptome sind sehr belastend für die Angehörigen/Pflegenden und führen häufig zu einer Überforderung (Aupperle, 2006; Shih et al., 2005) und schließlich zur Heimeinweisung der Erkrankten.

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)

Letzteres ist mit sehr hohen sozialen und ökonomischen Kosten verbunden und steht im Widerspruch zum Wunsch von Betroffenen, im eigenen sozialen Umfeld zu bleiben. Zudem geht eine Heimeinweisung häufig mit einem schnelleren Verlauf der Erkrankung einher. Aus diesen Gründen hat in den letzten Jahren diese Problematik immer mehr an Bedeutung gewonnen. Die Verhaltensstörungen sind aus der Sicht der Patientenrelevanz sogar höher einzustufen als die Kognition.

Im IQWiG-Vorbericht zu den Acetylcholinesterasehemmern wird (vermutlich eben aus diesen Gründen) eine Wirkung auf die begleitende Psychopathologie (Verhaltensauffälligkeiten) noch vor der Kognition beurteilt. Die Bewertung basiert auf Daten von Studien, in denen entsprechende Skalen eingesetzt wurden. Bei Donepezil und Galantamin handelt es sich hierbei ausschließlich um das Neuropsychiatric Inventory (NPI) oder um dessen Pflegeheim-Version (NPI-NH). Bei Rivastigmin konnte keine Beurteilung gemacht werden, da keine NPI-Daten verfügbar waren, Studien mit Einbezug der NOSGER Skala eine hohe Anzahl an Studienabbruchern vorwies, und bezüglich der BEHAVE-AD Skala keine Auswertung durchgeführt wurde.

Bei der Beurteilung der NPI-Ergebnisse zu Donepezil und Galantamin berücksichtigte das Institut lediglich die Gesamtpunktzahl der NPI-Skala. Die NPI-Skala besteht hingegen aus 12 unterschiedlichen Symptombereichen, die jeweils nach Schwere und Häufigkeit der Störung beurteilt werden (Cummings et al., 1994). Eine Zusammenfassung aller 12 Größen kann nur Anhaltspunkte über das generelle Vorliegen von Verhaltensstörungen liefern, erlaubt jedoch keine differenzierte Betrachtung der einzelnen Bereiche und damit keine Gewichtung nach Symptomen. Hingegen sind die Verhaltensprofile der Patienten individuell sehr unterschiedlich; ebenfalls ist die Bedeutung bestimmter Symptome von unterschiedlicher klinischer Relevanz. Um diese Besonderheiten zu berücksichtigen, tendieren Experten immer mehr dazu, in erster Linie die einzelnen Bereiche der Skala zu betrachten anstelle der Gesamtpunktzahl aus allen Bereichen (Aupperle, 2006; Cummings et al., 2006; Engelborghs et al., 2005; Mirakhur et al., 2004; Shih et al., 2005). Die Relevanz dieser Einzelsymptombetrachtung wird zusätzlich verstärkt durch die Beobachtung, dass z.B. die Symptomatik Agitation/Aggression einer der häufigsten Gründe für eine Pflegeheimeinweisung ist.

Die unterschiedlichen Erkrankungsmuster spiegeln sich in den unterschiedlichen Symptomausprägungen wider. Daher sind diese Unterschiede allein im Vergleich

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)

der verschiedenen Demenzformen deutlich feststellbar. Während bei der gemischten und bei der Alzheimer Demenz die Agitation/Aggression zu den häufigsten (ca. 80%) Verhaltensstörungen gehört, sind es bei der frontotemporalen Demenz die Apathie, bei der Lewy Body–Demenz die Halluzinationen (Engelborghs et al., 2005). Da eine medikamentöse Intervention nicht bei allen Demenzformen vergleichbar effektiv wirkt, können daher nicht alle 12 Symptomen-Bereiche mit nur einem Medikament therapeutisch signifikant verbessert werden. Eine zusammenfassende, globale Betrachtung, wie vom Institut vorgenommen, die ausschließlich die Gesamtpunktzahl betrachtet, führt zu einer Verzerrung relevanter und signifikanter Einzel-Symptomatik-Ergebnisse und unterbindet differenziertere Aussagen.

Daher erscheint uns die Betrachtung einzelner Symptom-Bereiche oder deren krankheitsspezifische Gruppenzusammenfassung anhand von Clustern klinisch sinnvoller. Auch sollten weitere differenziertere Betrachtungen hinsichtlich Neu-Auftreten von bisher nicht beobachteten Symptomatiken, oder Verbesserung bereits existierender Symptomatiken unter der medikamentösen Therapie, mitbetrachtet werden.

Prof. Cummings, der Entwickler der NPI-Skala, bezieht sich in seiner neuesten Publikation ebenfalls auf diese Methodik der Einzelbetrachtung (Cummings et al., 2006).

Aus diesen Gründen sollte die Beurteilung der Verhaltenseffekte von AChEIs anhand der NPI-Summscore-Ergebnisse durch das IQWiG kritisch hinterfragt, ausführlicher diskutiert und durch eine differenzierte Betrachtung der Einzelsymptome ergänzt werden.

Frankfurt/Main, 06. Oktober 2006

Dr. med. Yvonne Wirth

Dr. med. Albrecht Stöffler

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)

- (1) Aupperle P. Management of aggression, agitation, and psychosis in dementia: focus on atypical antipsychotics. *Am J Alzheimers Dis Other Demen* 2006; 21(2): 101-108.
- (2) Craig D, Mirakhor A, Hart DJ, McIlroy SP, Passmore AP. A cross-sectional study of neuropsychiatric symptoms in 435 patients with Alzheimer's disease. *Am J Geriatr Psychiatry* 2005; 13(6): 460-468.
- (3) Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994; 44(12): 2308-2314.
- (4) Cummings JL, Schneider E, Tariot PN, Graham SM. Behavioral effects of memantine in Alzheimer disease patients receiving donepezil treatment. *Neurology* 2006; 67(1): 57-63.
- (5) Engelborghs S, Maertens K, Nagels G, Vloeberghs E, Marien P, Symons A, et al. Neuropsychiatric symptoms of dementia: cross-sectional analysis from a prospective, longitudinal Belgian study. *Int J Geriatr Psychiatry* 2005; 20(11): 1028-1037.
- (6) Hope T, Keene J, Fairburn CG, Jacoby R, McShane R. Natural history of behavioural changes and psychiatric symptoms in Alzheimer's disease. A longitudinal study. *Br J Psychiatry* 1999; 174: 39-44.
- (7) Mirakhor A, Craig D, Hart DJ, McIlroy SP, Passmore AP. Behavioural and psychological syndromes in Alzheimer's disease. *Int J Geriatr Psychiatry* 2004; 19(11): 1035-1039.
- (8) Shih CS, Yan SH, Ho YH, Lin YT, Li JY, Lo YK. Psychiatric morbidity in dementia patients in a neurology-based memory clinic. *Acta Neurol Taiwan* 2005; 14(4): 179-186.

Stellungnahme Merz Pharmaceuticals GmbH (Fortsetzung)

- (9) Suh GH, Kim SK. Behavioral and Psychological Signs and Symptoms of Dementia (BPSD) in antipsychotic-naive Alzheimer's Disease patients. *Int Psychogeriatr* 2004; 16(3): 337-350.
- (10) Wynn ZJ, Cummings JL. Cholinesterase inhibitor therapies and neuropsychiatric manifestations of Alzheimer's disease. *Dement Geriatr Cogn Disord* 2004; 17(1-2): 100-108.

Stellungnahme Novartis Pharma GmbH



Formblatt Stellungnahme zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

Projektbezeichnung: Cholinesterasehemmer bei Alzheimer Demenz
Auftragsnummer: A05/19-A

Name, Vorname des/der Stellungnehmenden:

1. Dr. med. F. Tracik
2. Dr. M. Baier
3. Dr. med. D. Götte

Kontakt-Informationen des Trägers:

Anschrift: Novartis Pharma GmbH, z.Hd. Dr. med. Ferenc Tracik; Roonstr. 25, 90429
Nürnberg
Tel./Fax-Nr.: 0911-273-12532; -12588
E-Mail-Adresse: Ferenc.Tracik@novartis.com

Stellungnahme Novartis Pharma GmbH (Fortsetzung)



L. Folgende Originalstudien haben im Vorbericht (Bitte beachten Sie die Ein- und Auschlusskriterien für Studien) die Bewertung:

Psychopathologische Symptome (PPS)

Im Vorbericht wird erwähnt, dass es zur Wirksamkeit von Rivastigmin auf PPS keine Daten gibt. Dieser Bewertung ist in Bezug auf Rivastigmin nicht haltbar. Außerdem muss die Wertigkeit dieses Parameters für leichte bis mittelschwere Krankheitsstadien gegenüber anderen Messgrößen wie der Kognition in Relation gesehen werden, denn PPS treten in frühen Stadien der Alzheimer-Erkrankung „gelegentlich“ auf (IQWiG-Vorbericht, S. 1, zweiter Absatz). Erst mit zunehmender Schwere der Erkrankung werden sie häufiger (z.B. Srikanth 2005).

Bezüglich der Wirksamkeit von Rivastigmin auf PPS möchten wir u.a. auf den Artikel von Burns 2004 hinweisen, in dem gepoolte Daten aus den vom IQWiG anerkannten randomisierten kontrollierten Studien (RCTs) publiziert wurden (Corey-Bloom 1998; Schneider 1998; Rösler 1999). In dieser Arbeit wird die Wirksamkeit von Rivastigmin auf PPS bei Patienten mit mittelschwerer Alzheimer Demenz (AD), gemessen durch die BEHAVE-AD dargestellt. In 5 von 7 untersuchten Domänen konnte ein Vorteil unter der Gabe von Rivastigmin (6-12 mg/die) gegenüber Placebo gezeigt werden. Von den Zulassungsbehörden wurden psychopathologische Symptome nicht explizit als primärer oder sekundärer Endpunkt bei der leichten bis mittelschweren Alzheimer Erkrankung gefordert und daher auch in den pivotalen Studien nicht als solche definiert. In den pivotalen Studien wurde die BEHAVE-AD, eine Skala zur Einschätzung von Verhaltensauffälligkeiten, in erster Linie zur Validierung der CIBIC-plus Skala eingesetzt und dementsprechend nicht in den oben genannten Artikeln publiziert. Die akkumulierten Daten und Analysen zeigen jedoch einen positiven Effekt von Rivastigmin auf psychopathologische Symptome.

Nicht akzeptabel ist, wie die Daten der RCT von Forette 1999 beurteilt werden; insbesondere in Bezug auf die NOSGER. Das primäre Endziel der Studie war, die maximal tolerierte Dosis für Rivastigmin zu ermitteln und deshalb stand die per protocol Population im Fokus der Analyse. Es wird auch dementsprechend keine Begründung für eine Fallzahl angeführt. Ein sekundäres Endziel war es, die Wirksamkeit in der höchsten individuellen tolerierten Dosis zwischen Verum und Placebo zu vergleichen, unter anderem auch durch die NOSGER Skala. Der Anteil der Studienabbrecher war vergleichbar mit anderen Studien die im Bericht beachtet wurden. Da nur wenige Daten für begleitende Psychopathologie in RCTs vorliegen, ist es wichtig, jegliche Daten in Betracht zu ziehen auch wenn die Analyse nicht in der ITT Population vorgenommen wurde und es auch nur ein sekundäres Endziel war. Daher halten wir es für angebracht, die vorliegenden Resultate für NOSGER dementsprechend zu beurteilen.

Eine weitere nicht adäquat beachtete Studie ist die von Bullock 2005 publizierte, RCT (Head-to-Head Studie; Rivastigmin vs Donepezil). Hier wurde der NPI (Neuropsychiatric Inventory) als Messinstrument zur Untersuchung der Wirksamkeit von Rivastigmin auf PPS eingesetzt. In der Untersuchung an 955 Patienten konnte nachgewiesen werden, dass Rivastigmin über eine Studiendauer von 2 Jahren zu einer klinisch relevanten Veränderung führte. In der Subgruppe der Patienten, die jünger als 75 Jahre waren, konnte darüber hinaus ein statistisch signifikanter Unterschied gegenüber Donepezil nachgewiesen werden, dies sollte auch in der Bewertung der Cholinesteraseinhibitoren (ChE-I) untereinander berücksichtigt werden.

Weiterhin wurden einige Studien zur Wirksamkeit auf PPS nicht beachtet, da es sich nicht um RCTs handelt. Dieses Vorgehen entspricht nicht den internationalen Standards der Nutzenbewertung. Nicht zuletzt das NICE (National Institute for Clinical Excellence) weist darauf hin, dass in eine Nutzenbewertung auch Kohortenstudien und Studien zum Verlauf der Erkrankung eingehen sollten. Weiterhin fordert das NICE hochqualitative Beobachtungsstudien, um Daten aus RCTs zu ergänzen. Dieser Position des NICE schliessen wir uns an. Hier sei insbesondere auf die Studien von Cummings 2005 und Aupperle 2004 hingewiesen. In beiden Studien werden über 24 bzw. 52 Wochen Patienten mit mittelschwerer und schwerer AD hinsichtlich psychopathologischer Symptome unter einer Rivastigmin-Therapie untersucht. In mittelschweren bis schweren Krankheitsstadien ist die Auftretenswahrscheinlichkeit psychopathologischer Symptome größer als in leichten Stadien der Erkrankung (vgl. S. 1 Vorbericht). Rivastigmin zeigt in beiden Studien einen positiven Effekt, in der Studie von Aupperle sogar über einen Beobachtungszeitraum von 52 Wochen.

Stellungnahme Novartis Pharma GmbH (Fortsetzung)



Langzeitstudien

Zum einen werden berechtigterweise Langzeitstudien zu den ChE-I gefordert, zum anderen werden vorliegende Publikationen über eine Beobachtungsdauer von 5 Jahren nicht beachtet. Allein aus ethischen Gründen sind RCTs über eine Studiendauer von mehr als 6 Monaten nicht durchführbar. Ergänzende Arbeiten werden daher auch in Zukunft nicht vorgelegt werden können. Deshalb wurde dem IQWiG eine einarmige 5-Jahres Studie (Small 2005) vorgelegt, die zeigt, dass Patienten, die über einen Zeitraum von 5 Jahren mit Rivastigmin behandelt werden, im Mittel nicht in das schwere Krankheitsstadium mit einem MMSE unter 12 Punkte eintreten.

Notwendigkeit der vollstationären Pflege

Auch zur Untersuchung der Notwendigkeit einer vollstationären Pflege sind Langzeitstudien notwendig, so dass obige Argumentation analog gilt. Eine mögliche Alternative stellt die Analyse bereits vorhandener Daten aus Datenbanken dar. Bezüglich Rivastigmin verweisen wir auf Beusterien 2004: „This study suggests that rivastigmine treatment may be associated with a decreased risk of NHP (Nursing Home Placement) in the US.“ Zu beachten ist in diesem Zusammenhang auch die Studie von Hauber 2000, die auf Basis einer anderen Datenbank-Analyse zeigt, dass die Wahrscheinlichkeit der Institutionalisierung mit fallenden MMSE-Punkten stetig zunimmt. Eine Behandlung mit Rivastigmin, die zu einer Verzögerung der Krankheitsprogression führt, hat somit mittelbar auch eine Verzögerung der Notwendigkeit vollstationärer Pflege zur Folge.

Lebensqualität

Die geforderte Erfassung von Lebensqualität bei AD-Patienten anhand valider Indikatoren erweist sich als schwierig (vgl. Methodenkritik unten). Trotz bestehender Probleme möchten wir in diesem Zusammenhang auf die Studie von Jönsson 2006 hinweisen, die zeigt, dass es klare Zusammenhänge zwischen Veränderungen von Surrogatparametern und Indikatoren der Lebensqualität gibt. So besteht ein signifikanter Zusammenhang zwischen den MMSE-Werten von Alzheimer-Patienten und der Beurteilung der Lebensqualität durch Angehörige anhand des EuroQoL EQ-5D. Eine Verbesserung kognitiver Fähigkeiten ist somit auch mit einer signifikanten Verbesserung der Lebensqualität – gemessen durch valide Indikatoren – verbunden.

Mortalität

Als Zielgröße für die Untersuchung der Wirksamkeit wird unter anderem die Reduktion der Mortalität genannt. Die durchschnittliche Überlebenszeit nach Diagnosestellung der Alzheimer-Krankheit beträgt 10 bis 15 Jahre. In leichten und mittleren Krankheitsstadien ist die Mortalität durch die Alzheimer-Erkrankung nicht unterschiedlich von der Gesamtbevölkerung. Mortalität ist daher keine geeignete Messgröße im Zulassungsbereich von Rivastigmin (leichte bis mittelschwere Alzheimer-Erkrankung) im Sinne der Überprüfung von Therapieeffektivität sondern ausschließlich unter Sicherheitsgesichtspunkten. Richtigerweise wird darauf hingewiesen, dass unter der Therapie mit Rivastigmin (und anderen cholinergen Medikamenten) keine Steigerung der Mortalität beobachtet wurde. Insbesondere in Tabelle 44 des Vorberichtes sollte daher Mortalität aus der Rubrik der Patientenrelevanten Therapieziele zu entfernen.

Karaman et al. 2005

Die 12-Monatsstudie von Karaman 2005 ging in die Bewertung nicht ein, da die Dauer der Placebo-Gabe nicht genannt sei. Diesbezüglich möchten wir auf folgendes Zitat aus der Publikation (Seite 53; Results) verweisen: „In this study, 24 patients with moderate advanced AD received rivastigmine for 12 months, while another 20 patients received placebo for the same period.“

2. Folgende Bewertungen von Originalstudien im Vorbericht sind nicht korrekt.
(siehe Begründung...)

Stellungnahme Novartis Pharma GmbH (Fortsetzung)



Aktivitäten des täglichen Lebens

In der Studie von Bullock 2005 sind hohe Studienabbrüche in beiden Gruppen vermerkt, wobei die Therapieabbrüche höher in der Rivastigmin Gruppe waren. In beiden Gruppen wurde jedoch die ITT-LOCF Population für die Auswertung in Betracht gezogen. Der Vergleich mit einem t-test, den das IQWiG eigenhändig erstellte, muss hierbei bemängelt werden. Die in der Publikation dargestellten LSMEANS der Behandlungsgruppen sind die adjustierten Mittelwerte, die aus dem Varianzanalyse-Modell kommen (z.B. baseline adjustiert). Im Gegensatz zu den (deskriptiven) Mittelwerten der Behandlungsgruppen sind diese LSMEANS nicht unabhängig, und daher führt die Anwendung eines einfachen t-tests auf diesen LSMEANS zu verfälschten Ergebnissen. Die Schätzung der Variabilität der Differenz zwischen den Behandlungsgruppen wird durch die Nichtbeachtung der Korrelation aufgebläht und führt dementsprechend zu einem zu hohen p-Wert. In dem genannten Falle der ADL-Skala beträgt die Kovarianz z.B. 0.31, und der Standard Error beim einfachen t-test wird um ca. 25% aufgebläht. Ein einfacher t-test auf den (unkorrelierten) deskriptiven Mittelwerten der gleichen Skala (13.1 und 15.8 mit SD von 16.4 und 17.1) führt hingegen zu einem p-Wert von 0.016 und bestätigt das Ergebnis der Kovarianzanalyse.

Therapieassoziierte unerwünschte Ereignisse

Zur Auftretenshäufigkeit von unerwünschten Nebenwirkungen unter Rivastigmin, die sich im Wesentlichen auf gastrointestinale Symptome wie Übelkeit und Erbrechen beschränken, ist anzumerken, dass in den begutachteten Studien (Corey-Bloom 1998; Schneider 1998; Rösler 1999) Aufdosierungsschemata verwendet wurden, die eine deutlich schnellere Aufdosierung vorsahen, als dies in der aktuellen Fachinformation der Fall ist. Große Beobachtungsstudien sowie Umstellungsstudien (Auriacomb 2002) zeigen, dass in der klinischen Praxis die Auftretenshäufigkeit von gastrointestinalen Nebenwirkungen sowie die Abbruchraten aufgrund unerwünschter Ereignisse deutlich niedriger liegen als in den oben zitierten RCTs. Des Weiteren muss darauf hingewiesen werden, dass in den oben zitierten RCTs alle Studienpatienten so lange aufdosiert wurden, bis sie die maximale Zieldosis erreichen oder nicht-tolerable Nebenwirkungen auftraten. Ein Vorgehen, das in der klinischen Praxis keine Anwendung findet. Ein weiterer Grund, weshalb Studien wie von Auriacomb 2002 unbedingt berücksichtigt werden sollten. Im Vergleich Rivastigmin vs Donepezil wurden die Ergebnisse von Fuschillo 2001 nicht beachtet, da die zu den einzelnen unerwünschten Ereignissen dargestellten Prozentangaben nicht plausibel seien. Dies ist nicht nachvollziehbar. Vor dem Hintergrund der Ergebnisse der Studien von Fuschillo 2001 und Bullock 2005, die zeigen konnten, dass sich die Anzahl an Patienten mit unerwünschten Ereignissen während der Erhaltungsphase unter Rivastigmin und Donepezil nicht unterscheiden, kann die Einschränkung im Vorbericht, dass Rivastigmin ein höheres Schadenspotential habe nicht nachvollzogen werden. Insbesondere auch deshalb, weil in den Subgruppen der Studie gezeigt werden konnte, dass sich für Patienten mit einer vollständigen Butyrylcholinesterase-Expression eine deutliche Überlegenheit im Bereich der Alltagsaktivitäten zeigte (Bullock 2005). Immerhin zeigten rund 2/3 aller untersuchten Patienten diesen Genotypus, weshalb er als relevanter Parameter anzusehen ist.

Vergleich der Cholinesteraseinhibitoren untereinander

Wie oben bereits erwähnt ist nicht nachvollziehbar, warum die Überlegenheit von Rivastigmin gegenüber Donepezil nicht in dem Maße akzeptiert wird, wie es in den entsprechenden Studien dargestellt wird (Bullock 2005, Bullock 2006, Touchon 2006). In allen genannten Publikationen zum Head-to-Head Vergleich Rivastigmin vs Donepezil konnte zumindest eine numerische Überlegenheit von Rivastigmin bei psychopathologischen Symptomen, Krankheitsstadium, Kognition und den Aktivitäten des täglichen Lebens gezeigt werden (Bullock 2005). In den post-hoc Analysen wurden die Patienten mit vollständiger Butyrylcholinesterase-Expression sowie die Patienten unter 75 Jahren untersucht. In der Gruppe der Patienten mit vollständiger Butyrylcholinesterase-Expression zeigten sich statistisch signifikante Effekte hinsichtlich der Aktivitäten des täglichen Lebens zugunsten der Rivastigmin-behandelten Patienten. Ein Befund, der mit den pharmakodynamischen / neurobiologischen Eigenschaften von Rivastigmin erklärbar ist, da Rivastigmin der einzige Cholinesteraseinhibitor ist, der neben der Acetylcholinesterase auch die Butyrylcholinesterase im Sinne eines dualen Wirkansatzes hemmt (Darvesh 2003). Auch in der Subpopulation der unter 75-jährigen, die die Patienten mit höherer Progredienz repräsentiert, können statistisch signifikante Unterschiede zugunsten Rivastigmin im NPI gesehen werden. Bei Patienten unter 75 Jahre mit vollständiger Butyrylcholinesterase-Expression können



auch in der Kognition signifikante Unterschiede gegenüber Donepezil nachgewiesen werden (Touchon 2006).

Anmerkungen zur projektspezifischen Methodik

Interaktionspotential

Ein wesentlicher Aspekt der Beurteilung des Nutzens eines Medikaments, nämlich das Interaktionspotential wurde in dem vorliegenden Vorbericht nicht beachtet. Rivastigmin wird, im Gegensatz zu anderen Antidementiva, nicht über das Cytochrom P450-System metabolisiert, was einen klinisch hoch relevanten Vorteil eines niedrigeren Interaktionspotentials bietet; von besonderer Bedeutung für ältere Patienten, die sehr häufig einer Polytherapie bedürfen. Auch dieser Aspekt sollte explizit im Abschlussbericht erwähnt werden.

Studien- und Berichtsqualität

Im Vorbericht wird erwähnt, dass die Rater hinsichtlich des Auftretens von unerwünschten Ereignissen nicht ausreichend verblindet waren und sich somit ein Bias auf die Zielparameter ergeben könnte. Hierzu ist anzumerken, dass eine ausreichende Verblindung gewährleistet wurde. So hatten beispielsweise die CIBIC-plus Rater keinen Zugriff auf Patientendaten oder irgendwelche anderen Messdaten, die in den Studien erhoben wurden. Zumindest in Bezug auf die pivotalen Studien sei auch darauf hingewiesen, dass die Zulassungsbehörden keine Verletzung von GCP Guidelines und unzureichende Verblindung bemängelten.

Studien bei Patienten mit Parkinson Demenz und Lewy Körperchen Demenz

Wesentliche RCTs zur Wirksamkeit, Sicherheit und Verträglichkeit von Rivastigmin bei Patienten mit Parkinson Demenz und Lewy Körperchen Demenz finden in dem vorliegenden Vorbericht keine Beachtung (Emre 2004, McKeith 2000). Die Nichtbeachtung dieser Studien erscheint willkürlich, da Studien zu Mischformen einer Alzheimer Demenz und vaskulären Schädigungen des Gehirns zugelassen werden. Außerdem unterscheiden sich die Parkinson Demenz und die Lewy Körperchen Demenz auf der syndromalen Ebene der psychopathologischen Symptome nicht von der Alzheimer Demenz, weshalb sie zumindest in Beurteilung des NPI eingehen sollten.

Lebensqualität

Die Erhebung der krankheitsbezogenen Lebensqualität bei Patienten mit Alzheimer Erkrankung ist methodisch – anders als kognitive und behaviorale Skalen – unzureichend validiert. Gerade weil hierauf in der Einleitung des Vorberichtes verwiesen wird (S. 7/8) sollte aus methodischen Gründen darauf verzichtet werden, dieses Maß mit anderen patientenrelevanten Therapiezielen, wie z.B. Kognition, in gleichen Kontext zu setzen wie in der Übersichtstabelle 44. Die im Vorbericht genannten Arbeiten von Albert beziehen sich in Bezug auf die Lebensqualität auf weit vorangeschrittenen Krankheitsstadien mit Messgrößen wie Null-Aktivität und Null-Affekt, sind sehr alt (Cotrell 1993 und Magaziner 1988) oder noch nicht repliziert (Smith 2005). Auch über die genannten Arbeiten hinaus sind bei den bisherigen Lebensqualitätsskalen die testtheoretischen Gütekriterien nicht erfüllt oder untersucht. Externe und interne Einflussgrößen auf die Bewertung durch Patient und Caregiver sind nicht geklärt. Vorliegende Studien zeigen bereits bei unbehandelten Patienten ein uneinheitliches Bild der Bewertung der krankheitsbezogenen Lebensqualität durch Erkrankte und Angehörige. Die Hälfte der Caregiver von Patienten mit leichter bis mittelschwerer Alzheimer-Erkrankung bewertet die Lebensqualität des Patienten anders als sie einschätzen, dass der Patient dies selbst tun würde (Karlavish 2001). Die Einschätzung der Patienten-Lebensqualität wird beeinflusst von externen und internen Faktoren auf Seiten des Caregivers und nicht nur des Patienten (Karlavish 2004; Zanetti 1999. Anders als mit dem ADAScog zur Messung der kognitiven Leistungen ist daher in Bezug auf die Lebensqualität keine der bisher verfügbaren Skalen als Standard etabliert, so dass diese Messgröße bisher nicht ausreichend evaluiert ist, um in eine Bewertung therapeutischer Effizienz eines Medikamentes einzugehen.

Stellungnahme Novartis Pharma GmbH (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Umfang der Stellungnahme ist nicht größer als **5 DIN A4-Seiten**
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst Maximal **10 DIN A4-Seiten**
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt
- Formblatt zur Stellungnahme vollständig ausgefüllt und **von allen Stellungnehmenden** unterschrieben
- Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln ausgefüllt vor
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Nürnberg, 6.10.06
Ort/Datum

Nürnberg, 6.10.06
Ort/Datum

Nürnberg, 6.10.06
Ort/Datum

Unterschrift 1

Unterschrift 2

Unterschrift 3

Stellungnahme Pfizer Pharma GmbH



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:
*Cholinesterasehemmer bei Alzheimer Demenz
A05-19/A*

Name, Vorname; Titel und Funktion des/der Stellungnehmenden:

- 1. Dr. med. Michael Warmbold, Vice President Medical, Pfizer Pharma GmbH, Karlsruhe*
- 2. Dr. med. Rainer Ziefel, Senior medical advisor ZKS, Pfizer Pharma GmbH, Karlsruhe*
- 3. Mehlis, Hilmar Koordinator Health Care Management, Eisai GmbH, Frankfurt*

Kontakt für weitere Nachfragen:

Name:
Institution:
Anschrift:
Tel./Fax-Nr.:
E-Mail-Adresse:

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

- 1. Im Vorbericht fehlende Originalstudien**
- 2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht**
- 3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)**

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für **alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Karlsruhe / 4.10.06
Ort/Datum

[Handwritten Signature]
Unterschrift 1

[Handwritten Signature]
Unterschrift 2

[Handwritten Signature]
Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

Stellungnahme der Pfizer Pharma GmbH zum IQWiG-Vorbericht A05/19-A: Cholinesterasehemmer bei Alzheimer-Demenz

Sehr geehrte Damen und Herren,

zu den von Ihnen am 8.9.2006 publizierten Vorbericht A05/19-A „Cholinesterasehemmer bei Alzheimer-Demenz“ möchten wir gerne wie folgt Stellung nehmen:

Die Studie von Holmes²⁰ ist eine 12-wöchige, randomisierte, doppelblinde Studie von Donepezil an Patienten mit leichter bis mittelschwerer Alzheimer-Demenz (NINCDS-ADRDA-Kriterien; MMSE 10-26) und neuropsychiatrischen Symptomen (Baseline-NPI-Wert > 11). Der Doppelblind-Phase ging eine 12-wöchige open-label-Behandlung mit Donepezil voraus. 134 Patienten erhielten 6 Wochen lang 5 mg Donepezil, anschließend 10 mg Donepezil für 6 weitere Wochen über einen Gesamtzeitraum von 24 Wochen. Diese Studie liefert wichtige Ergebnisse für die Therapieziele „begleitende Psychopathologie (Verhaltensstörungen)“ und „Betreuungsaufwand“ (im Sinne von „Caregiver Distress“) (siehe unten).

Fazit: Die Studie von Holmes²⁰ muss in die Bewertung eingehen, da sie die formalen, vom IQWiG festgelegten Kriterien einer 24-wöchigen Behandlungsdauer erfüllt.

Beachtungen zum Bewertungsvorbericht

A. Bedeutung der Skalen zur Erfassung der kognitiven Leistungsfähigkeit und des klinischen Gesamteindrucks

Relevante Therapieziele unter Alltagsbedingungen können sich von solchen Kriterien unterscheiden, die für Responder-Analysen herangezogen werden; sie basieren aber auf Veränderungen in denselben zugrunde liegenden Symptombefeldern. Ein klinisch signifikantes Ansprechen auf die Therapie der Alzheimer-Demenz, einer chronisch-degenerativen Erkrankung, umfasst den Erhalt oder die Verzögerung des fortschreitenden Abbaus von Fähigkeiten in den oft voneinander abhängigen Symptombefeldern der Demenz, darunter Kognition, Aktivitäten des täglichen Lebens und Verhalten⁴⁷. Das Ziel der Therapie besteht darin, die Symptomatik in einer für den Patienten relevanten Form zu verbessern oder die Symptomprogression zu verzögern. Von großer Bedeutung ist hierbei der möglichst lange Erhalt der noch vorhandenen Alltagsfähigkeiten der Patienten⁴⁷.

Es gibt mehrere mögliche Arten, den klinischen Nutzen zu erfassen. Eine besteht darin, die Effekte in den einzelnen Symptombefeldern – Kognition, Aktivitäten des täglichen Lebens und Verhalten – möglichst detailliert auszuwerten. Eine andere besteht darin, Skalen zur Erfassung des klinischen Gesamteindrucks zu nutzen, um alle Symptombfelder zu erfassen und um deren klinische Relevanz beurteilen zu können. Obwohl der Nutzen einer Therapie manchmal gleichbedeutend mit Verbesserung gesehen wird, ist der geringer als erwartete Abbau auch ein wichtiges und erfolgreiches Therapieziel bei einer chronisch-degenerativen Erkrankung wie der Alzheimer-Demenz. Obwohl die Verzögerung des Abbaus aufgrund der Behandlung durch den Vergleich mit einer Placebo-Gruppe deutlich beobachtet werden kann, kann der Therapieeffekt im Einzelfall weniger offensichtlich sein; allerdings kann beim einzelnen Patienten eine Reihe von Zielgrößen bewertet werden, wenn der Kliniker mit dem natürlichen Verlauf der Alzheimer-Demenz vertraut ist und sich Kliniker und Patient realistische und klare Therapieziele gesetzt haben. Solche Auswertungen der Therapieziele des einzelnen Patienten wurden in Instrumenten formalisiert wie der Goal Attainment Scale (GAS)³⁵. Mittels der GAS wurde eine statistisch signifikante Erreichung der Therapieziele bei mit Donepezil behandelten Patienten demonstriert³⁸.

A1. Bedeutung der ADAS-Cog und MMSE zur Messung des „patientenrelevanten Nutzens“ einer medikamentösen Therapie bei der Alzheimer-Demenz

Im Vorbericht wird für die Cholinesterasehemmer ein Nutzen hinsichtlich des Therapieziels „Besserung bzw. Erhalt der kognitiven Leistungsfähigkeit“ bestätigt. Zudem wird der Nutzen zum Therapieziel „Besserung bzw. Verhinderung der Beeinträchtigungen der Aktivitäten des täglichen Lebens“ festgestellt. Für einen Einfluss auf andere Bereiche des patientenorientierten Nutzens, z.B. dem „Erhalt der krankheitsbezogenen Lebensqualität“, gebe es hingegen keine Hinweise. Es entsteht durch die Darstellung der Eindruck, dass die Kognition lediglich einen gewissen Teil der Demenzsymptomatik ausmacht und dass andere Bereiche der relevanten Demenzsymptome von der Kognition unabhängig seien. Hierdurch wird unseres Erachtens der elementaren Bedeutung der kognitiven Symptomatik (gemessen z.B. mit der ADAS-Cog) im Gesamtkomplex der Demenzerkrankung nicht hinreichend Rechnung getragen.

Begründung

Der Begriff Kognition stellt eine allgemeine Bezeichnung für einen Komplex von geistigen Fertigkeiten dar, die der Aneignung, der Organisation, der Repräsentation und der Anwendung von Wissen dienen. Die Kognition stellt die Grundlage allen zielgerichteten Handelns dar. Es gibt keinen Bereich des menschlichen Handelns und Fühlens, der von kognitiven Fertigkeiten bzw. Defiziten unabhängig ist.

Die ADAS-Cog, die die kognitive Leistungsfähigkeit erhebt, erfasst verschiedene kognitive Funktionsbereiche, die sich bei der Erkennung und Beschreibung von Demenz-Syndromen als wesentlich erwiesen haben. Diese sieben kognitiven Funktionsbereiche stehen jeweils im Zusammenhang zu bestimmten alltäglichen Verhaltensweisen bzw. stellen die Grundlage für die Möglichkeit des Handelns dar. Zusammenhänge zwischen kognitiver Leistungsfähigkeit und Alltagsverhalten können als gesichert gelten^{7, 10, 11, 21, 33, 34, 37}. Laut Galasko¹³ ist die kognitive Leistungsfähigkeit der wichtigste Faktor, der die Alltagsfähigkeiten bestimmt. Das mittels der Disability Assessment for Dementia Skala (DAD-Skala) bestimmte Niveau der Alltagsfähigkeiten korreliert mit dem ADAS-Cog-Wert ($r = -0,52$)¹². Daher beeinflusst die mittels der ADAS-Cog erfasste kognitive Leistungsfähigkeit ein weiteres, patientenrelevantes Therapieziel, nämlich den Erhalt oder Verlust der Alltagsfähigkeiten. Dies stimmt überein mit der Auffassung, dass Defizite in den Alltagsfähigkeiten in einem bedeutenden Maße eine Folge der eingeschränkten kognitiven Leistungsfähigkeit sind.

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

1. Gedächtnis: Die Fähigkeit zur Informationsspeicherung macht Erlebnisse und Erfahrungen unserem Bewusstsein dauerhaft zugänglich. Die Erinnerung und das Gedächtnis verleihen dem Menschen so sein Empfinden für Kontinuität, für seine Identität und dadurch seine einzigartige Persönlichkeit.

Bei der Alzheimer-Demenz weist der Bereich des Gedächtnisses in der Regel die ersten Anzeichen einer Störung auf. Der Parameter der verzögerten Wiedergabe stellt den sensitivsten Indikator für das Vorliegen einer Demenz dar^{44, 45}.

Insgesamt haben Gedächtnisstörungen gravierende Auswirkungen auf das soziale Leben und die soziale Eingebundenheit der Betroffenen³⁶.

2. Befolgen von Anweisungen: (Überprüfung verschiedener kognitiver Teilleistungsbereiche (Sprachverständnis und Fähigkeit zu einfachen praktischen Handlungen und deren kombinierte Anwendung)). Die Kommunikationsfähigkeit hängt mit der Praxis zusammen und interferiert mit der Selbständigkeit im Alltag¹⁶. Die Beeinträchtigungen haben zur Folge, dass Personen auf dauerhafte Hilfe angewiesen sind. Einfache Aufforderungen oder Erläuterungen beispielsweise bei einem Arztbesuch können so nicht verstanden und umgesetzt werden, es bedarf eines an den Demenzpatienten angepassten Kommunikationsstils²⁵.

3. Abzeichnen: (Überprüfung der Umsetzung von Gesehenem in koordinierte Handlungen (das Konstruieren)). Im Alltag können bereits leichte visuokonstruktive Störungen zu Gefahren beim Autofahren führen. In schwerer Ausprägung können sich die Patienten nicht mehr selbständig anziehen, Kleidungsstücke zusammenzulegen oder den Tisch decken. Eine Einschränkung der selbständigen Lebensführung ist die Folge.

4. Sprachliche Fertigkeiten: Da ein soziales menschliches Miteinander in der Kommunikationsfähigkeit seinen Ausdruck findet, haben Einschränkungen in den sprachlichen Fertigkeiten Einfluss auf den gesamten Lebensalltag der Betroffenen, sie können zum Ausschluss von gesellschaftlicher Teilhabe und zu sozialer Isolierung führen.

5. Ideatorische Praxis: (Überprüfung von mehrschrittigen Handlungen mit mehreren Objekten). Bei einer Störung liegt ein deutlicher Hinweis auf Beeinträchtigungen zur selbständigen Durchführung alltagspraktischer Handlungen vor, die mehrere Teilschritte erfordern, wie eigenständiges Waschen und Duschen oder Kochen.

6. Orientierung: (Überprüfung der zeitlichen, räumlichen und personalen Orientierung) Orientierungsstörungen können gravierende Folgen haben wie folgendes Beispiel zeigt: aufgrund von Gedächtnis- und Orientierungsproblemen hatte sich ein Alzheimer-Patient in einem Berliner Krankenhaus verirrt und war tot im Heizungsraum des Krankenhauses aufgefunden worden⁷.

7. Konzentration/Ablenkbarkeit: Eine gestörte Konzentrationsfähigkeit wirkt sich im Alltagsleben auf verschiedenste Bereiche, v.a. die Durchführung einer beruflichen Tätigkeit, aus.

Über alle aufgeführten Teilleistungsbereiche hinweg kann ein Punktwert von 70 erreicht werden, der die maximale Ausprägung der Störung bedeutet. Wesentlich erscheint, dass die klinische oder Alltagsrelevanz von Veränderungen der ADAS-Cog stets im Kontext des individuellen Leistungsbereichs zu sehen ist. Veränderungen um einzelne Punktwerte sind im Einzelfall immer dann von Relevanz, wenn die Störung einer weiteren kognitiven Funktion hinzutritt, die besonderer Beachtung bedarf, wie beispielsweise eine räumliche Orientierungsstörung. Umgekehrt können einzelne Funktionen diskrete Besserungen aufweisen, die für den Betroffenen bereits spürbar sein können. In den obigen Ausführungen ist für jeden Teilleistungsbereich die alltagspraktische Relevanz dargestellt worden. Jegliche Verbesserung und der Erhalt in den kognitiven Funktionen trägt daher ganz entscheidend zur Erleichterung des alltäglichen Lebens der Patienten bei.

Wie für die ADAS-Cog kann man auch für die MMSE zeigen, dass die mit ihr erfassten kognitiven Funktionsbereiche im Zusammenhang zu bestimmten alltäglichen Verhaltensweisen stehen bzw. die Grundlage für die Möglichkeit des Handelns darstellen. Abbildung 1 (siehe Anhang) zeigt die Korrelation zwischen dem MMSE und den Alltagsfähigkeiten. Geringere Verschlechterungen des MMSE-Werts führen bereits zu spürbaren Verschlechterungen der Alltagsfähigkeiten¹². Die Bedeutung der mit der MMSE erfassten kognitiven Leistungsfähigkeit zeigt sich auch dadurch, dass es gängige Praxis ist, die MMSE nicht nur als reines Kognitionsmaß, sondern auch zur Schweregrad-Einstufung der Alzheimer-Patienten zu verwenden.

Fazit

Die kognitiven Funktionen als Komplex verschiedener geistiger Fähigkeiten bilden die Grundlage aller selbständigen Lebensgestaltung, sie sind insbesondere bei Demenzkranken die Basis für das soziale und praktische Leben. Jede Funktion zeigt dabei spezifische Zusammenhänge zu bestimmten alltäglichen Tätigkeiten. Somit spielen die kognitiven Funktionen im Gesamtkomplex der Symptomatik der Alzheimer-Demenz die zentrale Rolle. Diese zentrale Rolle wird ihnen auch in den Klassifikationssystemen eingeräumt. In den weltweit vorherrschenden Klassifikationssystemen ICD-10 und DSM-IV werden zur Diagnostik als zentrales Kriterium explizit Beeinträchtigungen in mehreren kognitiven Funktionen (Gedächtnis und mindestens eine weitere) gefordert. Der Erhalt der kognitiven Leistungsfähigkeit stellt aus unserer Sicht das prominente patientenrelevante Therapieziel dar.

Die ADAS-Cog war in vielen der entscheidenden klinischen Studien mit Donepezil ein primärer Endpunkt. Dieser leistungsorientierte Test misst die verschiedenen kognitiven Funktionen, die gewöhnlich bei Patienten mit Alzheimer-Demenz eingeschränkt sind, z.B. Gedächtnis, Sprache, Praxis und Orientierung. Die Verwendung des ADAS-Cog als primärer Endpunkt wurde empfohlen, um die für die Zulassung benötigten Anforderungen für den Wirksamkeitsnachweis zu erfüllen. In der klinischen Praxis erfolgt die Beurteilung der Therapieziele globaler und umfasst Verhaltenssymptome, Aktivitäten des täglichen Lebens, die Fähigkeit zur Selbstversorgung und Interaktion mit Familienangehörigen – ein umfassendes klinisches Bild. Dennoch sind diese Bereiche, wie ausführlich dargestellt, nie getrennt von der Kognition, gemessen z.B. mit der ADAS-Cog oder dem MMSE, zu sehen.

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

A2. Bedeutung von Skalen zur Erfassung des klinischen Gesamteindrucks (CIBIC-Skalen) zur Messung von Therapieeffekten bei der Alzheimer-Demenz

In Wirksamkeitsuntersuchungen zur Alzheimer-Demenz kommen Skalen zur globalen klinischen Beurteilung (klinischer Gesamteindruck) neben Skalen zur Bewertung von kognitiven Funktionen und Aktivitäten des täglichen Lebens zum Einsatz. Der Vorteil der zusätzlichen Verwendung von Skalen zum klinischen Gesamteindruck in klinischen Studien ist, dass mit ihnen die Beurteilung des Patienten durch den Kliniker unter Alltagsbedingungen nachgeahmt wird. Grundlage für die Skalen zur globalen klinischen Beurteilung ist die Notwendigkeit, ein Instrument zur Messung einer klinisch aussagekräftigen Veränderung zur Verfügung zu haben. Mit Hilfe der Skalen schätzen erfahrene Kliniker ein, ob und wie sich das klinische Erscheinungsbild eines Patienten aufgrund der Behandlung verändert hat. Bezugspunkt der Beurteilung ist die Einschätzung des Patienten zu Beginn der Behandlung.

Wesentlich ist, dass es sich um klinisch aussagekräftige Veränderungen handelt, d.h., es geht nicht darum, kleine Veränderungen in verschiedenen Bereichen möglichst genau zu erfassen, deren Relevanz aber möglicherweise in Frage gestellt werden könnte. Wird ein Patient auf einer Skala als verändert eingestuft, so ist diese Änderung per definitionem klinisch signifikant⁴⁰.

Nach Vorliegen der ersten Skalen zum klinischen Gesamteindruck, denen die Einschätzung durch den Kliniker zugrunde liegen sollten (Clinical global Impression of Change (CGIC))¹⁸, wurden die globalen Einschätzungen verbessert, indem Struktur und Dokumentation durch den Prüfarzt vorgegeben wurden. In der CIBIC (Clinician's Interview Based Impression of Change)^{27, 28} wird die Einschätzung der Wirksamkeit einer Arzneimittelbehandlung aufgrund eines Interviews mit dem Patienten vorgenommen. Beiträge von anderen Personen werden nicht berücksichtigt, um eine Unabhängigkeit der Beurteilung zu gewährleisten.

Die Firma Parke-Davis entwickelte in Zusammenarbeit mit der FDA im Vorfeld der Zulassungsstudien von Tacrin die CIBI (Clinical Interview Based Impression)²⁴. Ziel war es, ein solides Instrument zu schaffen, das eine klinisch aussagekräftige Veränderung bei dem komplexen Demenz-Syndrom messen kann. Die Methode sollte unabhängig sein von einzelnen Skalen, deren Ziel das genaue Messen von Einzelsymptomen, nicht aber die Einschätzung der klinischen Bedeutung der Symptome für den Patienten ist. Die CIBIC-Skalen sind nicht darauf angelegt, kleine Veränderungen auf der Symptomebene zu messen, die unter Umständen nicht klinisch aussagekräftig sind. In einer CIBIC-Skala stellt der erfahrene Kliniker, der einen Patienten als verändert einstuft, eine „klinisch aussagekräftige und deutliche Änderung“ fest. Es handelt sich demnach um Verfahren, bei denen es keinen Zweifel über die klinische Relevanz der Therapieeffekte gibt, da die klinische Relevanz die eigentliche Messgröße ist.

Signifikante Wirkungen unter anderem auf der CIBI trugen zur Zulassung von Tacrin in den USA bei²³.

Seither sind vergleichbare Maße zur Einschätzung des klinischen Gesamteindrucks Standard in Untersuchungen zur Arzneimittelwirksamkeit bei Demenz. Der Nachweis eines Vorteils gegenüber Placebo, gemessen mit einer Skala zur Messung des klinischen Gesamteindrucks (zumeist CIBIC) ist Voraussetzung für die Zulassung von Antidementiva.

Der Begriff CIBIC+ bezeichnet all jene CIBIC-Skalen, die als Grundlage zur Einschätzung der Wirksamkeit durch den Prüfarzt zusätzlich zu dem Patienteninterview ein Interview mit einem pflegenden Angehörigen bzw. einer betreuenden Person erfordern. Auch hier ist das zugrundeliegende Rational, dass die von einem unabhängigen, erfahrenen Kliniker, der verblindet gegenüber den Ergebnissen der anderen Testverfahren und Untersuchungen ist, beobachteten Veränderungen mit hoher Wahrscheinlichkeit klinisch bedeutsam sind.

Ein Beispiel einer CIBIC+-Skala ist die ADCS-CGIC, die im Rahmen der „Alzheimer's Disease Cooperative Study“ (ADCS) im Auftrag des National Institute on Aging entwickelt wurde⁴¹. Diese CIBIC+-Skala basiert auf zwei ausführlichen, halbstrukturierten Interviews, die die Themenbereiche Kognition, Verhalten, Aktivitäten des Täglichen Lebens und einen allgemeinen Bereich abdecken. Die Dokumentationsbögen enthalten Vorschläge für abzufragende Stichpunkte und ausreichend Raum für Notizen des Interviewers. Festgehaltene Informationen können so zur Verlaufsbeurteilung herangezogen werden. Die Interviews sind unabhängig von anderen Testverfahren und Untersuchungen durchzuführen. Gute Reliabilitäten und Validitäten wurden nachgewiesen⁴².

Die Clinician's Global Impression of Change ist eine Methode, die häufig in klinischen Studien quer durch alle Therapiegebiete eingesetzt wird. Definitionsgemäß ist die globale Veränderung klinisch bedeutsam³⁹. Im Wesentlichen bewertet der Kliniker den Patienten im Vergleich zum Anfangszustand und bestimmt, ob der Zustand des Patienten sich (minimal, mäßig oder deutlich) gebessert hat, unverändert geblieben ist bzw. sich (minimal, mäßig oder deutlich) verschlechtert hat. Diese klinischen Beurteilungen lassen sich mit einer Likert-Skala beschreiben⁴², wobei statistische Tests durchgeführt werden können, um die Verteilung der Punktwerte auf die Studienpopulationen zu vergleichen oder die Trends im Zeitverlauf auszuwerten. Die wesentliche Stärke dieser Bewertungen besteht darin, dass sie sich dem Verfahren annähern, das Kliniker in der realen Praxis zur Bewertung ihrer Patienten anwenden. Solche Instrumente wurden für den Einsatz bei Alzheimer-Patienten validiert^{24, 42} und sind – gemeinsam mit spezifischeren Messungen, wie z.B. der ADAS-Cog, der MMSE, den Skalen zur Bestimmung der Alltagsfähigkeiten und dem NPI –, eine wertvolle Ergänzung des Instrumentariums klinischer Studien.

Fazit

Die Alzheimer-Demenz besteht aus einer komplexen Symptomatik, die ausgehend von Einschränkungen der geistigen Leistungsfähigkeit keinen Lebensbereich unbeeinträchtigt lässt. Mit Skalen zur Messung des klinischen Gesamteindrucks wurden Instrumente geschaffen, die Zustandsveränderungen valide messen können und unabhängig sind von Tests zur Messung spezifischer Symptome. Es handelt sich um Verfahren, bei denen es keinen Zweifel über die klinische Relevanz der Therapieeffekte gibt, da die klinische Relevanz die eigentliche Messgröße ist.

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

Die CIBIC korreliert mit den mittels der ADAS-Cog und der MMSE beobachteten Veränderungen, ist aber weniger sensitiv gegenüber Veränderungen²⁴. Daher ist sie möglicherweise ein konservativeres Messinstrument als die spezifischeren Skalen.

Die CIBIC-Skalen stellen das Ergebnis einer Diskussion um die klinische Relevanz von Testergebnissen dar, die Anfang der 90er Jahre zwischen der FDA und der Firma Parke-Davis begonnen wurde. Seither sind die CIBIC-Skalen mit ihren Weiterführungen (CIBIC+, ADCS-CGIC) Standard, nach dem Alzheimer-Medikamente geprüft werden.

Eine CIBIC-Skala ist ein Verfahren, das den Anforderungen zur Nutzenmessung des IQWiG mit am besten entspricht, denn es wird die klinische Relevanz von Therapieeffekten valide gemessen. Es ist daher unverstänlich, dass zu einem Zeitpunkt, als die Nutzenbewertung der Cholinesterasehemmer bereits sehr weit fortgeschritten war, durch das Amendement zum Berichtsplan A05/19A eine Herabstufung der Bedeutung des klinischen Gesamteindrucks dahingehend erfolgte, dass dieser nicht primär in die Nutzenbewertung des IQWiG eingeht.

Die Verbesserung des klinischen Gesamteindrucks durch Donepezil ist eindeutig nachgewiesen. Dadurch wird für ein weiteres patientenrelevantes Therapieziel der Nutzen von Donepezil deutlich. Dies muss in das Fazit des IQWiG-Berichtes aufgenommen werden.

A3. Klinische Bedeutung der Wirksamkeitsergebnisse anhand der Number Needed to Treat (NNT; Anzahl der notwendigen Behandlungen)

Die klinische Bedeutung der Ergebnisse zur Wirksamkeit von Donepezil lässt sich ebenfalls zeigen, wenn man die NNT betrachtet. In einer Metaanalyse betrug die NNT 12 (95%-KI 9-16) für Cholinesterasehemmer im Vergleich zu Placebo bei Alzheimer-Demenz für ein globales Ansprechen in Form einer „Verbesserung“ (ohne Berücksichtigung einer Stabilisierung oder eines geringer als erwartet ausfallenden Abbaus)²⁵. Für eine Verbesserung auf der ADAS-Cog um 4 Punkte (wiederum ohne Berücksichtigung einer Stabilisierung oder eines verzögerten Abbaus) war die NNT = 10 (95%-KI 8-15). Eine weitere Analyse zeigte, dass Donepezil eine sehr kleine NNT aufweist, um die Verschlechterung der kognitiven oder globalen Funktion hinauszuzögern oder umzukehren: für Donepezil 5 mg NNT = 5-15; für Donepezil 10 mg NNT = 4-8³⁰. Als Vergleichswerte können die folgenden NNTs für Therapieformen, die aktuell den Standard in der Versorgung bei anderen Krankheiten darstellen, herangezogen werden: Trizyklische Antidepressiva bzw. Selektive Serotonin-Wiederaufnahme-Hemmer (SSRI) zur ambulanten Behandlung der Depression: NNT = 4 bzw. 6³¹; Fluoxetin (ein SSRI) bei einer Reihe von psychiatrischen Erkrankungen: NNT = 4-16³²; Antihypertensiva zur 5-Jahres-Prävention von Schlaganfall, Herzinfarkt oder Tod: NNT = 29-86³³; Betablocker zur Mortalitätsprävention bei Patienten mit Herzinsuffizienz: NNT = 38³⁴; Influenza-Impfung zur Prävention der Grippeerkrankung: NNT = 23³⁵. Der Vergleich zeigt, dass Donepezil NNT-Werte in Bezug auf seine Wirksamkeit aufweist, die bei einer Reihe von Erkrankungen als weitgehend akzeptabel beurteilt werden würden.

Die Ergebnisse der Metaanalysen des IQWiG zeigen, dass die Effektgröße bei Donepezil bezüglich der kognitiven Funktion (Cohen's d = -0,51) und der Aktivitäten des täglichen Lebens (Cohen's d = -0,28) im mittleren Bereich liegt.

Fazit: Die Therapie mit Donepezil führt zu einem relevanten klinischen Nutzen.

B. Wirksamkeit von Donepezil bezüglich der einzelnen Therapieziele

B1. Lebensqualität

Wir stimmen dem IQWiG zu, dass die Erfassung der Lebensqualität der Alzheimer-Patienten in klinischen Studien schwierig ist, da bisher kaum geeignete und validierte Messinstrumente existieren⁴³ und die Patienten sich nur eingeschränkt selbst beurteilen können. Um den Hinweisen aus den vorliegenden klinischen Studien bezüglich des Einflusses auf die Lebensqualität der Erkrankten und ihrer Angehörigen nachzugehen, wurden 179 Alzheimer-Patienten und 301 Angehörige gefragt, welche Faktoren ihre Lebensqualität beeinflussen¹. Die seit September 2006 vorliegenden Ergebnisse dieser vom Kompetenznetz Demenzen und der Alzheimer Gesellschaft Baden-Württemberg durchgeführten Befragung lauten: Die Lebensqualität der Patienten wird von den Befragten umso höher eingeschätzt, je geringer die Krankheitsprogression ist. Die Lebensqualität der Angehörigen wird besonders stark durch die kognitiven Fähigkeiten, Alltagsfähigkeiten und depressiven Symptome der Erkrankten beeinflusst¹.

Der Nutzen von Donepezil bezüglich der Verbesserung bzw. dem Erhalt der kognitiven Fähigkeiten und den Alltagsfähigkeiten sowie der Besserung der begleitenden psychopathologischen Symptomen wurde in den klinischen Studien gezeigt. Daraus lässt sich ableiten, dass Donepezil auch die Lebensqualität der Angehörigen positiv beeinflusst.

Bei einer Befragung von Alzheimer-Patienten und deren Angehörigen durch die Alzheimer Gesellschaft Großbritanniens^{3,4,5} bzw. Australiens² war die Verzögerung der Krankheitsprogression der am häufigsten bzw. am zweithäufigsten genannte Nutzen einer Behandlung mit Antidementiva. Die Mehrheit der britischen und australischen Befragten hatte Erfahrung mit Donepezil (77 % bzw. 70 %), die übrigen Befragten mit den beiden anderen Cholinesterasehemmern und Memantine (in Australien hatten 2 % der Befragten Erfahrung nur mit weiteren Demenzmitteln). Die Verbesserung der Lebensqualität der Patienten durch die Antidementiva spiegelt sich v.a. in den folgenden, häufig genannten Freitext-Antworten der Befragten wider (die Freitext-Antworten wurden in Kategorien zusammengefasst): Die Patienten sind glücklicher und aktiver, zeigen Interesse an bestimmten Dingen und können besser kommunizieren und schlafen (Anzahl der Befragten mit den einzelnen Antworten: siehe Anhang: Tabelle 1)^{2,3,4,5}. **Fazit:** Aus den Ergebnisse der drei Befragungen^{1,2,3,4,5} zur Krankheitsprogression und zu den einzelnen Lebensqualitätsaspekten lässt sich ableiten, dass Donepezil (und die anderen drei Antidementiva) unter Alltagsbedingungen einen Nutzen bezüglich der Lebensqualität der Patienten haben.

Zusätzlich zur Lebensqualität zeigen die Antworten der britischen und australischen Befragten den Nutzen von Donepezil (und den anderen drei Antidementiva) unter Alltagsbedingungen auch bezüglich Kognition, Alltagsfähigkeiten und begleitenden psychopathologischen Symptomen. Insgesamt hielten mindestens 70 % der Befragten die Behandlung mit Donepezil (und den anderen drei Antidementiva) für wirksam^{2,3,4,5}.

Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

Darüber hinaus zeigt eine drei-monatige Anwendungsbeobachtung in Deutschland, dass nach Meinung der behandelnden Ärzte Donepezil in 70 % der Fälle die Lebensqualität der Patienten und /oder Angehörigen verbessert²².

Fazit: Die bisher vorhandene Evidenz weist drauf hin, dass Donepezil die Lebensqualität der Patienten und /oder Angehörigen verbessert.

B2. Notwendigkeit einer Pflegeheimübersiedlung

Anders als im Vorbericht dargestellt liegen Daten vor, die eine positive Beeinflussung der Notwendigkeit einer Pflegeheimübersiedlung durch Donepezil zeigen.

In der Studie von Winblad⁴⁶ wurden weniger Donepezil-Patienten (2 %) als Patienten der Placebogruppe (6 %) in ein Pflegeheim eingewiesen. Dieser Unterschied war noch nicht signifikant ($p=0,13$), da die statistische Power für diesen sekundären Endpunkt zu gering war und weil die Wahrscheinlichkeit einer Pflegeheimweisung für Patienten in einem frühen Stadium der Alzheimer-Demenz (Baseline-MMSE 19) gering ist¹⁹ und gleichzeitig der Zeitraum, in dem das Fortschreiten der Erkrankung beobachtet wird (1 Jahr), für die Beurteilung der Zielgröße „Pflegeheimweisung“ relativ kurz ist. Da mit der Zulassung nachgewiesen wurde, dass Donepezil ein wirksames Medikament ist, ist es aus ethischen Gesichtspunkten fragwürdig, eine placebokontrollierte Studie über einen Zeitraum durchzuführen, der eine Abschätzung des Einflusses der Cholinesterasehemmer-Behandlung auf das Therapieziel „Institutionalisierung“ zulässt. Daher wurden in der Studie von Geldmacher 2003¹⁴ die Daten von Patienten verwendet, die zuerst an einem von drei randomisierten, doppelblinden, placebokontrollierten Studien (Dauer 12-24 Wochen) und anschließend an zwei Open-Label-Fortsetzungen dieser Studien über einen Zeitraum von bis zu 254 Wochen teilnahmen. Durch die konsequente Donepezil-Behandlung wurde eine erste demenz-bedingte Heimübersiedlung im Vergleich zu unzureichend behandelten Patienten im Median um 21,4 Monate – also fast zwei Jahre – verzögert.

Patienten der REAL.FR-Studie¹⁵, die einen beliebigen Cholinesterasehemmer für mindestens 1 Jahr einnahmen, mussten nur mit 1/5 der Wahrscheinlichkeit unbehandelter Patienten im ersten Jahr der Nachbeobachtungsphase in ein Pflegeheim übersiedeln (OR = 0,20; 95%-KI 0,08-0,48; $p < 0,001$).

Die Studie von Lopez^{31, 32} zeigt, dass die Behandlung mit Cholinesterasehemmern das Risiko einer Heimübersiedlung dramatisch senkt (RR=0,06 für 2 Jahre Nachbeobachtungsphase, RR=0,12 für 3 Jahre Nachbeobachtungsphase)

Die drei letztgenannten Studien^{14, 15, 31, 32} wurden jedoch aufgrund des vom IQWiG geforderten Evidenzlevels (ausschließlich RCTs) vom Bewertungsprozess ausgeschlossen. Da aber aufgrund der ethischen Bedenken RCTs für diese Fragestellung praktisch nicht durchführbar sind, sollte für die Beurteilung auf die vorliegenden Daten eines niedrigeren Evidenzlevels zurückgegriffen werden, wie es auch bei der Nutzenbewertung von nichtmedikamentösen Verfahren praktiziert wird.

Fazit: Ein RCT und drei zu berücksichtigende Studien eines niedrigeren Evidenzlevels zeigen die günstige Beeinflussung der Notwendigkeit einer Pflegeheimübersiedlung durch Donepezil.

B3. Betreuungsaufwand

Wir bitten Sie, die missverständliche Darstellung zum Betreuungsaufwand in der Studie von Gauthier zu ändern. Die Pflegezeit war zu Studienbeginn in der Donepezil-Gruppe um 49 Minuten länger als in der Placebo-Gruppe, allerdings wurde dies in der statistischen Auswertung berücksichtigt. Zum Studienende war die Pflegezeit in der Donepezil-Gruppe um 52 Minuten ($p=0,004$) kürzer als in der Placebo-Gruppe. Der Betrag der Differenz entspricht dem zu Beginn beobachteten Unterschied zwischen den Gruppen, aber die Differenz hat ein anderes Vorzeichen.

Das IQWiG interpretiert die Ergebnisse der Winblad-Studie zum Betreuungsaufwand nur eingeschränkt, da für eine Hospitalisierung der Patienten die Betreuungszeit der pflegenden Angehörigen auf den Wert 0 gesetzt wurde und somit die Ermittlung der Betreuungszeit verzerrt werden könnte. Während der Winblad-Studie gab es bei 9 von 96 Donepezil-Patienten und 12 von 94 Placebo-Patienten eine Veränderung der Wohnverhältnisse. Die Mehrheit dieser Patienten, 5 der 9 Donepezil-Patienten und 7 der 12 Placebo-Patienten brach die Studie zu diesem Zeitpunkt ab (siehe „Response to reviewer’s comments“). Sowohl bezüglich der Gesamtzahl als auch bezüglich der Nicht-Studienabbrechern (Donepezil: 4 von 94, Placebo: 5 von 96) ist die Anzahl der Patienten mit einer Veränderung der Wohnverhältnisse in der Placebogruppe größer. Somit wurde in der Placebogruppe die Betreuungszeit öfter auf den Wert 0 gesetzt. Die daraus resultierende Verzerrung ist also zum Nachteil von Donepezil. Die ermittelte Verringerung der Pflegezeit gegenüber Placebo um 64 Minuten **pro Tag** stellt somit eine Unterschätzung des positiven Effekts von Donepezil dar.

Auch die Ergebnisse der Gauthier-Studie werden nur eingeschränkt interpretiert, da in der Publikation von Gauthier das Problem der Hospitalisierung nicht dargestellt wird. In der Gauthier-Studie wurde für die Fälle, für die keine Pflegezeit berichtet worden war (z.B. weil sich der Patient im Krankenhaus befunden hatte), der LOCF-Ansatz verwendet (Persönliche Mitteilung des für die Analyse der Betreuungszeit zuständigen Statistikers). Der LOCF-Ansatz stellt eine geeignete Methode dar.

Darüber hinaus wurde auch in der Studie von Holmes²⁰ gezeigt, dass der Betreuungsstress der pflegenden Angehörigen, gemessen mit dem NPI-D, für die mit Donepezil behandelten Patienten signifikant abnimmt.

Fazit: Die Einschränkung der Interpretation der Studien von Winblad und Gauthier ist nicht gerechtfertigt. Diese beiden Studien und die Studie von Holmes zeigen, dass Donepezil den Betreuungsaufwand verringert.

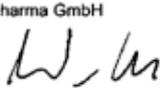
Stellungnahme Pfizer Pharma GmbH (Fortsetzung)

A. Nicht berücksichtigte Evidenz

Aufgrund der vom IQWiG gewählten Ein-/Ausschlusskriterien sind wichtige Studienergebnisse nicht berücksichtigt (siehe Punkt 2: bei den Therapiezielen „Lebensqualität“, „Pflegeheimübersiedlung“ und „Betreuungsaufwand“ weisen wir auf weitere, relevante Studien und Befragungen von Patienten und Angehörigen hin).

Pfizer Pharma GmbH

ppa.



Dr. med. M. Warmbold
Vice President Medical

i. V.



Dr. R. Zerfaß
Senior Medical Advisor ZNS

Eisai GmbH, Frankfurt



H. Mehlig
Kordinator Health Care Management

Stellungnahme privat, Grass-Kapanke



EINGEGANGEN 06. Okt. 2006
Eingangsstempel

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

Vorbericht AChE-Hemmer bei Alzheimerdemenz
A05119-A

Name, Vorname, Titel und Funktion des/der Stellungnehmenden

1. Grass-Kapanke, Brigitte Dr. phil. Dipl. - Psych.
- 2.
- 3.

Kontakt für weitere Nachfragen:

Name: Dr. B. Grass-Kapanke
Institution:
Anschrift: Lerchenweg 20, 40822 Mettmann
Tel./Fax-Nr.:
E-Mail-Adresse: grass_kapanke@yahoo.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

1

Stellungnahme privat, Grass-Kapanke (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für **alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

MeHmann, 05.10.06
Ort/Datum

BKA
Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme privat, Grass-Kapanke (Fortsetzung)

**Stellungnahme zum Vorbericht (Vorläufige Nutzenbewertung, Auftrag A05/19-A)
Cholinesterasehemmer bei Alzheimerdemenz v.08.09.2006**

Aufgrund meiner langjährigen Arbeit mit Demenzpatienten als Psychologin in einer Gedächtnissprechstunde begrüße ich die Auseinandersetzung des IQWiG mit der Substanzgruppe der Antidementiva.

Durch die Diskussionen der letzten Zeit sind niedergelassene Ärzte, insbesondere die Allgemeinmediziner, zunehmend hinsichtlich der Wirksamkeit und Verträglichkeit der Substanzen verunsichert. Dies führt zu einer erheblichen Belastung von Patienten und Angehörigen.

Im Rahmen meiner Arbeit konnte und kann ich die positiven Effekte der Antidementiva erfahren.

Im Folgenden möchte ich einige Kritikpunkte und Anregungen zu dem Vorbericht äußern:

- Grundsätzlich sei darauf verwiesen, dass die kritische Bewertung von NICE nicht die Wirksamkeit der Antidementiva bezweifelt, sondern deren Wirtschaftlichkeit in Frage stellt.
- Aufgrund der anerkannten Wirksamkeit von Antidementiva genehmigen deutsche Ethikkommissionen keine placebokontrollierten Studien mehr, da davon auszugehen ist, dass den Patienten unter Placebo ein wirksames Medikament zu Behandlung ihrer Erkrankung vorenthalten wird.
- Die Alzheimerkrankheit ist eine neurodegenerative Erkrankung. Unbehandelt ist mit einer kontinuierlichen Progression der Beschwerden zu rechnen. Anders als bei den meisten internistischen Erkrankungen ist es unmöglich die Wirkung der Therapie mit nur einem einfachen Parameter (z.B. Blutdruck, Blutzucker) zu messen. Deshalb erfordert die Alzheimerkrankheit eine spezielle Methodik und die Ergebnisse sind vor diesem besonderen theoretischen Hintergrund zu interpretieren. Einzelfallbeobachtungen erlauben keine Beurteilung der Wirksamkeit. Selbst eine Verschlechterung der Patienten unter Therapie besagt nichts über eine vermeintliche Wirkungslosigkeit. Hier können nur Gruppenvergleiche zu validen Aussagen führen. In den entsprechenden Studien konnte diese Wirksamkeit – auch bei bestehenden methodischen Mängeln – belegt werden.
- Wenig nachvollziehbar erscheint mir der Ausschluss von 3-Monatsstudien für die Bewertung der Substanzen. Effekte, die in diesem relativ kurzen Zeitraum nachgewiesen werden können, sind nicht weniger bedeutsam. Dies umso mehr als aus Studien bekannt ist, dass zu Beginn einer Studie oft auch die Placebogruppe eine Verbesserung der Symptomatik zeigt, die dann im Verlauf wieder abfällt. Unter diesem Aspekt sind Behandlungseffekte zwischen Placebo und Verum in kurzzeitig angelegten Studien eher schwieriger nachzuweisen.
- Auf Seite 3 des Berichts wird eine Stadieneinteilung der Demenz in drei Stadien anhand von MMSE bzw. GDS erwähnt. Der MMSE wird in Studien zwar häufig als Einschlusskriterium und auch als Einstufungsinstrument für die Schwere einer Demenz verwendet, seine Eignung hierzu ist aber umstritten. Die GDS wiederum ist ein anerkanntes Maß zur Beschreibung des Schweregrads demenzieller Erkrankungen, basiert jedoch auf 7 Schweregraden von 1 (unauffällig) über 2 (subjektive Beschwerden) zu 3 – 7 (beginnende bis sehr schwer ausgeprägte Demenz). Entsprechend der GDS werden demenzielle Erkrankungen also in 5 Stadien unterteilt, in Studien werden in der Regel die Stadien 3-4 (leichte bis mittelschwere Demenzerkrankungen) eingeschlossen.

Stellungnahme privat, Grass-Kapanke (Fortsetzung)

- Sicher ist es wünschenswert den Einfluss der Behandlung mit Antidementiva auf die Lebensqualität betroffener Patienten zu erfassen (Vorbericht S. 171). Allerdings setzt die Symptomatik der Erkrankung selbst hier enge Grenzen (vgl. Vorbericht S. 8). Aus der psychologischen Forschung ist jedoch bekannt, dass Lebenszufriedenheit ein subjektives Konstrukt ist und nicht mit objektiven Maßen wie Krankheit, finanziellem Status usw. korreliert. Demnach wäre die Objektivierung der Lebenszufriedenheit nicht nur ein angemessenes Outcome für medikamentöse Studien, sondern auch für die dringend erforderliche Evaluation psychologischer Interventionen.

Eine kürzlich in unserem Hause durchgeführte Studie zur Lebenszufriedenheit, die wir demnächst auf der DGPPN vorstellen werden, von Patienten mit MCI bzw. leichter Demenz zeigte keinen Zusammenhang der Lebenszufriedenheit mit dem Ausmaß der kognitiven Defizite.

- Die möglicherweise deutlicheren Therapieeffekte bei schwerer erkrankten Patienten sind am ehesten auf die messtechnischen Umstände zurückzuführen. So kann nicht davon ausgegangen werden, dass die zugrunde liegenden psychometrischen Skalen eine lineare Abbildung der Erkrankungsschwere liefern. Vielmehr ist davon auszugehen, dass ab einer gewissen Ausprägung der Demenz die Verschlechterung psychometrisch besser abgebildet werden kann und damit auch ein größerer Punkterange für Verbesserungen zur Verfügung steht. Geringere Punktwertverbesserungen im frühen Demenzbereich spiegeln also u. U. nur die besondere „Wertigkeit“ dieser Punkte wieder.

Der Vorbericht lässt erkennen, dass die Eigenschaften der verwendeten Skalen und damit die Interpretation ihrer Ergebnisse oft falsch eingeschätzt werden, deshalb hier eine kurze Charakteristik, der wichtigsten Messinstrumente:

- Die ADAS-cog ist eine Skala, die speziell zur Erfassung der Defizite bei Alzheimerdemenz entwickelt wurde. Sie greift mit ihren Items die typischen Symptome der Erkrankung ab und ist allgemein anerkanntes Verfahren zur Behandlungsevaluation von Demenzen. Die ADAS wurde von der FDA als „Goldstandard“ für die Erfassung kognitiver Veränderungen im Rahmen klinischer Studien vorgegeben.

Der Punkterange der ADAS geht von 1-70, wobei 0 Punkte einer perfekten Leistung und 70 Punkte dem schlechtesten Ergebnis entsprechen. Da in Studien üblicherweise Patienten der GDS-Schweregrade 3-4 eingeschlossen werden, liegt der Punkterange von Studienpatienten beim Einschluss üblicherweise zwischen 15 und 25 Punkten. Keinesfalls kann bei der ADAS von einem linearen Verfahren ausgegangen werden. Selten werden Punktwerte von über 30 oder gar 40 erreicht. Patienten dieser Demenzausprägung sind an der Grenze der Fähigkeit zur Durchführung kognitiver Testverfahren und werden mit der ADAS nicht mehr adäquat erfasst. Somit ist der Punkterange von 1-70 als theoretisch anzusehen, reale Werte schwanken überwiegend im Bereich 15 – 30, wobei 30 bereits einer erheblichen Ausprägung demenzieller Beschwerden entspricht (Deckeneffekt des Verfahrens). Auf diese Problematik wird im Bericht in einer Fußnote auf Seite 166 auch hingewiesen, sie soll hier nochmals deutlich unterstrichen werden.

Aus dem o.g. ergibt sich die Bedeutsamkeit der in den Studien erfassten Veränderungswerte der ADAS von etwa 3 Punkten. Dies entspricht im realen Punkterange der Skala einer Verbesserung um 15 Prozent. Die von Kaduszkiewicz et al. (2005) als bedeutungslos dargestellte Veränderung beruht vermutlich auf mangelnder Kenntnis in der Anwendung der diskutierten Skalen.

Stellungnahme privat, Grass-Kapanke (Fortsetzung)

• Der MMST wurde von Folstein et al. (1975) als erstes standardisiertes Verfahren vorgestellt. Der MMST fand aufgrund seiner pragmatischen Vorteile schnell weite Verbreitung, die psychometrische Qualität allerdings ist nicht zufriedenstellend. Die geringe Sensitivität für beginnende Demenzen ist vielfach belegt, mit ihm gelingt weder die Schweregradmessung noch die Verlaufsbeurteilung überzeugend.

• Globale Skalen wie die CIBIC sind sinnvoll und in der klinischen Forschung weit verbreitet. In Studien zu Demenzerkrankung wird häufig die um ein Angehörigeninterview erweiterte CIBIC-plus verwandt. Wesentliche Begrenzung dieses Instrumentes liegt in dem notwendigen „Expertenwissen“ zur reliablen Bewertung. Die Qualität der Daten steht somit in direktem Zusammenhang mit der Qualifikation der Rater.

Für die Diskussion ist zu beachten, dass der Range dieser Skala von 1 (viel besser) über 4 (unverändert) bis 7 (viel schlechter) reicht. Bei einer primär degenerativen Erkrankung wie der Alzheimerdemenz ist bei einem natürlichen Verlauf nicht mit einer Verbesserung zu rechnen, der Range der möglichen Bewertungen ist also auch in dieser Skala deutlich reduziert. Dadurch erhält die in Studien erfasste mittlere klinische Verbesserung um .26-.54 Punkte einen deutlich höheren Stellenwert als von Kaduszkiewicz et al. (2005) wahrgenommen.

• ADL-Skalen werden üblicherweise von den Caregivern der betroffenen Patienten ausgefüllt. Hierzu ist zu vermerken, dass diese häufig Probleme mit der Beantwortung der Fragen haben. Regelmäßig sind unsichere Rückfragen zu erwarten, was mit den einzelnen Items erfasst wird. Auch die Bewertungsmaßstäbe sind nicht so selbsterklärend wie gewünscht. Bei der B-ADL beispielsweise soll die Ausprägung der Schwierigkeiten auf einer Skala von 1 – 10 geratet werden. Hier ist ein häufiger Fehler von Caregivern die Invertierung der Wertung bei der Beantwortung, teilweise auch nur bei einzelnen Items. Hinzu kommt die Tatsache, dass bei Studieneinschluss die Angehörigen oftmals erst kürzlich mit der Diagnose einer Alzheimerdemenz konfrontiert wurden und häufig noch verdrängende Bewältigungsstrategien anwenden. Dies führt gegebenenfalls zu einer zu positiven Beschreibung der Defizite. Mit zunehmender Auseinandersetzung mit der Erkrankung kommt es in der Folge zu realistischeren Einschätzungen der Erkrankungsfolgen und damit höheren Scores in ADL-Skalen, die nicht zwingend für eine tatsächliche Verschlechterung der Alltagsfähigkeiten sprechen. Die in den Studien gemessenen positiven Effekte auf die Alltagsaktivitäten sind also gegebenenfalls unterschätzt.

Auf die möglichen Auswirkungen durch Wahrnehmungsstörungen oder –verzerrungen durch Patienten, Angehörige und Behandler auf die Messinstrumente wird entsprechend im Bericht auf Seite 6 hingewiesen.

• Das NPI ist als beschreibendes Verlaufsinstrument entwickelt worden, welches die begleitende Psychopathologie erfasst. Die mit diesem Verfahren erfassten Auffälligkeiten treten in verschiedenen Krankheitsphasen auf. So ist bei beginnenden Demenzen eher mit depressiven Symptomen zu rechnen, während in späteren Krankheitsphasen psychotisch gefärbtes Erleben und Halluzinationen auftreten können. Gesamtscores spiegeln diesen Verlauf aber nicht wieder und lassen entsprechend keine reliablen Aussagen zu, empfehlenswert sind hingegen Analysen der Subskalen.

Die diskutierten Schwierigkeiten der verwendeten psychometrischen Verfahren lassen erkennen, dass es sich um Messinstrumente handelt, zu deren reliabler Durchführung Expertenwissen notwendig ist. Ein entscheidendes Qualitätsmerkmal klinischer Studien stellen somit Ratertrainings dar. Ein darüber hinausgehendes Mittel zur Erhöhung der

Stellungnahme privat, Grass-Kapanke (Fortsetzung)

Reliabilität ist der Einsatz „externer Rater“. Hierbei handelt es sich um entsprechend geschulte Neuropsychologen, die nicht in den Zentren beschäftigt sind sondern sie lediglich zur Durchführung der Visiten aufsuchen.

In dem Schlussabsatz der Diskussion fordert das IQWiG Studien längerer Laufzeit unter Einsatz valider Erfassungsinstrumente. Die Entwicklung und Validierung solcher Instrumente ist aufwändig und wird in der derzeitigen Forschungsförderung nicht entsprechend unterstützt. Wenn die Initiative des IQWiG als Appell zur gezielten Forschungsförderung psychometrischer Testverfahren bei der Demenzdiagnostik anzusehen ist, ist dies natürlich zu begrüßen.

Auch wenn die Forderung nach längeren Studien mit besseren Instrumenten zu unterstützen ist, so darf im Umkehrschluss jedoch nicht davon ausgegangen werden, dass die vorhandenen Methoden keine validen Ergebnisse liefern.

Sie belegen sehr deutlich die Wirksamkeit der AChE-Hemmern wie eine ganze Reihe von Meta-Analysen zeigen, auch jene der unabhängigen Chochrane-Gruppe.



Dr. phil. Dipl.-Psych. Brigitte Grass-Kapanke



EINGEGANGEN: 06. Okt. 2006

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer:

A 05 / 19 - A

Name, Vorname, Titel und Funktion des/der Stellungnehmenden:

1. Kaduszkiewicz, Hanna, Dr. med. wiss. Angestellte
2. Zimmermann, Thomas, Dipl.-Psych. D. phil. wiss. Angestellter
3. Van den Bussche, Hendrik, Prof. Dr. med. Direktor Institut für Allgemeinmedizin

Kontakt für weitere Nachfragen:

Name: Kaduszkiewicz, Hanna
Institution: Institut für Allgemeinmedizin
Anschrift: Martinistr. 52, 20246 HAMBURG
Tel./Fax-Nr.: 040-42803-3247 Fax: -3681
E-Mail-Adresse: kaduski@uke.uni-hamburg.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme Universitätsklinikum Hamburg-Eppendorf, Kaduszkiewicz (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

- Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.
- Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.
- Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.
- Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.
- Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln **ausgefüllt** vor.
(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Hamburg, den 4. 10. 06
Ort/Datum

K. Uederschick
Unterschrift 1
H. Zimmermann
Unterschrift 2
R. Kater
Unterschrift 3

Vielen Dank für Ihre Stellungnahme.



Universitätsklinikum
Hamburg-Eppendorf

Institut für Allgemeinmedizin

Prof. Dr. Hendrik van den Bussche
Direktor

Zentrum für Psychosoziale Medizin

Martinstraße 52
20246 Hamburg
Telefon: (040) 42803-2400
Telefax: (040) 42803-3681
bussche@uke.uni-hamburg.de

Stellungnahme zum Vorbericht des IQWiG „Cholinesterasehemmer bei Alzheimer Demenz“

Anmerkungen zur projektspezifischen Methodik¹

Wesentliche Kommentare

1. Bewertung der methodischen Qualität der Studien

Wir begrüßen grundsätzlich den vom IQWiG gewählten Weg einer Bewertung der methodischen Qualität der in das Review eingeschlossenen Studien. Diese Bewertung ist transparent und nachvollziehbar, was sie positiv heraushebt. Allerdings werden keine einheitlichen Konsequenzen aus der Bewertung der methodischen Qualität gezogen (vgl. Tabelle 1). So werden z.B. die beiden großen Donepezilstudien (Burns 1999 und Homma 2000) trotz grober Mängel in die Metaanalysen zu kognitiven Fähigkeiten einbezogen (S. 90-92). Auf der anderen Seite wird die Studie von Forette 1999 in den Metaanalysen nicht berücksichtigt. Wenn aber grobe Mängel nach Definition des IQWiG „die Gesamtaussage grundsätzlich in Frage stellen“ (S. 19), müsste sich zumindest eine einheitliche Konsequenz aus der Feststellung eines Mangels ergeben, z.B. eine Nichtberücksichtigung in Metaanalysen. Nach Durchsicht der Unterlagen würde der Nichteinbezug der Studien mit groben Mängeln nicht zu wesentlich anderen Ergebnissen der Metaanalysen führen, dennoch würde deutlich, dass die Datenbasis geringer ausfällt als im Vorbericht dargestellt. Bei gleich bleibenden Vorgaben des IQWiG, dass Meta-Analysen nur erfolgen sollten, „wenn mindestens drei Studien mit validen statistischen Maßzahlen“ vorliegen (S. 22), würde aber auch die Zahl der durchführbaren Metaanalysen bezüglich der Therapieziele sinken.

2. Verletzung des ITT-Prinzips

Das IQWiG nimmt an, dass eine Verletzung des ITT-Prinzips im Bereich von insgesamt weniger als 10,5 Prozentpunkten nicht relevant ist. Ebenso sei es nicht relevant, wenn der Unterschied in

¹ Alle Seitenangaben beziehen sich auf den IQWiG-Vorbericht. Die Bezeichnung der in dieser Stellungnahme erwähnten Studien folgt den vom IQWiG verwendeten Kurznamen.

den Nichtberücksichtigungsraten zwischen den Behandlungsgruppen weniger als 4,5% beträgt (S. 18). Es ist unklar, auf welcher Basis diese Grenzziehung beruht. Aus unserer Sicht muss man annehmen, dass auch eine Verletzung des ITT-Prinzips von 9% oder eine Differenz zwischen den Behandlungsgruppen von 4,3% bedenklich ist (Tabelle 2, oben).

Im Hinblick auf die Studien von Tariot 2001 und Tune 2003 (Tabelle 2, unten) sollte die Klassifizierung des IQWiG überprüft werden, denn es liegt aus unserer Sicht selbst nach IQWiG-Kriterien eine relevante Abweichung vom ITT-Prinzip vor.

Im Zusammenhang mit der Feststellung einer relevanten Abweichung vom ITT-Prinzip ist weiterhin zu kritisieren, dass diese sich lediglich auf „die Rate der in den primären Auswertungen unberücksichtigt gebliebenen Patienten“ bezieht. Da in den IQWiG-Metaanalysen kein Unterschied zwischen als primäre und sekundäre Endpunkte erhobenen Daten gemacht wird, wäre es konsequent, die Frage der relevanten Abweichung vom ITT-Prinzip für alle interessierenden Endpunkte separat zu untersuchen. Aus Tabelle 3 geht hervor, dass bei Betrachtung der zur Verfügung stehenden Daten zu alltagspraktischen Fähigkeiten eine relevante Abweichung vom ITT-Prinzip insgesamt häufiger vorkommt als bei Betrachtung der primären Endpunkte dieser Studien. Eine nach Endpunkt differenzierte Beurteilung einer relevanten Abweichung vom ITT-Prinzip macht natürlich nur Sinn, wenn, wie oben bereits erwähnt, aus der Feststellung einer relevanten Abweichung auch Konsequenzen gezogen werden.

3. Dropoutraten und LOCF

Das IQWiG nimmt an, dass die Verwendung der LOCF-Methode die Ergebnisse nicht wesentlich verfälscht. Diese Auffassung wird damit begründet, dass ein Zusammenhang zwischen den Abbruchraten und den Therapieeffekten hinsichtlich Kognition in einer eigenen Analyse nicht ersichtlich wird (S. 169, Abbildung 47). Diese Analyse ist allerdings nicht gut nachvollziehbar, z.B. ist nicht klar, welche Studien in diese Betrachtung eingeschlossen wurden und wonach gewichtet wurde. Auch stellt diese Analyse keinen Beweis dar, dass LOCF die Ergebnisse nicht verfälscht hat, da viele andere Faktoren die Ergebnisse einer jeden Studie beeinflussen und damit den Zusammenhang zwischen Cohen's d und Dropoutraten maskieren können. So haben wir entsprechende Regressionsanalysen in Abweichung zum IQWiG ohne die Studien durchgeführt, die laut IQWiG grobe Mängel aufweisen, denn diese Mängel beruhen vor allem auf Unklarheiten bzw. einer relevanten Abweichung vom ITT-Prinzip, was die Ergebnisse deutlich verfälscht haben kann. Auch haben wir Studien ausgeschlossen, bei denen die Ergebnisse nicht auf LOCF-Analysen beruhen bzw. bei denen

nicht klar ist, ob LOCF durchgeführt wurde (vgl. Tabellen 4 und 5). Sowohl wenn man nur die hohen Dosisgruppen betrachtet wie auch bei Betrachtung aller Dosisgruppen zeigt sich ein deutlicher Zusammenhang zwischen der Differenz der Dropoutraten und dem Studienergebnis zur Kognition im Sinne der Annahme, dass das Effektmaß mit der Zunahme der Differenz der Dropoutraten (Verum-Placebo) steigt. Die p-Werte betragen ungewichtet 0,03 und 0,034, nach N gewichtet 0,002 und 0,006 (vgl. Tabelle 6 sowie Abbildung 1). Diese Analysen liefern einen Hinweis darauf, dass LOCF die Ergebnisse verfälscht hat.

Weiterhin wird bezüglich der Dropoutraten und der LOCF-Strategie auf S. 169 des Berichts geäußert, dass Sensitivitätsanalysen mittels anderer Ersetzungsstrategien wünschenswert wären, um die Robustheit der Ergebnisse besser einschätzen zu können. Dabei wird auf die Studie von Unnebrink und Windeler, 2001, verwiesen. Dieser Hinweis gilt jedoch nicht für Studien, in denen die Dropoutraten größer sind als 20%. Denn bei so hohen Dropouts könne nach Unnebrink und Windeler gar keine Ersetzungsstrategie empfohlen werden:

„For LOCF and RegP the level of significance is violated already for only slightly different courses of disease in the groups. (...) To sum up, for drop-out rates less than 20 per cent and not too different courses of disease in the treatment groups, missing values might be replaced by the mean of the other group, or counted as treatment failures after dichotomization of the endpoint. MOTH² seems to be more robust regarding the type I error. For larger drop-out rates or less similar courses of disease, no adequate recommendations can be given. (...) Unreflected use of universal ‚standard methods,‘ like, for example, LOCF, is no adequate approach to the analysis of clinical trials with missing values.“

In 16 der 22 vom IQWiG analysierten Studien finden sich aber in mindestens einem der Behandlungsarme Dropoutraten von 20% und mehr, in weiteren 2 Studien finden sich zu Dropoutraten keine Angaben (vgl. Tabelle 7).

Bezüglich der Dropoutraten und der LOCF-Methode verlangt das IQWiG indirekt, dass die Kritiker der LOCF-Methode den Nachweis erbringen, dass ihre Anwendung die Ergebnisse verfälscht hat (S. 168). Unbeschadet der Frage, ob diese Zuweisung der Beweispflicht angemessen ist, kann ohne Vorliegen der Originaldaten dieser Beweis mit hinreichender Genauigkeit nicht geführt werden.

Bei den Ergebnissen zum CIBIC-plus kann man hingegen für einige Studien anhand der vorliegenden Daten Worst-Case-Szenarios berechnen, indem man bei den in den Placebogruppen fehlenden Patienten eine „Verbesserung“ annimmt (CIBIC-plus ≤ 4 bzw. ≤ 3), bei den in den Verumgruppen fehlenden Patienten hingegen eine Verschlechterung. Bei dieser – zugegebenermaßen sehr strengen Übung – verringert sich die Zahl der Studien mit statistisch signifikantem Ergebnis in mindestens

² Imputation of the mean of the other group

einer der Dosisgruppen von 9 auf 5 (vgl. Tabelle 8). Angesichts dieses Worst-Case-Szenarios sowie fünf weiterer Studien, bei denen Worst-Case-Szenarios zu diesem Endpunkt nicht berechnet werden können, erscheint das Fazit, dass ein Beleg für einen günstigen Effekt hinsichtlich des klinischen Krankheitsstadiums vorliegt, als eine zu positive Annahme (S. 161).

4. Binäre Responseraten

Auf den Seiten 166-167 berechnet das IQWiG auf Basis der vorliegenden Daten binäre Responseraten. Dies erfolgt unter der Annahme einer Normalverteilung der Veränderungen im ADAS-cog. Die Annahme einer Normalverteilung widerspricht aber z.B. der Beobachtung von Klinikern, die von 5-15% „high responders“ berichten¹. Gibt es diese „high responder“ tatsächlich, so vergrößert die Annahme einer Normalverteilung den Effekt. Eine gültige Berechnung von Responseraten ist daher nur mit Verteilungsdaten möglich, die in den Originalpublikationen meistens nicht zu finden sind – und von der pharmazeutischen Industrie trotz mehrfacher Bitten unabhängiger Wissenschaftler bisher nicht ausgehändigt wurden².

Weitere Kommentare

Die weiteren Kommentare haben sicherlich keinen entscheidenden Einfluss auf das Gesamtergebnis der Untersuchung des IQWiG. Dennoch erachten wir sie als für die Konsistenz des Berichtes bedeutsam und führen Sie im Folgenden auf.

5. Darstellung der Ergebnisse im Fazit

Als Ziel der Untersuchung formuliert das IQWiG die „Nutzenbewertung einer langfristigen Behandlung mit einem Cholinesterasehemmer“ (S. 9). Der Aspekt der Dauer der Behandlung wird im Bericht in weiten Teilen jedoch nicht berücksichtigt. Am Ende der Diskussion findet sich hierzu ein Halbsatz (S. 170), im Fazit wird die Dauer der Behandlung gar nicht erst erwähnt. Angesichts der Tatsache, dass die große Mehrzahl der Studien maximal 6 Monate dauerte, halten wir die Suggestion eines zeitlich unbeschränkten Nutzens für falsch. Ebenso wird im Fazit weder das Ausmaß des Nutzens thematisiert noch die Einschränkung des Nutzens auf „Responder“. Da das Fazit aber das meist wahrgenommene Kapitel des Berichtes sein dürfte, sollten auch diese Aspekte berücksichtigt werden.

6. Eingeschlossene Studien

Das Ziel der Untersuchung war die Nutzenbewertung der Cholinesterasehemmer bei Alzheimer Demenz. Es ist daher nicht nachvollziehbar, warum die Studie von Prasher 2002 zu Patienten mit Down-Syndrom und Demenz in die Analysen einbezogen wurde, da diese Gruppe unter allen Patienten mit Alzheimer Demenz eine extreme Minderheit darstellt – auch wenn diese Studie mit 31 Patienten sehr klein ist.

7. Kosten-Nutzen-Erwägungen

Eine Kosten-Nutzen-Erwägung der Cholinesterasehemmer fehlt im Bericht. Man könnte das damit begründen, dass entsprechende RCTs fehlen und dass Kosten/Nutzen-Analysen auch nicht der Auftrag gewesen seien. Dennoch stellt das angesichts der Diskussionen um Wirtschaftlichkeit im Gesundheitswesen (national und international) ein gewisses Manko dar.

8. In eigener Sache

Auf S. 165 wird die „Wertigkeit einzelner methodischer Kritikpunkte“ unserer Arbeit mit Hinweis auf das Gutachten von Windeler, 2005 bezweifelt. Angesichts der vielfach ungerechtfertigten Kritik, die dem Gutachten zu entnehmen ist, sollte zumindest auch ein Hinweis auf unsere Entgegnung in den Literaturangaben nicht fehlen (<http://www.uke.uni-hamburg.de/institute/allgemeinmedizin/> weiter unter „Forschung“ sowie „kognitive Störungen und Demenzen“).

Zusammenfassung

1. Grundsätzlich begrüßen wir das Vorgehen des IQWiG, die methodische Qualität der in das Review eingeschlossenen Studien zu bewerten. Die Bewertung ist transparent dargestellt und hebt sich damit von anderen Arbeiten positiv ab. Allerdings sollte das IQWiG aus der Feststellung eines groben methodischen Mangels eine für alle Studien einheitliche Konsequenz ziehen, die z.B. darin besteht, diese Studien nicht in Metaanalysen einzuschließen.
2. Es ist sicher notwendig, eine relevante Verletzung des ITT-Prinzips von einer irrelevanten zu unterscheiden. Ebenso ist es notwendig, die Relevanz von Dropouts und Ersetzungsstrategien für die Ergebnisse von Studien zu beurteilen. Auch die Darstellung von Studienergebnissen in Form von binären Responseraten ist zu wünschen. All diese wichtigen Berechnungen und Betrachtungen sollten aber nicht pauschal erfolgen, sondern konkret auf der Basis von Originaldaten.
3. Im Fazit sollte differenzierter auf die Evidenzlage zur Dauer der Behandlung, dem Ausmaß des Nutzens sowie der Einschränkung des Nutzens auf „Responder“ eingegangen werden.

ANHANG

Tabelle 1: Einschluss von Studien mit groben Mängeln in die Metaanalysen des IQWiG-Vorberichtes

Studien mit groben Mängeln	Einschluss in Metaanalysen
Burns 1999	Kognitive Leistungsfähigkeit (S. 90-92)
Homma 2000	ADL (S. 85)
	Kognitive Leistungsfähigkeit (S. 90-91)
Mohs 2001	Nein
Moraes 2006	Nein
Prasher 2002	Nein
Erkinjuntti 2002	ADL (S. 113)
	Begleitende Psychopathologie (S. 115)
	Kognitive Leistungsfähigkeit (S. 117)
Forette 1999	Nein
Rösler 1999	ADL (S. 133)
	Kognitive Leistungsfähigkeit (S. 135)

Tabelle 2: Relevante Abweichung vom ITT-Prinzip: Studien, die nah am IQWiG-Grenzwert liegen (oben) sowie Studien, bei denen das Kriterium erfüllt ist ohne im Vorbericht berücksichtigt zu sein (unten)

Studie	N ITT/N randomised (Verletzung ITT in %) Verum	N ITT/N randomised (Verletzung ITT in %) Placebo	Bemerkung
Rogers 1998 5mg	152/154 (1,3%)	153/162 (5,6%)	Differenz zwischen den Gruppen nähert sich 4,5%
Brodaty 2005 – galantamine PRC	291/320 (9,1%)	296/324 (8,6%)	Dropoutraten nähern sich 10,5%
Brodaty 2005 – galantamine	296/327 (9,5%)		
Tariot 2000 – 8 mg	126/140 (10%)	255/286 (10,8%)	Dropoutraten nähern sich 10,5%
Tariot 2000 – 16 mg	253/279 (9,3%)	255/280 (8,9%)*	Dropoutrate für alle Verumgruppen gemeinsam: 632/660 (4,2%)*
Tariot 2000 – 24 mg	253/273 (7,3%)		
Tariot 2001	103	105	Keine Angaben bzgl. N beim primären Endpunkt (NPI), aus unserer Sicht „grober Mangel“
Tune 2003	14/14 (0%)	13/14 (7,1%)	Differenz zwischen den Gruppen: 7,1%, aus unserer Sicht „grober Mangel“

* In der Publikation wird angegeben, dass 40 Patienten eines Studienzentrums wegen Verstößen des dortigen Studienleiters gegen das Studienprotokoll vor Beginn der Datenanalysen ausgeschlossen wurden, davon 32 Patienten aus den Galantamingruppen und 6 aus der Placebogruppe. Über die Zuordnung von 2 Patienten wird nicht berichtet. Insofern lässt sich eine reduzierte Verletzung des ITT-Prinzips nur für alle Verumgruppen gemeinsam berechnen (4,2%). Die Differenz zu Placebo beträgt 4,7%.

Tabelle 3: Relevante Abweichung vom ITT-Prinzip: Unterschiede bei Betrachtung des prim. Endpunktes und der alltagspraktischen Fähigkeiten

Studie	Relevante Abweichung vom ITT-Prinzip – prim. Endpunkt(e)	Relevante Abweichung vom ITT-Prinzip – alltagspraktische Fähigkeiten	Seiten im IQWiG-Vorbericht
Burns 1999	unklar	unklar	S. 49-50 sowie 83
Gauthier 2002	nein	ja (Differenz zwischen Gruppen 6%)	
Homma 2000	ja	unklar	
Tariot 2001	nein	unklar	
Winblad 2001	nein	ja (Daten nur für Observed Cases dargestellt)	
Brodsky 2005	nein	unklar	S. 61 sowie 112
Erkinjuntti 2002	ja	nein	
Raskind 2000	nein	unklar	
Rockwood 2006	nein	unklar	
Tariot 2000	nein	nein	
Wilcock 2000	nein	nein	
B 304 1998	nein	nein	S. 70 sowie 132
Corey-Bloom 1998	nein	nein	
Rösler 1999	ja	nein	

Tabelle 4: Studien und Daten, die in die Meta-Regressionen eingeflossen sind; grau unterlegt sind die Werte für die niedrig-Dosis-Behandlungsarme

	Standardisierte Differenz der Mittelwerte*	Differenz der Abbruchraten (Verum- Placebo)	Gesamt N (Verumarm + Placebo)
Corey-Bloom 1998 - 6-12 mg	0,632652	18	466
Gauthier 2002 - flex. Dosis	0,452371	8	207
Krishnan 2003 - flex. Dosis	0,512738	-12	67
Rogers 1998 - 10 mg	0,467087	12	319
Seltzer 2004 - 10 mg	0,702581	8	153
Tariot 2001 - flex. Dosis	0,226723	-8	208
Winblad 2001 - flex. Dosis	0,45057	0	286
Brodaty 2005 - PRC 16-24	0,454625	4	644
Brodaty 2005 - 16-24	0,470504	5	651
Raskind 2000	0,666408	13	425
Rockwood 2006	0,374801	2	130
Tariot 2000 - 24 mg	0,499552	6	559
Wilcock 2000	0,499285	7	435
Rogers 1998 - 5 mg	0,402718	-5	316
Tariot 2000 - 8 mg	0,213843	-7	426
Tariot 2000 - 16 mg	0,523904	6	565
Corey-Bloom 1998 - 1-4 mg	0,28303	-2	468

* In Abweichung vom IQWiG-Vorbericht wurde die standardisierte Differenz der Mittelwerte mit N randomised berechnet. Der Unterschied zum vom IQWiG berechneten Cohen's d zeigt sich aber nur ab der dritten Stelle hinter dem Komma und ist daher irrelevant.

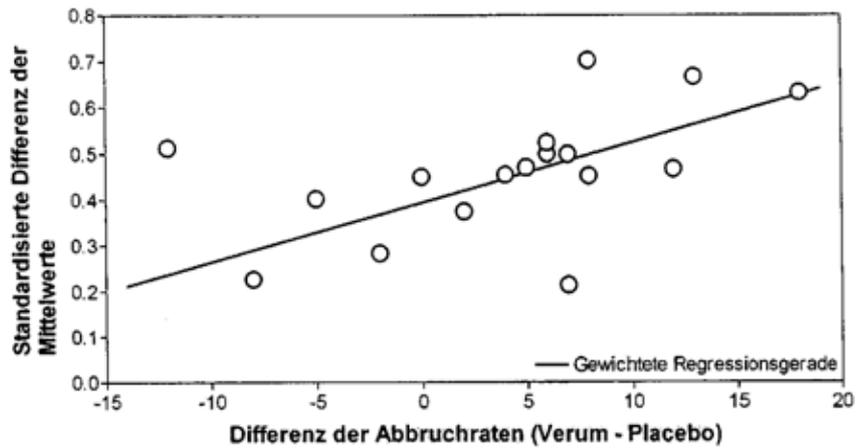
Tabelle 5: Studien, die nicht in die Meta-Regression eingeflossen sind

Studie	Begründung für Nicht-Einschluss
Burns 1999	Studie mit groben Mängeln (Verletzung des ITT-Prinzips unklar)
Homma 2000	Studie mit groben Mängeln (relevante Abweichung vom ITT-Prinzip)
Mohs 2001	Studie mit groben Mängeln (Verletzung des ITT-Prinzips unklar)
Moraes 2006	Keine Angaben zu Dropoutraten, kein LOCF durchgeführt
Prasher 2002	Kein LOCF durchgeführt
Tune 2003	Kein LOCF durchgeführt
Erkinjuntti 2002	Studie mit groben Mängeln (relevante Abweichung vom ITT-Prinzip)
B 304	Unklar, ob LOCF durchgeführt wurde (lässt sich dem IQWiG-Vorbericht nicht entnehmen)
Forette 1999	Studie mit groben Mängeln (relevante Abweichung vom ITT-Prinzip), LOCF wurde nicht durchgeführt
Rösler 1999	Studie mit groben Mängeln (relevante Abweichung vom ITT-Prinzip)

Tabelle 6: Ergebnisse der Linearen logistischen Regressionen in Abhängigkeit von Anzahl der eingeschlossenen Studien und Wichtung

Einschluss in Regression	alpha bzw. gewichtetes alpha (±Standardfehler)	95% Konfidenzintervall	p-Wert	Regressionsgerade
Nur hohe Dosisgruppen (13 Datensätze aus 12 Studien)	0,009 (± 0,004)	0,001 - 0,017	0,030	$y = 0,449 + 0,009x$
Hohe und niedrige Dosisgruppen (17 Datensätze aus 12 Studien)	0,009 (± 0,004)	0,001 - 0,018	0,034	$y = 0,423 + 0,009x$
Nur hohe Dosisgruppen (13 Datensätze aus 12 Studien), nach N gewichtet	0,013 (± 0,003)	0,006 - 0,019	0,002	$y = 0,417 + 0,013x$
Hohe und niedrige Dosisgruppen (17 Datensätze aus 12 Studien), nach N gewichtet	0,013 (± 0,004)	0,004 - 0,021	0,006	$y = 0,394 + 0,013x$

Abbildung 1: Streudiagramm zum Zusammenhang zwischen der Differenz der Dropoutraten und der standardisierten Differenz der Mittelwerte bei 12 Studien (mit 17 Datensätzen), in denen LOCF verwendet wurde; nach N gewichtete Regressionsgerade*



* Erläuterungen zur benutzten Datenbasis in Tabellen 4-6

Tabelle 7: Dropoutraten in den 22 RCT, die vom IQWiG zur Nutzenbewertung der Cholinesterasehemmer herangezogen wurden (grau unterlegt sind Dropoutraten $\geq 20\%$)

Studie	Dropout Verum (%)	Dropout Placebo (%)
Burns 1999 - 5 mg	22	20
Burns 1999 - 10 mg	26	20
Gauthier 2002 - flex. Dosis	19	11
Homma 2000 - 5 mg	keine Angaben	keine Angaben
Krishnan 2003 - flex. Dosis	18	30
Mohs 2001	28	26
Donepezil Moraes 2006	keine Angaben	keine Angaben
Prasher 2002	13	13
Rogers 1998 - 5 mg	15	20
Rogers 1998 - 10 mg	32	20
Seltzer 2004 - 10 mg	27	19
Tariot 2001 - flex. Dosis	18	26
Tune 2003 - 10mg	0	7
Winblad 2001 - flex. Dosis	33	33
Galantamin Brodaty 2005 - PRC 16-24	22	18
Brodaty 2005 - 16-24	23	18
Erkinjuntti 2002	26	17
Raskind 2000	32	19
Rockwood 2006	17	15
Tariot 2000 - 8 mg	23	16
Tariot 2000 - 16 mg	22	16
Tariot 2000 - 24 mg	22	16
Wilcock 2000	20	13
Rivastigmin B 304 - 2x/d	24	15
B 304 - 3x/d	17	15
Corey-Bloom 1998 - 1-4 mg	15	17
Corey-Bloom 1998 - 6-12 mg	35	17
Forette 1999 - 2x/d	36	8
Forette 1999 - 3x/d	24	8
Rösler 1999 1-4 mg	14	13
Rösler 1999 6-12 mg	33	13

Tabelle 8: Ergebnisse zum klinischen Gesamteindruck, CIBIC-plus, Berechnung eines Worst Case Szenarios

Studien	N ran-	N in	N mit	N feh-	N mit	N mit	N mit	N mit	p-Wert (Verum/	Anteil (%) mit	Richtung des	p-Wert in Studie
	do-	Aus-	Score	lend	Score	Score	Score	Score	Placebo), worst <td>Score \leq 4, worst <td>worst case <td>(\dagger=nicht dichotomi-</td> </td></td>	Score \leq 4, worst <td>worst case <td>(\dagger=nicht dichotomi-</td> </td>	worst case <td>(\dagger=nicht dichotomi-</td>	(\dagger =nicht dichotomi-
	mi-	wertung	\leq 4		\leq 4,	\leq 4,	\leq 4,	\leq 4,	case, Vierfelder-	case	Effektes	siert, \S =van Elteren
	stiert				worst case	worst case	worst case	worst case	test	fest		Test)
												< 0,001+
Gauthier 2002 - flex. Dosis	102	98	69	4	69	67	0,002	67	0,002	0,002	†	< 0,001+
Gauthier 2002 - Placebo	105	105	49	0	49	47	0,0001	47	0,0001	0,0001	†	< 0,001+
Homma 2000 - 5 mg	136	133	108	3	108	79	0,0001	79	0,0001	0,0001	†	< 0,001+
Homma 2000 - Placebo	132	128	73	4	77	58	0,06	58	0,06	0,06	†	0,005+
Rogers 1998 - 5 mg*	154	149	39	5	39	25	0,12	25	0,12	0,12	†	< 0,001+
Rogers 1998 - 10 mg*	157	149	37	8	37	24	0,16	24	0,16	0,16	†	< 0,001+
Rogers 1998 - Placebo*	162	152	17	10	27	16	0,27	16	0,27	0,27	†	n.s.§
Brodaty 2005 - PRC 16-24	320	291	180	29	180	56	0,42	56	0,42	0,42	†	n.s.§
Brodaty 2005 - 16-24	327	302	190	25	190	58	0,046	58	0,046	0,046	†	n.s.§
Brodaty 2005 - Placebo	324	301	172	23	195	60	0,35	60	0,35	0,35	†	0,001§
Erkinjuntti 2002	188	172	129	16	129	69	0,25	69	0,25	0,25	†	< 0,05§
Erkinjuntti 2002 - Placebo	97	92	50	5	55	56	0,48	56	0,48	0,48	†	0,2600
Raskind 2000	212	186	136	26	136	64	0,03	64	0,03	0,03	†	0,0020
Raskind 2000 - Placebo	213	196	112	17	129	60	0,14	60	0,14	0,14	†	< 0,05
Tariot 2000 - 8 mg	140	128	68	12	68	48	0,09	48	0,09	0,09	†	< 0,001§
Tariot 2000 - 16 mg	279	256	169	23	169	61	0,16	61	0,16	0,16	†	< 0,001§
Tariot 2000 - 24 mg	273	253	162	20	162	59	0,25	59	0,25	0,25	†	< 0,05§
Tariot 2000 - Placebo	286	261	128	25	153	53	0,48	53	0,48	0,48	†	< 0,001
Wilcock 2000	220	206	128	14	128	58	0,25	58	0,25	0,25	†	< 0,05§
Wilcock 2000 - Placebo	215	203	102	12	114	53	0,48	53	0,48	0,48	†	< 0,05
B 304 - 2x/d*	229	222	51	7	51	22	0,03	22	0,03	0,03	†	< 0,001
B 304 - 3x/d*	227	222	69	5	69	30	0,01	30	0,01	0,01	†	< 0,001
B 304 - Placebo*	222	216	41	6	47	21	0,14	21	0,14	0,14	†	< 0,05
Rösler 1999 1-4 mg*	243	233	70	10	70	29	0,01	29	0,01	0,01	†	< 0,001
Rösler 1999 6-12 mg*	243	219	81	24	81	33	0,01	33	0,01	0,01	†	< 0,001
Rösler 1999 - Placebo*	239	230	46	9	55	23	0,14	23	0,14	0,14	†	< 0,001

*: alle Berechnungen gelten für CIBIC-plus \leq 3

Hamburg, den 4.10. 2006

H. Kaduszkiewicz

Dr. med. Hanna Kaduszkiewicz

*

Dipl.-Psych. Dr. Thomas Zimmermann

*

Prof. Dr. med. Hendrik van den Bussche

Beigefügte Literatur

¹ Lemstra AW, Eikelenboom P, van Gool WA. The cholinergic deficiency syndrome and its therapeutic implications. Gerontology 2003; 49: 55-60.

² Richard E, van Gool WA. Cholinesteraseremmers bij dementie: een wankel evenwicht tussen werkzaamheid en veiligheid. Ned Tijdschr Geneeskd. 2006; 150 (10): 530-532.

*

5.10.06

Sehr geehrter Prof. Sawicki,

mir ist leider erst jetzt aufgefallen, daß auch dieser Bogen von allen Unterschriften sein muss.

Da beide bis zum 6.10. außer Haus sind, kann ich Ihnen diese beide Unterschriften heute nicht "liefern".

Ich versichere Ihnen, daß sowohl Herr Zimmermann als auch Herr van den Bussche die Stellungnahme gelesen, diskutiert, korrigiert und freigegeben haben. Daher hoffe ich, daß die Unterschriften hier vorzulegen sind.

Mit freundlichen Grüßen,
H. Kaduszkiewicz

Stellungnahme University of Manchester, Burns



Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

Auftrag des IQWiG/Projektbezeichnung mit Auftragsnummer

A05/19A

Cholinesterase inhibitor in Alzheimer's disease

Name, Vorname, Titel und Funktion des/der Stellungnehmenden

1. Professor Alistair Burns
Professor of Old Age Psychiatry

2. University of Manchester

3.

Kontakt für weitere Nachfragen

Name: Professor Alistair Burns
Institution: The University of Manchester
Anschrift: Department of Psychiatry, 2nd Floor, Education and Research Centre, Wythenshawe Hospital, Manchester, M23 9LT
Tel./Fax-Nr.: +44 161 291 5887 Mobile: +44 7801 358231
E-Mail-Adresse: alistair.burns@manchester.ac.uk

Bitte strukturieren Sie Ihre Stellungnahme wie folgt

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Stellungnahme University of Manchester, Burns (Fortsetzung)



Substanzielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt für **alle Stellungnehmenden** einzeln **ausgefüllt** vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substanziellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Margareta 6.07.2006
Ort/Datum

[Signature]
Unterschrift 1

Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

2

Stellungnahme University of Manchester, Burns (Fortsetzung)

I have produced this report because I have become aware that the validity and the relevance of the Clinicians' Interview Based Impression of Change, with caregiver input (CIBIC-plus) scale has been questioned by the IQWiG. Clinical Global Impression of Change (CGIC) is a technique that is widely used in clinical trials across therapeutic areas. In AD studies, regulatory agencies such as the United States FDA consider these assessments essential to demonstrate clinically meaningful benefit.¹ The goal of this type of assessment is to detect clinically meaningful change as it would be judged by a physician in clinical practice. The standard instrument used in trials of AD treatments is the CIBIC-plus, which was developed based on a validated, reliable instrument developed by the Alzheimer's Disease Cooperative Study.¹⁻³ Whereas performance-based assessments such as the ADAS-cog are often performed by clinicians and clinical researchers trained in neuropsychological assessment, CGICs in general and the CIBIC-plus in particular are administered by a physician. After a semistructured interview with the patient and the caregiver, the clinician performing the assessment is asked to rate, on a Likert scale, whether the patient's overall condition has improved, stayed the same, or worsened.¹ Improvement and worsening are graded by the terms minimally, moderately, and markedly. This process reflects the judgment a physician would make in clinical practice after seeing the patient and caregiver.

Because the CIBIC-plus was designed to be insensitive to small, clinically irrelevant changes, any change recorded on the CIBIC-plus can be regarded as clinically meaningful.^{1,2} In AD, which is a chronic degenerative disease, a clinically meaningful benefit would be detected not only in the number of patients rated as improved, but also as a reduced proportion of patients who worsened, or a reduced degree of worsening, relative to untreated or placebo patients. In contrast to some other conditions in which CGICs are utilized, such as schizophrenia, in AD an outcome of "no change" after a period of 3 or 6 months would suggest successful treatment.

Stellungnahme University of Manchester, Burns (Fortsetzung)

The CIBIC-plus and other CGICs might be seen as limited because they are based on objective, performance-based criteria. This might be a valid concern if the CIBIC-plus was the only assessment performed. On the contrary, the CIBIC-plus is meant to complement performance based assessments and verify the clinical relevance of the changes detected by sensitive neuropsychological tests. This combination of objective, performance-based measurement and real-world clinical assessment is a very powerful way to evaluate interventions in AD. The assessment of CIBIC-plus and other CGICs is therefore both important in clinical research and in clinical practice. I therefore consider that the decision by the IGWIG to downgrade the CIBIC to only providing "supplementary information" as misjudged.

A. Burns

Professor Austin Burns

6 027 0866 2006. References

1. Schneider LS, Olin JT, Doody RS, et al. Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord.* 1997;11(suppl 2):S22-S32.
2. McLendon BM, Doraiswamy PM. Defining meaningful change in Alzheimer's disease trials: the donepezil experience. *J Geriatr Psychiatry Neurol.* 1999;12(1):39-48.
3. Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT, Donepezil Study Group. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. *Neurology.* 1998;50(1):136-145.

Stellungnahme VFA, Wahler



Eingangskontrolle

Formblatt zur Anhörung (schriftlichen Stellungnahme) zum Vorbericht

A05/19-A- Cholinesterasehemmer bei Alzheimer Demenz

Name, Vorname, Titel und Funktion des/der Stellungnehmenden

1. Wahler, Steffen; Dr. med./VFA/Geschäftsführer Gesundheitsökonomie
2. Binek, Matthias; Dr. med./VFA/Nutzenbewertung - Arzneimittelsicherheit
- 3.

Kontakt für weitere Nachfragen:

Name: Dr. med. Matthias Binek
Institution: VFA
Anschrift: Hausvogteiplatz 13, 10117 Berlin
Tel./Fax-Nr.: +49 30 20604-355
E-Mail-Adresse: m.binek@vfa.de

Bitte strukturieren Sie Ihre Stellungnahme wie folgt:

1. Im Vorbericht fehlende Originalstudien
2. Anmerkungen zu Bewertungen von Originalstudien im Vorbericht
3. Anmerkungen zur projektspezifischen Methodik (bei Änderungsvorschlägen einschließlich Begründung unter Angabe entsprechender wissenschaftlicher Belege)

Formblatt zur Abgabe schriftlicher Stellungnahmen; Version 4, Stand: 08/2006

1

Stellungnahme VFA, Wahler (Fortsetzung)



Substantielle Stellungnahmen, die der Form genügen und rechtzeitig eingereicht werden, finden Berücksichtigung. Anhand folgender Checkliste können Sie prüfen, ob Ihre Stellungnahme die formalen Kriterien erfüllt. Bitte beachten Sie darüber hinaus die Hinweise im „Leitfaden zur Abgabe von Stellungnahmen“.

Der Umfang der Stellungnahme umfasst maximal **5 DIN-A4-Seiten**.



Der Umfang ggf. beigefügter Anhänge zur Stellungnahme umfasst maximal **10 DIN-A4-Seiten**.



Für alle Zitate sind **Originalpublikationen (Volltext)** beigelegt.



Die Stellungnahme ist **von allen Stellungnehmenden** unterschrieben.



Das **Formblatt zur Darlegung potenzieller Interessenskonflikte** liegt **für alle Stellungnehmenden** einzeln ausgefüllt vor.



(Anm.: Dies gilt auch für Stellungnahmen von Institutionen / Gesellschaften. Die Abgabe dieses Formblatts ist für die Teilnahme an der ggf. stattfindenden wissenschaftlichen Erörterung der schriftlichen Stellungnahmen zwingend erforderlich)

Mir / uns ist bekannt, dass alle substantiellen Stellungnahmen im Rahmen des Abschlussberichts im Internet veröffentlicht werden können. Einer solchen Veröffentlichung stimme ich / stimmen wir mit meiner / unseren Unterschrift(en) ausdrücklich zu. Den „Leitfaden zur Abgabe von schriftlichen Stellungnahmen im Rahmen der Anhörung zum Vorbericht“ habe ich / haben wir zur Kenntnis genommen.

Berlin, 06.10.2006
Ort/Datum

Wahler
Unterschrift 1

[Signature]
Unterschrift 2

Unterschrift 3

Vielen Dank für Ihre Stellungnahme.

Stellungnahme VFA, Wahler (Fortsetzung)



Verband Forschender
Arzneimittelhersteller e.V.

**Stellungnahme zum Vorbericht
„Cholinesterasehemmer bei Alzheimer Demenz“ des
Instituts für Qualität und Wirtschaftlichkeit im
Gesundheitswesen [Auftrag A05/19-A] mit speziellem
Fokus auf projektspezifische Methodik**

Einleitung

Am 08.09.2006 wurde vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) der Vorbericht „Cholinesterasehemmer bei Alzheimer Demenz“ [Auftrag A05/19-A] veröffentlicht. Der Vorbericht des IQWiG ist Teil des Bewertungsprozesses, der erst nach einer öffentlichen Kommentierung in einen Abschlussbericht mündet. Der Verband Forschender Arzneimittelhersteller e.V. (VFA) begrüßt die Möglichkeit, zu diesem Bericht eine Stellungnahme abgeben zu können, die gegebenenfalls in einer wissenschaftlichen Diskussionsrunde hinsichtlich ihrer Relevanz für eine Ergänzung und/oder Überarbeitung des Berichts führen kann.

Seite 1/5

Projektspezifische Methodik

Frühzeitiger Dialog

Die stark begrenzte Form der Zusammenarbeit zwischen dem IQWiG und der Industrie ist nach Auffassung des VFA nicht sachgerecht. Statt die Datenlage, die Bewertungskriterien und die Vergleichsmaßstäbe bei Nutzenbewertungen zu Cholinesterasehemmern frühzeitig mit den Produktherstellern und deren Experten zu erörtern, hat das Institut zunächst allein auf das eigene Know-how und die selbst selektierten Fachkreise vertraut. Der VFA hat Zweifel, ob in einem fortgeschrittenen Stadium (Anhörung zum Vorbericht) tiefergehende oder grundsätzliche Einwände der Hersteller noch adäquat berücksichtigt werden können. Dieses Verfahren widerspricht im Übrigen den etablierten Formen eines strukturierten Dialogs mit den Arzneimittelherstellern, wie er routinemäßig von internationalen HTA- bzw. Erstattungsprüfinstanzen in anderen Ländern praktiziert wird.

Hausvogelplatz 13
10117 Berlin
Telefon 030 20604-0
Telefax 030 20604-222
www.vfa.de

Hauptgeschäftsführerin
Cornelia Yzer

Stellungnahme VFA, Wahler (Fortsetzung)



Verband Forschender
Arzneimittelhersteller e.V.

Zeitpunkt der Literaturrecherche

Die Literaturrecherche ("Erstrecherche") wurde vor der Erstellung des Berichtsplans durchgeführt, was sowohl Zweifel an der Unabhängigkeit des Berichtsplans aufkommen lässt als auch im Widerspruch zu den IQWiG-Methoden (S. 94/95, Abb. 4) und allen gängigen HTA-Methoden steht, da erst im Research-Protokoll (Berichtsplan) die Suchstrategien definiert werden.

Stellungnahmeverfahren

Seite 2/5

Ein weiterer kritischer Punkt besteht darin, dass die Stellungnahmen zum Vorbericht zwar eingeholt, jedoch nicht berücksichtigt werden müssen. Es wäre im Rahmen eines transparenten und fairen Verfahrens sicherlich sachgerecht, eine fachliche Begründung zur Berücksichtigung oder Nicht-Berücksichtigung einer Stellungnahme im Vorfeld der Anhörung zum Vorbericht zu geben.

Für die Akzeptanz des Stellungnahmeverfahrens zum Vorbericht wäre es sinnvoll gewesen, auf die vom IQWiG verlangte „Darlegung potenzieller Interessenskonflikte“ zu verzichten oder sie zumindest erst bei feststehender Berücksichtigung der Stellungnahme im Anhörungsverfahren zu verlangen. Die Übertragung der Vorgaben für externe Sachverständige im Rahmen der IQWiG-Forschungsaufträge auf Vorbericht-Stellungnehmende ist juristisch nicht nachvollziehbar und mit einem abschreckenden bzw. präselektierenden Effekt verbunden. Eine einfache qualitative Angabe zum Interessenskonflikt, wie sie häufig in wissenschaftlichen Publikationen verwendet wird, ist in diesem Zusammenhang ausreichend. Steuerbehördlich relevante Detailfragen sind völlig inadäquat und deren Beantwortung führt nicht zu einer Objektivierung der wissenschaftlichen Aussagen.

Studienselektion

Nach Durchsicht des Vorberichtes stellen wir fest, dass das IQWiG zwar eine ausführliche Analyse der vorliegenden Studien vorgenommen hat, jedoch hat die vom IQWiG angewandte und bereits kritisierte Methode leider dazu geführt, dass zahlreiche klinisch relevante Publikationen nicht bei der Nutzenbewertung berücksichtigt wurden.

Das IQWiG bemängelt in seinem Vorbericht, dass für mehrere Therapieziele bei bestimmten Substanzen keine interpre-

Stellungnahme VFA, Wahler (Fortsetzung)



Verband Forschender
Arzneimittelhersteller e.V.

tierbaren Daten vorlägen. Bei einer Reduzierung der in die Nutzenbewertung eingeschlossen Studien/Publikationen auf n=27/48 ist dies nicht verwunderlich. Umso überraschender ist das Ausbleiben einer weiterführenden Recherche bzw. Anfrage bei den Autoren zu den Studien, die in die Nutzenbewertung eingeflossen sind, jedoch bezüglich deren Qualität vom IQWiG kritisiert wurden. In diesem Zusammenhang hätte das Institut die wissenschaftliche Pflicht gehabt zumindest den Versuch zu unternehmen, die etwaigen biometrischen Mängel zu klären, um die bereits sehr stringent ausselektierten Studien adäquat auswerten zu können.

Seite 3/5

Das IQWiG akzeptiert für die Nutzenbewertung von Cholinesterasehemmern ausschließlich Daten aus randomisierten, kontrollierten klinischen Studien (randomized controlled trials - RCTs). Dieses methodische Vorgehen wird dem eigentlichen Ziel, den Patientennutzen zu erfassen und zu bewerten, nicht gerecht und ist international unüblich.

Durch die engen methodischen Vorgaben des Instituts wird die Nutzenbewertung zu einer Zweitprüfung der Wirksamkeit von Medikamenten in klinischen Settings verkürzt („efficacy“). Die Prüfung der Wirksamkeit, Verträglichkeit und Sicherheit sind Gegenstand der Arzneimittelzulassung. Aus den für den Zulassungsprozess relevanten Studien lassen sich jedoch oft keine Nutzenbewertungen ableiten, so dass das Ziel, die „Effectiveness“ eines Arzneimittels zu messen, die über eine bereits dokumentierte „Efficacy“ hinausgeht, kaum erreicht werden kann. Die praktischen Anwendungserfahrungen der Ärzte und Patienten mit Medikamenten können im Rahmen von RCTs, also kontrollierter experimenteller Studiendesigns, nur begrenzt abgebildet werden. Der Behandlungsalltag beinhaltet genau jene methodischen „Störgrößen“, die in RCTs konzeptionell ausgeschlossen werden. Deshalb ist es unabdingbar, neben RCTs weitere Studienformen bei der Nutzenbewertung zu berücksichtigen, die die Wirksamkeit einer Therapie in der Versorgungsrealität belegen („effectiveness“) und die gewonnene Erfahrung von Patienten, Angehörigen, Betreuenden und Ärzten mit Cholinesterasehemmern wissenschaftlich zu würdigen.

Ausschlaggebend für die Berücksichtigung der Ergebnisse darf nicht die Studienform per se, sondern allein die Frage sein, ob die Studien transparent konzipiert und wissenschaftlich kompetent durchgeführt wurden und geeignet sind, die im Berichtsplan spezifizierte Fragestellung zu beantworten.

Leider hat der rein methodisch bedingte Ausschluss klinisch relevanter Studien dazu geführt, dass die Patientengruppen,

Stellungnahme VFA, Wahler (Fortsetzung)



Verband Forschender
Arzneittelhersteller e.V.

die von den positiven Aspekten einer Therapie mit Cholinesterasehemmern profitieren, nicht eindeutiger identifiziert wurden und das IQWiG in dem Fazit des Vorberichts keine klare Empfehlung verfasst hat.

Studienmindestbeobachtungszeit

Für die Bewertung des Nutzens einer Therapie mit Cholinesterasehemmern werden vom IQWiG die zu berücksichtigenden Studiencharakteristika auf „eine Mindestbeobachtungszeit von 16 Wochen festgelegt, da innerhalb dieses Zeitraumes ein Ansprechen auf die Therapie erwartet und ein mehr als kurzfristig anhaltender Effekt beobachtet werden kann“. Als Begründung wird die Empfehlung der Arzneimittelkommission der Deutschen Ärzteschaft für die Ersteinstellung von Medikamenten zur Pharmakotherapie kognitiver Störungen angegeben, die die Kontrolle eines Therapieerfolgs nach 12 Wochen befürwortet. Zusätzlich bezieht sich das IQWiG auf die Zulassungsanforderungen der European Medicines Agency (EMA), die eine Studiendauer von mindestens 24 Wochen als sinnvoll betrachtet. Die Schlussfolgerung bzw. Vorgabe, nur Studien mit einer minimalen Laufzeit von 16 Wochen einzuschließen, ist vom Institut nicht weiter wissenschaftlich begründet und vermittelt den Eindruck einer willkürlichen Festsetzung. Sie folgt damit keinem internationalen Standard in der Nutzenbewertung.

Seite 4/5

Endpunkte

Ein weiterer kritikwürdiger Punkt ist die künstliche Trennung zwischen den Therapiezielen Institutionalisierung (der als "patientenrelevant" eingestuft wird) und Pflegeaufwand (der als "angehörigenrelevant" eingestuft wird). Eine Heimeinweisung hängt nicht nur von der Pflege- bzw. Betreuungs- und Unterstützungsbedürftigkeit eines Alzheimerkranken ab, sondern u.a. auch davon, ob jemand (ob angehörig oder nicht) vorhanden ist, der die außerhalb einer Institution leisten kann, und ob eine Heimplatz verfügbar ist. Ein solcher Endpunkt stellt daher immer nur eine grobe Hilfsgröße ab, mit der eine (Vollzeit-) Pflege-, Betreuungs- und Unterstützungsbedürftigkeit gemessen werden kann. Es ist aber eigentlich kein patientenrelevanter Endpunkt, sondern zuallererst ein ökonomisch relevanter Endpunkt (da häusliche Pflege/Unterstützung die Pflegekassen weniger stark belastet als institutionelle Hilfe).

Stellungnahme VFA, Wahler (Fortsetzung)



Verband Forschender
Arzneimittelhersteller e.V.

Betreuungsaufwand bzw. Zeit, die mit der Assistenz von Patienten bei Alltagsaktivitäten verbracht wird spiegelt hingegen viel präziser den Grad der (Un-) Abhängigkeit und Hilfsbedürftigkeit des Patienten wider und ist damit zuallererst patientenrelevant, nicht unbedingt angehörigerelevant. Methodisch und aus Patientenbedarfs-Perspektive sollte es nämlich keine Rolle spielen, ob diese Unterstützung durch Angehörige oder Nicht-Verwandte geleistet wird, und letztlich auch nicht, ob sie im häuslichen Umfeld oder einer Institution geleistet wird, denn auch in Pflegeheimen gibt es je nach Patient unterschiedliche Pflege- und Betreuungsintensitäten.

Ferner ist anzumerken, dass die vom IQWiG gewählte Semantik eine etwas diskriminierende Wirkung hat. Sie impliziert, dass nur Angehörige solche Aufgaben übernehmen, oder dass, wenn diese von dritten (Nicht-Verwandten) übernommen werden, sie offenbar weder als patienten- noch als angehörigerelevant eingestuft werden.

Seite 5/5

06. Oktober 2006