

IQWiG-Berichte - Jahr: 2011 Nr. 84

# Responderanalysen zu Memantin bei Alzheimer **Demenz**

# **Rapid Report**

Auftrag: A10-06 Version: 1.0

Stand: 28.03.2011

## **Impressum**

#### Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

#### Thema:

Responderanalysen zu Memantin bei Alzheimer Demenz

## Auftraggeber:

Gemeinsamer Bundesausschuss

#### **Datum des Auftrags:**

04.11.2010

#### **Interne Auftragsnummer:**

A10-06

### **Anschrift des Herausgebers:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen Dillenburger Str. 27 51105 Köln

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

ISSN: 1864-2500

Dieser Rapid Report wurde einem externen Review unterzogen. Das Review wurde durchgeführt von:

Prof. Dr. Hans J. Trampisch
 Ruhr-Universität Bochum, Abteilung für Medizinische Informatik, Biometrie und
 Epidemiologie

Die Offenlegung der potenziellen Interessenkonflikte des externen Reviewers findet sich in Anhang A.

#### Kurzfassung

#### Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 04.11.2010 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) in Ergänzung zum Abschlussbericht A05-19C und zum Arbeitspapier "Memantin bei Alzheimer Demenz" mit der Bewertung der von der Firma Merz im Oktober 2010 übermittelten Responderanalysen beauftragt.

#### **Fragestellung**

Die Fragestellung der vorliegenden Untersuchung ist:

Welchen Einfluss haben die von der Firma Merz im Nachgang berechneten und an den G-BA im 4. Quartal 2010 übermittelten Responderanalysen auf das Fazit des Abschlussberichts A05-19C ("Memantin bei Alzheimer Demenz")?

#### Methoden

Die übermittelten Responderanalysen wurden mit der für den Abschlussbericht A05-19C und das Arbeitspapier "Memantin bei Alzheimer Demenz" festgelegten Methodik bewertet.

Eine darüber hinausgehende Informationsbeschaffung wurde nicht durchgeführt.

#### **Ergebnisse**

Das Schreiben der Firma Merz enthielt Responderanalysen zu den 2 patientenrelevanten Endpunkten Aktivitäten des täglichen Lebens und kognitive Leistungsfähigkeit. Darüber hinaus fanden sich Informationen zu dem im Abschlussbericht A05-19C als nicht patientenrelevant eingestuften und als ergänzende Information dargestellten Endpunkt Krankheitsstadium gemäß klinischem Eindruck. Untersucht wurde jeweils der Anteil Patienten, der eine höchstens nicht relevante Verschlechterung erfahren hat.

Insgesamt gingen alle 9 sowohl bereits im Abschlussbericht A05-19C als auch im Arbeitspapier "Memantin bei Alzheimer Demenz" als relevant eingestuften Studien in die Bewertung ein. Von diesen Studien verglichen 7 Memantin in der Monotherapie mit Placebo. Die übrigen 2 Studien verglichen Memantin in Kombination mit einem Cholinesterasehemmer vs. Cholinesterasehemmer plus Placebo.

Sowohl für die beiden patientenrelevanten Endpunkte Aktivitäten des täglichen Lebens und kognitive Leistungsfähigkeit als auch für das Krankheitsstadium gemäß klinischem Eindruck zeigte die Meta-Analyse der Studien einen statistisch signifikanten Effekt zugunsten von Memantin. Damit ergibt sich für den Endpunkt kognitive Leistungsfähigkeit ein Beleg für einen Nutzen von Memantin hinsichtlich des Anteils der Patienten, die eine höchstens irrelevante Verschlechterung erfahren haben. Für den Endpunkt Aktivitäten des täglichen Lebens wurden ausschließlich Responsekriterien herangezogen, die mit einer relativ hohen Unsicherheit behaftet sind. Unter gleichzeitiger Berücksichtigung der geringen Größe des Effekts ergibt sich hier ein Hinweis auf einen Nutzen von Memantin.

#### **Fazit**

Aufgrund der von der Firma Merz im Nachgang berechneten und an den G-BA übermittelten Responderanalysen ergibt sich folgende Änderung des Fazits des Abschlussberichts A05-19C:

Hinsichtlich der Vermeidung einer relevanten Verschlechterung im Bereich der kognitiven Leistungsfähigkeit ergibt sich der Beleg für einen Nutzen von Memantin bei Patienten mit Alzheimer Demenz. Im Bereich der alltagspraktischen Fähigkeiten ergibt sich bei Beachtung der unsicheren Responsekriterien und der gleichzeitig geringen Größe des Effekts ein Hinweis auf einen Nutzen von Memantin.

**Schlagwörter:** Memantin, Alzheimer, Demenz, Nutzenbewertung, Health Technology Assessment, systematische Übersicht

## Inhaltsverzeichnis

			Seite
T	abellenv	erzeichnis	vi
A	bbildun	gsverzeichnis	vii
A	bkürzur	ngsverzeichnis	ix
1	Hinte	ergrund	1
2	Ziele	der Untersuchung	2
3	Proje	ktbearbeitung	3
4	Meth	oden	4
	4.1	Datengrundlage	4
	4.2	Bewertung der Responderanalysen	4
	4.3	Änderungen der Methodik im Vergleich zur Projektskizze	4
5	Erge	bnisse	
	5.1	Der Bewertung zugrunde liegende Datenquellen	
	5.2	Studiencharakteristika	
	5.3	Studien- und Publikationsqualität	6
	5.4	Ergebnisse zu patientenrelevanten Endpunkten	7
	5.4.1	Aktivitäten des täglichen Lebens	7
	5.4.2	Kognitive Leistungsfähigkeit	13
	5.4.3	Krankheitsstadium gemäß klinischem Eindruck	19
	5.4.4	Subgruppenmerkmale und andere Effektmodifikatoren	22
	5.4.5	Sensitivitätsanalyse	23
6	Disk	ıssion	24
7	Fazit		28
8	Liste	der eingeschlossenen Studien	29
9	Liter	atur	30
A	nhang A	– Sensitivitätsanalysen	31
A	nhang B	– Offenlegung potenzieller Interessenkonflikte des externen Reviewers	40
A	nhang (	C – Schreiben der Firma Merz vom 22.10.2010	42

## **Tabellenverzeichnis**

	Seite
Tabelle 1: Ergebnisse zu alltagspraktischen Fähigkeiten	9
Tabelle 2: Ergebnisse zur kognitiven Leistungsfähigkeit	15
Tabelle 3: Ergebnisse zum klinischen Globalurteil	20

## Abbildungsverzeichnis

Abbildung 1: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala
Abbildung 2: Meta-Analyse zu kognitiver Leistungsfähigkeit; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala 19
Abbildung 3: Meta-Analyse zu klinischem Globalurteil; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala
Abbildung 4: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)
Abbildung 5: Meta-Analyse zur kognitiven Leistungsfähigkeit; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)
Abbildung 6: Meta-Analyse zum klinischen Globalurteil; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)
Abbildung 7: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen
Abbildung 8: Meta-Analyse zur kognitiven Leistungsfähigkeit; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen
Abbildung 9: Meta-Analyse zum klinischen Globalurteil; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen

Dagaadagaal	1	1	La: A1.	-1:	D
Responderanaly	sen zu r	viemanum	Dei Aiz	zneimer	Demenz

20	n	9	1	n	١1	1
28.	u		) . ZI	u	ш	

Abbildung 10: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: Ausschluss der Studie MEM-MD-22	37
Abbildung 11: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen; Ausschluss der Studie MEM-MD-22	38
Abbildung 12: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen; Ausschluss der Studie	
MEM-MD-22	39

## Abkürzungsverzeichnis

Abkürzung	Bedeutung
ADAS-cog	Alzheimer's Disease Assessment Scale, cognitive subscale
ADCS-ADL	Alzheimer's Disease Cooperative Study – Activities of Daily Living -sev: for severe impairment
	-J: japanische Version
ADCS-CGIC-CIBIC-plus	Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change – Clinician's Interview-Based Impression of Change (mit Input des Betreuenden)
ADL	activities of daily living (Aktivitäten des täglichen Lebens, alltagspraktische Fähigkeiten)
ANCOVA	Analysis of Co-Variance (Kovarianzanalyse)
BGP	Behavioral Rating Scale for Geriatric Patients
	-C: Cognitive Subscale
	-D: Care Dependency Subscale
СНЕ-Н	Cholinesterasehemmer
FDA	Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
KI	Konfidenzintervall
MID	Minimal Important Difference
MMSE	Mini Mental State Examination
NYU	New York University
OR	Odds Ratio
SD	Standardabweichung
SE	Standardfehler
SIB	Severe Impairment Battery

#### 1 Hintergrund

Im Juli 2009 hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) einen Abschlussbericht zum Thema "Memantin bei Alzheimer Demenz" erstellt (Auftrag A05-19C [1]). Bei der damaligen Bewertung konnten die Ergebnisse der relevanten Studien IE2101 und MEM-MD-22 nicht berücksichtigt werden. Hintergrund war, dass diese beiden Studien nicht vollständig publiziert waren und die Firma Merz die entsprechenden Studienberichte nicht bereitstellte.

Im Zuge der Beratungen zur Umsetzung der Ergebnisse des Abschlussberichts hat der Gemeinsame Bundesausschuss (G-BA) die Nichtübermittlung der unpublizierten Daten kritisiert und die Firma Merz zur Datenübermittlung aufgefordert. Im 1. Quartal 2010 hat die Firma Merz schließlich die Studienberichte der Studien IE2101 und MEM-MD-22 an den G-BA übermittelt. Darüber hinaus hat die Firma Merz mit Schreiben vom 04.03.2010 dem G-BA im Nachgang berechnete und zuvor unpublizierte Responderanalysen vorgelegt. Der G-BA hat das IQWiG gebeten, in einem Arbeitspapier den Einfluss der beiden Studien IE2101 und MEM-MD-22 sowie der Responderanalysen auf das Fazit des Abschlussberichts A05-19C zu prüfen [1]. Diese Prüfung ergab, dass die Responderanalysen in der damals vorgelegten Form das Fazit der Nutzenbewertung A05-19C nicht geändert haben. Wesentliche Gründe dafür waren, dass nicht der vollständige Pool relevanter Studien in die Analysen eingeschlossen war, dass die Ergebnisse der Einzelstudien nicht in adäquaten Metazusammengefasst worden waren und dass zum gewählte Teil Responsekriterium selbst bzw. die Rationale für das Responsekriterium unklar war.

Daraufhin hat die Firma Merz im 4. Quartal 2010 erneut selbst berechnete Responderanalysen vorgelegt. Der G-BA hat das IQWiG damit beauftragt, den Einfluss dieser nachgereichten Responderanalysen auf das Fazit des Abschlussberichts A05-19C zu bewerten.

## 2 Ziele der Untersuchung

Die Fragestellung der vorliegenden Untersuchung ist:

Welchen Einfluss haben die von der Firma Merz im Nachgang berechneten und an den G-BA im 4. Quartal 2010 übermittelten Responderanalysen auf das Fazit des Abschlussberichts A05-19C ("Memantin bei Alzheimer Demenz")?

## 3 Projektbearbeitung

#### Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 04.11.2010 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Auswertung der von der Firma Merz mit Schreiben vom 22.10.2010 beim G-BA eingereichten ergänzenden Responderanalysen zur Nutzenbewertung von Memantin bei Alzheimer Demenz (Auftrag A05-19C) beauftragt.

Die Bearbeitung des Projekts erfolgte durch eine IQWiG-interne Projektgruppe. Auf Basis der Projektskizze wurde ein Rapid Report erstellt. Der Rapid Report wurde einem externen Review unterzogen. Nach Fertigstellung wird der Bericht an den G-BA übermittelt und 4 Wochen später auf der Website des IQWiG veröffentlicht.

#### 4 Methoden

#### 4.1 Datengrundlage

Für die vorliegende Bewertung wurden die von der Firma Merz an den G-BA übermittelten Dokumente herangezogen. Dabei handelte es sich um ein Schreiben der Firma Merz an den G-BA vom 22.10.2010 zu Responderanalysen. Zusätzlich wurden die Informationen herangezogen, die bereits für den Abschlussbericht A05-19C [1] oder für das Arbeitspapier "Memantin bei Alzheimer Demenz" [2] vorlagen. Eine darüber hinausgehende Informationsbeschaffung wurde nicht durchgeführt.

#### 4.2 Bewertung der Responderanalysen

Die Anlage des Schreibens der Firma Merz an den G-BA vom 22.10.2010 enthielt Responderanalysen zu im Abschlussbericht A05-19C aufgeführten patientenrelevanten Endpunkten.

Die Ergebnisse der Responderanalysen wurden gemeinsam mit den sonstigen Aussagen im Schreiben unter Berücksichtigung der Zielsetzung und der Methodik des Abschlussberichts A05-19C [1] bewertet. Dabei wurden auch die im Arbeitspapier benannten Kritikpunkte zu den am 04.03.2010 eingereichten Responderanalysen bei der Bewertung berücksichtigt (siehe Kapitel 1 und [2]).

Bei der Bewertung wurde dargestellt, inwieweit die vorgelegten Informationen das Fazit des Abschlussberichts A05-19C ändern.

#### **Datenextraktion**

Alle für die Bewertung notwendigen Informationen wurden aus den eingereichten Unterlagen in die entsprechenden Tabellen des Arbeitspapiers zu Memantin bei Alzheimer Demenz extrahiert.

#### 4.3 Änderungen der Methodik im Vergleich zur Projektskizze

Im Vergleich zu der in der Projektskizze dargestellten Methode ergaben sich keine Änderungen.

#### 5 Ergebnisse

#### 5.1 Der Bewertung zugrunde liegende Datenquellen

Die Anlage des Schreibens der Firma Merz an den G-BA vom 22.10.2010 enthält Responderanalysen zu den im Abschlussbericht A05-19C aufgeführten patientenrelevanten Endpunkten Beeinträchtigung der Aktivitäten des täglichen Lebens (alltagspraktische Fähigkeiten) und kognitive Leistungsfähigkeit sowie zum im Abschlussbericht A05-19C als ergänzende Information dargestellten (und für das Fazit als nicht relevant angesehenen) Endpunkt Krankheitsstadium gemäß klinischem Eindruck (klinisches Globalurteil). Die Analysen lagen in Form von Meta-Analysen zum einen für die Gesamtstudien und zum anderen für die jeweiligen Subgruppen, stratifiziert nach Schweregrad, vor.

Für jede Studie wurde pro Behandlungsgruppe der Anteil der Non-Responder bestimmt, also der Anteil Patienten, der anhand der für die jeweilige Skala festgelegten Minimal Important Difference (MID) eine relevante klinische Verschlechterung im Vergleich zu Baseline erfahren hat. Da das Odds Ratio als Distanzmaß für den Gruppenunterschied gewählt wurde, sind die Ergebnisse äquivalent zu solchen Auswertungen, in denen der Anteil der Patienten verglichen wird, die im Verlauf der Studie eine höchstens irrelevante Verschlechterung erfahren haben.

Darüber hinaus enthalten die Dokumente auch sogenannte Triple-Responderanalysen, die eine Zusammenfassung der oben genannten Endpunkte darstellen. Da diese Analysen auch den nicht fazitrelevanten Endpunkt klinisches Globalurteil enthalten, sind sie für die vorliegende Bewertung nicht relevant und werden deshalb nicht weiter betrachtet.

Im Arbeitspapier zu Memantin bei Alzheimer Demenz wurde nach der Bewertung der im März 2010 eingereichten Responderanalysen gefordert, dass geeignete zusätzliche Sensitivitätsanalysen berechnet werden, in denen Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) ausgewertet wurden. In den für die vorliegende Bewertung eingereichten Analysen wurden sowohl zu den 3 unter 5.4.1 bis 5.4.3 beschriebenen Endpunkten als auch zu den jeweiligen Subgruppenanalysen, stratifiziert nach Schweregrad der Erkrankung, Sensitivitätsanalysen berichtet, in denen Studienabbrecher, bei denen keine Messwerte zu Studienende vorlagen (im Folgenden Studienabbrecher genannt), als Non-Responder gewertet wurden.

Weiterhin wurde im Arbeitspapier gefordert, dass geeignete Responsekriterien für die verwendeten Skalen verwendet werden. Für solche Skalen, für die keine validierte oder etablierte MID vorliegt, könnte z. B. eine Differenz von 0,5 Standardabweichungen gewählt werden, die sich aus dem Median der gepoolten Standardabweichung in den vorliegenden Studien ergibt. Zumindest für den Bereich Lebensqualität gibt es empirische Evidenz, dass eine MID häufig in diesem Bereich von 0,5 Standardabweichungen liegt [3]. Dies wurde in den vorliegenden Responderanalysen von der Firma Merz berücksichtigt Dennoch muss beachtet werden, dass Responsekriterien, die auf diese Weise ermittelt werden, eine höhere

Unsicherheit aufweisen als solche, die validiert oder durch langjährigen Gebrauch etabliert sind.

#### 5.2 Studiencharakteristika

Das Schreiben der Firma Merz lieferte keine zusätzlichen Informationen zum Studiendesign der eingeschlossenen Studien, zu den jeweiligen Ein- und Ausschlusskriterien sowie zu den Studienpopulationen. Deshalb gelten die Tabellen 1 bis 3 des Arbeitspapiers zu Memantin bei Alzheimer Demenz unverändert.

In die Meta-Analysen gingen alle 9 bereits im Arbeitspapier eingeschlossenen Studien ein. In der Studie LU-10116 wurde der Endpunkt klinisches Globalurteil nicht erhoben, sodass für diesen Endpunkt 8 Studien in die Auswertung eingingen.

#### 5.3 Studien- und Publikationsqualität

Grundsätzlich gelten die in Tabelle 4 des Arbeitspapiers zu Memantin bei Alzheimer Demenz dargestellten Bewertungen zur Studien- und Publikationsqualität der relevanten Studien unverändert.

Darüber hinaus wurde für jeden der 3 berichteten Endpunkte geprüft, ob eine relevante Verletzung des ITT-Prinzips vorlag. Die Auswertungen beziehen sich jeweils auf die ITT-Population, die für jede Studie als adäquat angesehen wurde, da keine relevanten Abweichungen vom ITT-Prinzip vorlagen. Um dem Problem der Studienabbrecher zu begegnen, wurde von der Firma Merz zu jeder Analyse eine Sensitivitätsanalyse mitgeliefert, in der Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) gewertet wurden, jeweils bezogen auf die ITT-Population. In allen Studien war die ITT-Population definiert als die randomisierten und behandelten Patienten, die einen Messwert nach Baseline hatten. Ob dies für die vorliegenden Responderanalysen ebenfalls adäquat ist, ist fraglich. Aus diesem Grund wurden Sensitivitätsanalysen in Bezug auf die jeweilige Zahl der randomisierten Patienten durchgeführt, d. h. sowohl im Nenner als auch im Zähler wurde jeweils die Differenz zwischen der ITT-Population und der Population der randomisierten Patienten addiert.

Weiterhin wurden die übermittelten Responderanalysen auf ihre Plausibilität überprüft. Dazu wurden einerseits die Informationen innerhalb des Schreibens der Firma Merz selbst sowie andererseits die Informationen aus den bereits für den Abschlussbericht A05-19C und das Arbeitspapier zu Memantin bei Alzheimer Demenz vorliegenden Dokumenten herangezogen. In einzelnen Abbildungen der Anlage des Schreibens der Firma Merz gab es zum Teil Diskrepanzen zwischen den grafischen Darstellungen der Effektschätzer mit Konfidenzintervallen und den in Zahlen angegebenen Effektschätzern mit Konfidenzintervallen. So war zum Beispiel in Abbildung 4 der Anlage der Effekt in der Studie 10116 mit OR = 1,10 angegeben, in der grafischen Darstellung lag dieser bei ca. OR = 0,5. Die eigene Berechnung des OR und der zugehörigen 95 %-Konfidenzintervalle entsprach dabei den in Zahlen angegebenen Schätzern. Aufgrund dieser Diskrepanzen wurden alle für diesen Bericht als

relevant angesehenen Meta-Analysen anhand der Angaben zu den Patienten und der Anzahl der Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) pro Behandlungsgruppe und Studie nach den Methoden des Abschlussberichts A05-19C und des Arbeitspapiers zu Memantin bei Alzheimer Demenz neu berechnet. Darüber hinaus fanden sich keine Hinweise auf Inplausibilitäten in den Responderanalysen.

Insgesamt wurde die Wahrscheinlichkeit für eine systematische Verzerrung der berichteten Analysen als niedrig bewertet.

#### 5.4 Ergebnisse zu patientenrelevanten Endpunkten

Im Folgenden werden die Ergebnisse zu den von der Firma Merz vorgelegten Responderanalysen für die 3 genannten Endpunkte dargestellt. Dazu werden zunächst die für die einzelnen Skalen herangezogenen Responsekriterien bewertet und anschließend die jeweiligen Ergebnisse besprochen.

#### 5.4.1 Aktivitäten des täglichen Lebens

#### Bewertung der gewählten Responsekriterien

In den Studien wurden folgende Skalen für die Erhebung der alltagspraktischen Fähigkeiten verwendet: ADCS-ADL-19 (= ADL-sev, ADL-J), ADCS-ADL-23 und BGP-D.

Laut dem Schreiben der Firma Merz fanden sich in der Literatur für keine dieser Skalen Hinweise auf eine validierte oder etablierte Irrelevanzschwelle im Sinne einer MID. Aus diesem Grund wurde von der Firma Merz für diese Skalen gemäß dem Vorschlag im Arbeitspapier jeweils der Median der gepoolten Standardabweichung aus den in die Auswertung eingeschlossenen Studien als Näherung an eine MID verwendet.

Diese Berechnung ergab laut dem Schreiben der Firma Merz für die ADCS-ADL-19 eine MID von 3 Punkten, für die ADCS-ADL-23 eine MID von 5 Punkten und für die BGP-D eine MID von 3 Punkten. Eine Überprüfung dieser Angaben anhand der Informationen aus den eingeschlossenen Studien ergab für die Skalen ADCS-ADL-19 und ADCS-ADL-23 keine abweichende Einschätzung, sodass die gewählten Responsekriterien für die vorliegende Bewertung akzeptiert wurden. Für die BGP-D wurde für die Berechnung von der Firma Merz lediglich die Studie MEM-MD-22 eingeschlossen, obwohl die BGP-D auch in den Studien MEM-MD-01 und MEM-MD-02 erhoben wurde (siehe dazu auch Tabelle 5 des Arbeitspapiers). Allein auf Basis der Studie MEM-MD-22 ergibt sich eine gepoolte Standardabweichung von 2,6, was gerundet die von der Firma Merz als Grenzwert festgelegten 3 Punkte ergibt. Bezieht man die Studien MEM-MD-01 und MEM-MD-02 in die Berechnung mit ein, ergibt sich gerundet ein Grenzwert von 2 Punkten (der Median der gepoolten Standardabweichung beträgt 2,3). Da der Einfluss der unterschiedlichen Grenzwerte auf den Behandlungseffekt unklar ist, wurden Sensitivitätsanalysen zu den alltagspraktischen Fähigkeiten unter Ausschluss der Studie MEM-MD-22 durchgeführt.

## Ergebnisse zu alltagspraktischen Fähigkeiten

Die nachfolgende Tabelle 1 zeigt die Ergebnisse zu den alltagspraktischen Fähigkeiten. Diese Tabelle entspricht der auf Tabelle 15 des Abschlussberichts A05-19C aufbauenden Tabelle 6 des Arbeitspapiers "Memantin bei Alzheimer Demenz" mit Ergänzung der Ergebnisse der Responderanalysen.

Tabelle 1: Ergebnisse zu alltagspraktischen Fähigkeiten

Studie	Gruppe	End-	N <sup>a</sup>	Aı	nalysen auf stetig	ger Ebene		Res	ponderanalysen	
		punkt		Mittlere Differenz zu	Gruppenunter	schied zu Pl	lacebo	Anteil Patienten mit relevanter	Gruppenunte Placebo	rschied zu
				Baseline (SD)	Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
Intervention: N	Memantin vs. 1	Placebo								
10116	Memantin	ADCS- ADL-	124	-0,3 (6,7)	0,47 <sup>b</sup> [-1,08; 2,02]	7	0,55 <sup>b</sup>	37 (29,8)	0,84 [0,49; 1,43]	7
	Placebo	sev	125	-0,8 (6,2)				42 (33,6)		
99679	Memantin	ADCS- ADL <sub>23</sub>	163	-2,76 <sup>b</sup> (10,34 <sup>c</sup> )	1,68 <sup>b</sup> [-0,93; 4,29] <sup>d</sup>	7	$0,206^{b}$	60 (36,8)	1,01 [0,57; 1,80]	ν
	Placebo		71	-4,45 <sup>b</sup> (9,69 <sup>c</sup> )				26 (36,6)		
IE2101	Memantin	ADCS- ADL-J	100	-1,65 (6,18)	n. g.	7	0,838 <sup>e</sup>	40 (40,0)	0,88 [0,51; 1,54]	7
	Placebo		107	-1,64 (5,61)				46 (43,0)		
MEM-MD-01	Memantin	ADCS- ADL-	171	-2,0 <sup>b</sup> (7,85 <sup>c</sup> )	0,7 <sup>b</sup> [-0,6; 2,0]	7	0,282 <sup>b</sup>	62 (36,3)	0,70 [0,45; 1,08]	7
	Placebo	sev	165	-2,7 <sup>b</sup> (7,71 <sup>c</sup> )				74 (44,8)		
MEM-MD-10	Memantin	ADCS- ADL <sub>23</sub>	126	-4,12 <sup>b</sup> (10,78 <sup>c</sup> )	1,14 <sup>b</sup> [-1,09; 3,37] <sup>d</sup>	7	0,321 <sup>b</sup>	53 (42,1)	0,69 [0,42; 1,11]	7
	Placebo		140	-5,26 <sup>b</sup> (10,65 <sup>c</sup> )				72 (51,4)		
MEM-MD-22	Memantin	BGP-D	131 <sup>f</sup>	0,6 <sup>b</sup> (4,80 <sup>g</sup> )	-0,6 [-1,8; 0,6]	7	0,315 <sup>b</sup>	34 (26,0)	0,66 [0,39; 1,13]	7
	Placebo		$130^{\rm f}$	1,2 <sup>b</sup> (5,70 <sup>g</sup> )				45 (34,6)		

(Fortsetzung)

Responderanalysen zu Memantin bei Alzheimer Demenz

Tabelle 1: Ergebnisse zu alltagspraktischen Fähigkeiten (Fortsetzung)

Studie	Gruppe	End-	$N^a$	Aı	nalysen auf stetig	ger Ebene	Responderanalysen			
		punkt		Mittlere Differenz zu	Gruppenunterschied zu Placebo			Anteil Patienten mit relevanter	Gruppenunterschied zu Placebo	
				Baseline (SD)	Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
MRZ-9605	ADL-		126	-3,02 (6,75)	2,06 [0,4; 3,7]	7	0,022 <sup>e</sup>	68 (54,0)	0,70 [0,42; 1,15]	7
	Placebo	sev	126	-5,08 (6,30)				79 (62,7)		
Intervention: N	Memantin + CH	IE-H vs. P	lacebo	+ СНЕ-Н						
MEM-MD-02	Memantin + Donepezil	ADCS- ADL-		-1,4 <sup>b</sup> (6,73 <sup>c</sup> )	1,3 <sup>b</sup> [-0,46; 3,06] <sup>h</sup>	7	0,100 <sup>b</sup>	37 (35,6)	0,90 [0,53; 1,55]	7
	Placebo + Donepezil	sev	124	-2,7 <sup>b</sup> (6,79 <sup>c</sup> )				47 (37,9)		
MEM-MD-12	Memantin + CHE-H	ADCS- ADL <sub>23</sub>	151	-4,15 <sup>b</sup> (8,97 <sup>c</sup> )	0,36 <sup>b</sup> [-1,44; 2,16] <sup>d</sup>	7	0,700 <sup>b</sup>	65 (43,0)	1,05 [0,66; 1,67]	7
	Placebo + CHE-H		146	-4,51 <sup>b</sup> (9,18 <sup>c</sup> )				61 (41,8)		

(Fortsetzung)

28.03.2011

## Responderanalysen zu Memantin bei Alzheimer Demenz

28.03.2011

## Tabelle 1: Ergebnisse zu alltagspraktischen Fähigkeiten (Fortsetzung)

**Fettdruck:** Endpunkte, die in den Studien als primär erklärt wurden, *kursiv*: Angaben laut von Herstellern zur Verfügung gestellten Dokumenten (Studienberichte, Subgruppenanalysen, Stellungnahmen).

grau hinterlegt: Ergänzung im Vergleich zum Abschlussbericht A05-19C und zum Arbeitspapier "Memantin bei Alzheimer Demenz".

Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (\*) oder eine Verschlechterung (\*) gegenüber Placebo bedeutet; er enthält keine Informationen über die Größe oder Signifikanz des Effektes.

- a: Anzahl ausgewerteter Patienten.
- b: ANCOVA adjustiert nach Baseline und Zentrum.
- c: Eigene Berechnung aus Standardfehler.
- d: Eigene Berechnung aus Angaben zu Effekt und Standardfehler.
- e: Wilcoxon-Mann-Whitney-Test.
- f: Je Behandlungsgruppe ein Patient weniger als im Arbeitspapier "Memantin bei Alzheimer Demenz" angegeben. Laut Schreiben der Firma Merz beruht dies auf einer missverständlichen Formulierung in einer Tabelle des Studienberichts der Studie MEM-MD-22.
- g: Unadjustiert.
- h: Eigene Berechnung, asymptotisch.

ADCS-ADL: Alzheimer's Disease Cooperative Study – Activities of Daily Living [-sev: for severe impairment; -J: japanische Version], BGP: Behavioral Rating Scale for Geriatric Patients [-D: Care Dependency Subscale], CHE-H: Cholinesterasehemmer, KI: Konfidenzintervall, Mem: Memantin, n. g.: nicht genannt, SD: Standardabweichung.

Im Vergleich zu den Analysen auf stetiger Ebene war die Richtung der Effektschätzer aus den Responderanalysen in 3 der 9 Studien umgekehrt (99679, IE2101, MEM-MD-12). In diesen 3 Studien zeigten sich jedoch weder auf stetiger Ebene noch in den Responderanalysen statistisch signifikante Unterschiede zwischen den Behandlungsgruppen. Ein im Gegensatz zu der jeweiligen Analyse auf stetiger Ebene nicht signifikanter Effekt zeigte sich in der Studie MRZ-9605.

In keiner der 9 Studien war der Unterschied zwischen den Behandlungsgruppen statistisch signifikant. In der Meta-Analyse der Studien zeigte sich hingegen ein statistisch signifikanter Unterschied zugunsten von Memantin (Abbildung 1). Da bei der betrachteten Indikation und den vorliegenden Responderanalysen davon ausgegangen wird, dass die Responsedefinition eine Bewertung der klinischen Relevanz bereits auf individueller Ebene erlaubt, wurde dieser Unterschied zwischen den Behandlungsgruppen auch gleichzeitig als klinisch relevant gewertet. Im Rahmen einer von der Firma Merz durchgeführten Sensitivitätsanalyse wurden alle Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) ausgewertet. Hier zeigte sich kein abweichendes Ergebnis (Anhang A, Abbildung 4).

Der p-Wert des Interaktionstests zwischen den Subgruppen der Studien, in denen Memantin in Monotherapie gegeben wurde, und den Studien, in denen Memantin in Kombination mit einem Cholinesterasehemmer gegeben wurde, lag bei 0,198. Damit zeigte sich der Hinweis auf eine Effektmodifikation durch die Art der zusätzlichen Behandlungen. Aufgrund der Tatsache, dass die jeweiligen Meta-Analysen der Subgruppen und auch die Gesamtanalyse keine relevante Heterogenität aufwiesen und die beiden Studien zur Kombinationstherapie nur ein Gesamtgewicht von ca. 23 % aufwiesen, wurde dies nicht als relevant angesehen und im Folgenden nicht weiter betrachtet.

Bei der Bewertung des Ergebnisses muss zum einen beachtet werden, dass es für keine der verwendeten Skalen durch langjährigen Gebrauch etablierte oder sogar validierte Responsekriterien gibt. Aus diesem Grund musste für alle Studien auf ein Hilfskonstrukt zur Bestimmung des Responsekriteriums zurückgegriffen werden, nämlich den Median der halben Standardabweichung aus den eingeschlossenen Studien. Daraus ergibt sich eine höhere Unsicherheit der Ergebnisse. Bezieht man in die Bewertung zusätzlich noch die geringe Größe des Effekts in der Meta-Analyse mit ein (die obere Grenze des Konfidenzintervalls des Odds Ratios beträgt 0,96), könnte das statistisch signifikante Ergebnis in hohem Maße von der Wahl des Responsekriteriums abhängen.

In der Gesamtschau ergibt sich insgesamt ein Hinweis auf einen Nutzen von Memantin bezüglich der Vermeidung einer relevanten Verschlechterung im Bereich alltagspraktische Fähigkeiten.

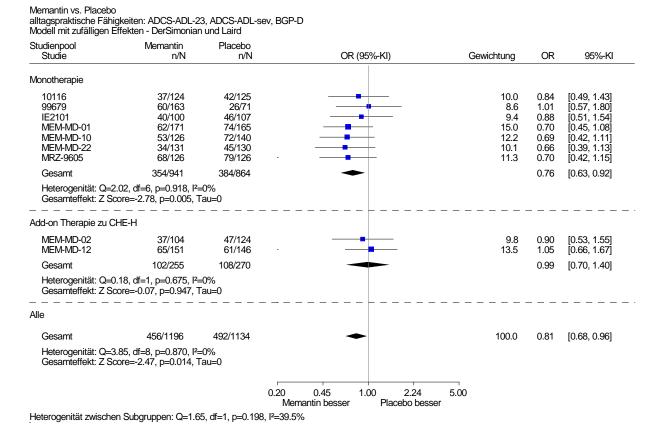


Abbildung 1: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala

#### 5.4.2 Kognitive Leistungsfähigkeit

#### Bewertung der gewählten Responsekriterien

In den Studien wurden folgende Skalen für die Erhebung der kognitiven Leistungsfähigkeit verwendet: ADAS-cog, SIB / SIB-J und BGP-C.

Laut Schreiben der Firma Merz gilt für die ADAS-cog-Skala und die SIB-Skala jeweils eine individuelle Verschlechterung von jeweils 4 bzw. 5 Punkten als relevant, was durch die Zulassungsbehörden anerkannt wird. Entsprechende Literatur wurde von der Firma Merz nicht genannt. Für die BGP-C-Skala, die eine Subskala der BGP darstellt, liegt laut Schreiben der Firma Merz keine solche MID vor.

Für die ADAS-cog-Skala wurde mehrfach als MID für die individuelle Veränderung eine Grenze von 4 Skalenpunkten vorgeschlagen. Allerdings ist fraglich, ob dies der dazu zitierten Expertenrunde der FDA [4] zu entnehmen ist (siehe dazu [5]). Dennoch wurde der Grenzwert von 4 Punkten als MID für die vorliegende Bewertung akzeptiert. Eine eigene Berechnung des Medians der halben Standardabweichung anhand der Informationen aus den eingeschlossenen Studien ergab, dass der Grenzwert von 4 Punkten im Bereich einer halben Standardabweichung lag.

Für die SIB-Skala fanden sich keine Hinweise auf eine validierte oder etablierte MID. Die eigene Berechnung des Medians der halben gepoolten Standardabweichung aus den eingeschlossenen Studien ergab 5 Punkte, sodass der von der Firma Merz verwendete Grenzwert (5 Punkte) auch für die vorliegende Bewertung als Näherung an eine MID akzeptiert wurde.

Für die BGP-C-Skala fanden sich ebenfalls keine Hinweise auf eine MID. Aus der Berechnung des Medians der gepoolten Standardabweichung aus den eingeschlossenen Studien ergab sich laut dem Schreiben der Firma Merz eine MID von 1 Punkt. Eine Überprüfung dieser Angaben anhand der Informationen aus den eingeschlossenen Studien, in denen die BGP-C erhoben wurde, ergab keine abweichende Einschätzung.

#### Ergebnisse zur kognitiven Leistungsfähigkeit

Die nachfolgende Tabelle 2 zeigt die Ergebnisse zur kognitiven Leistungsfähigkeit. Diese Tabelle entspricht der auf Tabelle 16 des Abschlussberichts A05-19C aufbauenden Tabelle 7 des Arbeitspapiers "Memantin bei Alzheimer Demenz" mit Ergänzung der Ergebnisse der Responderanalysen.

Tabelle 2: Ergebnisse zur kognitiven Leistungsfähigkeit

Studie	Gruppe		$N^a$	Aı	nalysen auf stetig	ger Ebene		Res	ponderanalysen	ı
		punkt		Mittlere Differenz zu	Gruppenunter	Gruppenunterschied zu Placebo			Gruppenunterschied zu Placebo	
				Baseline (SD)	Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
Intervention: N	Memantin vs. 1	Placebo								
10116	Memantin	SIB	124	1,4 (10,4)	-0,13 <sup>b</sup> [-2,73; 2,47]	7	0,923 <sup>b</sup>	24 (19,4)	0,83 [0,45; 1,53]	7
	Placebo		125	1,4 (10,9)				28 (22,4)		
99679	Memantin	ADAS- cog	166	-0,89 <sup>b</sup> (7,86 <sup>c</sup> )	-1,11 <sup>b</sup> [-3,01; 0,79] <sup>d</sup>	7	0,251 <sup>b</sup>	30 (18,1)	0,71 [0,37; 1,38]	7
	Placebo		76	0,22 <sup>b</sup> (7,32 <sup>c</sup> )				18 (23,7)		
IE2101	Memantin	SIB-J	100	0,10 (6,54)	n. g.	7	0,005 <sup>e</sup>	24 (24,0)	0,49 [0,27 0,89]	7
	Placebo		107	-3,42 (9,84)				42 (39,3)		
MEM-MD-01	Memantin	SIB	170	-2,0 <sup>b</sup> (13,04 <sup>c</sup> )	0,6 <sup>b</sup> [-1,7; 2,8]	7	0,616 <sup>b</sup>	41 (24,1)	0,53 [0,33; 0,85]	7
	Placebo		165	-2,5 <sup>b</sup> (12,85 <sup>c</sup> )				62 (37,6)		
MEM-MD-10	Memantin	ADAS- cog	125	-0,53 <sup>b</sup> (7,49 <sup>c</sup> )	-2,42 <sup>b</sup> [-3,99; -0,85] <sup>d</sup>	7	0,003 <sup>b</sup>	30 (24,0)	0,52 [0,30; 0,88]	7
	Placebo		140	1,89 <sup>b</sup> (7,45 <sup>c</sup> )				53 (37,9)		
MEM-MD-22	Memantin	BGP-C	131 <sup>f</sup>	0,3 <sup>b</sup> (1,87 <sup>g</sup> )	-0,1 <sup>b</sup> [-0,5; 0,3]	7	$0,699^{b}$	46 (35,1)	0,87 [0,52; 1,43]	7
	Placebo		130 <sup>f</sup>	$0.3^{\rm b} (1.95^{\rm g})$				50 (38,5)		

(Fortsetzung)

Responderanalysen zu Memantin bei Alzheimer Demenz

28.03.2011

Tabelle 2: Ergebnisse zur kognitiven Leistungsfähigkeit (Fortsetzung)

Studie	Gruppe	End-	N <sup>a</sup>	Aı	nalysen auf stetig	ger Ebene		Responderanalysen		
	punkt			Mittlere Differenz zu	Gruppenunterschied zu Placebo			Anteil Patienten mit relevanter	Gruppenunterschied zu Placebo	
				Baseline (SD)	Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
MRZ-9605	Memantin	SIB	126	-3,93 (11,26)	5,91 [2,84; 8,99]	7	< 0,001 <sup>e</sup>	51 (40,5)	0,43 [0,26; 0,72]	7
	Placebo		126	-9,84 (13,43)				77 (61,1)		
Intervention: N	Memantin + CH	IE-H vs. P	lacebo	+ СНЕ-Н						
MEM-MD-02	Memantin + Donepezil	SIB	104	1,6b (8,06°)	1,7 <sup>b</sup> [-0,41; 3,81] <sup>h</sup>	7	0,074 <sup>b</sup>	9 (8,7)	0,46 [0,20; 1,05]	7
	Placebo + Donepezil		123	-0,1b (8,10°)				21 (17,1)		
MEM-MD-12	Memantin + CHE-H	ADAS- cog	149	0,65b (7,20°)	-1,05 <sup>b</sup> [-2,5; 0,40] <sup>d</sup>	7	0,157 <sup>b</sup>	37 (24,8)	0,69 [0,41; 1,15]	7
	Placebo + CHE-H		145	1,71b (7,35°)				47 (32,4)		

(Fortsetzung)

Rapid Report A10-06 Version 1.0

#### Responderanalysen zu Memantin bei Alzheimer Demenz

28.03.2011

## Tabelle 2: Ergebnisse zur kognitiven Leistungsfähigkeit (Fortsetzung)

**Fettdruck:** Zielparameter, die in den Studien als primär erklärt wurden, *kursiv*: Angaben laut von Herstellern zur Verfügung gestellten Dokumenten (Studienberichte, Subgruppenanalysen, Stellungnahmen).

grau hinterlegt: Ergänzung im Vergleich zum Abschlussbericht A05-19C und zum Arbeitspapier "Memantin bei Alzheimer Demenz".

Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (\*\*) oder eine Verschlechterung (\*\*) gegenüber Placebo bedeutet; er enthält keine Informationen über die Größe oder Signifikanz des Effektes.

- a: Anzahl ausgewerteter Patienten.
- b: ANCOVA adjustiert nach Baseline und Zentrum.
- c: Eigene Berechnung aus Standardfehler.
- d: Eigene Berechnung aus Angaben zu Effekt und Standardfehler.
- e: Wilcoxon-Mann-Whitney-Test.
- f: Je Behandlungsgruppe 1 Patient weniger als im Arbeitspapier "Memantin bei Alzheimer Demenz" angegeben. Laut Schreiben der Firma Merz beruht dies auf einer missverständlichen Formulierung in einer Tabelle des Studienberichts der Studie MEM-MD-22.
- g: Unadjustiert.
- h: Eigene Berechnung, asymptotisch.

ADAS-cog: Alzheimer's Disease Assessment Scale – cognitive subscale, BGP: Behavioral Rating Scale for Geriatric Patients [-C: Cognitive Subscale], CHE-H: Cholinesterasehemmer, KI: Konfidenzintervall, Mem: Memantin, n. g.: nicht genannt, SD: Standardabweichung, SIB: Severe Impairment Battery [-J: japanische Version].

In 1 der 9 Studien (10116) zeigten die Responderanalysen und die Auswertungen auf stetiger Ebene numerisch entgegengesetzte Richtungen der Effektschätzer. In dieser Studie zeigte sich aber weder in der Responderanalyse noch in der Auswertung auf stetiger Ebene ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Ein im Gegensatz zu der jeweiligen Analyse auf stetiger Ebene statistisch signifikanter Effekt zeigte sich in 1 Studie (MEM-MD-01).

In insgesamt 4 der 9 Studien war der Unterschied zwischen den Behandlungsgruppen in den Responderanalysen statistisch signifikant zugunsten von Memantin. In der Meta-Analyse der Studien zeigte sich ebenfalls ein statistisch signifikanter und damit auch klinisch relevanter Unterschied zugunsten von Memantin (Abbildung 2). Im Rahmen einer von der Firma Merz durchgeführten Sensitivitätsanalyse wurden alle Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) ausgewertet. Hier zeigte sich kein abweichendes Ergebnis (Anhang A, Abbildung 5).

Ähnlich wie im Bereich alltagspraktische Fähigkeiten muss auch hier bei der Bewertung der kognitiven Leistungsfähigkeit beachtet werden, dass bei etwa 2 Drittel der Studien Skalen verwendet wurden, für die nur unsichere Responsekriterien verwendet wurden (Ableitung aus den Standardabweichungen der vorliegenden Studien). Aufgrund der Größe des Effekts in der Meta-Analyse ist die Wahrscheinlichkeit jedoch gering, dass sich das Ergebnis durch die Wahl eines geringfügig geänderten Responsekriteriums ändern würde.

Insgesamt ergibt sich der Beleg für einen Nutzen von Memantin bezüglich der Vermeidung einer relevanten Verschlechterung im Bereich kognitive Leistungsfähigkeit.

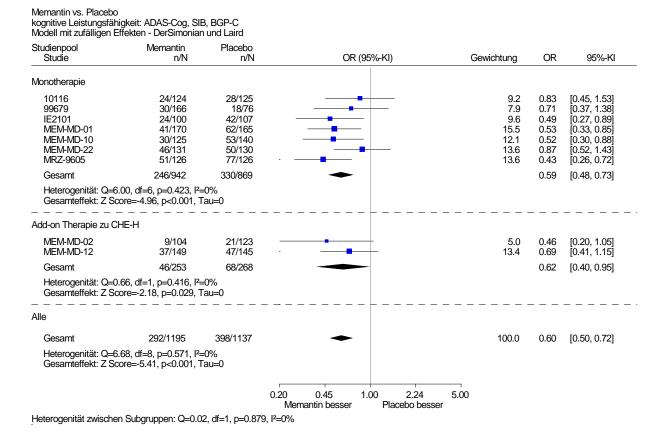


Abbildung 2: Meta-Analyse zu kognitiver Leistungsfähigkeit; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala

#### 5.4.3 Krankheitsstadium gemäß klinischem Eindruck

Dieser Endpunkt wird analog zum Abschlussbericht A05-19C und zum Arbeitspapier "Memantin bei Alzheimer Demenz" als nicht patientenrelevant angesehen und an dieser Stelle lediglich als ergänzende Information dargestellt.

#### Bewertung der gewählten Responsekriterien

In den Studien wurde die CIBIC-plus-Skala zur Erhebung der Veränderung des Krankheitszustandes der Patienten aus Sicht des behandelnden Arztes oder des Betreuers verwendet. Bei CIBIC-plus-Skalen ergibt sich per Definition eine MID für die individuelle Veränderung von  $\geq 1$  Punkt, um ein klinisch relevantes Globalurteil zu fällen. Dieser Grenzwert wurde auch in den vorliegenden Responderanalysen verwendet.

## Ergebnisse zum Krankheitsstadium gemäß klinischem Eindruck

Die nachfolgende Tabelle 3 zeigt die Ergebnisse zum klinischen Globalurteil zur Feststellung des Krankheitsstadiums gemäß klinischem Eindruck. Diese Tabelle entspricht der auf Tabelle 21 des Abschlussberichts A05-19C aufbauenden Tabelle 10 des Arbeitspapiers "Memantin bei Alzheimer Demenz" mit Ergänzung der Ergebnisse der Responderanalysen.

Tabelle 3: Ergebnisse zum klinischen Globalurteil

Studie	Gruppe	End-	N <sup>a</sup>	Aı	nalysen auf stetig	er Ebene		Res	ponderanalysen	
		punkt		Mittlere Differenz zu	Gruppenunter	Gruppenunterschied zu Placebo			Gruppenunterschied zu Placebo	
				Baseline (SD)	Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
Intervention: N	Memantin vs.	Placebo								
99679	Memantin	ADCS- CGIC-	166	4,2 <sup>b</sup> (1,42 <sup>c</sup> )	-0,25 <sup>b</sup> [-0,58; 0,08] <sup>d</sup>	7	0,131 <sup>b</sup>	66 (39,8)	0,77 [0,45; 1,34]	7
	Placebo	CIBIC -plus	76	$4,5^{\rm b} (1,31^{\rm c})$				35 (46,1)		
IE2101	Memantin	NYU- CIBIC-	100	4,38 (1,55)	n. g.	7	0,339 <sup>e</sup>	51 (51,0)	0,76 [0,44; 1,31]	7
	Placebo	plus-J	107	4,62 (1,26)				62 (57,9)		
MEM-MD-01	Memantin	ADCS- CGIC-	171	4,3 (1,0)	-0,30 [-0,51; -0,09]	7	0,182 <sup>f</sup>	70 (40,9)	0,77 [0,50; 1,19]	7
	Placebo	CIBIC- plus	163	4,6 (1,0)				77 (47,2)		
MEM-MD-10	Memantin	ADCS- CGIC-	126	4,28 <sup>b</sup> (1,12 <sup>c</sup> )	-0,30 <sup>b</sup> [-0,54; -0,07] <sup>d</sup>	7	0,014 <sup>b</sup>	49 (38,9)	0,48 [0,30; 0,79]	7
	Placebo	CIBIC -plus	139	4,58 <sup>b</sup> (1,18 <sup>c</sup> )				79 (56,8)		
MEM-MD-22	Memantin	ADCS- CGIC-	132	3,6 (1,36)	-0,1 <sup>b</sup> [-0,5; 0,2]	7	0,445 <sup>b</sup>	30 (22,7)	0,95 [0,54; 1,68]	7
	Placebo	CIBIC- plus	131	3,7 (1,29)				31 (23,7)		
MRZ-9605	Memantin	NYU- CIBIC	126	4,48 (1,09)	-0,25 [-0,5; 0,0]	7	0,064 <sup>e</sup>	57 (45,2)	0,66 [0,40;1,09]	7
	Placebo	-plus	126	4,73 (1,07)				70 (55,6)		

(Fortsetzung)

Responderanalysen zu Memantin bei Alzheimer Demenz

28.03.2011

Tabelle 3: Ergebnisse zum klinischen Globalurteil (Fortsetzung)

Studie	Gruppe	End- punkt	Nª	Analysen auf stetiger Ebene				Responderanalysen		
				Mittlere Differenz zu Baseline (SD)	Gruppenunterschied zu Placebo			Anteil Patienten mit relevanter	Gruppenunterschied zu Placebo	
					Differenz [95 %-KI]	Effekt- richtung	p-Wert	Verschlechterung n (%)	Odds Ratio [95 %-KI]	Effekt- richtung
Intervention: Memantin + CHE-H vs. Placebo + CHE-H										
MEM-MD-02	Memantin + Donepezil	ADCS- 104 CGIC- CIBIC- plus 123	4,2 (1,01)		0,65 [0,38; 1,10]	7				
	Placebo + Donepezil		123	3 4,5 (0,99)				58 (47,2)		
MEM-MD-12	Memantin + CHE-H	CGIC-	151	4,49 <sup>b</sup> (1,11 <sup>c</sup> )	-0,06 <sup>b</sup> [-0,30; 0,18] <sup>d</sup>	7	0,585 <sup>b</sup>	67 (44,4)	0,68 [0,43; 1,07]	7
	Placebo + CHE-H	CIBIC -plus	146	4,55 <sup>b</sup> (1,21 <sup>c</sup> )				79 (54,1)		

**Fettdruck:** Zielparameter, die in den Studien als primär erklärt wurden, *kursiv*: Angaben laut von Herstellern zur Verfügung gestellten Dokumenten (Studienberichte, Subgruppenanalysen, Stellungnahmen).

grau hinterlegt: Ergänzung im Vergleich zum Abschlussbericht A05-19C und zum Arbeitspapier "Memantin bei Alzheimer Demenz".

Der Pfeil stellt dar, ob die numerische Veränderung auf der jeweiligen Skala eine Verbesserung (\*) oder eine Verschlechterung (\*) gegenüber Placebo bedeutet; er enthält keine Informationen über die Größe oder Signifikanz des Effektes.

- a: Anzahl ausgewerteter Patienten.
- b: ANCOVA adjustiert nach Zentrum (entspricht nicht der geplanten Auswertung laut Studienprotokoll).
- c: Eigene Berechnung aus Standardfehler.
- d: Eigene Berechnung aus Angaben zu Effekt und SE.
- e: Wilcoxon-Mann-Whitney-Test.
- f: Van-Elteren-Test adjustiert nach Zentrum.

ADCS-CGIC-CIBIC-plus: Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change - Clinician's Interview-Based Impression of Change (mit Input des Betreuenden), CHE-H: Cholinesterasehemmer, KI: Konfidenzintervall, n. g.: nicht genannt, NYU-CIBIC-plus: New York University - Clinician's Interview-Based Impression of Change (mit Input des Betreuenden) [-J: japanische Version], SD: Standardabweichung.

Der Endpunkt klinisches Globalurteil wurde in 8 der 9 Studien untersucht.

Im Vergleich zu den Ergebnissen der Analysen auf stetiger Ebene zeigten die Effektschätzer aus den Responderanalysen in jeder der 8 Studien in die gleiche Richtung.

In 1 Studie (MEM-MD-10) zeigte sich ein statistisch signifikanter Effekt zugunsten von Memantin, in den übrigen 7 Studien war der Unterschied zwischen den Behandlungsgruppen nicht statistisch signifikant. Eine Meta-Analyse der Studien zeigte einen statistisch signifikanten Effekt zugunsten von Memantin (Abbildung 3). Im Rahmen einer von der Firma Merz durchgeführten Sensitivitätsanalyse wurden alle Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) ausgewertet. Hier zeigte sich kein abweichendes Ergebnis (Anhang A, Abbildung 6).

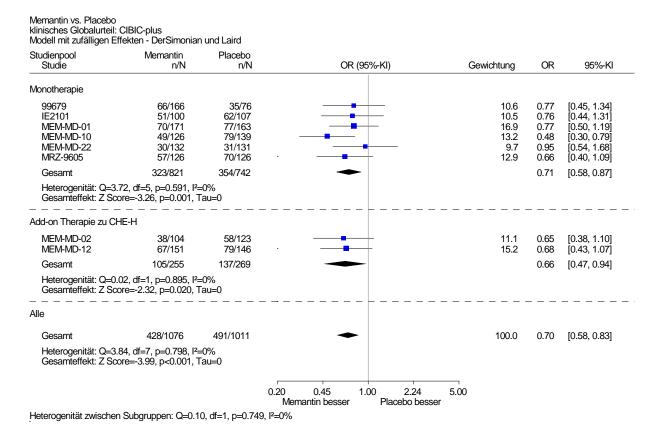


Abbildung 3: Meta-Analyse zu klinischem Globalurteil; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala

#### 5.4.4 Subgruppenmerkmale und andere Effektmodifikatoren

#### Subgruppe "Schweregrad der Erkrankung"

Zu jedem der unter 5.4.1 bis 5.4.3 beschriebenen Endpunkte lagen auch stratifizierte Auswertungen nach dem Merkmal "Schweregrad der Erkrankung" (gemessen an der MMSE) vor (siehe Anhang C). Die in die Studien eingeschlossenen Patienten wurden eingeteilt in

solche mit moderater Alzheimer Demenz (MMSE zwischen 10 und 19) und solche mit schwerer Alzheimer Demenz (MMSE < 10). Für keinen der 3 Endpunkte war der auf Basis einer Meta-Regression berechnete Interaktionstest statistisch signifikant (jeweils p > 0,2).

Damit gibt es keinen Hinweis auf spezifische Behandlungseffekte für die unterschiedlichen Schweregrade der Alzheimer Demenz.

#### 5.4.5 Sensitivitätsanalyse

Über die bereits in den Abschnitten 5.4.1 bis 5.4.3 besprochenen Sensitivitätsanalysen unter Auswertung aller Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) hinaus wurden weitere, eigene Sensitivitätsanalysen durchgeführt. Alle Sensitivitätsanalysen finden sich in Anhang A.

- Auswertung aller Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) in Bezug auf alle randomisierten Patienten (Abbildung 7 bis Abbildung 9)
- Analysen für den Endpunkt alltagspraktische Fähigkeiten jeweils unter Ausschluss der Studie MEM-MD-22 aufgrund der abweichenden Bewertung des Grenzwertes für eine relevante individuelle Verschlechterung (Abbildung 10 bis Abbildung 12)

Keine dieser Sensitivitätsanalysen zeigte ein abweichendes Ergebnis im Vergleich zu den in den Abschnitten 5.4.1 bis 5.4.3 besprochenen Analysen.

#### 6 Diskussion

Bereits im 1. Quartal 2010 wurden von der Firma Merz Responderanalysen zu Memantin bei Alzheimer Demenz übermittelt. Aufgrund von erheblichen Mängeln in diesen Analysen wurde im Arbeitspapier "Memantin bei Alzheimer Demenz" trotz eines statistisch signifikanten Effektes zugunsten von Memantin hinsichtlich alltagspraktischer Fähigkeiten und der kognitiven Leistungsfähigkeit kein Nutzen von Memantin ausgesprochen. Diese Mängel waren im Wesentlichen, dass nicht der vollständige Pool relevanter Studien in die Analysen eingeschlossen war, dass die Ergebnisse der Einzelstudien nicht in adäquaten Metazusammengefasst worden waren und dass zum Teil Responsekriterium selbst bzw. die Rationale für das Responsekriterium unklar war. In den infolgedessen im 4. Quartal 2010 nachgelieferten Responderanalysen, die die Grundlage für die vorliegende Bewertung darstellen, wurden diese Mängel im Wesentlichen berücksichtigt. Hinsichtlich des Anteils der Patienten, die im Studienverlauf eine höchstens nicht relevante Verschlechterung in der kognitiven Leistungsfähigkeit erfahren haben, ergibt sich aus diesen Responderanalysen nun der Beleg für einen Nutzen von Memantin bei Alzheimer Demenz. Im Bereich der alltagspraktischen Fähigkeiten ergibt sich in der Gesamtschau unter Beachtung der unsicheren Responsekriterien und der gleichzeitig geringen Größe des Effekts ein Hinweis auf einen Nutzen von Memantin.

#### Vergleich der Responderanalysen mit denjenigen auf stetiger Ebene

Sowohl die alltagspraktischen Fähigkeiten als auch die Kognition wurden jeweils mit komplexen Skalen erhoben. Insbesondere bei solchen Endpunkten ist es notwendig, neben der statistischen Signifikanz der Effekte die Relevanz der beobachteten Wirkungen der untersuchten Interventionen zu bewerten, da aufgrund der Komplexität der Skalen häufig eine sinnvolle Interpretation schwierig ist. Gerade bei geringfügigen Gruppenunterschieden ist oft nicht klar, was Änderungen auf der Skala für die Patienten bedeuten.

Die interessierende Frage für eine Nutzenbewertung ist daher, ob der beobachtete Unterschied zwischen 2 Behandlungsgruppen überhaupt für die Patienten eine Bedeutung hat. Der Idealfall wäre hier, dass eine validierte Irrelevanzschwelle für die Differenz der Mittelwerte der Änderung des Scores von Studienbeginn bis -ende vorliegt. In diesem Fall könnte man den Behandlungseffekt (inklusive des 95 %-Konfidenzintervalls) mit dieser Relevanzschwelle vergleichen. Läge das Konfidenzintervall jenseits dieser Schwelle, wäre der Effekt relevant. Solche Irrelevanzschwellen lagen aber für die in der vorliegenden Bewertung verwendeten Skalen nicht vor. In einer systematischen Übersicht kann die Relevanz nur anhand der vorliegenden Informationen aus den Studien beurteilt werden. Im Abschlussbericht A05-19C wurde als Annäherung an eine solche Irrelevanzschwelle deshalb eine Grenze von 0,2 Standardabweichungen angewendet. Ein Effekt galt als relevant, wenn das 95 %-Konfidenzintervall der standardisierten Mittelwertdifferenz jenseits dieser 0,2 Standardabweichungen lag. Von verschiedenen Seiten wurde kritisiert, dass es sich dabei um ein rein statistisches Maß handele. In einer systematischen Übersicht wurden allerdings über verschiedene Indikationen hinweg 0,5 Standardabweichungen als relevante individuelle Schwelle beobachtet [3]. Eine feste Schwelle über Indikationen hinweg kann also durchaus inhaltlich sinnvoll sein. Als eine mögliche Alternative einer Annäherung können Responderanalysen dienen [6]. Hier ist es erforderlich, dass ein validiertes bzw. etabliertes Responsekriterium angewendet wurde (im Sinne einer individuellen MID) [7]. Bei den vorliegenden Auswertungen zu Memantin wurde ein statistisch signifikanter Unterschied bezüglich der Anteile der Responder zwischen den Gruppen als Nachweis eines relevanten Effekts angesehen werden, da davon ausgegangen wurde, dass die Responsedefinition bereits eine Schwelle der Relevanz auf individueller Ebene beinhaltet. Auch wenn beiden Arten der Auswertung, also Analysen auf der Ebene von Gruppenunterschieden mit Anwendung einer verschobenen Nullhypothese und Responderanalysen, verschiedene statistische Konzepte zugrunde liegen, dienen sie in der vorliegenden Bewertung gleichsam als Annäherung an den oben beschriebenen Idealfall für die Bewertung der Relevanz des Gruppenunterschieds.

Aus diesem Grund können die vorliegenden Responderanalysen mit den für das Arbeitspapier "Memantin bei Alzheimer Demenz" vorliegenden Analysen auf stetiger Ebene verglichen werden.

Bei Betrachtung der Forestplots zu den einzelnen Endpunkten zeigte sich jeweils ein ähnliches Bild. Hinsichtlich der Richtung der Effektschätzer und der statistischen Signifikanz gab es zumeist keine wesentlichen Unterschiede zwischen den Analysen, in denen die Mittelwerte der Behandlungsgruppen verglichen wurden, und den Responderanalysen. Nur in wenigen Einzelstudien kam es bei den untersuchten Endpunkten zu einer umgekehrten Richtung der Effektschätzer. War dies der Fall, waren jedoch beide Effekte nicht statistisch signifikant, häufig lagen beide Effektschätzer nahe am jeweiligen Nulleffekt. Traten in Einzelstudien statistisch signifikante Effekte auf, dann zumeist in beiden Analysen. Einen auffälligen Unterschied zwischen den Analysen auf stetiger Ebene und Responderanalysen gab es in der Studie MEM-MD-01 für den Endpunkt Kognition. In den Analysen auf stetiger Ebene war der Gruppenunterschied nicht statistisch signifikant, der Effektschätzer lag sehr nahe beim Nulleffekt. In den Responderanalysen zeigte sich hingegen ein statistisch signifikanter Unterschied zugunsten von Memantin mit einem OR = 0,53. Eine Erklärung für diese Diskrepanz könnte sein, dass für diesen Endpunkt eine schiefe Verteilung zugrunde liegt. Diese Vermutung wird unterstützt durch Aussagen im Studienbericht zu MEM-MD-01 [8]. Die Autoren berichteten, dass sich bei der Datenanalyse die Annahme der Normalverteilung der Residuen des ANCOVA-Modells für die SIB-Werte als nicht gerechtfertigt erwies und eine post hoc durchgeführte explorative Analyse mithilfe des Wilcoxon-Mann-Whitney-Tests zu einem statistisch signifikanten Ergebnis zugunsten von Memantin führte.

Vergleicht man die Schlussfolgerungen, die sich in der Nutzenbewertung aus den Analysen auf stetiger Ebene und den vorliegenden Responderanalysen, auch unter Berücksichtigung der klinischen Relevanz der Effekte, ergeben haben, zeigte sich auf stetiger Ebene zwar ein statistisch signifikanter Effekt zugunsten von Memantin, dessen klinische Relevanz wurde aber als unklar bewertet, da die 95 %-Konfidenzintervalle der standardisierten Mittelwertdifferenzen nicht jenseits von 0,2 Standardabweichungen lagen. Für die vorliegenden Responderanalysen, bei denen die Bewertung der klinischen Relevanz bereits

auf der Ebene des individuellen Patienten vorgenommen wurde, wurde bei einem statistisch signifikanten Effekt gleichzeitig auch ein klinisch relevanter Effekt angenommen. Diese Nutzenbewertung zeigt, dass die Anwendung von Responderanalysen ein deutlich milderes Kriterium für die Bewertung der klinischen Relevanz darstellen kann.

Insgesamt unterstreicht diese Bewertung noch einmal die Wichtigkeit, dass für eine Bewertung von Studienergebnissen von komplexen Skalen adäquate Analysen vorliegen, um auch die klinische Relevanz adäquat beurteilen zu können, zum Beispiel durch Responderanalysen auf Basis einer validierten MID. Hinzu kommt, dass der Anteil von Patienten, die eine relevante Änderung ihres Krankheitszustandes erfahren, deutlich besser interpretierbar ist als Mittelwerte über die Differenz des Scores auf einer komplexen Skala.

Dennoch muss beachtet werden, dass für keine der in die Bewertung eingeschlossenen Skalen eine validierte und nur für einen Bruchteil etablierte MID vorliegen. Deshalb stützt sich auch bei diesen Skalen die Bewertung der klinischen Relevanz auf ein Hilfskonstrukt, nämlich die Verwendung des Medians der halben Standardabweichung aus den eingeschlossenen Studien. Dieses Problem stellte sich insbesondere für den Endpunkt alltagspraktische Fähigkeiten, bei dem in keiner der Studien ein Instrument verwendet wurde, bei dem auf ein etabliertes oder validiertes Responsekriterium zurückgegriffen werden konnte.

#### **BGP-C in Studie MEM-MD-22**

Im Arbeitspapier "Memantin bei Alzheimer Demenz" wurde die kognitive Subskala der BGP (BGP-C) als Instrument zur Messung der kognitiven Leistungsfähigkeit eingeschlossen. Diese wurde lediglich in der Studie MEM-MD-22 erhoben. In ihrem Schreiben wies die Firma Merz darauf hin, dass es sich bei der BGP-C nicht um ein adäquates Instrument für diesen Endpunkt handele, da sie nur sehr allgemeine Fragen zur örtlichen und zeitlichen Orientierung enthält und die Einschätzung durch das Pflegepersonal erfolgt. Auch die US-amerikanische Zulassungsbehörde FDA äußerte Zweifel, dass diese Subskala ein adäquates Instrument zur Messung der Kognition darstellt [4]. Aus diesem Grund hat die Firma Merz in ihrem Schreiben zusätzliche Sensitivitätsanalysen zum Endpunkt kognitive Leistungsfähigkeit unter Ausschluss der Ergebnisse der Studie MEM-MD-22 übermittelt (siehe Anhang C).

Diese Skala ist im Arbeitspapier in Ermangelung eines alternativen Instruments zur Messung der Kognition für die Studie MEM-MD-22 eingeschlossen worden, um eine möglichst vollständige Datenbasis zu haben.

Doch unabhängig davon hatten die eingeschlossenen Ergebnisse zur BGP-C keinen relevanten Einfluss auf das Gesamtergebnis in der Kognition. Zum einen führte der Einschluss der Studie MEM-MD-22 in der Meta-Analyse nicht zum Auftreten von Heterogenität. Zum anderen änderte sich auch unter Ausschluss der Ergebnisse zur BGP-C im Rahmen der Sensitivitätsanalysen das Ergebnis zur Kognition nahezu nicht (siehe Anhang C).

# Abschließende Bemerkung

Insgesamt muss festgehalten werden, dass seit der Veröffentlichung keine neuen Daten zum Nutzen von Memantin vorliegen, sondern in die vorliegenden Responderanalysen nur Daten aus Studien eingeflossen sind, die bereits zur Zeit der Veröffentlichung des Vorberichts A05-19C finalisiert waren. Daraus folgt, dass die vorliegenden Analysen auch bereits früher und nicht erst ca. 2 Jahre nach Veröffentlichung des Vorberichts hätten vorgelegt werden können.

Darüber hinaus gilt nach wie vor, dass die eingeschlossenen Studien nur einen Zeitraum von höchstens 6 Monaten abdecken, sodass über den Langzeitnutzen von Memantin keine Aussagen getroffen werden können.

# 7 Fazit

Aufgrund der von der Firma Merz im Nachgang berechneten und an den G-BA übermittelten Responderanalysen ergibt sich folgende Änderung des Fazits des Abschlussberichts A05-19C:

Hinsichtlich der Vermeidung einer relevanten Verschlechterung im Bereich der kognitiven Leistungsfähigkeit ergibt sich der Beleg für einen Nutzen von Memantin bei Patienten mit Alzheimer Demenz. Im Bereich der alltagspraktischen Fähigkeiten ergibt sich bei Beachtung der unsicheren Responsekriterien und der gleichzeitig geringen Größe des Effekts ein Hinweis auf einen Nutzen von Memantin.

# 8 Liste der eingeschlossenen Studien

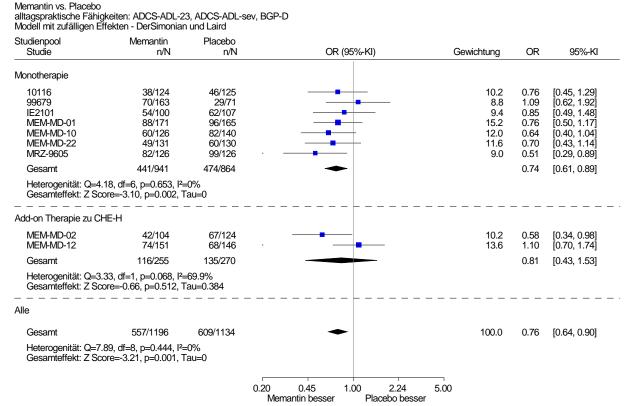
Die Datenquellen für die in die Bewertung aufgenommenen Studien können dem Abschlussbericht A05-19C und dem Arbeitspapier "Memantin bei Alzheimer Demenz" sowie dem Anhang C des vorliegenden Berichts entnommen werden.

#### 9 Literatur

- 1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Memantin bei Alzheimer Demenz: Abschlussbericht; Auftrag A05-19C [online]. 08.07.2009 [Zugriff: 15.06.2010]. (IQWiG-Berichte; Band 59). URL: <a href="http://www.iqwig.de/download/A05-19C\_Abschlussbericht\_Memantin\_bei\_Alzheimer\_Demenz.pdf">http://www.iqwig.de/download/A05-19C\_Abschlussbericht\_Memantin\_bei\_Alzheimer\_Demenz.pdf</a>.
- 2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Memantin bei Alzheimer Demenz: Ergebnisse der unpublizierten Studien IE2101 und MEM-MD-22 sowie unpublizierter Responderanalysen; Arbeitspapier [online]. 01.07.2010 [Zugriff: 02.08.2010]. (IQWiG-Berichte; Band 74). URL: <a href="http://www.iqwig.de/download/Arbeitspapier\_zu\_Memantin\_bei\_Alzheimer\_Demenz.pdf">http://www.iqwig.de/download/Arbeitspapier\_zu\_Memantin\_bei\_Alzheimer\_Demenz.pdf</a>.
- 3. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003; 41(5): 582-592.
- 4. Food and Drug Administration. Final briefing document for Advisory Committee Meeting: efficacy review of new drug application [online]. 19.08.2003 [Zugriff: 24.02.2011]. URL: <a href="http://www.fda.gov/ohrms/dockets/ac/03/briefing/3979B1\_05">http://www.fda.gov/ohrms/dockets/ac/03/briefing/3979B1\_05</a> FDA-Efficacy%20Review.pdf.
- 5. Kreis J, Kaiser T, Wieseler B. If shrimps had lips, they could whistle. J Am Geriatr Soc 2009; 57(10): 1947.
- 6. Schünemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcome measures in clinical trials: the clinician's perspective. Health and quality of life outcomes 2006; 4: 62.
- 7. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008; 61(2): 102-109.
- 8. Forest Laboratories. A randomized, double-blind, placebo-controlled evaluation of the safety and efficacy of memantine in patients with moderate to severe dementia of the Alzheimer's type: abbreviated study report; study no. MEM-MD-01 [unveröffentlicht]. 2004.

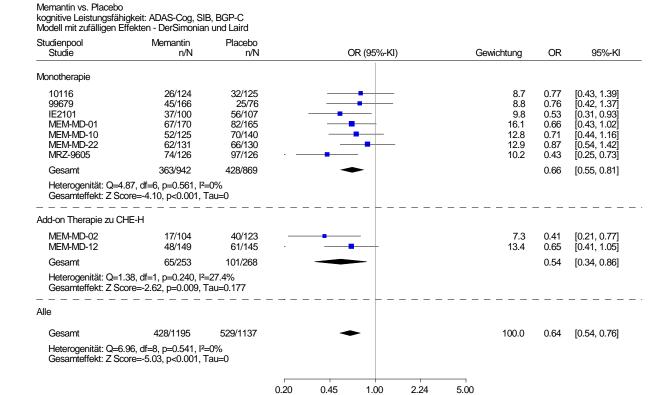
# Anhang A – Sensitivitätsanalysen

# Auswertung aller Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) bezogen auf die ITT-Population



Heterogenität zwischen Subgruppen: Q=0.39, df=1, p=0.534, l²=0%

Abbildung 4: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)



Heterogenität zwischen Subgruppen: Q=0.71, df=1, p=0.398, l2=0%

Abbildung 5: Meta-Analyse zur kognitiven Leistungsfähigkeit; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)

Memantin besser

Placebo besser

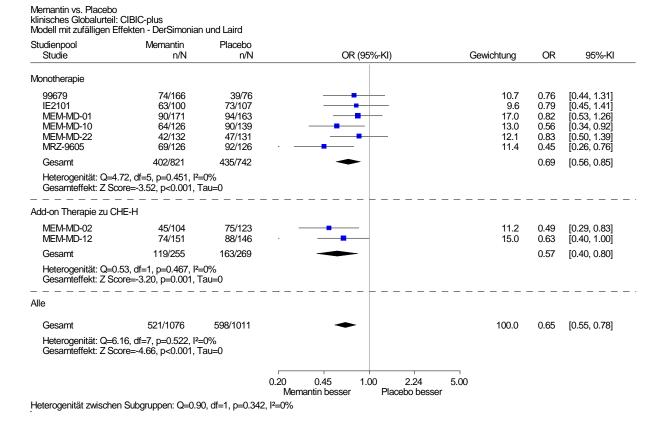


Abbildung 6: Meta-Analyse zum klinischen Globalurteil; Anteil der Patienten mit einer relevanten klinischen Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: alle Studienabbrecher wurden als Non-Responder eingestuft (im Sinne einer relevanten Verschlechterung)

# Auswertung aller Studienabbrecher als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) in Bezug auf alle randomisierten Patienten

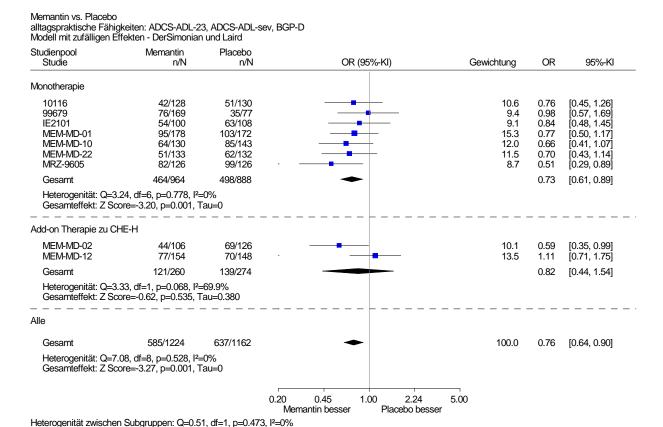


Abbildung 7: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen

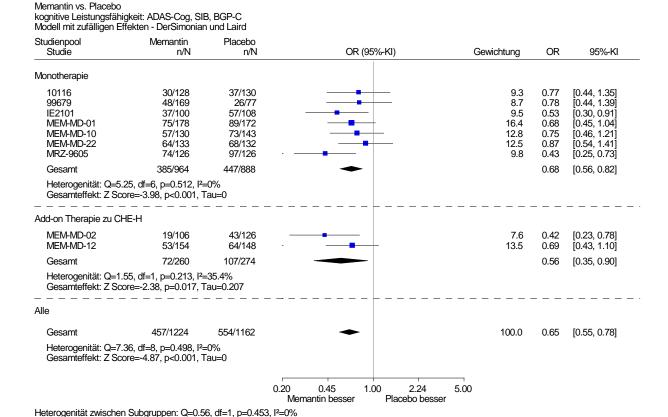


Abbildung 8: Meta-Analyse zur kognitiven Leistungsfähigkeit; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen

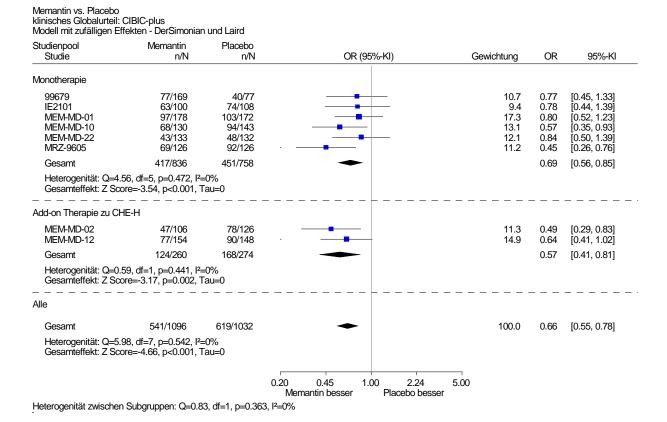


Abbildung 9: Meta-Analyse zum klinischen Globalurteil; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen

# Sensitivitätsanalysen für den Endpunkt alltagspraktische Fähigkeiten jeweils unter Ausschluss der Studie MEM-MD-22

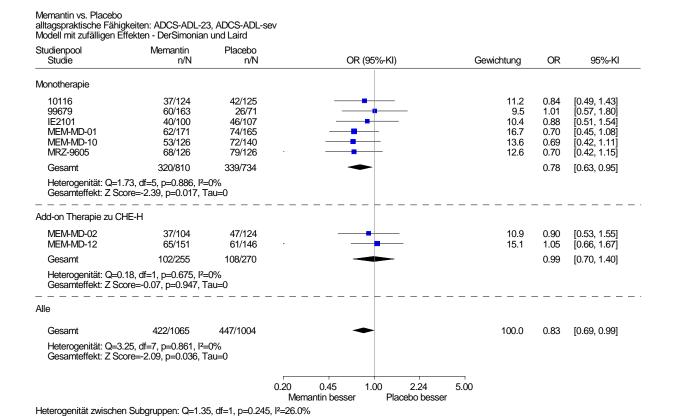


Abbildung 10: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: Ausschluss der Studie MEM-MD-22

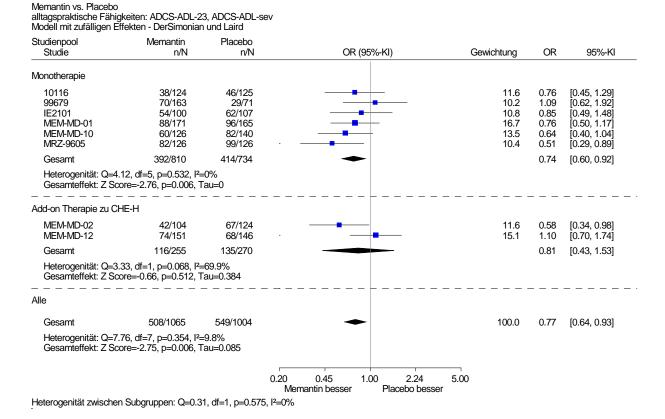


Abbildung 11: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen; Ausschluss der Studie MEM-MD-22

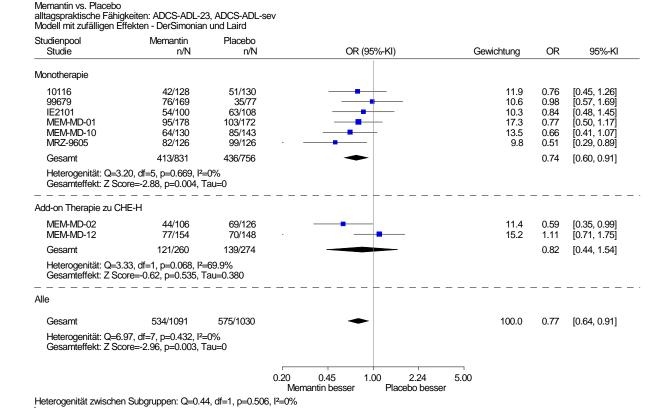


Abbildung 12: Meta-Analyse zu alltagspraktischen Fähigkeiten; Anteil der Patienten mit relevanter Verschlechterung (jeweils ≥ MID) auf der jeweiligen Skala; Sensitivitätsanalyse: die jeweilige Zahl der randomisierten Patienten wird als Grundgesamtheit angenommen und alle Studienabbrecher werden als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) angesehen; Ausschluss der Studie MEM-MD-22

# Anhang B - Offenlegung potenzieller Interessenkonflikte des externen Reviewers

Im Folgenden sind potenziellen Interessenkonflikte des externen Reviewers dargestellt. Alle Informationen beruhen auf Selbstangaben an Hand des "Formblatts zur Offenlegung potenzieller Interessenkonflikte". Das Formblatt ist unter <a href="http://www.iqwig.de">http://www.iqwig.de</a> abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

#### **Externer Reviewer**

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Trampisch, Hans J.	nein	nein	nein	nein	nein	nein

Im "Formblatt zur Offenlegung potenzieller Interessenkonflikte" wurden folgende 6 Fragen gestellt:

*Frage 1:* Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter<sup>1</sup> abhängig (angestellt) beschäftigt? Falls ja, wo und in welcher Position?

*Frage 2:* Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor einen Interessenverband im Gesundheitswesen oder einen vergleichbaren Interessenvertreter direkt oder indirekt beraten? Falls ja, wen und wie hoch ist / war die Zuwendung / das Honorar?

Frage 3: Haben Sie abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres oder der 3 Kalenderjahre davor im Auftrag eines Interessenverbands im Gesundheitswesen oder eines vergleichbaren Interessenvertreters Honorare für Vorträge, Stellungnahmen, Ausrichtung und / oder Teilnahme an Kongressen und Seminaren – auch im Rahmen von Fortbildungsveranstaltungen, oder für (populär-)wissenschaftliche oder sonstige Aussagen oder Artikel erhalten? Falls ja, von wem, für welche Tätigkeiten und wie hoch war die Zuwendung / das Honorar?

Frage 4: Haben Sie abseits einer Anstellung oder Beratungstätigkeit und / oder hat die Institution<sup>2</sup>, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten? Falls ja, von wem, für welche Tätigkeit und in welcher Höhe?

*Frage 5:* Haben Sie und / oder hat die Institution, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres oder der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Reisekostenunterstützung ohne wissenschaftliche Gegenleistungen) von einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter erhalten? Falls ja, von wem, aus welchem Anlass und in welcher Höhe?

*Frage 6:* Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile (auch in Fonds) von einer Firma oder Institution, die zu einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter gehört? Falls ja, von wem und welchen Wert haben diese aktuell?

<sup>2</sup> Sofern Sie in einer ausgedehnten Institution tätig sind, ist es ausreichend, die geforderten Angaben auf Ihre Arbeitseinheit (z. B.: Klinikabteilung, Forschungsgruppe etc.) zu beziehen.

\_\_\_

<sup>&</sup>lt;sup>1</sup> Dieses Formblatt erfasst finanzielle Beziehungen zu Interessenverbänden im Gesundheitswesen oder vergleichbaren Interessenvertretern, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie.

Rapid	Report	A10-06
-------	--------	--------

Version 1.0

Responderanalysen zu Memantin bei Alzheimer Demenz

28.03.2011

# Anhang C – Schreiben der Firma Merz vom 22.10.2010

#### MERZ PHARMACEUTICALS



Merz Pharmaceuticals GmbH · Postfach 1113 53 D-60048 Frankfurt am Main

Herrn Dr. jur. Rainer Hess Vorsitzender des Gemeinsamen Bundesausschusses Herrn Thomas Müller Abteilungsleiter Arzneimittel Wegelystr. 8 10623 Berlin

Eckenheimer Landstraße 100 60318 Frankfurt am Main Telefon: 0 69 / 15 03 - 6461 Telefax: 0 69 / 15 03 - 96461 claus.goebel@merz.de www.merz.de

Cc: IQWiG (Herr Dr. T. Kaiser Herr PD Dr. Lange Herr Dr. Skipka)

22.10.2010

Nutzenbewertungsverfahren Memantin – Vorlage der ergänzenden Responderanalysen

Sehr geehrter Herr Dr. Hess, sehr geehrter Herr Müller,

wie zwischen Herrn Dr. Zügel und Herrn Müller vereinbart, senden wir Ihnen die im Arbeitspapier Memantin vom IQWiG angesprochenen, ergänzenden Responderanalysen in der Anlage zu. Diese Analysen sind auch Teil unserer offiziellen Stellungnahme, welche wir in der kommenden Woche bei der Abteilung Nutzenbewertung des G-BA einreichen werden.

Wichtige methodische Fragen, insbesondere zu den Relevanzgrenzen einiger in den Studien verwendeter Skalen und zu Sensitivitätsanalysen der Studienabbrecher, konnten wir dankenswerter Weise in einem wissenschaftlichen Gespräch am 1. Oktober 2010 beim IQWiG erörtern.

Wir möchten an dieser Stelle auch unser Bedauern zum Ausdruck bringen, dass unsere Aussage in der Korrespondenz mit Ihrem Hause, dass es in Studie MEM-MD-22 keinen Endpunkt zur kognitiven Leistungsfähigkeit gab, missverständlich und zu undifferenziert war. Zur Beurteilung der Kognition wurde in dieser Studie eine Subskala aus einigen wenigen Items einer an sich für die Beurteilung von Alltagsaktivitäten verwendeten Skala (BGP) verwendet. Wir halten diese Subskala für nicht vergleichbar mit den sehr gut validierten und auf psychometrischer Testung beruhenden Skalen ADAS-cog und SIB, welche in Zulassungsstudien akzeptiert werden, und haben die Subskala daher nicht als einen adäquaten bzw. etablierten Endpunkt für die Kognition angesehen.

Wir haben diesen Punkt auch im Fachgespräch mit dem IQWiG erörtert und vorgeschlagen, den Endpunkt Kognition sowohl mit dieser Studie als auch in einer Sensitivitätsanalyse ohne diese Studie zu untersuchen.

Die ergänzenden, erweiterten Triple-Responderanalysen bestätigen die Überlegenheit von Memantin über Plazebo, in vergleichbarer Größenordnung wie seinerzeit die für die EMA-Zulassung geforderte Triple-Responderanalyse. Die Überlegenheit von Memantin wird dabei sowohl in der zulassungskonformen Triple-Responderanalyse wie auch in der modifizierten Triple-Responderanalyse (neue Relevanzgrenzen gemäß IQWiG-Vorschlag / Abstimmung im Fachgespräch) als auch in den drei separaten Mono-Responderanalysen zu den Zielgrößen alltagspraktische Fähigkeiten (Relevanzgrenzen neu gemäß IQWiG-Vorschlag), kognitive Leistungsfähigkeit und klinisches Globalurteil nachgewiesen.

Auch die zusätzlichen vom IQWiG geforderten, weiteren Sensitivitätsanalysen, alle Studienabbrecher (Drop outs) als "deutlich klinisch verschlechtert" zu werten, sind unserer Stellungnahme beigefügt und bestätigen die positiven Ergebnisse für Memantin.

Wir kopieren Herrn PD Dr. Lange, Herrn Dr. Kaiser und Herrn Dr. Skipka vom IQWiG in dieses Schreiben ein, da wir annehmen, dass Sie das IQWiG mit einer Bewertung unserer ergänzenden Analysen beauftragen werden.

Sollten sich bezüglich der dargestellten Analysen noch Rückfragen, auch von Seiten des IQWiG ergeben, so stehen wir zur Klärung oder eventueller Nachlieferung von Analysen jederzeit gerne zur Verfügung.

Wir sind überzeugt, dass die vorgelegten Responderanalysen den Nutzen von Memantin bestätigen, wie dies auch in der Zulassung und anderen evidenzbasierten Bewertungen wie den S3-Leitlinien der Fachgesellschaften zum Ausdruck kommt.

Berücksichtigt der G-BA zusätzlich das sehr günstige Verträglichkeitsprofil von Memantin, welches auch vom IQWIG im Arbeitspapier erneut bestätigt wurde, sind wir sehr zuversichtlich, dass Memantin sowohl für die moderate als auch für schwere Alzheimer Demenz als zweckmäßige und klinisch sinnvolle Therapie erhalten bleibt.

Mit freundlichen Grüßen

Merz Pharmaceuticals GmbH

Dr. Martin Zügel

CEO Merz Pharmaceuticals GmbH

Dr. Claus Göbel

Head Global Clinical Development Operations

#### **Anlage**

Ergänzende Responderanalysen von Merz Pharmaceuticals

# Ergänzende im IQWiG Arbeitspapier Memantin geforderte Responderanalysen

#### 1 Einleitung

Das IQWiG bewertet die bisher von Merz vorgelegten Responderanalysen aus verschiedenen Gründen als nicht ausreichend für eine Nutzenbewertung ("selektive" Studienauswahl von 6 der 9 relevanten Studien unter Einbeziehung nicht zulassungskonform behandelter Patienten, zusammenfassende Analyse ohne Darstellung der Ergebnisse der Einzelstudien, unklares Responsekriterium für den Endpunkt alltagspraktische Fähigkeiten). Das IQWiG (vgl. Arbeitspapier Seite -v-) hält es aber für möglich, "dass sich bei adäquat durchgeführten und berichteten Analysen ein Nutzen für Memantin im Bereich der kognitiven Leistungsfähigkeit zeigt." Darüber hinaus fordert das IQWiG "zusätzliche Sensitivitätsanalysen (...), bei denen Patienten, für die keine Messwerte bei Studienende vorliegen, als Non-Responder (im Sinne einer relevanten klinischen Verschlechterung) gewertet werden."

Merz legt die geforderten Analysen unter Einbeziehung der Studien IE-2101 und MEM-MD-22 vor; Methodik und Ergebnisse sind nachfolgend dargestellt. Die Methodik wurde in einem wissenschaftlichen Gespräch mit dem IQWiG am 01.10.2010 erörtert, dabei wurde unter anderem die Herleitung der Relevanzgrenzen der psychometrischen Skalen diskutiert.

#### 2 Methodik der ergänzenden Responderanalysen

# 2.1 Auswahl der Studien und Patientenkollektive für die Responderanalyse

Für die ergänzenden Responderanalysen wurden die neun auch vom IQWiG im Arbeitspapier berücksichtigten Studien herangezogen, siehe Tabelle 1.

Zusätzlich zu den sechs in der europäischen Zulassung enthaltenen Studien wurde die bereits im IQWiG-Abschlussbericht Memantin berücksichtigte Studie LU-10116 sowie die im IQWiG Arbeitspapier behandelten Studien IE-2101 und MEM-MD-22 in die Analyse einbezogen.

Bei allen Studien wurden die "zulassungskonform" behandelten Patienten berücksichtigt. Insbesondere wurden bei Studie MEM-MD-02 nur die Patienten mit moderater AD eingeschlossen, da die Patienten mit schwerer AD Donepezil außerhalb der EU-Zulassungsindikation erhielten. Bei der Studie IE-2101 wurden in der Memantin-Gruppe wie

vom IQWiG nur die Patienten mit 20 mg Memantin entsprechend der Zulassung berücksichtigt.

Tabelle 1: Übersicht der Studien und Patientenkollektive für die Responderanalysen

Studien- nummer	Stadium AD in Studie	Stadium AD in Analyse	MMSE- Bereich in Ana- lyse	Endpunkt COG	Endpunkt ADL	Endpunkt GLOB	Triple- Respon- der-Ana- lyse	Kommen -tar
IE-2101	moderat schwer	moderat schwer	5-14	SIB	ADL-J (=ADL19)	CIBIC plus	Ja	
LU-10116	moderat schwer	moderat schwer	5-18	SIB	ADLsev (=ADL19)	nein	nein	
LU-99679	leicht moderat	moderat	11-19	ADAScog	ADL23	CIBIC plus	Ja	EMA- Analyse
MEM-MD-01	moderat schwer	moderat schwer	5-14	SIB	ADLsev (=ADL19)	CIBIC plus	Ja	EMA- Analyse
MEM-MD-02	moderat schwer	moderat	10-14	SIB	ADLsev (=ADL19)	CIBIC plus	Ja	EMA- Analyse
MEM-MD-10	leicht moderat	moderat	10-19	ADAScog	ADL23	CIBIC plus	Ja	EMA- Analyse
MEM-MD-12	leicht moderat	moderat	10-19	ADAScog	ADL23	CIBIC plus	Ja	EMA- Analyse
MRZ-9605	moderat schwer	moderat schwer	3-14	SIB	ADLsev (=ADL19)	CIBIC plus	Ja	EMA- Analyse
MEM-MD-22	moderat schwer	moderat schwer	5-18	BGPc	BGPdep	CIBIC plus	Ja	Limitierte Validität BGPc

AD = Alzheimer Demenz; MMSE = Mini Mental State Examination; COG = kognitive Leistungsfähigkeit; ADL = alltagspraktische Fähigkeiten (activities of daily living); GLOB = klinisches Globalurteil; ADL-J = japanische Version der ADL-Skala; ADLsev = ADL-Skala mit 19 Items für moderate bis schwere Alzheimer Demenz; ADL23 = ADL-Skala mit 23 Items für leichte bis mittelschwere Alzheimer Demenz; BGPc = Behavioral Rating Scale for Geriatric Patients – cognitive subscale; BGPdep = Rating Scale for Geriatric Patients – dependency subscale; ADAScog = Alzheimer's Disease Assessment Scale-cognitive subscale; SIB = Severe Impairment Battery

Die Studie LU-10116 enthielt keinen Endpunkt zum klinischen Globalurteil. Die Mono-Responderanalyse Globalurteil und die Triple-Responderanalysen wurden deshalb ohne diese Studie durchgeführt.

Die Studie MEM-MD-22 unterscheidet sich in wesentlichen Punkten von den anderen Studien (explorative nicht konfirmatorische Studie, einzige Studie in Pflegeheimen, deutlich höheres Durchschnittsalter) und enthält aus Sicht von Merz keinen adäquaten Endpunkt für die kognitive Leistungsfähigkeit. Das IQWiG hat im Arbeitspapier als kognitiven Endpunkt für diese Studie die Behavioral Rating Scale for Geriatric Patients (BGP) herangezogen. Von den 35 Items wurden in der Studie fünf für kognitive Verhaltensaspekte ausgewertet. Im Gegensatz zu ADAS-cog und SIB, den beiden etablierten Skalen zur psychometrischen Beurteilung der kognitiven Leistungsfähigkeit, handelt es sich bei der BGP nur um sehr

allgemeine Fragen zur örtlichen und zeitlichen Orientierung. Die Beurteilung erfolgt nicht durch psychometrische Testung, sondern aufgrund der Beobachtung des Patienten <u>durch das Pflegepersonal</u>. Zwar hatte die Firma Forest diese kognitive Subskala der BGP - wie vom IQWiG im Arbeitspapier zitiert - im FDA-Zulassungsdossier eingereicht. Die FDA zweifelte allerdings im Assessment die Wertigkeit dieser Subskala für den Endpunkt Kognition an [1, Seite 84]. Aus diesen Gründen halten wir eine Sensitivitätsanalyse für den Endpunkt Kognition und die Triple-Responderanalysen ohne Einbeziehung der Studie MEM-MD-22 für notwendig.

Sämtliche Analysen wurden im Intention to Treat (ITT) Kollektiv entsprechend IQWiG Abschlussbericht und Arbeitspapier durchgeführt.

#### 2.2 Methodik der Responderanalysen

Die Responderanalysen wurden als Metaanalysen mittels eines Modells mit zufälligen Effekten (nach DerSimonian und Laird) mit Angabe der Einzelergebnisse pro Studie gerechnet. Dargestellt werden die Odds Ratio (Quoten- oder Chancenverhältnis zugunsten Memantin oder Plazebo) mit 95% Konfidenzintervallen (graphisch und numerisch), die Gewichtung, die Heterogenität und der Gesamteffekt.

Die Analysen werden entsprechend IQWiG Abschlussbericht und Arbeitspapier dargestellt, d.h. mit einer Subgruppenaufteilung in Studien mit Monotherapie Memantin und Studien mit einer zusätzlichen Therapie mit AChEI. Aufgrund der vom G-BA vorgenommenen unterschiedlichen Einstufung von Memantin bezüglich der moderaten und schweren Alzheimer Demenz stellen wir die Responderanalysen zusätzlich stratifiziert nach Krankheitsstadium dar, also separat für die moderate und schwere Alzheimer Demenz. Für die moderate Alzheimer Demenz wurden dabei die Patienten mit einem MMSE von 10 bis 19 ausgewertet, für die schwere Alzheimer Demenz die Patienten mit einem MMSE < 10.

Die Ergebnisse der Metaanalysen der jeweiligen Gesamtgruppe (Mono-/Zusatztherapie bzw. moderat/schwer) variieren geringfügig bezüglich der Odds Ratio und Konfidenzintervalle aufgrund der unterschiedlichen Subgruppeneinteilungen.

Es wird darauf hingewiesen, dass die pivotalen Studien darauf ausgelegt waren, Patienten mit moderater bis schwerer Alzheimer Demenz zu untersuchen. Bezüglich der Zahl der teilnehmenden Patienten (Power) war diese nicht für die Subgruppenanalysen nach Krankheitsstadium ausgelegt und insgesamt sind die Patientenzahlen für das Stadium der schweren Alzheimer Demenz deutlich geringer. Dies limitiert die Interpretation der Subgruppenanalysen.

Die Analysen wurden post hoc durchgeführt, es erfolgte keine Adjustierung für multiples Testen.

### 2.2.1 Responderdefinition in Analogie zur EMA Zulassung 2005

Die Responderdefinition erfolgte in Analogie zur EMA Zulassung 2005. Dabei wird entsprechend dem klinischen Verlauf der Alzheimer Erkrankung mit einer zunehmenden Progredienz das Vermeiden einer relevanten deutlichen Verschlechterung als Behandlungserfolg für Memantin gewertet.

Entsprechend werden in den Responderanalysen die Anteile der Patienten mit einer deutlichen klinischen Verschlechterung zwischen Memantin- und Plazebogruppe verglichen.

#### 2.2.2 Wahl der Grenzwerte für die Relevanzschwellen

Die Grenzwerte für eine relevante, deutliche klinische Verschlechterung wurden für die verwendeten Skalen hergeleitet aus der Literatur bzw. aus den Hinweisen des IQWiG im Arbeitspapier Memantin, siehe Tabelle 2. Solche Grenzwerte werden vom IQWiG als Ausdruck einer "Minimal Important Difference" gefordert und wurden bei dem wissenschaftlichen Gespräch zwischen Merz / Lundbeck und IQWiG am 01.10.2010 ausführlich diskutiert. Für die Skalen, für die keine validierten bzw. etablierten Grenzwerte vorlagen, wurde dem Ansatz des IQWiG gefolgt, die Relevanzschwelle aus der halben gepoolten Standardabweichung herzuleiten. Wurde die Skala in mehreren Studien verwendet mit entsprechend unterschiedlichen Standardabweichungen, so wurde für die betreffende Skala der Median aus den einzelnen Studien für die Wahl der Relevanzschwelle herangezogen. Für die Endpunkte kognitive Leistungsfähigkeit und alltagspraktische Fähigkeiten wurden unterschiedliche Testinstrumente verwendet, entsprechend ergeben sich hier jeweils verschiedene Relevanzschwellen.

Zusätzlich zu den vom IQWiG geforderten, teilweise neu definierten Grenzwerten werden auch die in der zulassungsrelevanten Triple-Responderanalyse verwendeten Relevanzschwellen für eine separate Analyse verwendet.

Für den G-BA dürfte diese Analyse wegen der Bindungswirkung der Zulassungsentscheidung von besonderer Bedeutung sein, da sie den von der Zulassungsbehörde im Zulassungsverfahren gemachten methodischen Annahmen entspricht. Die Abweichung betrifft ausschließlich den Endpunkt alltagspraktische Fähigkeiten, für den in der zulassungsrelevanten Triple-Responderanalyse ein Grenzwert von ≥ 1 Punkt Verschlechterung verwendet wurde.

Tabelle 2: Grenzwerte für die gewählten Relevanzschwellen (Mono-Responderanalyse)

Skala	Grenzwert für	Begründung		
	Responderanalyse	99		
Instrumente	Instrumente zur Beurteilung der kognitiven Leistungsfähigkeit			
ADAScog	Verschlechterung	Die angegebenen Werte gelten als Ausdruck einer		
_	um ≥ 4 Punkte	Verschlechterung ohne Behandlung über einen Zeitraum von		
SIB	Verschlechterung	6 Monaten und werden von Zulassungsbehörden als relevant		
	um ≥ 5 Punkte	anerkannt.		
BGPc	Verschlechterung	Die BGPc besteht aus 5 Fragen mit einem Antwort-Score		
	um ≥ 1 Punkt	von 0 bis 2. Die Spanne der Skala beträgt 0 bis 10. Die		
		gepoolte Standardabweichung der Mittelwertsdifferenz für		
		Memantin und Plazebogruppe in Studie MEM-MD-22 betrug		
		1,91. Aus der Literatur gibt es keine Hinweise für Relevanz-		
		schwellen, aus den Hinweisen des IQWiG Arbeitspapiers		
		ergibt sich der hier vorgeschlagene Grenzwert von ≥ 1.		
		alltagspraktischen Fähigkeiten		
ADL-19	Verschlechterung	Die ADL-19 (=ADLsev, ADL-J) wurde in 5 Studien bei		
	um ≥ 3 Punkte	moderater bis schwerer AD verwendet, die Spanne der		
		Skala beträgt 0 – 54. Aus der Literatur gibt es keine		
		Hinweise für Relevanzschwellen, aus den Hinweisen des IQWiG Arbeitspapiers ergibt sich der hier vorgeschlagene		
		Grenzwert von ≥ 3 (entsprechend dem Median der halben		
		gepoolten Standardabweichung der 5 Studien von 3,12).		
ADL-23	Verschlechterung	Die ADL-23 wurde in 3 Studien bei leichter bis moderater AD		
7.52.20	um ≥ 5 Punkte	verwendet, die Spanne der Skala beträgt 0 – 78. Aus der		
		Literatur gibt es keine Hinweise für Relevanzschwellen, aus		
		den Hinweisen des IQWiG Arbeitspapiers ergibt sich der hier		
		vorgeschlagene Grenzwert von ≥ 5 (entsprechend dem		
		Median der halben gepoolten Standardabweichung der 3		
		Studien von 4,64).		
BGPdep	Verschlechterung	Die BGPdep wurde in Studie MEM-MD-22 verwendet, die		
	um ≥ 3 Punkte	Spanne der Skala beträgt 0 – 46. Aus der Literatur gibt es		
		keine Hinweise für Relevanzschwellen, aus den Hinweisen		
		des IQWiG Arbeitspapiers ergibt sich der hier		
		vorgeschlagene Grenzwert von ≥ 3 (entsprechend der		
		halben gepoolten Standardabweichung der		
ADL-19,	Verschechterung	Mittelwertsdifferenz von 2,64).		
ADL-19, ADL-23,	um ≥ 1 Punkt	In der EMA-Zulassung wurde in der kombinierten Triple- Responderanalyse der Grenzwert für den Endpunkt		
BGPdep	uni z i r uliki	alltagspraktische Fähigkeiten mit ≥ 1 festgelegt. Damit kann		
		dieser Grenzwert zumindest von Seiten der Zulassungs-		
		behörden als etabliert angesehen werden.		
		Dieser Grenzwert wurde ausschließlich für die Triple-		
		Responderanalysen gemäß Zulassungsverfahren		
		verwendet.		
Instrumente zur Beurteilung des klinischen Globalurteils				
CIBICplus	Verschechterung	Bei der CIBICplus ergibt sich die Grenze von ≥ 1 Punkt per		
	um ≥ 1 Stufe	Definition, ein klinisch relevantes Globalurteil zu fällen, vgl.		
		IQWIG-Abschlussbericht, Seite 104. Auch in der Literatur ist		
		dieser Grenzwert beschrieben.		

# 2.2.3 Zusätzliche Sensitivitätsanalysen

Wegen der unter 2.1. beschriebenen limitierten Validität der BGPc wurden für die Mono-Responderanalyse Kognition und die Triple-Responderanalysen eine Sensitivitätsanalyse durchgeführt, um eine mögliche Verzerrung des Gesamtergebnisses durch diese Studie zu beurteilen.

Das IQWiG hatte im Arbeitspapier Memantin auch Sensitivitätsanalysen Studienabbrecher gefordert, bei denen keine Messwerte bei Studienende vorlagen. Diese sollten als Non-Responder (im Sinne einer relevanten Verschlechterung) gewertet werden. Bei den entsprechenden Analysen wurden daher alle Patienten sowohl der Memantin- als auch der Plazebogruppen, die die Studie vorzeitig abgebrochen hatten, so behandelt, als hätten sie sich entsprechend um die in Tabelle 2 gelisteten Grenzwerte verschlechtert.

### 2.2.4 Erläuterung zu Diskrepanzen bei den Patientenzahlen

Bei Studie MEM-MD-22 liegen die Patientenzahlen der Analysen in der Memantin- und Plazebogruppe um jeweils einen Patienten niedriger als im IQWIG Arbeitspapier. Dies ist auf eine missverständliche Darstellung im Studienbericht zurückzuführen. In einer Tabelle im Studienbericht (11.4.1.3.3-1., Seite 63) wurde versäumt, darauf hinzuweisen, dass für den BGPc für je einen Patienten aus beiden Gruppen keine Beobachtungen vorlagen (vgl. Tabelle 14.4.4.3.2.A, Seite 353 im Studienbericht [liegt dem IQWiG und G-BA vorl).

Im IQWiG-Abschlussbericht wurden zum Teil kleine Diskrepanzen bei den Patientenzahlen zwischen den tabellarischen Aufstellungen der Studien und den Metaanalysen festgestellt. Dadurch können geringfügige Abweichungen bei den Patientenzahlen zwischen den IQWiG-Berichten und den hier vorgestellten Analysen erklärt werden.

#### 3 Ergebnisse der ergänzenden Responderanalysen

#### 3.1.1 Alltagspraktische Fähigkeiten

Die alltagspraktischen Fähigkeiten wurden in allen 9 Studien untersucht. Für die Mono-Responderanalyse wurden die in Tabelle 2 genannten Grenzwerte für die Definition "deutliche klinische Verschlechterung" verwendet (ADL-19 ≥ 3, ADL-23 ≥ 5, BGPdep ≥ 3). Abbildung 1 zeigt das Ergebnis zu den alltagspraktischen Fähigkeiten mit einer Subgruppenaufteilung der Studien mit Monotherapie Memantin und Studien mit einer zusätzlichen Therapie mit AChEI. Abbildung 2 zeigt das entsprechende Ergebnis mit einer Subgruppenaufteilung der Patientenpopulation nach Schweregrad der Alzheimer Erkrankung.

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv alltagspraktische Fähigkeiten = Verschlechterung ≥ MID Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

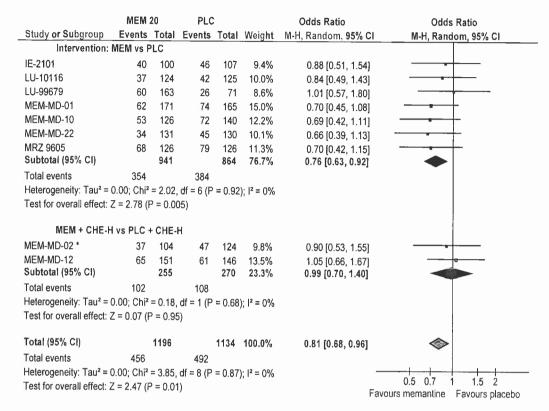


Abbildung 1: Mono-Responderanalyse alltagspraktische Fähigkeiten (Mono-/Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv alltagspraktische Fähigkeiten = Verschlechterung  $\geq$  MID (ADL-19  $\geq$  3, ADL-23  $\geq$  5, BGPdep  $\geq$  3) Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

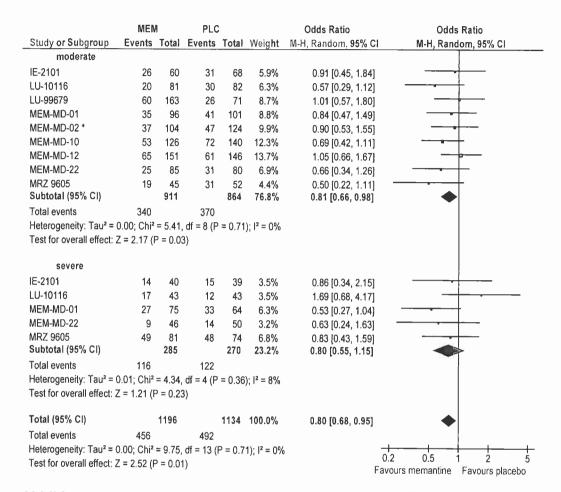


Abbildung 2: Mono-Responderanalyse alltagspraktische Fähigkeiten (moderat - schwer)

#### 3.1.2 Kognitive Leistungsfähigkeit

Die kognitive Leistungsfähigkeit wurde in allen 9 Studien untersucht. Für die Mono-Responderanalyse wurden die in Tabelle 2 genannten Grenzwerte für die Definition "deutliche klinische Verschlechterung" verwendet. Abbildung 3 zeigt das Ergebnis zur kognitiven Leistungsfähigkeit mit einer Subgruppenaufteilung der Studien mit Monotherapie Memantin und Studien mit einer zusätzlichen Therapie mit AChEI. Abbildung 4 zeigt das entsprechende Ergebnis mit einer Subgruppenaufteilung der Patientenpopulation nach Schweregrad der Alzheimer Erkrankung.

Abbildung 5 und 6 zeigen die entsprechenden Sensitivitätsanalysen ohne Studie MEM-MD-22, da in dieser Studie mit der BGPc eine kognitive Skala mit limitierter Validität verwendet wurde.

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit = Verschlechterung ≥ MID Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

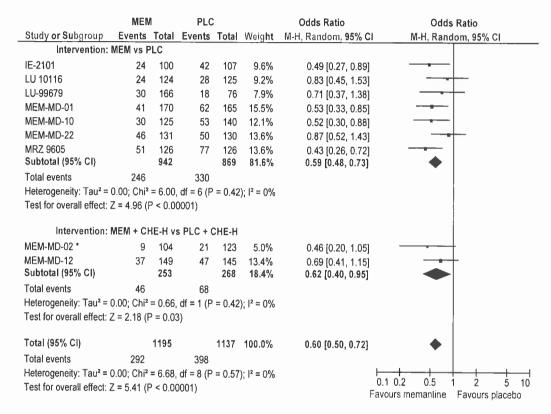


Abbildung 3: Mono-Responderanalyse kognitive Leistungsfähigkeit (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit = Verschlechterung ≥ MID Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

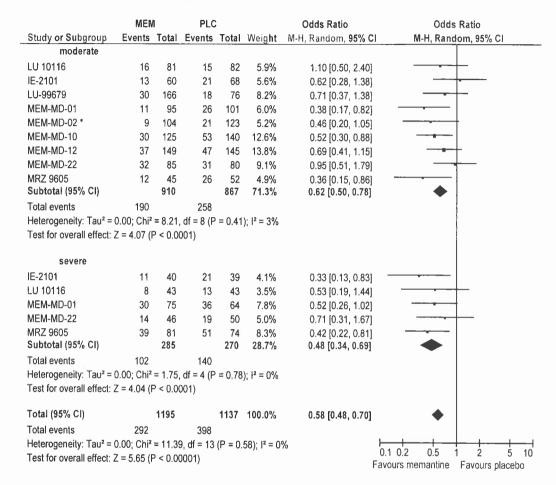


Abbildung 4: Mono-Responderanalyse kognitive Leistungsfähigkeit (moderat - schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit = Verschlechterung ≥ MID Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

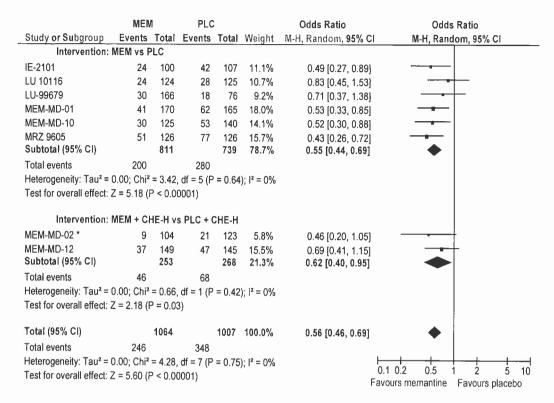


Abbildung 5: Mono-Responderanalyse kognitive Leistungsfähigkeit (Mono- / Zusatztherapie) – Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit = Verschlechterung ≥ MID Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

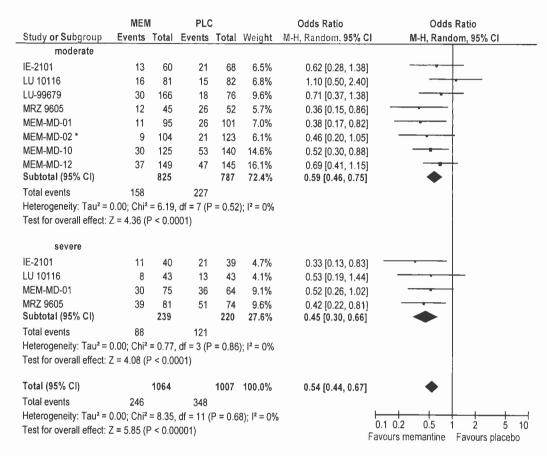


Abbildung 6: Mono-Responderanalyse kognitive Leistungsfähigkeit (moderat - schwer) – Sensitivitätsanalyse ohne Studie MEM-MD-22

### 3.1.3 Klinisches Globalurteil

Das klinische Globalurteil wurde in 8 der 9 Studien untersucht, Studie LU-10116 enthielt keinen Endpunkt für diese Zielgröße. Für die Mono-Responderanalyse wurde der in Tabelle 2 genannte Grenzwert für die Definition "deutliche klinische Verschlechterung" verwendet. Abbildung 7 zeigt das Ergebnis zum klinischen Globalurteil mit einer Subgruppenaufteilung der Studien mit Monotherapie Memantin und Studien mit einer zusätzlichen Therapie mit AChEI. Abbildung 8 zeigt das entsprechende Ergebnis mit einer Subgruppenaufteilung der Patientenpopulation nach Schweregrad der Alzheimer Erkrankung.

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Klinisches Globalurteil = Verschlechterung ≥ MID Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEl Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

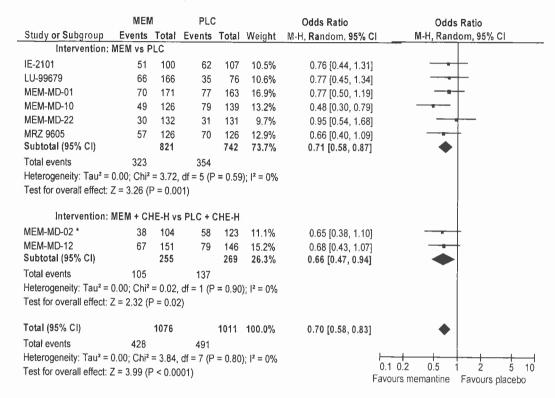


Abbildung 7: Mono-Responderanalyse klinisches Globalurteil (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Klinisches Globalurteil = Verschlechterung ≥ MID Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

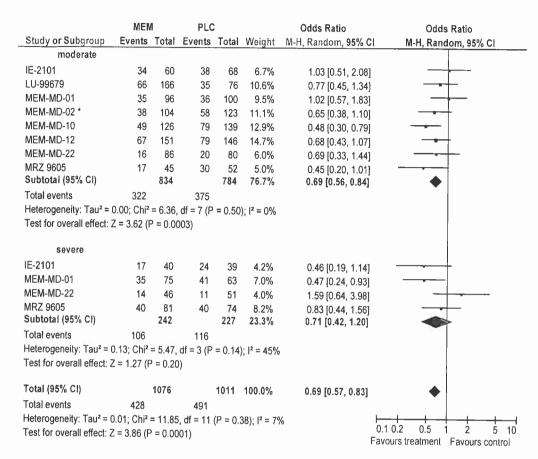


Abbildung 8: Mono-Responderanalyse klinisches Globalurteil (moderat - schwer)

# 3.1.4 Triple-Responderanalysen

Für die Triple-Responderanalysen mit den kombinierten Endpunkten Kognition, alltagspraktische Fähigkeiten und Globalurteil konnten 8 der 9 Studien untersucht werden, Studie LU-10116 enthielt keinen Endpunkt für die Zielgröße Globalurteil. Die Triple-Responderanalysen wurde

 einmal mit den Grenzwerten aus dem Zulassungsverfahren (Kognition: Verschlechterung ≥ 4 Punkte für ADAScog, ≥ 5 Punkte für SIB; alltagspraktische Fähigkeiten: jegliche Verschlechterung in allen verwendeten ADL-Skalen; klinisches Globalurteil: Verschlechterung um mindestens eine Stufe),  zum anderen mit den in Tabelle 2 neu definierten Grenzwerten entsprechend Anforderung aus dem IQWiG-Arbeitspapier (Änderung bei ADL: ADL-19 ≥ 3, ADL-23 ≥ 5, BGPdep ≥ 3)

#### durchgeführt.

Der Grenzwert für die BGPc (kognitive Leistungsfähigkeit) in Studie MEM-MD-22 wurde für beide Triple-Responderanalysen gemäß der vom IQWiG vorgeschlagenen Methode bestimmt (mindestens ein Punkt Verschlechterung). Abbildung 9 und 10 zeigen die Ergebnisse der Triple-Responderanalysen gemäß Zulassungsverfahren, jeweils mit Subgruppenaufteilung der Studien nach Monotherapie Memantin und zusätzlicher Therapie mit AChEI und Subgruppenaufteilung nach Schweregrad der Alzheimer Erkrankung.

Abbildung 11 und 12 zeigen die entsprechenden Sensitivitätsanalysen ohne Studie MEM-MD-22, da in dieser Studie mit der BGPc eine kognitive Skala mit limitierter Validität verwendet wurde.

Abbildung 13 bis 16 zeigen die entsprechenden Triple-Responderanalysen mit den neu definierten Grenzwerten entsprechend dem Vorschlag aus dem IQWiG-Arbeitspapier.

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv

Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID

Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI

Modell mit zufälligen Effekten - DerSimonian und Laird

(\*nur Patienten mit MMSE 10-14)

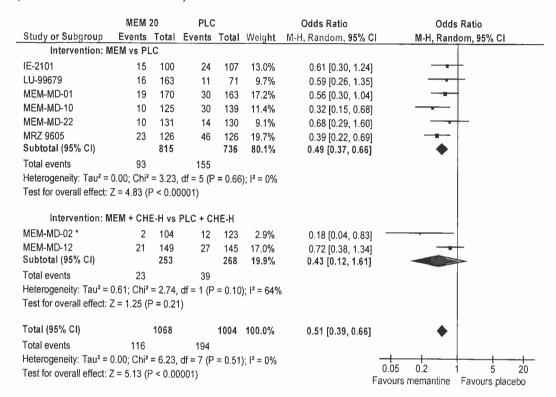


Abbildung 9: Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

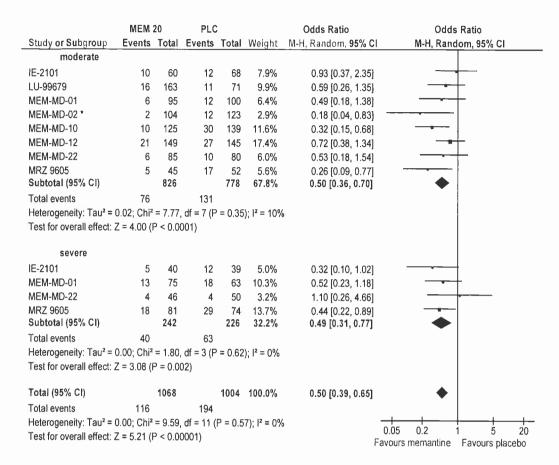


Abbildung 10: Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (moderat - schwer)

Memantin vs. Placebo -- Intention to Treat (ITT) Kollektiv

Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID

Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI

Modell mit zufälligen Effekten - DerSimonian und Laird

(\*nur Patienten mit MMSE 10-14)

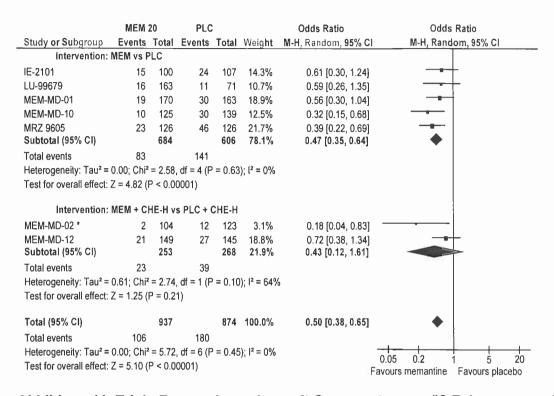


Abbildung 11: Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (Mono- / Zusatztherapie) - Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

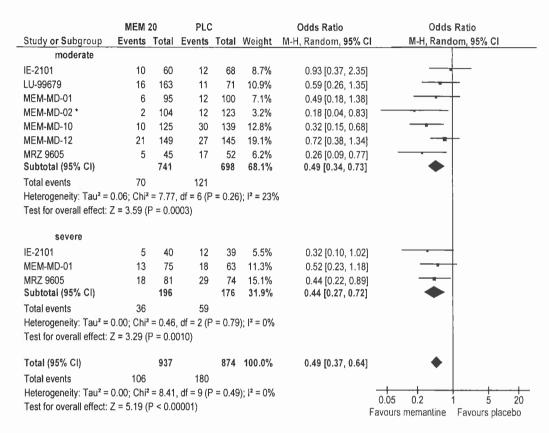


Abbildung 12: Triple-Responderanalyse mit Grenzwerten gemäß Zulassungssverfahren (moderat - schwer) - Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

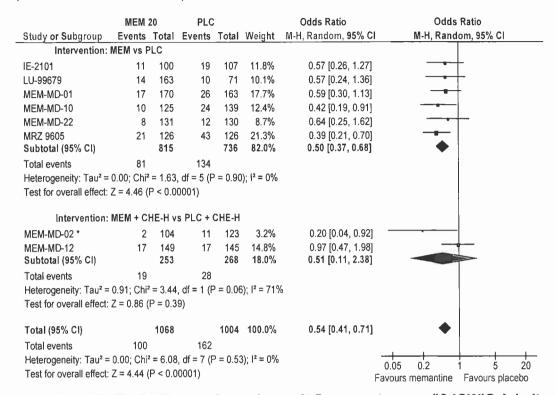


Abbildung 13: Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

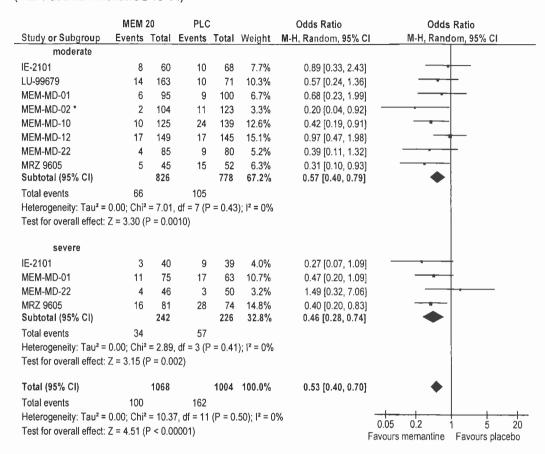


Abbildung 14: Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (moderat - schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEl
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

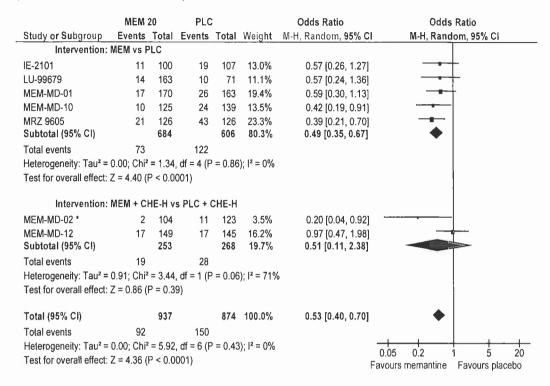


Abbildung 15: Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (Mono- / Zusatztherapie) - Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

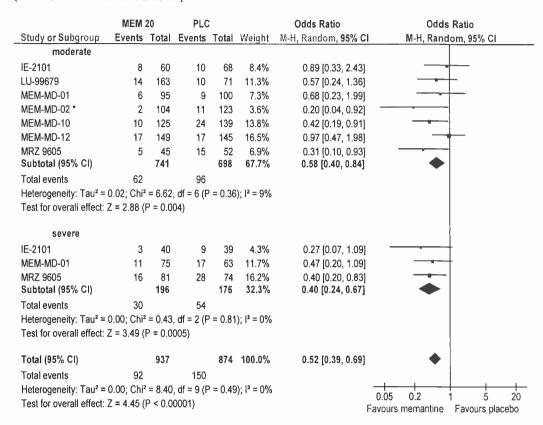


Abbildung 16: Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (moderat - schwer) - Sensitivitätsanalyse ohne Studie MEM-MD-22

## 3.1.5 Sensitivitätsanalysen für Studienabbrecher

Im Folgenden werden die im Arbeitspapier Memantin vom IQWiG geforderten Sensitivitätsanalysen für Studienabbrecher dargestellt. Dabei wurden alle Studienabbrecher in den Memantin- und Plazebogruppen als Non-Responder im Sinne einer deutlichen klinischen Verschlechterung bewertet. Abbildung 17 – 24 zeigen die Ergebnisse für die Mono-Responderanalysen in den drei Domänen alltagspraktische Fähigkeiten, kognitive Leistungsfähigkeit und klinisches Globalurteil, jeweils stratifiziert nach Monotherapie / Zusatztherapie mit AChEI bzw. nach Schweregrad moderate / schwere AD. Für den Endpunkt kognitive Leistungsfähigkeit werden zusätzlich die Sensitivitätsanalysen ohne Studie MEM-MD-22 wegen der limitierten Validität der in dieser Studie verwendeten BGPc dargestellt.

Abbildungen 25 – 28 zeigen die entsprechenden Triple-Responderanalysen gemäß Zulassungskriterien, Abbildungen 29 bis 32 die Triple-Responderanalysen gemäß den im Arbeitspapier Memantin vom IQWiG geforderten Grenzwerten.

Um den zeitlichen Verlauf der Studienabbrüche besser beurteilen zu können, haben wir zusätzlich aus unserer Arzneimittelsicherheitsdatenbank für die 9 hier analysierten Studien Kaplan-Meier-Kurven für die Studienabbrecher allgemein (alle Ursachen) bzw. Studienabbrecher aufgrund von unerwünschten Ereignissen erstellt, siehe Abbildungen 33 und 34.

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv alltagspraktische Fähigkeiten Completer = Verschlechterung ≥ MID alltagspraktische Fähigkeiten Studienabbrecher = immer Verschlechterung Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

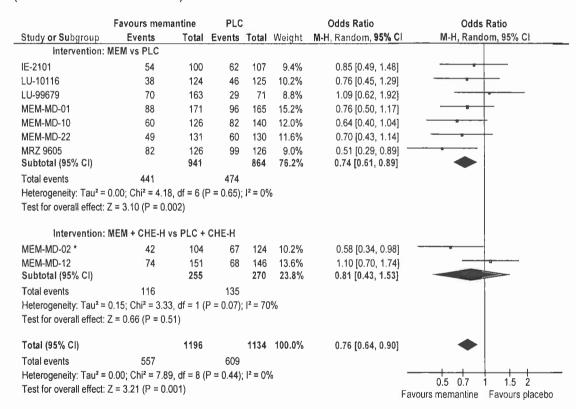


Abbildung 17: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse alltagspraktische Fähigkeiten (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv alltagspraktische Fähigkeiten Completer = Verschlechterung ≥ MID alltagspraktische Fähigkeiten Studienabbrecher = immer Verschlechterung Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

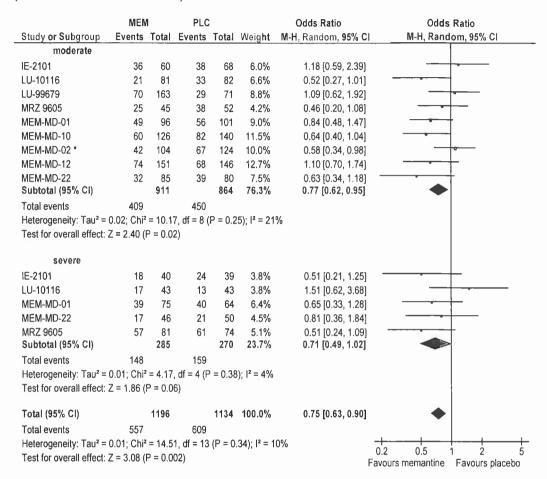


Abbildung 18: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse alltagspraktische Fähigkeiten (moderat - schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Kognitive Leistungsfähigkeit Completer = Verschlechterung ≥ MID
Kognitive Leistungsfähigkeit Studienabbrecher = immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

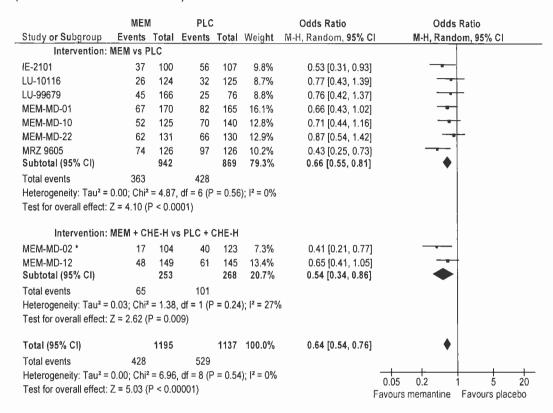


Abbildung 19: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse kognitive Leistungsfähigkeit (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit Completer = Verschlechterung ≥ MID Kognitive Leistungsfähigkeit Studienabbrecher = immer Verschlechterung Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

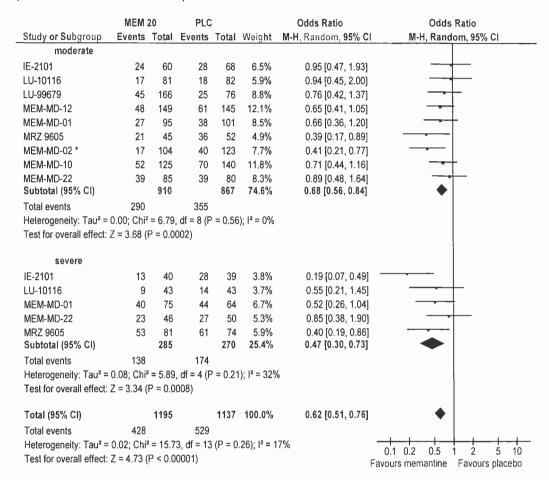


Abbildung 20: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse kognitive Leistungsfähigkeit (moderat - schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Kognitive Leistungsfähigkeit Completer = Verschlechterung ≥ MID
Kognitive Leistungsfähigkeit Studienabbrecher = immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

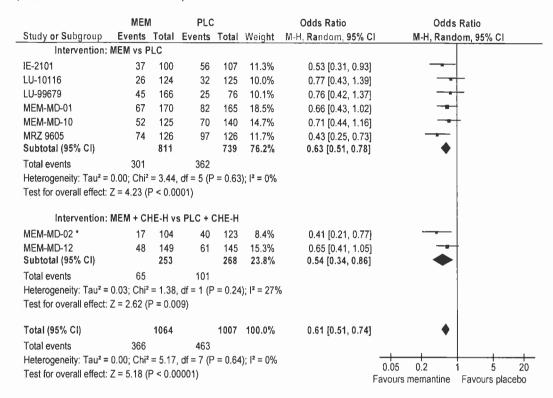


Abbildung 21: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse kognitive Leistungsfähigkeit (Mono- / Zusatztherapie) - Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Kognitive Leistungsfähigkeit Completer = Verschlechterung ≥ MID Kognitive Leistungsfähigkeit Studienabbrecher = immer Verschlechterung Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

	MEN	1	PLC			Odds Ratio	Odds Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% C	M-H, Random, 95% Cl		
moderate									
IE-2101	24	60	28	68	7.6%	0.95 [0.47, 1.93]			
LU-10116	17	81	18	82	6.9%	0.94 [0.45, 2.00]			
LU-99679	45	166	25	76	10.2%	0.76 [0.42, 1.37]			
MEM-MD-01	27	95	38	101	9.9%	0.66 [0.36, 1.20]			
MEM-MD-02 *	17	104	40	123	8.9%	0.41 [0.21, 0.77]			
MEM-MD-10	52	125	70	140	13.5%	0.71 [0.44, 1.16]	<del></del>		
MEM-MD-12	48	149	61	145	13.8%	0.65 [0.41, 1.05]			
MRZ 9605	21	45	36	52	5.8%	0.39 [0.17, 0.89]			
Subtotal (95% CI)		825		787	76.5%	0.66 [0.54, 0.82]	•		
Total events	251		316						
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi <sup>2</sup>	= 5.99	df = 7 (F	0.54	$\frac{1}{2} = 0\%$				
Test for overall effect:	Z = 3.77 (	P = 0.00	002)						
severe									
	40	40	28	20	4.5%	0.40.10.07.0.401			
IE-2101	13	40		39		0.19 [0.07, 0.49] 0.55 [0.21, 1.45]	1		
LU-10116	9 40	43 75	14 44	43	4.4% 7.8%	0.55 [0.21, 1.45]			
MEM-MD-01	53	81	61	64 74	6.8%	0.40 [0.19, 0.86]			
MRZ 9605 Subtotal (95% CI)	53	239	01	220	23.5%	0.40 [0.19, 0.66]	•		
Total events	115	200	147	220	20.070	0.40 [0.20, 0.02]	•		
		! = 2 27		) = 0 25	() 12 = 00/				
Heterogeneity: Tau <sup>2</sup> = Test for overall effect:			,	- 0.30	7), 1 - 070				
restroi overan enect.	2-4.14	1 ~ 0.0	001)						
Total (95% CI)		1064		1007	100.0%	0.59 [0.48, 0.73]	<b>•</b>		
Total events	366		463						
Heterogeneity: Tau <sup>2</sup> =	0.03; Chi	2 = 13.6	4, df = 11	(P = 0)	.25); I <sup>2</sup> = 19	1%	0.1 0.2 0.5 1 2 5 10		
Test for overall effect:	7 = 4 84 (	P < 0.0	0001)				0.1 0.2 0.5 1 2 5 10 Favours memantine Favours placebo		

Abbildung 22: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse kognitive Leistungsfähigkeit (moderat - schwer) - zusätzliche Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Klinisches Globalurteil Completer = Verschlechterung ≥ MID
Klinisches Globalurteil Studienabbrecher = immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

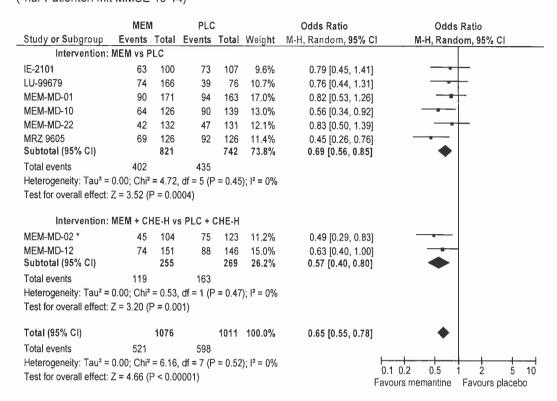


Abbildung 23: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse klinisches Globalurteil (Mono- / Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv Klinisches Globalurteil Completer = Verschlechterung ≥ MID Klinisches Globalurteil Studienabbrecher = immer Verschlechterung Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird (\*nur Patienten mit MMSE 10-14)

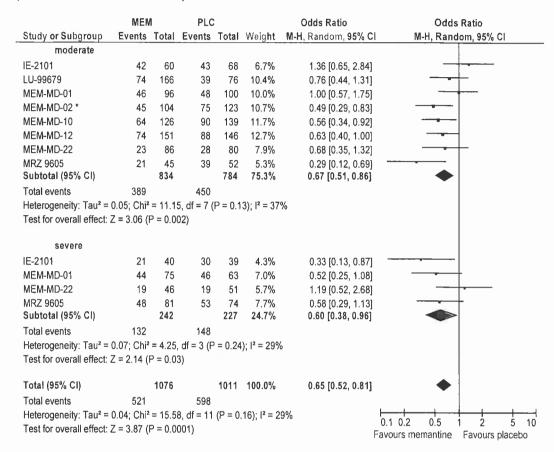


Abbildung 24: Sensitivitätsanalyse Studienabbrecher Mono-Responderanalyse klinisches Globalurteil (moderat - schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

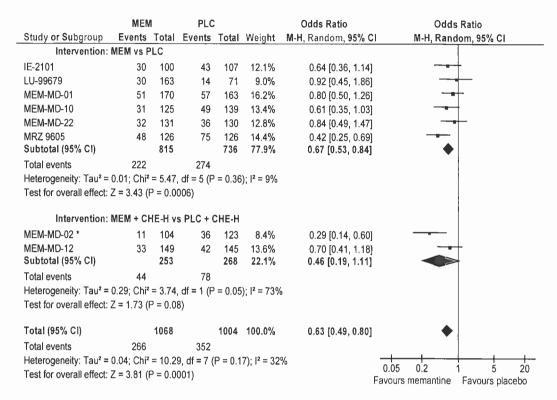


Abbildung 25: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (Mono-/Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

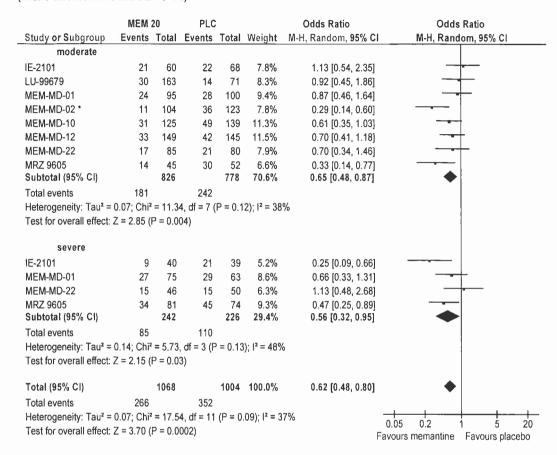


Abbildung 26: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (moderat schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

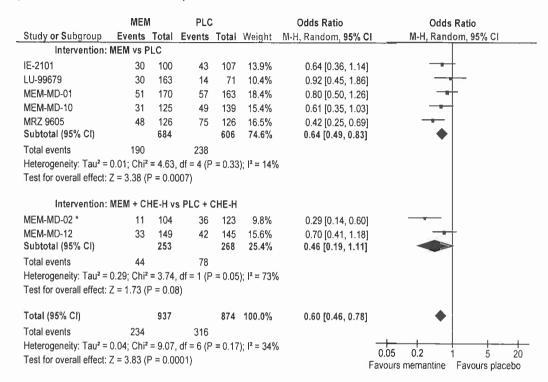


Abbildung 27: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren Zulassungsverfahren (Mono- / Zusatztherapie) – Zusätzliche Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

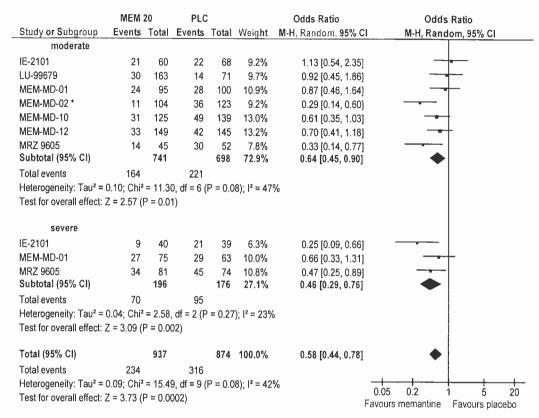


Abbildung 28: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß Zulassungsverfahren (moderat schwer) – zusätzliche Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEl
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

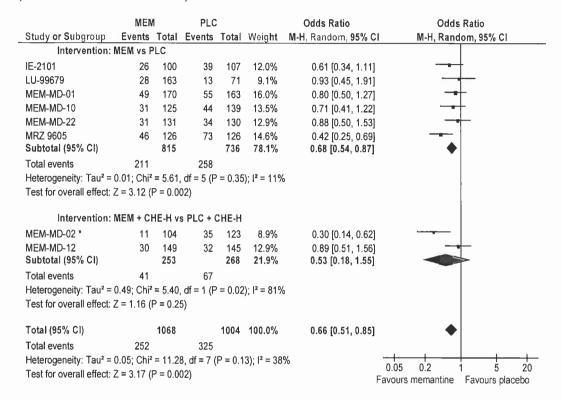


Abbildung 29: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (Mono-/Zusatztherapie)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung moderate AD – schwere AD
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

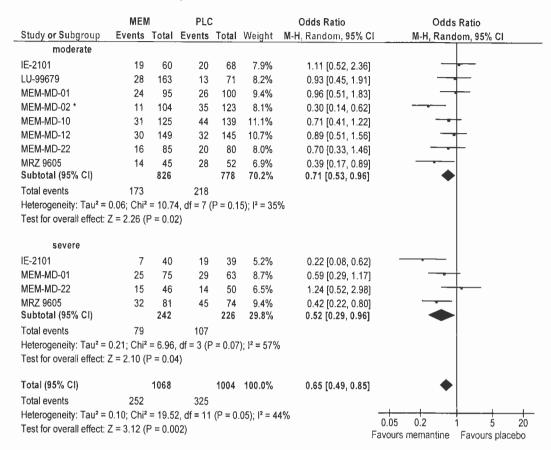


Abbildung 30: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (moderat / schwer)

Memantin vs. Placebo – Intention to Treat (ITT) Kollektiv
Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID
Studienabbrecher: immer Verschlechterung
Subgruppenaufteilung Monotherapie Memantin und zusätzliche Therapie mit AChEI
Modell mit zufälligen Effekten - DerSimonian und Laird
(\*nur Patienten mit MMSE 10-14)

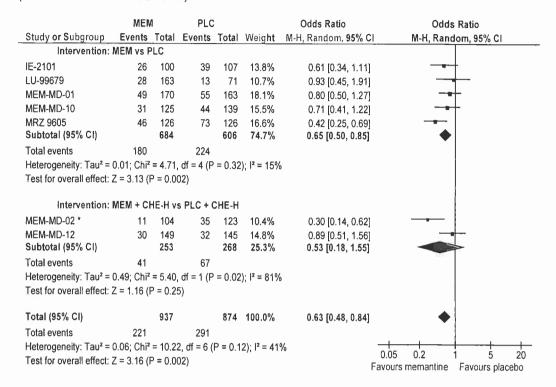


Abbildung 31: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (Mono-/Zusatztherapie) - zusätzliche Sensitivitätsanalyse ohne Studie MEM-MD-22

Memantin vs. Placebo - Intention to Treat (ITT) Kollektiv

Completer: Verschlechterung Kognition UND alltagspraktische Fähigkeiten UND Globalurteil ≥ MID

Studienabbrecher: immer Verschlechterung Subgruppenaufteilung moderate AD – schwere AD Modell mit zufälligen Effekten - DerSimonian und Laird

(\*nur Patienten mit MMSE 10-14)

	MEN	ſ.	PLC		Odds Ratio		Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% C	I M-H, Random, 95% CI
moderate							
IE-2101	19	60	20	68	9.3%	1.11 [0.52, 2.36]	
LU-99679	28	163	13	71	9.6%	0.93 [0.45, 1.91]	
MEM-MD-01	24	95	26	100	10.9%	0.96 [0.51, 1.83]	
MEM-MD-02 *	11	104	35	123	9.5%	0.30 [0.14, 0.62]	<del></del>
MEM-MD-10	31	125	44	139	12.8%	0.71 [0.41, 1.22]	<del></del>
MEM-MD-12	30	149	32	145	12.4%	0.89 [0.51, 1.56]	<del></del>
MRZ 9605	14	45	28	52	8.2%	0.39 [0.17, 0.89]	
Subtotal (95% CI)		741		698	72.7%	0.71 [0.50, 1.00]	•
Total events	157		198				
Heterogeneity: Tau <sup>2</sup> =	0.09; Chi <sup>2</sup>	= 10.73	3, df = 6	P = 0.1	0); $I^2 = 44$	%	
Test for overall effect:	Z = 1.99 (I	P = 0.08	5)		,		
severe							
IE-2101	7	40	19	39	6.2%	0.22 [0.08, 0.62]	<del></del>
MEM-MD-01	25	75	29	63	10.2%	0.59 [0.29, 1.17]	<del></del>
MRZ 9605	32	81	45	74	10.9%	0.42 [0.22, 0.80]	<del></del>
Subtotal (95% CI)		196		176	27.3%	0.42 [0.26, 0.68]	<b>•</b>
Total events	64		93				
Heterogeneity: Tau <sup>2</sup> =	0.03; Chi <sup>2</sup>	= 2.34,	df = 2 (P	= 0.31	);   <sup>2</sup> = 14%		
Test for overall effect:	Z = 3.59 (I	o = 0.00	003)		,		
Total (95% CI)		937		874	100.0%	0.61 [0.45, 0.83]	•
Total events	221		291				•
Heterogeneity: Tau <sup>2</sup> =		= 17 3°		D – n n	4\· 12 - 4Q	0/_	
Test for overall effect:			,	- 0.0	4), 1 - 40	70	0.05 0.2 1 5 20
restrui overali ellect.	L - J. 10 (I	- 0.00	(17				Favours memantine Favours placebo

Abbildung 32: Sensitivitätsanalyse Studienabbrecher Triple-Responderanalyse mit Grenzwerten gemäß IQWiG Arbeitspapier (moderat – schwer) – zusätzliche Sensitivitätsanalyse ohne Studie MEM-MD-22

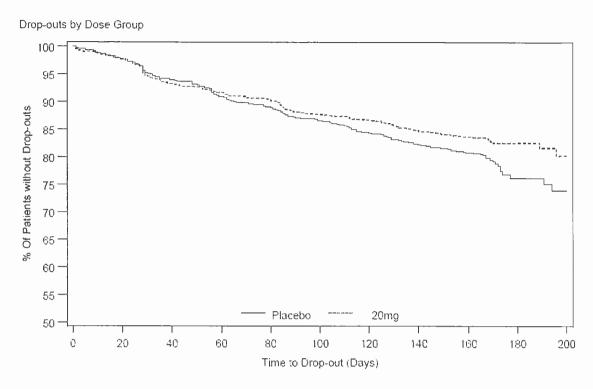


Abbildung 33: Studienabbrecher – alle Ursachen (9 Studien)

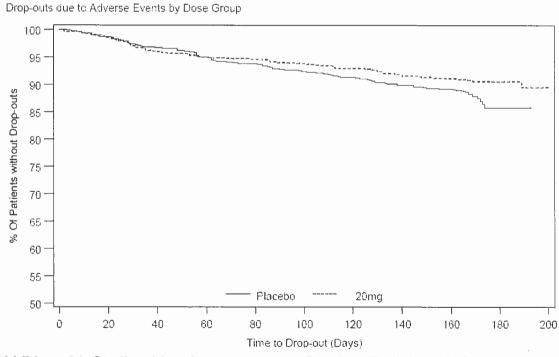


Abbildung 34: Studienabbrecher wegen unerwünschter Ereignisse (9 Studien)

3.1.6 Zusammenfassung der Ergebnisse der Mono- und Triple-Responderanalysen Die Ergebnisse der Responderanalysen über alle Studien aus der Subgruppenanalyse nach Krankheitsstadium (moderat – schwer) sind in Tabelle 3 zusammengefasst.

Tabelle 3: Zusammenfassung der Ergebnisse der Mono- und Triple-Responderanalysen (aus Subgruppenanalyse nach Krankheitsstadium)

Responsekriterium	Memanti	n	Plazebo		Odds-ratio	P-Wert		
	NV/n	%	NV/n	%				
Alltagspraktische Fähigkeiten (moderat – schwer) (Abb. 2)								
Mono-Responderanalyse 9 Studien	456/1196	38,1%	492/1134	43,4%	0.80 (0.68, 0.95)	0.01		
Kognitive Leistungsfähigkeit	(moderat –	schwe	r) (Abb. 4 u.	6)				
Mono-Responderanalyse 9 Studien	292/1195	24,4%	398/1137	35,0%	0.58 (0.48, 0.70)	< 0.001		
Sensitivitätsanalyse 8 Studien (ohne MEM-MD-22)	246/1064	23,1%	348/1007	34,6%	0.54 (0.44, 0.67)	< 0.001		
Klinisches Globalurteil (Abb. 8)								
Mono-Responderanalyse 8 Studien	428/1076	39,8%	491/1011	48,6%	0.69 (0.57, 0.83)	< 0.001		
Triple-Responderanalysen mi	t zulassun	gskonfo	rmer Definit	ion der				
(moderat - schwer) (Abb. 10 u								
Triple-Responderanalyse EMA 9 Studien	116/1068	10,9%	194/1004	19,3%	0.50 (0.39, 0.65)	< 0.001		
Triple-Responderanalyse EMA 8 Studien (ohne MEM-MD-22)	106/937	11,3%	180/874	20,6%	0.49 (0.37, 0.64)	< 0.001		
Triple-Responderanalysen mit neuer Definition der Grenzwerte für alltagspraktische Fähigkeiten (moderat – schwer) (Abb. 14 u. 16)								
Triple-Responderanalyse neu 9 Studien	100/1068	9,4%	162/1004	16,1%	0.53 (0.40, 0.70)	< 0.001		
Triple-Responderanalyse neu 8 Studien (ohne MEM-MD-22)	92/937	9,8%	150/874	,	0.52 (0.39, 0.69)	< 0.001		
NV = Anzahl Patienten mit V	erschlecht	erung /	n = Gesamt	zahl Pa	atienten,			

## 3.2 Diskussion der ergänzenden Responderanalysen

Die ergänzenden Responderanalysen bestätigen die Überlegenheit von Memantin über Plazebo in vergleichbarer Größenordnung wie die seinerzeit für die EMA-Zulassung geforderte Triple-Responderanalyse. Die Überlegenheit für Memantin wird dabei sowohl in der zulassungskonformen Triple-Responderanalyse wie auch in der modifizierten Triple-Responderanalyse (gemäß IQWiG-Methode) als auch in der Mono-Responderanalyse zur kognitiven Leistungsfähigkeit sowohl für das Gesamtkollektiv als auch für die Subgruppen moderate und schwere Alzheimer Demenz nachgewiesen. Damit ist – wie vom IQWiG im Abschlussbericht gefordert – auch gezeigt, dass die Ergebnisse der zulassungskonformen Triple-Responderanalyse nicht durch die CIBICplus Ergebnisse dominiert werden.

Für die Mono-Responderanalysen alltagspraktische Fähigkeiten und klinisches Globalurteil sind die Ergebnisse für die Subgruppe moderate Alzheimer Demenz und für das Gesamtkollektiv signifikant, während für die Subgruppe schwere Alzheimer Demenz hier aufgrund der deutlich geringeren Patientenzahlen die Signifikanz verfehlt wird.

Auch die zusätzlichen vom IQWiG geforderten Sensitivitätsanalysen, bei denen alle Studienabbrecher (Drop outs) als "deutlich klinisch verschlechtert" gewertet wurden, bestätigen diese Ergebnisse. Die Kaplan-Meier-Kurven zeigen, dass in der Memantin-Gruppe numerisch weniger Patienten die Studie abbrechen als in der Plazebo-Gruppe. Dies gilt sowohl für die Betrachtung aller Abbruchursachen als auch für Abbrüche wegen unerwünschter Ereignisse.

Die Analysen zeigen, dass jeweils mehr Patienten unter Plazebo eine deutliche Verschlechterung aufweisen als unter Memantin, die Unterschiede sind signifikant.

Damit werden mit den hier vorgelegten Mono- und Triple-Responderanalysen, bei denen die Grenzwerte auch entsprechend den vom IQWiG im Arbeitspapier Memantin gemachten Vorgaben gewählt wurden, die Ergebnisse der EMA-Zulassung bestätigt.

Eine Einschränkung der Memantin-Kostenerstattung der GKV stünde damit im Widerspruch zur Beurteilung der europäischen Zulassungsbehörde und würde somit die höchstrichterliche Feststellung der Vorgreiflichkeit von behördlichen Zulassungsentscheidungen verletzen.

## Literatur

1	AU:	FDA Assessment Report
	TI:	Final Briefing Document for Advisory Committee Meeting
		Efficacy Review Of New Drug Application
	SO:	http://www.fda.gov/ohrms/dockets/ac/03/briefing/3979B1_05_FDA-Efficacy%20Review.pdf