

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
20. Mai 2010 (20.05.2010)

(10) Internationale Veröffentlichungsnummer  
WO 2010/055043 A2

- (51) Internationale Patentklassifikation:  
G06F 19/00 (2006.01)
- (21) Internationales Aktenzeichen: PCT/EP2009/064951
- (22) Internationales Anmeldedatum:  
11. November 2009 (11.11.2009)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:  
01765/08 12. November 2008 (12.11.2008) CH
- (72) Erfinder; und
- (71) Anmelder : SPIEGEL, René [CH/CH]; Oberalpstrasse  
55, CH-4054 Basel (CH).
- (74) Anwälte: MÜLLER, Christoph et al.; Friedtalweg 5,  
CH-9500 Wil (CH).
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für  
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,  
AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY,  
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM,  
GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN,  
KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA,  
MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG,  
NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC,  
SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR,  
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für  
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,  
GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,  
ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU,  
TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE,  
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,  
LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI,  
SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,  
GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu  
veröffentlichen nach Erhalt des Berichts (Regel 48 Absatz  
2 Buchstabe g)

(54) Title: METHOD FOR PERFORMING CLINICAL STUDIES AND METHOD FOR ESTABLISHING A PROGNOSIS MODEL FOR CLINICAL STUDIES

(54) Bezeichnung : METHODE ZUR DURCHFÜHRUNG KLINISCHER STUDIEN UND METHODE ZUR ETABLIERUNG EINES PROGNOSEMODELLS FÜR KLINISCHE STUDIEN

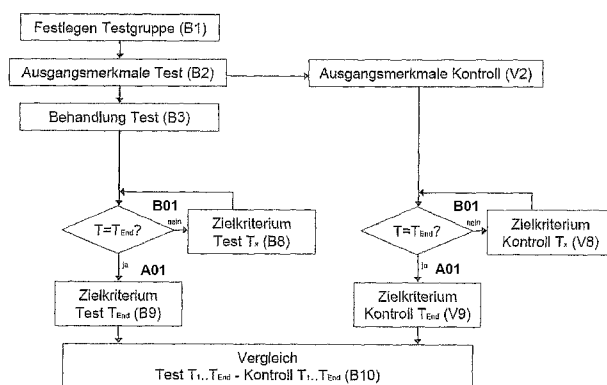


Fig. 1

- A01 yes B01 no
- B1 Test group determination
- B2 Test initial characteristics
- B3 Test treatment
- B8 Test target criterion Tx
- B9 Test target criterion TEnd
- B10 Test T1..TEnd - control T1..TEnd comparison
- V2 Control initial characteristics
- V8 Control target criterion Tx
- V9 Control target criterion TEnd

(57) Abstract: The invention relates to a method for planning and performing clinical studies, wherein during and after expiration of a predetermined time period, values of a predefined target criterion of control persons of a virtual control group are compared with values of the predefined target criterion of test persons of a test group that has received a treatment to be examined. The values of the predefined target criterion of the control persons of the virtual control group during and after expiration of the predetermined time period are determined from values of initial characteristics associated with the control persons.

(57) Zusammenfassung: Eine Methode zur Planung und Durchführung klinischer Studien, wobei Werte eines vordefinierten Zielkriteriums von Kontrollpersonen einer virtuellen Kontrollgruppe mit Werten des vordefinierten Zielkriteriums von Testpersonen einer Testgruppe, welche eine zu untersuchende Behandlung erhalten hat, während und nach Ablauf eines vorherbestimmten Zeitraumes verglichen werden. Die Werte des vordefinierten Zielkriteriums der Kontrollpersonen der virtuellen Kontrollgruppe während und nach Ablauf des vorherbestimmten Zeitraumes werden dabei aus den Kontrollpersonen zugeordneten Werten von Ausgangsmerkmalen bestimmt.

WO 2010/055043 A2

## Methoden zur Durchführung klinischer Studien und Methoden zur Etablierung eines Prognosemodells für klinische Studien

Die vorliegende Erfindung betrifft eine Methode zur Durchführung klinischer Studien mit neuen Wirkprinzipien, insbesondere Medikamenten, für Patienten mit progredienten, letztlich fatalen Krankheiten, beispielsweise der Alzheimer Krankheit, sowie eine Methode zur Etablierung eines Prognosemodells für klinische Studien in solchen Indikationen.

Einer klinischen, d.h. symptomatischen Phase, beispielsweise der Alzheimer Krankheit (AD, Alzheimer Disease), die sich in fortgeschrittenen Stadien als Demenz manifestiert, geht eine jahrelange präklinische Phase voraus, in der sich Teile des Gehirns langsam und zunehmend pathologisch verändern, ohne dass diese Veränderungen im Alltag für den Betroffenen zunächst spürbar bzw. für andere Personen ersichtlich werden.

Zahlreiche Verlaufsstudien an älteren und alten Personen, in denen anhand verschiedenartiger Testmethoden und beispielsweise genetischer, biochemischer, elektrophysiologischer, bildgebender und neuropsychologischer Indikatoren früheste Veränderungen im Gehirn und Frühindikatoren der Alzheimer Krankheit erfassbar gemacht werden, sind im Gang. Von diesen Studien liegen erste Resultate vor und es sind in den nächsten Jahren weitere Erkenntnisse über den natürlichen (spontanen) Verlauf dieser Krankheit zu erwarten. Beispiel einer solchen Studie ist die ADNI (Alzheimer Disease Neuroimaging Initiative; <http://www.ADNI-info.org>), eine von 56 Instituten in den USA getragene longitudinale Untersuchung an grossen Stichproben von gesunden Alterspersonen, Patienten mit sog. Mild Cognitive Impairment sowie Patienten mit AD.

Bereits heute bestehen Aussichten, dass das pathologische Geschehen der AD nicht nur frühzeitig erfasst, sondern durch rechtzeitige pharmakologische Interventionen auch verlangsamt oder aufgehalten werden kann, so dass die klinische, d.h. symptomatische Phase der Krankheit verzögert, abgeschwächt oder ganz verhindert werden kann.

Unter DMDs (Disease Modifying Drugs) versteht man Substanzen bzw. Wirkprinzipien, die sich für die Behandlung präklinischer (präsymptomatischer) oder auch sehr früher, milder Stadien neurodegenerativer Krankheiten, im Fall der AD auch als MCI (Mild Cognitive Impairment) bezeichnet, eignen. Solche Substanzen befinden sich in vielen pharmazeutischen Firmen und akademischen Labors weltweit in Entwicklung.

DMDs sollen auf den Krankheitsprozess im Gehirn einwirken und haben im präklinischen Stadium zunächst keine klinisch erkennbare, in milden Stadien einer Krankheit, beispielsweise des MCI im Fall der AD, vermutlich keine unmittelbar eintretende Wirkung auf die Symptome; ihre Wirkung wird erst mit längerer Anwendungsdauer (Jahre) auch klinisch erkennbar. „Erkennbar“ heisst traditionell, anhand von Vergleichen zwischen behandelten und unbehandelten oder mit Placebo behandelten Probanden erfass- und quantifizierbar.

Die Anwendungsdauer von DMDs beträgt im Falle der AD bzw. des MCI nach heutigen Annahmen mindestens 18 Monate, bevor mit einiger Sicherheit eine Wirkung auf den Verlauf des präklinischen bzw. des sehr frühen klinischen Stadiums der Krankheit feststellbar wird (Vellas B, Andrine S, Sampaio C et al (2007): Disease-modifying trials in Alzheimer's disease: a European task force consensus. LancetNeurol 6: 56-62). Klinische Studien der Phasen 2 und 3 mit neuen DMDs müssen also eine Behandlungsdauer

von mindestens 18 Monaten pro Patient vorsehen, um - falls positiv - den Anspruch eines „disease modifying effect“ erheben zu können. In der Phase 2 nehmen normalerweise mehrere Hundert, in der Phase 3 mehrere Tausend Patienten an den klinischen Studien teil.

Obwohl die Zielpopulation für DMDs, namentlich Patienten mit präklinischer oder sehr früher AD, entweder keine oder nur sehr milde klinische Symptome (MCI) aufweist, ist eine Anwendung von Placebo in klinischen Prüfungen bei Hunderten oder Tausenden dieser Patienten angesichts des vorhersehbar malignen Verlaufs solcher Krankheiten ethisch schwer vertretbar (Council for International Organizations of Medical Sciences: International Ethical Guidelines for Biomedical Research Involving Human Subjects; Geneva 2002). Andererseits verlangen Zulassungsbehörden wie FDA oder EMEA heute doppelblinde, Placebo-kontrollierte Studien zum Wirkungsnachweis vieler Medikamente, insbesondere auch von Neuro-Psychopharmaka, zu denen die DMDs gehören.

Ethisch besser vertretbare Studiendesigns, die zum Schutz der Patienten bei fehlender Wirkung oder schlechter Verträglichkeit einer Versuchssubstanz bzw. von Placebo eine so genannte Escape Clause - häufig mit Umstellung des Patienten auf eine symptomatische Behandlung - enthalten, stellen keine befriedigende Lösung dar. Sie bringen - ähnlich wie Placebo-kontrollierte Studien allgemein - einen Selektionsbias a priori mit sich, da nicht alle Patienten bzw. Probanden bereit sind, an einer Studie teilzunehmen, in der sie möglicherweise über lange Zeit ein pharmakologisch unwirksames Scheinmedikament erhalten. Ausserdem fehlt in vielen Fällen, beispielsweise der AD, bisher der Nachweis, dass symptomatisch wirksame Medikamente, wie beispielsweise Cholinesterase-Hemmer, in präklinischen Stadien der Krankheit eine günstige Wirkung zeigen und somit als Escape Medication von

Nutzen sein können (Aisen PS: Treatment for MCI: Is the evidence sufficient? Neurology 70: 2020-2021, 2008).

Auch so genannte Add-on Designs, d.h. die Testung neuer Präparate an Patienten, die zusätzlich eine in der betreffenden Indikation als wirksam anerkannte Grundmedikation erhalten, sind problematisch. Erstens weil in vielen Fällen, zum Beispiel bei der AD, der Nachweis fehlt, dass sich die heutigen, vorwiegend symptomatisch wirksamen Medikamente zur Behandlung präklinischer Stadien der betreffenden Krankheit eignen (Aisen 2008, siehe oben), zweitens weil bei allen Add-on Studien das grundsätzliche Problem pharmakokinetischer und pharmakodynamischer Wechselwirkungen besteht.

Die "International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use (ICH)" erwähnt in ihrer Publikation vom 20. Juli 2000 zur "ICH Harmonised Tripartite Guideline" unter der Nummer E10 „Choice Of Control Group And Related Issues In Clinical Trials" die Möglichkeit, eine externe oder historische Kontrollgruppe für die Durchführung eines klinischen Tests zu verwenden. Hierbei wird jedoch darauf hingewiesen, dass wenn die Kontrollpersonen nicht aus derselben Gruppe wie die Testpersonen ausgewählt sind, die Resultate nur mit Vorsicht zu verwenden sind. Ebenso sollte darauf geachtet werden, dass bei der Verwendung von externen (historischen) Kontrollpersonen diese Kontrollpersonen und auch die Umgebungsbedingungen sehr genau mit denjenigen der Testpersonen übereinstimmen.

Es bestehen also widersprüchliche Anforderungen hinsichtlich des Einsatzes von (Placebo-)Kontrollgruppen von Zulassungsbehörden einerseits und den Patienten, ihren Vertretern und den ethischen Richtlinien andererseits. Auch durch die Einführung einer Escape Clause oder Anwendung eines Add-on Designs sind diese nicht in Einklang zu bringen. Vor diesem Hintergrund ist es eine Aufgabe

der Erfindung, die Nachteile der bekannten Studiendesigns zu vermeiden und insbesondere eine Methode zur Durchführung klinischer Studien, beispielsweise mit DMDs gegen präklinische Stadien der AD, zur Verfügung zu stellen, welche

- a) für alle an der Studie teilnehmenden Probanden die Behandlung mit einem als aussichtsreich betrachteten Wirkprinzip, beispielsweise einem neuartigen Medikament, sicherstellt,
- b) gleichzeitig die Nachteile, insbesondere Placebo-Effekte, Spontanverbesserungen und andere Zeiteffekte, einer offenen, unkontrollierten Studie vermeidet,
- c) den besonders bei Langzeitstudien mit gleichzeitig geführten Placebo-Gruppen relevanten Selektionsbias verhindert bzw. reduziert, und
- d) kostengünstiger gegenüber bestehenden Studien ist.

Diese Aufgabe wird durch die in den unabhängigen Patentansprüchen definierten Verfahren gelöst. Weitere Ausführungsformen ergeben sich aus den abhängigen Patentansprüchen.

Bei einer erfindungsgemässen Methode zur Durchführung klinischer Studien wird zuerst eine der Hypothese der zu testenden Behandlung und dem heutigen Wissen über die in Frage stehende Krankheit, beispielsweise die AD, entsprechende Testgruppe von Testpersonen in präklinischen (MCI) oder sehr milden Stadien der Krankheit festgelegt. Anschliessend an die Bestimmung der relevanten (medizinischen, neurophysiologischen, neuropsychologischen, biochemischen etc.) Ausgangswerte wird den Testpersonen eine Behandlung, insbesondere eine medikamentöse Behandlung, über einen vorher bestimmten Zeitraum verabreicht. Die Werte

mindestens eines vordefinierten Zielkriteriums der Testpersonen der Testgruppe werden bis zum Ablauf des vorher bestimmten Zeitraumes mindestens einmal bestimmt. Hier und im Folgenden wird jeweils von einem Zielkriterium gesprochen, wobei selbstverständlich jeweils mehrere Zielkriterien genauso bestimmt werden können.

Beim Zielkriterium kann es sich dabei insbesondere um ein neuropsychologisches Messinstrument und/oder ein Verhaltenskriterium und/oder auch um neurobiologische oder -physiologische Variablen handeln.

Parallel dazu werden die Werte des mindestens einen vordefinierten Zielkriteriums von Kontrollpersonen, welche die Behandlung nicht erhalten haben, bis zum Ablauf des vorherbestimmten Zeitraumes mindestens einmal bestimmt. Gemäss der Erfindung gibt es aber keine realen Kontrollpersonen. Die Werte der Zielkriterien werden nicht an einer realen Kontrollperson bestimmt, sondern mittels eines Prognosemodells (Algorithmus) aus den den Kontrollpersonen zugeordneten Werten von Ausgangsmerkmalen bestimmt. Die Kontrollpersonen sind virtuell und bilden eine virtuelle Kontrollgruppe. Schliesslich werden die bis zum Ablauf des vorherbestimmten Zeitraumes bestimmten Werte des vordefinierten Zielkriteriums der virtuellen Kontrollpersonen mit denjenigen der realen Testpersonen verglichen.

Dank dem erfindungsgemässen Verfahren kann eine Vergleichsstudie gemacht werden, ohne dass reale von der Krankheit betroffene Personen auf eine Behandlung mit dem zu untersuchenden Wirkprinzip verzichten müssen.

Bis zum Ablauf des vorherbestimmten Zeitraumes kann der Wert des Zielkriteriums der Testpersonen nicht nur vor Beginn und nach

Ablauf des vorbestimmten Zeitraumes der Behandlung, sondern auch mehrmals während des Zeitraumes der Behandlungsdauer bestimmt werden. Ebenso kann auch bei der Kontrollgruppe die Berechnung des Wertes des Zielkriteriums bis zum Ablauf der vorbestimmten Zeitdauer auf der Basis des Prognosemodells mehrmals erfolgen. Idealerweise stimmt die Anzahl der Bestimmungen der Werte der Zielkriterien der Testpersonen mit der Anzahl der Bestimmungen bei den Kontrollpersonen überein.

Parallel zur virtuellen Kontrollgruppe können zusätzliche reale Kontrollpersonen einer realen Kontrollgruppe über eine ethisch akzeptable, begrenzte, initiale Zeitspanne eine Placebo-Behandlung erhalten. Auf diese Weise kann den in der ICH-Publikation vom 20. Juli 2000 (vgl. Seite 4) geäußerten Bedenken gegenüber Vergleichen mit externen und historischen Kontrollgruppen Rechnung getragen werden. Die Werte des mindestens einen vordefinierten Zielkriteriums der realen Kontrollpersonen der realen Kontrollgruppe werden dann nach Ablauf der begrenzten, initialen Zeitspanne bestimmt und mit den nach Ablauf dieser Zeitspanne bestimmten Werten des mindestens einen vordefinierten Zielkriteriums der virtuellen Kontrollpersonen der virtuellen Kontrollgruppe verglichen.

Es versteht sich von selbst, dass nicht nur vor und nach Ablauf der begrenzten, initialen Zeitspanne die Werte des mindestens einen vordefinierten Zielkriteriums bestimmt werden können, sondern auch eine oder mehrere Bestimmungen während dieser Zeitspanne möglich sind.

Die begrenzte, initiale Zeitspanne dauert vorzugsweise 6 Monate, andere Zeitspannen sind jedoch ebenfalls denkbar.

Basierend auf oben genanntem Vergleich der Werte des Zielkriteriums der virtuellen Kontrollpersonen mit den Werten des Zielkriteriums der zusätzlichen realen Kontrollpersonen kann das Prognosemodell, falls nötig, korrigiert werden.

Die Werte der Ausgangsmerkmale der Kontrollpersonen entsprechen bevorzugt den realen Werten der Ausgangsmerkmale der Testpersonen. Somit weist die virtuelle Kontrollgruppe zum Startzeitpunkt der klinischen Studie identische Werte der Ausgangsmerkmale auf wie die Testgruppe.

Somit wird ermöglicht, dass basierend auf Werten der Ausgangsmerkmale der Testpersonen eine virtuelle Kontrollgruppe gebildet wird. Anhand dieser Kontrollgruppe lassen sich dann die Werte der vorherbestimmten Zielkriterien der virtuellen Kontrollpersonen im Verlaufe der vorbestimmten Zeitdauer mit den realen Werten der vorherbestimmten Zielkriterien der Testpersonen vergleichen. Alternativ kann die Gruppe der Testpersonen auch aufgeteilt werden, so dass beispielsweise die eine Hälfte der Testpersonen eine Behandlung A und die zweite Hälfte der Testpersonen eine Behandlung B erhält. Die Kontrollgruppe kann aus der ungeteilten Testgruppe gebildet werden, oder es können zwei oder mehr Kontrollgruppen gebildet werden. Die Werte der Zielkriterien im Verlaufe der vorbestimmten Zeitdauer lassen sich unabhängig davon mit denjenigen der virtuellen Kontrollpersonen vergleichen. Es ist aber auch denkbar, dass die Zielkriterien der virtuellen Kontrollpersonen ausgehend von einem Personenpool bestimmt werden, der nicht oder nur teilweise mit der Testgruppe übereinstimmt.

Die Testgruppe sollte eine der Fragestellung der Studie, insbesondere der Art der Krankheit und der Intervention angepasste Anzahl Testpersonen umfassen. Diese Anzahl kann abhängig sein

von der Häufigkeit des Vorkommens der interessierenden Krankheit, aber auch von der erwarteten Wirksamkeit (effect size) und/oder der Verträglichkeit des untersuchten Wirkprinzips. Ebenfalls kann diese Anzahl auch von den entsprechenden Zulassungsbehörden vorgeschrieben werden. Natürlich ist auch jede andere Anzahl möglich.

Ein Verhältnis von 1:1 der Anzahl Kontrollpersonen zur Anzahl Testpersonen findet sich in vielen klinischen Studien mit neuen therapeutischen Wirkprinzipien. Es sind aber durchaus andere Verhältnisse, beispielsweise mit mehreren Kontrollpersonen je Testperson, möglich, zum Beispiel wenn aufgrund der Ausgangsmerkmale die Zielkriterien parallel mit mehr als einem Rechenmodell ermittelt werden, oder wenn weitere virtuelle Kontrollpersonen auf der Grundlage von Daten von weiteren realen Personen berücksichtigt werden, deren Ausgangsmerkmale vergleichbar mit den Ausgangsmerkmalen der Testperson sind. Dies erlaubt ein Ausloten möglicher Toleranzen des Prognosemodells und/oder der Ungenauigkeiten der Bestimmung der Zielkriterien bzw. von Ausgangsmerkmalen, welche auf das Prognosemodell einen Einfluss haben können.

Bei den zu untersuchenden therapeutischen Prinzipien handelt es sich vorzugsweise um eine Behandlung von Krankheiten mit voraussichtlich malignem Verlauf, beispielsweise Demenzen, insbesondere der AD. Selbstverständlich können auch Studien zur Behandlung anderer Krankheiten, bei denen eine längere Placebo-Verabreichung ethisch und medizinisch problematisch ist, mit dieser Methode durchgeführt werden.

Bei dem vordefinierten Zielkriterium kann es sich auch um eine Mehrzahl von vordefinierten Zielkriterien handeln, welche bestimmt und verglichen werden. Ausschlaggebend für die Art und

Anzahl Zielvariablen ist die Fragestellung der Studie, insbesondere der von der untersuchten Intervention erwartete Effekt.

Im Fall von neurodegenerativen Krankheiten, beispielsweise der AD, werden die Zielkriterien typischerweise ausgewählt aus folgenden Gruppen:

- Parameter des zentralen Nervensystems, insbesondere neuropsychologische Leistungsmerkmale, wie sie beispielsweise mit der CERAD-Batterie (Morris JC, Heyman A, Mohs RC et al (1989): The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): I: clinical and neuropsychological assessment of Alzheimer's disease. Neurology 39: 1159-1169) und ähnlichen Instrumenten erfasst werden,
- Daten von Tests gemäss der ADNI Datenbank (Alzheimer's Disease Neuroimaging Initiative, [www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)), nachfolgend ADNI Datenbank genannt,
- Subjektive und objektive Demenzsymptome, z.B. Behinderungen im Alltag, die sich mit ADL- und IADL-Skalen (Galasko D, Bennett D, Sano M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. Alzheimer Dis Assoc Disord. 1997;11(Suppl 2):S33-S39.) erfassen und quantifizieren lassen,
- Veränderungen im Gehirn, insbesondere Schrumpfungen des Hippocampus insgesamt und/oder Schrumpfungen bestimmter Abschnitte des Hippocampus oder anderer Gehirnbereiche, eine vaskuläre Pathologie und ähnliche Veränderungen,
- Marker in der Cerebro-Spinal Flüssigkeit, wie beispielsweise Abeta, Tau, Quotienten von Abeta und/oder andere Indikatoren,
- das Vorhandensein pathologischer Proteine, wie Abeta und/oder Tau in anderen Körperflüssigkeiten,
- weitere biochemische und genetische Marker,
- psychiatrische und motorische Auffälligkeiten.

Ein in Studien an präsymptomatischen Patienten häufig verwendetes Zielkriterium ist „time to event“, d.h. die Zeit in Tagen, Wochen oder Monaten, welche vom Eintritt des Patienten in die Studie bis zum erstmaligen Auftreten eines relevanten klinischen Ereignisses, beispielsweise eines Herzinfarktes oder Schlaganfalls, im Fall von MCI vorzugsweise der erstmaligen Diagnose „Demenz“, durch die zuständigen Kliniker verstreicht.

Die Ausgangsmerkmale umfassen typischerweise mindestens ein Zielkriterium und/oder sind im Fall der AD ausgewählt aus der Gruppe:

- Alter des Patienten,
- das Vorhandensein des E-4 Allels des Apolipoprotein-Gens und/oder anderer genetischer bzw. biochemischer Marker,
- die in der Kindheit oder der Adoleszenz erfasste Intelligenz oder ein Äquivalent dazu, die Ausbildungsdauer und/oder die berufliche Tätigkeit,
- Life-style Faktoren im mittleren Alter, beispielsweise Ernährung, körperliche Aktivität, Sport und/oder andere Aktivitäten,
- Klinische Merkmale und Diagnosen wie Adipositas, Hypertonie, Diabetes und/oder Hyperlipidämie, die auch in einem klinischen Gesamtscore zusammengefasst sein können,
- Traumatische Ereignisse wie Hirnschlag und Unfälle, die die Hirnfunktion zeitweise oder dauernd beeinflussen können
- Funktionelle Parameter des menschlichen Herzens oder anderen Organen des menschlichen Körpers,
- Charaktermerkmale wie Neurotizismus und/oder Gewissenhaftigkeit.

Weitere Ausgangsmerkmale sind auch denkbar.

Eine erfindungsgemässe Methode zur Etablierung eines Prognosemodells für klinische Studien umfasst die Schritte:

- a) Sammeln von repräsentativen Resultaten aus klinischen Studien mit unbehandelten Patienten betreffend den Verlauf einer interessierenden Erkrankung,
- b) Auswertung dieser Resultate,
- c) Ermittlung mindestens eines Zielkriteriums, das für den Verlauf der interessierenden Erkrankung klinisch relevant ist und in einer nachfolgenden Studie als vordefiniertes Zielkriterium verwendet werden kann,
- d) Bestimmung der Veränderung des Wertes dieses mindestens einen Zielkriteriums für eine bestimmte Anzahl von Patienten und über eine bestimmte Zeitdauer,

so dass anhand dieser Bestimmung im Schritt d) eine Prognose für die zeitliche Veränderung des Wertes dieses Zielkriteriums basierend auf Werten von Ausgangsmerkmalen bei einem Patienten oder bei Patientengruppen erstellt werden kann.

Zur Etablierung des Prognosemodells sollten Resultate einer ausreichenden Zahl von Personen, die während einer genügend langen Dauer der präklinischen Phase der betreffenden Krankheit beobachtet wurden, gesammelt und ausgewertet werden. Diese Zahl kann abhängig von der Art und Häufigkeit der interessierenden Krankheit variieren, so dass möglicherweise mit einer eher geringen oder grösseren Anzahl von Resultaten ein Prognosemodell etabliert werden kann. Ein geeigneter, umfangreicher Datensatz steht aus der Eingangs genannten ADNI (Alzheimer Disease Neuroimaging Initiative) zur Verfügung.

Es kann je nach Krankheit eine Mehrzahl von Zielkriterien, aber auch nur ein Zielkriterien ermittelt werden.

Die Zielkriterien und die Ausgangsmerkmale sind dabei gleich definiert, wie vorab beschrieben für die Methode zur Durchführung klinischer Studien.

Obwohl im vorliegenden Fall die Erfindung mehrheitlich anhand der AD gezeigt wurde, ist die Methode keinesfalls darauf eingeschränkt, sondern gilt auch für die Durchführung klinischer Studien und die Etablierung eines Prognosemodells für klinische Studien anderer Krankheiten.

Die Erfindung wird nachfolgend anhand von Ausführungsbeispielen und Figuren näher erläutert. Es zeigen:

**Figur 1:** ein Flussdiagramm einer Methode zur Durchführung klinischer Studien,

**Figur 2:** ein Flussdiagramm einer alternativen Methode zur Durchführung klinischer Studien mit einer initialen, zeitlich begrenzten Placebo-Kontrollgruppe,

**Figuren 3a-3d:** eine graphische Darstellung von Scores der neuropsychologischen Testbatterie aus der ADNI Datenbank im Verlauf der Zeit in Abhängigkeit von 4 verschiedenen Ausgangsmerkmalen, ermittelt anhand eines ersten Rechenbeispiels,

**Figuren 4a-4c:** die Verteilungen der beobachteten und simulierten Daten basierend auf einem ersten Rechenbeispiel in zwei Histogrammen und in QQ-Plots,

**Figuren 5a-5f:** eine graphische Darstellung von Scores der neuropsychologischen Testbatterie aus der ADNI Datenbank im Verlauf der Zeit in Abhängigkeit von 6 verschiedenen Ausgangsmerkmalen, ermittelt anhand eines zweiten Rechenbeispiels und

**Figuren 6a-6c:** die Verteilungen der beobachteten und simulierten Daten basierend auf einem zweiten Rechenbeispiel in zwei Histogrammen und in QQ-Plots,

In **Figur 1** ist ein Flussdiagramm dargestellt, welches eine erfindungsgemässe Methode zur Durchführung klinischer Studien zeigt. In einem ersten Schritt B1 wird die Testgruppe der Testpersonen festgelegt, welche eine Behandlung zur interessierenden Erkrankung erhalten soll. Die Ausgangsmerkmale der Testpersonen werden im Schritt B2 bestimmt. Basierend auf diesen Ausgangsmerkmalen der Testpersonen werden nun Ausgangsmerkmale einer virtuellen Gruppe von Kontrollpersonen im Schritt V2 definiert. Eine Bestimmung der Kontrollpersonen basierend auf anderen Ausgangsmerkmalen ist ebenfalls denkbar.

Die Testpersonen erhalten nun während eines vorherbestimmten Zeitraumes bis zum Zeitpunkt  $T_{\text{End}}$  die entsprechende Behandlung B3. Während dieser Zeit wird im Schritt B8 in regelmässigem Abstand der Wert mindestens eines Zielkriteriums bestimmt, so dass eine Reihe von Werten zu korrespondierenden Zeiten  $T_x$  entsteht.

Gleichzeitig wird ebenfalls in regelmässigen Abständen im Schritt V8 anhand der Ausgangsmerkmale der Kontrollpersonen mit-

tels eines Prognosemodells (Algorithmus) der Wert mindestens eines Zielkriteriums der Kontrollpersonen errechnet. Es entsteht ebenfalls eine Reihe von Werten, welche mit den entsprechenden Zeiten  $T_x$  korrespondieren.

Nach Ablauf des vorherbestimmten Zeitraumes, bzw. zum Zeitpunkt  $T_{\text{End}}$  werden die Zielkriterien sowohl der Testpersonen als auch der Kontrollpersonen in den Schritten B9 und V9 nochmals bestimmt bzw. berechnet.

Im letzten Schritt B10 werden dann die Werte der Zielkriterien zu den Zeitpunkten  $T_1..T_{\text{End}}$  der Testpersonen mit den berechneten Werten der virtuellen Kontrollpersonen verglichen und die Studienresultate können ausgewertet werden. Hierbei kommen die für Vergleichsstudien zwischen zwei oder mehreren Parallelgruppen geeigneten statistischen Verfahren (z.B. t-Test, Varianzanalyse mit Adjustierung, gemischte Modelle, logistische Regression, GEE-Modelle, Cox-PH-Regression) zur Anwendung. Das Testniveau ist nach Methoden für sequentielle Verfahren zu adjustieren.

**Figur 2** zeigt in einem Flussdiagramm eine modifizierte Methode zur Durchführung klinischer Studien. Neben der in Figur 1 beschriebenen Test- und virtuellen Kontrollgruppe wird im Schritt P1 auch eine Placebogruppe festgelegt. Unter einer Placebogruppe wird eine reale, zeitlich begrenzte Kontrollgruppe verstanden, welche über eine initiale Zeitspanne, bis zum Ablauf von  $T_{\text{placebo}}$ , eine Placebo-Behandlung erhält.

Die Ausgangsmerkmale einer virtuellen Gruppe von Kontrollpersonen wird im Schritt V2 basierend auf den Ausgangsmerkmalen der Testpersonen, welche im Schritt B2 bestimmt wurden, und den Ausgangsmerkmalen der Placebogruppe, aus Schritt P2, bestimmt. Die

Ausgangsmerkmale der Kontrollpersonen können auch anders bestimmt werden.

Während der Behandlung der Testgruppe in Schritt B3, erhält die Placebogruppe über die initiale Zeitspanne bis zum Erreichen von  $T_{\text{placebo}}$  entsprechend eine Placebo-Behandlung P3. Die Zeitspanne  $T_{\text{placebo}}$  ist dabei kleiner als der vorherbestimmte Zeitraum bis zum Zeitpunkt  $T_{\text{End}}$ .

In regelmässigem Abstand wird in nun in den Schritten V4 und P4 jeweils der Wert mindestens eines Zielkriteriums bestimmt, so dass je eine Reihe von Werten zu korrespondierenden Zeitpunkten  $T_x$  entsteht. Hierbei ist zu beachten, dass die Werte der Kontrollgruppe mittels eines Prognosemodells aus den Werten der Ausgangsmerkmale berechnet werden.

Nach Ablauf von  $T_{\text{placebo}}$  werden in den Schritten V5 bzw. V6 nochmals die Werte des entsprechenden Zielkriteriums bestimmt, bevor im Schritt V6 diese Werte der Zielkriterien  $T_1..T_{\text{placebo}}$  der Kontrollpersonen mit denjenigen der Placebogruppe verglichen werden. Anschliessend wird entschieden, ob eine Korrektur des Prognosemodells nötig ist, oder ob die Werte eine ausreichende Übereinstimmung aufweisen. Falls eine Korrektur benötigt wird, wird im Schritt P7 das Prognosemodell entsprechend angepasst und die Werte der Zielkriterien  $T_1..T_{\text{placebo}}$  der Kontrollpersonen werden nochmals berechnet. Nach Ablauf der vorbestimmten Placebophase können die Probanden dieser Gruppe ebenfalls das untersuchte Wirkprinzip (Medikament) erhalten. Die in dieser Gruppe erhobenen Messwerte können separat statistisch ausgewertet werden, sind aber nicht zwingend Teil des Vergleichs zwischen Test- und virtueller Kontrollgruppe.

Im Schritt V8 werden anschliessend in regelmässigen Abständen bis zum Ablauf des vorherbestimmten Zeitraumes, bis zum Zeitpunkt  $T_{\text{End}}$ , die Werte des Zielkriteriums weiter berechnet.

Parallel zu den Schritten V4-V8 bzw. P3-P5 erhält die Testgruppe die entsprechende, zu untersuchende Behandlung B3. Auch hier werden in regelmässigen Abständen, idealerweise zu denselben Zeitpunkten wie in den Schritten V5 und V8 die Werte mindestens eines Zielkriteriums der Testpersonen bestimmt, siehe Schritt B8.

Nach Ablauf des vorherbestimmten Zeitraumes, bzw. zum Zeitpunkt  $T_{\text{End}}$  werden die Zielkriterien sowohl der Testpersonen als auch der Kontrollpersonen in den Schritten B9 und V9 nochmals bestimmt bzw. berechnet.

Im letzten Schritt B10 werden dann die Werte der Zielkriterien zu den Zeitpunkten  $T_1..T_{\text{End}}$  der Testpersonen mit den berechneten Werten der virtuellen Kontrollpersonen verglichen und die Studienresultate können ausgewertet werden. Hierbei kommen die für Vergleichsstudien zwischen zwei oder mehreren Parallelgruppen geeigneten statistischen Verfahren (z.B. t-Test, Varianzanalyse mit Adjustierung, gemischte Modelle, logistische Regression, GEE-Modelle, Cox-PH-Regression) zur Anwendung.

Typische Zielkriterien können sein

- Endpunkte (end points), d.h. zuvor festgelegte Werte von Tests, Skalen oder auch biologischen Messgrössen, welche für die untersuchte Krankheit als wesentlich angesehen werden;
- Slopes, d.h. Merkmale von Verlaufskurven, welche die Entwicklung der untersuchten Krankheit in der Zeit charakterisieren;
- Time to event (vgl. Seite 10), d.h. die Zeit, die vom Eintritt der Patienten in die Studie bis zum Eintreten

von klinisch als relevant betrachteten Ereignissen verstreicht.

Zum rechnerischen Ermitteln der Zielkriterien der Kontrollgruppe werden geeignete Prognosemodelle verwendet. Auswirkungen von einzelnen Faktoren (d.h. Ausgangsmerkmalen) auf den Verlauf von AD sowie Modelle zur quantitativen Schätzung von Risiken, an AD zu erkranken, sind bekannt. Die vorliegende Erfindung kann als mögliche Ausgangspunkte Modelle gemäss den Erkenntnissen von Kivipelto (Prognostic models for cognitive impairment and mid-life risk, S4-03-04), Wilson (Personality and risk of cognitive impairment in old age, S4-03-01), Bennett (Brain reserve: the epidemiological perspective, PL-05-01), alle aus Abstracts zu „Alzheimer's Association International Conference on Alzheimer's disease“, July 26-31, 2008, Chicago Illinois, The Journal of the Alzheimer's Association, Vol. 4, Issue 4, Suppl. 2, July 2008) verwenden.

Die Zielkriterien können typischerweise, wie in den zwei folgenden Rechenbeispielen mit Daten aus der ADNI Datenbank gezeigt, ermittelt werden.

Gemäss einem ersten Modell werden bei Patienten mit MCI Veränderungen des globalen Scores der ADNI Neuropsychologischen Testbatterie („Neurobat“) mit der Zeit geschätzt, unter Adjustierung für demographische Variablen und cognitive Variablen zu Studienbeginn (Prädiktoren bei Baseline=bl). Die Veränderungen werden als Steigungen (Slopes) bestimmt.

Zum Endpunkt wird als Zielvariable der Globale Score der Neurobat (Variable: nbatglob) ermittelt.

Als Faktor innerhalb der Patienten wird die Zeit berücksichtigt, gemessen in Halbjahren (Variable `nviscode` (0 = bl, 1 = 6 mon, 2 = 12 mon, ... 6 = 36 mon)).

Die folgenden Prädiktoren werden als Ausgangsmerkmale berücksichtigt:

- Geschlecht (Variable `ptgender`, 1 = male, 2 = female)
- Ausbildung (Variable `pteducation`, in Jahren )
- Alter (Variable `ptage` , in Jahren)
- BMI zu Studienbeginn, (Variable `vsbmicat` , kategorisiert als normal (<25, reference), hoch/high (25-30) und fettleibig/obese (>30)
- Modifizierter Hachinski Score zu Studienbeginn (Variable `hmscore`)
- Quotient  $\text{Abeta}_{1-42} / \text{Total tau}$  zu Studienbeginn (Variable `be-taovtau`)
- Anzahl ApoE4 Gene (Variable `napgen`, 0-2)
- Total Score des Functional Assessment Questionnaire (FAQ) zu Studienbeginn (Variable `faqtotal`, max = 50)
- MMS (Mini Mental State) zu Studienbeginn (Variable `mmscore`)
- ADAScog (Alzheimer Disease Assessment Scale cognitive part) modifizierter Total Score zu Studienbeginn (Variable `totalmodbl`, max = 85)
- Globaler Score der Neurobat zu Studienbeginn (Variable: `nbatglobbl`)

Alle Berechnungen wurden in R durchgeführt (R Development Core Team (2007)).

In einem ersten Schritt wurde mit einem Startmodell mit der Zeit, allen Prädiktorvariablen und allen Interaktionen mit der Zeit als festen Effekten und mit zufälligen Effekten für die Patienten gearbeitet. Dies erlaubt es, für jeden Patienten eine zufällige Abweichung vom Mittelwert, der aus den festen Effekten geschätzt wird, zu ermitteln. Diese Abweichung wird als Zufallsvariable behandelt. Ihre Standardabweichung wird neben der Standardabweichung der Residuen geschätzt. Darüber hinaus lässt das Modell für jeden Patienten eine zufällige Abweichung der Steigung von der über alle Patienten ermittelten Steigung zu. Auch zu diesen Abweichungen wird die Standardabweichung geschätzt.

Die Zerebro-spinale Flüssigkeit (CSF) geht als Quotient von Abetal-42 über totales Tau ein. Dieser Quotient hat sich in anderen Studien bewährt.

Konfidenzintervalle werden für die Standardabweichung und die Korrelation der zufälligen Effekte wiedergegeben. Konfidenzintervalle der festen Effekte können aus der Parametertabelle leicht ermittelt werden und sind von sekundärem Interesse, da hier Waldtests vorliegen.

Quadratische Effekte der Zeit lassen sich nicht modellieren, weil zu viele Prädiktoren im Modell sind. Alternative Analysen haben nur einen schwachen quadratischen Effekt gezeigt, der erst im dritten Jahr deutlicher wurde. Gegenwärtig werden daher keine quadratischen Effekte der Zeit analysiert.

Nur etwa 50% der MCI-Patienten gemäss der ADNI Datenbank hatten CSF-Werte zu Studienbeginn. Bei einzelnen Patienten fehlten andere Baseline Daten. Daher wurden in dem ersten Rechenbeispiel nur 189 Patienten in die Analysen einbezogen.

Linear mixed-effects model fit by REML

Data: MCIba\_0  
 AIC BIC logLik  
 651.1486 775.2715 -298.5743

Random effects:

Formula: ~1 + nviscode | rid  
 Structure: General positive-definite, Log-Cholesky parametrization  
 StdDev Corr  
 (Intercept) 0.18518363 (Intr)  
 nviscode 0.07060092 0.514  
 Residual 0.25075384

Fixed effects: nbatglob ~ vsbmicat2 + nviscode + hmscore + betaovtau + napgen + pteducacat + ptage + ptgender + faqtotal + mmscore + totalmodbl + nbatglobbl + nviscode:vsbmicat2 + nviscode:hmscore + nviscode:betaovtau + nviscode:napgen + nviscode:pteducacat + nviscode:mmscore + nviscode:totalmodbl + nviscode:nbatglobbl

	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.8560546	0.5964127	557	1.435339	0.1518
vsbmicat2high	-0.0715724	0.0553691	176	-1.292641	0.1978
vsbmicat2obese	-0.0054378	0.0758238	176	-0.071716	0.9429
nviscode	-0.0885050	0.1734763	557	-0.510185	0.6101
hmscore	0.0315156	0.0340266	176	0.926203	0.3556
betaovtau	0.0042881	0.0176467	176	0.242998	0.8083
napgen	0.0232389	0.0408216	176	0.569280	0.5699
pteducacat	-0.0003740	0.0087864	176	-0.042563	0.9661

ptage	-0.0058720	0.0032846	176	-1.787750	0.0755
ptgender	-0.0502078	0.0511771	176	-0.981060	0.3279
faqttotal	-0.0093692	0.0051685	176	-1.812738	0.0716
mmscore	-0.0052058	0.0156676	176	-0.332269	0.7401
totalmodbl	-0.0021895	0.0048945	176	-0.447336	0.6552
nbatglobbl	0.9686973	0.0541730	176	17.881546	0.0000
vsbmicat2high:nviscode	0.0166253	0.0196038	557	0.848068	0.3968
vsbmicat2obese:nviscode	0.0467125	0.0274983	557	1.698740	0.0899
nviscode:hmscore	0.0046822	0.0124132	557	0.377198	0.7062
nviscode:betaovtau	0.0153686	0.0062597	557	2.455182	0.0144
nviscode:napgen	-0.0242458	0.0143613	557	-1.688268	0.0919
nviscode:pteducat	-0.0013945	0.0031599	557	-0.441323	0.6592
nviscode:mmscore	0.0019046	0.0056013	557	0.340025	0.7340
nviscode:totalmodbl	-0.0026544	0.0017326	557	-1.532055	0.1261
nviscode:nbatglobbl	0.0232797	0.0197777	557	1.177067	0.2397

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.4769035	-0.5375011	0.0116638	0.5243880	2.8531692

Number of Observations: 756

Number of Groups: 189

Random Effects:

Level: rid	lower	est.	upper
sd((Intercept))	0.1249367	0.18518363	0.27448291
sd(nviscode)	0.0506308	0.07060092	0.09844777
cor((Intercept),nviscode)	-0.5154828	0.51433737	0.93632619

Within-group standard error:

lower	est.	upper
0.2336142	0.2507538	0.2691510

Die folgenden Interaktionen trugen nicht zur Vorhersage bei und wurden in einem schrittweisen Verfahren eliminiert:

- nviscode:pteducat
- nviscode:mmscore
- nviscode:hmscore
- nviscode:napgen
- nviscode:nbatglobbl

Außerdem wurden die Kategorien "normal" und "hoch" des BMI zur Kategorie "not.obese" zusammengefasst, weil zwischen diesen beiden Kategorien keine Unterschiede in den Parameterschätzungen erkennbar waren.

Das schrittweise Eliminationsverfahren schloss anschließend die folgenden Terme (Interaktion und Variablen) aus:

nviscode:vsbmicat3

mmscore

apgen

pteducation

ptgender

Hachinski Score

Age

BMI (not obese versus obese)

Aus den zuvor beschriebenen vereinfachenden Schritten ergibt sich folgendes einfacheres Modell:

Linear mixed-effects model fit by REML

Data: MCIba\_0

	AIC	BIC	logLik
	539.4807	594.8896	-257.7404

Random effects:

Formula: ~1 + nviscode | rid

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.18163785	(Intr)
nviscode	0.06979345	0.56
Residual	0.25115344	

Fixed effects: nbatglob ~ nviscode + betaovtau + faqttotal + totalmodbl + nbatglobbl + nviscode:betaovtau + nviscode:totalmodbl

	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.1983170	0.09375083	564	2.115363	0.0348
nviscode	-0.0716350	0.03346308	564	-2.140717	0.0327
betaovtau	0.0032773	0.01535459	184	0.213440	0.8312
faqttotal	-0.0101478	0.00507821	184	-1.998296	0.0472
totalmodbl	0.0005787	0.00463211	184	0.124926	0.9007
nbatglobbl	0.9997337	0.04373146	184	22.860741	0.0000
nviscode:betaovtau	0.0222897	0.00548279	564	4.065402	0.0001
nviscode:totalmodbl	-0.0039768	0.00146246	564	-2.719227	0.0067

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.54951632	-0.53529203	0.01501314	0.52888491	2.89043920

Number of Observations: 756

Number of Groups: 189

Random Effects:

Level: rid	lower	est.	upper
sd((Intercept))	0.12361124	0.18163785	0.26690380
sd(nviscode)	0.05061054	0.06979345	0.09624724
cor((Intercept),nviscode)	-0.50343814	0.55983715	0.94874790

Within-group standard error:

lower	est.	upper
0.2342006	0.2511534	0.2693335

Interaktion von Abeta<sub>1-42</sub> /Total tau mit der Zeit bedeutet: Die (negative) Steigung mit der Zeit wird abgeschwächt (i.e. weniger steil) für Patienten mit höherem Abeta<sub>1-42</sub> beziehungsweise niedrigerem Total tau.

Interaktion des ADAScog modifizierter Total Score mit der Zeit bedeutet: Patienten mit höherem Score zu Studienbeginn haben steilere Abnahmen.

Die Residuen dieses Modells weichen nicht erkennbar von der Normalverteilung ab. Daher ist keine Transformation der Zielvariablen angezeigt.

Der gesamte Modellfehler (Standardabweichung einer einzelnen Beobachtung) liegt zwischen 0.31 (6 Monate) und 0.52 (36 Monate).

Aus den unteren Konfidenzgrenzen ergeben sich die entsprechenden Modellfehler 0.26 (6 Monate) und 0.40 (36 Monate)

Aus den oberen Konfidenzgrenzen ergeben sich die entsprechenden Modellfehler 0.38 (6 Monate) und 0.69 (36 Monate)

Zum Vergleich: Die Standardabweichung aller Global-Scores (6 bis 36 Monate) beträgt 0.86

In den Graphiken gemäss **Figuren 3a bis 3d** wird die Auswirkung jedes Prädiktors über die Zeit veranschaulicht. Die jeweils anderen Prädiktoren werden auf ihrem Median oder ihrem häufigsten Wert als Referenzwert fixiert. Diese Referenzwerte sind nachfolgend gelistet:

$$\text{Abeta}_{1-42} / \text{Total tau} = 1.6$$

```
faqtotal = 2
totalmodbl = 18.3
nbatglobbl = -1
```

**Figur 3a** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von Abeta1-42 /Total tau.

**Figur 3b** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von FAQtotal.

**Figur 3c** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von ADAScog total mod.

**Figur 3d** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte der Baseline.

Zu jedem Probanden mit vollständigen Prädiktordaten (insb. CSF vorhanden) werden Werte nach dem gemischten Modell simuliert. Dabei werden pro Proband zunächst Logarithmen der Standardabweichungen für Achsenabschnitt und Steigung simuliert. Mit den so gewonnenen Standardabweichungen werden Zufallseffekte für den Achsenabschnitt und die Steigung generiert. Ferner wird für jeden Probanden ein Vektor von festen Effekten gemäß den Parameterschätzungen und deren Kovarianzmatrix erzeugt. Zusätzlich wird für jeden Probanden und jeden Untersuchungszeitpunkt ein Fehlerterm generiert.

Diese Daten wurden vollständig für die Visiten 1-4 erzeugt. Die beobachteten Daten sind nicht vollständig. Ein Vergleich mit jedem Datenpunkt ist insofern nicht möglich.

Die simulierten Daten stimmen zwar an keiner Stelle vollständig mit den beobachteten Daten überein. Ihre Gesamtverteilung ist aber ähnlich der bei den beobachteten Daten. Folgende Verteilungen ergaben sich rechnerisch:

Beobachtete Daten gemäss ADNI Daten:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-3.9520	-1.4630	-0.9516	-1.0180	-0.4975	1.2780	14.0000

SD: 0.7839431

Simulierte Daten gemäss erstem Modell:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-3.2510	-1.5170	-1.0570	-1.0510	-0.6136	1.1020

SD: 0.7399358

**Figuren 4a und 4b** zeigen die Verteilungen der beobachteten und simulierten Daten des Global Score in zwei Histogrammen (jeweils alle Visiten enthaltend) im Vergleich. **Figur 4c** zeigt die Verteilungen der beobachteten und simulierten Daten des Global Score pro Visite in QQ-Plots.

Die simulierten Daten werden aus den Modellparametern und den zu Studienbeginn verfügbaren Prädiktoren ermittelt. Bei der Anwendung auf eine neue Studie werden die gleichen Modellparameter, aber die Prädiktoren der neuen Studienteilnehmer, zur Simulation verwendet.

In einem zweiten Rechenbeispiel werden bei einer im Vergleich zum ersten Rechenbeispiel grösseren Zahl von Patienten mit MCI Veränderungen des globalen Scores mit der Zeit unter Adjustierung für demographische Variablen und cognitive Variablen zu Studienbeginn (Prädiktoren) geschätzt. Das zweite Rechenbeispiel entspricht im wesentlichen dem ersten Beispiel, wobei der Quotient Abeta1-42 / Total tau zu Studienbeginn (Variable betaovtau) als möglicher Prädiktor nicht verwendet wird.

Auch das zweite Rechenbeispiel beginnt mit einem wie vorstehend beschriebenen Startmodell mit der Zeit, allen Prädiktorvariablen und allen Interaktionen mit der Zeit als festen Effekten und mit zufälligen Effekten für die Patienten, d.h. es erlaubt für jeden Patienten eine zufällige Abweichung vom Mittelwert, der aus den festen Effekten geschätzt wird.

Data: MCIba\_0  
 AIC BIC logLik  
 1073.724 1205.254 -511.8618

Random effects:

Formula: ~1 + nviscode | rid  
 Structure: General positive-definite, Log-Cholesky parametrization  
 StdDev Corr  
 (Intercept) 0.19018619 (Intr)  
 nviscode 0.06718466 0.661  
 Residual 0.24998144

Fixed effects: nbatglob ~ vsbmicat2 + nviscode + hmscore + napgen + pteducat + ptage + ptgender + faqttotal + mmscore + totalmodbl + nbatglobbl + nviscode:vsbmicat2 + nviscode:hmscore + nviscode:napgen + nviscode:pteducat + nviscode:mmscore + nviscode:totalmodbl + nviscode:nbatglobbl

	Value	Std. Error	DF	t-value	p-value
(Intercept)	0.2186482	0.4046705	1061	0.540312	0.5891
vsbmicat2high	-0.0693033	0.0393589	363	-1.760802	0.0791
vsbmicat2obese	-0.0405157	0.0553292	363	-0.732266	0.4645
nviscode	0.0474761	0.1218900	1061	0.389500	0.6970
hmscore	0.0167155	0.0256497	363	0.651684	0.5150
napgen	0.0355239	0.0263572	363	1.347787	0.1786
pteducat	0.0051555	0.0064076	363	0.804598	0.4216
ptage	-0.0039488	0.0023259	363	-1.697763	0.0904
ptgender	-0.0244486	0.0353680	363	-0.691264	0.4898
faqttotal	-0.0091341	0.0037908	363	-2.409531	0.0165
mmscore	0.0095854	0.0109641	363	0.874254	0.3826
totalmodbl	-0.0027840	0.0035731	363	-0.779138	0.4364
nbatglobbl	0.9521837	0.0355300	363	26.799457	0.0000
vsbmicat2high:nviscode	0.0199568	0.0137361	1061	1.452874	0.1466
vsbmicat2obese:nviscode	0.0488740	0.0203208	1061	2.405120	0.0163
nviscode:hmscore	-0.0016343	0.0091400	1061	-0.178805	0.8581
nviscode:napgen	-0.0362322	0.0093205	1061	-3.887381	0.0001
nviscode:pteducat	-0.0050203	0.0023201	1061	-2.163857	0.0307
nviscode:mmscore	0.0018276	0.0039647	1061	0.460954	0.6449
nviscode:totalmodbl	-0.0049473	0.0012680	1061	-3.901703	0.0001
nviscode:nbatglobbl	0.0142686	0.0126101	1061	1.131520	0.2581

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.510351267	-0.553693504	-0.002053127	0.540254540	3.400516216

Number of Observations: 1445  
 Number of Groups: 375

Random Effects:

Level: rid	lower	est.	upper
sd((Intercept))	0.1466417	0.19018619	0.24666092
sd(nviscode)	0.0529942	0.06718466	0.08517496
cor((Intercept), nviscode)	-0.2795052	0.66135055	0.95427510

Within-group standard error:  
 lower est. upper  
 0.2376837 0.2499814 0.2629155

Die folgenden Interaktionen tragen nicht zur Vorhersage bei und wurden in einem schrittweisen Verfahren eliminiert:

nviscode:hmscore

nviscode:mmscore  
 nviscode:nbatglobbl  
 nviscode:pteducat

Folgende Prädiktoren tragen nicht zur Vorhersage bei und wurden schrittweise eliminiert:

pteducation  
 hmscore (Hachinski score)  
 ptgender  
 mmscore

Nach den zuvor beschriebenen Schritten ergibt sich folgendes einfaches Modell:

Linear mixed-effects model fit by REML

Data: MCIba\_0  
 AIC BIC logLik  
 1005.021 1094.557 -485.5104

Random effects:

Formula: ~1 + nviscode | rid  
 Structure: General positive-definite, Log-Cholesky parametrization  
 StdDev Corr  
 (Intercept) 0.18855028 (Intr)  
 nviscode 0.06763452 0.665  
 Residual 0.24999246

Fixed effects: nbatglob ~ vsbmicat2 + nviscode + napgen + ptage + faqtotal + totalmodbl + nbatglobbl + vsbmicat2:nviscode + nviscode:napgen + nviscode:totalmodbl

	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.5517950	0.18082015	1065	3.05162	0.0023
vsbmicat2high	-0.0671974	0.03829876	367	-1.75456	0.0802
vsbmicat2obese	-0.0404127	0.05358203	367	-0.75422	0.4512
nviscode	0.0121559	0.02183583	1065	0.55670	0.5779
napgen	0.0340903	0.02620448	367	1.30093	0.1941
ptage	-0.0038468	0.00227450	367	-1.69125	0.0916
faqtotal	-0.0087170	0.00375371	367	-2.32225	0.0208
totalmodbl	-0.0028959	0.00340943	367	-0.84939	0.3962
nbatglobbl	0.9747272	0.03045660	367	32.00381	0.0000
vsbmicat2high:nviscode	0.0211988	0.01365817	1065	1.55209	0.1209
vsbmicat2obese:nviscode	0.0546617	0.02002082	1065	2.73024	0.0064
nviscode:napgen	-0.0353962	0.00931502	1065	-3.79991	0.0002
nviscode:totalmodbl	-0.0055600	0.00103651	1065	-5.36418	0.0000

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.5597357723	-0.5603156605	-0.0002501770	0.5339628395	3.3359298640

Number of Observations: 1445  
 Number of Groups: 375

Random Effects:

```

Level: rid
              lower      est.      upper
sd((Intercept))      0.14500809 0.18855028 0.24516706
sd(nviscode)         0.05361047 0.06763452 0.08532715
cor((Intercept),nviscode) -0.28172862 0.66478431 0.95556777

```

```

Within-group standard error:
      lower      est.      upper
0.2376749 0.2499925 0.2629484

```

Interaktion von BMI mit der Zeit bedeutet: Patienten mit Übergewicht fallen weniger schnell ab oder werden sogar besser. Bei fettleibigen Patienten tritt dieser Effekt noch stärker auf.

Interaktion von APOE mit der Zeit bedeutet: Patienten mit E4 Allelen fallen schneller ab.

Interaktion von ADAScog Score mit der Zeit bedeutet: Patienten mit höhern Ausgangswerten des ADAScog Score haben steileren Abfall.

Die Residuen des Modells zeigen keine Abweichung von der Normalverteilung, eine Transformation der abhängigen Variablen ist nicht angezeigt.

Der Gesamtfehler dieses Modells (Standardabweichung eines Einzelwertes) beträgt:

0.32 (6 Monate) bis 0.51 (36 Monate)

Untere Grenzen für diesen Gesamtfehler:

0.28 (6 Monate) bis 0.43 (36 Monate)

Obere Grenzen für diesen Gesamtfehler:

0.37 (6 Monate) bis 0.63 (36 Monate)

Zum Vergleich: Die Standardabweichung aller Global-Scores (6 bis 36 Monate) beträgt 0.86

In den Graphiken gemäss **Figuren 5a bis 5f** wird die Auswirkung jedes Prädiktors über die Zeit veranschaulicht. In jeder Graphik wird neben

der Zeit (in Halbjahren) eine Variable variiert (verschiedene Geraden). Bei der variierenden Variablen werden Quartile und Median oder typische Werte eingesetzt. Bei allen anderen, in der betreffenden Graphik nicht variierenden Variablen werden die folgenden Mediane beziehungsweise typischen Werte eingesetzt.

BMI: normal  
APOE: 0 E4-Allele  
Alter:  
faqttotal = 2  
ADAScog: totalmodbl = 18.3  
Neurobat: nbatglobbl = -1

**Figur 5a** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte des BMI.

**Figur 5b** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von APO E4.

**Figur 5c** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte des Alters.

**Figur 5d** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von FAQtotal

**Figur 5e** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte von ADAScog total (modifiziert)

**Figur 5f** zeigt den vorausgesagten Score in Abhängigkeit der Zeit für drei verschiedene Werte der Baseline.

Zu jedem Probanden mit vollständigen Prädiktordaten werden Werte nach dem gemischten Modell simuliert.

Dabei werden pro Proband zunächst Logarithmen der Standardabweichungen für Achsenabschnitt und Steigung simuliert. Mit den so gewonnenen Standardabweichungen werden Zufallseffekte für den Achsenabschnitt und die Steigung generiert. Ferner wird für jeden Probanden ein Vektor von festen Effekten gemäß den Parameterschätzungen und deren Kovarianzmatrix erzeugt. Zusätzlich wird für jeden Probanden und jeden Untersuchungszeitpunkt ein Fehlerterm generiert.

Diese Daten wurden vollständig für die Visiten 1-4 erzeugt. Die beobachteten Daten sind nicht vollständig. Ein Vergleich für jeden Datenpunkt ist insofern nicht möglich.

Die simulierten Daten stimmen zwar an keiner Stelle vollständig mit den beobachteten Daten überein. Ihre Gesamtverteilung ist

aber ähnlich der bei den beobachteten Daten. Folgende Verteilungen ergaben sich rechnerisch:

Beobachtete Daten gemäss ADIN Daten:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-3.9520	-1.4960	-0.9681	-1.0110	-0.4673	1.7420	20.0000

SD: 0.8296862

Simulierte Daten mit zweitem Rechenmodell:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-3.3970	-1.5340	-0.9858	-0.9937	-0.3990	1.4310

SD: 0.7875737

Die Verteilungen der beobachteten und simulierten Daten werden In **Figur 6a** und **Figur 6b** in zwei Histogrammen (jeweils alle Visiten enthaltend) und in **Figur 6c** in QQ-Plots pro Visite verglichen.

Die simulierten Daten werden aus den Modellparametern und den zu Studienbeginn verfügbaren Prädiktoren ermittelt. Bei der Anwendung auf eine neue Studie werden die gleichen Modellparameter, aber die Prädiktoren der neuen Studienteilnehmer zur Simulation, verwendet.

## Patentansprüche

1. Methode zur Durchführung klinischer Studien zur Behandlung einer interessierenden Erkrankung mit den Schritten:

- Festlegen einer Testgruppe von Testpersonen und Bestimmung von Ausgangsmerkmalen,
- Verabreichen einer zu untersuchenden Behandlung über einen vorherbestimmten Zeitraum an die Testpersonen der Testgruppe,
- mindestens einmaliges Bestimmen von Werten mindestens eines vordefinierten Zielkriteriums der Testpersonen der Testgruppe bis zum Ablauf des vorherbestimmten Zeitraumes,

gekennzeichnet durch die weiteren Schritte

- mindestens einmaliges Bestimmen von Werten des mindestens einen vordefinierten Zielkriteriums von Kontrollpersonen einer virtuellen Kontrollgruppe, welche die Behandlung nicht erhalten hat, bis zum Ablauf des vorherbestimmten Zeitraumes, wobei die Werte des mindestens einen vordefinierten Zielkriteriums der Kontrollpersonen im vorherbestimmten Zeitraum mittels eines Prognosemodells aus Werten von Ausgangsmerkmalen der Kontrollpersonen bestimmt werden.
- Vergleichen der bis zum Ablauf des vorherbestimmten Zeitraumes bestimmten Werte des mindestens einen vordefinierten Zielkriteriums der Kontrollpersonen mit denjenigen der Testpersonen.

2. Methode gemäss Anspruch 1, **dadurch gekennzeichnet, dass** zusätzliche Kontrollpersonen einer realen Kontrollgruppe über eine initiale Zeitspanne eine Placebo-Behandlung erhalten, dass die Werte des mindestens einen vordefinierten Zielkriteriums der zusätzlichen Kontrollpersonen der realen Kontrollgruppe nach Ablauf der initialen Zeitspanne bestimmt werden und mit nach Ablauf dieser Zeitspanne bestimmten Werten des mindestens einen vordefinierten Zielkriteriums der Kontrollpersonen der virtuellen Kontrollgruppe verglichen werden.
3. Methode gemäss Anspruch 2, **dadurch gekennzeichnet, dass** basierend auf dem Vergleich der Werte das Prognosemodell korrigiert wird.
4. Methode gemäss einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** die Werte der Ausgangsmerkmale der Kontrollpersonen anhand von Werten der Ausgangsmerkmale der Testpersonen festgelegt werden.
5. Methode gemäss einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet, dass** das die Testgruppe eine der Fragestellung der Studie, insbesondere der Art der Krankheit und der Intervention und optional der Toleranz des Prognosemodells angepasste Anzahl Testpersonen umfasst.
6. Methode gemäss einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet, dass** ein Verhältnis von mindestens 1:1 der Anzahl Kontrollpersonen zur Anzahl Testpersonen verwendet wird.
7. Methode gemäss einem der Ansprüche 1 bis 6, **dadurch gekennzeichnet, dass** es sich bei der zu untersuchenden Behandlung um eine Behandlung von Krankheiten mit voraussichtlich malign-

nem Verlauf, beispielsweise Demenzen, insbesondere der Alzheimer-Krankheit, handelt.

8. Methode gemäss einem der Ansprüche 1 bis 7, **dadurch gekennzeichnet, dass** Werte von mindestens einem oder mehreren vordefinierten Zielkriterien bestimmt und verglichen werden.
9. Methode gemäss einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet, dass** die Zielkriterien typischerweise ausgewählt sind aus der Gruppe bestehend aus:
  - Parameter des zentralen Nervensystems, insbesondere neuropsychologische Leistungsmerkmale,
  - Subjektive und objektive Demenzsymptome,
  - time to event (Progression to Dementia)
  - Veränderungen im Gehirn, insbesondere Schrumpfungen des Hippocampus insgesamt und/oder Schrumpfungen bestimmter Abschnitte des Hippocampus oder anderer Gehirnbereiche, eine vaskuläre Pathologie und ähnliche Veränderungen,
  - Marker in der Cerebro-Spinal Flüssigkeit, wie beispielsweise Abeta, Tau, Quotienten von Abeta und/oder andere Indikatoren,
  - das Vorhandensein pathologischer Proteine, wie Abeta und/oder Tau in anderen Körperflüssigkeiten,
  - weitere biochemische und genetische Marker,
  - psychiatrische und motorische Auffälligkeiten
10. Methode gemäss einem der Ansprüche 1 bis 9, **dadurch gekennzeichnet, dass** die Ausgangsmerkmale typischerweise mindestens ein Zielkriterium umfassen und/oder ausgewählt sind aus der Gruppe bestehend aus:
  - Alter des Patienten,
  - das Vorhandensein des E-4 Allels des Apolipoprotein-Gens und/oder anderer genetischer bzw. biochemischer Marker,

- die in der Kindheit oder der Adoleszenz erfasste Intelligenz oder ein Aequivalent dazu, die Ausbildungsdauer und/oder die berufliche Tätigkeit,
- Life-style Faktoren im mittleren Alter, beispielsweise Ernährung, körperliche Aktivität, Sport und/oder andere Aktivitäten,
- Klinische Merkmale und Diagnosen wie Adipositas, Hypertonie, Diabetes und/oder Hyperlipidämie, die auch in einem klinischen Gesamtscore zusammengefasst sein können,
- Traumatische Ereignisse wie Hirnschlag und Unfälle, die die Hirnfunktion zeitweise oder dauernd beeinflussen können
- Funktionelle Parameter des menschlichen Herzens oder anderen Organen des menschlichen Körpers,
- Charaktermerkmale wie Neurotizismus und/oder Gewissenhaftigkeit.

11. Methode zur Etablierung eines Prognosemodells für klinische Studien, gekennzeichnet durch die Schritte:

- a) Sammeln von repräsentativen Resultaten aus klinischen Studien mit unbehandelten Patienten betreffend den Verlauf einer interessierenden Erkrankung,
- b) Auswertung dieser Resultate,
- c) Ermittlung mindestens eines Zielkriteriums, das für den Verlauf der interessierenden Erkrankung klinisch relevant ist und in einer nachfolgenden Studie als vordefiniertes Zielkriterium verwendet werden kann,

d) Bestimmung der Veränderung des Wertes dieses mindestens einen Zielkriteriums für eine bestimmte Anzahl von Patienten,

so dass anhand dieser Bestimmung im Schritt d) eine Prognose für die zeitliche Veränderung des Wertes dieses Zielkriteriums basierend auf Werten von Ausgangsmerkmalen bei einem Patienten erstellt werden kann.

12. Methode gemäss Anspruch 11, **dadurch gekennzeichnet, dass** Resultate einer ausreichenden Zahl von Personen, die während einer genügend langen Dauer der präklinischen Phase der betreffenden Krankheit beobachtet wurden, gesammelt und ausgewertet werden.
13. Methode gemäss einem der Ansprüche 11 oder 12, **dadurch gekennzeichnet, dass** mindestens ein oder mehrere Zielkriterien ermittelt werden.
14. Methode gemäss einem der Ansprüche 11 bis 13, **dadurch gekennzeichnet, dass** die Zielkriterien typischerweise ausgewählt sind aus der Gruppe bestehend aus:
- Parameter des zentralen Nervensystems, insbesondere neuropsychologische Leistungsmerkmale,
  - Subjektive und objektive Demenzsymptome,
  - time to event (Progression to Dementia)
  - Veränderungen im Gehirn, insbesondere Schrumpfungen des Hippocampus insgesamt und/oder Schrumpfungen bestimmter Abschnitte des Hippocampus oder anderer Gehirnbereiche, eine vaskuläre Pathologie und ähnliche Veränderungen,
  - Marker in der Cerebro-Spinal Flüssigkeit, wie beispielsweise Abeta, Tau, Quotienten von Abeta und/oder andere Indikatoren,

- das Vorhandensein pathologischer Proteine, wie Abeta und/oder Tau in anderen Körperflüssigkeiten,
- weitere biochemische und genetische Marker,
- psychiatrische und motorische Auffälligkeiten

15. Methode gemäss einem der Ansprüche 11 bis 14, **dadurch gekennzeichnet, dass** die Ausgangsmerkmale typischerweise mindestens ein Zielkriterium umfassen und/oder ausgewählt sind aus der Gruppe bestehend aus:

- Alter des Patienten,
- das Vorhandensein des E-4 Allels des Apolipoprotein-Gens und/oder anderer genetischer bzw. biochemischer Marker,
- die in der Kindheit oder der Adoleszenz erfasste Intelligenz oder ein Aequivalent dazu, die Ausbildungsdauer und/oder die berufliche Tätigkeit,
- Life-style Faktoren im mittleren Alter, beispielsweise Ernährung, körperliche Aktivität, Sport und/oder andere Aktivitäten,
- Klinische Merkmale und Diagnosen wie Adipositas, Hypertonie, Diabetes und/oder Hyperlipidämie, die auch in einem klinischen Gesamtscore zusammengefasst sein können,
- Traumatische Ereignisse wie Hirnschlag und Unfälle, die die Hirnfunktion zeitweise oder dauernd beeinflussen können
- Funktionelle Parameter des menschlichen Herzens oder anderen Organen des menschlichen Körpers,
- Charaktermerkmale wie Neurotizismus und/oder Gewissenhaftigkeit.

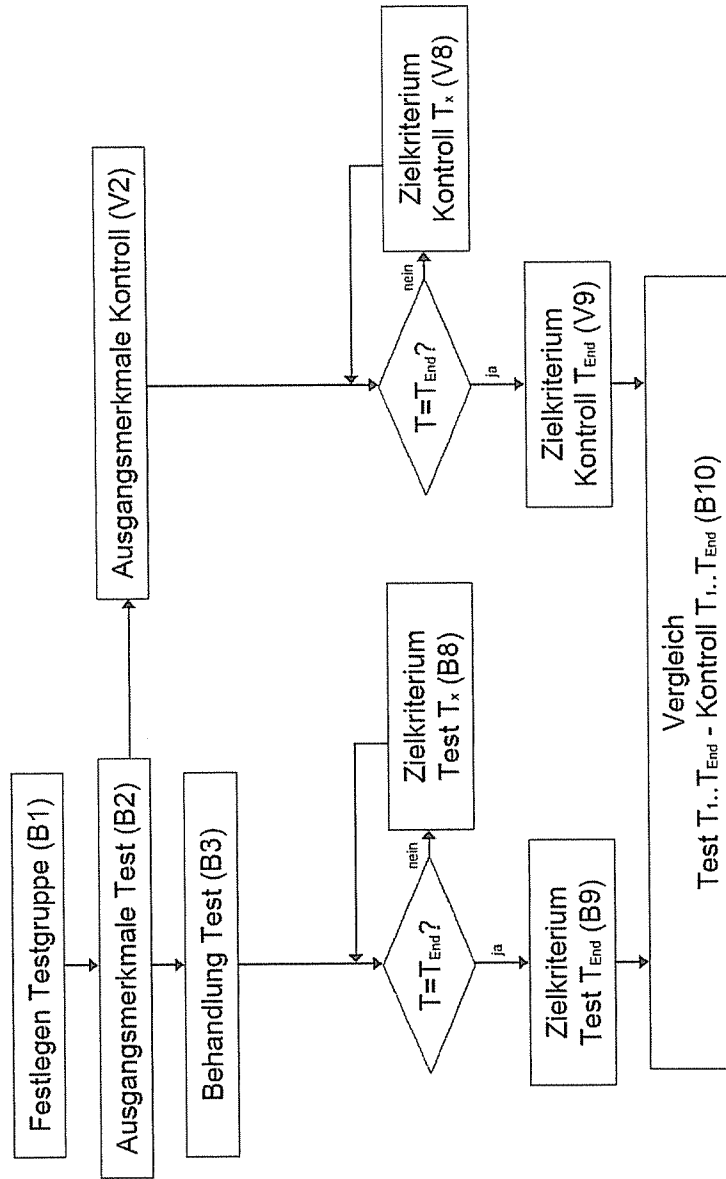


Fig. 1

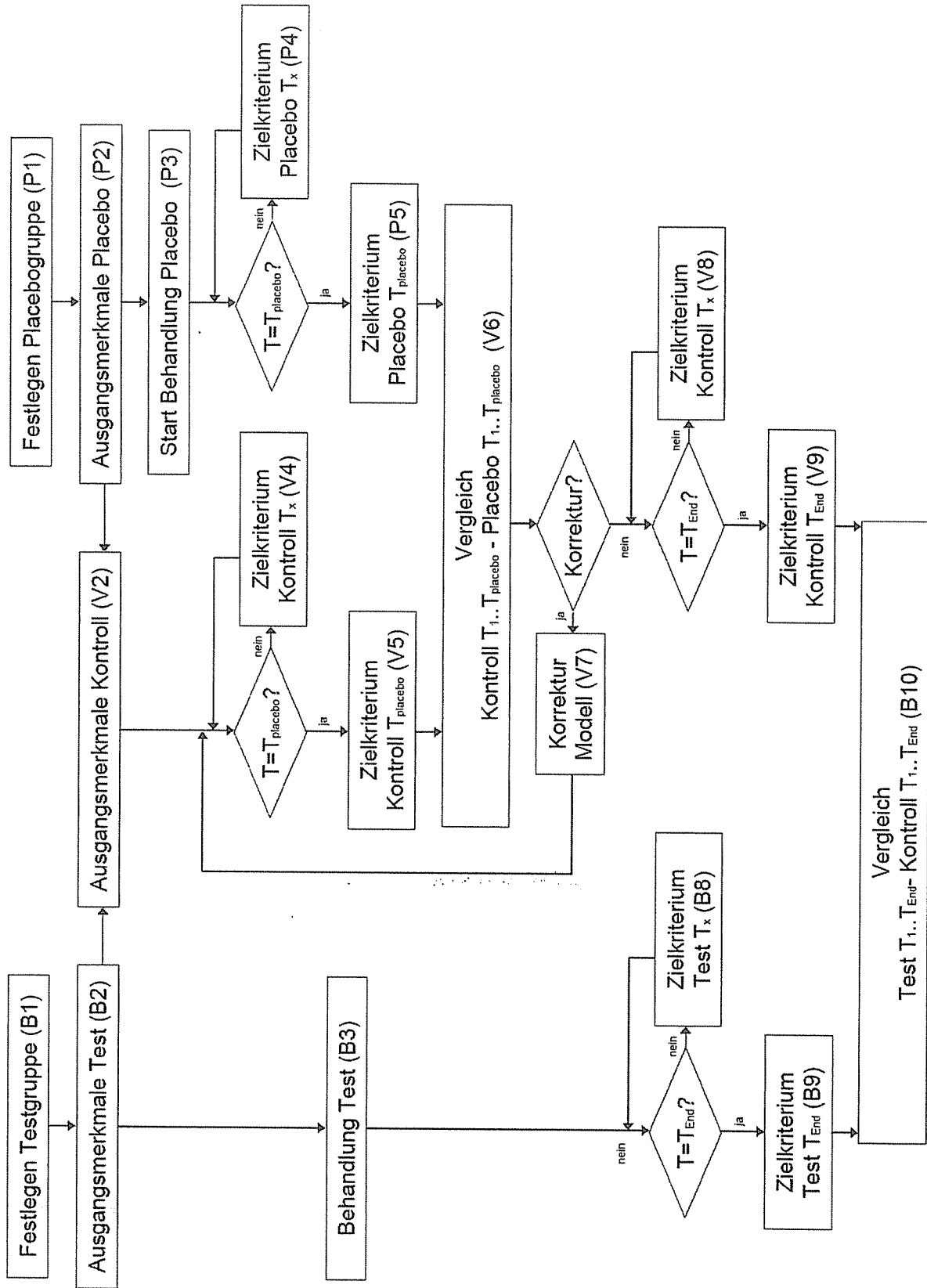


Fig. 2

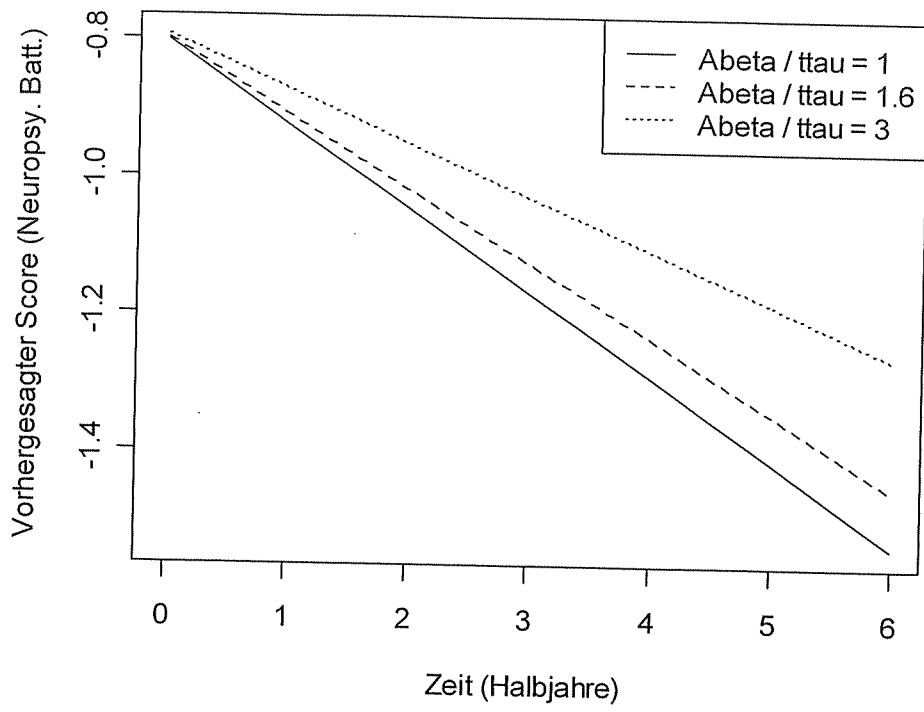


FIG. 3a

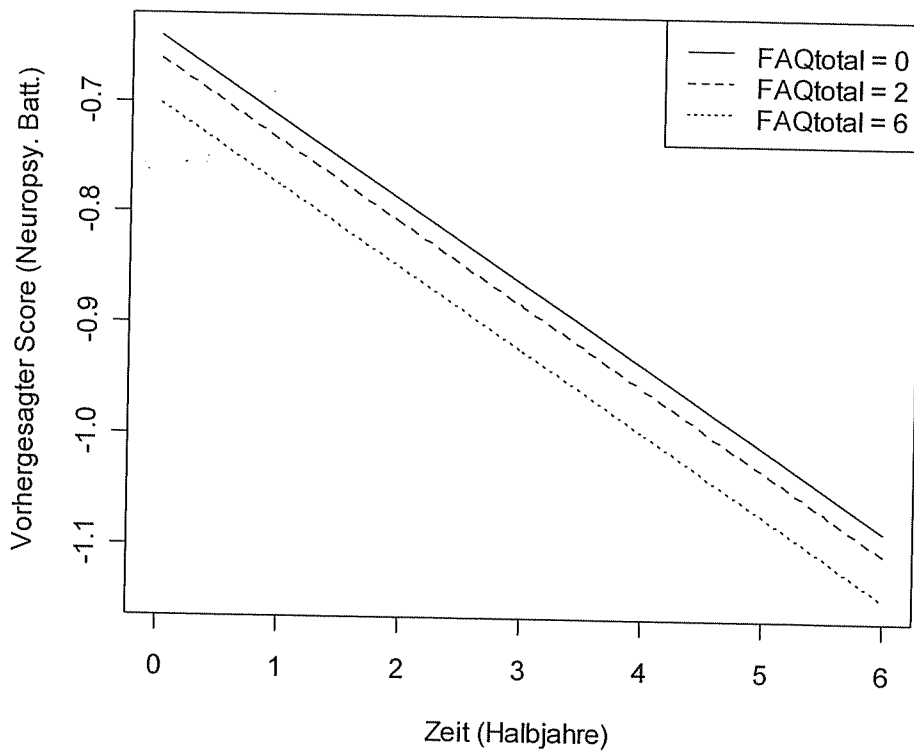


FIG. 3b

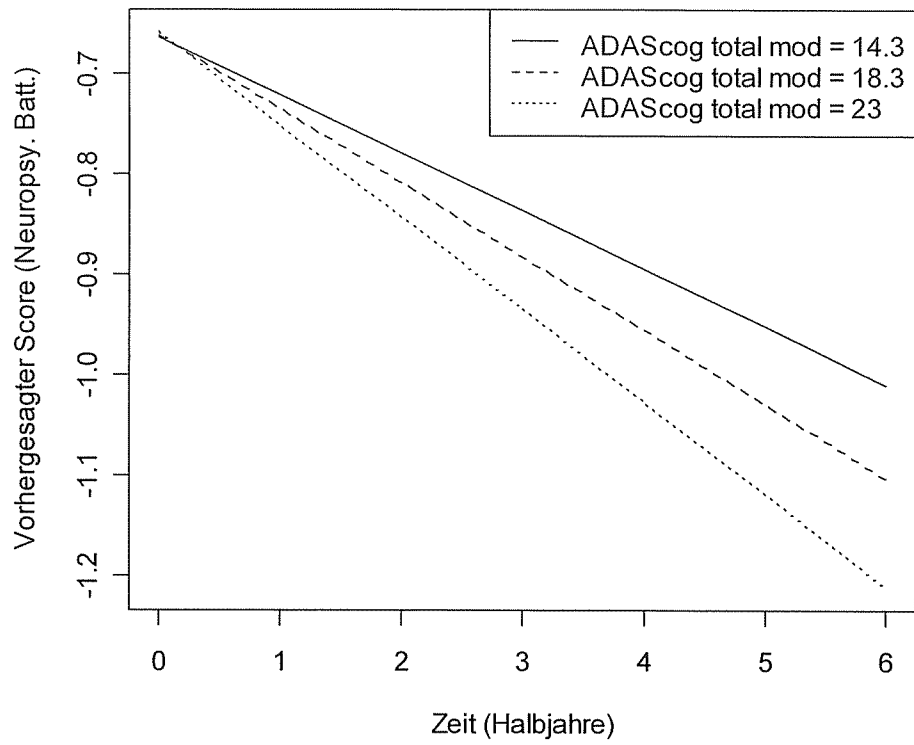


FIG. 3c

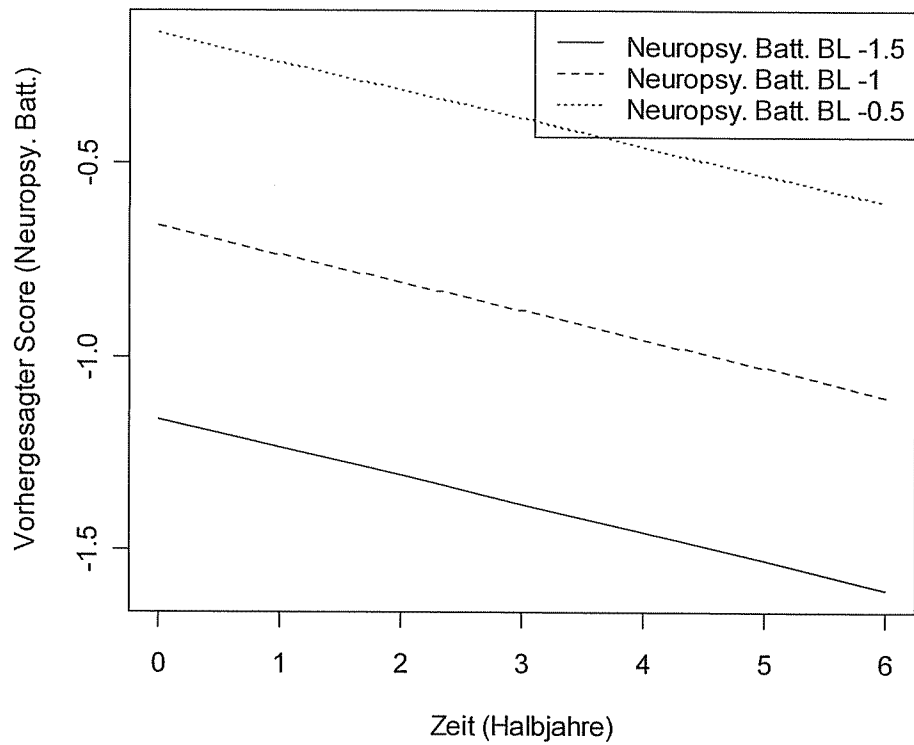


FIG. 3d

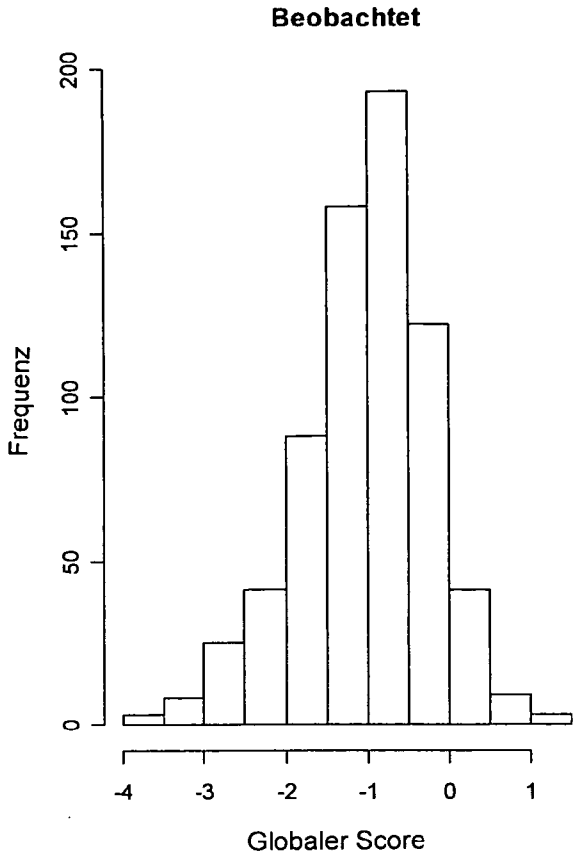


FIG. 4a

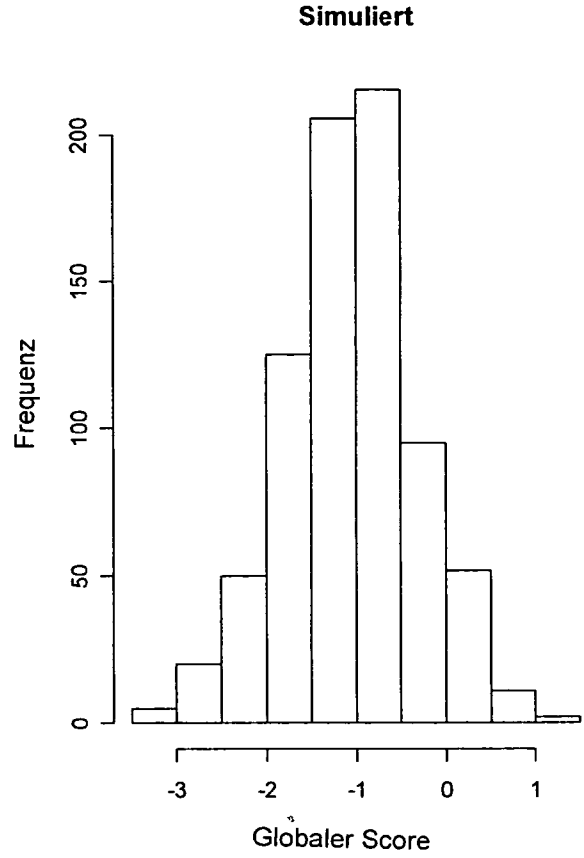


FIG. 4b

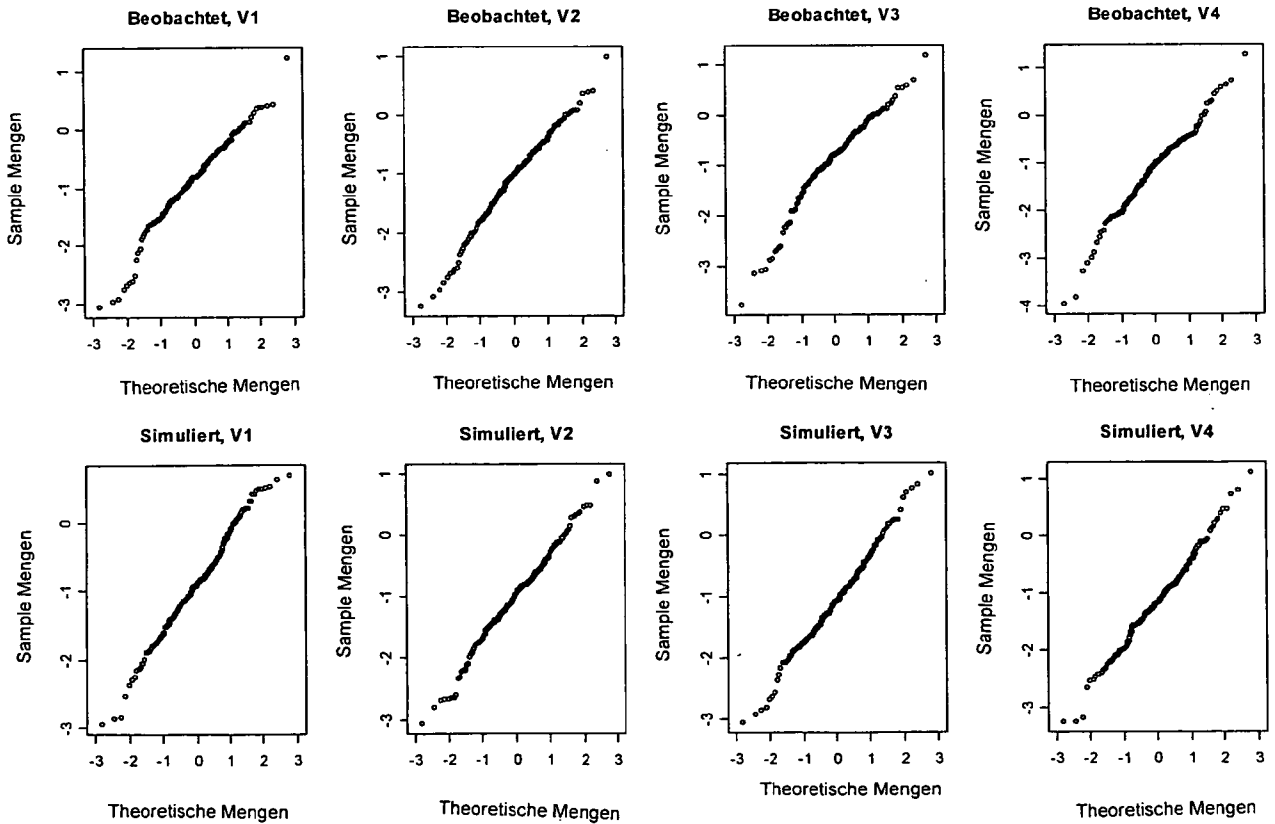


FIG. 4c

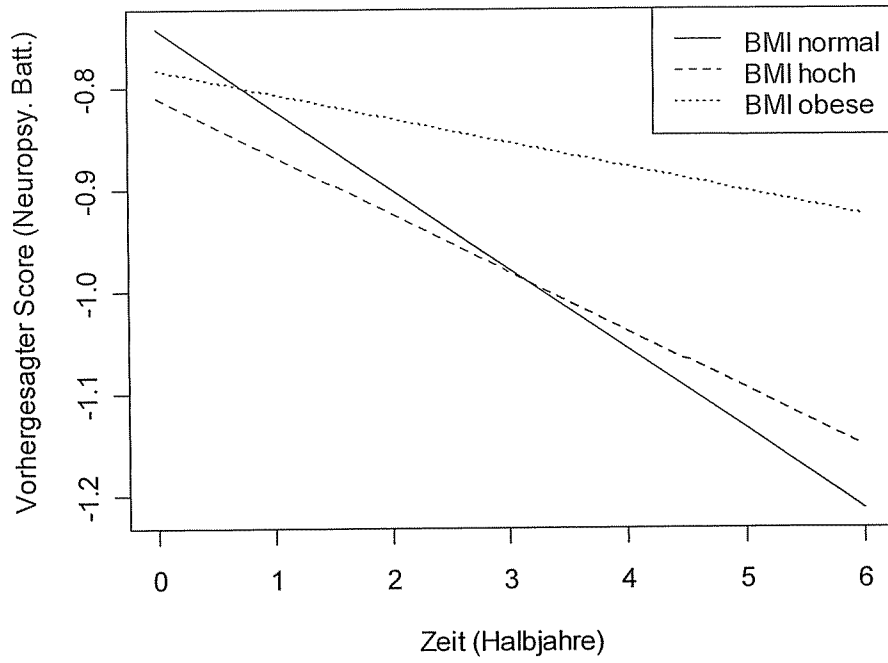


FIG. 5a

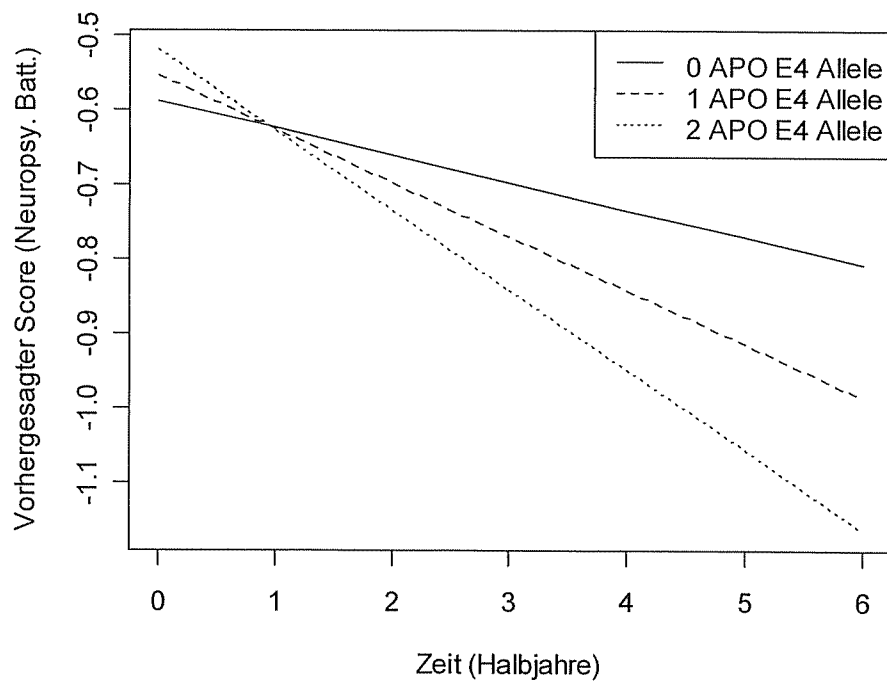


FIG. 5b

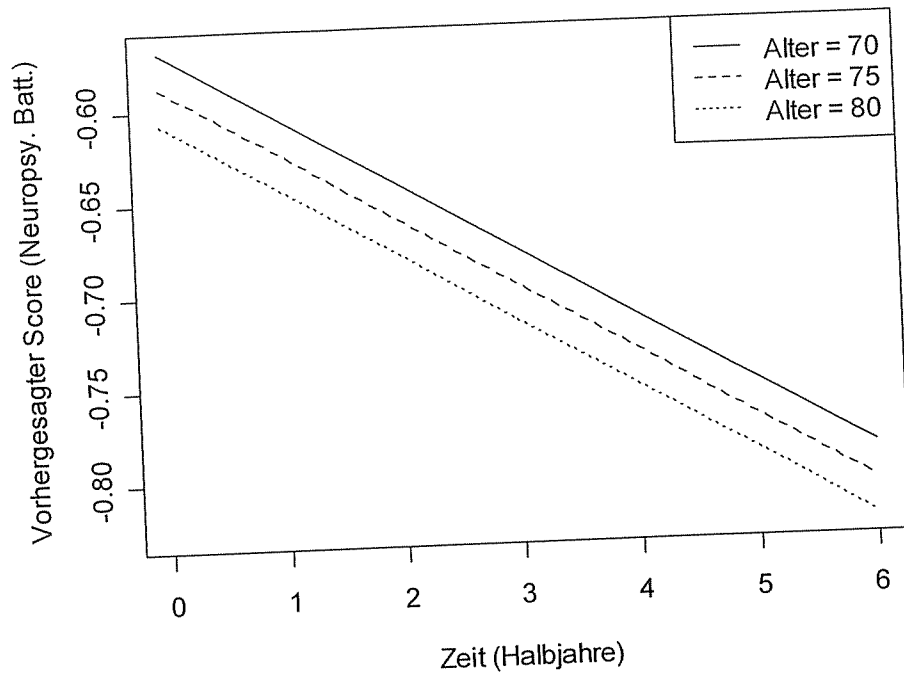


FIG. 5c

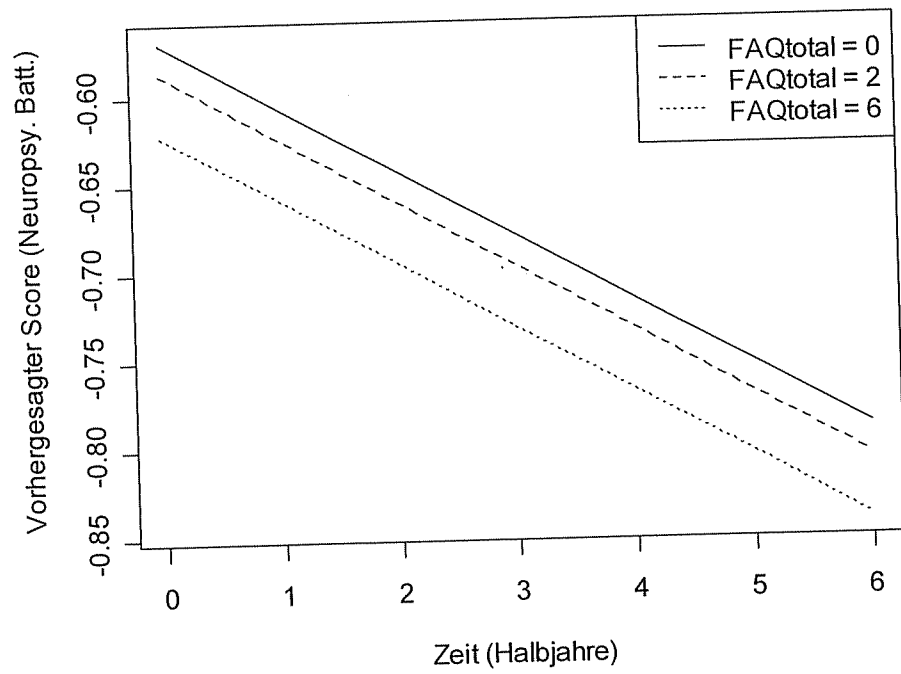


FIG. 5d

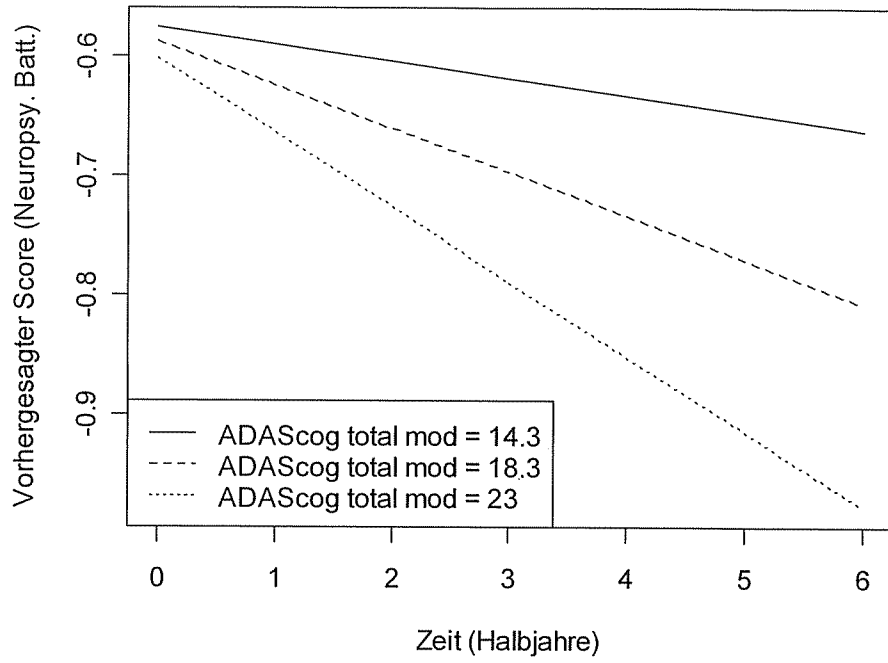


FIG. 5e

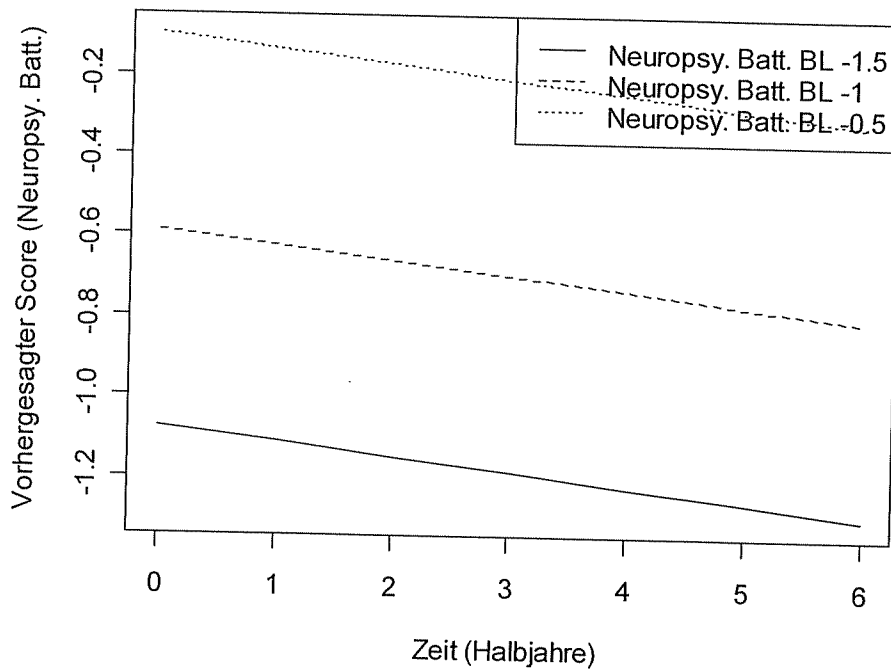


FIG. 5f

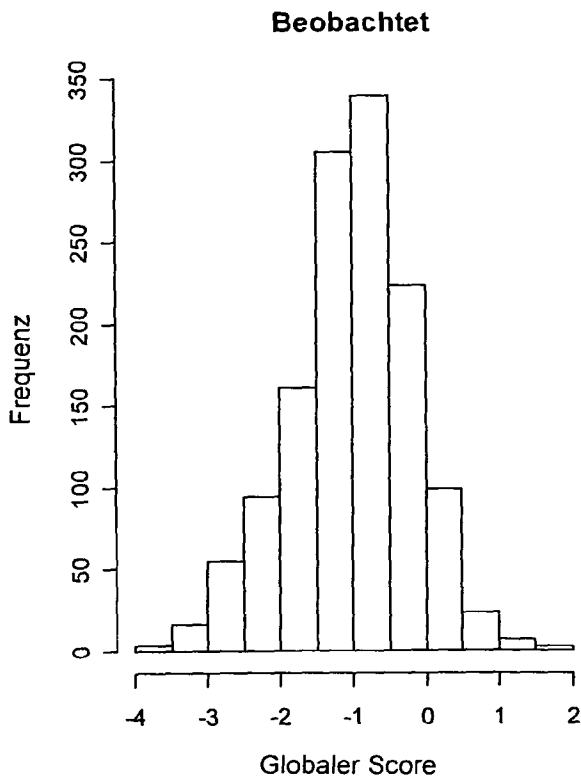


FIG. 6a

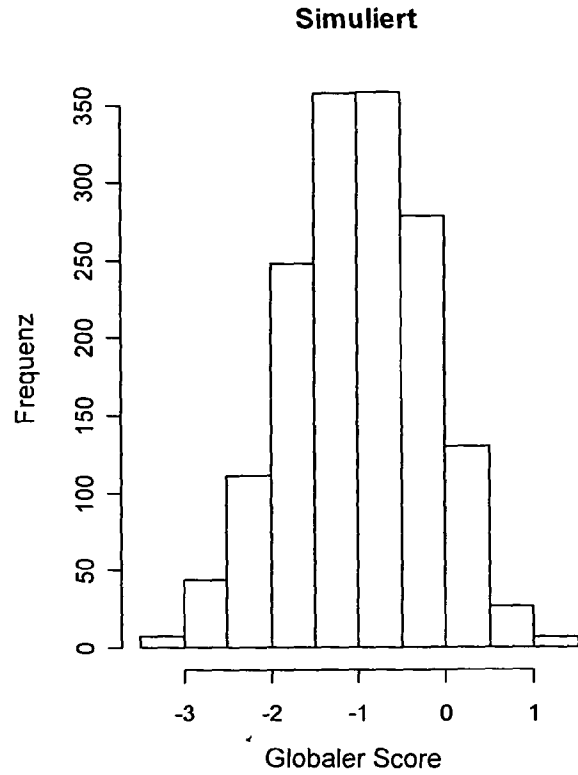


FIG. 6b

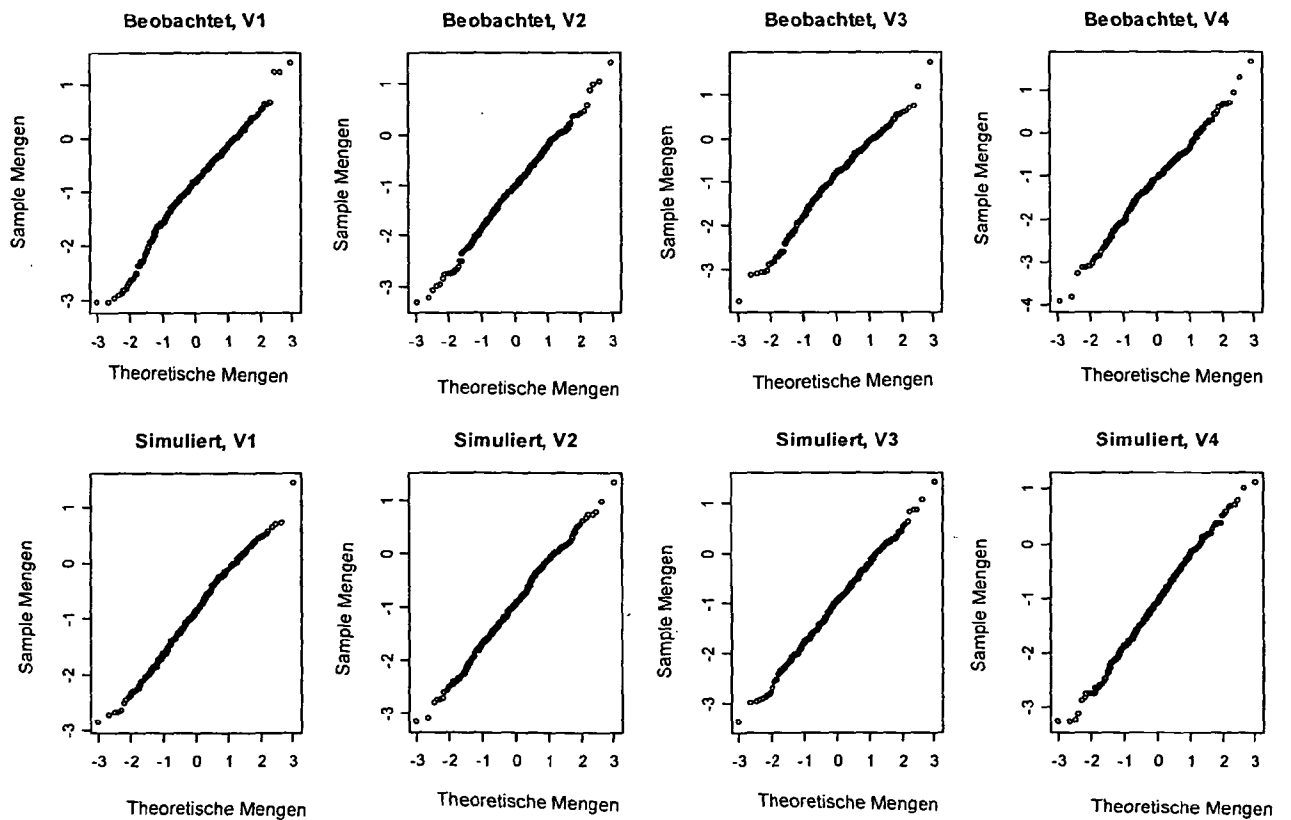


FIG. 6c