

# **Positronenemissionstomographie (PET) und PET/CT bei Alzheimer-Demenz**

**Dokumentation und Würdigung der  
Anhörung zum vorläufigen Berichtsplan**

Auftrag: D06-01N  
Version: 1.0  
Stand: 05.06.2013

# Impressum

**Herausgeber:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

**Thema:**

Positronenemissionstomographie (PET) und PET/CT bei Alzheimer-Demenz

**Auftraggeber:**

Gemeinsamer Bundesausschuss

**Datum des Auftrags:**

21.12.2006

**Interne Auftragsnummer:**

D06-01N

**Anschrift des Herausgebers:**

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

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: [berichte@iqwig.de](mailto:berichte@iqwig.de)

Internet: [www.iqwig.de](http://www.iqwig.de)

# Inhaltsverzeichnis

	Seite
<b>Abkürzungsverzeichnis.....</b>	<b>iv</b>
<b>1 Dokumentation der Anhörung .....</b>	<b>1</b>
<b>2 Würdigung der Anhörung .....</b>	<b>2</b>
<b>2.1 Anmerkungen zum Kapitel „Hintergrund“ des Berichtsplans .....</b>	<b>2</b>
<b>2.2 Fragestellung des Berichts.....</b>	<b>2</b>
2.2.1 Umgang mit Tracern.....	2
2.2.2 Vergleich von PET und SPECT .....	3
<b>2.3 Methoden .....</b>	<b>3</b>
<b>2.4 Grundsätzliche Kommentare zur Bewertungsmethodik.....</b>	<b>3</b>
2.4.1 Meta-Analysen .....	6
<b>2.5 Einschlusskriterien.....</b>	<b>6</b>
2.5.1 Population.....	6
2.5.2 Indextest .....	6
2.5.3 Referenzstandard .....	7
2.5.4 Studientypen .....	9
<b>2.6 Literaturverzeichnis.....</b>	<b>11</b>
<b>3 Offenlegung potenzieller Interessenkonflikte .....</b>	<b>14</b>
<b>3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen,         Institutionen und Firmen .....</b>	<b>15</b>
<b>3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der         wissenschaftlichen Erörterung (externer Sachverständiger).....</b>	<b>15</b>
<b>4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll.....</b>	<b>17</b>
<b>4.1 Teilnehmerliste der wissenschaftlichen Erörterung .....</b>	<b>17</b>
<b>4.2 Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen         Erörterung teilgenommen hat.....</b>	<b>17</b>
<b>4.3 Tagesordnung der wissenschaftlichen Erörterung .....</b>	<b>18</b>
<b>4.4 Protokoll der wissenschaftlichen Erörterung.....</b>	<b>18</b>
4.4.1 Begrüßung und Einleitung.....	18
4.4.2 Tagesordnungspunkt 1: Studientypen (auch retrospektive Studien einschließen?).....	19
4.4.3 Tagesordnungspunkt 2: Gemeinsame Bewertung von PET und Radiopharmaka (widerspricht dies dem Auftrag des G-BA?) .....	24

4.4.4	Tagesordnungspunkt 3: Einschluss von Patientinnen ohne Risiko, Verdacht oder Symptome (sollen Screeningstudien berücksichtigt werden?).....	26
4.4.5	Tagesordnungspunkt 4: Nachbeobachtungszeit für das Referenzverfahren bei Diagnosestudien (18 Monate adäquat?) .....	34
4.4.6	Tagesordnungspunkt 5: Verschiedenes .....	38
<b>Anhang A – Dokumentation der Stellungnahmen .....</b>		<b>46</b>



### Abkürzungsverzeichnis

<b>Abkürzung</b>	<b>Bedeutung</b>
AA	Alzheimer's Association
AD	Alzheimer Demenz
ADRDA	Alzheimer's Disease and Related Disorders Association Work Group
CCT	Controlled Clinical Trial (kontrollierte klinische Studie)
CDR	Clinical Dementia Rating
CT	Computertomographie
DGN	Deutsche Gesellschaft für Nuklearmedizin e. V.
FDG	Fluor-18-2-Fluoro-2-Deoxy-D-Glukose
G-BA	Gemeinsamer Bundesausschuss
HTA	Health Technology Assessment
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
M. Alzheimer	Morbus Alzheimer
MCI	Mild Cognitive Impairment (leichte kognitive Beeinträchtigung)
MRT	Magnetresonanztomographie
NIA	National Institute on Aging
NINCDS	National Institute of Neurologic, Communicative Disorders and Stroke
PET	Positronenemissionstomographie
PET/CT	Integration von PET und CT in einem Untersuchungsgerät
QUADAS	QUality Assessment of studies of Diagnostic Accuracy included in Systematic reviews
RCT	Randomized Controlled Trial (randomisierte kontrollierte Studie)
RI	Reagan Institute
SMI	Subjective Memory Impairment (subjektive Gedächtnisstörungen)
SPECT	Single-Photon-Emissions-Computertomographie
SUV	Standardized Uptake Value

## **1 Dokumentation der Anhörung**

Der vorläufige Berichtsplan in der Version 1.0 vom 09.01.2012 wurde am 18.01.2012 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 15.02.2012 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 5 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 16.04.2012 im IQWiG diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 4.

Eine Würdigung der in der Anhörung vorgebrachten Aspekte befindet sich in Kapitel 2. Im überarbeiteten Berichtsplan sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der überarbeitete Berichtsplan ist auf der Website des IQWiG unter [www.iqwig.de](http://www.iqwig.de) veröffentlicht.

## **2 Würdigung der Anhörung**

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Berichtsplans überprüft. Die wesentlichen Argumente werden im Folgenden diskutiert. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, zum Beispiel zu rechtlichen Vorgaben für das Institut, angesprochen. Auf solche Punkte wird im Rahmen dieser projektspezifischen Würdigung der Anhörung nicht weiter eingegangen.

### **2.1 Anmerkungen zum Kapitel „Hintergrund“ des Berichtsplans**

*In einer Stellungnahme wird angemerkt, dass bestimmte Ausführungen im Berichtsplan offenbar aus älteren IQWiG-Produkten stammten und sich daraus „Ungenauigkeiten“ für den aktuellen Bericht ergäben. So sei beispielsweise die SUV-Quantifizierung der zerebralen FDG-Aufnahme unüblich. Ebenso spiele die Bildfusion mit der CT kaum eine Rolle; vielmehr sei die Koregistrierung der PET mit MRT-Bildern mitunter hilfreich.*

Eine bewertende Aussage dazu, was in der PET-Diagnostik als üblich angesehen wird oder eine Rolle spielt, kann im Rahmen des vorliegenden Berichts nicht geleistet werden. Das IQWiG orientiert sich bei der Festlegung seiner methodischen Vorgehensweise an den Vorgaben aus der Konkretisierung und stellt die Daten der eingeschlossenen Studien systematisch und vollständig dar. Die Konkretisierung sieht für die Bewertung explizit die Bewertung der PET und der PET/CT vor. In Vorabrecherchen wurden mehrere Arbeiten gefunden, die eine Hybridtechnologie einsetzen und /oder Angaben zum SUV machten. Auch diese Arbeiten werden daher berücksichtigt. Da die PET/MRT erst Ende 2010 eingeführt wurde und bislang an nur sehr wenigen Standorten in Deutschland verfügbar ist [1], ist das Vorhandensein von Studien zur PET/MRT und deren theoretische Übertragbarkeit auf die medizinische Versorgungsroutine in Deutschland fraglich. Die Auftragskonkretisierung durch den G-BA sieht eine Berücksichtigung der PET/MRT nicht vor.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

### **2.2 Fragestellung des Berichts**

In mehreren Stellungnahmen wurden verschiedene Vorschläge für eine Erweiterung bzw. Änderung der Fragestellung eingebracht. Diese werden im Folgenden diskutiert.

#### **2.2.1 Umgang mit Tracern**

*In einer Stellungnahme wird bemängelt, dass die Formulierung „keine Eingrenzung auf bestimmte Tracer“ zu unklar sei. Die stellungnehmende Organisation hält einen Vergleich unterschiedlicher Tracer für unbedingt erforderlich. Sie schlägt einen „gesonderten Technologievergleich aller zur Diagnose der Alzheimer-Demenz eingesetzten Tracer unter Berücksichtigung der verschiedenen gemessenen Parameter“ vor.*

Da der Berichtsplan keine Eingrenzung auf bestimmte Tracer vorsieht, bedeutet dies, dass alle Studien, die den Einschlusskriterien genügen – unabhängig vom eingesetzten Tracer –, eingeschlossen werden. Ein direkter Vergleich unterschiedlicher Tracer innerhalb einer Studie wird selten zu finden sein. Sofern genügend Einzelstudien zu unterschiedlichen Tracern eingeschlossen werden, sind indirekte Vergleiche möglich, welche grundsätzlich mit einem hohen Verzerrungspotenzial behaftet sind. Unterschiedliche Tracer werden daher als potenzielle Effektmodifikatoren angesehen.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

### **2.2.2 Vergleich von PET und SPECT**

*In einer Stellungnahme wird darauf hingewiesen, dass die PET im Vergleich zur SPECT „unabhängig vom Einsatz radioaktiver Medikamente grundsätzliche messtechnische Vorteile gegenüber herkömmlichen nuklearmedizinischen Messverfahren wie der planaren Szintigraphie oder der SPECT („Single photon emission computed tomography“)" aufweise. „Der Vergleich der PET gegenüber der SPECT [sei] recht einfach über technische Abbildungsparameter zu bewerkstelligen, die unter anderem in der propädeutischen Literatur zur Nuklearmedizin zu finden sind. Auch im klinischen intraindividuellen Vergleich [seien] die Unterschiede zwischen FDG-PET und FDG-SPECT so offensichtlich, dass umfangreiche klinische Studien ethisch nicht vertretbar [seien], um diesen Technologievergleich nach [den] formalen Vorgaben [des IQWiG] durchzuführen.“*

In der Auftragskonkretisierung durch den G-BA wird der Vergleich mit Standardverfahren explizit gefordert. Den Vergleich mit der SPECT zu vernachlässigen erscheint daher für den vorliegenden Bericht nicht adäquat. Bei anderen Erkrankungen wird der Vergleich zwischen der PET und SPECT ebenfalls angestellt (z. B. bei der koronaren Herzerkrankung, beim Mammakarzinom etc.). Auch wenn der Vorteil der PET aus nuklearmedizinischer Sicht offensichtlich erscheinen mag, ist es der Zweck einer Evaluation, diesen Vorteil empirisch zu bestätigen. Genau dieser Vergleich wird auch in einer Meta-Analyse angestellt, die von der stellungnehmenden Organisation zitiert wird (Yuan 2009 [2]).

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

## **2.3 Methoden**

### **2.4 Grundsätzliche Kommentare zur Bewertungsmethodik**

*In einer Stellungnahme wird das Vorgehen der Bewertung der gesamten diagnostisch-therapeutischen Kette infrage gestellt. Hier wird vorgeschlagen, alleine die diagnostische Güte zu bewerten, da der Nutzen von therapeutischen Verfahren (z. B. medikamentöse Behandlung mit der Gabe von Memantin) bereits in vorangegangenen Berichten des Instituts nachgewiesen worden sei. Da die PET-Diagnostik in die Patientenselektion einfließe, sei*

*dieser für das Medikament Memantin erbrachte Nutzenbeleg nicht vom Nutzen der PET-Diagnostik zu trennen.*

Der Bericht des IQWiG zu Memantin schloss 5 Studien ein [3]. In keiner dieser Studien finden sich Hinweise auf eine Patientenselektion mithilfe der PET. Als diagnostische Verfahren werden in allen 5 Studien die kraniale Computertomographie und die Magnetresonanztomographie angegeben. Aus diesem Grund ergeben sich aus dem vorhandenen Bericht bislang keine Schlussfolgerungen zum Nutzen der PET bei der Selektion von Patienten, die von einer Memantintherapie profitieren könnten. Sollten jedoch derartige Studien bei den Recherchen zum vorliegenden Bericht identifiziert werden, so werden sie im Rahmen der Fragestellung 1 (patientenrelevanter Nutzen) eingeschlossen und berücksichtigt.

Die Notwendigkeit der Bewertung der gesamten diagnostisch-therapeutischen Kette zur Ermittlung eines patientenrelevanten Nutzens oder Schadens der PET wurde im Rahmen eines wissenschaftlichen Austausches zwischen der DGN und dem IQWiG bereits ausführlich diskutiert [4,5]. Auch der aktuelle internationale methodische Diskurs legt eine Bewertung der gesamten diagnostisch-therapeutischen Kette nahe [6-9].

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

*In einer Stellungnahme wird vorgeschlagen, die Einschlusskriterien für Studien zur diagnostischen Güte nicht auf jene mit unbekanntem Krankheitsstatus (Phase 3 nach Köbberling [10]) einzuschränken, sondern auch Studien der Phase 2 zu akzeptieren (bekannter Krankheitsstatus; diagnostische Fall-Kontroll-Studien). Es wird befürchtet, dass durch diese Einschränkung möglicherweise entscheidende Informationen zur Bewertung der diagnostischen Güte nicht berücksichtigt würden.*

Diagnostische Fall-Kontroll-Studien verwenden ein stark selektiertes Probandenkollektiv, das mit Kollektiven in realen klinischen Situationen kaum vergleichbar ist. Entsprechend werden die Ergebnisse zur diagnostischen Güte kaum auf die klinische Praxis übertragbar sein. Informationen zur diagnostischen Güte der PET in artifiziellen Kollektiven sind daher zur Bewertung der diagnostischen Güte in realen klinischen Situationen nicht geeignet. So wurden beispielsweise im Rahmen der Überarbeitung der QUADAS-Checkliste 7 Studien gefunden, die zeigten, dass in diagnostischen Fall-Kontroll-Studien die diagnostische Güte systematisch überschätzt wird [11].

Die Beschränkung der Einschlusskriterien auf Studien der Phase 3 (und 4) erscheint daher berechtigt.

*In einer Stellungnahme wird vorgeschlagen, auch retrospektive Studien und jene mit einer Obduktion als Referenzstandard zu akzeptieren. Andernfalls würden wichtige Studien übersehen.*

Dieser Punkt wurde auch in der Erörterung zu diesem Projekt diskutiert. Im Anschluss an die Erörterung wurde dem Institut von der DGN eine Liste von Studien zur Verfügung gestellt, die nach Auffassung der DGN für die Berichtsfragestellung hochrelevant, aber retrospektiv sind. Das IQWiG versicherte, die Liste genau zu prüfen.

Aus der Liste der 66 genannten Primärstudien waren bereits 13 im Rahmen der bibliografischen Recherchen eingeschlossen und 46 ausgeschlossen worden. Bei allen ausgeschlossenen Studien war unabhängig davon, ob das Studiendesign prospektiv oder retrospektiv war, ein weiteres Einschlusskriterium des vorliegenden Berichts verletzt. Zusätzlich gab es 6 Referenzen, die erst nach dem Zeitraum der Primärrecherche publiziert wurden. Diese werden erst im Rahmen der Nachrecherche (nach der Publikation des Vorberichts) gewürdigt. Die verbleibende Studie von Silverman et al. [12] ist eine retrospektive Studie mit einer Nachbeobachtung von mehr als 10 Jahren. Sie schließt 128 Patienten ein, stellt jedoch keinen direkten Vergleich der PET mit einer anderen diagnostischen Modalität dar.

Die von der DGN zur Verfügung gestellte Liste lieferte daher keine Hinweise auf retrospektive Studien, deren Einschluss relevante Erkenntnisse zur Fragestellung des vorliegenden Berichts geliefert hätte. Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

*In einer Stellungnahme wird darauf hingewiesen, dass die Bewertung des Verzerrungspotenzials eingeschlossener Studien mithilfe des QUADAS-Instruments „beim Einsatz für die Meta-Analyse von Studien mit bildgebender Diagnostik [...] nicht unumstritten“ ist. Als Referenz wird hier eine Arbeit von Weber et al. angeführt [13].*

Die Argumente von Herrn Prof. Weber wurden in einem Leserbrief der QUADAS-Autoren zu seinem Artikel widerlegt [14]. Darüber hinaus wurden die Kritikpunkte im Rahmen einer Erörterung zu einem anderen PET-Bericht am 29.08.2012 im IQWiG mit Herrn Prof. Weber persönlich und ausführlich diskutiert (die Dokumentation folgt im Rahmen des Abschlussberichts D06-01H). Der Standpunkt des IQWiG ist dabei, dass es durchaus Fragestellungen geben kann, zu denen eine diagnostische Studie ohne hohes Verzerrungspotenzial nicht durchführbar ist. Herr Prof. Weber nennt hier als Beispiel die Detektion von Fernmetastasen. Die Tatsache jedoch, dass zu dieser Fragestellung keine Studie ohne hohes Verzerrungspotenzial durchführbar ist, kann nicht den vorhandenen Instrumenten zur Erfassung des Verzerrungspotenzials angelastet werden. In diesen Fällen kann daher die methodische Lösung keinesfalls darin bestehen, die Anforderungen an methodisch hochwertige Diagnosestudien aufzulockern. Vielmehr sollte man sich in solchen Situationen die Frage stellen, ob nicht andere Studiendesigns (z. B. vergleichende Interventionsstudien) eine hilfreiche und informative Alternative darstellen würden.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

### **2.4.1 Meta-Analysen**

*Eine stellungnehmende Organisation schlägt vor, Meta-Analysen, die keine bedeutsame Heterogenität aufweisen, statt mit Modellen mit zufälligen Effekten mit solchen mit festen Effekten durchzuführen.*

Die Methodik des vorliegenden Berichts sieht bivariate Meta-Analysen vor. Bisher ist es nicht möglich, für bivariate Meta-Analysen Heterogenitätstests durchzuführen, daher kann nicht ermittelt werden, in welchen Fällen eine bedeutsame Heterogenität vorliegt. Bei homogener Datenlage unterscheiden sich die metaanalytischen Ergebnisse von Modellen mit zufälligen und festen Effekten ohnehin nur marginal. Daher kann in jedem Fall ein Modell mit zufälligen Effekten berechnet werden.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

## **2.5 Einschlusskriterien**

### **2.5.1 Population**

*In einer Stellungnahme wird vorgeschlagen, die Primärstudienrecherche auf die Differenzialdiagnostik aller demenziellen Krankheiten auszuweiten. Außerdem ist den Stellungnehmenden nicht klar, ob die Einschlusskriterien auch Personen mit MCI (= Mild Cognitive Impairment) einschließen. In der Erörterung wurde des Weiteren darauf hingewiesen, dass auch eine sogenannte SMI (= Subjective Memory Impairment) ein Risiko für die Herausbildung einer MCI und im weiteren Verlauf einer Alzheimer-Demenz darstelle.*

Der Auftrag des G-BA bezieht sich eindeutig auf die Diagnostik der Alzheimer-Demenz und nicht grundsätzlich aller Demenzformen. Das Institut bezieht sich bei der Festlegung der Einschlusskriterien auf den Auftrag bzw. die Konkretisierung des G-BA. Die Eignung der PET zur Differenzialdiagnostik wird daher im Rahmen des vorliegenden Berichts nur mit Fokus auf die Alzheimer-Demenz erfolgen.

Da Patienten mit MCI ein erhöhtes Risiko haben, eine Alzheimer-Demenz zu entwickeln, fallen auch Studien, die Personen mit MCI untersuchen, in diese Kategorie. Selbiges gilt für Studien zu Patienten mit SMI, sofern sie allen anderen Einschlusskriterien genügen.

Die Definition der einzuschließenden Population wurde im Berichtsplan präzisiert.

### **2.5.2 Indextest**

*In einer Stellungnahme wird darauf hingewiesen, dass die PET ein nicht invasives diagnostisches Verfahren darstelle und aufgrund der geringeren Komplikationsraten bzw. Ablehnungs- und Misserfolgsquoten einen Vorteil gegenüber den invasiven Verfahren (z. B. Liquordiagnostik) habe.*

Der Endpunkt der unerwünschten Arzneimittelwirkungen wurde allgemeiner als „andere unerwünschte Wirkungen“ gefasst, sodass mit dieser Formulierung Komplikationen sowohl der Diagnostik als auch der Therapie umfasst werden.

### 2.5.3 Referenzstandard

*In 2 Stellungnahmen wird vorgeschlagen, zusätzlich zum kognitiven Status auch die histopathologische Untersuchung im Rahmen einer Obduktion als Referenzstandard aufzunehmen.*

Wie im Berichtsplan bereits dargestellt, wird der Stellenwert des postmortalen Nachweises übermäßiger neurofibrillärer Verklumpungen und neuritischer Plaques, welche zur Diagnosesicherung – etwa im Rahmen der NINCDS-ADRDA-Kriterien – eingesetzt werden kann [15], von neueren Arbeiten kritisch betrachtet [16-18]. In einer Arbeit wird berichtet, dass lediglich 40 % bis 57 % der untersuchten Gehirne von Patienten, die zu Lebzeiten eine wahrscheinliche Alzheimer-Demenz aufwiesen, die Pathologie einer reinen AD aufwiesen [19]. Dabei bezogen sich die Autoren auf die Kriterien des US-amerikanischen National Institute on Aging und des Reagan Institute (NIA-RI) [20]. Auch die überarbeiteten Kriterien des National Institute on Aging und der Alzheimer's Association (NIA-AA) für die neuropathologische Begutachtung basieren auf Daten, die lediglich eine diffuse Korrelation zwischen klinischen Symptomen einer AD, neurofibrillären Verklumpungen und neuritischen Plaques aufweisen [21].

Seit der Publikation der NINCDS-Kriterien aus 1984 [15] gab es Empfehlungen, die Definition der „sicheren AD“ auszudehnen. So haben Dubois et al. vorgeschlagen, dass klinische und genetische Belege einer AD für eine „sichere“ AD-Diagnose schon zu Lebzeiten ausreichend seien [22]. Eine Reihe von kritischen Arbeiten zur Gewichtung der postmortalen neuropathologischen Begutachtung des Gehirns in der NINCDS-Diagnose der sicheren AD folgte [16-18], darunter auch Studien, die zeigten, dass mit zunehmendem Alter die Korrelation zwischen dem Vorhandensein von AD-typischen neuropathologischen Läsionen und klinischer Demenz abnimmt [23]. 2011 wurde empfohlen, der Neuropathologie eine bescheidenere Rolle in der AD-Diagnose beizumessen [17]. Sogar die Empfehlungen der Autoren der NINCDS-ADRDA-Kriterien zeigen in diese Richtung: Das, was in 1984 als „definite AD“ bezeichnet wurde, wird in den NIA-AA-Empfehlungen aus 2011 lediglich als „pathophysiologisch bestätigte AD“ definiert [24].

Insgesamt besitzt die histopathologische Untersuchung des Gehirns post mortem keine hinreichende Validität für die Fragestellung des IQWiG-Berichts. Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

*In einer Stellungnahme wird angemerkt, dass die klinische Definition des M. Alzheimer mittlerweile an die PET-Diagnostik gekoppelt sei. Dies gehe aus den aktuellen NINCDS-Kriterien hervor. Im Berichtsplan finde sich weder ein Bezug auf die „Revised NINCDS-Criteria“ [22] noch auf die neuen AD-Diagnosekriterien der Arbeitsgruppe des National*



*Institute on Aging und der Alzheimer's Association (NIA-AA) [24], in denen explizit die Nutzung von Biomarkern wie der PET empfohlen werde.*

Die Publikation, die von den Stellungnehmenden als „Revised NINCDS-Criteria“ bezeichnet wird, trägt den Untertitel „revising the NINCDS-ADRDA criteria“ [22]. Auch der Text spiegelt weniger eine feststehende Diagnoserichtlinie, sondern vielmehr den aktuellen Stand eines Forschungsprozesses wider und macht einen Vorschlag zur Modifikation der Diagnostik. Insgesamt bezieht sich die genannte Publikation auch nicht auf den Einsatz der PET im klinischen Alltag, sondern allein auf die Rolle der PET im Rahmen klinischer Studien. Die PET wird hier nicht als obligatorischer Bestandteil der Diagnostik, sondern lediglich als „supportive feature“ angeführt [22]. Die Arbeitsgruppe des National Institute on Aging und der Alzheimer's Association (NIA-AA) kommt zu dem Ergebnis, dass zum aktuellen Zeitpunkt der routinemäßige Einsatz von Biomarkern zur Diagnostik der Alzheimer-Demenz nicht empfohlen werden kann. Biomarker sollten weiterhin lediglich im Rahmen von Forschungsprojekten eingesetzt werden [24]. Des Weiteren wird betont, dass die Diagnose der Alzheimer-Demenz keinen Einsatz von Biomarkern erfordere und, solange keine neueren Ergebnisse vorlägen, ein konservativer Umgang mit Biomarkern der vernünftige Weg sei. Diese Empfehlung steht der Behauptung in der Stellungnahme entgegen.

Die Arbeiten von Dubois et al. [22] und McKhann et al. [24] werden in den Berichtsplan aufgenommen. Es ergeben sich keine Konsequenzen für den Referenztest.

*In 2 Stellungnahmen wird darüber hinaus nach der Begründung für eine 18-monatige Nachbeobachtung als Teil des Referenzstandards gefragt. Die Länge der Nachbeobachtung war auch ein Thema der Erörterung.*

Die Nachbeobachtungsdauer von 18 Monaten ist anhand diverser Kohortenstudien gewählt worden, die zeigen, dass nur ein sehr geringer Anteil der Personen mit MCI unterhalb dieser Zeitspanne eine Alzheimer-Demenz entwickelt. Beispielsweise zeigte die Studie von Morris et al. [25], dass es sogar bei Patienten mit einer Alzheimerdiagnose konform mit psychometrischen Testergebnissen des CDR von 0,5/DAT noch durchschnittlich 10 Jahre dauerte, bis sie einen deutlichen kognitiven Abbau (von CDR 0,5 auf CDR 1,0) aufwiesen. Nach 12 Monaten erlebten in dieser Gruppe etwa 15 % der Patienten einen deutlichen kognitiven Abbau, nach 18 Monaten waren es etwa 30 %. Patienten, bei denen zu einem frühen Zeitpunkt eine unklare („uncertain“) oder beginnende („incipient“) Demenz diagnostiziert wurde, entwickelten bei einer Nachbeobachtung von 10 Jahren zu etwa 28 % bzw. 42 % einen deutlichen Abbau der kognitiven Fähigkeiten. Nach 1 Jahr waren in beiden Gruppen weniger als 5 % betroffen, nach 18 Monaten etwa 10 %. Der Referenzstandard klinische Diagnostik nach 12 Monaten (oder weniger) ist daher für diese Gruppen nicht ausreichend valide.

Zu ähnlichen Ergebnissen kam auch die Studie von Hansson et al. [26]. In dieser Kohorte entwickelten 14 bis 41 % der MCI-Patienten eine klinisch diagnostizierte Alzheimer-Demenz

in einem Zeitraum von 2 bis 7 Jahren. Auch hier wäre also eine Nachbeobachtung von lediglich einem Jahr vermutlich zu kurz gewesen.

Eine Nachbeobachtung von 5 Jahren oder mehr zu fordern, wäre unter studentechischen Gesichtspunkten kaum realistisch. Vor diesem Hintergrund bietet eine Nachbeobachtung von durchschnittlich 18 Monaten zwar keinen 100%ig validen Referenzstandard, was jedoch als pragmatischer Kompromiss hinzunehmen ist. Sollten Studien mit längeren Nachbeobachtungszeiten identifiziert werden, so werden sie aufgrund der höheren Validität ihrer Referenzstandards besonders berücksichtigt.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

*In einer Stellungnahme wird vorgeschlagen, CCTs, welche die Sensitivität und Spezifität eines Tests untersuchen, keinesfalls als Studien der zweiten Kategorie zu klassifizieren, sondern als Standard für Studien zur Bewertung von Diagnostika zu betrachten.*

Der Berichtsplan sieht vor, CCTs für die Nutzenbewertung heranzuziehen, sofern die Datenlage aus RCTs nicht hinreichend ist, um den patientenrelevanten Nutzen und Schaden der PET bzw. PET/CT mit ausreichender Ergebnissicherheit schätzen zu können (vgl. Abschnitt 4.1.2.4 des Berichtsplans). Die Aussagekraft von RCTs bezüglich des kausal auf die PET bzw. PET/CT zurückzuführenden patientenrelevanten Nutzens oder Schadens ist – wie auch bei der Bewertung medikamentöser Interventionen – erheblich höher als jene von CCTs. Daher erscheint die aktuelle Hierarchisierung der Evidenz aus unterschiedlichen Studientypen gerechtfertigt. Darüber hinaus werden auch die Ergebnisse von (nicht interventionellen) Kohortenstudien zur Ermittlung der diagnostischen und prognostischen Güte der PET im Rahmen der Fragestellung 2 des Berichts eingeschlossen und gewürdigt.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

#### **2.5.4 Studientypen**

*In einer Stellungnahme wird angeregt, auch indirekte Vergleiche zwischen der PET bzw. PET/CT und anderen diagnostischen Verfahren zuzulassen. Die Stellungnehmenden gehen davon aus, dass „der G-BA [diese] Informationen sehr wohl bei seiner weiteren Begutachtung benötigen wird, z. B. um festzustellen, ob es sich um Untersuchungsverfahren mit Potenzial für eine ‚Erprobung‘ gemäß §137e SGB V handelt“. Die Frage, ob auch indirekte Vergleiche angestellt werden sollten, wurde auch im Rahmen der Erörterung diskutiert.*

Indirekte Vergleiche bergen ein erhebliches Verzerrungspotenzial. Einerseits finden diese Vergleiche kaum zeitlich parallel statt. Zeitliche Verschiebungen können aber in Veränderungen in der Durchführung der Indextests oder Referenztests oder der verabreichten Medikamente, durchgeführten Operationen etc. resultieren, was zu Verzerrungen in den Ergebnissen zur diagnostischen Güte führen kann [27].

Noch wesentlicher ist aber die Tatsache, dass indirekte Vergleiche immer auf unterschiedlichen Patientenkollektiven basieren, was insbesondere bei den typischerweise kleinen Stichproben in diesem Bereich ein erhebliches Verzerrungspotenzial birgt. In der Erörterung wurde darauf hingewiesen, dass Gütekriterien für diagnostische Verfahren – die Sensitivität und die Spezifität – viel stärker von den Populationen abhängen als die Therapieeffekte in Therapiestudien. Das Problem der Heterogenität, das für die indirekten Vergleiche von zentraler Relevanz ist, wird daher bei einarmigen Diagnosestudien noch viel größer als bei indirekten Vergleichen von Therapiestudien.

Ein aktueller HTA-Bericht des NHS zur Rezidiverkennung bei Brustkrebs wendet die gleiche Methode wie das IQWiG an und kommentiert die Problematik indirekter Vergleiche ebenfalls ausführlich [27]. Das Problem der nicht vorhandenen Strukturgleichheit in indirekten Vergleichen diagnostischer Studien führt zu einem praktisch unkontrollierbaren Einfluss mannigfaltiger Störgrößen. Insbesondere bei kleineren diagnostischen Studien (Fallzahlen häufig kleiner 30) erscheinen daher indirekte Vergleiche unzulässig.

Ob der G-BA bei der Bewertung eines Potenzials indirekte Vergleiche berücksichtigen wird, ist zum Zeitpunkt der Erstellung des vorliegenden Berichtsplans noch nicht abzusehen. Das Institut kann sich daher nur auf die methodischen Argumente, wie sie in den vorhergehenden Absätzen dargestellt sind, berufen und stellt daher keine indirekten Vergleiche dar.

Es ergibt sich daher keine Änderung für die Methodik des Berichtsplans.

## 2.6 Literaturverzeichnis

1. Catana C, Drzezga A, Heiss WD, Rosen BR. PET/MRI for neurologic applications. *J Nucl Med* 2012; 53(12): 1916-1925.
2. Yuan Y, Gu ZX, Wei WS. Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. *AJNR Am J Neuroradiol* 2009; 30(2): 404-410.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Memantin bei Alzheimer Demenz: Abschlussbericht; Auftrag A05-19C [online]. 08.07.2009 [Zugriff: 17.04.2013]. (IQWiG-Berichte; Band 59). URL: [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).
4. Scheibler F, Raatz H, Suter K, Janssen I, Grosselfinger R, Schröer-Günther M et al. Nutzenbewertung der PET bei malignen Lymphomen: Standpunkt des IQWiG. *Nuklearmedizin* 2010; 49(1): 1-5.
5. Kotzerke J, Dietlein M, Grünwald F, Bockisch A. Vorbehalte der DGN zur Nutzenbewertung der PET durch das IQWiG. *Nuklearmedizin* 2010; 49(1): 6-12.
6. Buyse M, Michiels S, Sargent DJ, Grothey A, Matheson A, De Gramont A. Integrating biomarkers in clinical trials. *Expert Rev Mol Diagn* 2011; 11(2): 171-182.
7. Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, Bossuyt PM, Deeks JJ. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ* 2012; 344: e686.
8. Lijmer JG, Bossuyt PMM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.
9. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008; 336(7653): 1106-1110.
10. Köbberling J, Trampisch HJ, Windeler J. Memorandum for the evaluation of diagnostic measures. *J Clin Chem Clin Biochem* 1990; 28(12): 873-879.
11. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-536.
12. Silverman DH, Truong CT, Kim SK, Chang CY, Chen W, Kowell AP et al. Prognostic value of regional cerebral metabolism in patients undergoing dementia evaluation: comparison to a quantifying parameter of subsequent cognitive performance and to prognostic assessment without PET. *Mol Genet Metab* 2003; 80(3): 350-355.

13. Meyer PT, Hellwig S, Amtage F, Rottenburger C, Sahm U, Reuland P et al. Dual-biomarker imaging of regional cerebral amyloid load and neuronal activity in dementia with PET and <sup>11</sup>C-labeled Pittsburgh compound B. *J Nucl Med* 2011; 52(3): 393-400.
14. Kleijnen J, Westwood M, Wolff R, Whiting P, Raatz H, Kunz R. Evidence-based assessment of PET in Germany. *J Nucl Med* 2012; 53(7): 1166-1167.
15. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34(7): 939-944.
16. Jellinger KA. Con: can neuropathology really confirm the exact diagnosis? *Alzheimers Res Ther* 2010; 2(2): 11.
17. Scheltens P, Rockwood K. How golden is the gold standard of neuropathology in dementia? *Alzheimers Dement* 2011; 7(4): 486-489.
18. Tyas SL, Snowdon DA, Desrosiers MF, Riley KP, Markesbery WR. Healthy ageing in the Nun Study: definition and neuropathologic correlates. *Age Ageing* 2007; 36(6): 650-655.
19. Jellinger KA. Criteria for the neuropathological diagnosis of dementing disorders: routes out of the swamp? *Acta Neuropathol* 2009; 117(2): 101-110.
20. Varma AR, Snowden JS, Lloyd JJ, Talbot PR, Mann DM, Neary D. Evaluation of the NINCDS-ADRDA criteria in the differentiation of Alzheimer's disease and frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 1999; 66(2): 184-188.
21. Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement* 2012; 8(1): 1-13.
22. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007; 6(8): 734-746.
23. Savva GM, Wharton SB, Ince PG, Forster G, Matthews FE, Brayne C. Age, neuropathology, and dementia. *N Engl J Med* 2009; 360(22): 2302-2309.
24. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; 7(3): 263-269.
25. Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH et al. Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol* 2001; 58(3): 397-405.
26. Hansson O, Zetterberg H, Vanmechelen E, Vanderstichele H, Andreasson U, Londos E et al. Evaluation of plasma A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Neurobiol Aging* 2010; 31(3): 357-367.

27. Pennant M, Takwoingi Y, Pennant L, Davenport C, Fry-Smith A, Eisinga A et al. A systematic review of positron emission tomography (PET) and positron emission tomography/computed tomography (PET/CT) for the diagnosis of breast cancer recurrence. Health Technol Assess 2010; 14(50): 1-103.

### **3 Offenlegung potenzieller Interessenkonflikte**

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden sowie weiterer Teilnehmer an der wissenschaftlichen Erörterung zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter [www.iqwig.de](http://www.iqwig.de) abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

**3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen**

<b>Organisation / Institution / Firma</b>	<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>
BPI	Wilken, Matthias	ja	nein	nein	nein	nein	nein
DGGPP	Adler, Georg	nein	nein	nein	nein	nein	nein
	Haupt, Martin	nein	ja	ja	nein	nein	nein
DGN	Drzezga, Alexander	nein	ja	ja	ja	ja	nein
	Hellwig, Dirk	nein	ja	nein	nein	nein	nein
	Kotzerke, Jörg	nein	nein	nein	nein	nein	nein
	Sabri, Osama	nein	ja	ja	ja	nein	nein
	Schreckenberger, Mathias	nein	nein	ja	ja	nein	nein
DKG	Acikgöz, Özlem	ja	nein	nein	nein	nein	nein
	Dippmann, Anja	ja	nein	nein	nein	nein	nein
	Rüther, Klaus	ja	nein	nein	nein	nein	nein
	Schlottmann, Nicole	ja	nein	nein	nein	nein	nein
Lilly	Boess, Frank	ja	nein	nein	nein	nein	ja
	Lieb, Matthias	ja	nein	nein	nein	nein	nein
	Stahl, Oliver	ja	nein	nein	nein	nein	ja

**3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externer Sachverständiger)**

<b>Name</b>	<b>Frage 1</b>	<b>Frage 2</b>	<b>Frage 3</b>	<b>Frage 4</b>	<b>Frage 5</b>	<b>Frage 6</b>
Wolff, Robert	nein	nein	nein	ja	nein	nein



Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version 06/2011):

*Frage 1:* Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

*Frage 2:* Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten?

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

*Frage 4:* Haben Sie und / oder hat die Einrichtung<sup>1</sup>, die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

*Frage 5:* Haben Sie und/oder hat die Einrichtung<sup>1</sup>, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

*Frage 6:* Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

---

<sup>1</sup> Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

#### 4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

##### 4.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma / privat
Beckmann, Lars	IQWiG
Boess, Frank	Lilly-Pharma
Brockhaus, Catherina	IQWiG
Drzezga, Alexander	Deutsche Gesellschaft für Nuklearmedizin
Haupt, Martin	Deutsche Gesellschaft für Gerontopsychiatrie und -psychotherapie
Hausner, Elke	IQWiG
Hellwig, Dirk	Deutsche Gesellschaft für Nuklearmedizin
Lieb, Matthias	Lilly-Pharma
Menig, Matthias	IQWiG
Rörtgen, Thilo	Sitzungsdokumentarischer Dienst, Landtag NRW
Sauerland, Stefan	IQWiG
Scheibler, Fülöp	IQWiG
Schreckenberger, Mathias	Deutsche Gesellschaft für Nuklearmedizin
Slanger, Tracy	IQWiG
Windeler, Jürgen	IQWiG
Wolff, Robert	Kleijnen Systematic Reviews

##### 4.2 Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat

In der folgenden Tabelle werden Stellungnahmen genannt, zu denen trotz Einladung kein Stellungnehmender oder Vertreter zur wissenschaftlichen Erörterung erschienen ist.

Organisation / Institution / Firma / Privatperson
Bundesverband der Pharmazeutischen Industrie
Deutsche Krankenhausgesellschaft

### 4.3 Tagesordnung der wissenschaftlichen Erörterung

	<b>Begrüßung und Einleitung</b>
<b>TOP 1</b>	Studientypen (auch retrospektive Studien einschließen?)
<b>TOP 2</b>	Gemeinsame Bewertung von PET und Radiopharmaka (widerspricht dies dem Auftrag des G-BA?)
<b>TOP 3</b>	Einschluss von Patientinnen ohne Risiko, Verdacht oder Symptome (sollen Screeningstudien berücksichtigt werden?)
<b>TOP 4</b>	Nachbeobachtungszeit für das Referenzverfahren bei Diagnosestudien (18 Monate adäquat?)
<b>TOP 5</b>	Verschiedenes

### 4.4 Protokoll der wissenschaftlichen Erörterung

Datum: 16.04.2012, 11:00 bis 13:00 Uhr

Ort: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG),  
Im Mediapark 8, 50670 Köln

Moderation: Prof. Dr. Jürgen Windeler

#### 4.4.1 Begrüßung und Einleitung

**Moderator Jürgen Windeler:** Ich begrüße Sie in den neuen Räumen. Diejenigen, die schon einmal zu Erörterungen hier waren, erkennen das alles gar nicht wieder. Der Umstand, dass Sie an diesem vergleichsweise wenig großzügigen Tisch sitzen, zeigt, dass wir im Moment noch in der Anlaufphase des Umzugs des IQWiG in den Mediapark sind. Wir sind jetzt seit vier Wochen hier. Die Konferenzraumbestuhlung ist noch nicht ganz abgeschlossen. Beim nächsten Mal sieht das vielleicht ähnlich aus wie in der Dillenburger Straße, wo Sie einen etwas größeren Tisch und ein Boot an der Wand, das wir hier nicht mehr haben werden, gewöhnt waren.

Ich begrüße Sie zu der Erörterung und freue mich, dass Sie gekommen sind. Zunächst erläutere ich wie gewohnt und wie üblich am Anfang das, was erörtert werden muss.

Erstens. Die Erörterung wird per Tonband und per Protokollant aufgezeichnet. Das Gesprochene wird aufgeschrieben und im Internet veröffentlicht. Wir haben Sie bei der Einladung darauf hingewiesen. Sie haben – so verstehen wir das – mit Ihrem Erscheinen eingewilligt, dass diese Aufzeichnung stattfindet. Ich muss Sie darauf hinweisen, dass, wenn Sie jetzt noch anderer Meinung geworden sind, Sie die Erörterung verlassen müssen. Ansonsten können wir aber fortfahren. Da alle hier bleiben, gehen wir davon aus, dass wir so vollständig sein werden.

Im Zusammenhang mit der Aufzeichnung möchte ich Sie bitten, jedes Mal – nicht nur beim ersten Mal – Ihren Namen zu sagen, damit die Äußerungen zugeordnet werden können.

Abschließend möchte ich darauf hinweisen, dass die Erörterung der Klärung von Fragen dient, die wir aus der Befassung mit Ihren Stellungnahmen gewonnen haben, die sozusagen übrig geblieben sind. Gehen Sie also bitte davon aus, dass wir die Stellungnahmen sorgfältig gelesen haben, dass wir meinen, das meiste, was heute nicht thematisiert wird, verstanden zu haben, sodass wir uns jetzt auf einzelne relevante Fragen konzentrieren wollen. Das bedeutet auch, dass die Erörterung nicht dazu dient, insbesondere nicht dazu dienen muss, noch einmal die schriftlichen Stellungnahmen zu wiederholen. Es ist also anders als zum Beispiel bei Anhörungen im Gesundheitsausschuss oder an anderen Stellen. Das, was Sie schriftlich geäußert haben, ist bei uns angekommen.

Wir haben Ihnen eine Tagesordnung übermittelt. Ich frage Sie, ob Sie Streichungs- oder Ergänzungswünsche haben, die über spezifische und natürlich erlaubte Thematisierungen unter TOP 5 hinausgehen. – Das scheint nicht der Fall zu sein. Dann können wir in die Tagesordnung eintreten.

Ich rufe auf:

#### **4.4.2 Tagesordnungspunkt 1: Studientypen (auch retrospektive Studien einschließen?)**

**Fülöp Scheibler:** Guten Morgen. Die Deutsche Gesellschaft für Nuklearmedizin schlägt vor, für die Bewertung der PET bzw. PET/CT auch retrospektive Studien einzuschließen, und warnt davor, bei Berücksichtigung lediglich der prospektiven Designs wichtige Studien auszuschließen. Wir würden gerne wissen, auf welche Studien Sie sich im Einzelnen beziehen, um diese methodisch besser würdigen zu können. Wir hätten also gerne Beispiele von den Studien, wo Sie Sorge haben, dass die nicht berücksichtigt werden könnten.

**Moderator Jürgen Windeler:** Möchte sich jemand dazu äußern?

**Dirk Hellwig:** Schönen Dank für die Einladung. Wir als Deutsche Gesellschaft für Nuklearmedizin haben es gerne wahrgenommen, hier Stellung zu beziehen.

Die Frage, die für Sie offengeblieben ist, verwundert mich ehrlich gesagt ein bisschen, weil es ja hier erst einmal um den Berichtsplan geht. Das heißt, es geht um die Festsetzung der Methodik. Und Sie fragen uns jetzt schon direkt nach konkreten Studien. Das Ziel des Berichtsplans ist ja erst einmal, die Methodik festzulegen. Ihre Aufgabe besteht dann darin, mit dieser Methodik zu arbeiten und entsprechende Studien zu identifizieren. Wir haben ja schon in vorausgegangenen Stellungnahmeverfahren zu Ihren Produkten darauf hingewiesen und auch an anderer Stelle haben wir uns getroffen und darüber gesprochen, dass es Sinn macht, um den Kenntnisstand zusammenzutragen, den es zu einem zu bearbeitenden Auftrag gibt, dass das Wissen möglichst breit zusammengetragen wird.

Wir haben insbesondere in dieser Stellungnahme noch einmal darauf hingewiesen, weil es die erste Stellungnahme ist, welche im Rahmen der neuen Rechtslage verfasst wird. Schlussendlich geht es ja auch bei der Bewertungsarbeit des Gemeinsamen Bundesausschusses, der schließlich Ihren abgeschlossenen Nutzenbericht verwenden möchte, darum, zu entscheiden, ob, wenn es nicht zu einem direkten Einschluss des Verfahrens kommt, dieses Verfahren Potenzial aufweist. Auch wenn nach meinem Kenntnisstand die neue Verfahrensordnung des Gemeinsamen Bundesausschusses immer noch nicht abgeschlossen ist und wir eigentlich gar nicht wissen, was eine Definition für ein Verfahren mit Potenzial ist, sollte auf keinen Fall die Möglichkeit verwehrt werden, hier das Wissen breit zusammenzutragen, um eine Entscheidung dafür fällen zu können.

Deswegen, denke ich, macht es jetzt nicht unbedingt Sinn, ganz konkret Studien zu benennen, sondern Kriterien festzulegen, nach denen Ihre Recherche funktioniert. Wir haben ja auch früher schon darauf hingewiesen, dass beispielsweise andere Meta-Analysen – ich verstehe Ihre Arbeitsweise erst einmal so, dass eine Meta-Analyse erstellt wird, auch mit quantitativer Informationssynthese –, die Sie für Ihren weiteren Nutzenbericht verwenden, durchaus zulassen, dass retrospektive Studien verwendet werden. Nach meinem Kenntnisstand ist ja Ihre methodische Anforderung, dass ein Vergleich innerhalb der Meta-Analysen durchgeführt wird zwischen den Ergebnissen von Studien mit retrospektivem Design gegenüber Studien mit prospektivem Design. Das heißt also, wenn Sie Studien mit retrospektiver Erfassung von Daten zulassen und sich gleichzeitig prospektive Studien anschauen und das vergleichen, dann kann ja Ihr Nutzenbericht auch Ihren eigenen Qualitätskriterien für die Akzeptanz von Meta-Analysen standhalten.

**Fülöp Scheibler:** Ich stimme Ihnen weitestgehend zu, bis auf den letzten Punkt. Das Ziel des Berichtes ist nicht, einen Vergleich zwischen prospektiven und retrospektiven Studien zu machen. Das wäre ja eine methodische Fragestellung, die ja schon andernorts mehrfach durchgeführt wurde. Und es zeigt sich, dass retrospektive Studien sehr stark verzerrt sein können. Das ist ja der Grund, warum wir sagen, dass wir uns grundsätzlich, wenn wir schon auf die diagnostische Güte ausweichen müssen, weil es keine Studien zum patientenrelevanten Nutzen gibt, wenigstens von den Diagnosestudien nur die möglichst unverzerrtesten angucken, und das sind eben prospektive Studien. Jeder, der weiß, wie retrospektive Diagnosestudien zustande kommen, weiß auch, dass es auf jeden Fall ein riesengroßes Problem gibt, das ist eben dieser Selektionsbias. Man nimmt sich also Akten von Patienten, die eine PET und möglicherweise dann noch eine endgültige finale Diagnostik aufweisen, und diese Akten werden ausgewertet. Aber jeder, der schon einmal Akten ausgewertet hat, weiß, wie viele Akten nicht aufzufinden sind, in wie vielen Akten nicht genau steht, warum der Patient den Indextest und den Referenzstandard nicht bekommen hat. Jedenfalls ist dann die Ausfallquote enorm hoch. Man weiß noch nicht einmal, wie hoch sie ist, weil das ja gar nicht ermittelt werden kann. Man hat ja keine Grundgesamtheit, aus der man eine Stichprobe bezieht. Das ist nur ein Problem.

Ein ganz anderes Problem ist halt sozusagen auch der Case-Mix. Man weiß gar nicht, welche Patienten letztendlich berücksichtigt werden und welche Confounder da eine Rolle gespielt haben könnten und so weiter.

Wir sind also der Auffassung, dass eine methodische Mindestqualität, um eine diagnostische Güte einigermaßen vernünftig darstellen zu können, eine prospektiv geplante Diagnose- oder auch Prognosestudie sein sollte. Wir dachten jetzt, dass Sie in Ihrer Stellungnahme auf ganz spezifische Studien ausgerichtet sind, die Sie in Vorabrecherchen oder einfach durch Ihre Expertise bereits gekannt haben, und dass Sie bezüglich dieser Studien die Sorge haben, dass wir sie nicht berücksichtigen könnten.

**Moderator Jürgen Windeler:** So habe ich jetzt auch die Frage verstanden. Aber ich habe Ihre Antwort auch so verstanden, dass so spezifisch dieser Hinweis nicht war. Es war also jetzt nicht so, dass Sie spezielle Studien im Sinn hatten oder kennen, wo Sie die Sorge haben, dass die jetzt hier bei der Recherche nicht berücksichtigt werden.

**Dirk Hellwig:** Das, was wir im Hinterkopf haben, ist: Es gibt durchaus bereits Meta-Analysen zur diagnostischen Testgenauigkeit der PET-Untersuchungen im Zusammenhang mit der Alzheimererkrankung. Wenn ich mir die Meta-Analysen anschau und im Hinterkopf habe, nach welcher Methodik Sie arbeiten, dann ist mir von vornherein klar, dass viele der in diesen Meta-Analysen identifizierten Studien alleine aufgrund des Kriteriums der retrospektiven Studiendurchführung nicht für Sie akzeptabel sind. Deswegen kann man eigentlich sehr einfach auf aktuelle Meta-Analysen verweisen. Da gibt es durchaus in diesem Jahr schon Publikationen. Dann können Sie das durchsehen und finden da eigentlich eine gute Zusammenstellung dieser Literatur.

Ein Punkt, den wir auch im Hinterkopf hatten, als wir das in unserer Stellungnahme angegeben haben, ist das Problem, dass Sie ein Studiendesign fordern, in dem gleichzeitig zum einen die PET-Diagnostik untersucht wird, zum anderen an demselben Patientenkollektiv ein weiteres diagnostisches Verfahren durchgeführt wird und dann noch das Referenzverfahren herangezogen wird.

Das Vorgehen unterscheidet sich auch von dem bei sonstigen Meta-Analysen. Wir sind uns einig, dass es eine Reihe von Einflussmöglichkeiten, Confoundern gibt, welche die Aussagekraft von Meta-Analysen beeinflussen. Aber es macht ja Sinn, den verfügbaren Kenntnisstand zusammenzutragen. Nach meiner Auffassung ist es wirklich sehr gut geworden, wie methodisch in Ihrem Hause die quantitativen Informationssynthesen gemacht werden. Da haben alle etwas von. Aber wenn wir uns jetzt von vornherein auf derartige Diagnosestudien beschränken, in denen zwei Diagnoseverfahren an demselben Patientenkollektiv mit einem weiteren Referenztest verglichen werden, führt das wiederum zu einer sehr starken Einschränkung. Wir sind uns darüber einig, dass es wünschenswert ist, dieses Design zu haben, aber im Hinterkopf muss man sicherlich diese Confounder haben. Und dann kann man das durchaus verwerten. Man muss beispielsweise dann auch schauen,

wie die diagnostische Treffsicherheit nicht nur der PET-Untersuchung aussieht, sondern man sollte auch zusammentragen – es soll ja PET/CT bewertet werden –, wie die diagnostische Testgenauigkeit zum Beispiel von CT-Untersuchungen ist, auch wenn CT bei Alzheimer nicht wirklich das Standardverfahren ist, oder Morphometrie im Kernspin oder sonstige Biomarker. Das sollte sich durchaus wiederfinden, damit man eben auch einen Vergleich durchführen kann.

Wenn ich keinerlei Informationen zur diagnostischen Testgenauigkeit habe, weil ich so strenge Selektionskriterien für die akzeptablen Studien verwende, dann kann ich nachher gar nicht sagen, ob dieses Verfahren im Vergleich zu sonstigen diagnostischen Tests von Potenzial ist oder nicht.

**Fülöp Scheibler:** Herr Hellwig, ich habe einen Vorschlag zum Vorgehen. Das ist nämlich jetzt ein weiterer Punkt, den Sie ansprechen. Ich schlage vor, diesen Punkt unter „Verschiedenes“ ausführlich zu diskutieren. Direkte Vergleiche sind ein wichtiges Thema. Das haben wir ja auch schon öfters erörtert im Zusammenhang mit anderen Berichten. Ich schlage vor, dass wir dieses Thema nach hinten verschieben und unter „Verschiedenes“ aufnehmen. Denn ich würde gerne diesen ersten Tagesordnungspunkt mit den retrospektiven Studien abschließen.

Sie hatten am Anfang Ihres letzten Statements erwähnt, dass es aktuelle Meta-Analysen gibt, die sowohl prospektive als auch retrospektive Studien einschließen, und dass damit der IQWiG-Bericht nur schwer vergleichbar wäre, wenn er sich ausschließlich auf prospektive Studien konzentriert. Genau das sehe ich nicht so. Das Problem ist generell: In der PET-Literatur gibt es sehr viele systematische Übersichten, Meta-Analysen. Die unterscheiden sich natürlich durchaus von den systematischen Übersichten, Meta-Analysen des IQWiG. Beispielsweise gibt es auch systematische Übersichten, die läsionsbasierte Auswertungen berücksichtigen. Das tun wir auch nicht. Es gibt halt viele methodische Mindestkriterien, wo wir einfach sagen, wenn man daruntergeht, dann ist es nicht mehr diagnostische Güte, sondern Kaffeesatzleserei.

In unseren Berichten im Diskussionsteil findet sich immer der Vergleich unserer Ergebnisse mit den Ergebnissen aktueller systematischer Übersichten und Meta-Analysen. Das heißt, diese Sensitivitätsanalyse, die Sie gerne hätten, haben Sie ja, indem Sie unsere Ergebnisse mit den Ergebnissen der Meta-Analysen, die auch retrospektive Studien berücksichtigen, vergleichen. Wir diskutieren das eigentlich sehr ausführlich, dass wir sagen: Das ist unser Ergebnis. Das ist das Ergebnis der Meta-Analyse von XY. Manchmal unterscheiden die sich gar nicht. Dann ist es auch egal, ob da retrospektive oder prospektive Studien drin sind. Und wenn sich die Ergebnisse unterscheiden, dann muss man detaillierter gucken und sagen: Wie kann das denn sein? Sind da Studien eingeschlossen, die wir nicht haben? Liegt es womöglich genau an diesen Studien, dass das Ergebnis so unterschiedlich ist? Sprich: Wenn sich herausstellen würde, dass eine Meta-Analyse, die prospektive und retrospektive Studien einschließt, zu einem ganz anderen Ergebnis kommt als der IQWiG-Bericht, und man drösel

die dann auf und stellt fest, bezüglich der prospektiven Studien gibt es gar keinen Unterschied zwischen IQWiG-Bericht und der anderen Meta-Analyse, allein die retrospektiven Studien führen dann zu dem deutlichen Unterschied, dann ist es offensichtlich, dass dieses methodische Charakteristikum zu diesen unterschiedlichen Ergebnissen führt. Das wäre ja die Bestätigung dessen, was wir sagen, dass diese methodischen Unterschiede die Ergebnisse stark beeinflussen, warum wir sie gerne nicht berücksichtigen würden.

**Alexander Drzezga:** Ich möchte mich dazu gerne äußern. Zunächst würde ich sagen, generell in der Wissenschaft zu sagen, retrospektive Studien stellen Kaffeersatzleserei dar, das kann man sicherlich nicht aufrechterhalten.

**Fülöp Scheibler:** Das habe ich auch nicht gesagt.

**Alexander Drzezga:** Das ist durchaus ein Verfahren, das man einschließen kann.

Ein Punkt, der ganz wichtig ist und der den Bereich der Alzheimerbildung stark unterscheidet von der Tumorbildgebung, sind die extrem langen Follow-up-Zeiträume. Das ist natürlich ein Punkt, der es leichter macht, größere Fallzahlen, unterschiedliche Populationen retrospektiv zu analysieren. Sie könnten also durchaus prospektiv und retrospektiv Unterschiede haben, einfach aufgrund der Möglichkeiten, über 15 Jahre jemanden verlaufszukontrollieren. Das ist ein Punkt, den ich in dem Zusammenhang erwähnen wollte, dass es unter Umständen gerade für die Alzheimererkrankung ein wichtiger Punkt sein könnte, auch dieses Tor zur Retrospektivität zu öffnen und das nicht von vornherein auszuschließen.

**Fülöp Scheibler:** Da würde ich gerne zurückfragen. Sie sprachen von langen Verlaufszeiten. Gerade das macht ja die retrospektiven Studien aus meiner Sicht so schwierig. Ich habe dann 15 Jahre später ein Ergebnis. Weiß ich denn, wie viele Leute zu Beginn der Studie eingeschlossen wurden und wie viele im Verlauf dieser Beobachtung weggefallen sind? Das kann man identifizieren?

**Alexander Drzezga:** Das kommt natürlich auf das Studiendesign an. Ich würde gar nicht sagen, dass man nicht im Prinzip ein Qualitätsstandard definieren sollte auch für retrospektive Studien, der erfüllt sein muss. Aber ich würde nur das Tor nach hinten hin grundsätzlich nicht schließen wollen. Das ist gerade für den Bereich der Alzheimerforschung vielleicht auch insgesamt und nicht nur der Bildgebung ein Ansatz, der problematisch sein könnte und der unter Umständen die Bildgebung in jeglicher Richtung – das muss man jetzt gar nicht werten – mehr oder weniger positiv beeinflussen könnte. Nur in dem Bereich der Alzheimerforschung würde ich dieses Tor nicht schließen wollen. Das wäre aus meiner Sicht wissenschaftlich eine Limitation mehr als ein Qualitätszusatz.

**Fülöp Scheibler:** Dann würde ich doch gerne auf die eingangs formulierte Bitte zurückkommen: Können Sie uns in den folgenden Wochen ein Beispiel zusenden, damit wir uns das angucken können? Sie haben natürlich vollkommen recht: Grundsätzlich gibt es



durchaus für bestimmte Fragestellungen retrospektive Studien, die durchaus Sinn machen. Ich bin bisher von dem Standpunkt ausgegangen, dass es für die Fragestellung der diagnostischen Güte schwierig ist, mit retrospektiven Studien zu arbeiten. Ich lasse mich aber gerne eines Besseren belehren. Wenn Sie mir da ein Beispiel geben könnten – dann kann ich mir ja vorstellen, dass es vielleicht noch mehrere davon gibt –, dann können wir uns das in den nächsten Wochen angucken und das entsprechend würdigen.

**Alexander Drzezga:** Okay.

**Moderator Jürgen Windeler:** Das ist ein guter Vorschlag. Ich würde auch vom Grundsatz her sagen, dass es wie bei bildgebenden Verfahren natürlich gute Bilder, Performance, Interpretationen gibt und schlechte. Und bei Studien gibt es auch gute und schlechte. Und retrospektive können von den Qualitätsanforderungen nicht das erreichen, was prospektive erreichen können. – Wenn alle nicken und wir aufschreiben, dass alle nicken, dann ist das schon einmal eine wichtige Äußerung.

Und dann kann man sich Gedanken darüber machen, was Mindeststandards sind. Das tun wir. Und jetzt kann man sich Gedanken darüber machen, ob retrospektive Studien wider die Erwartungen, die man üblicherweise hat, in dieser speziellen Indikation bestimmte Qualitätsmindestanforderungen auch erfüllen können. Das können wir noch einmal prüfen. Wenn Sie uns da helfen, umso besser. Und dann gucken wir, was wir daraus machen können.

Gut, dann schließen wir TOP 1 ab.

Ich rufe auf:

#### **4.4.3 Tagesordnungspunkt 2: Gemeinsame Bewertung von PET und Radiopharmaka (widerspricht dies dem Auftrag des G-BA?)**

**Fülöp Scheibler:** Die DGN weist in ihrer Stellungnahme darauf hin, dass eine Bewertung der PET bzw. PET/CT gemeinsam mit entsprechenden Radiopharmaka möglicherweise über den Beratungsantrag des G-BA hinausginge. Wir sind nicht ganz sicher, ob wir diesen Hinweis richtig verstehen. Wir bitten Sie daher, ihn etwas näher zu erläutern.

**Dirk Hellwig:** Das kann ich gerne versuchen. Das ist natürlich eine komplizierte Frage, auf die wir als Mediziner sicherlich nicht so fundiert antworten können wie Juristen, die sich wohl schon an anderer Stelle damit befasst haben.

Für uns ist in dem ganzen Verfahren immer noch nicht verständlich, warum man überhaupt PET betrachtet. Ursprünglich war der Antrag ergangen, um PET gegenüber der Szintigraphie abzugrenzen. Da haben wir in unserer Stellungnahme ein klassisches Paper zitiert, welches diesen Methodenvergleich zwischen der Szintigraphie bzw. der Einzelphotonencomputer-tomographie, sprich SPECT, gegenüber der Positronenemissionstomographie durchführt. Da ist wirklich für jeden Nichtnuklearmediziner direkt klar, dass das methodisch messtechnisch

ein Quantensprung ist. Wenn man eigentlich den Methodenvergleich durchführen möchte zwischen der szintigrafischen Abbildungsmethodik, sprich SPECT, und der PET-Diagnostik, dann habe ich ernsthafte Zweifel, ob man einen derartigen Methodenvergleich durch irgendeine Strahlenschutzkommission oder Ethikkommission bringen würde, was ja eigentlich der Grundgedanke wäre, dass man eine klinische Studie durchführen müsste, in der man prospektiv untersucht, ob die diagnostische Treffsicherheit bei Verwendung der PET- oder der SPECT-Diagnostik besser wird oder nicht. Deswegen ist an der Stelle das Ganze sehr schwierig. Wir haben dieses Paper in unserer Stellungnahme zitiert. Daraus wird eigentlich sehr schnell klar, dass die PET das messtechnisch ausgereifere Verfahren ist.

**Stefan Sauerland:** Sie haben jetzt sehr stark auf den Vergleich PET versus SPECT abgehoben. Unsere Frage bezieht sich aber eher auf die Radiopharmaka. Sie hatten in Ihrer Stellungnahme darauf hingewiesen, es gäbe kurz vor der arzneimittelrechtlichen Zulassung stehende neue Tracer, die speziell für die Darstellung der Amyloidlast entwickelt wurden. Das ist eher das Argument, das uns interessieren würde.

**Dirk Hellwig:** Dazu kann man Folgendes sagen: Es hat in den letzten Wochen die Zulassung eines Amyloid darstellenden PET-Radiopharmakons in den USA gegeben. Da ist also das FDA-Approval erreicht worden. Es hat eine Einreichung bei der EMA, bei der Europäischen Arzneimittelbehörde, gegeben, um die Zulassung für diesen PET-Tracer zu erreichen.

Zum anderen gibt es bereits in Deutschland zugelassene PET-Radiopharmaka, zum Beispiel das Fluor-18-markierte Fluorodopa, mit dem man eine Differenzierung zwischen einer Demenz vom Lewy-Body-Typ gegenüber anderen Demenzformen durchführen kann. Das sind im Prinzip zugelassene Radiopharmaka bzw. Präparate, die kurz vor der Zulassung stehen. Es ist wirklich die Frage, inwieweit juristisch, formal gesehen, diese gemeinsame Betrachtung zulässig ist. Das ist was für Juristen.

**Fülöp Scheibler:** Soweit ich die Diskussion im G-BA mitverfolgen durfte – wir waren ja nicht immer anwesend –, war ganz am Anfang dieses Thema auf der Tagesordnung. Da hieß es, dass wir grundsätzlich die PET in allen beauftragten Erkrankungen und Indikationen bewerten sollen, ganz unabhängig vom eingesetzten Tracer. Wir sollen also so tun, als wären alle Tracer in Deutschland zugelassen, und sollen jetzt nicht bestimmte Tracer ausschließen, weil sie noch nicht in Deutschland zugelassen sind. Ich kann ja die PET in einer Studie nur gemeinsam mit irgendeinem Tracer durchführen und letztlich validieren. Es bleibt mir methodisch gar nichts anderes übrig. Deswegen haben wir jetzt darüber philosophiert, wie Sie die Frage wohl gemeint haben könnten. Sobald wir nämlich eine diagnostische Studie einschließen und auswerten, dann ist es immer die PET oder PET/CT, also eine bestimmte Technologie in Kombination mit einem bestimmten Tracer im Vergleich zu einem bestimmten Referenzstandard. Und daraus errechnen wir dann ganz einfach unsere Vierfeldertafel. Aber diese drei Komponenten müssen wir haben, sonst können wir nichts berechnen. Deswegen war ich mir nicht so sicher, wie ich Sie da verstehen soll.

**Dirk Hellwig:** An der Stelle sind wir uns einig. Es muss quasi eine separate Auswertung, Analyse für die verschiedenen verfügbaren oder im Zulassungsprozess befindlichen Radiopharmaka durchgeführt werden.

**Moderator Jürgen Windeler:** Das ist aber doch keine juristische Frage, sondern eine medizinische Frage. Die juristische Frage beschäftigt uns erst einmal nicht so sehr. Die wichtige Frage wäre, ob Sie aus medizinischer Sicht sagen, sie müssen – so habe ich Sie gerade verstanden – getrennt ausgewertet werden. Man kann nicht aus medizinisch-fachlichen Gründen alle Radiopharmaka, alle Tracer in einen großen Topf schmeißen. Deswegen müssen die getrennt betrachtet werden. Die Antwort würde uns im Moment interessieren. Die juristische Frage, speziell was Zulassung, Nichtzulassung angeht, können wir getrennt klären. Das ist nicht unser spezieller Fokus.

**Fülöp Scheibler:** Noch einmal anders gefragt: Wenn wir sämtliche Studien zur PET bei Demenzerkrankungen, die jetzt den Einschlusskriterien entsprechen, berücksichtigen und auswerten würden, käme am Ende wahrscheinlich eine ganze Palette an verschiedenen Radiopharmaka und auch an verschiedenen Technologien heraus. Wenn wir die alle darstellen würden, ist das das, was Sie von uns wollen, oder ist es genau nicht das, was Sie von uns wollen? Ich verstehe es wirklich inhaltlich nicht ganz. Ich würde so vorgehen: Ich würde das, was da ist, deskriptiv darstellen. Und die Frage, ob dann bestimmte Radiopharmaka in Deutschland aktuell zugelassen sind, ist ja eine nachgeordnete Frage, die dann der G-BA zu beantworten hat. Die würde ich nicht im Vorfeld schon in der Nutzenbewertung ausschließen, einschließen usw.

**Dirk Hellwig:** Ja, das kann gut sein. Schlussendlich muss man eine ganz breite Palette an verschiedenen Indikationen mit verschiedenen Radiopharmaka betrachten.

**Fülöp Scheibler:** Gut. Das werden wir tun.

**Moderator Jürgen Windeler:** Dann haben wir das geklärt. Wir schließen dann diesen Tagesordnungspunkt, bevor die Unklarheit möglicherweise wieder steigt.

Ich rufe auf:

#### **4.4.4 Tagesordnungspunkt 3: Einschluss von Patientinnen ohne Risiko, Verdacht oder Symptome (sollen Screeningstudien berücksichtigt werden?)**

**Fülöp Scheibler:** Die Firma Lilly weist in ihrer Stellungnahme darauf hin, dass die Einschränkung auf Studien zu Patienten mit erhöhtem Risiko bzw. Verdacht auf eine Alzheimer-Demenz und Patienten mit einer bereits diagnostizierten Alzheimer-Demenz nach ihrer Ansicht zu kurz greife. Sie fragt nach, ob Patienten mit Mild Cognitive Impairment eingeschlossen würden. Laut Berichtsplan werden Patienten mit Mild Cognitive Impairment explizit als Risikopopulation für Alzheimer-Demenz betrachtet und daher eingeschlossen.

Wir würden nun gerne von Ihnen wissen, ob diese Definition ausreicht oder ob Sie darüber hinaus auch den Einsatz der PET als Screeninginstrument, also bei tatsächlich gesunden, symptomfreien und auch risikofreien Patienten, von uns untersucht haben wollen.

**Frank Boess:** Ich möchte erst sicherstellen, dass ich Sie richtig verstanden habe. Patienten mit leichter kognitiver Einschränkung im Mild Cognitive Impairment sind im Auftrag erwähnt und werden auch in Ihrer Analyse eingeschlossen?

**Fülöp Scheibler:** Ja.

**Frank Boess:** Des Weiteren: Wie steht es mit Populationen, bei der eine Ausschlussdiagnostik von Alzheimer von Vorteil wäre, also zum Beispiel von Patienten mit Verdacht auf Lewy-Body-Demenz, Frontallappendemenz oder anderen Demenzformen? Sind diese auch in Ihrer Analyse berücksichtigt?

**Stefan Sauerland:** Ja, denn wenn ich bei einer demenziellen Erkrankung mir nicht darüber im Klaren bin, welche Erkrankung zugrunde liegt, und ich die Alzheimer-Demenz als eine Differenzialdiagnose ausschließen möchte, selbst wenn der primäre Verdacht einer Lewy-Body-Demenz besteht, dann ist das eine Untersuchung, die auch dazu dient, die Alzheimer-Demenz abzugrenzen. Das wäre vom Auftrag mit umfasst.

**Frank Boess:** Ich habe das jetzt so verstanden, dass alle Patienten mit Verdacht auf Demenz mit in Ihrer Analyse sein werden.

Des Weiteren wären für mich als konkrete Population alte Patienten mit Depression, die sich auch in kognitiver Beeinträchtigung äußern kann. Wären solche Populationen auch abgedeckt durch den Auftrag?

**Fülöp Scheibler:** Wir würden nicht grundsätzlich alle älteren Patienten mit Depression einschließen, lediglich wenn die Studienautoren der Auffassung sind, dass diese Depression ein Risikofaktor für eine Alzheimer-Demenz sein könnte. Dann ist es ja quasi wieder eine Population mit Risiko auf Alzheimer Demenz. Dann ist die Studie eingeschlossen. Aber es ist ja denkbar, dass es Studien gibt zu älteren Patienten mit leicht depressiven Symptomen. Die würden wir jetzt nicht automatisch dazu zählen, dass da ein Verdacht auf Alzheimer-Demenz besteht.

**Frank Boess:** Vielen Dank für die Klarstellung.

Jetzt zu Ihrer Frage, ob wir es als sinnvoll erachten, zum Beispiel junge Patienten ohne irgendwelche Anzeichen von Demenz, in denen PET-Diagnostik durchgeführt wurde, in Ihre Analyse mit einzubeziehen. Insoweit als es Studien betrifft, in denen die Sensitivität und Spezifität der PET-Analyse zur Detektion von Amyloidplaques betrachtet würde, würden wir einen Einschluss solcher Studien für sinnvoll halten.

**Fülöp Scheibler:** Noch einmal die Rückfrage: Wären dann die eingeschlossenen Patienten gesunde Patienten ohne jegliches Risiko, ohne jegliche Symptome?

**Frank Boess:** Wenn es um die Frage geht, Beurteilung der Spezifität und Sensitivität der Diagnostik zur Amyloiddarstellung: Ja.

**Stefan Sauerland:** Aber der reine Nachweis von Amyloid ist jetzt kein Berichtssinn. Das ist Ihnen schon klar, dass wir es hier mit einer Untersuchungsmodalität zu tun haben, die darauf abzielt, eine Erkrankung festzustellen und nicht ein pathophysiologisches Korrelat?

**Frank Boess:** Ich stimme Ihnen voll zu. Aber zur Beurteilung des pathophysiologischen Korrelats, das dann in die Diagnostik einfließt, und zur Beurteilung der Sensitivität und Spezifität eines Tests, um dieses Korrelat zu finden, wäre das gerechtfertigt, meiner Meinung nach.

**Martin Haupt:** Ich halte die Stichprobendiskussion für ganz wichtig. Meine Frage wäre hierbei nach der Unterteilung des sogenannten MCI, also der leichten kognitiven Störung. Wir haben jetzt aus dem letzten Jahr neuere Publikationen zu dem Subjective Memory Impairment. Und es zeigt sich, dass sie – ob man sie Patienten oder noch Personen oder Probanden, Teilnehmer an Studien nennt – innerhalb eines Zeitraums von ein bis zwei Jahren in einem beträchtlichen Prozentsatz schon in eine klare amnestische MCI-Situation und in eine prodromale Situation für Alzheimer-Demenz-Symptomatik geraten. Möchten Sie auch gezielt solche Personen mit einbeziehen, da sie wirklich mehr oder weniger an der Schwelle stehen, um ebenfalls in diesem Pfad gehen zu müssen, Subjective Memory Impairment?

**Fülöp Scheibler:** Ich hoffe, ich spreche jetzt wirklich für die Projektgruppe. Grundsätzlich würde ich sagen, dass, wenn das Ziel, das von den Autoren, von den Studienleitern definierte Ziel darin besteht, eine Subjective Memory Impairment, die mit irgendeinem Instrument erfasst wurde zum Studieneingang, um dann nachzuweisen, dass dieses Subjective Memory Impairment dazu führt, dass später eine klinisch nachweisbare Alzheimer-Demenz auftritt, dann ist das für mich wieder eine Risikopopulation, mit einem spezifischen Instrument erfasst. Dann würde ich sagen, diese Studie gehört eingeschlossen.

**Martin Haupt:** Sie kennen die Publikationen vom Deutschen Zentrum für neurodegenerative Erkrankungen hierzu von Herrn Frank Jessen, 2010 publiziert?

**Fülöp Scheibler:** Kenne ich noch nicht, aber ich werde sie mir besorgen.

**Stefan Sauerland:** Ich glaube, das Problem, das Sie beschreiben, ist relativ theoretisch. Ich fände es erstaunlich, dass dann in einer solchen Studie Patienten mit einer rein subjektiven Beschreibung von Gedächtnisstörungen einer Diagnostik zugeführt und in eine Studie eingeschlossen würden, ohne dass das in irgendeiner Weise objektiviert würde. Ich würde es für klinische Routine halten, dass dann dieses Subjective Memory Impairment in irgendeiner

Weise durch ärztliche Untersuchung konkretisiert wird. Insofern ist die Frage doch sehr theoretisch.

**Martin Haupt:** Dem würde ich insofern widersprechen, als die Kriterien für dieses Subjective Memory Impairment nicht einfach nur sind, dass jemand sagt: „Ich glaube, mit meinem Gedächtnis ist etwas nicht in Ordnung“, sondern es kommt auf eine Kombination an von „Ja, meine Leistungsfähigkeit des Gedächtnisses hat nachgelassen“ und „Ich bin in Sorge darüber“. Dieser Sorgenaspekt ist offenbar im prospektiven Zeitraum von ein bis zwei Jahren der entscheidende daran, dass der Übertritt in dann auch neuropsychologisch festzustellende oder erhebbare Einbußen hineingerät. Insofern ist es für uns jedenfalls vom Klinischen her eine wichtige Population, die wir gegen Depression abgrenzen müssen, damit wir nicht in eine Situation geraten, wo wir jemanden übersehen, der kognitive Störungen beklagt, sie neuropsychologisch nicht hat, aber dennoch Sorgen äußert und dahinter ja eine Depression stecken könnte, diejenigen aber auch abzugrenzen von denen, die sich nicht mit der Kummernis oder der Sorge über diese Einbußen äußern und nicht in dieser Engmaschigkeit oder Systematik im weiteren Verlauf betreut werden müssen. Wenn da mit weiterer Diagnostik festzustellen wäre, dass man gezieltere Muster entdecken kann, die für eine spätere Alzheimer-Demenz sprechen können, würde ich das für sehr wichtig halten, und nicht theoretisch.

**Robert Wolff:** Es ist tatsächlich relativ schwierig, sich eine Studie auszumalen. Da würde ich gerne eine solche Studie vor mir haben. Aber ich würde mir auch das Paper noch einmal konkret angucken. Aber prinzipiell sind die Einschlusskriterien, die wir gewählt haben, relativ breit. Wir würden uns, was die Frage der Mild Cognitive Impairment angeht, insbesondere an diese McKhann-Kriterien anlehnen. Seine Arbeitsgruppe hat ja im letzten Jahr – das ist das Update von diesem NINCDS-ADRDA ... Da gibt es auch die Anhänger, die sich mit MCI beschäftigen. Und daran lehnen wir uns insbesondere an.

**Martin Haupt:** Ich kann es im Moment nicht mit Bestimmtheit sagen. Aber ich würde Sie bitten, noch einmal zu überprüfen, ob auch in diesen Kriterien solche Personen mit erfasst würden, die noch keine definitiven objektivierbaren Leistungsabfälle in standardisierten neuropsychologischen Tests zeigen, aber schon mit einer nachhaltigen Klage über Gedächtniseinbußen aufgekommen sind. Das würde mich interessieren. Das kann ich im Moment nicht definitiv sagen, von der 2011er-Publikation.

**Robert Wolff:** Da schauen wir.

**Moderator Jürgen Windeler:** Wenn ich das jetzt richtig mitbekommen habe, können wir das doch zusammenfassen, indem wir sagen, dass Studien, in denen ein Anlass für eine alzheimerorientierte Diagnostik bestand, eingeschlossen werden, dass die Art des Anlasses beschrieben wird – das würde auch dem entsprechen, was Sie gesagt haben –, dass – umgekehrt – Studien, bei denen kein Anlass besteht, also wirklich reine Screeningstudien,

nicht eingeschlossen werden. Habe ich das so richtig mitgenommen? – Ich sehe weitreichendes Nicken und keine gegenteilige Wortmeldung.

**Alexander Drzezga:** Ich muss mich doch noch einmal melden, um das Verständnis zu vertiefen, auch im Sinne zum Beispiel der Amyloiddiagnostik: Man muss natürlich auch den Wert des Negativausschlusses zulassen, dass Personen einschließbar sind, die nicht an der Erkrankung leiden.

**Moderator Jürgen Windeler:** Das ist klar. Aber dass ein Anlass zu der Diagnostik besteht, das ist ja der Punkt.

**Alexander Drzezga:** Angenommen, man macht eine histopathologische Post-mortem-Studie und untersucht, was Patienten in Hospizen passiert bezüglich der Wertigkeit der Amyloiduntersuchung, der Anlass der Diagnostik ist natürlich die Evaluierung dieses Tools und nicht in irgendeiner Form eine klinische Indikation. Es werden Leute eingeschlossen, die keinen Hinweis auf eine Demenzerkrankung haben. Der Anlass ist dann die Wissenschaft in dem Sinne. Sind diese Personen dann aus Ihrer Sicht einschließbar oder würden Sie dann diese Ergebnisse ausschließen wollen?

**Fülöp Scheibler:** Bei den Populationen sagen wir ganz klar: Patienten mit unbekanntem Krankheitsstatus, aber mit einem Verdacht auf Alzheimer Demenz. Wenn ich Sie richtig verstanden habe, diese älteren Personen in Hospizen. Sofern das jetzt eine Population ist, die ausschließlich aus Patienten ohne Verdacht oder ohne Anlass zu einer Alzheimerdiagnostik wäre, dann müsste man Herrn Hellwig gleich fragen, ob das ethisch vertretbar, denn überhaupt strahlenschutztechnisch vertretbar wäre, ältere Patienten ohne Verdacht auf Alzheimer einer PET zu unterziehen. Aber abgesehen davon würden wir diese Studie nicht einschließen. Genauso wenig würden wir Studien einschließen bei einer Population von jungen Medizinstudenten, weil wir davon ausgehen würden, die sind zu 99,9 % nicht von Alzheimer betroffen. Das wäre dann tatsächlich eine Screeningstudie.

**Stefan Sauerland:** Ich möchte kurz ergänzen: Das würde dann ja in der klinischen Implikation bedeuten, wenn aus so einer Studie etwas herauskäme, dass man tatsächlich eine Art Screening im Altersheim mittels der PET durchführen würde. Das ist nicht Ziel des Auftrags und nicht Ziel der Auftraggeber.

**Alexander Drzezga:** Nein, da verstehen Sie mich falsch. Es geht nicht um das Screening von präfinalen Patienten im Altersheim, sondern es geht um die Evaluierung des tatsächlichen Nutzens der Methode. Zeigt das Bildgebungssignal tatsächlich das, was wir im Gehirn sehen wollen? Und das kann man natürlich nur durch Einschluss von Patienten, wie in den Studien passiert, die praktisch kurz vor dem Ableben sind, um das so zu formulieren, um dann an das Gehirnmaterial heranzukommen. Das ist die spezifische Problematik der Hirnbildgebungsstudien. Da können Sie sich natürlich nicht beschränken auf Patienten, die

aus klinischer Sicht zur Abklärung einer Alzheimererkrankung untersucht worden sind, denn darum geht es nicht. Es geht hier um die Evaluierung der Qualität des Messwerkzeuges.

**Moderator Jürgen Windeler:** Genau.

**Mathias Schreckenberger:** Ich möchte kurz ergänzen: Wir haben eine gewisse Unschärfe bei der Definition des Begriffs „Screening“. Natürlich geht man nicht davon aus, dass die Kontrollpersonen irgendwie erkrankt sind, sondern sie sind wirklich ein reines Kontrollkollektiv, um die Güte des Messinstruments, also hier FDG-PET oder Amyloid-PET, zu erfassen, aber es ist kein Screening wie bei der Mammographie im engeren Sinne.

**Fülöp Scheibler:** Wenn ich es richtig verstehe, sind es zwei verschiedene Arten von diagnostischen Designs. Nach Köbberling ist das eine Phase-2-Studie. Das ist eine technische Voruntersuchung. Die ist natürlich absolut wichtig. Aber für unsere Fragestellung, die letztendlich auf den klinischen Nutzen bei einer bestimmten Population abzielt, liefert so eine technische Voruntersuchung keine validen Ergebnisse, weil die Population von vornherein schon festgelegt ist, genauso wie die jungen Medizinstudenten. Auch da ist es so. Sie werden natürlich durchaus feststellen, ob Ihr Verfahren technisch funktioniert. Aber wir wollen ja eine Sensitivität, eine Spezifität in einer Population mit unbekanntem Krankheitsstatus, aber Risiko für die Erkrankung. Und das kann man in der Population nicht machen.

**Alexander Drzezga:** Ich muss Sie kurz unterbrechen. Wenn Sie davon ausgehen, dass 20 % von über 80-Jährigen per se statistisch schon an einer Demenz leiden, dann können Sie einfach, wenn Sie einen Schnitt durch die Bevölkerung von randomisiert ausgewählten älteren Personen machen, davon ausgehen, dass ein bestimmter Anteil davon an dieser Erkrankung leidet. Sie haben durchaus ein gemischtes Kollektiv, was jetzt nicht spezifisch nur Gesunde beinhaltet, sondern das ist eine randomisierte Untersuchung einer Per-se-Risikogruppe, weil im Grunde ältere Personen an sich schon eine Risikopopulation darstellen.

**Stefan Sauerland:** Kurz zur Begriffsklarstellung: Es ist keine randomisierte Studie, es ist eine Zufallsstichprobe.

**Alexander Drzezga:** Ja. Da bin ich statistisch nicht ...

**Dirk Hellwig:** Es ist eine Querschnittsstudie, um das richtig einzuordnen.

**Fülöp Scheibler:** Das auf jeden Fall.

**Dirk Hellwig:** Ich würde das noch nicht einmal als eine technische Voruntersuchung bezeichnen, sondern das ist wirklich dazu gemacht, um den Vorhersagewert zu sehen, um eben auch die Prävalenz zu bestimmen von Amyloiddarstellungen in dieser Patientengruppe, in diesem Altenkollektiv. Was wir haben wollen, ist eine Information darüber, wie hoch der Ausschlusswert dieser Amyloidbildung ist. Und dafür sind diese Studien sehr gut. Wir



brauchen diese Studien wirklich, um den verfügbaren Wissensstand darzustellen. Deswegen kann ich nur dafür plädieren, diese Studien in Ihrer Analyse zu berücksichtigen.

**Fülöp Scheibler:** Vielleicht hängt es tatsächlich von der Definition der Population ab. Querschnittsstudie, absolut d'accord. Diagnosestudien sind ja zumeist Querschnittsstudien. Aber es gibt innerhalb der Querschnittsstudien noch weitere Unterteilungen, nämlich nach der jeweiligen Population.

Da komme ich jetzt zurück auf die Frage von Herrn Schreckenberger. Was Sie, glaube ich, beschrieben haben, ist eine diagnostische Fall-Kontroll-Studie. Wenn ich jetzt zum Beispiel 20 Patienten mit bereits klinisch nachgewiesener Alzheimer-Demenz vergleiche mit zehn anderen Patienten, die definitiv keine Alzheimer-Demenz haben – das sind auch Voruntersuchungen. Aber auch aus so einer Studie werden Sie, weil ja die Population eine künstliche ist, Ergebnisse zur diagnostischen Güte bekommen. Die werden dann auch künstlich sein. Das heißt, die sind nicht übertragbar, auf keine tatsächliche Population von Patienten. Wir haben vorne im Berichtsplan die Einteilung, wo wir sagen, wir wollen Phase 3 nach Köbberling. Das müssen also Patienten mit unbekanntem Krankheitsstatus sein. Ansonsten bedeutet das, was wir zur diagnostischen Güte bekommen, einfach nur: Die PET funktioniert technisch. – Ich glaube, davon müssen wir den G-BA nicht mehr überzeugen. Die Frage ist: Wie gut funktioniert sie in einer ganz bestimmten Population? Und diese Frage kann man wirklich nur beantworten, wenn es eine Population ist, in der später klinisch diese diagnostische Frage tatsächlich relevant sein könnte. Wenn das in dieser Population der Fall wäre, dann wäre natürlich die Studie eingeschlossen.

**Moderator Jürgen Windeler:** Das scheint ein ganz guter Abschluss, eine ganz gute Zusammenfassung zu sein. Gibt es weitere Wortmeldungen dazu?

**Alexander Drzezga:** Ich habe noch einen anderen Punkt zu dem Thema. Den ApoE-Genotyp habe ich in dem vorläufigen Berichtsplan nicht gefunden. Wir wissen, dass es in einem relativ hohen Prozentsatz der Bevölkerung diesen ApoE-4-positiven Genotyp gibt und dass der mit einem höheren Risiko für die Alzheimererkrankung einhergeht. Ist das für Sie jetzt ein Risikofaktor oder würden Sie sagen, das ist kein Risikofaktor?

**Robert Wolff:** Er ist indirekt enthalten. Uns ist das bewusst, dass das ein Risikofaktor ist. Und im Hintergrund haben wir genetische Vorerkrankungen erwähnt. Insbesondere in der Fußnote oder in der Bildunterschrift zur Abbildung 2 heißt es, in der Originalpublikation wird genau auf diese Gruppe verwiesen. Und wir haben das als Personen mit genetischen Vorbelastungen umschrieben, einfach weil der Hintergrund wirklich nur eine Einführung sein soll und nicht das ganze Thema abarbeiten soll. Es ist uns bewusst, aber wir wollten nicht zu weit greifen.

**Alexander Drzezga:** Okay. Aber Sie würden das als Risikofaktor bewerten in dem Sinne?

**Robert Wolff:** Ja, genau.

**Alexander Drzezga:** Ergänzend wäre aus meiner Sicht wichtig zu diskutieren: Was wollen wir diagnostisch mit dem Werkzeug erreichen? Oder: Was soll überprüft werden? Geht es primär wirklich nur um die Diagnose einer manifesten Demenzerkrankung oder Differenzialdiagnose oder auch, wie Sie es schon in den Stadien einteilen, um diese präklinischen Phasen, die früheren Phasenerkrankungen, also den Nachweis der ablaufenden Pathologie?

Wir haben dazu einen Kommentar, wo ich hoffe, dass Herr Haupt mir zustimmen wird. Man geht ja heute davon aus, dass, wenn man überhaupt therapeutisch einen Einfluss wird nehmen können auf die Erkrankung, dieses in einem relevanten Ausmaß nur möglich sein wird in einer präklinischen Phase, also wenn Patienten noch nicht dement sind. Das heißt, die Diagnose der Demenz wird in dem Sinne als Spätstadium betrachtet. Insofern wird es sicherlich diagnostisch wichtig sein, gerade diese Population zu identifizieren. Und auch aus ökonomischen Aspekten wird es sicherlich so sein, dass man in diesem Stadium unter Umständen therapeutisch wird anfangen müssen, um eine zeitlich spätere Pflegebedürftigkeit zu erreichen oder die Erkrankung nach hinten herauszuschieben. Es ist mir in Ihrem Bewertungsauftrag noch nicht klar, auf was jetzt abgezielt wird, also den Nachweis der Erkrankung der Demenz, der manifesten klinischen Erkrankung. Oder ist der Nachweis der ablaufenden Pathologie auch etwas, was in Betracht gezogen wird?

**Fülöp Scheibler:** Wenn Sie mit Pathologie nur die Plaques meinen, dann sind wir da sehr zurückhaltend. Wir haben ja den Referenzstandard so definiert, dass er durch klinische Tests noch bei lebenden Patienten nachgewiesen sein muss. Der Nachweis von Plaques alleine würde uns nicht genügen, weil uns dieses ganze Plaqueskonzept nicht ganz einfach verständlich erscheint, während eine noch zu Lebzeiten klinisch nachgewiesene Alzheimer-Demenz ein eindeutig patientenrelevantes Krankheitsgeschehen ist. Dagegen sollten Diagnostik und Therapie vorgehen. Deswegen sind wir der Meinung, das ist unser Zielkriterium für den Referenzstandard.

Allein der Nachweis eines pathologischen Vorgangs, von dem aber nicht ganz klar ist, wie er mit Alzheimer-Demenz zusammenhängt, ist aus meiner Sicht zwar für die Grundlagenforschung sehr wichtig, aber für die spezifische Fragestellung dieses IQWiG-Berichts würde das zu weit führen.

**Alexander Drzezga:** Da müssten wir jetzt vielleicht doch mal Herrn Haupt zu Wort kommen lassen.

**Martin Haupt:** Zunächst wollte ich mich der Vollständigkeit halber melden und sagen: Ja, Herr Drzezga, ich stimme Ihnen zu. So steht es auch in unserer Stellungnahme.

Ich stimme aber auch Herrn Scheibler zu. Es ist sicherlich richtig: Das Paradigma Amyloid als die Entstehungsgeschichte für die klinischen Symptome der später auftretenden Alzheimer-Demenz ist insofern brüchig, als sie viele andere Aspekte außer Acht lässt, nicht

nur entzündlich oder inflammatorisch-immunologisch, sondern auch vor allen Dingen die gesamte Taubpathologie zunächst einmal unberücksichtigt lässt, die ja nun sehr viel enger mit fortschreitendem Stadium und schließlich auch der Progression der Symptomatik der Krankheit korreliert, als die Amyloidpathologie das tut. Aber dennoch: Man kommt natürlich nicht darum herum, wenn man ein diagnostisches Verfahren überprüfen will, welches die Amyloidpathologie entdecken kann, die Stichprobe – daher komme ich mit meinem Subjective Memory Impairment daher – so auszuweiten, dass man möglichst alle in Betracht kommenden, für eine spätere Ausbildung oder schon für die Symptomträger einer gegenwärtig bestehenden klinischen Symptomatik schon wirklich entdecken kann, um sie hineinzunehmen, aber nicht so weit auszuweiten, dass man dann tatsächlich ins Feld geht. Man müsste ja dann letztlich wieder Repräsentativität, epidemiologische Fragestellungen mit hineinnehmen. Ich denke, das sprengt den Rahmen und wäre für die Fragestellung wohl auch nicht sinnvoll.

**Moderator Jürgen Windeler:** Weitere Wortmeldungen zu TOP 3? Ich würde gerne, ohne jetzt den TOP ausweiten zu wollen, auf den TOP „Verschiedenes“ die Rückfrage verschieben, welche Informationen, welche Daten es zu der Aussage „Je früher, desto besser“ gibt, insbesondere um die Pflegebedürftigkeit zu vermeiden. Es sind ja nun harte Aussagen, die ich auch teilweise in Stellungnahmen gesehen habe. Da würde mich einfach eine Rückmeldung interessieren, wo die Daten für diese Aussagen sind. Aber das können wir unter „Verschiedenes“ thematisieren. Das war für uns keine Unklarheit, aber wenn das so explizit ausgesprochen wird, dann möchte ich das gerne lernen.

Ich rufe auf:

#### **4.4.5 Tagesordnungspunkt 4: Nachbeobachtungszeit für das Referenzverfahren bei Diagnosestudien (18 Monate adäquat?)**

**Fülöp Scheibler:** Die Firma Lilly fragt in ihrer Stellungnahme nach einer Begründung für eine Mindestnachbeobachtungszeit von 18 Monaten für den Referenzstandard bei diagnostischen Fragestellungen. Sie schlägt vor, alle Studien unabhängig von ihrer Nachbeobachtungszeit einzuschließen. Das IQWiG bezieht sich im vorläufigen Berichtsplan auf die European Task Force, diesen Consensus, publiziert im Lancet Neurology 5/2008 und jüngst noch einmal bestätigt im Journal of Nutrition Health and Aging 4/2012. Die Empfehlungen für Nachbeobachtungszeiten variieren jedoch in der Literatur. Wir würden daher gerne erfahren, welche Quellen oder Kriterien oder welcher Zeitraum für eine Nachbeobachtung in Diagnosestudien aus Sicht der anwesenden Experten sinnvoll erscheinen.

**Frank Boess:** Soll ich dazu Stellung nehmen?

**Moderator Jürgen Windeler:** Wenn Sie möchten. Wir zwingen hier keinen.

**Frank Boess:** Der Hintergrund meiner Nachfrage und der Abklärung ist, dass in dem zitierten EMA-Dokument die 18 Monate für die Entwicklung neuer Disease Modifying Treatments erwähnt sind, während im gleichen Dokument zum Beispiel für Medikamente, die eine symptomatische Verbesserung erzielen, ein sechsmonatiger Untersuchungszeitraum als adäquat gesehen wird, was auch in der Vergangenheit von FDA und EMA der Standpunkt war. In klinischen Studien wird teilweise schon nach einer sechswöchigen Behandlung mit Acetylcholinesterase-Inhibitoren eine Verbesserung der Patienten in verschiedenen Wee-douts nachgewiesen. Insofern konnte ich nicht nachvollziehen, warum gerade ein 18-Monate-Zeitraum zur Verbesserung eines Patient Outcomes, um den es hier geht, ausschlaggebend sein soll und warum eventuell Studien, die einen kürzeren Zeitraum betrachtet haben, ausgeschlossen werden, weil meiner Meinung nach möglicherweise wertvolle Informationen dabei verloren gehen könnten.

**Stefan Sauerland:** Der Zeitraum 18 Monate ist immer eine Zahl, die man auch anders interpretieren könnte. Warum nicht zwölf, warum nicht 15, warum nicht 24? Wir müssen aber auch bedenken, dass wir es zum Teil hier mit Studien zu tun haben werden, wo überhaupt keine Therapie stattfindet, wo also einfach nur der Patient nachbeobachtet wird. Ich denke, die Kollegen der Nuklearmedizin wären dann sehr enttäuscht und würden am Ende unseren Bericht kritisieren, wenn wir Studien mit drei oder sechs Monaten Nachbeobachtungszeit berücksichtigen würden. Sie würden uns dann – ich denke, zu Recht – vorwerfen, dass wir solche Studien berücksichtigen, wo ja gar nichts herauskommen kann, weil der Zeitraum zu kurz ist. Das heißt, die Testgüte würde dann, weil die Nachbeobachtungszeit so kurz ist, deutlich zu gering eingeschätzt werden. Daher muss man irgendwo eine Grenze festlegen. Die Grenze 18 Monate ist nun einmal in diesen Papieren eine wesentliche Grenze. Daher erscheint uns die ganz plausibel.

**Alexander Drzezga:** Ganz grundsätzlich: Der Wert der Diagnostik, der Bildgebung, was die Alzheimererkrankung angeht, liegt ja nicht nur in der Definition eines Status, sondern eben auch in der Prädiktion. Und da ist die Geschwindigkeit der Konversion oder die Zeit zur Konversion, was zunehmend als wichtig oder interessant betrachtet wird, weil man davon ausgehen kann, vielleicht bekommt jeder Mensch irgendwann eine Alzheimererkrankung. Wie schnell es sich entwickelt, ist von Interesse. Insofern wären Verlaufsstudien schon interessant, die im Grunde ab Beginn der Baseline beginnen, nachuntersuchen und dann involvieren, wann ist welcher Patient konvertiert oder nicht konvertiert. Da würde ich es für eine Limitation halten, zu sagen, wir müssen jetzt 18 Monate abwarten, wenngleich ich Ihnen recht gebe, dass es ein relativ realistischer Zeitraum ist. Ich sehe nur eigentlich nicht, warum man es überhaupt limitieren soll. Was ist überhaupt der Grund, es zu beschränken? PET-Studien, die nach drei Monaten ein Follow-up machen und sagen, die Diagnostik funktioniert nicht, die gibt es meines Wissens nicht. Was ist die Argumentation, es überhaupt zu limitieren? Was versprechen Sie sich davon?

**Fülöp Scheibler:** Wie gesagt, es könnte ja auch Studien geben, die eine andere Technologie, zum Beispiel das MRT, mit der PET vergleichen. Man kennt es ja aus Studienplanungen, dass

man dann die Studie so plant, dass die Technologie, die man selber bevorzugt, am besten abschneidet. Es würde dann vielleicht Sinn machen – das ist jetzt alles konstruiert –, bewusst eine sehr kurze Nachbeobachtungszeit zu wählen, weil man dann davon ausgeht, dass die PET möglicherweise noch nicht so viel zeigen kann. Für uns war die Überlegung, nicht zu kurz, aber auch nicht zu lang, weil man dann natürlich sehr hohe Ausfälle und Abbruchraten hat. Das ist natürlich ein Kompromiss zwischen „Die Erkrankung muss sich in der Zeit etablieren“ und „Man muss Patienten in der Studie haben“. Wir sind da natürlich offen, wenn Sie uns andere Zahlen nennen.

**Alexander Drzezga:** Ich halte es prinzipiell für nicht erforderlich, das zu limitieren. Ich würde sagen, dass man auf die Qualität der Studien achten sollte. Sie sprechen an, verschiedene Methoden zu vergleichen. Da kann natürlich umgekehrt genauso eine PET-Studie unter Umständen von einem kürzeren oder längeren Zeitraum profitieren. Ich würde das Fenster so offen wie möglich halten. Im Grunde führen diese Limitationen eigentlich immer zum Wegfall von bestimmten Studien. Ich glaube, es führt zu einem unselektierten Wegfall, der unter Umständen das Gesamtbild eher verzerrt als ein klares Gesamtbild hergibt.

**Stefan Sauerland:** Aber gerade sagten Sie, es gäbe solche Studien vermutlich gar nicht in dem Zeitraum von unter einem Jahr.

Was ich auch noch zu bedenken geben möchte ist, dass wir ja – das hatten wir eben diskutiert – uns vermutlich eher auf Studien beziehen werden, die nicht eine manifeste Demenz als Einschlusskriterien voraussetzen, sondern wo es eben um Mild Cognitive Impairment oder Subjective Cognitive Impairment gehen wird. Dementsprechend ist dann der Zeitraum bis zur Konversion, den Sie bezeichnen, ein etwas längerer. Und da macht es dann schon Sinn, durchaus auch längere Nachbeobachtungszeiten für solche Studien zu fordern, um eine Testgüte zu erzielen, die nicht nur auf ein oder zwei Patienten beruht.

**Alexander Drzezga:** Ich würde dem überhaupt nicht widersprechen, dass die Testgüte natürlich variieren kann bezüglich der Follow-up-Zeiten. Ich würde aber schon sehen, dass eine Studie, die meinerwegen nach zwölf Monaten ordentlich durchgeführt worden ist und zu klaren Ergebnissen führt, nicht zwingend ausgeschlossen werden sollte. Das wäre der Punkt meiner Betrachtung. Ich habe jetzt keine konkrete Studie im Auge, wo ich befürchte, dass die dadurch wegfällt. Ich sehe nur den Nutzen, es eher etwas breiter zu halten und es nicht zwingend zu beschränken.

**Fülöp Scheibler:** Noch ein zweites Problem, das wir im Hinterkopf hatten bei diesen 18 Monaten, gar nicht so sehr beim Indextest, sondern vielmehr beim Referenzstandard: Wenn man sich beim Referenzstandard auf diese klinischen Tests beschränkt, dann, wenn ich es richtig verstanden habe, ist es so, man braucht immer eine zweizeitige Messung, um wirklich eine Alzheimer-Demenz diagnostizieren zu können. Da jetzt auch die Rückfrage von mir: Ist es nicht da notwendig, dass man eine bestimmte Zeitspanne zwischen dem ersten Messzeitpunkt und dem zweiten hat, damit man sagen kann, wenn innerhalb dieses Zeitraums

sich eine deutliche Verschlechterung ergibt oder die Symptome gleich bleiben, dann ist man sicher, dass es sich nicht um eine Alzheimer-Demenz handelt? Gibt es da nicht eine Mindestzeit, die notwendig ist?

**Martin Haupt:** Zunächst einmal ist im klinischen Setting eine zweizeitige Beobachtung, um die Diagnose Alzheimer-Demenz zu erstellen, nicht notwendig. Allerdings glaube ich, dass, wenn wir jetzt über die Studie sprechen, umso mehr Ihre Überlegungen zum Tragen kommen, je weiter wir vor die tatsächlich klinische Symptomatik einer Demenz gehen. Je unklarer oder heterogener die Stichprobe dann wird, umso eher müssen Sie natürlich eine Zweizeitigkeit einbauen, die im Allgemeinen auf etwa sechs Monate im neuropsychologischen Bereich gelegt wird, wahrscheinlich im strukturellen Bildgebungsbereich mindestens ein Jahr, würde ich vermuten. Aber das wissen die Kollegen von der Nuklearmedizin besser. Hier wäre vielleicht zu überlegen, ob man dann auch tatsächlich so etwas in das Studienprotokoll mit hineinnimmt, dass diejenigen, die sich präklinisch befinden, was die Demenz anbetrifft, zur Bestätigung vor Aufnahme in die Studie eine zweite neuropsychologische Testung bekommen, weil ja auch immer wieder gesehen wird, dass sich MCI-Patienten, die nicht schon eine klare anamnestische Schwerpunktsituation haben, wieder verbessern können oder überhaupt nicht progradierend sind und dann zwei Jahre danach wieder gesund bzw. symptomfrei sind. Vielleicht bietet sich das hier eher an. Bei denen, die schon ein Demenzsyndrom haben und hier die Kriterien greifen, glaube ich, würde die einmalige Untersuchung völlig ausreichen.

**Moderator Jürgen Windeler:** Nun haben wir natürlich in dem TOP 3 den Einschluss von Betroffenen in den Studien relativ breit, auch relativ weit nach vorne angelegt, sodass eine Form von Nachbeobachtungszeit, also jetzt TOP 4, von Zweizeitigkeit absehbar relevant werden kann. Und dann verstehe ich Sie richtig, dass vielleicht nicht 18 Monate, sondern vielleicht ein bisschen weniger, zwölf Monate – wir haben bei den neuropsychologischen Verfahren sechs Monate gesagt –, als Mittelwert, ohne das jetzt festlegen zu wollen, durchaus Sinn machen würde.

Auf EMA-Kriterien zu verweisen, wenn ich das noch erwähnen darf, halte ich nicht für sinnvoll, weil das natürlich Therapieeffektkriterien sind, die mit den Diagnose- und Referenzverfahren nun gar nichts zu tun haben. Hier geht es wirklich um die „eigentlich“ spontane Progredienz, die wir abbilden wollen. Da mag insbesondere in den sehr frühen Situationen eben eine relativ mittelfristigere, jeweils in Monaten bemessene Nachbeobachtungszeit sinnvoll sein. Habe ich das so richtig sortiert?

**Martin Haupt:** Ich würde Ihnen so zustimmen.

**Moderator Jürgen Windeler:** Gut. Herr Hellwig möchte widersprechen.

**Dirk Hellwig:** Ich möchte nicht widersprechen, sondern ich möchte das Ganze dadurch unterstützen, dass es eine Meta-Analyse zur diagnostischen Testgüte verschiedener

bildgebender Verfahren gibt, publiziert von Yuan im American Journal of Neuroradiology aus dem Jahr 2009, wo als Grenze teilweise sechs Monate angenommen wurden. Für die PET-Verfahren existieren Studien ab zwölf Monaten Nachbeobachtungszeit. Ich denke, das, was die Kollegen dort in der Analyse durchgeführt haben, ist auch nicht unvernünftig.

**Moderator Jürgen Windeler:** Okay. Wenn es keine Wortmeldungen mehr gibt, würde ich TOP 4 beschließen.

Ich rufe auf:

#### **4.4.6 Tagesordnungspunkt 5: Verschiedenes**

Da hatten wir schon zwei Restanten aus der bisherigen Tagesordnung, einerseits den Punkt, den Herr Scheibler verschoben hat. Machen wir den zuerst?

**Fülöp Scheibler:** Gerne.

Ich komme noch einmal auf die direkten Vergleiche zurück. Sie hatten ja schon erwähnt, dass Sie das für wichtig, aber auch für notwendig halten, auf indirekte Vergleiche zurückzugreifen, sofern die direkten Vergleiche nicht vorhanden sind. Diese Diskussion haben wir schon im Zuge der anderen Berichte mehrfach geführt. Auch hier ist wieder das Problem, es sind zum Teil sehr kleine Studien. Der indirekte Vergleich bedeutet: Sie haben eine Studie zur MRT. Die wurde in Finnland mit 40 Patienten durchgeführt vor 20 Jahren. Und die würden Sie jetzt vergleichen mit einer Studie zu PET/CT aus Deutschland mit zum Beispiel 30 weiblichen Patientinnen. Dass diese Studie nicht besonders aussagekräftig sein kann, das, glaube ich, liegt auf der Hand. Und so sehen nun einmal diese indirekten Vergleiche diagnostischer Verfahren alle aus. Und es wird halt vielfach diskutiert – nicht nur von uns, sondern auch von anderen Methodikern –, dass indirekte Vergleiche grundsätzlich, aber insbesondere im diagnostischen Bereich, äußerst problematisch und fehleranfällig sind. Deswegen sind wir wirklich der deutlichen Auffassung, dass wir diese Vergleiche nicht anstellen wollen.

**Dirk Hellwig:** Ich kann Ihre Bedenken verstehen, dass aufgrund methodischer Unzulänglichkeiten oder Einflussmöglichkeiten das nicht unbedingt wünschenswert ist, aber wir müssen doch zusehen, dass wir den verfügbaren Wissensstand zusammentragen. Wir sind uns alle darüber einig, dass es wünschenswert ist, an ein und demselben Patientenkollektiv, das genau der klinischen Fragestellung und der Versorgungswirklichkeit entspricht, das alles zu überprüfen. Aber wir müssen auch der Wahrheit ins Auge schauen, dass diese Daten nicht unbedingt verfügbar sind.

Sie stellen jetzt anekdotisch irgendwelche Patientenkollektive vor. Ich weiß nicht, ob das tatsächlich die Kollektive sind, die in den entscheidenden Studien zur Debatte stehen, oder ob das ein Konstrukt von Ihnen ist, um das alles besonders drastisch zu schildern. Wir wissen, dass es diese Einflussmöglichkeiten gibt. Es bleibt doch nichts anderes übrig, als dass man

versucht, diese verschiedenen Verfahren zu bewerten, indem man das, was an Literatur vorhanden ist, zusammenträgt und dann schlussendlich kritisch wertet.

Darin würde ich wirklich Ihre Aufgabe sehen, zu helfen, dieses Wissen zusammenzutragen, indem die Recherche sich nicht nur auf Studien beschränkt, die tatsächlich mehrere diagnostische Testverfahren in einer Studie am selben Patientenkollektiv untersuchen, auch noch zusätzlich mit einem Referenzverfahren, sondern dass da getrennte Bewertungen stattfinden. Das ist ja auch das übliche Vorgehen, wenn man jetzt in die Literatur schaut: Was gucke ich mir ansonsten für Wissen an, das man zusammengetragen hat? Auch die Publikation von Yuan aus 2009, die ich vorhin zitiert habe, macht genau so etwas, dass eben das verfügbare Wissen aus der Literatur zusammengetragen wird.

Man kann ganz schnell sagen, ich möchte unbedingt PET-Diagnostik mit einem weiteren Testverfahren haben, und dann werde ich wahrscheinlich Schwierigkeiten haben, überhaupt irgendetwas an Literatur zu finden. Und dann sieht Ihr Nutzenbericht wahrscheinlich wieder so aus, dass kein Nutzen belegt ist nach den strengen Kriterien, was dann von Dritten gerne falsch interpretiert wird. Aber schlussendlich stehen wir doch vor dem Problem, wir wollen – ich wiederhole mich – verfügbares Wissen zusammentragen, damit wir das in der weiteren Bewertung des Verfahrens vernünftig verwenden können. Ich sehe wirklich den großen Vorteil, dass Sie die Methodik an der Hand haben, um das auf ganz hohem Niveau zu machen.

**Moderator Jürgen Windeler:** Zur Einordnung: Ich sehe unsere Aufgabe nicht darin, Wissen zusammenzutragen, sondern ich sehe die Aufgabe darin, Wissen zusammenzutragen, das eine konkrete Frage beantwortet, das geeignet ist, eine bestimmte Frage zu beantworten. Bei dieser Vorabüberlegung im Berichtsplan – das sind im Grunde immer wiederkehrende TOPs, die wir angucken – ist es nicht nur sinnvoll, sondern unserer Ansicht nach geboten, sich den Aufwand zu ersparen, Wissen im Sinne von Daten, Publikationen und irgendwelchen Fakten und Zahlen zusammenzutragen, das nicht geeignet ist, die Frage zu beantworten. Das kann man sich sparen. Das war das Moment von Herrn Scheibler, das er in der Tat an einem plakativen, eindrücklichen Beispiel, an einem theoretischen Beispiel, wie er auch gesagt hat, deutlich gemacht hat.

Wenn es so ist, dass unseres Erachtens mit guten Argumenten man zwar bestimmte Informationen sammeln kann, sie aber nicht in irgendeiner Weise zusammenführen kann, um die Frage beantworten zu können, dann macht es sehr viel Sinn, auf das Sammeln zu verzichten, nicht nur aus Ressourcengründen, sondern auch aus inhaltlichen Gründen und auch um der Kritik in einem halben Jahr zu entgegenen, dass wir Äpfel mit Birnen zusammengeschmissen hätten und das alles gar nicht vergleichbar wäre. Insofern bitte ich um Verständnis für diese Form von Vorfilter. Wo immer der Filter genau sein mag, das ist immer eine Frage, über die man nachdenken kann. Aber noch einmal allgemein: Es ist nicht unsere Aufgabe, nur Wissen zusammenzustellen. Das wäre mir etwas zu allgemein.



**Stefan Sauerland:** Vielleicht kann ich das in dem Feld PET und Alzheimer konkretisieren. Sie wünschen sich im Prinzip, dass wir Studien anschauen, wo alleine die PET genommen wird, und dann wird irgendwie der Verlauf der Erkrankung oder die Möglichkeit der korrekten Diagnosefindung untersucht. Bei diesen Studien ist es doch, wenn man das Ganze klinisch anwenden will, irgendwann sinnvoll, sich anzuschauen, welche Begleitinformationen ich schon vorliegen habe. Wenn die PET mir nur das bestätigt, was ich ohnehin zum Beispiel aufgrund des klinischen Verlaufes, zum Beispiel aufgrund der genetischen Belastung, aufgrund der Familienanamnese schon vermute, dass dieser Patient ein erhöhtes Risiko hat, eine Alzheimer-Demenz zu entwickeln, dann ist es doch in einer solchen Studie extrem wichtig, dass ich diese Kovariate mit betrachte. Und das ist alles, was wir hier wollen.

Sie hatten das immer noch dargestellt, als würden wir da sehr hohe Anforderungen stellen, erst einmal ein Indextest 1 und dann noch ein komplexer Indextest 2 und dann kommt noch eine Referenzdiagnostik. Aber das sind ja jetzt nicht unbedingt alles apparativ hochtechnische Untersuchungen, sondern wir würden als Indextest 2 eben auch einen Vergleich zu einer klinischen Diagnostik oder einer laborchemischen Diagnostik akzeptieren. Es ist also nicht so, dass wir da unbedingt sehr stark eingrenzen in dem Sinne, dass wir immer schon zwei hochkomplizierte Geräte, medizinische Untersuchungen fordern, sondern das kann auch durchaus einfacher sein. Und dann gewinnt es eben eine Praxisrelevanz, die auch für unseren Auftraggeber wichtig ist.

**Fülöp Scheibler:** Ein ganz kurzer Kommentar zu meiner anekdotischen Information von vorhin. Ich kann das jetzt ganz gut sagen zu den onkologischen Themen. Da ist es tatsächlich so, dass die Fallzahlen durchschnittlich eher niedrig sind. Die variieren manchmal zwischen zehn in einer Studie und einigen Hunderten. Aber das Gros liegt bei 30, 40 Patienten pro Studie. Die Länder und die Technologien, mit denen die Studien durchgeführt wurden, variieren über den gesamten Erdball. Es gibt zum Beispiel in China eine sehr große Aktivität im Bereich der PET. Wenn man bedenkt, dass die Therapieregimes – jetzt wieder im onkologischen Setting – durchaus unterschiedlich sind in USA, Europa und China, dann fragt man sich schon: Will man wirklich diese Studien miteinander indirekt vergleichen? Wie das jetzt bei der Diagnostik der Alzheimer-Demenz aussehen wird, kann ich noch nicht sagen. Ich kenne die Studien noch nicht. Aber nach dieser Erfahrung, die wir bisher gemacht haben, gibt es tatsächlich sehr große Variationen zwischen den eingeschlossenen Populationen, zwischen den eingesetzten Technologien, zwischen der Prämedikation in diesen Bereichen und den vorhandenen Therapien. Da denke ich tatsächlich, dass dieser Vergleich nicht weiterführt. Man bekommt dann irgendeine Zahl, die womöglich statistisch sogar unterschiedlich sein kann, aber ob man sich darauf verlassen will, halte ich für sehr unsicher.

**Mathias Schreckenberger:** Ich weiß nicht, ob dieser a priori kategorische Ausschluss von indirekten Vergleichen nicht doch dazu führt, dass man wichtige Informationen übersieht. Denn es mag ja, genauso wie Sie dieses Beispiel konstruiert haben mit diesem ganz extremen Missverhältnis bei dem indirekten Vergleich, auch indirekte Vergleiche geben, die sehr gut

miteinander vergleichbar sind. Und dann wäre es schon problematisch, die a priori auszuschließen.

**Fülöp Scheibler:** Die Frage ist: Wie kann ich a posteriori beurteilen, dass zwei Kollektive beispielsweise gut miteinander vergleichbar sind? Dann sind Sie in derselben Problematik, die Sie sonst bei nicht randomisierten Interventionsstudien haben, dass Sie immer Gefahr laufen, irgendeinen sehr wichtigen Confounder nicht berücksichtigt zu haben. Sie können jetzt natürlich eine Population aus Ungarn vergleichen mit einer Population aus Frankreich und grob sozialstrukturelle Charakteristika, Krankheitscharakteristika usw. vergleichen. Das ist aber alles theoretisch, weil praktisch haben Sie die Daten gar nicht in den Studien. Aber selbst wenn Sie die hätten, könnten Sie ganz grob diese Charakteristika miteinander abgleichen und sagen: Das sieht ungefähr vergleichbar aus, auch die Technologie ist ähnlich, auch die Behandlungsschemata sind in den beiden Ländern ähnlich. Vielleicht könnte ich jetzt die Daten vergleichen. – Aber Sie wissen natürlich immer noch nicht, ob Sie nicht irgendeinen ganz wesentlichen Confounder nicht berücksichtigt haben. Das führt dann letztendlich dazu, dass Sie irgendwelche Zahlen bekommen – das kann mal pro PET ausgehen, das kann aber auch mal kontra PET ausgehen –, und Sie müssen dann damit leben. Das ist ein bisschen ein Lotteriespiel. Wir wollen ja möglichst sachlich und objektiv die Unterschiede in der diagnostischen Güte bewerten. Sonst können wir uns gleich einigen: Wir würfeln. Alle geraden Zahlen sind für die PET, alle ungeraden dagegen.

**Moderator Jürgen Windeler:** Wir nehmen Ihre Hinweise noch einmal auf. Ich glaube, man muss wirklich deutlich machen, dass schon im Bereich von randomisierten Therapiestudien indirekte Vergleiche kritisch gesehen werden und an spezielle Anforderungen geknüpft sind, die nicht ganz trivial sind, und dann immer noch – so ist breiter und allgemeiner Konsens, nicht nur IQWiG-Meinung – zu einer vergleichsweise unsicheren Aussage führen. Diagnosestudien, die man üblicherweise hat – ich sage das in dieser allgemeinen Form –, sind, was die Qualität der Aussagen angeht, weit, um nicht „Lichtjahre“ zu sagen, von randomisierten Studien entfernt. Ich sage nicht, dass man sie randomisiert machen soll, nicht dass ich missverstanden werde, aber die sind von dem her, wie sie gemacht werden und was sie aussagen sollen, weit davon entfernt, sodass an der Stelle das Problem der Anforderung an indirekten Vergleichen sich verschärft und zusätzlich verschärft dadurch, dass die Gütekriterien für diagnostische Verfahren – darum geht es ja hier, Sensitivität, Spezifität – viel stärker von den Populationen abhängig sind als die Therapieeffekte in Therapiestudien, sodass Sie also in der Frage der Heterogenität, die bei den Studien für die indirekten Vergleiche von zentraler Relevanz ist, noch ein viel größeres Problem bekommen, ein mehrfach größeres Problem als bei solchen Therapiestudien. Das ist der Grund, warum wir so extrem vorsichtig sind dabei.

Natürlich kann man jetzt sagen: Wenn ihr aber doch einmal einen indirekten Vergleich findet, der suffizient ist, dann nehmt ihn doch. Dem würde ich im Grundsatz auch erst einmal so zustimmen. Die Frage ist nur, ob man einen halbwegs vernünftigen Weg findet, den überhaupt zu identifizieren. Da bin ich sehr skeptisch.

**Dirk Hellwig:** Ich finde es gut, dass Sie jetzt noch einmal spezifizieren, was sozusagen der sinnvolle Wissensgewinn ist, den Sie versuchen zusammenzutragen. Ich denke, wir sind uns alle einig, dass das prospektive Design wünschenswert ist im Vergleich der verschiedenen Verfahren an demselben Patientenkollektiv. Wir ahnen aber auch, dass es derzeit diese Studien kaum gibt. Deswegen ist ja unser Wunsch, dass hier geholfen wird, dass eben die diagnostische Testgüte der verschiedenen Verfahren zusammengetragen wird. Wenn wir uns anschauen, wie das an anderer Stelle gemacht wird – ich zitiere mal wieder die Meta-Analyse von Yuan: Da ist beispielsweise der in den Studien dokumentierte Mini Mental State als ein einfaches Kriterium herangezogen worden, um wenigstens diese Vergleichbarkeit dem Leser zu überlassen, inwieweit diese Studien das Patientenkollektiv widerspiegeln, welches entsprechend untersucht werden soll.

Wir sind uns sicherlich auch darüber einig, dass es problematisch sein kann, diesen Mini Mental State zu verwenden. Das ist eine andere Diskussion. Aber man kann sich gewiss auf irgendwelche Basisparameter einigen, anhand derer man die Studien charakterisieren kann. Denn ansonsten ist zu befürchten, dass es wegen geringer Studienzahlen wenig informativ ist, was schlussendlich bei dem Nutzenbericht herauskommt. Ich habe bereits eingangs gesagt, es ist das Problem offen, zu sagen, welches Verfahren Potenzial hat. Das wissen wir nicht. Da wird vielleicht die diagnostische Testgüte in irgendeiner Form eine Rolle spielen. Das wäre wünschenswert, wenn dann wenigstens das Wissen über die diagnostische Testgüte anhand solcher Studien mit berichtet wird.

**Moderator Jürgen Windeler:** Herr Scheibler noch einmal, oder können wir das abschließen?

**Fülöp Scheibler:** Ist gut.

**Moderator Jürgen Windeler:** Dann würde ich es bei dieser Klärung im Sinne von nicht nur Meinungs austausch, sondern schon einer gewissen Annäherung und Verständigung, was die Meinungen angeht, belassen wollen.

Ich möchte nun meinen Punkt, den ich vorhin angesprochen habe, aufrufen wollen. Wir haben in einer ganzen Reihe von Stellungnahmen – ich würde sogar behaupten, in fast jeder – den Hinweis gefunden, dass insbesondere bei frühzeitig eingeleiteten Maßnahmen – das steht beispielsweise in der BPI-Stellungnahme – man besonders effektiv therapieren kann. Hier steht, dass durch gezielte Maßnahmen der Zeitpunkt, zu dem man erkrankt und zu einem Pflegefall wird, um sechs Monate verschoben werden kann. In einer anderen Stellungnahme – das ist die der Deutschen Gesellschaft für Gerontopsychiatrie und -psychotherapie – steht, dass ein möglichst frühzeitiger Therapiebeginn wünschenswert ist, und auch andere Äußerungen. Sie haben vorhin noch einmal ausgeführt, dass schon in der präklinischen Phase die Therapie besonders effektiv sei, dass Pflegebedürftigkeit verhindert werden könne. Ich würde gerne wissen, auf welche Informationen, Daten, Studien Sie diese Aussagen stützen, weil ich sie in keiner der Stellungnahmen gefunden habe.

**Martin Haupt:** Da ich direkt angesprochen bin, möchte ich zunächst einmal Ihre kurze Zitation aus unserer Stellungnahme darin gehend korrigieren, dass dort steht: Der Therapiebeginn sollte möglichst frühzeitig sein. – Im Satz davor steht aber: Wir wissen natürlich auch, dass der Krankheitsverlauf im Wesentlichen nur gemildert oder etwas verzögert werden kann. – Ich würde mich nicht den beiden anderen Reden anschließen, dass wir, wenn wir präklinisch behandeln, effektiv behandeln können und wissen das heute schon, oder dass wir, wenn wir präklinisch effektiv behandeln, was wir im Moment noch nicht können, dann schon die Pflegeheimeinweisung vermeiden oder hinauszögern könnten.

Der erste Teil des zweiten Parts ist schlichtweg nicht belegt. Wir haben keine klaren Belege dafür, dass irgendetwas hilft, was wir bisher zur Verfügung haben im MCI-Stadium. Man kann jetzt Subgruppen bilden und sagen: Da gibt es aber noch bei Donepezilstudien eine Subgruppe, die etwas besser reagiert hat. Und es gibt auch Studien, die wir mit Ginkgo biloba heranziehen könnten. Zum Beispiel in der GUIDAGE-Studie gibt es doch auch schon Hinweise dafür, dass sich irgendetwas nach vier Jahren bessern könnte oder jedenfalls zu einem geringeren Prozentsatz in eine Demenz konvertieren würde. Den Punkt würde ich also außer Acht lassen.

Ich halte den Punkt mit der Pflegeheimeinweisung für hochproblematisch in der Diskussion, denn er stigmatisiert zunächst einmal Seniorenheime. Das finde ich völlig unredlich. Es gibt Familien, die sind sehr froh darüber, dass es die Möglichkeit gibt, den Kranken in ein Seniorenheim, Pflegeheim weiterzugeben, dort zu besuchen und die furchtbare Belastung zu Hause nicht mehr haben zu müssen. Es ist wahrscheinlich für diese besondere Familie dann auch besser, wenn es so ist. Und in anderen Familien ist etwas anderes ein Ziel. Also Kriterium Survival, Kriterium Pflegeheim ja oder nein finde ich sehr problematisch sui generis.

Ansonsten: Wenn Sie klinische Studien oder randomisierte klinische Studien ansprechen, dann haben wir meist nur Sechs- bis Zwölfmonatszeiträume, zu denen wir etwas sagen können. Manchmal, wenn wir dann Open-Label-Behandlung betrachten, die fortgeführt worden ist, sehen wir, dass vielleicht auch in den dann weiterbehandelten Stichproben sich günstigere Verläufe zeigen, als man sie erwarten würde durch einen naturalistischen Verlauf, aber es gibt dann keine direkte Gegenüberstellung durch eine Placebogruppe.

Aber ich würde Ihnen gerne aus einer klinischen Erfahrungssituation antworten und sagen: Vielleicht können Sie ermessen – es würde sich in jedem Fall lohnen, diese Berichte immer wieder zu hören –, wie wichtig es für Familien ist, dass mit der Diagnose Behandlungsmöglichkeiten zur Verfügung stehen, die zumindest für einen gewissen Zeitraum stabilisieren oder verzögern können, gemessen an dem, was vorher passierte. Die Beratung, Begleitung, die dann erfolgt, sorgt dafür, dass wieder Mut gefasst wird, weiter zu versorgen, auch Mut gefasst wird, sich der Krankheit zu stellen. In dem Sinne ist, glaube ich, wo immer man früher beginnen kann, wenn man sicher ist, dass die Diagnose – wohlgermerkt: die Demenzdiagnose – steht, das von unschätzbarem Wert, auch wenn wir gegenwärtig keine

Daten darüber haben, ob das in zwei Jahren tatsächlich eine klare Verzögerung gegeben hat oder nicht. Da würde ich rein aus ärztlichen Gesichtspunkten antworten.

**Moderator Jürgen Windeler:** Dann bin ich nach Ihrer Äußerung froh, dass mein Wissensstand nicht völlig falsch war. – Gibt es weitere Wortmeldungen dazu?

**Alexander Drzezga:** Ich möchte kurz auf etwas hinweisen wollen, was in dem Zusammenhang von Interesse ist. Es ist gerade der Welt-Alzheimer-Bericht herausgekommen. Ich weiß nicht, ob der Ihnen ein Begriff ist. Das ist ein international erstellter Bericht, der sich fast hauptsächlich mit den Vorteilen von frühzeitiger Diagnose und Intervention der Alzheimererkrankung befasst, sicherlich nicht unbedeutend. Gerade letzte Woche hat in der WHO-Sitzung der stellvertretende Präsident der WHO darauf hingewiesen, eine bessere Früherkennung der Alzheimererkrankung wäre entscheidend. Das ist also schon ein Bereich, der sich momentan wissenschaftlich sehr stark als Thema darstellt.

Was ich ergänzen möchte zu dem, was Herr Haupt gesagt hat: Der Ausschluss der Alzheimererkrankung ist auch etwas, was man in diesen Endpunkten nicht wiederfindet, die Sie hier definieren. Ich weiß, dass das schwer fassbar ist. Es ist aber natürlich so, dass es für Patienten gerade mit MCI-Symptomatik, die zum Beispiel durch eine PET-Untersuchung eine relativ hohe Sicherheit bekommen, nicht an einer Alzheimererkrankung zu leiden, ein ungeheurer Wert darstellt, den man sehr schwer abbilden kann mit den Endpunkten, die Sie hier definiert haben. Auch das ist ein Punkt, den man hier dringend einbringen sollte.

**Moderator Jürgen Windeler:** Bei allem massiven Respekt vor der WHO habe ich allergrößte Probleme, mir Früherkennung ohne adäquate frühestmögliche Therapie vorzustellen. Aber das mag ja die Zukunft uns zeigen. Ich habe das auch eher als Forschungsthema, also Versorgungsthema verstanden.

**Alexander Drzezga:** Nein, das muss man vielleicht etwas konkretisieren. Ich hoffe, dass mir Herr Haupt wieder zur Seite springt. Es gibt schon eine ganze Reihe von Interventionen, die man in einem frühen Stadium beginnen kann. Ein wichtiger Punkt, der auch noch nicht zur Diskussion gekommen ist, ist, dass frühe Diagnostik auch bedeutet, dass die Leute noch in dem Besitz ihrer geistigen Kräfte sind im wesentlichen Maße, also auch in der Lage sind, ihre Angelegenheiten zu ordnen. Das ist sicher etwas ganz Relevantes. Und es gibt natürlich therapeutische, medikamentöse Interventionsmöglichkeiten, die jetzt nicht konkret zur Sprache gekommen sind, und auch andere Formen der sozialen Intervention, der Umstellung des Lebenswandels, eine ganze Reihe von Risikofaktoren, Hypertonie, Ernährung etc., die man adaptieren kann und von denen man weiß, dass die einen Einfluss haben auf den Erkrankungsverlauf.

**Martin Haupt:** Ja, ich springe Ihnen bei. Das ist richtig. Das Problem bei den auf wissenschaftlicher Ebene untersuchten Verfahren, die man im nichtmedikamentösen Bereich dort einsetzt, seien es kognitive Trainingsverfahren, seien es körperliche Ausdauerbildung,

Bewegungsverfahren, kardiovaskuläre Ausdauer oder ähnliche Interventionen, ist natürlich meist, dass sie bestmöglich über neun bis zwölf Monate gehen, meist aber nur über sechs bis zwölf Wochen. Das sind natürlich von der langfristigen Anlage her, wenn wir hier über krankheitsmodifizierende oder progressionsverzögernde Entwicklungen oder Einflüsse sprechen, fast vernachlässigbare Zeiträume, aber für die Familien natürlich ein ganz wichtiger Hoffungsanker und auch für die Zeit, in der das überprüft wurde, mit Nutzen belegt.

Ich glaube aber, auch aus einem anderen Grund, wenn ich das noch anschließen darf, bleibt Frühdiagnostik so wichtig, nämlich dass wir tatsächlich dann die Stichproben haben und auch sicher sind, dass wir die Richtigen identifiziert haben, wenn dann tatsächlich einmal eine Intervention zur Verfügung stehen sollte, die mehr macht als bisherige symptomatische Behandlungsmöglichkeiten.

**Moderator Jürgen Windeler:** Dieses Thema schließen wir auch.

Gibt es weitere Wortmeldungen zum Tagesordnungspunkt „Verschiedenes“?

**Martin Haupt:** Können Sie Ihren Zeitplan kurz nennen? Wie geht es jetzt weiter?

**Fülöp Scheibler:** Wir erstellen jetzt den abschließenden Berichtsplan, wo eben die Kommentare aus den Stellungnahmen und aus der heutigen Erörterung berücksichtigt werden. Parallel zu dem endgültigen Berichtsplan erscheint ein separates Dokument, das sich „Dokument und Würdigung der Stellungnahmen“ nennt, wo Sie nachlesen können, wie wir mit den einzelnen Argumenten umgegangen sind. Danach fangen wir an, den Vorbericht zu erstellen.

Den aktuellen Zeitplan habe ich nicht dabei. Insofern erwischen Sie mich auf dem falschen Fuß. Aber ich kann Ihnen den gerne nachreichen. Er steht aber auch im Internet. Auf der Seite, wo auch der Berichtsplan publiziert wird, stehen auch die weiteren Deadlines oder die weiteren Meilensteine im Projekt.

**Moderator Jürgen Windeler:** Weitere Wortmeldungen? – Alles geklärt.

Dann bedanke ich mich dafür, dass Sie da waren, dass Sie uns geholfen haben, dass wir einige Punkte wirklich gut und konstruktiv erörtert haben. Ich lade Sie nun zu einem abschließenden Imbiss ein und wünsche Ihnen eine gute Heimreise nach dem Imbiss. Danke schön.

**Anhang A – Dokumentation der Stellungnahmen**

# Inhaltsverzeichnis

	Seite
<b>1 Stellungnahmen von Organisationen, Institutionen und Firmen.....</b>	<b>A 2</b>
<b>1.1 Bundesverband der Pharmazeutischen Industrie e. V. ....</b>	<b>A 2</b>
<b>1.2 Deutsche Gesellschaft für Gerontopsychiatrie und -psychotherapie e. V. ....</b>	<b>A 29</b>
<b>1.3 Deutsche Gesellschaft für Nuklearmedizin e. V. ....</b>	<b>A 32</b>
<b>1.4 Deutsche Krankenhaus Gesellschaft .....</b>	<b>A 40</b>
<b>1.5 Lilly Deutschland GmbH.....</b>	<b>A 45</b>



## **1 Stellungnahmen von Organisationen, Institutionen und Firmen**

### **1.1 Bundesverband der Pharmazeutischen Industrie e. V.**

#### **Autoren:**

Wilken, Matthias

BPI e.V. | Friedrichstraße 148 | 10117 Berlin

Institut für Qualität und  
Wirtschaftlichkeit im Gesundheitswesen  
-IQWiG Stellungnahme zum Vorbericht D06-01N-  
Prof. Dr. med. Jürgen Windeler  
Dillenburger Str. 27  
51105 Köln

Ressort	Info	Rückspr.	Bearb.	Verbleib	Ressort	Info	Rückspr.	Bearb.	Verbleib
IL					GI				
AM					Komm				
NMV					VW				
VQ					GÖ				
Biometrie									

Arzneimittelzulassungen Europa

Telefon: (0 30) 2 79 09-1  
Telefax: (0 30) 2 79 09-3  
E-Mail: [\[REDACTED\]@bpi.de](mailto: [REDACTED]@bpi.de)  
Zeichen: MW  
Datum: 15.02.2012

**EINGEGANGEN**  
21. Feb. 2012

BPI-Stellungnahme vorab per e-Mail an [berichte@iqwig.de](mailto:berichte@iqwig.de)

**Stellungnahme zum Vorbericht D06-01N „Positronen-Emissions-Tomographie (PET) bei Alzheimer Demenz“**

Sehr geehrter Herr Professor Windeler,

beiliegend erhalten Sie die Stellungnahme des BPI Stellungnahme Vorbericht D06-01N „Positronen-Emissions-Tomographie (PET) bei Alzheimer Demenz“ inkl. der Formblätter, mit Bitte um Berücksichtigung.

Mit freundlichen Grüßen  
Bundesverband der Pharmazeutischen Industrie e.V.



Dr. Matthias Wilken  
Geschäftsfeldleiter

## Stellungnahme

Friedrichstraße 148  
10117 Berlin  
[www.bpi.de](http://www.bpi.de)

Telefon: (0 30) 2 79 09-1 72

Telefax: (0 30) 2 79 09-3 72

E-Mail: [REDACTED]@bpi.de

**EINGEGANGEN**

**21. Feb. 2012**

**Stellungnahme**  
**zum Vorläufigen Berichtsplan**  
**zur Positronenemissionstomographie**  
**(PET) und PET/CT**  
**bei Alzheimer Demenz**

**Auftrag: D 06-01N;**

**Version: 1.0;**

**Stand: 15.02.2012**

## Stellungnahme

### A. Sachverhalt

#### I. Hintergrund

Der Gemeinsame Bundesausschuss (nachfolgend „**G-BA**“) hat in seiner Sitzung am 19.12.2006 beschlossen, das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (nachfolgend „**IQWiG**“) wie folgt zu beauftragen:

*„Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen soll gemäß § 139a Abs. 3 Nr. 1 SGB V die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes zu PET bei folgenden Indikationen und in der angegebenen Reihenfolge durchführen:*

*(...)*

*14: Alzheimer Demenz.“*

*(...)*

#### Weitere Auftragspflichten:

*Mit dem Auftrag wird das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen verpflichtet,*

*a) die Verfahrensordnung des Gemeinsamen Bundesausschuss zu beachten,*

*(...)*“

In der Anlage I (Anerkannte Untersuchungs- und Behandlungsmethoden) der Richtlinie des G-BA zu Untersuchungs- und Behandlungsmethoden der vertragsärztlichen Versorgung (nachfolgend „**Richtlinie Methode vertragsärztliche Versorgung**“), werden für PET einige zugelassene Indikationen aufgeführt, bei denen diese zu Lasten der GKV erfolgt. Für die Indikation Alzheimer Demenz (nachfolgend „**AD**“) ist bislang keine Aufnahme in die Anlage I erfolgt.

## Stellungnahme

In der Richtlinie des G-BA zu Untersuchungs- und Behandlungsmethoden im Krankenhaus (nachfolgend „**Richtlinie Methoden Krankenhausbehandlung**“) sind PET und PET/CT bei einigen Indikationen gemäß § 5 Ziff. 4, von der Versorgung der gesetzlichen Krankenversicherung (nachfolgend „**GKV**“) im Rahmen von Krankenhausbehandlung ausgeschlossen. PET und PET/CT sind bei AD nicht durch die Richtlinie in der Krankenhausbehandlung ausgeschlossen.

Das IQWiG legt gemäß Ziff. 2 des Vorläufigen Berichtsplans folgende Ziele seiner Untersuchung fest:

*„Die vorliegende Untersuchung soll als Erstes prüfen, welcher patientenrelevante Nutzen oder Schaden mit der bildgebenden Diagnostik mittels PET bzw. PET/CT im Vergleich oder in Ergänzung zur konventionellen Diagnostik bei der (früheren) Detektion einer Alzheimer Demenz oder einem Therapie-Monitoring verbunden ist.*

(...)

*Folgende Technologievergleiche werden betrachtet:*

- *Diagnostische Strategie mit PET versus diagnostische Strategie ohne PET (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik.*
- *Diagnostische Strategie mit PET/CT versus diagnostische Strategie ohne PET/CT (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik.*
- *Diagnostische Strategie mit PET/CT versus diagnostische Strategie mit PET.*

(...)

## Stellungnahme

*Im Rahmen des vorliegenden Berichts soll sowohl der patientenrelevante Nutzen (Fragestellung 1) als auch die diagnostische und prognostische Güte (Fragestellung 2) für diese Indikationen untersucht werden:*

- *Diagnose einer Alzheimer Demenz (insbesondere Genauigkeit und Zeitpunkt der Diagnostik). Insbesondere folgende Unterindikationen:*
  - *Diagnostische Abgrenzung einer Alzheimer Demenz von anderen demenziellen Erkrankungen (z. B. vaskuläre Demenz)*
  - *Diagnostische Abgrenzung einer Alzheimer Demenz von anderen nicht-demenziellen Erkrankungen (z. .B. Depression)*
  - *Frühzeitige Abgrenzung von Patienten mit Risiko für Alzheimer Demenz (z. B. Patienten mit MCI), die eine AD entwickeln.*
  
- *Monitoring einer Therapie einer Alzheimer Demenz“*

Darüber hinaus wird unter Ziff. 4.1 „Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT“ Folgendes ausgeführt:

*„Um den patientenrelevanten Nutzen und Schaden der bildgebenden diagnostischen Technologien PET bzw. PET/CT für die Diagnose der Alzheimer Demenz zu bewerten, genügt es nicht festzustellen inwieweit die PET bzw. die PET/CT den gängigen diagnostischen Methoden überlegen ist. Vielmehr geht es bei einer erweiterten Bewertung darum, festzustellen, inwieweit sich eine präzisere Diagnosestellung auf die Behandlung und das weitere Management des Patienten auswirkt und inwieweit, in letzter Konsequenz, das veränderte Management patientenrelevante Endpunkte wie z. B. Mortalität, Morbidität oder die Lebensqualität beeinflusst.“*

Unter Ziff. 4.1.2.3 Patientenrelevante Endpunkte, werden für die Untersuchung folgende patientenrelevante Endpunkte verwendet, die eine Beurteilung der Interventionsziele ermöglichen:



## Stellungnahme

„(...)

- *Aktivitäten des täglichen Lebens (alltagspraktische Fähigkeiten),*
- *kognitive Leistungsfähigkeit*
- *gesundheitsbezogene Lebensqualität,*
- *andere, mit der Erkrankung verbundene Symptome (z. B. Depression, Schlaf-Wach-Umkehr, Wahrhaftigkeit, Agitiertheit (begleitende psychopathologische Symptome),*
- *Notwendigkeit einer vollstationären oder institutionalisierten ambulanten Pflege (Institutionalisierung),*
- *Mortalität,*
- *unerwünschte Arzneimittelwirkungen.*

*Zusätzlich werden auch Zielgrößen verwendet, die eine Beurteilung angehöriger relevanter Therapieziele ermöglichen:*

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

Das IQWiG beabsichtigt demzufolge offensichtlich, eine zweistufige Prüfung zur Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT durchzuführen (vgl. insbesondere Seite 17 des Vorläufigen Berichtsplans).

Auf einer ersten Stufe will das IQWiG feststellen, inwieweit PET bzw. PET/CT den gängigen Methoden überlegen ist. Auf einer zweiten Stufe soll eine erweiterte Bewertung dazu erfolgen, inwiefern sich die präzisere Diagnosestellung auf die Behandlung bzw. auch auf die patientenrelevanten Endpunkte der Ziff. 4.1.2.3 auswirkt.

Hierin sieht der Verband eine rechtswidrige Vermischung zwischen der Bewertung des Nutzens der Diagnostik und dem Nutzen der Therapie bei AD. Eine klare Grenzziehung erfolgt gerade nicht. Sofern die Therapie bei AD beurteilt werden sollte,

## Stellungnahme

müsste dies in einem eigenständigen Verfahren stattfinden. Es wird nachfolgend unter B. bis D. der rechtlich zulässige Beurteilungsrahmen für den Nutzen der Diagnostik mittels PET bzw. PET/CT bei AD dargestellt. Unter E. findet sich anschließend das zusammenfassendes Ergebnis.

### II. PET und PET/CT

Zum besseren Verständnis werden nachfolgend im Überblick die Vorteile der PET und PET/CT zusammenfassend dargestellt:

Werden bei einem Patienten kognitive Leistungsstörungen festgestellt, kann dies vielfältige Ursachen haben. AD ist dabei nur eine denkbar mögliche Ursache. Im Wege der klinischen Diagnostik kann eine AD nicht immer ausreichend festgestellt oder ausgeschlossen werden. PET und PET/CT sind indes geeignet, um Ablagerungen, sog. „Plaques“, im Gehirn festzustellen, die typisch für eine AD sind. Sog. „Amyloide Plaques“ können sogar bis zu 15 Jahren vor einer klinische manifesten Erkrankung mit AD im Gehirn erkannt werden, was dem Patienten die Chance gibt, seine weitere Lebensplanung aktiv zu gestalten und seine zukünftige Versorgung autark bestimmen zu können. Somit sind diese beiden diagnostischen Maßnahmen den klinischen häufig überlegen. Wird eine AD zuverlässig festgestellt, besteht ggf. die frühzeitige Möglichkeit, den Krankheitsverlauf eines Patienten zu entschleunigen. Insbesondere kann ggf. durch gezielte Maßnahmen der Zeitpunkt, zu dem ein Erkrankter zum Pflegefall wird, über weit mehr als 6 Monate (bis zu mehreren Jahren) nach hinten verschoben werden, was u.a. die Pflegeversicherung und sein Umfeld entlastet. Auch wird ein Patient, der um seine Erkrankung mit AD weiß, ggf. andere Lebensentscheidungen treffen wollen, als ohne Erkrankung. Inwiefern AD das Leben eines Patienten beeinflussen kann, wird dieser erst ersehen können, wenn er sich über die Erkrankung informiert, wofür er ohne Diagnose keinen Anlass haben wird.

Darüber hinaus profitieren aber auch diejenigen Patienten, die nicht an AD erkrankt sind, davon, wenn durch eine gezieltere Diagnostik festgestellt werden kann, dass ihre



## Stellungnahme

Gehirnleistungsstörung keine AD ist. Denn in der Vergangenheit wurden postmortal bei ca. 1/5 von Patienten festgestellt, dass sie, entgegen der „klinischen Diagnose AD“, gerade nicht an AD erkrankt waren. (Vgl. In J Neurol Neurosurg Psychiatry 1998;64:18–24, „Fallacies in the pathological confirmation of the diagnosis of Alzheimer’s disease“ von John V Bowler, David G Munoz, Harold Merskey, Vladimir Hachinski).

Hätte man bei diesen Patienten sogleich mittels PET oder PET/CT feststellen können, dass sie nicht an AD erkrankt waren, hätte man weitere Diagnosemaßnahmen zur Feststellung der eigentlichen Gehirnleistungsstörungen durchgeführt und die Patienten ggf. spezifischen Behandlungen zugeführt.

### **B. Rechtlich zulässiger Beurteilungsrahmen für den Nutzen der Diagnostik zu PET bzw. PET/CT bei AD**

#### **I. Zusammenfassung**

- Der Bewertungsrahmen zu PET bzw. PET/CT bei AD endet mit dem Ergebnis, das mit der jeweiligen Diagnosemethode erzielt werden kann.
- Das Ergebnis muss generell medizinisch verwertbar sein, wobei die Erkenntnisse nicht umfassend für zur Verfügung stehende Anschlusstherapien notwendig sein müssen, sondern es ausreichend ist, dass diese für künftige Therapiemöglichkeiten nützlich sein können.
- Die patientenrelevanten Endpunkte für die Bewertung der Diagnosestellung sind nicht gleichzusetzen mit denjenigen, die bei der Bewertung der Therapie von AD herangezogen werden. Ein Nachweis zum Vorliegen eines Mehrwerts der Diagnosemethode für solche Endpunkte anhand von Studien ist nicht erforderlich.

## Stellungnahme

### II. Verfahrensordnung des Gemeinsamen Bundesausschusses

#### 1. Übersicht über relevante Regelungen

Zur Vorbereitung seiner Entscheidungen kann der G-BA Aufträge an das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen („IQWiG“) nach Maßgabe der Bestimmungen der Verfahrensordnung des G-BA (nachfolgend „**Verf-O G-BA**“) vergeben, § 15 Abs. 1 des 4. Abschnitts des 1. Kapitels der VerfO G-BA. Das IQWiG erarbeitet sodann nach § 139 b SGB V Empfehlungen im Auftrag des G-BA, die vom G-BA im Rahmen seiner Aufgabenstellung sodann zu berücksichtigen sind, § 16 Abs. 1 VerfO G-BA. Absatz 3 des § 16 VerfO G-BA sieht in seinem Satz 1 ferner vor, dass die Zusammenarbeit zwischen dem IQWiG und dem G-BA bei einer Beauftragung auf der Grundlage der VerfO G-BA und nach Maßgabe der vom G-BA formulierten Aufträge erfolgt.

Die VerfO G-BA regelt in seinem 2. Kapitel die Bewertung medizinischer Methoden, u. a. zur Vorbereitung für Maßnahmen des G-BA für Richtlinien nach § 137 c und § 135 Abs. 1 Satz 1 SGB V, so in § 1 Abs. 1 VerfO G-BA des 1. Abschnitts. In § 3 des 1. Abschnitts des 2. Kapitels werden die gesetzlichen Grundlagen für die Bewertung medizinischer Methoden dargestellt, der wie folgt lautet:

#### **„§ 3 Gesetzliche Grundlagen**

- (1) *Rechtsgrundlagen für die Bewertung medizinischer Methoden sind für die vertragsärztliche und vertragszahnärztliche Versorgung die §§ 135 Abs. 1 und 138 SGB V sowie für die Versorgung mit Krankenhausbehandlung § 137 c SGB V.*
- (2) *Maßgeblich für das Bewertungsverfahren sind darüber hinaus insbesondere §§ 2 Abs. 1, 2 a, 12 Abs. 1 und 70 Abs. 1 SGB V, nach denen die Versorgung*

## Stellungnahme

*der Versicherten ausreichend, zweckmäßig und wirtschaftlich sein muss, auch in ihrer Qualität dem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen und dem medizinischen Fortschritt sowie den besonderen Belangen behinderter und chronisch kranker Menschen Rechnung zu tragen hat.“*

Im 3. Abschnitt „Bewertungsverfahren“ der VerFO G-BA setzt § 10 Abs. 2 fest, nach welchen Kriterien diagnostische und therapeutische Leistungen überprüft werden. Die Bestimmungen des § 10 lauten auszugswise wie folgt:

### **„§ 10 Unterlagen zur Bewertung der medizinischen Methoden**

*(2) Diagnostische und therapeutische Leistungen werden wie folgt überprüft:*

- 1. Die Überprüfung des Nutzens einer Methode erfolgt insbesondere auf der Basis von Unterlagen*
  - a) zum Nachweis der Wirksamkeit bei den beanspruchten Indikationen,*
  - b) zum Nachweis der therapeutischen Konsequenz einer diagnostischen Methode,*
  - c) zur Abwägung des Nutzens gegen die Risiken,*
  - d) zur Bewertung der erwünschten und unerwünschten Folgen (outcomes) und*
  - e) zum Nutzen im Vergleich zu anderen Methoden gleicher Zielsetzung.“*

Im Hinblick auf die Bewertung von Unterlagen sieht § 11 aus dem 3. Abschnitt des 2. Kapitels der VerFO G-BA in seinem Abs. 7 weiter vor, dass bei der Bewertung

*„ (...) insbesondere auch die Verwendung patientenrelevanter Zielgrößen (wie z.B. Mortalität, Morbidität, Lebensqualität), Versorgungsaspekte von Alter, biologischem und sozialem Geschlecht sowie lebenslangen spezifischen Besonderheiten, besondere Belange behinderter und chronisch kranker Menschen und die einer gesetzten Maßnahmen zur Vermeidung von verzerrten Studienergebnissen berücksichtigt werden“*

## Stellungnahme

sollen.

### **2. Auslegung der Regelung gemäß § 10 Abs. 2 Nr. 1, b) im 3. Abschnitt des 2. Kapitels VerfO G-BA: „Nachweis der therapeutischen Konsequenz einer diagnostischen Methode“**

Maßgeblich für die Bewertung des Nutzens der diagnostischen Methoden zur Bestimmung von AD sind im Wesentlichen die in § 10 Abs. 2 des 3. Abschnitts aus dem 2. Kapitel der VerfO G-BA genannten Kriterien. Diese sind im Kontext der der VerfO G-BA zugrunde liegenden rechtlichen Bestimmungen, besonders des SGB V, zu verstehen.

Der „Nachweis der therapeutischen Konsequenz einer diagnostischen Methode“ nach § 10 Abs. 2 Nr. 1 b) im 3. Abschnitt des 2. Kapitels der VerfO G-BA kann insofern nur dahingehend verstanden werden, dass die diagnostische Methode ein **Ergebnis** bringt, das **generell medizinisch verwertbar** ist. Es wäre zu weit gegriffen, Aspekte der sich anschließenden Therapie als therapeutische Konsequenz zu fordern. Sowohl die sich mit der Diagnose konkret stellenden Therapieoptionen als auch die Qualität der späteren Therapie dürfen bei der Bewertung der Methode der Diagnosestellung keine Rolle spielen. Die Bewertung des Nutzens der Diagnosestellung kann nur bis zu dem Punkt reichen, an dem das Ergebnis feststeht. Folgerungen oder sonstige Konsequenzen, die das jeweilige Ergebnis bewirken kann, müssen hierbei außen vor bleiben. Diese reduzierende Auslegung der Regelung des § 10 Abs. 2 Nr. 1 b) im 3. Abschnitt des 2. Kapitels der VerfO G-BA ergibt sich insbesondere aus den Regelungen des SGB V, was nachfolgend im Einzelnen erläutert wird.



## Stellungnahme

### III. Rechtliche Rahmenbedingungen des SGB V

Die Regelungen des SGB V sind Grundlage für die Regelungen der VerfO G-BA und dienen somit ebenso als rechtlicher und inhaltlicher Maßstab für die Nutzenbewertung des IQWiG. Insoweit wird auf diese auch in den Bestimmungen der VerfO G-BA, insbesondere in § 3 des 1. Abschnitts des 2. Kapitels, Bezug genommen.

Der Prüfraum des IQWiG wird festgelegt durch:

- die Richtlinienkompetenz des G-BA nach den §§ 135, 137 c SGB V (hierzu unter **C.I.1.**),
- den gesetzlich festgelegten Leistungsanspruch des Versicherten (hierzu unter **C.I.2.**) sowie
- die gesetzlich bestimmte Leistungspflicht der GKV (hierzu unter **C.I.3.**).

#### 1. Ermächtigungsgrundlagen zum Richtlinienenerlass

##### a) Die Regelungen der §§ 135, 137c SGB V

Wie dem Beschluss des G-BA vom 19.12.2006 zu entnehmen ist, soll die Bewertung des IQWiG der Vorbereitung für Maßnahmen zu der Richtlinie zu Untersuchungs- und Behandlungsmethoden im Krankenhaus (Richtlinie Methoden Krankenhausbehandlung) nach § 137 c SGB V sowie der Richtlinie des G-BA zu Untersuchungs- und Behandlungsmethoden der vertragsärztlichen Versorgung (Richtlinie Methoden vertragsärztliche Versorgung) nach § 135 SGB V dienen. Rechtliche Rahmenbedingungen zur Nutzenbewertung des IQWiG stellen daher die Vorschriften der §§ 135 und 137 c SGB V dar, die auszugsweise wie folgt lauten:

<sup>4</sup> Sitz des Verbandes - Friedrichstraße 148 - 10117 Berlin - Telefon-Zentral-Nr. (0 30) 2 79 09 - 0

## Stellungnahme

### **„§ 135 SGB V Bewertung von Untersuchungs- und Behandlungsmethoden.**

*(1) Neue Untersuchungs- und Behandlungsmethoden dürfen in der vertragsärztlichen und vertragszahnärztlichen Versorgung zu Lasten der Krankenkassen nur erbracht werden, wenn der gemeinsame Bundesausschuss auf Antrag eines Unparteiischen nach § 91 Abs. 2, Satz 1, einer kassenärztlichen Bundesvereinigung, einer kassenärztlichen Vereinigung oder des Spitzenverbandes Bund der Krankenkassen in Richtlinie nach § 92 Abs. 1 Satz 2 Nr. 5 Empfehlungen abgegeben hat über*

*1. die Anerkennung des diagnostischen und therapeutischen Nutzens der neuen Methode sowie deren medizinische Notwendigkeit und Wirtschaftlichkeit – auch im Vergleich zu bereits zu Lasten der Krankenkassen erbrachte Methoden – nach dem jeweiligen Stand der wissenschaftlichen Erkenntnisse in der jeweiligen Therapierichtung, (...)*“

### **„§ 137 c Bewertung von Untersuchungs- und Behandlungsmethoden im Krankenhaus.**

*(1) Der gemeinsame Bundesausschuss nach § 91 überprüft auf Antrag des Spitzenverbandes Bund, der Deutschen Krankenhausgesellschaft oder eines Bundesverbandes der Krankenhausträger Untersuchungs- und Behandlungsmethoden, die zu Lasten der gesetzlichen Krankenkassen im Rahmen einer Krankenhausbehandlung angewandt werden oder angewandt werden sollen, daraufhin, ob sie für eine ausreichende, zweckmäßige und wirtschaftliche Versorgung der Versicherten unter Berücksichtigung des allgemein anerkannten Standes der medizinischen Erkenntnisse erforderlich sind.*“

Sowohl § 135 als auch § 137 c SGB V sprechen von Untersuchungs- und Behandlungsmethoden sowie § 135 Abs. 1 in Nr. 1 SGB V von diagnostischem und therapeutischem Nutzen. Bereits die Wortwahl „Untersuchung“, „Behandlung“,

## Stellungnahme

„Methode“ sowie „diagnostischer Nutzen“ „und“ „therapeutischer Nutzen“ legt eine klar vorzunehmende Trennung fest.

Die vorliegende Bewertung des Nutzens von PET bzw. PET/CT bei AD als eine Methode zur Erstellung einer Diagnose begrenzt sich somit im Hinblick auf einen Richtlinienenerlass nach §§ 135, 137 c SGB V auf die Merkmale diagnostischer Nutzen einer Untersuchungsmethode.

### **b) Der medizinische Fortschritt im stationären Sektor**

Die Gesetzssystematik sieht vor, dass im ambulanten Sektor eine Erstattung der Kosten neuer Untersuchungs- und Behandlungsmethoden erst nach der Aufnahme einer solchen Methode in die Richtlinie Methoden vertragsärztliche Versorgung nach § 135 SGB V erfolgt. Im stationären Bereich hingegen gilt der Verbotsvorbehalt. Eine neue Methode ist solange grundsätzlich erstattungsfähig, bis sie aufgrund der Richtlinie Methoden Krankenhausbehandlung nach § 137 c SGB V ausgeschlossen ist.

Hieran zeigt sich die klare Entscheidung des Gesetzgebers, im stationären Sektor dem medizinischen Fortschritt Priorität einzuräumen. Soll eine Nutzenbewertung einer Diagnosemethode somit der Vorbereitung einer Maßnahme nach § 137 c SGB V dienen, so hat dieser gesetzgeberische Wille bei der Bewertung angemessen gewichtet zu werden.

Im Hinblick auf die Auslegung des „*Nachweis der therapeutischen Konsequenz einer diagnostischen Methode*“ nach § 10 Abs. 2 Nr. 1 b) im 3. Abschnitt des 2. Kapitels der VerfO G-BA bedeutet dies Folgendes: Die zu bewertende Diagnosemethode muss ein Ergebnis erzeugen, das auf irgendeine Weise medizinisch verwertbar ist. Nicht verlangt werden kann hiervon, dass das Ergebnis nur Informationen bringt, die für die Auswahl der bereits zur Verfügung stehenden Therapiemöglichkeiten erforderlich sind. Haben die zusätzlichen Informationen das Potential, dass aus ihnen weitere medizinische Methoden/ Arzneimittel etc. entwickelt werden können, liegt ein anzuerkennender Mehrwert vor.



## Stellungnahme

### c) Systematik der Richtlinien zur Regelung des Leistungsumfangs der GKV

Eine zwingend erforderliche Differenzierung zwischen Nutzenbewertung von Diagnose und Therapie ist ferner einem Rückschluss zu der gesamten Systematik der Richtlinienbestimmungen zur Regelung der Erstattung von Leistungen durch die GKV zu entnehmen.

Die Differenzierung zwischen einer Methode zur Diagnostizierung von AD und einer möglichen folgenden medikamentösen Therapie zeigt sich an der Aufteilung der vom G-BA zu erlassenden Richtlinien zur Kostenpflicht der GKV. So bestimmen die Richtlinie Methoden Krankenhausbehandlung nach § 137 c SGB V sowie Methoden vertragsärztliche Versorgung nach § 135 SGB V im Hinblick auf AD die Kostenerstattung von Methoden, hier möglicherweise künftig PET bzw. PET/CT zur Diagnosestellung von AD. Die Frage der Kostenpflicht von Therapiemaßnahmen mittels Medikation bestimmt wiederum die Richtlinie des G-BA über die Verordnung von Arzneimitteln in der vertragsärztlichen Versorgung („**AM-RL**“). Die Anlage III zur AM-RL enthält beispielsweise in Nr. 10 eine Bestimmung zur medikamentösen „*Therapie*“ von AD mit Antidementiva.

Der Bewertungsauftrag des IQWiG zum Nutzen von PET und PET/CT bei Alzheimer als eine Methode zur Erstellung einer Diagnose kann sich somit vor dem Hintergrund der möglichen zu erlassenden Richtlinien für den G-BA eben nur auf die Phase der Diagnosestellung beziehen und muss Nutzenaspekte einer sich möglichen anschließenden Therapie außen vor lassen.



## Stellungnahme

### **2. Leistungsanspruch eines Versicherten nach §§ 11 Abs. 1 Nr. 4, 27 Abs. 1 S. 1 SGB V**

In Bezug auf die Diagnosestellung von AD definiert sich der Leistungsanspruch der GKV im Wesentlichen nach den Regelungen der §§ 11 Abs. 1 Nr. 4, 27 Abs. 1 S. 1 SGB V. Hiernach hat ein Versicherter

*„Anspruch auf Krankenbehandlung, wenn sie notwendig ist, um eine Krankheit zu erkennen (...)“.*

Unter der Erkennung von Krankheit ist die Gesamtheit der diagnostischen Maßnahmen, d. h. der Feststellung von Art und Ursache der Krankheit zu verstehen (BeckOK SGB V § 27 Rn 24, KassKomm/Höfler SGB V § 27 Rn 49). Abzugrenzen hiervon ist die Früherkennung von Krankheiten im Sinne von § 25 SGB V (Gesundheitsuntersuchungen) und § 26 SGBV (Kinderuntersuchung) als Präventivmaßnahmen.

Eine gesetzliche Definition von Krankheit in diesem Zusammenhang existiert nicht. Die herrschende Auffassung versteht darunter zunächst einen regelwidrigen, vom Leitbild eines gesunden Menschen abweichenden Körper- oder Geisteszustand (BeckOK SGB V § 27 Rn 5). Teilweise wird im Rahmen des Krankheitsbegriffes darüber hinausgehend noch eine Behandlungsbedürftigkeit oder Arbeitsunfähigkeit verlangt (a.a.O.). Behandlungsbedürftigkeit liegt nach höchstrichterlicher Rechtsprechung vor, wenn durch den regelwidrigen Körper- oder Geisteszustand körperliche, geistige oder seelische Funktionen in einem solchen Maß beeinträchtigt sind, dass ihre Wiederherstellung ohne ärztliche Hilfe nicht erreichbar erscheint und nach den Regeln der ärztlichen Kunst der anomale Zustand einer Behandlung mit dem Ziel einer Diagnosestellung bzw. der Heilung, Linderung oder Verhütung einer Verschlimmerung zugänglich ist (vgl BSG 10.7.1979 – [3 RK 21/78](#) NJW 1980, [1919](#) = SozR 2200 § 182 Nr [47](#); 6.8.1987 – [3 RK 15/86](#) NJW 1988, [1550](#) = SozR 2200 § 182 Nr [106](#)). Die Rechtsprechung verlangt somit einerseits die Erforderlichkeit einer medizinischen

## Stellungnahme

Therapie, andererseits lässt sie das Ziel der Verhütung einer Verschlimmerung des Zustandes ausreichen.

Zusammengefasst hat ein Patient somit einen Anspruch darauf, dass eine Erkrankung zu Lasten der GKV diagnostiziert wird, die zur Linderung ärztlicher Hilfe bedarf.

Vor dem Hintergrund, dass die Nutzenbewertung des IQWiG der Vorbereitung einer möglichen Maßnahme zur Konkretisierung der Leistungspflicht der GKV dient, ist zur Auslegung der Regelungen der VerfO G-BA auch die zitierte sozialgerichtliche Rechtsprechung heranzuziehen. Die „*therapeutische Konsequenz einer diagnostischen Methode*“ nach § 10 Abs. 2 Nr. 1, b) im 3. Abschnitt des 2. Kapitels der VerfO G-BA kann somit allenfalls so weit reichen, als dass das **Ergebnis der Diagnose eine behandlungsbedürftige Krankheit in diesem Sinne ergibt**. Die therapeutische Konsequenz ist als deckungsgleich mit der Notwendigkeit der ärztlichen Behandlung und der potentiellen Möglichkeit des Erreichens eines Erfolges – wenn dieser auch nur in der Verhütung einer Verschlimmerung – besteht.

Die Behandlungsbedürftigkeit bedeutet hingegen nicht, dass der Patient sich auch zwingend nach Erhalt der Diagnose behandeln lassen muss. Das Selbstbestimmungsrecht des Patienten nach Art. 2 GG überlasst es insoweit dem Patienten selbst zu entscheiden, welche ärztliche Behandlung er in Anspruch nehmen möchte. Um aber eine Entscheidung treffen zu können, welche Behandlungsmöglichkeiten ergriffen werden sollen, benötigt der Patient eine umfassende Aufklärung über seinen Krankheitszustand.

Ein Recht auf Diagnose leitet sich somit auch aus dem verfassungsrechtlich verankerten Selbstbestimmungsrecht eines Patienten her. Nachdem bei einer Diagnosestellung mit PET bzw. PET/CT bekanntermaßen ein zuverlässigeres und genaueres Ergebnis erzielt werden kann, das für die Entscheidung der weiteren Behandlung maßgeblich ist, hat ein Patient ein Recht auf diese entscheidungsrelevanten Informationen.

## Stellungnahme

### 3. Leistungspflicht der GKV nach § 2 Abs. 1 S. 3 SGB V

Gemäß § 2 Abs. 1 S. 3 SGB V haben

*„Qualität und Wirksamkeit der Leistungen (...) dem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen und den medizinischen Fortschritt zu berücksichtigen.“*

Hiernach sind von der GKV grundsätzlich die Methoden anzuerkennen, die dem bereits bestehenden medizinischen Standard entsprechen. Die Grenze darf des Weiteren nicht so eng gezogen werden, dass ein medizinischer Fortschritt keine Beachtung findet. Zwar kann ein Versicherter nur verlangen und beanspruchen, was im Sinne von § 2 Abs. 1 S. 3 SGB V anerkannt ist, umgekehrt darf ein Versicherter aber auch nicht auf weniger wirksame Behandlungsmethoden verwiesen werden, sobald effiziente Heilmittel zur Verfügung stehen, die hiernach als anerkannt zu gelten haben (BeckOK SGB V § 2 Rn 4).

Die gesetzliche Verankerung der Passage *„den medizinischen Fortschritt zu berücksichtigen“* zeigt klar den gesetzgeberischen Willen, dass auch die Medizin im GKV-Wesen neuen Entwicklungen offen gegenüber steht. Mit der gleichzeitigen Verpflichtung, zu dem allgemein anerkannten Stand der medizinischen Erkenntnisse den *„medizinischen Fortschritt“* zu berücksichtigen, impliziert das Gesetz in den (medizinischen) Standard eine Dynamik (vgl. jurisPK-SGB V, Plagemann zu § 2 Rn 51). Auch unter dem Aspekt der zwingenden Berücksichtigung von medizinischen Fortschritten bei der Ermittlung des Leistungsumfangs der GKV bestätigt sich die Auslegung, dass sich auch die Nutzenbewertung von Diagnosemaßnahmen durch das IQWiG nicht auf nur bestehende Erkenntnisse begrenzen darf. Aus diesem Grund kann die Zielgröße bei der Bewertung einer Methode zur Findung einer Diagnose nur ein generell medizinisch verwertbares Ergebnis sein, in dem Sinne, dass das Ergebnis nicht zwingend einer bereits verfügbaren sich anschließenden Therapie entsprechen muss. So muss auch eine Erkenntnis als Mehrwert einer Diagnosemaßnahme



## Stellungnahme

anerkannt werden, die beispielsweise für eine sich noch in der Entwicklung befindlichen Therapie nützen kann.

#### 4. Grundsatz der Therapiefreiheit des Arztes

Unter Therapiefreiheit versteht man, dass ein Arzt die Freiheit hat, zu entscheiden, welche Diagnostik und welche Therapie medizinisch notwendig sind. Diesbezüglich stehen dem Arzt somit zwei gänzlich unterschiedliche Entscheidungsfreiheiten zu. Für welche Therapie sich der Arzt entscheidet, kann u.a. vom Willen des Patienten, der durch das Selbstbestimmungsrecht aus Art. 2 GG geschützt ist, abhängen. Ein Patient kann sich indes gänzlich gegen eine Therapie entscheiden, weshalb sich eine solche ggf. überhaupt nicht an eine Diagnoseentscheidung anschließt. Auch vor diesem Hintergrund muss eine getrennte Beurteilung von der PET und PET/CT sowie einer AD-Therapie erfolgen.

Darüber hinaus muss bei der Beurteilung der PET und PET/CT beachtet werden, dass ein Ausschluss dieser Leistungen aus der GKV, eine unzulässige Einschränkung der Diagnosefreiheit eines Arztes, der diese diagnostischen Verfahren als medizinisch notwendig erachtet, bedeuten würde. Damit wäre er in seinem Grundrecht aus Art. 12 GG (Berufsausübungsfreiheit) verletzt.

Zudem hätte der Ausschluss der vorbenannten diagnostischen Verfahren zur Folge, dass es einem Arzt nicht möglich ist, eine sichere (als durch klinische Verfahren mögliche) Diagnose zu stellen und nachfolgend eine entsprechend richtige Therapieentscheidung zu treffen.

Die Therapiefreiheit ist Bestandteil der ärztlichen Berufsausübungsfreiheit, wie sie durch Art. 12 I GG geschützt wird. Damit wird in erster Linie ein Abwehrrecht beschrieben. (BVerfGE 102, 26 (36) - Frischzellen; Dahm, in: Rieger, Lexikon des Arztrechts, Nr. 5090; Laufs, Arztrecht, Rn. 41 ff.; Laufs, in: Laufs/Uhlenbruck, Handbuch des Arztrechts, Rn. 11 zu § 3; ders., Zur Freiheit des Arztberufs, in: FS f. Deutsch, 625

## Stellungnahme

ff.; Deutsch/Spickhoff, MedR, Rn. 19; Grupp, MedR 1992, 256 f.; OVG Münster, MedR 1998, 130; ausführlich Schnitzler, Das Recht der Heilberufe, 242 ff.; Zuck, Das Recht der anthroposophischen Medizin, 2007 Rn. 102.) Der Arzt verantwortet seine Therapie nach professionellen Standards vor seinem Gewissen (§ 1 II BÄO, § 2 I MBO), kann also nicht zu einer diesen Grundsätzen widersprechenden Behandlungsmethode gezwungen werden. (Siehe dazu Daniels/Bulling, Bundesärzte-ordnung 1963, Rn. 14 ff. zu § 1 BÄO; Lippert, in: Ratzel/Lippert, MBO, Rn. 16 ff. zu § 2 MBO.) Dem Arzt steht aufgrund seiner fachlichen Kompetenz ein weiter Ermessensspielraum für die Entscheidung zu einer medizinisch notwendigen Therapie und Diagnostik zu. Dies bedeutet, dass der Arzt eine Diagnostische Maßnahme wählen darf, wenn diese nach den objektiven medizinischen Befunden und wissenschaftlichen Erkenntnissen zum Zeitpunkt der Durchführung als medizinisch notwendig anzusehen war. Die Therapiefreiheit steht grundsätzlich in einem Spannungsverhältnis zum Wirtschaftlichkeitsgebot, wobei dieses indes die Therapiefreiheit nicht unzulässig einschränken darf.

Auch im Krankenhaus darf das Wirtschaftlichkeitsgebot die ärztliche Therapiefreiheit nicht ersticken. (Laufs/Kern, Handbuch des Arztrechts, 4. Auflage 2010, § 3 Rn. 22.) Andernfalls würde dies das Ende jeglicher Innovationen bedeuten.

Im Hinblick auf das Wirtschaftlichkeitsgebot ist insbesondere zu beachten, dass es eine frühzeitige, sichere Diagnose von AD, wie sie durch PET und PET/CT gestellt werden kann, ermöglicht, den Krankheitsverlauf zu verlangsamen. Idealerweise kann durch einen frühen Behandlungsbeginn der Zeitpunkt, zu dem ein Patient mit AD pflegebedürftig wird, weit hinaus-geschoben werden. Bekanntermaßen entstehen insbesondere durch eine Unterbringung von Erkrankten mit AD in Pflegeeinrichtungen enorme Kosten für die GKV bzw. Pflegeversicherung.

## Stellungnahme

### **IV. Verstoß gegen Art. 14 GG zum Nachteil der Hersteller von PET- und PET/CT-Geräten**

Vor dem Hintergrund, dass die PET und PET/CT zur sicheren, genauen Diagnose bzw. zum Ausschluss von bestimmten Formen von AD beim Patienten bestimmt sind, ist ergänzend zu berücksichtigen, dass sich deren Ergebnis in der Diagnose erschöpft. Will das IQWiG in die Bewertung von PET und PET/CT, die ggf. anschließenden Therapiemöglichkeiten nach der Diagnose AD mit einbeziehen, wäre hierin ein Verstoß gegen Art. 14 GG zum Nachteil der Hersteller von PET- und PET/CT-Geräten zu sehen. Dies resultiert insbesondere daraus, dass der Hersteller diese Geräte nicht mehr herstellen bzw. absetzen könnte.

Etwaige Therapien, die im Anschluss an die Diagnose von AD zu einem positiven/milderen Verlauf der Krankheit beitragen können, kann ein Hersteller von PET bzw. PET/CT nicht initiieren, da mit der präzisen Erstellung der Diagnose die "Aufgabe" der PET und PET/CT erfüllt ist. Von einem Diagnoseverfahren zu verlangen, dass dieses auf die nachfolgenden Therapiemöglichkeiten Einfluss nehmen soll, würde diesem indes Unmögliches abverlangen. Denn es unterliegt letztendlich der Therapiefreiheit des jeweiligen Arztes, der die Diagnose mittels PET und PET/CT veranlasst hat, die jeweiligen Konsequenzen aus dem Ergebnis zu ziehen und das weitere Vorgehen mit dem Patienten abzusprechen. Auch deshalb ist die Beurteilung des Nutzens der PET bzw. PET/CT streng davon zu trennen, inwieweit daraus im Folgenden bestimmte Therapien abgeleitet werden. Entscheidend kann deshalb allenfalls nur sein, inwieweit die hier zu betrachtenden Methoden anderen Diagnostikmethoden gegenüber überlegen, gleichwertig oder unterlegen sind. Da jedoch aus dem Berichtsplan des IQWiG nicht hervorgeht, dass an dieser Stelle die Überprüfung des Nutzens der PET und PET/CT enden soll, sei nochmals darauf hingewiesen, dass dies zu einer unzulässigen Kompetenzüberschreitung führt.

Würde man von einem Hersteller von PET und PET/CT mehr als die Diagnose (Outcome), für die die Geräte entwickelt wurden, abverlangen, wäre darin eine Verletzung von Art. 14 Abs. 1 GG zu sehen. Denn die Entscheidung des



## Stellungnahme

Geräteherstellers zur Herstellung seines gewählten Produktes ist eine individuelle Entscheidung im Rahmen seiner betrieblichen Tätigkeit. Die Zielsetzung des Produktes ist ein sehr wichtiger Bestandteil dessen. Der Hersteller eines Gerätes kann selbst entscheiden, welche diagnostischen Ergebnisse das Gerät erzielen können soll. Die bevorstehende Prüfung des medizinischen Nutzens der PET und PET/CT darf deshalb nicht so weit gehen, dass es dazu führt, dass dem Hersteller abverlangt wird, dem Gerät (nachträglich) eine völlig andere Zielrichtung zu geben, die er nie erreichen kann. Erst recht darf die Freiheit eines Herstellers im Rahmen von Art. 14 GG, die grundsätzlich zwar durch Gesetze inhaltlich eingeschränkt werden darf, nicht mit derartigen Beschränkungen Unmögliches abfordern. Genau das würde aber verlangt, wenn die Diagnostikgeräte auch die Therapie beeinflussen können müssten.

### **C. Hilfsweise: Vorliegen einer therapeutischen Konsequenz bei PET bzw. PET/CT**

Selbst wenn jedoch nach § 10 Abs. 2 Nr. 1, b) im 3. Abschnitt des 2. Kapitels VerfO G-BA eine weitergehende therapeutische Konsequenz gefordert würde, ist deren Vorliegen zu bejahen. Wie eingangs unter Punkt **A.I.** dargestellt, kann bekanntermaßen mit der PET bzw. PET/CT-Methode ein genaueres Ergebnis gegenüber einer klinischen Diagnosestellung erzielt werden.

Dies führt beispielsweise dazu, dass Fehldiagnosen vermieden werden können, eine fehlende Therapie eines an AD erkrankten Patienten ergriffen bzw. optimiert werden kann und die genaueren Daten zu einer weiteren Fortentwicklung der sich anschließenden Therapie nützen. Eine medizinische Konsequenz bzw. eine Kausalität des bei Diagnosestellung von AD mittels PET und PET/CT auf die folgende Therapie, im Übrigen auch auf die herangezogenen Endpunkte, liegt somit vor.

Wir weisen jedoch auch an dieser Stelle nochmals darauf hin, dass eine Bewertung der therapeutischen Konsequenz selbst Gegenstand einer Nutzenbewertung der sich anschließenden Therapie wäre. Ebenso kann ein Nachweis eines Mehrwerts einer

## Stellungnahme

anschließenden Therapie nicht zum Nachweis des Mehrwerts der Methode zur Diagnosestellung gefordert werden. Demgemäß kann die Nachweispflicht auch nur soweit führen, bis ein Nachweis für ein medizinisch verwertbares Ergebnis der jeweiligen diagnostischen Methode erbracht worden ist.

### **D. Verwendung patientenrelevanter Zielgrößen gemäß § 11 Abs. 7 des 3. Abschnitts des 2. Kapitels VerfO G-BA**

Das IQWiG zieht nach dem Vorläufigen Berichtsplan zur Bewertung des Nutzens der Diagnosestellung patientenrelevante Endpunkte heran. Nach Ziff. 3.5 der „Allgemeinen Methoden“ des IQWiG, Version 4.0 vom 23.09.2011 (nachfolgend **„Allgemeine Methoden“**) führt das IQWiG

*„(...) Nutzenbewertungen zu diagnostischen Verfahren in erster Linie auf der Grundlage solcher Studien mit patientenrelevanten Endpunkten durch. (...) Unter patientenrelevanten Endpunkten sind in diesem Zusammenhang dieselben Nutzenkategorien zu verstehen wie bei der Bewertung therapeutischer Interventionen, nämlich Mortalität, Morbidität und gesundheitsbezogene Lebensqualität.“*

Dabei werden im Vorläufigen Berichtsplan noch detailliert weitere Endpunkte in Ziff. 4.1.2.3 benannt.

Die Endpunkte für ein Diagnostisches Verfahren können dabei nicht identisch mit denen der Therapie, die sich ggf. an eine Diagnose anschließt, sein. Vielmehr ist eine Trennung von Endpunkten für die Diagnose und für die Therapie vorzunehmen, wobei im Rahmen der Beurteilung der PET und PET/CT keine Bewertung der Therapien für AD erfolgen kann.

Formal betrachtet sind daher auch die patientenrelevanten Endpunkte für die Bewertung von diagnostischen Maßnahmen anders zu definieren als diejenigen der Bewertung einer darauffolgenden therapeutischen Maßnahme. Dies aus dem Grunde,



## Stellungnahme

dass die Bewertung des Nutzens einer therapeutischen Maßnahme sozusagen die „Phase 2“ betrifft, die erst in dem Zeitpunkt der bereits vorliegenden Diagnose beginnt. Die Bewertung der Maßnahme der Diagnosestellung hingegen betrifft demgemäß die „Phase 1“, die folglich in dem Zeitpunkt endet, in dem die Diagnose feststeht und die Phase 2 beginnt. Würden somit Aspekte der Phase 2 in die Nutzenbewertung der Phase 1 einbezogen, würde die Bewertung der Phase 1 insoweit verfälscht. Die Bewertung des Nutzens der Diagnosestellung muss somit mit dem Zielpunkt enden, dass die Methode ein generell medizinisch verwertbares Ergebnis bringt, in dem festgestellt wird, dass die Diagnose eine Konsequenz hat; d. h. dass grundsätzlich Therapiemöglichkeiten bereits existent sind bzw. entwickelt werden können.

Die reine Diagnose einer Erkrankung kann zudem niemals eine unmittelbare Auswirkung auf Endpunkte wie Mortalität und Morbidität haben. Eine direkte Auswirkung kann lediglich die sich an die Diagnose anschließende Therapie entfalten. Zutreffend kann es daher allenfalls sein, der Diagnose einer Erkrankung andere Zielgrößen als Mortalität und Morbidität zuzumessen, da die Diagnosestellung alleine keine Beeinflussung auf die Mortalität eines Patienten hat.

Anders hat die Diagnose von AD beispielsweise eine unmittelbare Auswirkung für die gesundheitsbezogene Lebensqualität des Patienten und seiner Angehörigen. Durch die Diagnose hat der Patient überhaupt erst die Möglichkeit eine Entscheidung darüber zu treffen, ob er sich wegen der AD behandeln lässt oder nicht. Auch für den Fall, dass bei einem Patient mittels PET bzw. PET/CT ausgeschlossen werden konnte, dass eine AD vorliegt, schließen sich hierdurch positive Konsequenzen für einen Patienten an. Denn durch den Ausschluss von AD, entgeht ein Patient der Gefahr, dass er bei einer fehlerhaften Diagnose keiner gezielten Behandlung für eine ganz andere Gehirnleistungsstörung zugeführt wird.

Bei der Auswahl der heranzuziehenden patientenrelevanten Endpunkte kann es somit nur darum gehen, zu beurteilen, inwieweit die Patienten die Erkenntnis des Vorliegens ihrer AD oder nicht beeinflussen kann. Denn die zur Verfügung stehenden Therapiemöglichkeiten von AD sind unabhängig von der Art und Weise der

## Stellungnahme

Diagnosegewinnung quasi immer gleich. Entscheidend ist jedoch, dass und wann ein verwertbares und detailliertes Ergebnis über das Vorliegen der Erkrankung gewonnen wird. Insoweit ist die PET bzw. PET/CT-Methode anderen Diagnostischen Verfahren überlegen.

§ 11 Abs. 7 des 3. Abschnitts des 2. Kapitels der Verfahrensordnung des G-BA bestimmt demgemäß:

*„Bei der Bewertung sollen insbesondere auch die Verwendung patientenrelevanter Zielgrößen (wie z. B. Mortalität, Morbidität, Lebensqualität), Versorgungsaspekte von Alter, biologischem und sozialem Geschlecht sowie lebenslagenspezifischen Besonderheiten, besondere Belange behinderter Menschen und die eingesetzten Maßnahmen zur Vermeidung von verzerrten Studienergebnissen berücksichtigt werden.“*

Im Rahmen der Gesamtbewertung sollen daher für die konkrete Bewertung angemessene Zielgrößen Berücksichtigung finden. Nicht jedoch ist bei jeder Bewertung – ob der Bewertung einer diagnostischen Methode oder einer therapeutischen Maßnahme – zwingend auf das durch Studien nachgewiesene Vorliegen jeweils entsprechend definierter patientenrelevanter Endpunkte abzustellen.

Vor diesem Hintergrund sind die in dem Vorläufigen Berichtsplan dargestellten patientenrelevanten Endpunkte zu prüfen und deren Nachweis allenfalls gemindert zu verlangen.

### **D. Ergebnis**

Der rechtlich zulässige Beurteilungsrahmen für den Nutzen der Diagnostik zu PET bzw. PET/CT bei AD beschränkt sich darauf zu prüfen, ob die Methode ein Ergebnis bringt, das generell medizinisch verwertbar ist. Ist die Diagnose AD beispielsweise in einem früheren Stadium zu erkennen oder bringt sie detailliertere Informationen, liegt

## Stellungnahme

darin ein anzuerkennender Mehrwert. Soweit bei der Bewertung auf patientenrelevante Endpunkte abgestellt wird, so sind diese nicht gleichzusetzen mit denjenigen, die bei der Bewertung der Therapie von AD herangezogen werden.

Der Verband sieht die Notwendigkeit, den vorliegenden Vorläufigen Berichtsplan unter Berücksichtigung der dargestellten Argumentation anzupassen.

## **1.2 Deutsche Gesellschaft für Gerontopsychiatrie und -psychotherapie e. V.**

### **Autoren:**

Adler, Georg  
Haupt, Martin



## Stellungnahme der DGGPP e.V.

### **Stellungnahme zum vorläufigen Berichtsplan „Positronen-Emissions-Tomographie (PET) bei Alzheimer-Demenz“**

Die Diagnostik und Therapie der Alzheimer-Demenz unterliegt derzeit einer raschen Weiterentwicklung, die zur Milderung der immensen individuellen und gesellschaftlichen Belastung dringend erforderlich ist. Neue therapeutische Prinzipien, die eine kausale Therapie dieser Erkrankung möglich erscheinen lassen, wurden und werden entwickelt. Einige dieser Therapieverfahren werden bereits in Phase 2- und Phase 3-Studien untersucht. Bekannt ist, wie auch im Berichtsplan angedeutet, dass die der Erkrankung ursächlichen pathologischen Prozesse schon lange Jahre vor der klinischen Manifestierung der Erkrankung beginnen.

Mit den derzeit zur Verfügung stehenden medikamentösen Therapieverfahren kann der Krankheitsverlauf im Wesentlichen nur gemildert und verzögert werden. Entsprechend ist zur Ausnutzung ihres therapeutischen Potentials ein möglichst frühzeitiger Therapiebeginn, im Stadium der beginnenden Demenz, wünschenswert.

Aus diesen Gründen kommt der Frühdiagnostik der Alzheimer-Krankheit eminente Bedeutung zu. Auf diesem Gebiet liegt der Nutzen der PET-Untersuchungen - ob es sich dabei nun um die Darstellung charakteristischer Veränderungen des Gehirnmotabolismus mittels FDG-PET oder um die Darstellung von Amyloid-Ablagerungen, z.B. mittels PIB-PET, handelt.

Eine diesbezügliche Stellungnahme, die unsere Einschätzung wiedergibt, findet sich in der im Berichtsplan erwähnten S3-Leitlinie „Demenzen“ der DGPPN und DGN, die auch unter Beteiligung der Deutschen Gesellschaft für Gerontopsychiatrie und -psychotherapie e.V. (DGGPP) entwickelt wurde. Demnach können FDG-PET und HMPAO-SPECT insbesondere „in klinisch unklaren Fällen“ „diagnostisch wertvolle Informationen geben und die ätiologische Zuordnung einer Demenz ermöglichen“. Ein regelhafter Einsatz in der Diagnostik wird nicht empfohlen.

Diese Empfehlung stützt sich in erster Linie auf Untersuchungen zur Sensitivität und Spezifität dieser Verfahren bei der Unterscheidung zwischen Alzheimer-Patienten und Kontrollpersonen oder Patienten, die an anderen Demenzformen erkrankt sind.

Die Erhebung der Daten, wie sie in Teilen des Berichtsplans gefordert werden, erscheint mit den Erfordernissen einer angemessenen Patientenversorgung – auch in Forschungskontexten – nicht vereinbar. Es ist beispielsweise kaum machbar und auch ethisch kaum vertretbar, bei Patienten, bei denen die klinischen Kriterien eines Demenzsyndroms erfüllt sind und die bei ansonsten untypischer klinischer Präsentation (z.B. sehr frühem Erkrankungsalter) anhand einer PET-Untersuchung einer diagnostischen Gruppe zugeordnet werden, die dann angemessene Behandlung in einem doppelblinden randomisierten Design zu untersuchen.

Daher ist zu befürchten, dass die im Berichtsplan anvisierten Kriterien zu ambitioniert sind und auch der aktuellen stetigen Weiterentwicklung in Diagnostik und Therapie der Demenzen nicht Rechnung tragen, insbesondere dem unbestreitbaren Wert der PET in der Früherkennung, auch gegenüber der SPECT. Die PET wird damit als sensitives und spezifisches Verfahren mit hohem prädiktiven Wert für die zutreffende Erkennung sehr leichtgradiger kognitiver Störungsprozesse nicht angemessen bewertet.

Für die DGGPP 15. Februar 2012

PD Dr. Martin Haupt

Prof. Dr. Georg Adler

### **1.3 Deutsche Gesellschaft für Nuklearmedizin e. V.**

**Autoren:**

Drzezga, Alexander

Hellwig, Dirk

Kotzerke, Jörg

Sabri, Osama

Schreckenberger, Mathias





EINGEGANGEN

17. Feb. 2012

DGN e. V. • Nikolaistraße 29 • D-37073 Göttingen

Geschäftsstelle  
Nikolaistraße 29  
D-37073 Göttingen

**Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen**

– IQWiG Stellungnahme zum Berichtsplan D06-01N  
Prof. Dr. med. Jürgen Windeler  
Dillenburger Straße 27  
51105 Köln

	Info	Rückspr.	Bearb.	Verbleib	Ressort	Info	Rückspr.	Bearb.	Verbleib
IL					GI				
AM					Komm				
NMV			/	/	Ynt				
VQ					GO				
Biometria									

Tel. (Mitgliederversammlung) +49 (0)551.48857.  
Tel. (Kassen) +49 (0)551.48857.  
Tel. (Sekretariat) +49 (0)551.48857.  
E-Mail: [info@nuklearmedizin.de](mailto:info@nuklearmedizin.de)  
Homepage: [www.nuklearmedizin.de](http://www.nuklearmedizin.de)

Ihre Zeichen

Ihre Nachricht vom

Unser Zeichen

Göttingen

gj/uf

15. Februar 2012

**Stellungnahme der Deutschen Gesellschaft für Nuklearmedizin zum Vorläufigen Berichtsplan des IQWiG zum Auftrag D06-01N: Positronenemissionstomographie (PET) und PET/CT bei Alzheimer Demenz, Version 1.0 vom 9. Januar 2012**

Sehr geehrter Herr Professor Windeler,  
sehr geehrte Damen und Herren,

für die Möglichkeit, zu oben genanntem Berichtsplan vom 9. Januar 2012 Stellung nehmen zu können, bedanken wir uns recht herzlich. Unsere Hinweise erstrecken sich auf grundlegende formale sowie spezielle fachliche Aspekte.

**Klinische Aspekte der Rolle der PET-Diagnostik**

Die PET-Diagnostik im Zusammenhang mit der Alzheimer-Demenz (AD) ist unter anderem bei folgenden Fragestellungen wichtig: Frühdiagnose, Differentialdiagnose und Ausschluss einer AD. Klinisch geht es dabei aber nicht nur um die Diagnosesicherung einer Alzheimer-Demenz, sondern auch um die Abgrenzung dieser gegenüber anderen Erkrankungen, vornehmlich der Depression (insbesondere bei präsenilen Patienten mit fortschreitenden kognitiven Funktionseinbußen unter Antidepressiva) sowie selteneren Demenzformen, die spezifischen Therapieformen zugänglich sind (z.B. Lewy-Body-Demenz, DLB).

Was die Frühdiagnose betrifft, so konnten zahlreiche Studien einen hohen Wert der FDG-PET in der Vorhersage des Auftretens der Alzheimer-Demenz bei Patienten im Stadium MCI (Mild Cognitive Impairment) zeigen. Der Diagnose MCI können verschiedene Ursachen zugrunde liegen. Einige Patienten bleiben stabil, bei manchen sind nicht-neurodegenerative Pathologien (z.B. vaskulär, Depression) Ursache der Symptomatik. Ein relevanter Anteil der Patienten zeigt aber eine Progression zur Alzheimer-Demenz oder anderen Demenzformen. Klinisch ist aufgrund der überlappenden unspezifischen Symptomatik eine Prädiktion nicht zuverlässig möglich. Das Stadium MCI ist dadurch gekennzeichnet, dass die Patienten noch in der Lage sind, den Alltagsaktivitäten nachzukommen. Eine frühe Prognose einer bevorstehenden Demenz erlaubt diesen Patienten also, ihre Lebensumstände auf diese schwerwiegende Situation rechtzeitig selbstbestimmt einzurichten. Gleiches gilt auch für die Angehörigen. Zudem erlaubt die zuverlässige frühe Prädiktion der AD die frühe Einleitung geeigneter therapeutischer Verfahren. Umgekehrt stellt der Ausschluss einer bevorstehenden Demenzerkrankung eine ausgesprochene Erleichterung für diese zumeist sehr stark beunruhigten Patienten dar. Dies ist somit als hoher diagnostischer Wert an sich bzw. patientenbezogener Nutzen zu betrachten. Nicht zuletzt ermöglicht der zuverlässige Ausschluss einer bevorstehenden AD die Suche nach anderen potentiell behandelbaren Ursachen

IQWiG\_D0601N\_vorl\_berichtsplan\_stellungnahme\_dgn\_120215.doc



der kognitiven Beschwerden (Depression? Vaskuläre Genese?) und vermeidet fälschlich eingeleitete kostenintensive und nebenwirkungsbehaftete Therapien.

In der aktuellen Metaanalyse [1] konnte gezeigt werden, dass die FDG-PET im Vergleich zur Perfusions-SPECT bzw. MRT die höchste diagnostische Testgüte bei der Vorhersage der schnellen Konversion vom MCI zur AD aufweist. Dies ist von erheblicher Bedeutung für die Selbstbestimmung des Patienten und den Arbeitsaufwand seiner Angehörigen. In der Meta-Analyse zeigten 88,9% der Älteren mit isolierten Gedächtnisstörungen (MCI), die nach durchschnittlich 28 Monaten eine Demenz entwickelt hatten, zur Baseline Alzheimer-typische FDG-PET Befunde. Andererseits hatten 84,9% der MCI-Patienten, die im Beobachtungszeitraum von durchschnittlich 28 Monaten stabile kognitive Leistungen aufwiesen, einen negativen FDG-PET Befund bzgl. neurodegenerativer Erkrankung zur Baseline [1].

Weiterhin ist die Differential-Diagnose dementieller Syndrome wichtig, beispielsweise die Unterscheidung der AD von einer Lewy-Body-Demenz. Der Einsatz von Neuroleptika, der bei der AD häufig stattfindet, aber bei der Lewy-Body-Demenz fatale Folgen haben kann (malignes Neuroleptika-induziertes Syndrom), muss vermieden werden. Dazu kann z.B. die PET-Diagnostik mit F-18-Fluoro-DOPA dienen. Ebenfalls können wiederum mittels der FDG-PET andere behandelbare Ursachen der dementiellen Symptome identifiziert werden, wie z.B. die „Pseudodemenz“ als Symptom der Depression.

### **Warum kein Technologie-Vergleich PET versus SPECT?**

Die PET ist ein zeitaufgelöstes szintigraphisches Verfahren zur Messung räumlicher Radioaktivitätsverteilungen. Es weist unabhängig vom Einsatz radioaktiver Medikamente grundsätzliche messtechnische Vorteile gegenüber herkömmlichen nuklearmedizinischen Messverfahren wie der planaren Szintigraphie oder der SPECT („Single photon emission computed tomography“) auf. Der Vergleich der PET gegenüber der SPECT ist recht einfach über technische Abbildungsparameter zu bewerkstelligen, die unter anderem in der propädeutischen Literatur zur Nuklearmedizin zu finden sind. Auch im klinischen intraindividuellen Vergleich sind die Unterschiede zwischen FDG-PET und FDG-SPECT so offensichtlich [2 (Abb. 6)], dass umfangreiche klinische Studien ethisch nicht vertretbar sind, um diesen Technologievergleich nach Ihren formalen Vorgaben durchzuführen.

### **Warum Bewertung von PET zusammen mit Radiopharmaka?**

Ihrem Berichtsplan zufolge wollen Sie die PET bzw. PET/CT zusammen mit entsprechenden Radiopharmaka bewerten. Diese Bewertung geht möglicherweise über den Beratungsantrag hinaus.

### **Warum Bewertung der diagnostisch-therapeutischen Kette statt des diagnostischen Verfahrens PET?**

Sie wollen die gesamte diagnostisch-therapeutische Kette der PET-basierten Alzheimer-Diagnostik einschließlich der damit verbundenen Therapie analysieren. Die Trennung der Einflüsse von Diagnostik und Therapie ist kaum möglich.

Die PET ist mittlerweile eine wichtige Komponente im Rahmen der Etablierung der Diagnose einer Alzheimer-Demenz (siehe unten). In vorausgegangenen Nutzenberichten Ihres Hauses (z.B. Projekt A10-06, Memantin bei Alzheimer-Demenz) konnten Sie den Nutzen von Therapieverfahren bei M. Alzheimer belegen. Da die PET-Diagnostik in die Patientenselektion einfließt, ist dieser für das Medikament Memantin erbrachte Nutzenbeleg nicht vom Nutzen der PET-Diagnostik zu trennen.

### **Problem der Diagnosesicherung der Alzheimer-Demenz**

Die sichere Diagnose des M. Alzheimer ist nur postmortal anhand einer neuropathologischen Begutachtung des Gehirns möglich. Es wird verkannt, dass die klinische Definition des M. Alzheimer mittlerweile an die PET-Diagnostik gekoppelt ist (aktuelle NINCDS-Kriterien, siehe unten). Somit ist der Referenzstandard bei diagnostischen Studien nicht unabhängig von der PET-Diagnostik zu definieren.

### **Kriterien für akzeptable Primär-Literatur und Evidenz-Synthesen**

Ihre geplante Literatur-Selektion ist im Hinblick auf das oben erwähnte Fehlen eines Goldstandards, der in vivo erhoben werden könnte, problematisch. In Abschnitt 4.2.1.2 geben Sie im letzten Absatz eine Arbeitsdefinition an. Gibt es dafür eine wissenschaftliche Quelle?

Die Bewertung der im geplanten Nutzenbericht zu betrachtenden Evidenz-Synthesen soll anhand der QUADAS-Methodik (bzw. QUADAS-2-Methodik) erfolgen. Diese Vorgehensweise ist nicht unumstritten, insbesondere beim Einsatz für die Metaanalyse von Studien mit bildgebender Diagnostik [3]. Die Bilddaten enthalten aufgrund der räumlichen Darstellung von Struktur und Funktion mehr an Information als nur das Vorhandensein oder Fehlen einer bestimmten Erkrankung. So dient z.B. die PET-Diagnostik bei dementiellen Syndromen nicht nur der Erkennung bzw. dem Ausschluss einer Alzheimer-Erkrankung, sondern auch zur Differenzierung verschiedener, teilweise seltener Demenztypen, die charakteristische Verteilungsmuster der Radiopharmaka aufweisen.

Wir befürchten daher, dass der Stand der Wissenschaft gerade bei der Alzheimer-PET-Diagnostik nicht adäquat zusammengetragen wird. Wichtige Primärliteratur bzw. Metaanalysen sind wahrscheinlich nach Ihren Maßstäben für die weitere Betrachtung inakzeptabel, weil es sich zumindest teilweise um eine retrospektive Rekrutierung der Patientenkollektive handelt oder Vergleiche mit weiteren Untersuchungsverfahren wie der MRT fehlen. Als Beispiele seien die großen Studien mit Korrelation von PET-Befunden mit den Anreicherungsmustern der Radiopharmaka FDG bzw. Amyloid-Markern zur post-mortem begutachteten Neurohistopathologie genannt. Weder diese wichtigen Studien [4,5] noch relevante Meta-Analysen zu dem Themenbereich [6 (mit Anhängen), 7] können bei dem geplanten Vorgehen erfasst werden.

Da in der Vergangenheit wichtige Literatur bei der Erstellung von Nutzenberichten übersehen wurde, bitten wir in diesem Zusammenhang, die Literaturreferenzen in aktuellen Übersichtsarbeiten als Anhalt für relevante Arbeiten zu beachten.

#### **Parallele Prüfung von Vergleichstechnologien?**

Sie planen, nur Primärliteratur zu akzeptieren, die Vergleichstechnologien in derselben Studie untersucht. Durch diese Beschränkung nehmen Sie sich die Möglichkeit, den verfügbaren wissenschaftlichen Kenntnisstand, z.B. die diagnostische Testgüte der FDG-PET oder Amyloid-PET bei den beauftragten Indikationsfeldern, zusammenzustellen. Diese Informationen wird der G-BA sehr wohl bei seiner weiteren Begutachtung benötigen, z.B. um festzustellen, ob es sich um Untersuchungsverfahren mit Potential für eine „Erprobung“ gemäß §137e SGB V handelt. Mit dem momentan geplanten Vorgehen wird diese Aussage nicht aus Ihrem Bericht abzuleiten sein.

#### **Biomarker: Unterscheidung nicht-invasiver und invasiver Verfahren**

Die Betrachtung diagnostischer Verfahren muss unter dem Aspekt der Invasivität der einzelnen Methoden erfolgen. Unter den Biomarkern gibt es neben der nicht-invasiven PET-Diagnostik die Messung von Parametern, welche aus dem Liquor bestimmt werden. Die Liquorpunktion ist ein invasives Verfahren. Somit sind PET bzw. PET/CT als nicht-invasiver Biomarker mit den invasiv aus dem Liquor gewonnenen Biomarkern nicht ohne gleichzeitige Berücksichtigung von Komplikationsraten bzw. Ablehnungs- und Misserfolgsquoten zu vergleichen.

#### **Kein direkter Nutznachweis wegen fehlender Studien zum Zusatznutzen**

Vermutlich gibt es momentan keine Literaturdaten über den Zusatznutzen der PET bzw. PET/CT im Sinne Ihrer formalen Anforderungen (Einfluss der Durchführung einer PET auf patientenrelevante Endpunkte). Umso wichtiger ist es, dass die Erkenntnisse aus der verfügbaren Literatur nach dem Prinzip der „linked evidence“ zusammengeführt werden.

Daher ist die Zusammenfassung des aktuellen Kenntnisstandes in einer (quantitativen) Evidenzsynthese für die weitere Beratungsarbeit des G-BA wünschenswert, auch im Hinblick auf die erforderlichen Informationen, falls es in Teilbereichen eine „Erprobung“ gemäß §137e SGB V geben soll. In den USA sowie anderen europäischen Gesundheitssystemen, z.B. in Italien, ist man schließlich aufgrund der verfügbaren Literatur zur Zulassung der FDG-PET bei der Diagnostik des M. Alzheimer gekommen.

#### **Gesundheitsökonomisches Problem Demenz und Depression**

Demenzerkrankungen stellen wegen der demographischen Entwicklung eine der größten gesundheitspolitischen Herausforderungen dar, denn eine zunehmende Zahl der älteren Bevölkerung erkrankt an Demenzen. Die Aufwendungen für Betreuung und Pflege der Betroffenen nehmen zu und stellen unsere Gesellschaft vor eines der größten Probleme der Zukunft.

Es ist bekannt, dass die um fünf Jahre verzögerte Manifestation einer Demenzerkrankung die Behandlungskosten um 50% reduziert [8]. Es gibt weiterhin Literatur zur Kosteneffektivität, wenn die Therapie früh begonnen wird [9].

Medizin-ökonomische Aspekte müssen daher in Ihrem Nutzenbericht Beachtung finden, sowohl für die Früherkennung einer AD als auch für die Bestätigung der AD-Diagnose.

### **Beschreibung der PET und PET/CT im Zusammenhang mit der Alzheimer-Demenz**

Die Ausführungen in Kapitel 1.6 entstammen offenbar den Texten älterer IQWiG-Produkte zum Thema PET und PET/CT. Da nunmehr ein nicht-onkologischer Indikationsbereich bearbeitet wird, ergeben sich einige Ungenauigkeiten. So ist beispielsweise die SUV-Quantifizierung der zerebralen FDG-Aufnahme unüblich. Ebenso spielt die Bildfusion mit der CT kaum eine Rolle, vielmehr ist die Koregistrierung der PET mit MRT-Bildern mitunter hilfreich. Die seit Jahren üblichen, in der täglichen Routine eingesetzten Darstellungen mittels Oberflächen-Projektionen und die Darstellung von regionalen Anreicherungsunterschieden im Vergleich zu Kontroll-Kollektiven anhand farblich kodierten z-Scores [10] müssen erwähnt werden. Wir empfehlen die Einbindung von Experten, die mit diesem nuklearmedizinischen Versorgungsstandard vertraut sind.

### **PET/CT mit diagnostischem (Volldosis-)CT gegenüber PET/CT mit (nicht-diagnostischem) Niedrigdosis-CT**

Bei der Bewertung der PET/CT sei in diesem Zusammenhang darauf hingewiesen, dass man zwischen dem **diagnostischen** Kombinationsverfahren PET/CT, bei dem die CT mit voller Strahlendosis in diagnostischer Qualität zusammen mit der PET durchgeführt wird, und der PET/CT, bei der die CT in Niedrig- bzw. Niedrigst-Dosis-Technik durchgeführt wird und lediglich der Korrektur der Strahlungsabschwächung der Positronen-Vernichtungsstrahlung dient, unterschieden werden muss. Für die PET-Akquisition des Gehirns, also auch für die Alzheimer-Diagnostik, wird sogar bei modernen PET/CT-Geräten im Wesentlichen nur die PET-Komponente genutzt. Es handelt sich also nicht – wie bei onkologischen Fragestellungen üblich – um eine PET/CT-Untersuchung mit diagnostischer CT-Komponente. Vielmehr werden auch an PET/CT-Geräten eigentlich nur PET-Untersuchungen des Gehirns durchgeführt, aber keine PET/CT-Diagnostik im herkömmlichen Sinne.

### **Pathophysiologische Hypothesen zur Alzheimer-Demenz**

Abbildung 2 in Kapitel 1.5.2 des Berichtsplans ist in dem Sinne etwas irreführend, da sie impliziert, dass auch mit dem MRT erkennbare strukturelle Veränderungen diagnostisch wertbar wären, bevor die Patienten kognitiv auffällig werden (dunkel-lila Kurve kommt zeitlich vor hell-lila Kurve und vor brauner Kurve). Dies ist erwiesenermaßen nicht der Fall. In der Publikation, aus der diese Graphik stammt [11] wird dies sogar explizit erklärt. Siehe hierzu Abbildung 3 in dem Beitrag mit der Erläuterung, dass die (mittels FDG-PET messbare) neuronale Dysfunktion den MRT-fassbaren Veränderungen vorangeht. Dass sich teilweise sogar schon strukturelle Veränderungen finden, obwohl klinische Kriterien noch nicht erfüllt sind, ist vermutlich ein Effekt von Kompensationsmechanismen. Daher ist die FDG-PET schon aus pathophysiologischer Sicht der MRT ganz klar überlegen.

Die klinische Symptomatik der Erkrankung erlaubt daher keine zuverlässige Frühdiagnose. Somit besteht die Notwendigkeit für geeignete Biomarker (Imaging).

### **Bezug auf veraltete Literatur**

Sie beziehen sich mit Ihrer Referenz 41 bzw. mit Tabelle 3 auf die NINCDS-Publikation von 1984. Mittlerweile gibt es eine Überarbeitung. Es findet sich weder ein Bezug auf die "Revised NINCDS-Criteria" [12], noch auf die neuen AD-Diagnosekriterien der Arbeitsgruppe am National Institute on Aging-Alzheimer's Association (NIA-AA) [13], in denen explizit die Nutzung von Biomarkern wie PET empfohlen wird.

### **Radiopharmaka für die Demenz-Diagnostik**

Die konkretisierten Fragestellungen umfassen die Differenzierung der AD von der DLB. Dies ist mittels PET bzw. PET/CT über die Verwendung des in Deutschland arzneimittelrechtlich zugelassenen Radiopharmakons F-18-Fluoro-DOPA möglich.

Die F-18-markierten PET-Tracer, welche für die Darstellung der Amyloid-Last entwickelt wurden und kurz vor der arzneimittelrechtlichen Zulassung stehen, werden nicht erwähnt.

In ähnlicher Weise wird man für innovative Therapien zur Vermeidung oder Behandlung von M. Alzheimer die mit PET objektivierte Amyloid-Last als Einschlusskriterium heranziehen wollen. Das Potential der Amyloid-PET für derartige Fragestellungen können Sie nur durch die Berücksichtigung weiterer Radiopharmaka erfassen.

Wir hoffen, dass wir Ihnen einige Anregungen für die umfassende Bearbeitung dieses wichtigen Themas geben konnten.

In der Vergangenheit hatte sich gezeigt, dass Ihre Autorennachfragen bei offenen Punkten in Publikationen ohne Erfolg geblieben waren. Gerne sind wir dabei behilflich, den Kontakt zu den entsprechenden Arbeitsgruppen herzustellen.

Mit freundlichen Grüßen



Prof. Dr. med. Jörg Kotzerke  
Präsident der Deutschen Gesellschaft für Nuklearmedizin (DGN)



Prof. Dr. med. Dirk Hellwig  
Vorsitzender des Ausschusses Positronen-Emissions-Tomographie (PET) der DGN



Prof. Dr. med. Mathias Schreckenberger  
Vorsitzender der Arbeitsgemeinschaft Neuronuklearmedizin der DGN



Prof. Dr. med. Osama Sabri  
Klinik und Poliklinik für Nuklearmedizin, Universitätsklinikum Leipzig



Prof. Dr. med. Alexander Drzezga  
Nuklearmedizinische Klinik, Klinikum rechts der Isar der Technischen Universität München

**Anlagen:**

- Literaturverzeichnis
- Zitierte Literatur (eingereicht in elektronischer Form mit E-Mail vom 15. Februar 2012)
- Formblatt zur schriftlichen Stellungnahme zum vorläufigen Berichtsplan
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. J. Kotzerke
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. D. Hellwig
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. M. Schreckenberger
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. O. Sabri
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. A. Drzezga



## Literaturverzeichnis

1. Yuan Y, Gu ZX, Wei WS. Fluorodeoxyglucose-Positron-Emission Tomography, Single-Photon Emission Tomography, and Structural MR Imaging for Prediction of Rapid Conversion to Alzheimer Disease in Patients with Mild Cognitive Impairment: A Meta-Analysis. *Am J Neuroradiol.* 2009 Feb;30(2):404-10.
2. Martin WH, Delbeke D et al. FDG-SPECT: Correlation with FDG-PET. *J Nucl Med* 1995; 36(6): 988-995.
3. Weber W. Is There Evidence for Evidence-Based Medical Imaging? *J Nucl Med* 2011; 52: Suppl. 2, 74S-76S
4. Silverman DHS, Small GW, Chang CY et al. Positron Emission Tomography in Evaluation of Dementia: Regional Brain Metabolism and Long-term Outcome. *JAMA* 2001; 286: 2120-2127.
5. Clark CM, Schneider JA, Bedell BJ et al. Use of Florbetapir-PET for Imaging  $\beta$ -Amyloid Pathology. *JAMA* 2011; 305: 275-283.
6. Bohnen NI, Djang DSW, Herholz K, Anzai Y, Minoshima S. Effectiveness and Safety of  $^{18}\text{F}$ -FDG PET in the Evaluation of Dementia: A Review of the Recent Literature. *J Nucl Med* 2012; 53: 59-71.  
Anmerkung: Diese Publikation enthält zwei Textverweise zu "Methods and Review Process" sowie "Supplemental table 1". Diese Verweise sind an die Originalarbeit angehängt.
7. Drzezga A. Diagnosis of Alzheimer's disease with [ $^{18}\text{F}$ ]PET in mild and asymptomatic stages. *Behav Neurol* 2009; 21:101-115.
8. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's Disease in the United States and the Public Health Impact of Delaying Disease Onset. *Am J Public Health.* 1998; 88(9): 1337-1342.
9. Getsios D, Blume S, Ishak KJ, Maclaine GDH. Cost Effectiveness of Donepezil in the Treatment of Mild to Moderate Alzheimer's Disease: A UK Evaluation Using Discrete-Event Simulation. *Pharmacoeconomics.* 2010; 28 (5): 411-427.
10. Minoshima S, Frey KA, Koeppe RA, Foster NL, Kuhl DE. A Diagnostic Approach in Alzheimer's Disease Using Three-Dimensional Stereotactic Surface Projections of Fluorine-18-FDG PET. *J Nucl Med* 1995; 36: 1238-1248.
11. Jack CR, Knopman DS, Jagust WJ et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* 2010; 9: 119-28.
12. Dubois B, Feldman HH, Jacova C et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007; 6: 734-46.
13. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia.* 2011; 7: 263-269.

## **1.4 Deutsche Krankenhaus Gesellschaft**

### **Autoren:**

Acikgöz, Özlem

Dippmann, Anja

Rüther, Klaus

Schlottmann, Nicole



EINGEGANGEN

17. Feb. 2012

DEUTSCHE  
KRANKENHAUS  
GESELLSCHAFT

Institut für Qualität und  
Wirtschaftlichkeit im Gesundheitswesen  
- Stellungnahme zum Berichtsplan <D06-01N>  
Prof. Dr. med. Jürgen Windeler  
Dillenburger Str. 27  
51105 Köln

- vorab per E-mail -

Ressort	Info	Rückprf.	Bearb.	Verf. d. B.	Ressort	Info	Rückprf.	Bearb.	Verf. d. B.
IL					GI				
AM					Kr. u. V.				
NMV					W. u. G.				
VQ					GC				
Biometrie									

Bundesverband  
der Krankenhausträger  
in der Bundesrepublik  
Deutschland

Unser Zeichen

Dez. V/Aci

Telefon

Durchwahl +49(0) 30 3 98 01 -

Telefax

+49(0) 30 3 98 01 -

Datum

15.02.2012

### Stellungnahme der Deutschen Krankenhausgesellschaft zum vorläufigen Berichtsplan D06-01N „Positronenemissionstomographie (PET und PET/CT) bei Alzheimer Demenz“

Sehr geehrter Herr Professor Windeler,  
sehr geehrte Damen und Herren,

für die Möglichkeit, zu oben genanntem Berichtsplan Stellung nehmen zu können, bedanken wir uns recht herzlich.

Der Gemeinsamen Bundesausschuss (G-BA) hat das Institut gem. § 139a Abs. 3, 1. SGB V am 21.12.2006 mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes der PET-Diagnostik bei Alzheimer Demenz beauftragt. Der spätere Abschlussbericht soll eine Grundlage für die Beratungen des G-BA zur Bewertung der PET bei o. g. Indikation darstellen.

Der G-BA hat im Rahmen seiner Methodenbewertung zu prüfen, ob Untersuchungs- oder Behandlungsmethoden, die zu Lasten der gesetzlichen Krankenkassen angewendet werden oder angewendet werden sollen, für eine ausreichende, zweckmäßige und wirtschaftliche Versorgung der Versicherten unter Berücksichtigung des allgemein anerkannten Standes der medizinischen Erkenntnisse erforderlich sind.

Die Deutsche Krankenhausgesellschaft (DKG) als eine Trägerorganisation des G-BA nimmt zum vorliegenden vorläufigen Berichtsplan des Instituts wie folgt Stellung:

1. Wir begrüßen es, dass Aussagen verschiedener Leitlinien Eingang in Ihren Berichtsplan finden. In der 2008 veröffentlichten S2e-Leitlinie der Deutschen Gesellschaft für Allgemeinmedizin (DEGAM) (Ihre Referenz 19) wird zwar festgestellt, dass die PET-Untersuchung nur für sehr spezielle klinische Fragestellungen und die Forschung von Bedeutung sind und die PET „nach heutigem Wissensstand bzgl. der Demenzdiagnose keinen zusätzlichen Nutzen hat“. Wie Sie aber richtig darstellen, folgt die Leitlinie dabei allerdings den Empfehlungen der American Academy of Neurology (Ihre Referenz 42), die bereits 2001 veröffentlicht worden ist und daher nur den Stand der Dinge von vor 10

Postfach 120555 · D-10595 Berlin  
Wegelystraße 3 · D-10523 Berlin

Telefon +49(0) 30 39801-0  
Telefax +49(0) 30 39801-3000

dkg@mail.dkgev.de · www.dkgev.de

Rue d'Arlon 50  
B-1000 Brüssel

Telefon +32(0) 2 282-0594  
Telefax +32(0) 2 282-0598

europa@dkgev.de

Commerzbank AG  
Konto-Nummer

UST-IdNr. DE119355528



Jahren darstellen kann. Dabei ist auch festzuhalten, dass die Autoren dieser Empfehlungen eine höhere diagnostische Güte von PET feststellen (Referenzen 73,75,76 aus Ref42-Berichtsplan), jedoch aufgrund fehlender weiterführender Studien zu diesem Zeitpunkt keine klare Aussage zum tatsächlichen Nutzen machen können. Die DEGAM-Leitlinie sollte laut Angabe in der Leitlinie bis Ende 2011 revidiert werden. Es ist hier also in absehbarer Zeit mit einer aktuelleren Version zu rechnen, die dann bezüglich der PET auf neueren Erkenntnissen beruht.

2. Wie im Berichtsplan dargestellt ist das Ziel Ihrer Untersuchung die Prüfung des patientenrelevanten Nutzens oder Schadens. Wir möchten Sie nochmals darauf hinweisen, dass der Auftrag an das IQWiG darin besteht, „die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zu PET, PET/CT“ durchzuführen (G-BA-Beschluss vom 19.12.2006). Die Beurteilung eines etwaigen Schadens durch das IQWiG ist hingegen nicht Gegenstand der Beauftragung. Der G-BA wird nach Prüfung der publizierten Daten feststellen, mit welchem Nutzen und mit welchen Risiken die in Frage stehende Methode ggf. verbunden ist. Wir bitten, dieses im Berichtsplan zu berücksichtigen.
3. Wir begrüßen, dass bei der Beurteilung der Studien die gesundheitsbezogene Lebensqualität als patientenrelevanter Endpunkt mit einbezogen werden soll. Allerdings beschränken Sie sich hinsichtlich des patientenrelevanten Nutzens erneut auf die Stufe 5 nach Fryback & Thornbury. Die in der Auftragskonkretisierung des G-BA benannten zentralen Fragen beziehen sich aber auch auf die Frage, inwieweit sich eine frühere und präzisere Diagnosestellung auf die Behandlung und das weitere Management des Patienten und der Patientin auswirkt (4. Stufe). PET ermöglicht die differentialdiagnostische Abklärung von unterschiedlichen Demenzformen. Wie Sie in Ihrem Berichtsplan richtig feststellen, kann die Differenzierung aufgrund klinischer Kriterien zu einer Fehldiagnose führen. Dies ist beispielweise bei der Abgrenzung zur vaskulären Demenz von entscheidender Bedeutung, da hieraus eine andere therapeutische Konsequenz resultieren würde. Wir möchten in Bezug auf den Begriff der therapeutisch-diagnostischen Einheit gemäß Auftragskonkretisierung nochmals darauf hinweisen, dass eine *direkte* Beeinflussung der von Ihnen angeführten patientenrelevanten Zielgrößen durch ein diagnostisches Verfahren nicht möglich ist. Eine diagnostische Methode kann anhand ihrer Befunde nur *indirekt* in Form einer therapeutischen Konsequenz patientenrelevante Endpunkte beeinflussen.
4. Bei Ihrer Bewertung diagnostischer Tests berufen Sie sich auf die „Mehrzahl der westlichen Gesundheitssysteme“, die „Überlegungen zu den Konsequenzen und dem Nutzen und Schaden diagnostischer Tests für Arzt, Patient und Gesellschaft in Betracht zieht“. Dieser Sichtweise habe sich die Verfahrensordnung des G-BA angeschlossen. Es bleibt unklar, was das IQWiG mit dem Ausdruck „westliches Gesundheitssystem“ meint. Die Verfahrensordnung des G-BA gibt unter § 11 , Abs. 2 eindeutig Auskunft über die Klassifizierung und Bewertung von Unterlagen, die eine evidenzbasierte Beurteilung des medizinischen Wissensstandes ermöglichen sollen, die Beurteilung von Nutzen oder Schaden ist allerdings nicht Gegenstand dieses von Ihnen zitierten Paragraphen.
5. Darüber hinaus zitieren und beziehen Sie sich auf eine Verfahrensordnung, die nicht mehr gültig ist. Eine aktuelle Version <sup>1</sup>, in Kraft seit August 2011, ist online verfügbar.
6. Bei der Bewertung der PET als diagnostischem Test beziehen Sie sich u. a. auf das Klassifikationsschema von Fryback und Thornbury, welches die Evidenzstufen von Testgüte und Nutzen hierarchisch integriert (Abschnitt 4, „Methoden“, S. 12). Allerdings beziehen Sie sich hinsichtlich diagnostischer und prognostischer Güte lediglich auf die Stufe 2 und zu den patientenrelevanten Endpunkten auf die Stufe 5 des Klassifikationsschemas nach

Fryback und Thornbury. Angesichts der differenzierten, inhaltlich unmittelbar aufeinander aufbauenden Evidenzsteigerung wäre es sinnvoll und stringent, nicht nur Teilaspekte des Schemas, sondern die gesamte Evidenz von Stufe 2 bis 5 (einschließlich Stufe 3 und 4) in die Betrachtungen einzubeziehen und darzustellen.

7. Wir begrüßen die realistische Darstellung des Paradigmenwechsels bei der Bewertung diagnostischer Verfahren hin zu vergleichenden Studien mit patientenrelevanten Endpunkten, der sich langsam vollzieht und können nachvollziehen, dass nicht-randomisierte, kontrollierte Interventionsstudien möglicherweise Hinweise auf die Ausgestaltung zukünftiger aussagekräftigerer Studien erlauben. Allerdings halten wir die von Ihnen vorab anhand der Ein- und Ausschlusskriterien erfolgte Einschränkung der zu berücksichtigenden Primärliteratur auf Studien der Evidenzstufen I und II nicht für sachgerecht und nicht für konform mit den Vorgaben der Verfahrensordnung des G-BA. Daher bitten wir Sie, alle vergleichenden Studientypen zur Bewertung einzubeziehen, d.h. auch Studien auf Stufe III gemäß Verfahrensordnung des G-BA. Zu dieser Problematik haben wir bereits mehrfach im Rahmen von Stellungnahmen u. a. zur allgemeinen Methodik des Instituts hingewiesen und diesen Aspekt in der Auftragskonkretisierung (Juni 2008) bereits thematisiert. Die Einschränkungen führen gerade auch in diesem Fall dazu, dass ein Großteil wissenschaftlicher Erkenntnisse von vornherein ausgeschlossen bleibt. Damit kann insbesondere die „Darstellung und Bewertung des aktuellen medizinischen Wissensstandes“, die ausdrücklich Gegenstand des Auftrages vom 21.12.2006 ist, nicht sachgerecht erfolgen. Die Darstellungen des vorgelegten Berichtsplanes als Ausgangspunkt nehmend, wird der G-BA hier voraussichtlich wieder selbständig weitere Erkenntnisse ermitteln müssen.
8. Die von Ihnen geplante Verwendung einer Relevanzschwelle von 0,2 ist international keineswegs unumstritten. Nicht kontrovers ist hingegen, dass die Relevanzschwelle vom Kontext der in diesem Fall klinischen Situation abhängig ist. Die Festlegung einer Relevanzschwelle im Falle fehlender skalenspezifisch validierter bzw. etablierter Relevanzkriterien auf eine standardisierte Mittelwertdifferenz (SMD) von 0.2 ist willkürlich und durch die zitierte Literatur nicht hinreichend belegt. Insbesondere vor dem Hintergrund, dass die im Bericht behandelte Erkrankung – die Alzheimer-Demenz - potenziell lebensbedrohlich ist, ist der o.g. Grenzwert für die SMD nicht akzeptabel. Da zudem die Entscheidung über Relevanz und Irrelevanz ohnehin vom G-BA getroffen werden muss und nicht Gegenstand der Literaturdarstellung sein sollte, fordern wir Sie auf, die Mittelwertdifferenzen zu benennen und auf die Anwendung jedweder Schwellenkriterien zu verzichten.
9. Die Metaanalysen, die keine bedeutsame Heterogenität in dem von Ihnen definierten Sinne aufweisen, sollten statt mit den von Ihnen in erster Linie geplanten Modellen mit zufälligen Effekten mit solchen mit festen Effekten durchgeführt werden.
10. Wir begrüßen es, dass Sie Patientenvertreter bei der Erstellung des Berichtsplans eingebunden haben. Sie stellen in Aussicht, dass zudem externe Sachverständige in die Bearbeitung des Projekts eingebunden werden. Wir regen in Bezug auf das vorliegende Bewertungsverfahren an, bereits ab Erstellung des Berichtsplanes neurologische, neurowissenschaftliche und nuklearmedizinische Expertise einzubeziehen.
11. Gemäß den *Methoden 4.0* Ihres Instituts hoffen wir, dass sich im endgültigen Berichtsplan eine geschlechtergerechte Sprache wiederfinden wird.

Wir hoffen, dass wir Ihnen mit unseren Anmerkungen einige Anregungen für die Überarbeitung des Berichtsplans und die weitere Auftragsbearbeitung geben konnten. Im Übrigen behalten wir uns vor, ggf. ergänzende Aspekte in das weitere Verfahren einzubringen.

Mit freundlichen Grüßen  
Der Hauptgeschäftsführer  
Im Auftrag



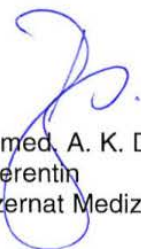
Dr. med. N. Schlottmann  
Geschäftsführerin  
Dezernat Medizin



Ö. Acikgöz  
Referentin  
Dezernat Medizin



Prof. Dr. med. K. Rüter  
Referent  
Dezernat Medizin



Dr. med. A. K. Dippmann  
Referentin  
Dezernat Medizin

*Quellenangaben:*

<sup>1</sup> [http://www.g-ba.de/downloads/62-492-548/VerfO\\_2011-01-20\\_2011-08-04.pdf](http://www.g-ba.de/downloads/62-492-548/VerfO_2011-01-20_2011-08-04.pdf)

## **1.5 Lilly Deutschland GmbH**

### **Autoren:**

Boess, Frank

Lieb, Matthias

Stahl, Oliver

**Stellungnahme zum  
vorläufigen Berichtsplan [D06-01N]**

*Lilly*  
EINGEGANGEN  
21. Feb. 2012

**Stellungnahme der  
Lilly Deutschland GmbH  
zum vorläufigen Berichtsplan der Bewertung D06-01N  
„Positronenemissionstomographie (PET) und PET/CT bei Alzheimer Demenz“  
vom 09. Januar 2012**

**A. GESETZLICHE GRUNDLAGE UND G-BA-AUFTRAG ..... 3**

**B. METHODEN ..... 3**

B.1 ZIELE DER UNTERSUCHUNG ..... 3

B.2 EINSCHLUSSKRITERIEN ZUR BEWERTUNG DES PATIENTENRELEVANTEN NUTZENS UND SCHADENS DER PET BZW. PET/CT ANHAND VON PRIMÄRSTUDIEN ..... 3

B.2.1 Populationen ..... 3

B.2.2 Prüf- und Vergleichsintervention – fehlende Technologievergleiche ..... 4

B.2.3 Auswahl von Studientypen ..... 4

B.2.4 Referenztest ..... 5

B.2.5 Nachbeobachtungszeitraum des Referenztests ..... 5

**ABKÜRZUNGSVERZEICHNIS ..... 2**

Ressort	Info	Rückspr.	Bearb.	Verfeilb	Ressort	Info	Rückspr.	Bearb.	Verfeilb
IL					GI				
AM					Komm				
NMV					VW				
VQ					GÖ				
Biometrie									

### **Abkürzungsverzeichnis**

AD	Alzheimer Demenz
CCT	Controlled Clinical Trial (kontrollierte klinische Studie)
CERAD	The Consortium to Establish a Registry for Alzheimer's Disease
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders
FDG	Fluor-18-2-Fluoro-2-Deoxy-D-Glukose
G-BA	Gemeinsamer Bundesausschuss
ICD-10	International Statistical Classification of Diseases and Related Health Problems
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MCI	Mild Cognitive Impairment
MRI	Magnetic Resonance Imaging (Magnetresonanztomographie)
NINCDS-ADRDA	National Institute of Neurologic, Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association Work Group
PET	Positronenemissionstomographie
PET/CT	Integration von PET und CT in ein Untersuchungsgerät
PIB	Pittsburgh compound-B
RCT	Randomisierte, kontrollierte Studie
VerfO	Verfahrensordnung

## A. Gesetzliche Grundlage und G-BA-Auftrag

Die wesentliche gesetzliche Grundlage für den Auftrag des G-BA, bildet § 139 a Abs. 3 Nr. 1 SGB V. Entsprechend § 139 a Abs. 4 SGB V gilt bei der Auftragsbearbeitung folgender Grundsatz:

„Das Institut hat zu gewährleisten, dass die Bewertung des medizinischen Nutzens nach den international anerkannten Standards der evidenzbasierten Medizin und die ökonomische Bewertung nach den hierfür maßgeblichen international anerkannten Standards, insbesondere der Gesundheitsökonomie erfolgt. Es hat in regelmäßigen Abständen über die Arbeitsprozesse und -ergebnisse einschließlich der Grundlagen für die Entscheidungsfindung öffentlich zu berichten.“

Auf Basis dieser Grundlage hat der Gemeinsame Bundesausschuss (G-BA) am 19. Dezember 2006 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der nachfolgenden Bewertung beauftragt:

„[...] die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zu PET, PET/CT bei folgenden Indikationen und in der angegebenen Reihenfolge durchführen.“

Nachdem für einige Tumorerkrankungen diese Bewertung durchgeführt wurde, veröffentlichte das IQWiG am 09.01.2012 den vorläufigen Berichtsplan zum Krankheitsbild der Alzheimer Demenz.

## B. Methoden

### B.1 Ziele der Untersuchung

Die vorliegende Untersuchung soll prüfen, welcher patientenrelevante Nutzen oder Schaden mit der bildgebenden Diagnostik mittels PET bzw. PET/CT im Vergleich oder in Ergänzung zur konventionellen Diagnostik bei der (früheren) Detektion einer Alzheimer Demenz oder einem Therapiemonitoring verbunden ist.

Außerdem soll untersucht werden, ob sich durch den Einsatz der PET bzw. PET/CT die prognostische bzw. diagnostische Güte der Diagnosestrategie verbessert.

### B.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Primärstudien

#### B.2.1 Populationen

In Abschnitt 4.1.2.1 werden die einzuschließenden Populationen wie folgt beschrieben:

„Studien zu Patienten mit erhöhtem Risiko für bzw. Verdacht auf eine Alzheimer Demenz und Patienten mit einer diagnostizierten Alzheimer Demenz werden eingeschlossen.“

Die Einschränkung, nur Studien zu Patienten mit erhöhtem Risiko für bzw. Verdacht auf eine Alzheimer Demenz und Patienten mit einer diagnostizierten Alzheimer Demenz einzuschließen, greift unserer Ansicht nach zu kurz.

Da es explizit Ziel des Berichts ist, die diagnostische und prognostische Güte der PET bzw. PET/CT für die „Diagnostische Abgrenzung einer Alzheimer Demenz von anderen demenziellen Erkrankungen (z. B. vaskuläre Demenz)“ zu untersuchen<sup>1</sup>, muss die Recherche nach Primärstudien auch diese Erkrankungen enthalten. Wir schlagen vor, die

<sup>1</sup> Vorläufiger Berichtsplan D06-01N, S.13



Primärstudienrecherche auf die Differentialdiagnostik aller demenziellen Krankheiten auszuweiten.

Es wird aus der gegenwärtigen Beschreibung der einzuschließenden Populationen nicht deutlich, ob Studien mit MCI Patienten in den zu berücksichtigten Studien eingeschlossen sind oder nicht.

Da als weiteres explizites Ziel des Berichts die „Frühzeitige Abgrenzung von Patienten mit Risiko für Alzheimer Demenz (z. B. Patienten mit MCI), die eine AD entwickeln“ untersucht werden soll, empfehlen wir, dass die „einzuschließenden Populationen“ auch Studien mit MCI-Patienten umfassen.

### **B.2.2 Prüf- und Vergleichsintervention – fehlende Technologievergleiche**

Laut Seite 12 des vorläufigen Berichtsplans soll im Zuge der Bewertung von PET bzw. PET/CT zur Diagnose der AD keine Eingrenzung auf bestimmte Tracer erfolgen.

Diese Formulierung ist unklar und sollte präzisiert und begründet werden. Ein Vergleich verschiedener Tracer ist unbedingt erforderlich, da die Qualität und der Informationswert eines PET Scans absolut abhängig von der eingesetzten Markierungssubstanz sind.

Nicht nachvollziehbar ist deshalb der Schritt, keinen Technologievergleich zwischen verschiedenen verfügbaren oder in naher Zukunft verfügbaren Tracern vorzunehmen. So erlauben neuere Substanzen, wie das im vorläufigen Berichtsplan genannte 11C-PIB oder das sich in der klinischen Erprobung befindliche Florbetapir F-18, eine Darstellung der Amyloid-Ablagerungen, was mit dem momentan in der klinischen Praxis eingesetzten FDG nicht möglich ist.<sup>2 3</sup>

Ein Vergleich „diagnostische Strategie mit PET“ versus „diagnostische Strategie ohne PET“ ohne Differenzierung zwischen den verwendeten Tracern lässt Unterschiede, hinsichtlich Art der erhaltenen Informationen sowie Spezifität und Sensitivität der verschiedenen Tracer außer Acht.

Um diesen methodischen Fehler zu vermeiden, ist ein gesonderter Technologievergleich aller zur Diagnose der Alzheimer Demenz eingesetzten Tracer unter Berücksichtigung der verschiedenen gemessenen Parameter (zerebrale Amyloidablagerung mit 11C-PIB und anderen Amyloid-bindenden Tracern versus Änderung des Glukosestoffwechsels in verschiedenen Hirnregionen mit FDG) nötig.

### **B.2.3 Auswahl von Studientypen**

Bei der Auswahl der einbezogenen Studientypen ist nicht nachvollziehbar, warum nur RCTs zur Beurteilung einer diagnostischen Methode primär Betrachtung finden sollen. In der klinischen Ökonomik sind die wichtigsten Kriterien zur Beurteilung diagnostischer Tests Sensitivität, Spezifität, sowie die Prävalenz. Aus diesen Größen lässt sich der positiv und negativ prädiktive Wert berechnen, welcher die Validität und damit die Güte eines Tests beschreibt.<sup>4</sup>

Aus diesem Grund sind CCTs, welche die Sensitivität und Spezifität eines Tests

---

<sup>2</sup> Grimmer, Drzezga, Kurz; Amyloiddarstellung mittels Positronenemissionstomographie (2010) in der Nervenarzt 05/2010

<sup>3</sup> Vorläufiger Berichtsplan D06-01N; siehe Sperling et al., Referenz 28

<sup>4</sup> Du Pruel, Muttray; Validität von Diagnostik- und Screeningstudien (2011) in Porzolt, Grundlagen der klinischen Ökonomik, S.47.



untersuchen, keinesfalls als Studien der zweiten Kategorie zu klassifizieren, sondern als Standard für Studien zur Bewertung von Diagnostika.<sup>5 6</sup>

#### **B.2.4 Referenztest**

Wenn, wie in Abschnitt 4.2.1.2 behauptet, kein allgemein akzeptierter Goldstandard für die Diagnose der AD existiert, bedarf die Auswahl des Referenzstandards einer gut nachvollziehbaren Begründung. Aus den Angaben in Abschnitt 4.2.1.2. geht nicht klar hervor, wie die Güte eines Referenztests bewertet oder die Zusammensetzung eines Referenzstandards definiert werden soll, und ob alle Studien mit unterschiedlichen studienspezifischen Referenzstandards berücksichtigt werden sollen.

Ein möglicher Ansatz zur Beurteilung der Spezifität und Sensitivität diagnostischer Methoden ist die histopathologische Untersuchung. Diese ist in den auf Seite 7 des vorläufigen Berichtsplans zitierten NINCDS-ADRDA Kriterien erwähnt. Diesbezüglich liegen auch neue Empfehlungen (des National Institute of Aging) vor.<sup>7</sup>

Darüber hinaus existieren weitere wissenschaftlich fundierte Methoden zur histopathologischen Bewertung einer Alzheimer Diagnose, wie z.B. die CERAD Kriterien.<sup>8</sup>

#### **B.2.5 Nachbeobachtungszeitraum des Referenztests**

Im Abschnitt 4.2.1.2. des vorläufigen Berichtsplans wird im Rahmen der Diskussion eines potentiellen Referenzstandards vorgeschlagen, die Erhebung des Krankheitsstatus nach einer Nachbeobachtung von 18 Monaten durchzuführen.

Wie begründet das IQWiG die Wahl einer 18-monatigen Nachbeobachtungszeit? Durch die Begrenzung auf eine bestimmte Nachbeobachtungszeit werden möglicherweise qualitativ hochwertige Studien ausgeschlossen.

Bei der Beurteilung der diagnostischen Güte eines Tests sind in jedem Fall auch Studien ohne Nachbeobachtungszeit einzuschließen.

Bei der Beurteilung der prognostischen Güte sind alle Studien mit unterschiedlichen Nachbeobachtungszeiten in die Bewertung einzuschließen.

In Abschnitt 4.2.1.4 **Den systematischen Übersichten zugrunde liegende Studientypen** wird folgendes Kriterium zur Diagnostischen Güte vorgeschlagen: Prospektive Querschnittstudien, ggf. mit klinischer Nachbeobachtung, durchgeführt in der Anwendungssituation (gemäß Phase 3 nach Köbberling et al. [63]).

Dieses Kriterium ist nicht ausreichend, da Studien gemäß Phase 2 nach Köbberling et al. (diagnostische Genauigkeit bei bekanntem Krankheitsstatus) mit möglicherweise entscheidenden Informationen zur Bewertung der Diagnostischen Güte nicht berücksichtigt werden.

---

<sup>5</sup> Wehberg, Sauerbrei, Schuhmacher; Diagnosestudien:Wertigkeit der Sonographie bei der Differenzierung von gut- und bösartigen Brusttumoren bei Patientinnen mit klinischen Symptomen (2008) in Methodik klinischer Studien 3. Auflage, S.319ff.

<sup>6</sup> Bortz, Lienert; Kurzgefasste Statistik für die klinische Forschung 3.Auflage (2008), S.261ff.

<sup>7</sup> McKhann et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease in Alzheimers Dement. 2011 May;7(3):263-9.

<sup>8</sup> Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, Vogel FS, Hughes JP, van Belle G, Berg L. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease in Neurology. 1991 Apr;41(4):479-86.

In Abschnitt **4.2.1.4 Den systematischen Übersichten zugrunde liegende Studientypen** wird postuliert: „Für die Interpretation der zuvor genannten prognostischen Zielgrößen ist ein Follow-up von mindestens 18 Monaten entscheidend. Für die Beurteilung medizinischer Interventionen bei Alzheimer Demenz, die die Krankheitsprogression reduzieren oder verlangsamen sollen, wird sowohl in einer Richtlinie der European Medicines Agency (EMA) [71] als auch in einem Konsensbeschluss einer europäischen Arbeitsgruppe [72] ein Studienzeitraum von 18 Monaten für notwendig erachtet. Um die Krankheitsprogression im Rahmen von diagnostischen oder prognostischen Studien zu erfassen (siehe auch 4.2.1.2), erscheint eine Nachbeobachtung von 18 Monaten daher sinnvoll.“

Diese Argumentation ist nicht nachvollziehbar. Warum soll für die Beurteilung einer diagnostischen Methode der gleiche Maßstab wie zur Beurteilung einer medizinischen Intervention, die die Krankheitsprogression reduzieren oder verlangsamen soll, eingesetzt werden? Die EMA stellt in dem vom Berichtsplan zitierten Dokument klar, daß zur Zeit keine Behandlungsmethoden existieren, die diese Bedingungen erfüllen. (Up to now no clinical trial has led to a successful claim of disease modification in dementing conditions.) Der Berichtsplan geht also von einem Dokument aus, das sich auf die Bewertung zukünftiger, jetzt noch in der Entwicklung befindlicher Medikamente bezieht und legt diese Maßstäbe zur Bewertung derzeitiger Diagnosemethoden an.

Im zitierten Dokument der EMA wird eine Studiendauer von 6 Monaten als Effizienzkriterium für eine symptomatische Verbesserung als notwendig erachtet. („Controlled clinical trials aimed at demonstrating short term improvement in AD should last at least 6 months.“).

Mehrere Studien haben gezeigt, dass ein früher Einsatz von bereits zugelassenen Medikamenten, wie z.B. Acetylcholinesterase-Inhibitoren, schon nach 6 Wochen eine deutliche Verbesserung verschiedener kognitiver Parameter zur Folge haben kann.<sup>9 10</sup> Eine frühzeitige korrekte Diagnose könnte zu einer frühen adäquaten Behandlung führen.

---

<sup>9</sup> Seltzer et al.; Efficacy of Donepezil in Early-Stage Alzheimer Disease (2004) in Arch.Neurol. Vol. 61 S.1852-1856.

<sup>10</sup> Winblad et al.; 3-Year Study of Donepezil Therapy in Alzheimer's Disease: Effects of Early and Continuous Therapy (2006) in Dementia and Geriatric Cognitive Disorders Vol 21 S.353-363