

Aus der Klinik für Psychiatrie und Psychotherapie
Universitätsklinikum des Saarlandes, Homburg/Saar
Direktor: Prof. Dr. med. M. Riemenschneider

**Optimierte Früherkennung kognitiver Störungen mittels
automatisierter Sprachanalyse:
eine prospektiv-randomisierte Studie**

**Dissertation zur Erlangung
des Grades eines Doktors der Theoretischen Medizin
der Medizinischen Fakultät
der UNIVERSITÄT DES SAARLANDES
2022**

vorgelegt von: Hanna Isabella Zeuschner, geb. Mayer

geb. am: 18.01.1988 in Saarlouis

Meinem Sohn Maximilian.

Tag der Promotion:	03. Januar 2024
Dekan:	Univ.-Prof. Dr. med. Michael D. Menger
Berichterstatter:	Prof. Dr. Matthias Riemenschneider Prof. Dr. Thomas Volk

Inhaltsverzeichnis

Abkürzungsverzeichnis	5
Abbildungsverzeichnis	6
Tabellenverzeichnis	7
1. Zusammenfassung	8
1.1 <i>Deutsche Zusammenfassung</i>	8
1.2 <i>English Abstract</i>	9
2. Einleitung	10
2.1 <i>Etymologie, Semantik und Definition</i>	10
2.2 <i>Epidemiologie: Prävalenz, Inzidenz, Geschlechtsunterschiede</i>	10
2.3 <i>Risikofaktoren</i>	12
2.4 <i>Formen dementieller Erkrankungen</i>	14
2.4.1 <i>Demenz bei Alzheimer Krankheit</i>	14
2.4.2 <i>Vaskuläre Demenz</i>	16
2.4.3 <i>Frontotemporale Demenz</i>	17
2.4.4 <i>Gemischte Demenz</i>	17
2.4.5 <i>Demenz mit frühem Beginn</i>	18
2.4.6 <i>Demenz bei andernorts klassifizierten Krankheiten</i>	18
2.5 <i>Verlauf und Prognose</i>	19
2.6 <i>Kosten</i>	20
2.7 <i>Diagnostik</i>	22
2.7.1 <i>Früherkennung</i>	23
2.8 <i>Fragestellung</i>	24
3. Material und Methoden	26
3.1 <i>ELEMENT-Studie</i>	26
3.2 <i>Studienablauf</i>	27
3.3 <i>Einschluss- und Ausschlusskriterien</i>	27
3.4 <i>Datenerhebung/Materialien</i>	28
3.4.1 <i>Erhobene Patientendaten</i>	28
3.4.2 <i>Neuropsychologische Testung und Fragebögen</i>	29
3.4.3 <i>Δelta (Delta-App)</i>	32
3.5 <i>Entwicklung sprachbasierter Biomarker</i>	34
3.6 <i>Statistische Auswertung</i>	37
4. Ergebnisse	39

4.1 Analyse des Gesamtkollektivs	39
4.1.1 Patientencharakteristika.....	39
4.1.2 Vergleich Studiengruppe vs. Dropout-Gruppe	40
4.1.3 Vergleich Studiengruppe vs. Kontrollgruppe	43
4.2 Einflussfaktoren auf das Testergebnis innerhalb der Studiengruppe (Modalität & Zeitpunkt).....	45
4.2.1 Vergleich Testung mittels App vs. Paper Pencil.....	45
4.2.2 Vergleich Ergebnis Erst- vs. Zweittestung.....	46
4.3 Vergleich der verschiedenen Demenzformen untereinander.....	48
4.4 Extraktion sprachbasierter Biomarker.....	51
4.4.1 Analytierte Subgruppe.....	51
4.4.2 Analytische und klinische Validierung	52
4.4.3 Ermittlung der diagnostischen Testgenauigkeit vom SB-C Biomarker	53
4.5 Aussagegüte Paper-Pencil-Subtests vs. Sprachbiomarker (SB-C).....	54
4.5.1 Unterscheidung gesunde Kontrollen vs. MCI	55
4.5.2 Unterscheidung gesunde Kontrollen vs. AD.....	56
4.5.3 Unterscheidung gesunde Kontrollen vs. MCI/AD	56
4.5.4 Unterscheidung MCI vs. AD.....	57
5. Diskussion	59
5.1 Interpretation der erhobenen Testergebnisse	59
5.2 Entwicklung eines Sprachbasierten Biomarkers	64
5.3 Vergleich der Performance vom SB-C mit der konventionellen CERAD Testbatterie..	69
5.4 Limitationen	70
5.5 Schlussfolgerung und Ausblick.....	71
6. Literaturverzeichnis	72
7. Publikation	82
8. Danksagung.....	83
9. Lebenslauf.....	84
10. Eidesstattliche Erklärung.....	85
11. Appendix	86
11.1 Patienteneinwilligung ELEMENT-Studie	86
11.2 Flyer ELEMENT Studie.....	87
11.3 Studieninformationen auf Internetseite des UKS	88

Abkürzungsverzeichnis

AD	Alzheimer Demenz
ADI	Alzheimer Disease International
ASR	Automatic speech recognition/ automatisierte Spracherkennung
AUC	area under curve
BDI-II	Beck's Depressions Inventar (II)
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BNT	Boston Naming Test
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CSF	Cerebrospinal fluid
CT	Computertomographie
DemTect	Demenz Detection Test
Δelta	Delta App
EOD	early onset dementia
DFKI	Deutsches Forschungsinstitut für künstliche Intelligenz
DLB	Dementia with Lewy-bodies
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
FDG-PET	Positronenemissionstomographie mit 18F- Fluordesoxyglucose
FTD	Frontotemporale Demenz
HC	Healthy control/ gesunde Kontrollgruppe
ICD-10	International statistical classification of diseases and related health problems
ki:elements	Spin-off des DFKI
MRT	Magnetresonanztomographie
MMSE	Mini Mental Status Examination
MMST	Mini Mental Status Test
ROC	receiver operating characteristic
SB-C	Score digitaler Sprachbiomarker Cognition
TMT	Trail Making Test
PET	Positronenemissionstomographie
VaD	vaskuläre Demenz
VF	Wortflüssigkeit

Abbildungsverzeichnis

Abbildung 1. Destabilisierung der Mikrotubuli durch Hyperphosphorylierung des Tau-Proteins (aus https://en.wikipedia.org/wiki/Alzheimer%27s_disease#Causes).	14
Abbildung 2. Schematische Darstellung der Entstehung des β -Amyloids als Vorläufer der Amyloid-Plaques: die membranständigen β - und γ - Sekretase zerschneiden das APP-Protein (aus https://en.wikipedia.org/wiki/Beta-secretase_1).	15
Abbildung 3. Kontinuum der Alzheimer Erkrankung (angelehnt an Alzheimer's disease facts & figures, 2020).....	19
Abbildung 4. Schematischer Studienablauf der ELEMENT-Studie.....	26
Abbildung 5. Schematische Repräsentation der Entwicklung des ki:e SB-C sowie der Subscores (SB-C: speech biomarker for cognition, SVF: semantic verbal fluency, CERAD: Consortium to Establish a Registry for Alzheimer's Disease).	35
Abbildung 6. Semantische und zeitliche Aspekte des verbalen Outputs des Subtests "Wortliste lernen".	36
Abbildung 7. Semantische und zeitliche Aspekte des verbalen Outputs der verbalen Wortflüssigkeit (Semantic verbal fluency (SVF)).	36
Abbildung 8. V3 Rahmen zur Evaluierung von Sprach Biomarkern (in Anlehnung an Goldsack (2020) (158)).....	37
Abbildung 9. Verteilung der Subtypen dementieller Erkrankungen in der Studiengruppe.	39
Abbildung 10. Korrelation zwischen MMSE (Mini Mental Status Examination) und ki:e SB-C Score für Kognition.....	52
Abbildung 11. Gruppenvergleich ki:e SB-C Score zwischen HC, MCI und AD (HC: healthy control, MCI: mild cognitive impairment, AD: Alzheimer's disease).	53
Abbildung 12. Vergleichende Receiver operating characteristic curve (ROC) für die vier Modelle (HC: healthy control/gesunde Kontrollgruppe, MCI: mild cognitive impairment, AD: Alzheimer Demenz).....	54
Abbildung 13. Receiver Operating Characteristic (ROC) Gesund vs. MCI.....	55
Abbildung 14. Receiver Operating Characteristic (ROC) zur Unterscheidung Gesund vs. AD.	56
Abbildung 15. Receiver Operating Characteristic (ROC) zur Unterscheidung von Gesund vs. MCI/AD.	57
Abbildung 16. Receiver Operating Characteristic (ROC) zur Unterscheidung von MCI vs. AD.	58

Tabellenverzeichnis

Tabelle 1. Potenziell modifizierbare Risikofaktoren der Demenz (in Anlehnung an Dementia prevention, intervention, and care: 2020 report of the Lancet Commission).	12
Tabelle 2. Interpretation der z-standardisierten Werte.	32
Tabelle 3. Charakteristika des Gesamtkollektivs, angegeben als absolute (relative) Häufigkeit bzw. Median (Spannweite).	40
Tabelle 4. Vergleich Studien- vs. Dropout-Gruppe.	41
Tabelle 5. Vergleich Studiengruppe vs. Dropout-Gruppe hinsichtlich der Testergebnisse anhand der Z-Werte.	42
Tabelle 6. Vergleich der Charakteristika von Studiengruppe vs. Kontrollgruppe.	43
Tabelle 7. Vergleich Studiengruppe (erste Testung) vs. Kontrollgruppe hinsichtlich der Testergebnisse, angegeben als z-standardisierte Werte.	44
Tabelle 8. Vergleich Testung mittels App vs. Paper-Pencil Verfahren in Studiengruppe.	46
Tabelle 9. Vergleich erste versus zweite Testung in der Studiengruppe.	47
Tabelle 10. Vergleich der Patientencharakteristika von Probanden mit MCI vs. Alzheimer vs. Andere Demenz in der Studiengruppe.	48
Tabelle 11. Vergleich der Testergebnisse der 1. Testung zwischen den dementiellen Subgruppen.	50
Tabelle 12. Deskriptive Statistik des eingeschlossenen Kollektivs (Gesund, MCI: mild cognitive impairment).	52
Tabelle 13. Diagnostische Genauigkeit (Performanz Matrix) SB-C Biomarker.	53
Tabelle 14. Unterscheidung Gesund vs. MCI.	55
Tabelle 15. Unterscheidung Gesund vs. AD.	56
Tabelle 16. Gesund vs. MCI/AD.	57
Tabelle 17. Vergleich der Studienlage zur Früherkennung neurodegenerativer Erkrankungen durch sprachlicher Biomarker.	68

Anmerkung

In dieser Dissertation wird eine möglichst Gender-neutrale Sprache angewendet, zu Zwecken eines besseren Leseflusses wird aber auch teilweise das generische Maskulinum eingesetzt. In diesen Situationen werden weibliche und anderweitige Geschlechteridentitäten ausdrücklich mitadressiert.

1. Zusammenfassung

1.1 Deutsche Zusammenfassung

Einleitung: Dementielle Erkrankungen nehmen zu und stellen Erkrankte wie Angehörige vor komplexe Herausforderungen, zu denen auch eine Veränderung in der Sprachproduktion zählt. Sprachbasierte Methoden haben daher in der Diagnostik dementieller Erkrankungen deutlich an Bedeutung zugenommen, sind aber noch personalintensiv und aufwändig. Die ELEMENT-Studie zielte in diesem Zusammenhang darauf ab, die diagnostische Genauigkeit der elektronischen Delta-App mit einer integrierten automatisierten Spracherkennung im Rahmen einer neuropsychologischen Demenzabklärung mit dem bisherigen Goldstandard mittels Paper-Pencil zu vergleichen. Aus den Sprachdateien, die die App während der Testung erfasste, wurden durch Spracherkennung digitale Biomarker extrahiert und deren Wertigkeit für die Diagnostik neurodegenerativer Erkrankungen überprüft.

Material und Methoden: Jeder Proband wurde in einem randomisierten, einfach verblindeten, zweiarmigen Studiendesign mit cross-over einmalig mit beiden Testmodalitäten im Abstand von drei Monaten getestet. Die kognitiv Beeinträchtigten unterschiedlicher Ätiologie und Schweregrade wurden dazu einer CERAD (Consortium to Establish a Registry for Alzheimer's Disease) Testung unterzogen. Die Datenerhebung erfolgte unizentrisch von 03/2019 bis 05/2022 in der Klinik für Psychiatrie und Psychotherapie des Universitätsklinikums des Saarlandes. Die Delta-App wurde vom Deutschen Forschungsinstitut für künstliche Intelligenz entwickelt und zeichnete die Testungen mit Audiodateien auf, aus denen ein Sprachbiomarker generiert wurde. Die Ergebnisse wurden mit denen der konventionellen Testung und gesunden Kontrollprobanden verglichen.

Ergebnisse: Von 200 randomisierten Probanden wurden 148 (74 %) zwei Mal getestet und in die finale Studiauswertung eingeschlossen. Die Kontrollgruppe mit 27 Probanden erzielte in allen 18 CERAD-Untertests signifikant bessere Ergebnisse. Die Studiengruppe schnitt bei der Erst- im Vergleich zur Zweitestung in neun Subtests signifikant schlechter ab. Bei Testung mittels App erzielten Probanden verglichen mit Paper-Pencil in fünf Untertests schlechtere Ergebnisse. Der kognitive Sprachbiomarker SB-C (speech biomarker cognition) ermöglichte eine gute bis sehr gute Unterscheidung zwischen den klinischen Testgruppen gesund, mild cognitive impairment und Alzheimer Demenz ($p < 0,001$). Die Zuordnung zu den Subgruppen war mit der Aussagekraft der sprachbasierten Subtests des CERAD vergleichbar, teilweise überlegen.

Schlussfolgerung: Der mithilfe automatisierter Spracherkennung kreierte Sprachbiomarker SB-C war dazu in der Lage, Veränderungen der Sprache zu erkennen und die diagnostische Genauigkeit bei der Erkennung dementieller Erkrankungen zu verbessern. Durch die Implementierung weiterer Testmodalitäten, bspw. von Magnetresonanztomographien, könnte dessen diagnostische Performance zukünftig noch weiter gesteigert werden.

1.2 English Abstract

Title: Optimized early detection of cognitive impairment using automated speech analysis:
A prospective randomized study

Introduction: Dementia has a rapidly increasing prevalence throughout the world and poses great challenges to patients and their relatives, including an impaired speech production. Therefore, speech-based diagnostics have significantly gained importance, but are still personnel- and labor-intensive. The ELEMENT study aimed to compare the diagnostic accuracy of the electronic Delta-App offering implemented automated speech recognition with the gold standard of paper and pencil testing in the neuropsychological work-up of dementia.

Materials and Methods: In a prospective, randomized, single-blind, two arm study design with crossover, each subject underwent testing with both modalities, each three months apart. Each subject with cognitive impairment of different etiology and severity was tested with the CERAD (Consortium to Establish a Registry for Alzheimer's Disease) test battery. Data acquisition took place from 2019/03 to 2022/05 at the Department of Psychiatry and Psychotherapy at Saarland University (Homburg/Saar, Germany). The Delta-App was developed by the German Research Center for Artificial Intelligence recorded the test procedures in order to create a speech-based biomarker. The diagnostic performance of the speech biomarker was compared with the conventional testing and the results of healthy controls.

Results: Out of 200 randomized subjects, 148 (74 %) were tested twice and included in the final analysis. The control group comprising 27 subjects had significantly better results in all 18 CERAD-subtests. The test subjects had significantly worse results in their first compared to their second testing in nine subtests. Subjects undergoing CERAD testing by Delta-App had significantly worse results in five subtests compared to conventional paper-pencil testing. The speech biomarker cognition (SB-C) had a good to very good discriminatory power to distinguish between the clinical test groups healthy, mild cognitive impairment and Alzheimer dementia ($p < 0,001$). The performance of the SB-C to correctly classify each subject to the corresponding clinical subgroup was comparable with the speech-based subtests of the CERAD. In some cases, its performance was even superior.

Conclusions: The speech biomarker cognition (SB-C) based on automated speech recognition did not only detect changes in speech of patients with dementia, but also enhanced the diagnostic accuracy of neuropsychological testing. To further improve its performance, other test modalities such as magnet resonance imaging could also be integrated in this composite biomarker in the future.

2. Einleitung

2.1 Etymologie, Semantik und Definition

Das Wort „*Demenz*“ stammt aus dem Lateinischen und setzt sich aus zwei Wortteilen zusammen mit der Präposition „de“, welche „von“ bzw. „weg“ bedeutet und dem Wort „mens/mentis“, welches als Denkvermögen, Gedanken oder auch Charakter übersetzt werden kann. Wörtlich übersetzt bedeutet der Begriff folglich „fern vom Geist“ (1). Damit ist ein wesentliches Merkmal dementieller Erkrankungen bereits beschrieben, die Beeinträchtigung verschiedener kortikaler Funktionen. In der 10. Ausgabe der „International statistical classification of diseases and related health problems“ (ICD-10) wird Demenz als ein Syndrom definiert, welches aufgrund chronischer oder fortschreitender Erkrankungen des Gehirns zur Beeinträchtigung des Gedächtnisses, des Denkens, der Orientierung, der Auffassung, der Lernfähigkeit sowie der Sprache und des Urteilsvermögens führt (2). Neben Einbußen in Bezug auf die Verrichtung von Alltagsaktivitäten („ADL“ Activities of Daily Living), wie z. B. der Körperhygiene oder dem Einnehmen von Mahlzeiten, werden die kognitiven Beeinträchtigungen meist von Veränderungen in den Bereichen des Sozialverhaltens, der emotionalen Kontrolle und der Motivation begleitet (2).

2.2 Epidemiologie: Prävalenz, Inzidenz, Geschlechtsunterschiede

Laut dem World Alzheimer Report (2021) leben weltweit über 55 Millionen Menschen mit einer dementiellen Erkrankung, wobei nach Schätzungen der Alzheimers Disease International (ADI) 75 % der Demenzerkrankungen weltweit nicht diagnostiziert werden (3). In einigen Ländern und Regionen mit geringem bis mittlerem Einkommen, in denen die Stigmatisierung und das mangelnde Bewusstsein für dementielle Erkrankungen ein Hindernis für die Diagnose darstellen, könnte diese Quote sogar 90 % betragen (4).

Aufgrund der wachsenden Lebenserwartung nimmt die Anzahl dementieller Erkrankungen zu, was sich auch in der europäischen Bevölkerung widerspiegelt. Nach Angaben von EuroStat betrug der Anteil der europäischen Bevölkerung ≥ 65 Jahre im Jahr 2020 bereits mehr als 20 %, was einem Anstieg von 3 % in den letzten 10 Jahren entspricht. Die Prävalenz der dementiellen Erkrankungen steigt von etwas mehr als 1 % bei den 65- bis 69- Jährigen auf rund 40 % in der Altersgruppe der über 90-Jährigen an (5). Im Jahr 2018 schätzte Alzheimer's Disease International, dass circa 50 Millionen Menschen weltweit an einer Demenz erkrankt sind und prognostizierte eine Verdreifachung der Prävalenz bis zum Jahr 2050, mit einem großen Teil der Erkrankungen in Ländern mit niedrigem bis mittlerem Einkommen. Laut neuesten Daten wird sich beispielsweise in Europa die Anzahl der Demenzkranken aufgrund der sich verändernden Altersstruktur bis zum Jahr 2050 verdoppeln, sodass bei gleichbleibender Prävalenz fast 19 Mio. Menschen hieran erkrankt sein werden (6). Auch in

Deutschland folgt die Inzidenz dem allgemeinen europäischen Trend, der Anteil der Erkrankten an der Gesamtbevölkerung wird mit geschätzten 2,7 Mio. im Jahr 2050 voraussichtlich 3,4 % betragen (5). In letzter Zeit häufen sich die Belege dafür, dass die Inzidenz dementieller Erkrankungen in Ländern mit hohem Einkommen sinken wird. In diesem Zusammenhang konnten jedoch keine spezifischen Schutzfaktoren identifiziert werden, welche den beobachteten Trend erklären könnten. Es wird vermutet, dass Verbesserungen der Lebensbedingungen, wie z. B. im Bereich der Bildung oder der Versorgung im Gesundheitswesen, die körperliche und geistige Gesundheit positiv beeinflussen und damit das Risiko an einer Demenz zu erkranken verringert haben könnten (7).

Darüber hinaus unterliegt die Prävalenz geschlechterspezifischen Unterschieden. In einer Metaanalyse aus dem Jahr 2019, die die Demenz hinsichtlich ihrer geografischen, Alters- und Geschlechtsverteilung untersuchte, zeigte sich erneut dieser Trend (8). Vor allem im Alter von über 80 Jahren wiesen Frauen eine signifikant höhere Prävalenz im Vergleich zu altersgleichen Männern auf. Diese Dominanz bestand bei allen Demenzarten in fast allen Altersgruppen. Die Prävalenz beispielsweise einer Alzheimer Erkrankung im Alter von 60 bis 69 Jahren war bei Frauen 1,9-mal höher als bei Männern. Auch wenn die Unterschiede in der Prävalenz und Inzidenz dementieller Erkrankungen nicht immer eindeutig sind und je nach Altersgruppe und geografischer Region variieren, gibt es Hinweise darauf, dass Frauen ein höheres Lebenszeitrisko für die Entwicklung einer Demenz haben (9-11). Dieses Risiko ist bisher vor allem auf die höhere Lebenserwartung zurückgeführt worden, wobei die Auswirkungen des Geschlechts und der dadurch entstehenden Ungleichheit auf das Demenzrisiko vermutlich vielschichtiger sind (10, 12).

Hinsichtlich der Anzahl an Neuerkrankungen, also der Inzidenz dementieller Erkrankungen, besteht eine Heterogenität: Einige bevölkerungsbezogene Studien berichten über rückläufige Zahlen der Alzheimer- und Demenzinzidenz (13-15), während andere keine Veränderung des Krankheitsrisikos feststellten (16, 17). Die Studien, die von einem Rückgang der Inzidenz berichten, gehen davon aus, dass ein steigendes Bildungsniveau und besserer Zugang zum Gesundheitssystem zum Rückgang beitragen könnten (7, 18). Auch in einer Metaanalyse aus dem Jahr 2019 wurde in allen der drei untersuchten Altersgruppen (65-74 vs. 75-84 vs. ≥ 85 Jahre) ein signifikanter Zusammenhang zwischen späteren Geburtsjahren und einer geringeren Inzidenz dementieller Erkrankungen beschrieben. Dieser Rückgang wurde ebenfalls auf ein verbessertes Bildungsniveau als auch auf eine Verringerung des kardiovaskulären Risikos zurückgeführt. Zwischen dem Geburtsjahr und der Alzheimer-Inzidenz lag jedoch in keiner Altersgruppe ein signifikanter Zusammenhang vor. Insbesondere die aus westlichen Ländern gemeldeten Alzheimer-Inzidenzen blieben konstant. Dies könnte durch eine verbesserte Überlebensrate ebenso wie eine Verbesserung der Erkennung von Alzheimer bedingt sein (19).

2.3 Risikofaktoren

Da Demenzen typischerweise eine Erkrankung des höheren Alters sind, stellt das Alter einen der wichtigsten Risikofaktoren für eine Entwicklung einer dementiellen Erkrankung dar. Daneben bestehen weitere nicht-modifizierbare familiäre, genetische Prädispositionen, wie z. B. bei der Alzheimer Demenz (20). Es existieren auch diverse modifizierbare Risikofaktoren, wie z. B. geringerer Bildungsstand, kardiovaskuläre Faktoren, Rauchen, Übergewicht, Depression, körperliche Inaktivität, Diabetes und wenig soziale Kontakte, welche die Gefahr für einen kognitiven Abbau und eine Demenz auch beeinflussen können. Im Jahr 2017 wurden durch die Lancet-Kommission für Demenzprävention, -intervention und -versorgung drei weitere Faktoren hinzugefügt: übermäßiger Alkoholkonsum, traumatische Hirnverletzungen und Luftverschmutzung (21) (siehe Tabelle 1).

Tabelle 1. Potenziell modifizierbare Risikofaktoren der Demenz (in Anlehnung an Dementia prevention, intervention, and care: 2020 report of the Lancet Commission).

Lebensalter	Risikofaktor	Prävalenz
früh (< 45 Jahre)	niedriges Bildungsniveau	40 %
mittel (45-65 Jahre)	Gehörverlust	31,7 %
	Schädel-Hirn-Trauma	12,1 %
	Alkoholkonsum (> 21 Einheiten pro Woche)	11,8 %
	Bluthochdruck	8,9 %
	Übergewicht	3,4 %
spät (> 65 Jahre)	Luftverschmutzung	75 %
	Nikotinkonsum	27,4 %
	körperliche Inaktivität	17,7 %
	Depression	13,2 %
	soziale Isolation	11 %
	Diabetes	6,4 %

Kognitive Veränderungen stellen einen gewöhnlichen Prozess des **Alterns** dar. Während einerseits Erlerntes, wie z. B. der Wortschatz eher stabil gegenüber der Alterung des Gehirns sind, nehmen Fertigkeiten wie das konzeptionelle Denken, Gedächtnis und die Verarbeitungsgeschwindigkeit mit zunehmendem Alter ab (22, 23). Neben der normalen Alterung des Gehirns steigt das Risiko, an einer Demenz zu erkranken mit zunehmendem Lebensalter exponentiell an (24). So leiden 3 % der 65-74 Jährigen, 17 % der 75-84-Jährigen und 32 % der über 84-Jährigen an einer Alzheimer Demenz (25).

Die Alzheimer Demenz (AD), welche die häufigste Form dementieller Erkrankungen darstellt, hat darüber hinaus eine starke **genetische Komponente** (26). Bis heute sind mithilfe genomweiter Untersuchungen an sehr großen Kollektiven 34 genetische Risikofaktoren mit der Alzheimer Krankheit in Verbindung gebracht worden (27). Der stärkste genetische Risikofaktor bei sporadischen Formen der AD stellt die häufige Apolipoprotein E ϵ 4-Variante (APOE ϵ 4) dar, die mit einer Verdoppelung des Risikos bei Heterozygotie und etwa Vervierfachung bei Homozygotie einhergeht. (28). Bisher ist es noch nicht geklärt, ob das genetische Risiko die Wirkung der Lebensgewohnheiten auf die Demenz modifiziert (29, 30). **Bildung** ist ein weiterer früher und potenziell modifizierbarer Risikofaktor (21), entweder durch eine direkte Auswirkung auf die Struktur des Gehirns, die zur kognitiven Reserve beiträgt oder durch die Ausprägung gesünderer Lebens- und Verhaltensweisen (31). Die sogenannte kognitive Reserve, mit der intellektuell anregende Aktivitäten wie z. B. Bildung, berufliche Komplexität oder auch Freizeitaktivitäten bezeichnet werden, spielen eine Rolle bei der Vorbeugung von Symptomen einer degenerativen Veränderung (21, 32). Durch entsprechende Stimulation hervorgerufene neuronale Verbindungen und Plastizität vergrößern die zerebrale kognitive Reserve (13, 21). Zusätzlich ist Bildung ein guter Prädiktor für die Position am Arbeitsmarkt, welche wiederum den sozioökonomischen Status beeinflusst (33).

Die **sozioökonomischen Bedingungen** stehen wiederum in Verbindung mit weiteren Risikofaktoren im Erwachsenenalter, wie z. B. Rauchen, körperlicher Aktivität, übermäßigem Alkoholkonsum, Bluthochdruck und Diabetes, die neuropathologische Prozesse auslösen können (32, 34, 35).

Kardiovaskuläre Erkrankungen umfassen viele Pathologien, die das Herz und die Blutgefäße betreffen (36). Vaskuläre Risikofaktoren, wie z. B. Bluthochdruck und Diabetes schädigen das zerebrovaskuläre System und können eine vaskuläre Demenz induzieren. Zusätzlich können sie auch einen Einfluss auf die Neuropathologie der Alzheimer Krankheit mit intra- und extrazellulären Plaques haben (37). Es konnte gezeigt werden, dass Bluthochdruck im mittleren Erwachsenenalter das Risiko einer kognitiven Beeinträchtigung im Alter deutlich erhöht (38). Zusätzlich besteht ein Zusammenhang zwischen der Dauer des Bluthochdrucks und dem Auftreten kognitiver Beeinträchtigungen (39).

Studien über den Zusammenhang zwischen **körperlicher Aktivität** und dem Risiko der Entstehung einer dementiellen Erkrankung sind komplex und umfassen verschiedenste Variablen, wie Muster der körperlichen Aktivität, Alter, Geschlecht, soziale Schicht oder Kultur (21). Metaanalysen von Längsschnittstudien konnten dennoch belegen, dass körperliche Aktivität mit einem geringeren Demenzrisiko verbunden ist (21, 40).

Einen vielversprechenden Einfluss auf positive Verhaltensweisen als auch auf die Förderung kognitiver Reserven haben zuletzt **soziale Kontakte**. Trotz kultureller Unterschiede in der

Beurteilung und Wahrnehmung von zwischenmenschlicher Isolation kommen Metaanalysen weitestgehend konsistent zu dem Ergebnis, dass soziale Kontakte sowohl für Männer als auch Frauen einen Schutzfaktor hinsichtlich des dementiellen Risikos darstellen (41, 42).

2.4 Formen dementieller Erkrankungen

2.4.1 Demenz bei Alzheimer Krankheit

Die von dem deutschen Neurologen und Psychiater Alois Alzheimer erstmals im Jahr 1907 beschriebene Demenz bei Alzheimer Krankheit stellt die häufigste dementielle Erkrankung dar (43). Dafür kennzeichnend sind zwei pathophysiologische Veränderungen des Gehirns, mit intrazellulären Neurofibrillenbündeln (engl. intracellular neurofibrillary tangles) und extrazellulären Ablagerungen, s. g. Amyloid-Plaques (engl. extracellular amyloid plaques) (44).

Im gesunden Zustand spielt das Tau-Protein bei der Stabilität der Mikrotubuli, welche das Zytoskelett der Axone bilden, und im axonalen Transport innerhalb von Neuronen eine Rolle. Im krankhaften Zustand kommt es zu einer Hyperphosphorylierung des Tau-Proteins, welche zu einer Destabilisierung der Mikrotubuli und somit zu einem Zerfall des Zytoskeletts mit Bildung neurofibrillärer Bündel führt, die sich vor allem im neuronalen Zellkörper ablagern. Sowohl der Untergang des Zytoskeletts als auch die Neurofibrillen induzieren sodann ein neuronales Absterben, weswegen dieser Zustand auch als „Taupathie“ bezeichnet wird (45).

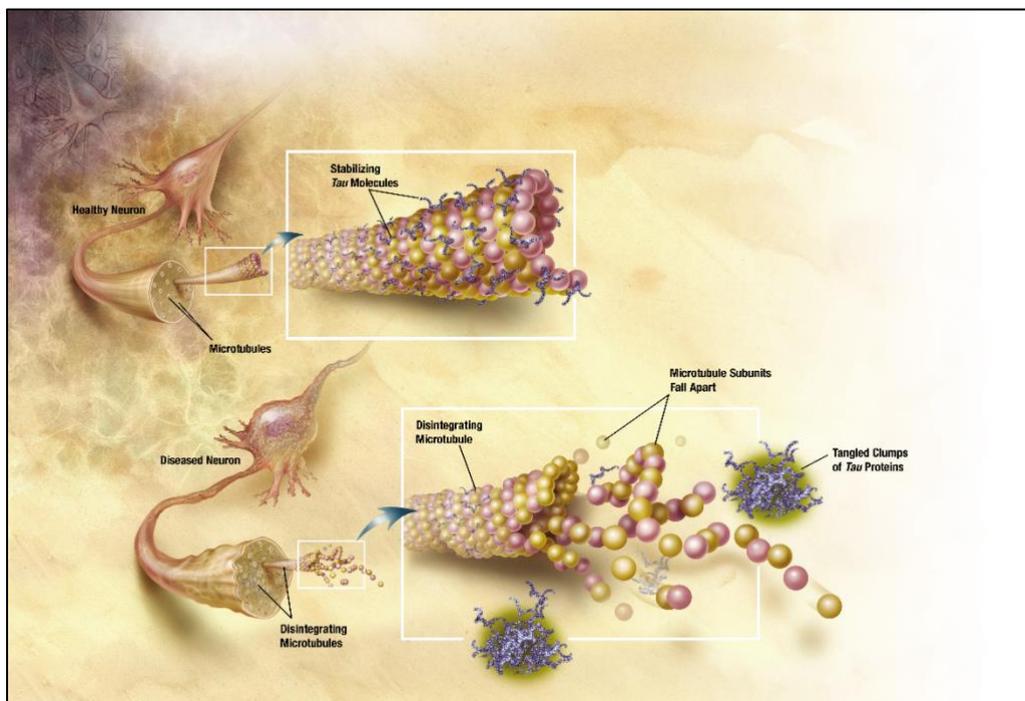


Abbildung 1. Destabilisierung der Mikrotubuli durch Hyperphosphorylierung des Tau-Proteins (aus https://en.wikipedia.org/wiki/Alzheimer%27s_disease#Causes).

Das Amyloid Vorläufer Protein (Amyloid-Precursor-Protein APP) ist ein Transmembranprotein, das u. a. in Synapsen gehäuft vorkommt und als Oberflächenrezeptor der Formation von Synapsen und neuronalen Plastizität dient (46). Im Rahmen der Entstehung einer Alzheimer Demenz wird APP von den ebenfalls physiologischerweise vorkommenden Transmembran-Proteasen beta (47)- und γ -Sekretase (48) gespalten. Aufgrund der Lokalisation der Schnittstellen (vgl. Abbildung 2) entsteht ein 39 bis 43 Aminosäuren langes Protein, Amyloid beta, das im extrazellulären Raum die Tendenz hat, zu verklumpen und s. g. beta-Amyloid Plaques zu bilden (vor allem Amyloid beta 42, auch A β 42) (49). Die Amyloid-Beta-Plaques können einerseits nicht von körpereigenen Prozessen abgebaut werden (s. g. gestörte Clearance) (50). Andererseits schädigen sie als s. g. Oligomere durch die Aktivierung einer neurotoxischen Kaskade die Neuronen und damit auch das gesamte Gehirn, was zum Verlust neuronaler Vernetzungen und zum Zelltod führt (51-53). Neben dem Verlust an Nervenzellen werden zusätzlich entzündliche Prozesse angeregt, die letztlich eine zelluläre Atrophie bedingen (53).

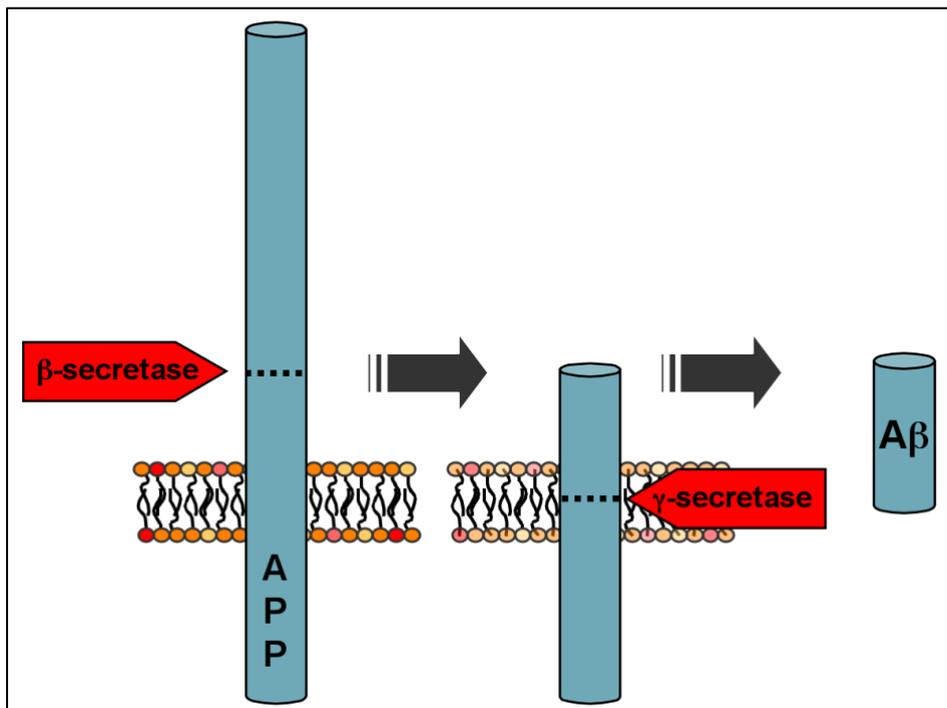


Abbildung 2. Schematische Darstellung der Entstehung des β -Amyloids als Vorläufer der Amyloid-Plaques: die membranständigen β - und γ - Sekretase zerschneiden das APP-Protein (aus https://en.wikipedia.org/wiki/Beta-secretase_1).

Die Alzheimer Erkrankung kann in familiäre Fälle mit Vererbung (familiäre Alzheimer-Krankheit, FAD) und sporadische Fälle ohne familiäre Häufung unterteilt werden. Bei den familiären Fällen handelt es sich vorwiegend um Fälle mit frühem Beginn (jünger als 65 Jahre, early onset), es werden aber auch Fälle mit spätem Auftreten beschrieben (54). Die familiär

vererbare Variante wird typischerweise autosomal-dominant vererbt (55) und kann auf das Vorhandensein von Mutationen direkt im APP oder Presenilin PSEN1 oder PSEN2, die einer Untereinheit der gamma-Sekretase entsprechen, zurückgeführt werden (56). Diese bedingen ein verstärktes Auftreten von Amyloid beta 42, der Hauptkomponente der Amyloid-Plaques (57). Insgesamt treten allerdings mehr als 90% der Fälle sporadisch auf und haben in der Regel einen späteren Erkrankungsbeginn (54).

2.4.2 Vaskuläre Demenz

Die vaskuläre Demenz (VaD) ist nach der Alzheimer Demenz die zweithäufigste Ursache dementieller Erkrankungen. Sie stellt eine heterogene Gruppe von Hirnerkrankungen dar, bei der kognitive Beeinträchtigungen auf zerebrovaskuläre Pathologien zurückzuführen sind (58). Eine VaD kann als Folge globaler oder fokaler Gefäßerkrankungen entstehen (59). Aus dem Konzept der vaskulären Demenz hat sich der Begriff der vaskulären kognitiven Beeinträchtigung entwickelt (VCI). Die VCI ist ein weiter gefasstes Konzept, das sowohl das Spektrum vaskulärer leichter kognitiver Beeinträchtigungen (vascular mild cognitive impairment, VMCI) bis zur VaD umfasst, aber auch gemischte Pathologien einschließt (60, 61). Verschiedene Studien, die sich sowohl mit der Bildgebung als auch der Pathologie dementieller Erkrankungen beschäftigten, konnten verschiedene vaskuläre Ursachen und Hirnparenchymläsionen identifizieren, die mit kognitiven Beeinträchtigungen einhergehen (62): Multi-Infarkt-Demenzen, strategische Infarkte, subkortikale ischämische VaD, Hirnblutungen (hämorrhagische Demenz), globale Hypoperfusionen (Hypoperfusionsdemenz) sowie gemischt vaskuläre Erkrankungen.

Multi-Infarkt Demenzen sind auf mehrere kleine oder große Hirninfarkte zurückzuführen, wobei größere Infarktvolumina und eine höhere Anzahl territorialer oder kleiner subkortikaler Infarkte mit schlechteren kognitiven Leistungen und einem höheren Risiko, eine Demenz zu entwickeln, in Verbindung gebracht werden (63). Bei einem s. g. „**strategischen Infarkt**“ kann ein einzelner kleiner Infarkt jedoch auch schwere kognitive Defizite verursachen. Die anatomischen Lokalisationen der strategischen Infarkte umfassen den Thalamus, den Gyrus angularis und die Basalganglien (64-66).

Die häufigste Ursache einer vaskulären kognitiven Beeinträchtigung ist eine **zerebrale Mikroangiopathie** (small vessel disease SVD), die durch Läsionen der weißen Substanz und lakunenartige Infarkte verursacht wird (67, 68). Sowohl makroskopische intrazerebrale Blutungen als auch Mikroblutungen werden mit kognitiven Beeinträchtigungen in Verbindung gebracht (69, 70). Bei der **Hypoperfusionsdemenz** können globale Reduktionen der zerebralen Perfusion zu vorübergehenden oder dauerhaften Ischämien und damit zu kognitiven Defiziten führen (71, 72).

2.4.3 Frontotemporale Demenz

Die Frontotemporale Demenz (FTD) ist eine heterogene Erkrankung mit unterschiedlichen Erscheinungsbildern. Sie umfasst klinische Störungen mit Veränderungen des Verhaltens und der Sprache sowie der exekutiven Kontrolle und Motorik. Zu den Störungen des FTD-Spektrums gehören die verhaltensbezogene FTD (behavioral variant frontotemporal dementia), nicht-flüssig / agrammatische und die semantische Variante der primär progressiven Aphasie (primär progressive aphasia, PPA) (73).

Die **verhaltensbezogene Variante** der FTD verursacht Veränderungen im Verhalten, in der Persönlichkeit, im Bereich der Emotionen und der exekutiven Kontrolle (74). Diese Abweichungen können Verhaltenssymptome wie Enthemmung, Zwänge oder Symptome wie auch Apathie und mangelnde Empathie umfassen. Da diese Symptome psychiatrischen Erkrankungen ähneln, besteht bei Patienten ein hohes Risiko für Fehldiagnosen (75).

Im Gegensatz dazu handelt es sich bei der **primär progressiven Aphasie** um ein Syndrom, bei dem Auffälligkeiten in der Sprache die primäre Beeinträchtigung darstellen (76). Bei der *semantischen Variante der PPA* hängt die Symptomatik davon ab, welcher Teil des Temporallappens betroffen ist. Wenn der linke Schläfenlappen betroffen ist, bestehen vor allem Probleme sprachlicher Natur mit einem langsamen Verlust semantischen Wissens (beeinträchtigt Objektwissen, beeinträchtigt Einzelwortverständnis bei erhaltener Syntax und Grammatik (77)). Bei Affektion der rechten Seite sind überwiegend Verhaltenssymptome (emotionale Distanz, Reizbarkeit, Schlaf-, Appetit- und Libidostörungen (78)) auffällig. Mit Fortschreiten der Erkrankung sind zumeist beide Gehirnhälften betroffen und die Symptome beginnen sich zu überschneiden (79). Bei der *nichtflüssigen/agrammatischen Variante* der PPA bestehen bei den Betroffenen zu Erkrankungsbeginn Probleme bei der Wortfindung und beim Sprechen. Die Sprache wird langsamer, undeutlicher und abgehackter mit veränderter Prosodie (80). Sprachliche Probleme treten oft vor einer ausgeprägten Sprechapraxie oder einem Agrammatismus auf (77). Das neuroanatomische Korrelat für die Symptomatik findet sich in den Brodmann-Arealen 44 und 45 im linken inferioren frontalen Gyrus und der anterioren Insula (81).

2.4.4 Gemischte Demenz

Bei der gemischten Demenz handelt es sich um eine heterogene Störung, die sich meist auf ein gleichzeitiges Vorliegen neurodegenerativer (wie bei der Alzheimer Demenz) und zerebrovaskulärer Schädigungen bezieht (82). Nach dem ICD-10 wird die Diagnose einer gemischten Demenz gestellt, wenn sowohl die Kriterien für eine Alzheimer Erkrankung als auch für eine Demenz vaskulärer Genese erfüllt sind (2). In der S3-Leitlinie „Demenzen“ werden die Kriterien der International Working Group (IWG) aufgeführt, die neben der o. g. Koexistenz auch ein gleichzeitiges Vorliegen einer Alzheimer- als auch einer Lewy-

Körperchen-Pathologie als gemischte Demenz bezeichnen (83). In der 5. Auflage des „Diagnostic and Statistical Manual of Mental Disorders (DSM-5)“ wird die Diagnose unter dem Begriff „Major neurocognitive disorder due to multiple etiologies“ eingeordnet, wenn aus der Anamnese und den Untersuchungen sowie den Befunden hervorgeht, dass die Störung auf mehrere pathologische Prozessen zurückzuführen ist (84). Vaskuläre Hirnläsionen erhöhen das Risiko der Entstehung einer Alzheimer-Pathologie (85) und umgekehrt erhöht die Alzheimer-Pathologie das Demenzrisiko nach einem Schlaganfall (86).

2.4.5 Demenz mit frühem Beginn

Eine Demenz mit frühem Beginn („early onset dementia“, EOD) ist typischerweise eine dementielle Erkrankung bei Menschen unter 65 Jahren (87, 88), unabhängig vom zugrunde liegenden Demenzsyndrom (89). Bei Auftreten vor dem 45. Lebensjahr wird zusätzlich von einer „young onset dementia, YOD“ gesprochen (88, 90). Bei diesen Patienten ist die korrekte Diagnosestellung erschwert (91), da sie häufig eine atypische Symptomatik mit diversen neurologischen Auffälligkeiten aufweisen können (92).

Der Begriff „früh einsetzende Demenzen“ bezieht sich auf eine Gruppe fortschreitender neurodegenerativer Erkrankungen, wie z. B. die Alzheimer Krankheit (AD), die frontotemporale Demenz (FTD) oder die Lewy-Body Demenz (93). Die Symptome der EOD ähneln denen der später einsetzenden AD oder FTD. Aufgrund der höheren Prävalenz frontotemporaler Demenzen ist das klinische Bild eher durch neuropsychiatrische Symptome, wie z. B. die Veränderung des sozio-emotionalen Verhaltens und des Verständnisses geprägt (94, 95). Da eine frühe Demenz erhebliche Auswirkungen auf die Patienten, deren Arbeitsplatz, Einkommen (96) und auch deren soziales Umfeld mit der Familie hat, zu denen auch kleine Kinder gehören können (97), ist es von großer Bedeutung, eine korrekte Diagnose zu stellen, um die Hilfestellungen und Bereitstellung von Ressourcen besser planen zu können (89).

2.4.6 Demenz bei andernorts klassifizierten Krankheiten

Die Demenz mit Lewy Körperchen (Lewy-Body-Dementia, LBD) und die Demenz bei Parkinson (Parkinson Disease Dementia, PD-D) sind beides neurodegenerative Erkrankungen, die mit dem Vorhandensein von Lewy-Körperchen in Verbindung gebracht werden, gemeinsame neuropathologische Befunde und ein ähnliches klinisches Bild aufweisen. Zu den klinischen Merkmalen beider dementieller Erkrankungen gehören kognitive Beeinträchtigungen, Verhaltensauffälligkeiten, autonome Dysfunktionen, Schlafstörungen und der Parkinsonismus. Die kognitiven Beeinträchtigungen sind durch Einschränkungen der exekutiven und visuell-räumlichen Funktionen sowie der Aufmerksamkeit gekennzeichnet (98). Die Unterscheidung beider Demenzformen anhand des klinischen Bildes basiert auf einer

willkürlichen Unterscheidung zwischen dem Zeitpunkt des Auftretens von Parkinsonismus und den kognitiven Symptomen (99). Zu den pathologischen Merkmalen der Lewy-Body-Demenz gehören Lewy-Körperchen (lewy body, LB) und Lewy-Neuriten (intraneuronale Einschlüsse). Sie können in verschiedenen Bereichen des Gehirns auftreten, wie z. B. in den Kernen des Hirnstammes, der Amygdala, den Basalganglien, den limbischen sowie paralimbischen Kortizes und dem Neokortex als auch in der Medulla und dem peripheren autonomen Nervensystem (100). Die pathologischen Veränderungen bei Morbus Parkinson sind durch eine Kombination einer Degeneration in den subkortikalen Kernen und einem kortikalen Zellverlust in Verbindung mit einer LB- und AD-Pathologie gekennzeichnet (101). Die Kombination beider Pathologien ist ein robustes pathologisches Korrelat der Demenz im Rahmen einer Parkinson Erkrankung (102).

2.5 Verlauf und Prognose

Das Fortschreiten von Veränderungen im Gehirn, die bei den Betroffenen zunächst unbemerkt bleiben, bis hin zu Veränderungen, die zu kognitiven Beeinträchtigungen und am Ende auch zu körperlichen Behinderungen führen, wird als **Kontinuum der Alzheimer-Krankheit** bezeichnet (s. Abbildung 3) (53). Der zugrundeliegende pathophysiologische Prozess der Erkrankung als auch die klinische Symptomatik können sich parallel, aber auch zeitlich versetzt entwickeln. Es wird davon ausgegangen, dass zwischen der Pathologie der Alzheimer-Demenz und dem Auftreten klinisch auffälliger Beeinträchtigungen eine Spanne von zehn Jahren oder mehr liegen kann (103).

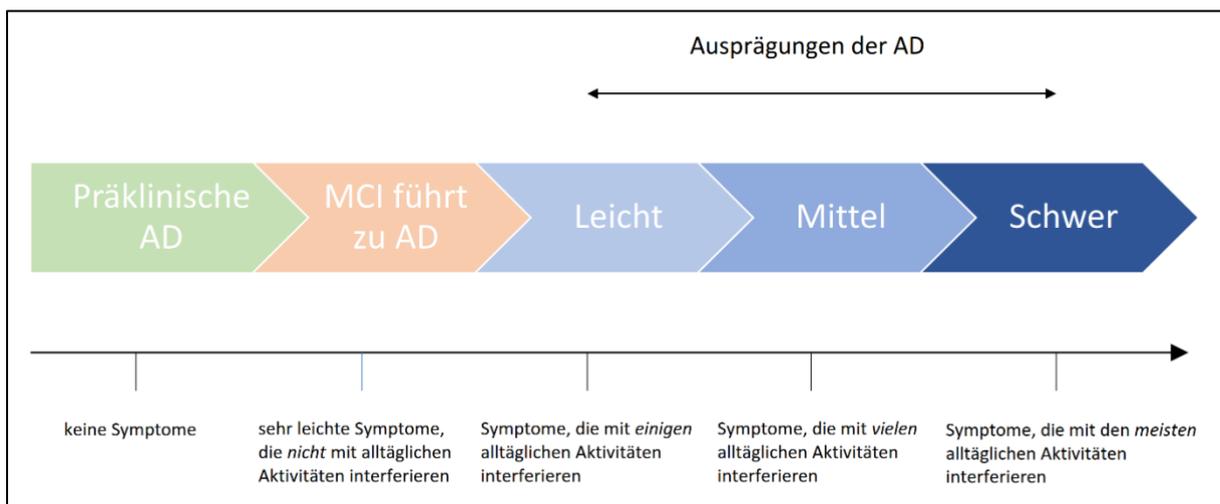


Abbildung 3. Kontinuum der Alzheimer Erkrankung (angelehnt an Alzheimer’s disease facts & figures, 2020).

Der Internationalen Arbeitsgruppe (International Working Group, IWG) zufolge erfordern **präklinische Zustände** der Alzheimer-Demenz das Vorhandensein eines Biomarkers, der für

die Alzheimer Pathologie typisch ist sowie das Fehlen klinischer Anzeichen (104). In die **Übergangsphase** lassen sich Patienten mit einer präklinischen Alzheimer Demenz einordnen, bei denen die Kognition nicht mehr vollständig normal ist, die aber auch noch nicht die Kriterien einer leichten kognitiven Störung erfüllen (105). Bei den Betroffenen tritt eine Verschlechterung der Leistungsfähigkeit ein, welche sich in der objektiven neuropsychologischen Untersuchung (noch) nicht widerspiegelt, da die Ergebnisse noch normwertig sind.

Das **prodromale Stadium** der Erkrankung wird auch als leichte kognitive Beeinträchtigung (**Mild cognitive impairment, MCI**) bezeichnet, welche durch kognitive Beeinträchtigung mit minimaler Einschränkung der Aktivitäten des täglichen Lebens gekennzeichnet ist. Die Bezeichnung amnestische MCI (aMCI) beschreibt ein Krankheitsbild, bei dem Störungen des Gedächtnisses dominieren. Bei der nicht-amnestischen MCI hingegen ist die Störung anderer kognitiver Funktionen, z. B. der Sprache oder der exekutiven Funktionen, deutlicher ausgeprägt (106). Patienten mit einer leichten kognitiven Beeinträchtigung (MCI) können eine Demenz entwickeln, andere verbleiben in diesem Stadium der Einschränkung oder kehren in den kognitiven Normalzustand zurück. Unter den Patienten mit einer MCI entwickeln zwischen 30-50 % innerhalb der nächsten fünf bis zehn Jahre eine Alzheimer Demenz (107), dabei haben Personen mit einer aMCI mit mehreren defizitären Domänen das größte Risiko (108). Alzheimer-spezifische Biomarker könnten sich zur Vorhersage einer Demenz vom Alzheimer Typ bei Patienten mit einer amnestischen sowie nicht-amnestischen MCI eignen. Jedoch sind die Biomarker zur frühen Diagnose einer AD bei nicht-amnestischer im Vergleich zu amnestischer Form möglicherweise weniger empfindlich (109).

Menschen im Alter ≥ 65 Jahren versterben nach der Erstdiagnose einer Alzheimer Demenz im Durchschnitt vier bis acht Jahre später (53). Einer Literaturübersichtsarbeit aus dem Jahr 2012 zufolge, die sich mit der Überlebenszeit bei dementiellen Erkrankungen befasst, nimmt das Risiko, an einer Demenz zu versterben, mit zunehmendem Alter bei allen untersuchten Subtypen der Demenz (AD, VaD, FTD, DLB) ab, es besteht also ein negativer Zusammenhang zwischen Alter und Sterberisiko. Es konnte kein geschlechterspezifischer Unterschied festgestellt werden. Die Reduktion der Lebenserwartung scheint bei einer frontotemporalen Demenz und einer vaskulären Demenz am stärksten zu sein. Dies ist darauf zurückzuführen, dass das Alter bei Patienten mit einer FTD im Allgemeinen niedriger ist und dass die vaskuläre Demenz häufiger mit Komorbiditäten in Verbindung steht im Vergleich zu anderen Demenz-Subtypen (110).

2.6 Kosten

Im Jahr 2010 hat Alzheimer Disease International (ADI) Schätzungen über die weltweiten wirtschaftlichen Folgen dementieller Erkrankungen vorgelegt (111). Die finanzielle Belastung im Jahr 2010 wird auf 604 Milliarden US-Dollar geschätzt, was auf schwerwiegende globale

sozioökonomische Folgen dieser Erkrankung hindeutet (112). Binnen der folgenden fünf Jahre, in denen die Prävalenz deutlich gestiegen ist, haben die Kosten um 35,4 % auf 818 Milliarden US-Dollar zugenommen. Die Ursachen für diese Kostenzunahme lassen sich sowohl auf die höhere Anzahl Demenzerkrankter als auch auf die höheren Kosten pro Person zurückführen (112).

Auch die vermehrte Inanspruchnahme des Gesundheitswesens infolge der Erkrankung wurden bereits untersucht. Demenzpatienten haben im Vergleich zu kognitiv Gesunden mehr Komorbiditäten (113), welche das Fortschreiten der Demenz verschlimmern können (114, 115). Der kognitive Abbau kann die Fähigkeiten, das Management chronischer Erkrankungen selbst zu übernehmen, mindern oder gänzlich verhindern (116). Dieses Zusammenspiel zwischen den Belastungen durch Komorbiditäten und Demenz kann zu häufigeren und längeren Krankenhausaufenthalten führen (117). Die Kosten der Gesundheitsfürsorge und der Langzeitpflege von Patienten mit einer Alzheimer-Demenz oder anderen Demenzerkrankungen sind erheblich (118), wobei der Wert der informellen Pflege häufig nicht berücksichtigt wird (53). Im Jahr 2019 leisteten in den Vereinigten Staaten fast 16 Mio. Menschen unbezahlte Pflege, sprich informelle Hilfe für Menschen mit einer Alzheimer-Demenz oder anderen dementiellen Erkrankungen (53). Diese Hilfe wird in der Regel von Mitgliedern der Familie, Freunden oder anderen unbezahlten Betreuungspersonen geleistet (119). Die Pflege Demenzkranker stellt dabei eine besondere Herausforderung dar: wenn sich die Symptomatik verschlimmert, kann die von Familienangehörigen geleistete Pflege zu erhöhtem emotionalem Stress, zu Problemen der eigenen Gesundheit sowie zur Unterbrechung der beruflichen Beschäftigung führen, was mit finanziellen Einbußen einhergeht (120-122).

Um das Voranschreiten der Erkrankung hinauszuzögern und die damit verbundenen Gesundheitskosten zu reduzieren ist es notwendig, die Diagnose frühzeitig zu stellen (123). Die Diagnosestellung einer dementiellen Erkrankung umfasst einerseits die Klärung der Ursache des demenziellen Syndroms als auch das Erkennen des Vorhandenseins einer Demenz. In einer Studie aus dem Jahr 2017 sind die Kosten für die Diagnosestellung bei Patienten mit Verdacht auf Demenz in einer deutschen Gedächtnisklinik analysiert worden. Die Gesamtkosten pro Patient mit Verdacht auf eine dementielle Erkrankung betrug für den gesamten Diagnoseprozess 501 € (inklusive 110 € für die klinische Beratung und 391 € für die technischen Verfahren). Bei den Patienten, die eine Demenzdiagnose erhielten, waren die durchschnittlichen Kosten bis zur Diagnosestellung mit 659 € deutlich höher (124).

2.7 Diagnostik

Gemäß den aktuell gültigen S3-Leitlinien „Demenzen“ (83) umfasst der diagnostische Prozess eine Anamnese, körperliche sowie psychopathologische Untersuchung, neuropsychologische Diagnostik und labordiagnostische Untersuchung sowie eine zerebrale Bildgebung.

Die diagnostische Aufarbeitung beginnt mit einer ausführlichen **Anamnese**, die den bisherigen Verlauf der Symptomatik im Zusammenhang mit zuvor bestehenden somatischen und psychischen Erkrankungen klären soll. Aufgrund von bestehenden kognitiven Einschränkungen des Patienten ist neben der Eigenanamnese eine Fremdanamnese z. B. zur Erfassung von Beeinträchtigungen alltagsbezogener Fähigkeiten von großer Bedeutung. Bei der anschließenden **körperlichen Untersuchung** werden insbesondere kardiovaskuläre, metabolische oder endokrinologische Veränderungen abgeklärt. Neben der allgemeinen internistischen körperlichen Untersuchung ist eine Erhebung möglicher neurologischer Ursachen einer Demenz (wie z. B. eine Parkinson-Symptomatik oder eine intrakranielle Blutung) sinnvoll. Der **psychopathologische Befund** kann Hinweise für psychiatrische Differenzialdiagnosen liefern.

Um die in der Anamnese gewonnenen Erkenntnisse und Informationen zu objektivieren, eignen sich zur orientierenden Einschätzung von kognitiven Defiziten **kognitive Kurztests oder Screening-Verfahren** wie z. B. der Mini-Mental Status-Test (MMST) oder der Demenz-Detektionstest (DemTect). Im Anschluss kann eine vertiefende **neuropsychologische Diagnostik** sinnvoll sein, um ein kognitives Profil zu erstellen, häufig wird hierfür die „Consortium to Establish a Registry for Alzheimer’s Disease Neuropsychological Battery“ (CERAD-NP) genutzt. Daneben finden weitere neuropsychologische Testverfahren wie die „Alzheimer’s Disease Assessment Scale-cognitive Subscale“ (ADAS-cog) oder auch Ratingskalen zur Charakterisierung des Schweregrades der Demenz wie die „Clinical Dementia Rating Scale“ (CDR) Anwendung. Die **Labordiagnostik** beinhaltet neben dem Blutbild, eine Bestimmung der Elektrolyte, des Blutzuckers und des Thyreoidea stimulierenden Hormons (TSH-Wert), der gamma-Glutamyltransferase (γ – GT) , des Serum-Kreatinins sowie des Serum-Harnstoffes, des Vitamin B12 und des Folsäure-Spiegels als Standarddiagnostik. Als **bildgebende Verfahren** im Rahmen der Diagnostik soll konventionell eine cranielle Computertomographie (cCT) oder eine cranielle Magnet-Resonanz-Tomographie (cMRT) durchgeführt werden.

Zur **ergänzenden Diagnostik** können eine *Liquordiagnostik* mit Bestimmung klinisch relevanter Parameter (beta-Amyloid, Gesamt-Tau und Phospho-Tau) sowie *nuklearmedizinische Verfahren* durchgeführt werden. Zu diesen Verfahren zählt die Messung des Glukosemetabolismus mittels Positronen-Emissionstomographie mit einer schwach radioaktiv markierten 18F-Fluordesoxyglucose (FDG-PET) und das Amyloid-PET zur Darstellung des zerebralen Amyloids, welche bei Unsicherheit in der Differenzialdiagnostik zur

Klärung nützlich sein können. Auch *humangenetische Untersuchungen* werden nicht standardmäßig durchgeführt. Bei früh beginnender Symptomatik oder bei einer familiären Häufung, die durch eine positive Familienanamnese festgestellt wurde, soll eine genetische Beratung angeboten werden (83).

2.7.1 Früherkennung

Eine Früherkennung bei voranschreitenden neurodegenerativen Erkrankungen wie der Demenz ist besonders bedeutsam, da sie auf mannigfaltige Lebensbereiche Auswirkungen haben kann. Bei einer frühzeitigen Diagnosestellung einer Alzheimer Demenz können das Nachlassen der kognitiven Funktion und die Verhaltenseinschränkungen in den anfänglichen Stadien außerdem durch den Einsatz von Arzneimitteln wie z. B. Acetylcholinesterase-Hemmern verlangsamt werden (125, 126). Zusätzlich verbessert eine frühzeitige Diagnosestellung die Organisation geeigneterer Unterstützungsmaßnahmen nicht nur für die Patienten, sondern auch für die Betreuer und Angehörigen. Es ist erwiesen, dass durch die frühzeitige Erkennung die psychischen Belastungen bei den Betroffenen und ihren Familien reduziert werden können (127). Ein rechtzeitiges, aufklärendes Gespräch mit den Familienangehörigen über den Verlauf und die Prognose der Erkrankung kann ihnen helfen, rechtliche Vorkehrungen zu treffen und im zwischenmenschlichen Zusammenleben darauf zu achten, mögliche Einbußen der Lebensqualität der Betroffenen abzumildern (128). Zusätzlich kann eine frühzeitige Diagnose von Vorteil sein, da bei einer geeigneten Behandlung eine Einweisung in eine Pflegeeinrichtung hinausgezögert werden kann (129).

Derzeit existieren verschiedene Arten von **Biomarkern**, die zur Früherkennung beitragen könnten. Hinsichtlich der genetischen Biomarker sind bei Alzheimer Erkrankungen Mutationen im Amyloid-Vorläuferprotein (APP) oder in den Presenilin-Enzymen 1 und 2 bekannt. Diese beiden Mutationen sind jedoch nur für einen kleinen Teil der familiären Alzheimer-Krankheit verantwortlich. Die meisten AD-Fälle treten jedoch später und sporadisch auf (130). Zu den flüssigen und bildgebenden Biomarkern zählen Amyloid- β und Tau im Liquor (CSF), die Amyloid-Positronenemissionstomographie (PET), die Atrophie des mesialen Temporallappens in der Magnetresonanztomographie (MRT) sowie der Tempoparietale-Hypometabolismus in der 18F-Fluorodeoxyglucose (FDG-PET; s.o.) (131). Diese AD-Biomarker, die z. B. mit Hilfe einer Liquor-Diagnostik oder eines Amyloid-PETs erhoben werden, sind aber invasiv, zeitaufwendig und kostenintensiv. Daher ist es nicht möglich und auch nicht sinnvoll, solche Biomarker bei allen Patienten bspw. im Rahmen einer Screening-Untersuchung anzuwenden (132).

Momentan erfolgt die Früherkennung von Demenzen neben den invasiven Methoden größtenteils durch **kognitive Tests** mittels Papier und Bleistift, die aber ebenfalls sehr zeitaufwendig und in der präklinischen Phase manchmal nicht sensitiv genug sind. Der

Dementia-Detection (DemTect) oder das Montreal Cognitive Assessment beispielsweise sind demgegenüber Screening-Instrumente, die von einem geschulten Prüfer innerhalb kurzer Zeit durchgeführt werden können. Beide weisen jedoch Lerneffekte auf Seiten des Patienten auf, die den Nutzen bei wiederholtem Einsatz begrenzen. Daher ist es notwendig, kostengünstige, nicht invasive Screening-Instrumente zu entwickeln, die wiederholt angewendet werden können und vielleicht auch aus der Ferne durchführbar sind (133). Bei dementiellen Störungsbildern kommt es neben den Gedächtnisstörungen auch zu erheblichen Veränderungen in der Sprachproduktion, vor allem in Bezug auf lexikalisch-semantische Merkmale. Aus diesem Grund haben in der letzten Zeit **sprachbasierte Methoden** in der Demenzdiagnostik, wie z. B. die automatische Spracherkennung (Automatic Speech Recognition, ASR), an Bedeutung gewonnen. Diese sind nicht invasiv, kostengünstig und eignen sich zur Diagnostik in den frühen Stadien einer Demenz (132). Bisher haben mehrere Forschungsstudien verschiedene Varianten der sprachlichen Analyse, z. B. die Erfassung linguistischer Veränderungen bei strukturierten Sprachaufgaben oder bei Spontansprachaufgaben (134-136) zur Erkennung von dementiellen Störungsbildern vorgeschlagen, die Ergebnisse sind allerdings sehr heterogen, nur schwer vergleichbar und noch nicht in den klinischen Alltag übertragbar.

2.8 Fragestellung

Aufgrund der steigenden Prävalenz, hohen Gesundheitskosten und der erheblichen Belastung dementieller Erkrankungen für Betroffene und das soziale Umfeld besteht derzeit ein großes Interesse an geeigneten, standardisierten und nicht-invasiven Früherkennungsverfahren. Daher ist das Ziel dieser Arbeit, das Potential digitaler, automatisch extrahierbarer Sprachbiomarker zur Verbesserung der Diagnostik bei dementiellen Erkrankungen zu untersuchen. Es handelt sich dabei um eine gemeinsame prospektive Erhebung der Klinik für Psychiatrie und Psychotherapie des Universitätsklinikums Homburg/Saar und der Forschungsgruppe Intelligente Benutzerschnittstellen (IUI) des Deutschen Forschungsinstituts für künstliche Intelligenz (DFKI GmbH).

Im Rahmen der s. g. ELEMENT-Studie wurden dazu prospektiv 200 Probanden mit einer dementiellen Erkrankung untersucht. In einem zeitlichen Abstand von drei Monaten erfolgte die Testung sowohl mit dem konventionellen Paper Pencil-Verfahren als auch einer neu entwickelten App auf einem Apple iPad® (Apple Inc., Cupertino, Kalifornien, USA), das zusätzlich zur Untersuchung eine Sprachdatei aufzeichnete, die danach weiter analysiert wurden. Die Ergebnisse wurden mit denen gesunder Probanden verglichen.

Konkret wurden folgende Fragestellungen untersucht:

1. Wodurch ist das Gesamtkollektiv charakterisiert? Unterscheidet sich die Studiengruppe im Vergleich zu den Probanden, die aus der Studie ausgeschlossen wurden (Dropout-Gruppe)? Unterscheiden sich Studien- und Kontrollgruppe hinsichtlich der Testergebnisse?
2. Wird das Testergebnis der Studienpatienten durch die Testmodalität (App vs. Paper-Pencil) oder den Testzeitpunkt beeinflusst?
3. Unterscheiden sich die Testergebnisse bei Patienten mit verschiedenen Demenzformen?
4. Können aus den Sprachdateien, die während der Durchführung sprachgebundener kognitiver Testverfahren erfasst werden, mittels Sprachanalyse digitale Biomarker extrahiert werden, die unterschiedliche Stadien einer dementiellen Erkrankung charakterisieren (insbesondere des MCI und der Alzheimer Krankheit)?
5. Wie ist die Aussagekraft des digitalen Sprachbiomarkers im Vergleich zu den sprachbasierten Subtests der traditionellen Paper-Pencil Testung des CERADs im Bereich des Screenings von MCI und Alzheimer?

3. Material und Methoden

3.1 ELEMENT-Studie

Die „ELEMENT-Studie: Zur Früherkennung kognitiver Störungen mittels Sprachanalyse“ ist eine prospektiv-randomisierte klinische Prüfstudie, die von März 2019 bis Mai 2022 Patienten rekrutierte. Alle Untersuchungen fanden in den Räumlichkeiten der Klinik für Psychiatrie und Psychotherapie der Universitätsklinik des Saarlandes (Homburg) statt. Die Studiendurchführung erfolgte gemäß den ethischen Richtlinien für die medizinische Forschung am Patienten der Deklaration von Helsinki. Sie wurde von der zuständigen Ethik-Kommission der Ärztekammer des Saarlandes (Kenn-Nr. 65/18), dem Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM, Bescheid vom 22.10.2018) und des Landesamtes für Umwelt und Arbeit im Saarland (Antragsnummer 11624) genehmigt.

Die ELEMENT-Studie zielte darauf ab, die diagnostische Genauigkeit, also Sensitivität und Spezifität der Computer-basierten Spracherkennung im Rahmen einer neuropsychologischen Demenzabklärung mit der „klassischen“, also auf Papier und Bleistift basierenden, neuropsychologischen Testdiagnostik bei 200 Probanden mit kognitiven Beeinträchtigungen unterschiedlicher Ätiologie und Schweregrade zu vergleichen. Um eine direkte Vergleichbarkeit beider Untersuchungsmethoden zu gewährleisten, wurde ein randomisiertes, einfach verblindetes, zweiarmiges Studiendesign mit cross-over Design gewählt (siehe Abbildung 4).

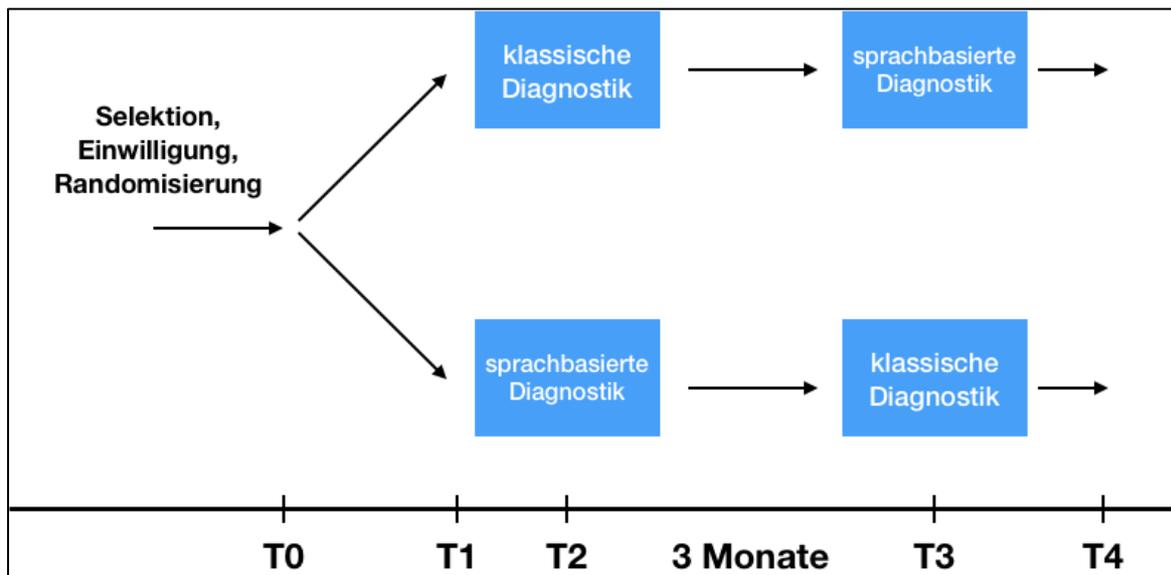


Abbildung 4. Schematischer Studienablauf der ELEMENT-Studie.

3.2 Studienablauf

Die Rekrutierung der Probanden erfolgte durch Aushänge bzw. Informationsflyer auf dem Klinikgelände, per Inserat auf der Klinik-Homepage sowie über die hausinterne Gedächtnisambulanz (s. Appendix). Es wurden nur Patienten mit subjektiven Gedächtnisbeeinträchtigungen sowie neurodegenerativen Grunderkrankungen in unterschiedlichem Krankheitsstadium und unterschiedlicher Ätiologie eingeschlossen.

Die Studienteilnehmer stellten sich zunächst zur Abklärung möglicher Ursachen neurokognitiver Beeinträchtigungen in der Gedächtnisambulanz des Universitätsklinikums des Saarlandes vor. Sie wurden sodann über die Untersuchung, Ein- sowie Ausschlusskriterien und Kontraindikationen der Studie aufgeklärt. Bei bestehenden Kontraindikationen wurden die Probanden nicht eingeschlossen. Eine erforderliche Wege- und Unfallversicherung für die Teilnehmer wurde im Vorfeld abgeschlossen und die Versicherungsbedingungen zusammen mit dem zweiten Exemplar der Einwilligungserklärung ausgehändigt.

Nach Einwilligung des Probanden erfolgte eine zufällige Zuteilung zu einer der beiden Studiengruppen durch eine Randomisierungsliste (Zeitpunkt T0) (angefertigt durch das Institut für Medizinische Biometrie, Epidemiologie und Medizinische Informatik; unter der Leitung von Herrn Prof. Dr. Wagenpfeil). Im Anschluss erfolgte die Durchführung des etwa 30-minütigen craniellen MRT (Zeitpunkt T1, s. Abbildung 4), welches in der Klinik für Neuroradiologie des UKS durchgeführt wurde.

Der erste Studienarm (n = 100) begann mit der neuropsychologischen Diagnostik mithilfe der Computer-basierten Spracherkennung „Delta-App“ (Zeitpunkt T2), nach drei Monaten gefolgt von der „klassischen“ Variante der Diagnostik „Paper-Pencil-Verfahren“ (T3). Bei den Probanden des zweiten Armes (n = 100) wurde mit der klassischen Variante der Neuropsychologie „Paper-Pencil-Verfahren“ begonnen (T2) und nach drei Monaten die computer-basierte Spracherkennungsvariante „Delta App“ (T3) durchgeführt.

Die Testung erfolgte in den Räumlichkeiten der Klinik für Psychiatrie und Psychotherapie (Universitätsklinikum Saarland) und dauerte circa eine Stunde pro Testung. Beide Arme endeten jeweils mit einer Abschlussbesprechung und Befundmitteilung durch ärztliches Personal (T4). Der Abstand von drei Monaten zwischen den jeweiligen Testungen wurde bewusst gewählt, um einerseits mögliche Übungs- bzw. Erinnerungseffekte bei den neuropsychologischen Tests zu minimieren, aber andererseits eine Verschlechterung der Untersuchungsergebnisse durch ein Voranschreiten der Grunderkrankung zu vermeiden.

3.3 Einschluss- und Ausschlusskriterien

In die Studie wurden 200 Probanden im Alter von 45-90 Jahren in einem nicht vorab definierten Geschlechterverhältnis eingeschlossen. Eingeschlossen wurden Patienten mit kognitiven Defiziten bei prodromaler Alzheimer-Krankheit, leichter kognitiver Beeinträchtigung vom

amnestischen Typ, leicht-, mittelgradige Demenz bei Alzheimer Krankheit, kognitive Beeinträchtigung und leicht-, mittelgradige Demenz bei Frontotemporaler Lobärdegeneration (FTLD) inkl. der Sprachvarianten Primär Progressive Aphasie (semantische Demenz, nichtflüssige/agrammatische und logopenische Variante).

Einschlusskriterien:

- ❖ Alter: 45-90 Jahre
- ❖ Einwilligungsfähigkeit zur Teilnahme an der Studie und Vorliegen eines Informed consent
- ❖ Mini Mental Status Examination (MMSE) \geq 18/30 Punkten
- ❖ keine wesentliche *körperliche* Grunderkrankung (fortgeschrittene Tumorerkrankung, instabile Herzerkrankung)
- ❖ keine wesentliche *neurologische* Erkrankung (Z. n. schwerem Schädel-Hirn-Trauma, Schlaganfall, etc.) oder relevante *psychiatrische* Begleiterkrankung (schwere depressive Episode, Schizophrenie)
- ❖ flüssige/s Sprache / Sprechen (keine komplette Aphasie)
- ❖ ausreichender Visus und Hörvermögen

Ausschlusskriterien:

- ❖ Fehlen eines Informed Consent bzw. Einwilligungsunfähigkeit
- ❖ MMSE $<$ 18/30 Punkten
- ❖ Vorliegen einer (weiteren) relevanten psychiatrischen oder einer neurologischen Erkrankung mit zerebraler Beteiligung bzw. Vorliegen einer relevanten internistischen Erkrankung
- ❖ akute Eigen- oder Fremdgefährdung
- ❖ mangelnde Sprachkenntnisse oder ausgeprägte formale Denkstörungen

Weitere Ausschlusskriterien waren etwaige Kontraindikationen gegen die Untersuchung mittels MRT infolge von nicht MRT-tauglichen Implantaten und/oder nicht beherrschbarer Angst- bzw. Panikzustände bei Klaustrophobie.

3.4 Datenerhebung/Materialien

3.4.1 Erhobene Patientendaten

Zur Studiauswertung wurden Alter (Geburtsdatum), Geschlecht, Händigkeit, Schul-, Berufsausbildung sowie psychiatrische und somatische Vorerkrankungen der Probanden erhoben. Bei der Schul- und Berufsausbildung zählte die Anzahl an Jahren, die üblicherweise für diesen Schultyp bzw. diese Ausbildung oder dieses Studium benötigt wird. Berufliche Weiterbildungen und Zusatzqualifikationen wurden bei den Ausbildungsjahren mitgezählt.

Hinsichtlich der psychiatrischen und somatischen Vorerkrankungen wurden alle Diagnosen aufgenommen, die im Rahmen der ambulanten bzw. stationären Behandlung der Patienten erhoben wurden. Als zusätzliche Informationen wurden mögliche Risikofaktoren dementieller Erkrankungen, insbesondere arterielle Hypertonie, Diabetes mellitus, Hyperthyreose und Hyperlipidämie erfasst.

3.4.2 Neuropsychologische Testung und Fragebögen

Es erfolgte eine neuropsychologische Testung mittels des „Consortium to Establish a Registry for Alzheimer’s Disease Neuropsychological Battery (CERAD-NP)“ sowie die Erhebung der derzeitigen depressiven Symptomatik mittels des „Becks Depressions Inventar (BDI-II)“.

CERAD

Das „Consortium to Establish a Registry for Alzheimer’s Disease (CERAD)“ ist im Jahr 1986 durch das „National Institute on Ageing (NIA)“ in den Vereinigten Staaten von Amerika gegründet worden. Der Zusammenschluss zielte auf die Entwicklung standardisierter Instrumente zur Erfassung und Beurteilung von klinischen und neuropsychologischen Anzeichen einer Demenz vom Alzheimer Typ ab, um die Zusammenführung von Informationen verschiedener Standorte zu ermöglichen (137). Die CERAD Testbatterie umfasst Tests, die die Bereiche, in denen Alzheimer Patienten Defizite aufweisen, wie z. B. das Gedächtnis, die Orientierung, die Sprache und die Praxis (138).

Die autorisierte deutschsprachige Version der CERAD-Testbatterie (CERAD-NP) ist an der Memory Clinic des Universitätsspitals Basel in der Schweiz entwickelt worden. Die ursprünglichen acht Untertests wurde durch drei zusätzliche Testverfahren erweitert, um auch die frontalen, subkortikalen und exekutiven Funktionen zu untersuchen. Die frontalen exekutiven Funktionen werden anhand der phonematischen Wortflüssigkeit (S-Wörter) und die subkortikalen Funktionen mit dem Trail-Making-Test A sowie mit dem Trail-Making-Test B gemessen (139).

Insgesamt besteht die Testbatterie aus 12 Untertests, die sich wie folgt zusammensetzen:

1. Verbale Wortflüssigkeit (Kategorie Tiere): Die verbale Wortflüssigkeit definiert die Fähigkeit, Elemente einer spezifischen Kategorie innerhalb eines festgelegten zeitlichen Rahmens abzurufen (140). Verwendet wird die kategoriale Wortflüssigkeit der „Tiere“, bei der der Proband innerhalb einer Minute so viele Tiere wie möglich aufzählen muss. Mit dieser Aufgabe werden die Geschwindigkeit, die verbale Produktionsfähigkeit, das semantische Gedächtnis, die Sprache, die exekutiven Funktionen und die kognitive Flexibilität getestet (141).
2. Boston Naming Test (BNT, modifizierte Version): In diesem visuellen konfrontativen Benennungstest wird die Fähigkeit, Wörter abzurufen, beurteilt. Dem Probanden werden

schwarz-weiße Strichzeichnungen von Objekten vorlegt, die von ihm identifiziert und benannt werden sollen (142). Die Bilder sind vom einfachsten zum schwersten hin geordnet. Die Originalversion enthält 60 Strichzeichnungen, die Kurzversion 15 Items (143). In der vorliegenden Studie wurde die Kurzversion verwendet.

3. Mini Mental State Examination (MMSE): Dieser Test stellt ein weitverbreitetes Screening-Instrument dar, mit dem das kognitive Leistungsniveau und Einschränkungen (älterer) Menschen erfasst werden können. Die Items umfassen Tests zu Orientierung, Gedächtnis, Rechenfähigkeit und Aufmerksamkeit, Sprache und Sprachverständnis sowie Visuokonstruktion. Wenn alle Aufgaben richtig beantwortet werden, beträgt die Punktzahl 30 (144).
4. Lernfähigkeit und Wortlisten Gedächtnis: Die Probanden lesen eine Liste von zehn gedruckten Wörtern laut vor und sollen diese nach jedem Durchgang frei abrufen (unmittelbares Erinnern). Es erfolgen drei Durchgänge, in denen jeweils die Wörter in anderer Reihenfolge präsentiert werden. Mit diesem Aufgabentyp wird die Fähigkeit der Testperson erfasst, nicht assoziiertes verbales Material zu erlernen. Die maximale Punktzahl für richtig erinnerte Wörter beträgt 30 Punkte (145).
5. Konstruktive Praxis: Um die visuokonstruktive Fähigkeit zu überprüfen, werden die Probanden gebeten, vier geometrische Figuren mit steigender Komplexität (Kreis, Raute/Rhombus, überschneidende Vierecke, Würfel) nachzuzeichnen (146).
6. Wortlistenabruf: Nach circa fünf Minuten, in denen eine nonverbale Aufgabe (konstruktive Praxis) durchgeführt wird, muss der Proband die zuvor gelernten zehn Wörter erneut frei abrufen (verzögerter Abruf). Die maximale Punktzahl beträgt zehn Punkte. Hier wird das verbal episodische Gedächtnis geprüft, was bedeutet, dass der Proband gelerntes verbales Material über einen Zeitraum von einigen Minuten behält (145).
7. Wortlisten Wiedererkennen: Direkt nach dem verzögerten Abruf wird eine Wiedererkennungsaufgabe durchgeführt und dem Probanden werden 20 Wörter vorgelegt. Die Aufgabe ist es, die zehn Wörter aus der ursprünglichen Wortliste und die zehn Distraktoren voneinander zu unterscheiden (145). Diese Aufgabe ermöglicht es zu testen, ob die Probanden von erleichterten Abrufbedingungen profitieren, um herauszufinden, ob vorrangig ein Defizit im Abruf oder im Abspeichern vorliegt (147).
8. Konstruktive Praxis Abruf: Bei dieser Aufgabe soll der Proband die vier Strichzeichnungen der Aufgabe Nummer fünf (Kreis, Raute, überlappende Rechtecke und Würfel) aus dem Gedächtnis abrufen. Sie misst das visuell-räumliche und konstruktive Gedächtnis (148).
9. Trail Making Test A (TMT A): In diesem Test werden die subkortikalen Funktionen mithilfe der psychomotorischen Geschwindigkeit und dem visuellen Scannen bewertet. Die Probanden müssen so schnell, wie möglich, die Zahlen 1 bis 25 in aufsteigender Reihenfolge mit einer Linie verbinden. Die Zahlen sind über eine DIN-A4 Seite verteilt. Die

Leistung wird anhand der Zeit bemessen, die der Proband zum Verbinden benötigt; maximal sind 180 Sekunden erlaubt. Fehler werden in der Bewertung nicht berücksichtigt, allerdings müssen alle Fehler während der Bearbeitung korrigiert werden, sodass sich die Bearbeitungszeit erhöht (139).

10. Trail Making Test B (TMT B): Die exekutiven Funktionen werden mit dem TMT-B bewertet, einem Maß für die kognitive Flexibilität und die Fähigkeit, zwischen zwei Modalitäten zu wechseln. Bei diesem Test werden sowohl Zahlen (1-13) als auch Buchstaben (A-L) durch eine Linie verbinden. Auch hier sind die beiden Modalitäten über ein DIN-A4-Blatt verteilt. Die Probanden sollen die Zahlen und Buchstaben in aufsteigender Reihenfolge verbinden und zusätzlich zwischen den beiden Modalitäten wechseln. Auch hier wird die Leistung durch die Gesamtbearbeitungszeit ermittelt, maximal 300 Sekunden sind erlaubt (139).
11. Phonematische Wortflüssigkeit (Buchstabe S): Die frontalen exekutiven Funktionen werden mithilfe der phonematischen Wortflüssigkeit gemessen. Der Proband soll innerhalb von 60 Sekunden so viele Wörter wie möglich generieren, die mit dem Buchstaben „S“ beginnen. Es werden folgende Einschränkungen auferlegt: alle Wörter sind erlaubt mit Ausnahme von Eigennamen, Zahlen und Wörtern mit dem gleichen Wortstamm. Die Leistung wird anhand der Gesamtzahl der genannten Wörter gemessen (139).

Die Durchführung dieser Testbatterie dauert circa 30-45 Minuten. Die deutschsprachige Version der CERAD Testbatterie ist an einer Stichprobe von 614 altersentsprechenden gesunden Probanden normiert worden (149). Im Rahmen einer Studie zeigte sich, dass die neuropsychologischen Maße durch das Alter, die Ausbildung sowie das Geschlecht beeinflusst werden (148). Daher werden die ermittelten Rohwerte in Standardskalen transformiert, wie z. B. einer Standardnormalverteilung. Diese ist eine besondere Form der Normalverteilung, die durch den Mittelwert 0 und eine Standardabweichung von 1 gekennzeichnet ist. Die Rohwerte werden dazu nach $z = \frac{x - Mx}{Sx}$ in s. g. „**z-Werte**“ transformiert, wobei „x“ den Rohwert, „Mx“ den Mittelwert einer Normalpopulation und „sx“ die Standardabweichung einer Normalverteilung beschreibt. Der z-Wert drückt somit aus, wie weit die Leistung einer Person in der Aufgabe von der durchschnittlichen Leistung der Normalpopulation differiert (150). Bei der Interpretation des z-transformierten Wertes bestehen verschiedene Möglichkeiten, weshalb die Kriterien, was einer durchschnittlichen, unter- oder überdurchschnittlichen Leistung entspricht, meist leicht differieren (151-153). In Übereinstimmung mit der Memory Clinic in Basel wurden die z-Werte in der vorliegenden Arbeit in Anlehnung an das Lehrbuch der Neuropsychologie bewertet (s. Tabelle 2) (154).

Tabelle 2. Interpretation der z-standardisierten Werte.

Standardisierter z-Wert	Beurteilung
> -1,28	unbeeinträchtigt
-1,28 bis -1,64	leicht beeinträchtigt
-1,64 bis -2,33	mittelschwer beeinträchtigt
< -2,33	schwer beeinträchtigt

Becks Depressions Inventar (BDI-II)

Dieser Selbstbeurteilungsfragebogen umfasst 21 Items, die in vier Bereiche untergliedert sind, den emotionalen, den kognitiven, den motivationalen und den physiologischen Bereich und zur Beurteilung des Schweregrades einer depressiven Symptomatik genutzt werden können. Jedes der Items wird von dem Probanden selbst auf einer 4-Punkte-Likert-Skala von 0-3 bewertet (155). Der Fragebogen kann bei Jugendlichen ab 13 Jahren und Erwachsenen angewendet werden. Die durchschnittliche Bearbeitungszeit beträgt fünf bis zehn Minuten. Eine Auswertung erfolgt durch eine Addition der Werte aller 21 Items, sodass der Summenscore zwischen 0 und 63 Punkten betragen kann. Zur Interpretation des Summenwertes werden fünf Kategorien gebildet: keine Depression (0 bis 8 Punkte), minimale (9 bis 13 Punkte), leichte (14 bis 19 Punkte), mittelschwere (20 bis 28 Punkte) und schwere Depression (29 bis 63 Punkte) (156).

3.4.3 Δ elta (Delta-App)

Δ elta ist eine iOS-App zur Unterstützung bei der Diagnose von neurokognitiven sowie psychiatrischen Störungen und wurde vom DFKI entwickelt. Δ elta soll Untersuchern (z. B. Psychologen, Neuropsychologen, Neurologen, Psychiatern, etc.) die Durchführung und Auswertung standardisierter Testverfahren, die auf gesprochener Sprache oder graphischer Stifteingabe beruhen, erleichtern. Dabei hilft Δ elta durch die Aufnahme von Sprache mittels automatischer Auswertung und Visualisierung von Testergebnissen und Generierung von Information bezüglich der Sprache (Wortproduktion, Benennensleistung, Semantik), die hinsichtlich ihres diagnostischen Potentials evaluiert werden können. Dazu greift Δ elta auf anerkannte, validierte, neurokognitive sowie psychiatrische Tests zurück, ohne diese in ihrer Durchführung oder Form zu verändern. Die App funktioniert wie folgt, wobei sich Patient und Untersucher physisch am gleichen Ort befinden:

1. Der Untersucher entsperrt das iPad, prüft die Internetverbindung und meldet sich mit persönlichem Benutzernamen und Passwort bei der App an.
2. Während des autobiografischen Interviews erfasst der Untersucher die charakteristischen Kerndaten eines Patienten (Alter, Geschlecht, Ausbildung, etc.) und legt sie in einer neuen Patientenmaske an.

3. Alternativ kann der Untersucher auch aus seinen ihm zugewiesenen und bereits im System angelegten Patienten wählen.
4. Der Untersucher wählt sodann einen neurokognitiven Test aus der Liste der verfügbaren Tests in Δ elta aus
5. Die App enthält eine Testanweisung, die vom Untersucher oder der App selbst laut vorgelesen werden
6. der Untersucher startet den Test und die Testbearbeitung des Patienten wird von Δ elta aufgezeichnet (d. h. Audioaufnahmen in sprachbasierten Tests, Pen-Tracking in grafischen Tests).
7. Nach oder während einer Diagnosesitzung greift der Untersucher entweder alleine oder zusammen mit dem Patienten auf die Analyse des kognitiven Tests zu, um die kognitive Testleistung zu beurteilen.
8. Falls die automatische Auswertung des Tests nicht akkurat ist, editiert der Untersucher manuell die Testauswertung durch Begutachtung der Rohdaten.

Je nach Test zeigt Δ elta während des Tests visuelle und akustische Reize an. Nach Abschluss des Tests sendet Δ elta diese Aufzeichnungen über eine sichere Verbindung an einen Analyseserver. Dieser Server berechnet die Testleistung und klinisch relevante Marker inkl. der sprachbasierten Marker. Diese Ergebnisse werden sodann an Δ elta zurückgegeben und dem Untersucher angezeigt. Dieser wird gebeten, die automatische Analyse zu überprüfen und gegebenenfalls zu korrigieren.

Spracherkennung

Um Sprachaufnahmen zu transkribieren und anhand des Transkripts weitere klinisch relevante Merkmale zu berechnen, nutzte Δ elta eine Back-End Automatische Spracherkennung (ASR). Aus Zertifizierungsgründen wurde die Google Speech API eingesetzt, da dieser Anbieter sowohl nach der neuen (ab 25.05.2018) EU General Data Protection Regulation (GDPR) als auch einer Vielzahl internationaler und nationaler Richtlinien zum Schutz von medizinischen personenbezogenen Daten zertifiziert war (s. a. <https://cloud.google.com/security/compliance/>).

Δ elta Analysis Engine

Das Δ elta Analysemodul extrahiert aus den Audio-Rohdaten medizinisch relevante Daten, die den kognitiven Test nach traditionellen und wissenschaftlich publizierten sowie neuropsychologisch anerkannten Metriken auswerten. Diese Auswertung (z. B. Anzahl korrekt genannter Tiere in der Semantischen Wortflüssigkeit mit Tieren) wird dann zusammen mit der Audiodatei und dem Transkript der Sprachdatei an die Δ elta iOS App zur Interpretation durch den Untersucher zurückgeschickt.

Datenbank

Die Δ elta Datenbank ist eine passwortgeschützte Datenbank, die auf einem Google Server in Deutschland abgelegt ist. Google stellt die beste physische und umweltbezogene Sicherheit bereit, einschließlich der Speicherung in einer Einrichtung mit einem streng eingeschränkten physischen Zugang. Die Server-Infrastruktur bietet eine ausreichende Netzwerksicherheit und sichere Server-Zugangspunkte.

Personenbezogene Daten

Bezüglich der Audio-Aufnahmen verarbeitet Δ elta sowohl anonymisierte Metadaten der Patienten (Medizinische Daten des Patienten, MDAT: Datum, Probanden-ID, Untersucher, Alter, Bildungsniveau, Diagnose, Ergebnisse der durchgeführten psychometrischen Testverfahren, wie z. B. CERAD-NP, MMSE, BDI-II). An personenbezogenen Daten wurden lediglich die erhobenen Audiodaten, sowie Testergebnisse aufgezeichnet. Personenbezogene Daten (Identifizierende Daten des Patienten, IDAT): Name, Anschrift, etc.) wurden nicht von Δ elta erfasst.

Aufgrund des Inhalts der aufgezeichneten Audiodateien musste eine Identifikation des jeweiligen Probanden bzw. Patienten ausgeschlossen werden, weswegen die Daten vor Auswertung pseudonymisiert wurden. Der Verwendung der Daten konnte jederzeit widersprochen werden, ohne dass den Probanden daraus ein Nachteil entstand.

Datenschutz

Jeder Teilnehmer wurde mit einer einzigartigen, zufällig generierten, nicht personenbezogenen Nummer ("Teilnehmer-ID") verknüpft und somit pseudonymisiert. Die Liste mit der Namen-ID Zuordnung (Kodierliste) war nur dem klinischen Principal Investigator zugänglich und verblieb immer in der Klinik für Psychiatrie und Psychotherapie der Universitätskliniken des Saarlandes. Die verwendeten iPads wurden ausschließlich zu Studienzwecken genutzt und nach jeder Testung in einem abschließbaren Schrank im Studienbüro im Universitätsklinikum des Saarlandes, Klinik für Psychiatrie und Psychotherapie, verwahrt. Nur das Studienteam hatte einen Schlüssel hierfür.

3.5 Entwicklung sprachbasierter Biomarker

Der sprachbasierte Biomarker wurde in Kooperation mit dem DFKI und ki:elements entwickelt. Ki:elements ist ein Ablegerunternehmen des DFKI, welches im Jahr 2017 gegründet wurde und sich mit der Entwicklung von digitalen Biomarkern, die die Symptomatik verschiedener Pathologien z. B. von Alzheimer Demenzen oder anderen psychischen Störungsbildern messen, beschäftigt. Der ki:elements Sprach Biomarker für Kognition ist ein zusammengesetzter Score, der aus drei verschiedenen neurokognitiven Subscores (SB-C

learning & memory (Lernen & Gedächtnis), SB-C executive function (Exekutive Funktionen), SB-C processing speed (Verarbeitungsgeschwindigkeit)) gebildet und für das Screening von dementiellen Störungsbildern genutzt wird (siehe Abbildung 5). Diese Scores setzen sich wiederum aus 50 automatisch extrahierten Sprachmerkmalen zusammen. Als Input dieser Scores fungieren verschiedene standardisierte neuropsychologische Testverfahren (z. B. der Trail Making Test, der Stroop Interferenz Test und die Copying Figure Recall) (siehe Abbildung 5). Aus den drei o. g. Subscores wurde der Gesamt-Score für Kognition abgeleitet.

Ki:elements entwickelte in einem Machine Learning Ansatz einen Classifier, der sowohl die drei Subscores als auch den globalen Score als Input verwendete. Mithilfe verschiedener Techniken (z. B. Support-Vector-Machine-Modell oder Decision Tree) wurde der Classifier ermittelt, der am besten zwischen den verschiedenen klinischen Subgruppen (Kontrollgruppe, MCI und AD) unterscheiden konnte.

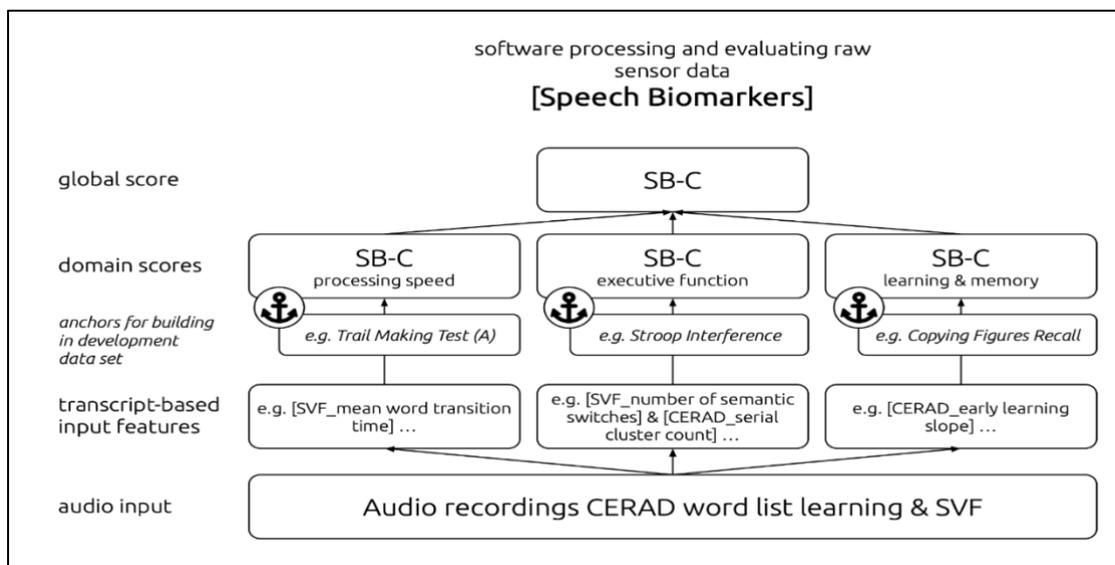


Abbildung 5. Schematische Repräsentation der Entwicklung des ki:e SB-C sowie der Subscores (SB-C: speech biomarker for cognition, SVF: semantic verbal fluency, CERAD: Consortium to Establish a Registry for Alzheimer’s Disease).

Im Rahmen der ELEMENT Studie dienten insbesondere der Untertest „Wortliste lernen“ des CERADs sowie die semantische Wortflüssigkeit (Semantic Verbal Fluency, SVF), die bei allen Studienpatienten erfasst wurden, als Input. Die Aufzeichnungen der Tests wurden automatisch mit der proprietären Sprachanalyse-Pipeline von ki:elements verarbeitet, welche eine automatische Spracherkennung beinhaltet. Diese Spracherkennung wurde zur Transkription von Sprache und weiterer Merkmale, wie Pausen oder Objektfolgen genutzt. Verarbeitet wurden die Aufgabenstimuli durch eine Analyse von subjektiven, seriellen oder semantischen Gruppierungen sowie die Analyse semantischer und zeitlicher Aspekte des verbalen Outputs der Probanden (siehe Abbildung 6 und Abbildung 7).

die Diagnose des Probanden, bewertet. Im letzten Schritt, der **klinischen Validierung**, wird getestet, ob der digitale Biomarker aussagekräftige klinische Informationen liefert (158).

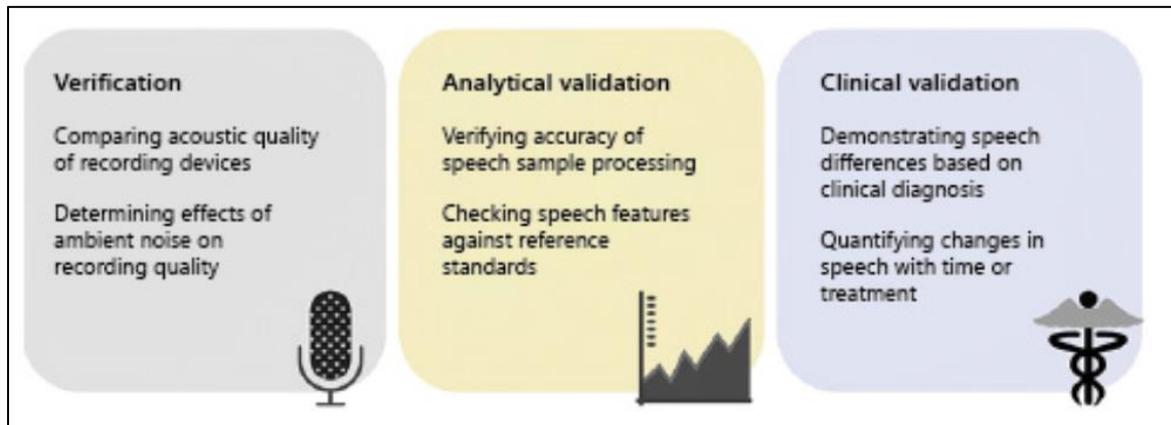


Abbildung 8. V3 Rahmen zur Evaluierung von Sprach Biomarkern (in Anlehnung an Goldsack (2020) (158)).

3.6 Statistische Auswertung

Die statistische Auswertung erfolgte mit dem Programm IBM SPSS Statistics System Version 28 (International Business Corporation, Armonk, New York, USA). Kategoriale Variablen wurden in absoluter und relativer Häufigkeit angegeben, kontinuierliche Variablen als Median und Range; es wurde angenommen, dass die Daten nicht normalverteilt sind.

Gruppenvergleiche wurden mithilfe des Chi-Quadrat-Tests, exakten Tests nach Fisher, Mann-Whitney-U- und Kruskal-Wallis-Test durchgeführt. Dazu erfolgte zunächst ein Vergleich von Studiengruppe vs. Dropout-Gruppe und Studiengruppe vs. Kontrollgruppe. Um zu überprüfen, ob Zeitpunkt oder Modalität der Testung einen Einfluss auf die Testergebnisse hatten, wurden die Ergebnisse der Testung mittels Paper Pencil und App in der Studiengruppe verglichen sowie das Testergebnis von der ersten vs. zweiten Testung. Außerdem erfolgte innerhalb der Studiengruppe ein Vergleich der Testergebnisse in Abhängigkeit von der jeweiligen Demenzform.

Die Entwicklung des Sprachbiomarkers (ki: elements SB-C Kognition Score) erfolgte analog zum statistischen Ansatz des Rahmenwerks der Gesellschaft für digitale Medizin (DiMe, s.o.) (159). Zur analytischen Validierung wurde der ki: e SB-C mit einem Standardmaß für Kognition, in diesem Fall dem MMSE, in der Population für Probanden mit kognitiver Beeinträchtigung verglichen. Dafür wurden Spearman-Rangkorrelationen zwischen dem digitalen Biomarker-Score und dem Mini Mental Examination Score (MMSE, (160)) berechnet. Im Rahmen der klinischen Validierung wurde ein Gruppenvergleich der Biomarker-Scores zwischen den verschiedenen Diagnosegruppen durchgeführt (nichtparametrischer Kruskal-

Wallis-Test zwischen der gesunden Kontrollgruppe (Healthy control), MCI – und der AD-Gruppe).

Mithilfe einer ROC-Analyse (Receiver Operating Characteristic) wurden beide Testmodalitäten hinsichtlich ihrer Fähigkeit, gesund vs. an MCI erkrankt, gesund vs. an Alzheimer erkrankt, gesund vs. an MCI oder Alzheimer erkrankt und an MCI vs. an Alzheimer erkrankt, analysiert. Die jeweiligen Cut-off Werte wurden mithilfe des Youden-Index bestimmt (Sensitivität + Spezifität – 1). Der optimale Cut-Off- Wert wurde anhand des Wertes definiert, bei dem der Youden-Index maximal ist. (161).

Alle statistischen Tests waren zweiseitig, $p < 0,05$ wurde als Signifikanzniveau definiert.

4. Ergebnisse

4.1 Analyse des Gesamtkollektivs

4.1.1 Patientencharakteristika

Das Patientenkollektiv umfasste insgesamt 200 Probanden, davon waren 100 (50 %) männlich (siehe Tabelle 3). Das mediane Alter betrug 71 Jahre (Range 49-94) bei einer Ausbildungsdauer von 12 Jahren (7-20). 193 (96,3 %) Probanden waren Rechtshänder. In der Studiengruppe lagen die **dementiellen Erkrankungen in unterschiedlichen Stadien** vor: fast die Hälfte der Probanden (48,5 %) hatte ein mild cognitive impairment (MCI), 65 (32,5 %) eine Demenz vom Alzheimer Typ. Bei 4 (2 %) Probanden bestand eine vaskuläre, bei 6 (3 %) eine frontotemporale, bei 9 (4,5 %) eine sonstige Demenz und bei 3 (1,5 %) eine Demenz bei andernorts klassifizierten Erkrankungen (siehe Abbildung 9). Bei 16 (8 %) Studienpatienten lag ausschließlich eine subjektive Beeinträchtigung vor.

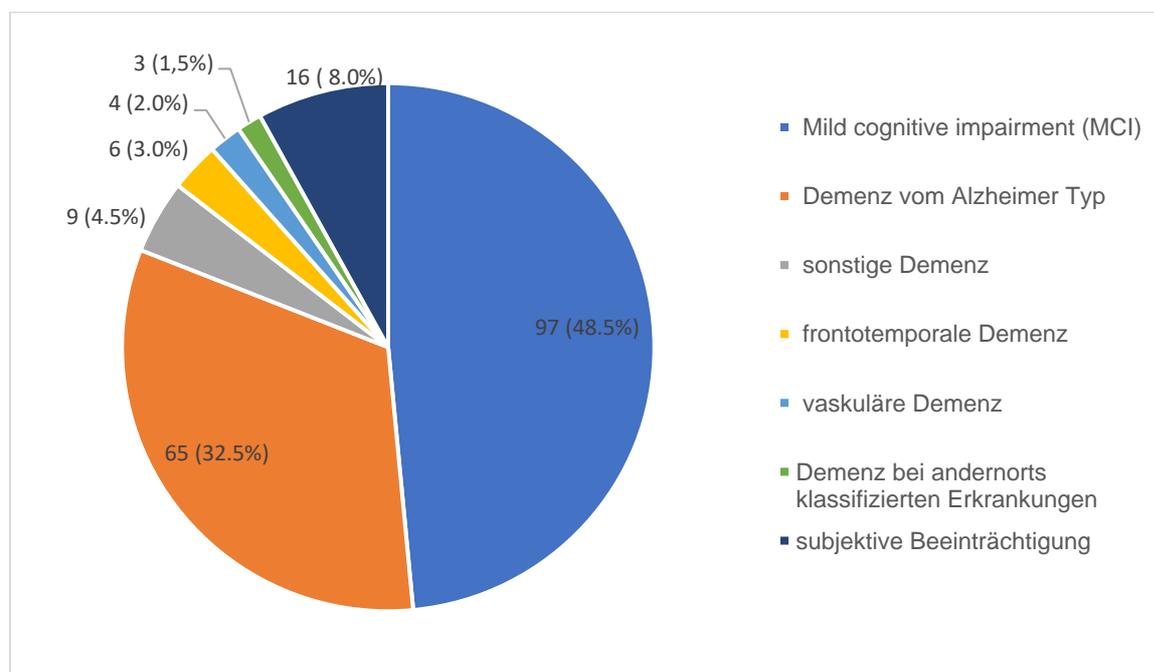


Abbildung 9. Verteilung der Subtypen dementieller Erkrankungen in der Studiengruppe.

Von den Probanden mit Alzheimer Demenz hatten mehr als die Hälfte eine Demenz mit spätem Beginn (56,9 %, s. Tabelle 3). Als mögliche **Risikofaktoren** einer dementiellen Erkrankung bestanden bei 85 (42,5 %) Probanden eine arterielle Hypertonie, bei 28 (14 %) eine Hyperthyreose, 25 (12,5 %) ein Diabetes Mellitus und bei 22 (11 %) eine Hyperlipidämie. Als Begleiterkrankung hatten 83 (41,5 %) Studienpatienten ein depressives Störungsbild in unterschiedlicher Ausprägung.

Tabelle 3. Charakteristika des Gesamtkollektivs, angegeben als absolute (relative) Häufigkeit bzw. Median (Spannweite).

Charakteristika	
Alter [J]	71 (49-94)
Geschlecht: männlich	100 (50%)
Bildungsjahre	12 (7-20)
Händigkeit rechts	193 (96,3%)
Risikofaktoren	
Arterielle Hypertonie	85 (42,5%)
Hyperthyreose	28 (14,0%)
Diabetes Mellitus	25 (12,5%)
Hyperlipidämie	22 (11,0%)
Alzheimer Demenz Subtyp	
Früher Beginn	13 (20%)
Später Beginn	37 (56,9%)
Gemischter Typ	15 (23,1%)
Depressionsdiagnose	83 (41,5%)
Ausprägung leicht	12 (14,5%)
Ausprägung mittel	45 (54,2%)
Ausprägung schwer	26 (31,3 %)
Zeitraum zwischen zwei Testungen	98 Tage (91-198)

4.1.2 Vergleich Studiengruppe vs. Dropout-Gruppe

148 (84 %) Studienteilnehmer wurden zwei Mal getestet und in die finale Studiauswertung eingeschlossen, gegenüber 52 (26 %) Probanden mit nur einfacher Testung als Dropout-Gruppe. Im direkten Vergleich beider Gruppen lag ein signifikanter Altersunterschied vor, denn Patienten aus der Dropout-Gruppe waren im Median 6 Jahre älter (69 vs. 75 Jahre, $p = 0,011$, siehe Tabelle 4). Demgegenüber hatten Patienten der Studiengruppe eine um zwei Jahre längere Ausbildungsdauer (13 vs. 11 Bildungsjahre, $p = 0,005$). Die Verteilung an Demenzdiagnosen war signifikant verschieden ($p = 0,008$), denn der Anteil an Probanden mit vaskulären Demenzen war in der Dropout-Gruppe signifikant höher (7,7 % vs. 0, $p = 0,004$). Auch wenn der Anteil an mittelgradigen Depressionen in der Studiengruppe mit 26,4 % mehr als doppelt so hoch wie in der Dropout-Gruppe (11,5 %) war, bestand hier kein signifikanter Unterschied. Alle weiteren Charakteristika waren vergleichbar und nicht signifikant verschieden (siehe Tabelle 4).

Tabelle 4. Vergleich Studien- vs. Dropout-Gruppe.

	Studiengruppe (n=148)	Dropout-Gruppe (n=52)	p-Wert
Alter	69 (50-91)	75 (49-94)	0,011
Männer	79 (53,4%)	21 (40,4%)	0,146
Bildungsjahre	13 (7-20)	11 (8-20)	0,005
Rechtshänder	143 (96,6%)	50 (96,2%)	1,000
Demenzdiagnose			0,008
MCI	71 (48%)	26 (50%)	0,872
Alzheimer Demenz	46 (31,1%)	19 (36,5%)	0,494
Vaskuläre Demenz	0 (0,0%)	4 (7,7%)	0,004
Frontotemporale Demenz	6 (4,1%)	0 (0%)	0,342
Sonstige Demenz	8 (5,4%)	1 (1,9%)	0,451
Demenz bei andernorts klassifizierten Erkr.	3 (2%)	0 (0%)	0,569
Subjektive Beeinträchtigung	14 (9,5%)	2 (3,8%)	0,248
Alzheimer Demenz			0,077
Früher Beginn	12 (8,1%)	1 (1,9%)	0,870
Später Beginn	26 (17,6%)	11 (21,2%)	1,000
Gemischter Typ	8 (5,4%)	7 (13,5%)	0,112
Risikofaktoren			
Arterielle Hypertonie	62 (41,9%)	23 (44,2%)	0,871
Hyperthyreose	17 (11,5%)	11 (21,2%)	0,104
Diabetes Mellitus	20 (13,5%)	5 (9,6%)	0,627
Hyperlipidämie	19 (12,8%)	3 (5,8%)	0,203
Depressionsdiagnose	65 (43,6%)	18 (34,6%)	0,257
Ausprägung leicht	9 (6,1%)	3 (5,8%)	0,717
Ausprägung mittel	39 (26,4%)	6 (11,5%)	0,062
Ausprägung schwer	17 (11,5 %)	9 (17,3%)	0,083

Beim Vergleich der Testergebnisse beider Gruppen bestand kein signifikanter Unterschied in der Selbstbeurteilung der depressiven Symptomatik mit dem BDI. Hinsichtlich der CERAD-Batterie Testung war der Boston Naming Test zwischen beiden Gruppen signifikant verschieden, und die Dropout-Gruppe schnitt mit einem z-Wert von -0,76 (Range -6,63; 2,20) schlechter als die Studiengruppe ab (-0,35 (-5,36; 1,74), $p = 0,019$, s. Tabelle 5), wobei beide Leistungen noch im Normbereich lagen (vergleiche Tabelle 2, S. 32). Im Demenz-Screening Test (MMST) hatten die Probanden der Dropout-Gruppe mit einem z-Wert von -2,00 (-7,61;

1,06), also einer mittelschweren Beeinträchtigung, ein signifikant schlechteres Ergebnis als die der Studiengruppe mit einem z-Wert von -1,59 (-5,96; 1,65; $p = 0,020$), was einer leichten Beeinträchtigung entspricht. Bei den Leistungen im Bereich des verbalen Gedächtnisses (Wortlisten Abruf) lagen die Leistungen der Studiengruppe im Normbereich (z-Wert: -1,13 (-4,35; 1,82)) und die Ergebnisse der Dropout-Gruppe (-1,78 (-4,28; 0,77)) waren mittelschwer beeinträchtigt ($p = 0,011$). Im Wiedererkennen der zuvor gelernten Wortliste bestand ebenso ein signifikanter Unterschied zwischen den Gruppen: Die Leistung der Studiengruppe lag im Normbereich (-0,97 (-4,68; 1,45)), wohingegen die Dropout-Gruppe leichtgradig beeinträchtigt war und damit ein signifikant schlechteres Ergebnis hatte (-1,44 (-4,28; 0,88), $p = 0,041$). Im Bereich des figuralen Abrufens waren die Leistungen beider Gruppen gleichermaßen mittelschwer beeinträchtigt, aber dennoch statistisch signifikant verschieden (Studiengruppe: -1,42 (-5,37; 2,24) vs. Dropout-Gruppe: -2,01 (-6,01; 1,27), $p = 0,046$). Auch hinsichtlich der Verarbeitungsgeschwindigkeit (Trail Making Test A) schnitt die Dropout-Gruppe mit einem z-Wert von -0,89 (-4,17; 2,02) im Vergleich zur Studiengruppe mit -0,36 (-3,80; 2,27) signifikant schlechter ab ($p = 0,003$), wobei beide Leistungen jedoch als unbeeinträchtigt zu interpretieren sind.

Dementsprechend schnitt die Dropout-Gruppe im Vergleich zur Studiengruppe in acht z-standardisierten Items signifikant schlechter ab. In drei Fällen fielen die Patienten der Dropout-Gruppe zusätzlich in eine schlechtere Gruppe nach z-Standardisierung (vgl. Tabelle 2, S. 32).

Tabelle 5. Vergleich Studiengruppe vs. Dropout-Gruppe hinsichtlich der Testergebnisse anhand der Z-Werte.

	Studiengruppe (n=148)	Dropout (n=52)	p-Wert
BDI	7 (0;19 Punkte)	7 (0;19 Punkte)	0,466
Z-Werte			
Semantische Wortflüssigkeit	-0,99 (-4,14; 2,19)	-0,90 (-3,72; 1,27)	0,787
Boston Naming	-0,35 (-5,36; 1,74)	-0,76 (-6,63; 2,20)	0,019
MMST	-1,59 (-5,96; 1,65)	-2,00 (-7,61; 1,06)	0,020
Wortliste Lernen Total	-1,68 (-6,83; 1,95)	-1,96 (-6,89; 0,98)	0,826
Wortliste Lernen #1	-1,40 (-4,39; 2,83)	-1,27 (-4,86; 1,35)	0,392
Wortliste Lernen #2	-1,40 (-5,39; 2,12)	-1,41 (-5,08; 0,34)	0,970
Wortliste Lernen #3	-1,32 (-7,20; 1,39)	-1,79 (-6,85; 2,37)	0,517
Wortliste Abruf	-1,13 (-4,35; 1,82)	-1,78 (-4,28; 0,77)	0,011
Wortliste Intrusionen	-0,67 (-3,99; 0,98)	-0,76 (-2,70; 0,95)	0,854
Wortliste Saving	-0,86 (-4,85; 4,70)	-1,64 (-4,13; 2,04)	0,034

Wortliste Wiedererkennen	-0,97 (-4,68; 1,45)	-1,44 (-4,28; 0,88)	0,041
Figuren Abzeichnen	0,43 (-4,26; 1,74)	-0,61 (-5,01; 1,73)	0,529
Figuren Abrufen	-1,42 (-5,37; 2,24)	-2,01 (-6,01; 1,27)	0,046
Figuren Saving	-1,15 (-3,82; 2,35)	-1,99 (-4,02; 2,49)	0,009
Trail Making A	-0,36(-3,80; 2,27)	-0,89 (-4,17; 2,02)	0,003
Trail Making B	-0,56 (-3,27; 2,34)	-1,02 (-3,62; 2,28)	0,602
Trail Making A/B	-0,29 (-3,04; 2,63)	0,15 (-3,75; 3,06)	0,015
Phonematische Wortflüssigkeit	-0,43 (-3,00; 2,35)	-0,68 (-2,29; 2,28)	0,266

4.1.3 Vergleich Studiengruppe vs. Kontrollgruppe

148 Studienteilnehmer wurden zwei Mal getestet und in die finale Studiauswertung eingeschlossen, gegenüber 27 Probanden in der Kontroll-Gruppe, die dem Studienprotokoll entsprechend nur einmalig getestet wurden. Im direkten Vergleich beider Gruppen lag ein signifikanter Altersunterschied vor, denn Patienten aus der Studien-Gruppe waren im Median 7 Jahre älter (69 vs. 62 Jahre, $p = 0,033$, siehe Tabelle 6). In der Studiengruppe waren signifikant mehr männliche Probanden (79 (53,4%) vs. 7 Männer (25,9%), $p = 0,011$). Die Verteilung der Depressionsdiagnosen war ebenso signifikant verschieden ($p < 0,001$), denn der Anteil an Probanden mit mittelschweren Depressionen war in der Studien-Gruppe höher (26,4 % vs. 22,2 %, $p = 0,022$). Der Anteil an schwergradigen Depressionen war in der Kontrollgruppe mehr als dreimal so hoch wie in der Studien-Gruppe (44,4 % vs. 11,5 %, $p = 0,016$). Auch wenn im Bereich der leichtgradigen Depressionen der Anteil an Probanden in der Kontrollgruppe fast doppelt so hoch war (11,1%) wie in der Studiengruppe (6,1%), bestand hier kein signifikanter Unterschied. Alle weiteren Charakteristika waren vergleichbar und nicht signifikant verschieden (siehe Tabelle 6).

Tabelle 6. Vergleich der Charakteristika von Studiengruppe vs. Kontrollgruppe.

	Studiengruppe (n=148)	Kontrollgruppe (n=27)	p-Wert
Alter	69 (50-91)	62 (51-86)	0,033
Männer	79 (53,4%)	7 (25,9%)	0,011
Bildungsjahre	13 (7-20)	13 (8-20)	0,357
Rechtshänder	143 (96,6%)	24 (88,9%)	0,108
Risikofaktoren			
Arterielle Hypertonie	62 (41,9%)	7 (25,9%)	0,138
Hyperthyreose	17 (11,5%)	2 (7,4%)	0,742
Diabetes Mellitus	20 (13,5%)	2 (7,4%)	0,535
Hyperlipidämie	19 (12,8%)	2 (7,4%)	0,537

Depressionsdiagnose	65 (43,6%)	21 (77,7%)	<0,001
Ausprägung leicht	9 (6,1%)	3 (11,1%)	1,00
Ausprägung mittel	39 (26,4%)	6 (22,2%)	0,022
Ausprägung schwer	17 (11,5 %)	12 (44,4 %)	0,016

Beim Vergleich der Testergebnisse beider Gruppen bestand ein signifikanter Unterschied in der Selbstbeurteilung der depressiven Symptomatik mit dem BDI ($p = 0,034$), wobei sich der mediane Rohwert nur um einen Punkt unterschied (Studiengruppe: 7 Punkte vs. Kontrollgruppe: 6 Punkte). Hinsichtlich der CERAD-Batterie Testung schnitt die Kontrollgruppe in *allen* Tests besser ab (siehe Tabelle 7), wobei in drei Fällen kein signifikanter Unterschied vorlag. Im ersten Durchgang des verbalen Lernverhaltens bestand kein signifikanter Unterschied, wobei die Studiengruppe (-1,53 (-3,87; 2,14)), welche eine leichtgradige Beeinträchtigung aufwies, schlechter als die Kontrollgruppe mit einem unbeeinträchtigten z-Wert von -1,21 (-2,84; 1,39) abschnitt ($p = 0,105$).

Im Bereich des Wortlisten Savings, welches einen Parameter für die Behaltensleistung zwischen dem 3. Durchgang des verbalen Lernverhaltens und dem Abruf des Materials (verbales Gedächtnis) darstellt, schnitt die Studiengruppe mit einem normwertigen z-Wert von -0,96 (-4,85; 4,70) schlechter ab als die Kontrollgruppe (-0,11 (-3,26; 1,50), $p = 0,057$), deren Leistung ebenfalls im unbeeinträchtigten Bereich lag; der Unterschied war aber nicht statistisch signifikant. Auch im Bereich der phonematischen Wortflüssigkeit schnitt die Kontrollgruppe mit einem z-Wert von 0,00 (-2,16; 2,21) schlechter ab als die Studiengruppe (0,00 (-2,16; 2,21), $p = 0,065$). Die Ergebnisse, beide im unbeeinträchtigten Bereich, waren aber nicht signifikant verschieden.

Dementsprechend schnitt die Kontrollgruppe im Vergleich zur Studiengruppe nach z-Standardisierung in allen 18 Items besser ab, wobei sich nur in 15 Items ein signifikanter Unterschied ergab. Eine Eingruppierung der Studiengruppe in eine schlechtere Gruppe erfolgte bei fünf Items anhand der z-Standardisierung.

Tabelle 7. Vergleich Studiengruppe (erste Testung) vs. Kontrollgruppe hinsichtlich der Testergebnisse, angegeben als z-standardisierte Werte.

	Studiengruppe (n=148)	Kontrollgruppe (n=27)	p-Wert
BDI (Rohwertpunkte)			
1. Testung	7 (0;19 Punkte)	6 (0;17 Punkte)	0,034
Z-Werte			
Semantische Wortflüssigkeit	-0,98 (-4,14;2,44)	-0,49 (-2,02; 2,24)	<0,001

Boston Naming	-0,40 (-5,36; 1,97)	0,56 (-2,66; 2,93)	<0,001
MMST	-1,61 (-5,96; 1,90)	-0,10 (-1,73; 1,10)	<0,001
Wortliste Lernen Total	-1,82 (-6,83; 1,99)	-0,93 (-1,99; 1,18)	<0,001
Wortliste Lernen #1	-1,53 (-3,87; 2,14)	-1,21 (-2,84; 1,39)	0,105
Wortliste Lernen #2	-1,51 (-5,93; 2,12)	-0,80 (-2,68; 1,48)	0,002
Wortliste Lernen #3	-1,29 (-7,20; 1,65)	-0,20 (-1,46; 1,50)	<0,001
Wortliste Abruf	-1,13 (-4,35; 1,52)	-0,29 (-2,25; 1,84)	<0,001
Wortliste Intrusionen	-0,80 (-3,99; 0,98)	0,41 (-1,27; 0,96)	0,005
Wortliste Saving	-0,96 (-4,85; 4,70)	-0,11 (-3,26; 1,50)	0,057
Wortliste Wiedererkennen	-0,88 (-4,68; 1,30)	0,23 (-1,61; 0,96)	0,005
Figuren Abzeichnen	0,37 /-3,47; 1,76)	0,59 (-11,60; 1,41)	0,016
Figuren Abrufen	-1,35 (-4,95; 2,26)	0,61 (-3,37; 2,11)	<0,001
Figuren Saving	-1,11 (-3,63; 2,35)	0,32 (-1,90; 1,45)	<0,001
Trail Making A	-0,27 (-3,75; 2,31)	0,03 (-1,53; 3,49)	0,021
Trail Making B	-0,49 (-3,27; 2,79)	0,65 (-1,77; 2,76)	<0,001
Trail Making A/B	-0,23 (-3,04; 2,63)	0,28 (-1,86; 1,72)	0,021
Phonematische Wortflüssigkeit	-0,44 (-3,83; 2,35)	0,00 (-2,16; 2,21)	0,065

4.2 Einflussfaktoren auf das Testergebnis innerhalb der Studiengruppe (Modalität & Zeitpunkt)

4.2.1 Vergleich Testung mittels App vs. Paper Pencil

Beim Vergleich der Testergebnisse beider Testmodalitäten schnitten die Probanden im Bereich des Boston Naming Tests bei Testung per App mit einem z-Wert von -0,40 (-5,36; 1,97) signifikant schlechter ab als bei Testung mittels Paper-Pencil (0,31 (-5,36; 1,78), $p < 0,001$), wobei beide Ergebnisse im unbeeinträchtigten Bereich anzusiedeln sind (siehe Tabelle 8). Im Wortlisten Lernen über alle drei Durchgänge hinweg erzielte die Testung mit App mit einem z-Wert von -1,82 (-6,83; 1,99), also einer mittelschweren Beeinträchtigung, ein signifikant schlechteres Ergebnis als mit Paper-Pencil Verfahren (-1,44 (-6,83; 1,93), $p = 0,021$), was einer leichten Beeinträchtigung entspricht. Auch im ersten Durchgang des verbalen Lernverhaltens wies die Testung per App (-1,53 (-3,87; 2,14)) mit einem Ergebnis, was im leichtbeeinträchtigten Bereich anzusiedeln ist, ein signifikant schlechteres Ergebnis auf, als die andere Testmodalität mit einem z-Wert von -1,16 (-4,29; 2,83, $p = 0,012$), welches dem Normbereich entspricht. Beim Wortlisten Abruf waren die Leistungen bei Testung mit beiden Modalitäten gleichermaßen unbeeinträchtigt, unterschieden sich jedoch signifikant (App: -1,13 (-4,35; 1,52) vs. Paper Pencil: -0,99 (-4,35; 1,82), $p = 0,022$). Auch im Bereich der Intrusionen während des verbalen Lern- und Abrufverhaltens schnitt die App mit einem z-Wert -0,80 (-

3,99; 0,98) signifikant schlechter ab als das Paper-Pencil Verfahren (-0,68 (-3,51; 0,98), $p = 0,020$), wobei auch hier beide Ergebnisse als normwertig zu beurteilen sind.

Dementsprechend schnitten die Probanden bei Testung mittels App im Vergleich zum Paper-Pencil-Verfahren in fünf z-standardisierten Items signifikant schlechter ab, wobei nach der z-Standardisierung in zwei Fällen eine Eingruppierung in eine schlechtere Gruppe erfolgte.

Tabelle 8. Vergleich Testung mittels App vs. Paper-Pencil Verfahren in Studiengruppe.

	App	Paper Pencil	p-Wert
BDI	7 (0;19)	7 (0;19)	0,910
Z-Werte			
Semantische Wortfl.	-0,98 (-4,14; 2,44)	-0,92 (-3,90; 1,85)	0,746
Boston Naming	-0,40 (-5,36; 1,97)	0,31 (-5,36; 1,78)	<0,001
MMST	-1,61 (-5,96; 1,90)	-1,51 (-7,26; 1,90)	0,706
Wortliste Lernen Total	-1,82 (-6,83; 1,99)	-1,44 (-6,83; 1,93)	0,021
Wortliste Lernen #1	-1,53 (-3,87; 2,14)	-1,16 (-4,29; 2,83)	0,012
Wortliste Lernen #2	-1,51 (-5,39; 2,12)	-1,23 (-5,93; 1,41)	0,082
Wortliste Lernen #3	-1,29 (-7,20; 1,65)	-1,29 (-7,20; 10,00)	0,370
Wortliste Abruf	-1,13 (-4,35; 1,52)	-0,99 (-4,35; 1,82)	0,022
Wortliste Intrusionen	-0,80 (-3,99; 0,98)	-0,68 (-3,51; 0,98)	0,020
Wortliste Saving	-0,96 (-4,85; 4,70)	-0,71 (-4,85; 4,92)	0,389
Wortliste Wiedererkennen	-0,88 (-4,68; 1,30)	-0,89 (-4,68; 1,45)	0,515
Figuren Abzeichnen	0,37 (-3,47; 1,76)	0,46 (-4,26; 1,74)	0,480
Figuren Abrufen	-1,35 (-4,95; 2,26)	-1,29 (-5,37; 2,26)	0,253
Figuren Saving	-1,11 (-3,36; 2,35)	-1,04 (-3,83; 2,32)	0,321
Trail Making A	-0,27 (-3,75; 2,31)	-0,45 (-3,80 – 1,73)	0,121
Trail Making B	-0,49 (-3,27; 2,79)	-0,57 (-3,27; 2,72)	0,140
Trail Making A/B	-0,23 (-3,04; 2,63)	-0,26 (-2,62; 2,34)	0,295
Phonematische Wortfl.	-0,44 (-3,83; 2,35)	-0,42 (-3,00; 3,16)	0,389

4.2.2 Vergleich Ergebnis Erst- vs. Zweittestung

Beim Vergleich der Erst- und Zweit-Testung innerhalb der Studiengruppe bestand ein signifikanter Unterschied in der Selbstbeurteilung der depressiven Symptomatik mit dem BDI, wobei sich der mediane Rohwert als auch die Range zwischen den Testungen nicht unterschied (siehe Tabelle 9). Hinsichtlich der CERAD-Testbatterie war der Boston Naming Test zwischen beiden Testungen signifikant verschieden, die Leistung der Probanden in der

ersten Testung war mit einem z-Wert von -0,35 (-5,36; 1,74) signifikant schlechter als bei der Zweittestung (0,15 (-5,36; 1,97, $p = 0,003$). Beide Ergebnisse sind als normwertig zu bewerten. Im Wortlisten-Lernen, über alle drei Durchgänge hinweg, lag in der ersten Testung mit einem z-Wert von -1,40 (-4,39; 2,83) ein signifikant schlechteres Ergebnis als in der zweiten Testung mit einem z-Wert von -1,52 (-6,83; 1,99), $p = 0,022$) vor, wobei beide Ergebnisse im Bereich der leichtgradigen Beeinträchtigung anzusiedeln sind. Auch im ersten Lerndurchgang lag im Rahmen der Ersttestung mit einem z-Wert von -1,40 (-4,39; 2,83) eine signifikant schlechtere Leistung als in der zweiten Testung (-1,22 (-3,87; 2,26), $p = 0,030$) vor (beide leichtgradige Beeinträchtigung). Im dritten Lerndurchgang bestand bei beiden Leistungen ebenfalls ein signifikanter Unterschied zwischen den Testungen, die der Ersttestung lag im leichtgradig beeinträchtigten Bereich (-1,32 (-7,20; 1,39)), die Zweittestung jedoch im Normbereich (1,26 (-7,20; 10,00), $p = 0,019$). Auch hinsichtlich des Wortlisten Abrufs war das Ergebnis zum ersten Testzeitpunkt mit einem z-Wert von -1,13 (-4,35; 1,82) im Vergleich zur zweiten Testung (-0,99 (-4,35; 1,73) signifikant schlechter ($p < 0,001$), beide Leistungen sind als unbeeinträchtigt zu interpretieren. Im Wortlisten Wiedererkennen wies die erste Testung ein signifikant schlechteres Ergebnis (-0,97 (-4,68; 1,45)) als die zweite Testung mit einem z-Wert von 0,46 (-3,45; 1,76), beide im Normbereich, auf. Die Ergebnisse des figuralen Abrufs schnitten mit einem z-Wert von -1,42 (-5,37; 2,24), leicht beeinträchtigt, zum Zeitpunkt der ersten Testung signifikant schlechter ab als zum zweiten Zeitpunkt (-1,22 (-5,34; 2,26), $p = 0,034$), die als unbeeinträchtigt zu bewerten sind. Im Bereich der kognitiven Flexibilität (Trail Making Test B) waren die Leistungen beider Testungen gleichermaßen unbeeinträchtigt, jedoch signifikant voneinander verschieden (1. Testung: -0,56 (-3,27; 2,34) versus 2. Testung: -0,43 (-3,27; 2,79), $p = 0,014$).

Folglich schnitt die die Studiengruppe bei der Ersten im Vergleich zur Zweittestung in neun Items signifikant schlechter ab, wobei in zwei Fällen nach der z-Standardisierung eine Abstufung in eine schlechtere Gruppe erfolgte.

Tabelle 9. Vergleich erste versus zweite Testung in der Studiengruppe.

	1. Testung	2. Testung	p-Wert
BDI	7 (0;19)	7 (0;19)	0,002
Z-Werte			
Semantische Wortfl.	-0,99 (-4,14; 2,19)	-0,93 (-3,97; 2,44)	0,908
Boston Naming	-0,35 (-5,36; 1,74)	0,15 (-5,36; 1,97)	0,003
MMST	-1,59 (-5,96; 1,65)	-1,50 (-7,26; 1,90)	0,746
Wortliste Lernen Total	-1,68 (-6,83; 1,95)	-1,52 (-6,83; 1,99)	0,022
Wortliste Lernen #1	-1,40 (-4,39; 2,83)	-1,22 (-3,87; 2,26)	0,030

Wortliste Lernen #2	-1,40 (-5,39; 2,12)	-1,25 (-5,93; 1,62)	0,056
Wortliste Lernen #3	-1,32 (-7,20; 1,39)	-1,26 (-7,20; 10,00)	0,019
Wortliste Abruf	-1,13 (-4,35; 1,82)	-0,99 (-4,35; 1,73)	<0,001
Wortliste Intrusionen	-0,67 (-3,99; 0,98)	-0,77 (-3,51; 0,98)	0,755
Wortliste Saving	-0,86 (-4,85; 4,70)	-0,80 (-4,85; 4,92)	0,523
Wortliste Wiedererkennen	-0,97 (-4,68; 1,45)	0,46 (-3,45; 1,76)	0,002
Figuren Abzeichnen	0,43 (-4,26; 1,74)	-1,22 (-5,34; 2,26)	0,089
Figuren Abrufen	-1,42 (-5,37; 2,24)	-1,22 (-5,34; 2,26)	0,034
Figuren Saving	-1,15 (-3,82; 2,35)	-0,98 (-3,83; 2,32)	0,051
Trail Making A	-0,36(-3,80; 2,27)	-0,45 (-3,75; 2,31)	0,120
Trail Making B	-0,56 (-3,27; 2,34)	-0,43 (-3,27; 2,79)	0,014
Trail Making A/B	-0,29 (-3,04; 2,63)	-0,07 (-2,88; 2,34)	0,001
Phonematische Wortfl.	-0,43 (-3,00; 2,35)	-0,40 (-3,83; 3,16)	0,835

4.3 Vergleich der verschiedenen Demenzformen untereinander

Von den 148 zwei Mal getesteten Studienteilnehmern hatte fast die Hälfte (48 %) ein mild cognitive impairment (MCI), 46 (31,1 %) eine Demenz vom Alzheimer-Typ. Bei 6 (4,1 %) lag eine frontotemporale, bei 14 (9,5 %) eine sonstige Demenz und 8 (5,4 %) eine Demenz bei andernorts klassifizierten Erkrankungen vor, die in der Gruppe „Andere“ zusammengefasst wurden. 14 (9,5%) Probanden hatten ausschließlich eine subjektive Beeinträchtigung und wurden hier ausgeschlossen. Beim Vergleich dieser drei Gruppen war die Häufigkeit an Depressionsdiagnosen signifikant verschieden, denn diese wurden in der Gruppe der „Anderen“ Demenzen mit 70,6 % mehr als dreimal so häufig wie bei Alzheimer-Demenzen gestellt ($p < 0,001$). Die Schwere der Depressionen unterschieden sich hingegen nicht signifikant. Alle weiteren Charakteristika waren vergleichbar und nicht signifikant verschieden (siehe Tabelle 10).

Tabelle 10. Vergleich der Patientencharakteristika von Probanden mit MCI vs. Alzheimer vs. Andere Demenz in der Studiengruppe.

	MCI (n=71)	Alzheimer (n=46)	Andere (n=17)	p-Wert
Alter	69 (50-85)	70 (52-91)	67 (51-85)	0,553
Männer	38 (53,5%)	26 (56,5%)	12 (70,6%)	0,443
Bildungsjahre	13 (7-19)	12,5 (8-20)	12 (10-18)	0,225
Rechtshänder	69 (97,2%)	44 (95,7%)	16 (94,1%)	0,805

Risikofaktoren				
Arterielle Hypertonie	27 (38%)	20 (43,5%)	9 (52,9%)	0,513
Hyperthyreose	10 (14,1%)	3 (6,5%)	1 (5,9%)	0,343
Diabetes Mellitus	7 (9,9%)	8 (17,4%)	4 (23,5%)	0,259
Hyperlipidämie	9 (12,7%)	5 (10,9%)	4 (23,5%)	0,410
Depressionsdiagnose	34 (47,9%)	9 (19,6%)	12 (70,6%)	<0,001
Ausprägung leicht	3 (4,2 %)	1 (2,2 %)	3 (17,6%)	0,347
Ausprägung mittel	18 (25,4%)	6 (13%)	8 (47,1%)	0,605
Ausprägung schwer	13 (18,3%)	2 (4,3%)	1 (5,9%)	0,129

Beim Vergleich der Testergebnisse der drei Subgruppen bestand ein signifikanter Unterschied in der Selbstbeurteilung der depressiven Symptomatik mit dem BDI, wobei die MCI-Patienten mit einem Median von 9 Punkten den höchsten Rohwert erreichten im Vergleich zu den Alzheimer Demenzen (5,5 Punkten) und den anderen Demenzen (7 Punkte, $p=0,004$ siehe Tabelle 11). Hinsichtlich der CERAD-Batterie Testung war die semantische Wortflüssigkeit zwischen den drei Gruppen signifikant verschieden ($p < 0,001$), die MCI-Gruppe schnitt mit einem z-Wert von -0,81 (-3,25; 2,19) besser ab als die Gruppe der Alzheimer Demenzen (-1,37 (-41,14; 1,26)), beides im unbeeinträchtigten Bereich, und die anderen Demenzen hatten mit einem z-Wert von -1,51 (-3,16; -0,25, $p<0,001$) das schlechteste Ergebnis, was im leicht beeinträchtigten Bereich lag.

Im Demenz-Screening Test (MMST) hatten die Probanden mit MCI mit einem z-Wert von -1,36 (-4,32; 1,65), also einer leichtgradigen Beeinträchtigung, ein signifikant besseres Ergebnis als die mit anderen Demenzformen (-2,38 (-3,88; 0,25) und die der Alzheimer Probanden mit einem z-Wert von -2,83 (-5,96; 1,20), bei denen die Leistung einer schwergradigen Beeinträchtigung entsprach ($p < 0,001$).

Bei den Leistungen im Bereich des verbalen Lernverhaltens (Wortlisten lernen total) lagen die Leistungen der MCI im leicht beeinträchtigten Bereich (-1,33 (-4,62; 1,93), die Ergebnisse der anderen Demenzformen (-2,00 (-5,21; 0,44)) waren mittelschwer beeinträchtigt und die der Alzheimer Demenzen hatten mit einem z-Wert von -3,05 (-6,83; 0,69) eine schwere Beeinträchtigung ($p < 0,001$). In den drei einzelnen Durchgängen bestand ebenfalls ein signifikanter Unterschied, wobei hier auch die MCI Probanden am besten abschnitten, vor denen mit anderen Demenzen und denen mit einer Alzheimer Demenz (siehe Tabelle 11).

Im verzögerten Abruf des verbalen Materials (Wortlisten Abruf) lag ein signifikanter Unterschied zwischen den Leistungen der drei Subgruppen vor, wobei Patienten mit MCI im Normbereich (-0,89 (-2,90; 1,82)), die mit anderen Demenzen mit einer leichtgradigen Beeinträchtigung (-1,31 (-3,73; 1,57) und die vom Alzheimer Typ mit einer mittelgradigen Beeinträchtigung (-2,28 (-4,35; 0,63) abschnitten ($p < 0,001$). Die gleiche, ebenso signifikant

verschiedene Verteilung bestand im Bereich des Wiedererkennens zwischen den drei Gruppen: Die Leistung der Patienten mit MCI lag im Normbereich (-0,67 (-3,85; 1,17)), wohingegen die der mit anderen Demenzen leicht beeinträchtigt (-1,40 (-4,68; 1,45)) und die der Alzheimer Probanden eine mittelgradige Beeinträchtigung (-2,05 (-4,68; 2,16) aufwiesen ($p < 0,001$). Im Bereich des figuralen Abrufes waren die Leistungen der drei Gruppen signifikant unterschiedlich beeinträchtigt, wobei Probanden mit MCI mit einem z-Wert von -1,04 (-4,17; 1,49) im unbeeinträchtigten Bereich lagen, und die beiden anderen Subgruppen mittelschwer beeinträchtigt waren (andere Demenzen (-1,94 (-4,74; 2,24)), Alzheimer Demenz (-2,26 (-5,37; 1,49), $p < 0,001$). Hinsichtlich der Verarbeitungsgeschwindigkeit (Trail Making Test A) schnitt die Alzheimer-Gruppe mit einem z-Wert von -0,87 (-3,80; 1,73) im Vergleich zur MCI-Gruppe mit 0,11 (-2,64; 2,27) und der Gruppe anderer Demenzen (-0,69 (-2,87; 1,60)) signifikant schlechter ab ($p = 0,003$), wobei alle Leistungen als unbeeinträchtigt zu interpretieren sind. Im Bereich der kognitiven Flexibilität unterschieden sich der Ergebnisse der drei Gruppen auch signifikant voneinander, wobei MCI (-0,12 (-2,67; 2,21)) und andere Demenzen (-1,04 (-2,81; 0,87)) als unbeeinträchtigt zu bewerten sind und die Alzheimer-Gruppe mit einem z-Wert von -1,54 (-3,27; 2,02) eine leichte Beeinträchtigung aufwies ($p < 0,001$). In der phonematischen Wortflüssigkeit bestand ebenso ein signifikanter Unterschied zwischen den Gruppen, wobei die Leistung der Probanden mit MCI (-0,28 (-2,21; 2,00)), denen mit anderen Demenzen (-0,95 (-2,07; 2,35)) sowie Alzheimer Demenzen (-0,81 (-3,00; 1,89)) alle im unbeeinträchtigten Bereich lagen ($p = 0,050$).

Dementsprechend waren die Testergebnisse zwischen Patienten mit MCI, Alzheimer und anderen Demenzen in 14 Items signifikant verschieden, wobei in 12 Fällen nach der z-Standardisierung eine Eingruppierung in verschiedene Gruppen vorlag. Insgesamt schnitten die Probanden mit einer Alzheimer Demenz in jedem Item (bis auf die semantische Wortflüssigkeit) am schlechtesten ab.

Tabelle 11. Vergleich der Testergebnisse der 1. Testung zwischen den dementiellen Subgruppen.

	MCI (n=71)	Alzheimer (n=46)	Andere (n=17)	p-Wert
BDI (Rohwertpunkte)				
1. Testung	9 (0;19)	5,5 (0;18)	7 (1;16)	0,004
Z-Werte (1.Testung)				
Semantische Wortflüssigkeit	-0,81 (-3,25; 2,19)	-1,37 (-4,14; 1,26)	-1,51 (-3,16; -0,25)	<0,001
Boston Naming	-0,14 (-3,68; 1,70)	-0,67 (-5,36; 1,39)	0,05 (-4,92; 1,74)	0,122
MMST	-1,36 (-4,32; 1,65)	-2,83 (-5,96; 1,20)	-2,38 (-3,88; 0,25)	<0,001

Wortliste Lernen Total	-1,33 (-4,62; 1,93)	-3,05 (-6,83; 0,69)	-2,00 (-5,21; 0,44)	<0,001
Wortliste Lernen #1	-1,20 (-4,39; 2,83)	-1,81 (-3,81; 0,92)	-1,63 (-3,87; 0,72)	0,008
Wortliste Lernen #2	-0,97 (-3,37; 1,23)	-2,68 (-5,93; 0,65)	-1,43 (-3,83; 0,15)	<0,001
Wortliste Lernen #3	-0,97 (-3,93; 1,17)	-3,26 (-7,20; 0,97)	-1,80 (-4,97; 0,20)	<0,001
Wortliste Abruf	-0,89 (-2,90; 1,82)	-2,28 (-4,35; 0,63)	-1,31 (-3,73; 1,57)	<0,001
Wortliste Intrusionen	0,35 (-2,92; 0,95)	-0,88 (-3,99; 0,98)	-0,79 (-2,73; 0,88)	0,288
Wortliste Saving	-0,60 (-3,65; 4,49)	-1,85 (-4,85; 1,15)	-0,66 (-4,64; 4,70)	0,003
Wortliste Wiedererkennen	-0,67 (-3,85; 1,17)	-2,05 (-4,68; 2,16)	-1,40 (-4,68; 1,45)	<0,001
Figuren Abzeichnen	0,45 (-2,82; 1,53)	0,32 (-4,26; 1,74)	-0,53 (-2,69; 1,50)	0,513
Figuren Abrufen	-1,04 (-4,17; 1,49)	-2,26 (-5,37; 1,49)	-1,94 (-4,74; 2,24)	0,001
Figuren Saving	-0,69 (-3,46; 2,35)	-1,85 (-3,82; 0,79)	-1,48 (-3,63; 1,43)	<0,001
Trail Making A	0,11 (-2,64; 2,27)	-0,87 (-3,80; 1,73)	-0,69 (-2,87; 1,60)	0,003
Trail Making B	-0,12 (-2,67; 2,21)	-1,54 (-3,27; 2,02)	-1,04 (-2,81; 0,87)	<0,001
Trail Making A/B	-0,28 (-2,62; 2,63)	-0,45 (-3,04; 1,76)	-0,52 (-2,12; 1,49)	0,274
Phonematische Wortflüssigkeit	-0,28 (-2,21; 2,00)	-0,81 (-3,00; 1,89)	-0,95 (-2,07; 2,35)	0,050

4.4 Extraktion sprachbasierter Biomarker

4.4.1 Analytierte Subgruppe

Bei insgesamt 138 (93,2%) Probanden der Studiengruppe lagen Audiofiles vor, die eine Analyse zur Generierung eines sprachbasierten Biomarkers ermöglichten, die anderen Probanden wurden bei dieser Analyse wegen fehlerhafter Aufnahmen oder schlechter Tonqualität ausgeschlossen. In diesem Kollektiv waren 71 (51,4%) Probanden weiblich, das mediane Alter betrug 67 (50-91) Jahre bei einer Ausbildungsdauer von 13 (7-20) Jahren. 131 (94,9%) Probanden waren Rechtshänder (siehe Tabelle 12). Von diesen Probanden waren 27 (19,5%) gesund, 67 (48,5%) litten an MCI und 44 (31,8%) an einer Alzheimer Demenz. Es lagen bis auf das Alter und die Geschlechterverteilung keine statistisch signifikanten Unterschiede hinsichtlich der demographischen Parameter zwischen den Gruppen vor (siehe Tabelle 12).

Tabelle 12. Deskriptive Statistik des eingeschlossenen Kollektivs (Gesund, MCI: mild cognitive impairment).

	Gesund (n=27)	MCI (n=67)	Alzheimer Demenz (n=44)	p-Wert
Geschlecht weiblich	20 (74,1%)	32 (47,8%)	19 (43,2%)	0,029
Alter	62 (51-86)	68 (50-85)	70 (52-91)	0,049
Ausbildungsdauer	13 (8-20)	13 (7-19)	12,5 (8-20)	0,174
Händigkeit rechts	24 (88,9%)	65 (97%)	42 (95,5%)	0,265
MMSE	29 (26-30)	28 (21-30)	25 (18-30)	<0,001
ki:e SB-C	0,47 (0,25-0,63)	0,39 (0,09-0,57)	0,22 (-0,13-0,59)	<0,001

4.4.2 Analytische und klinische Validierung

Sowohl der MMSE als auch der ki:e SB-C war zwischen den drei Gruppen signifikant verschieden, wobei im MMSE gesunde Probanden den höchsten Punktwert erreichten und Patienten mit AD den niedrigsten, hinsichtlich des ki:e SB-C in umgekehrter Assoziation (siehe Tabelle 12).

Zur analytischen Validierung des Ergebnisses des ki:e SB-C erfolgte eine Korrelationsanalyse mit dem MMSE Punktwert, es lag eine statistisch signifikante, starke Korrelation zwischen beiden Parametern vor ($r = 0,70$, $p < 0,001$, $d = 1,99$, siehe Abbildung 10).

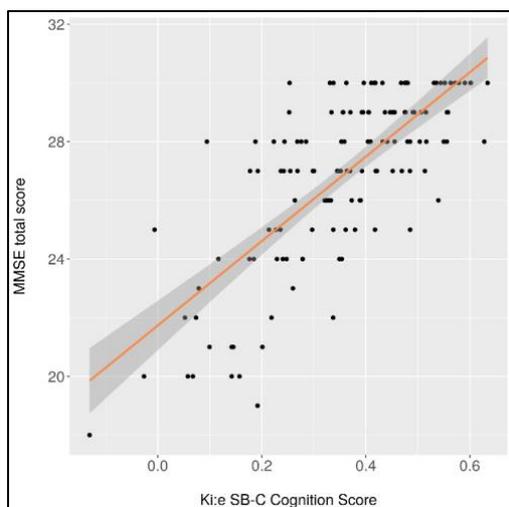


Abbildung 10. Korrelation zwischen MMSE (Mini Mental Status Examination) und ki:e SB-C Score für Kognition.

Zur klinischen Validierung erfolgte ein Gruppenvergleich der Biomarker-Scores zwischen den verschiedenen Demenzgruppen. Es bestand ein statistisch signifikanter Unterschied des ki:e SB-C-Scores zwischen den drei klinischen Gruppen ($p < 0,001$) (siehe Abbildung 11, siehe Tabelle 12).

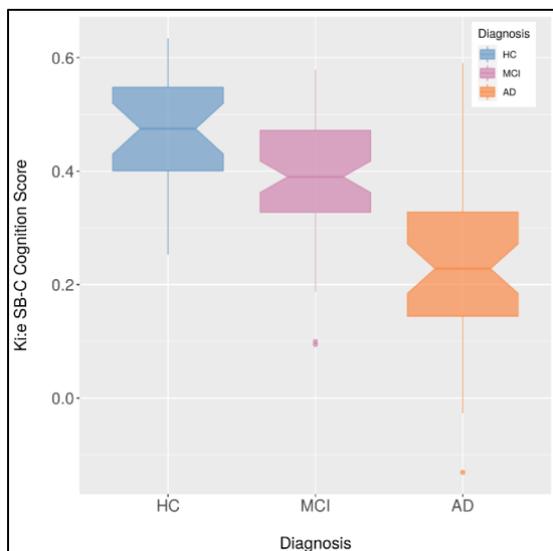


Abbildung 11. Gruppenvergleich ki:e SB-C Score zwischen HC, MCI und AD (HC: healthy control, MCI: mild cognitive impairment, AD: Alzheimer`s disease).

4.4.3 Ermittlung der diagnostischen Testgenauigkeit vom SB-C Biomarker

Zur Beurteilung des klinischen Nutzens erfolgte eine Analyse der diagnostischen Performance des SB-C Biomarkers hinsichtlich der Unterscheidung zwischen den klinischen Gruppen (gesund vs. MCI, gesund vs. AD, gesund vs. MCI / AD und MCI vs. AD).

Tabelle 13. Diagnostische Genauigkeit (Performanz Matrix) SB-C Biomarker.

	Gesund vs. MCI	Gesund vs. AD	Gesund vs. MCI/AD	MCI vs. AD
F1-Score	0,77	0,89	0,82	0,79
Accuracy	0,70	0,87	0,74	0,76
Balanced accuracy	0,71	0,89	0,75	0,76
ROC AUC	0,66	0,87	0,78	0,74
Sensitivität	0,69	0,82	0,73	0,75
Spezifität	0,74	0,96	0,78	0,77

PPV ¹	0,87	0,97	0,93	0,83
1-NPV ²	0,51	0,24	0,59	0,33

Die Modelle, die in diesem Rahmen auf Grundlage der ki: e SB-C- Merkmale trainiert wurden wiesen insgesamt mit einer balanced accuracy von 0,75 (ROC-AUC 0,78) sehr gute Ergebnisse beim Screening einer neurokognitiven Pathologie (MCI vs. AD) auf. Probanden mit einer Alzheimer Demenz hatten den schlechtesten SB-C Score (siehe Abbildung 11) und ließen sich durch diesen Score sehr gut gegenüber den gesunden Probanden abgrenzen (balanced accuracy 0,89; ROC-AUC 0,87). Der trainierte Algorithmus war mit einer balanced accuracy von 0,71 (ROC-AUC 0,66) gut in der Lage, zwischen gesunden Probanden und Personen mit einer MCI zu unterscheiden. Darüber hinaus konnte der Algorithmus mit einer balanced accuracy von 0,76 (ROC-AUC 0,74) gut zwischen MCI und einer Alzheimer Demenz unterscheiden (siehe Abbildung 11, Abbildung 12). Für die Sensitivität und Spezifität des Modells bestanden je nach Modell des Screenings leicht differierende Werte, wobei die Spezifität immer über 0,74 und die Sensitivität immer über 0,69 betrug (siehe Tabelle 13).

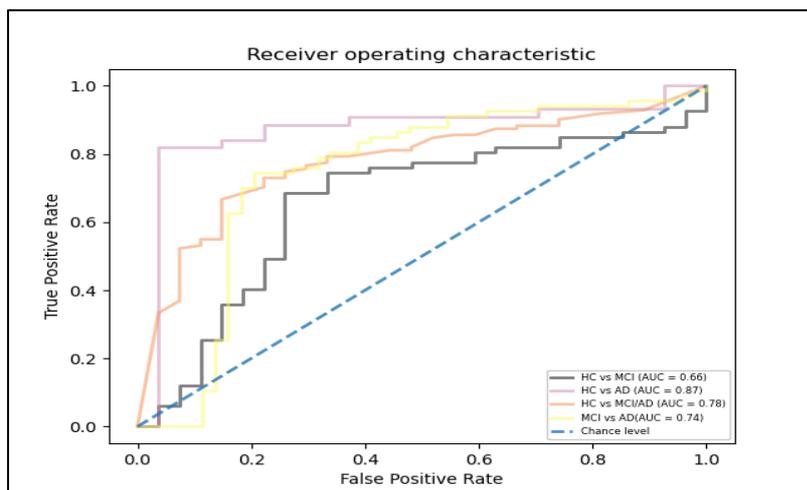


Abbildung 12. Vergleichende Receiver operating characteristic curve (ROC) für die vier Modelle (HC: healthy control/gesunde Kontrollgruppe, MCI: mild cognitive impairment, AD: Alzheimer Demenz).

4.5 Aussagegüte Paper-Pencil-Subtests vs. Sprachbiomarker (SB-C)

Die diagnostische Performance des SB-C Sprachbiomarkers wurde abschließend mit der der konventionell erhobenen CERAD-Testbatterie verglichen.

¹ PPV: positiv prädiktiver Wert

² NPV: negativ prädiktiver Wert

4.5.1 Unterscheidung gesunde Kontrollen vs. MCI

Bei der Unterscheidung zwischen gesunden Probanden der Kontrollgruppe (HC) und den Probanden mit einer MCI waren sowohl drei Subtests des CERAD (semantische Wortflüssigkeit, Wortlisten Lernen Total sowie der Wortlisten Abruf) als auch der Sprachbiomarker dazu in der Lage, statistisch signifikant zwischen den beiden Probandengruppen zu unterscheiden. Dabei war die Performance des Subtests „Wortlisten Lernen Total“ (AUC = 0,633; Sensitivität = 1,0; 1-Spezifität = 0,67) am besten. Der SB-C Score (AUC = 0,70; Sensitivität = 0,81; 1-Spezifität = 0,46) war im Vergleich zum Wortlisten Abruf (AUC = 0,65; Sensitivität = 0,48; 1-Spezifität = 0,17) besser dazu in der Lage zwischen den beiden Gruppen zu unterscheiden (siehe Tabelle 14, Abbildung 13).

Tabelle 14. Unterscheidung Gesund vs. MCI.

	Semantische Wortflüssigkeit	Wortlisten Lernen-Total	Wortlisten Abruf	Phonematische Wortflüssigkeit	SB-C Score
AUC	0,656	0,633	0,651	0,598	0,707
p-Wert	0,018	0,044	0,023	0,140	0,002
Sensitivität	0,407	1	0,481	0,778	0,815
1-Spezifität	0,09	0,672	0,179	0,493	0,463

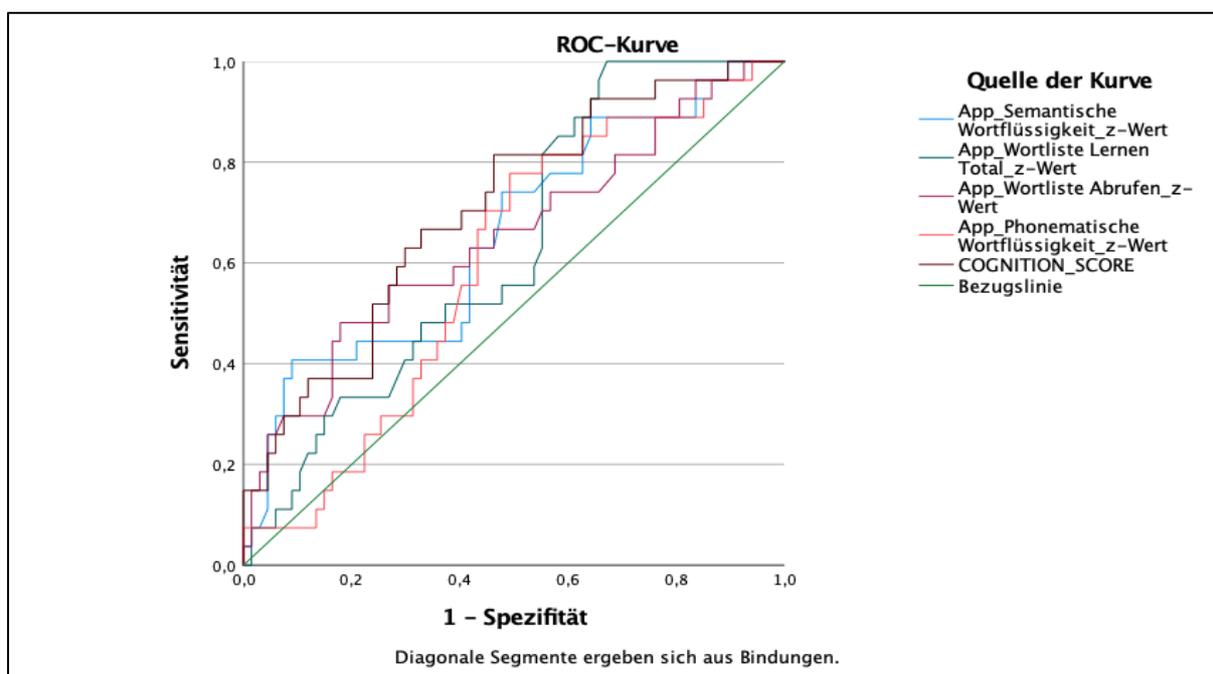


Abbildung 13. Receiver Operating Characteristic (ROC) Gesund vs. MCI.

4.5.2 Unterscheidung gesunde Kontrollen vs. AD

Im zweiten Gruppenvergleich zwischen gesunden Probanden der Kontrollgruppe (HC) und den Probanden mit einer AD konnten sowohl vier Subtests als auch der Sprachbiomarker zwischen den beiden Gruppen unterscheiden. Die Leistung des Wortlisten Lernen Total und des Wortlisten Abruf waren mit der des SB-C Scores vergleichbar, wobei die AUC des SB-C Score am höchsten war (AUC = 0,909). Jedoch hatten alle vier Subtests eine höhere Spezifität als der SB-C Score (siehe Tabelle 15, siehe Abbildung 14).

Tabelle 15. Unterscheidung Gesund vs. AD

	Semantische Wortflüssigkeit	Wortlisten Lernen- Total	Wortlisten Abruf	Phonematische Wortflüssigkeit	SB-C score
AUC	0,796	0,880	0,867	0,710	0,909
p-Wert	<0,001	<0,001	<0,001	0,003	<0,001
Sensitivität	0,778	0,963	0,963	0,778	0,926
1-Spezifität	0,25	0,25	0,25	0,364	0,159

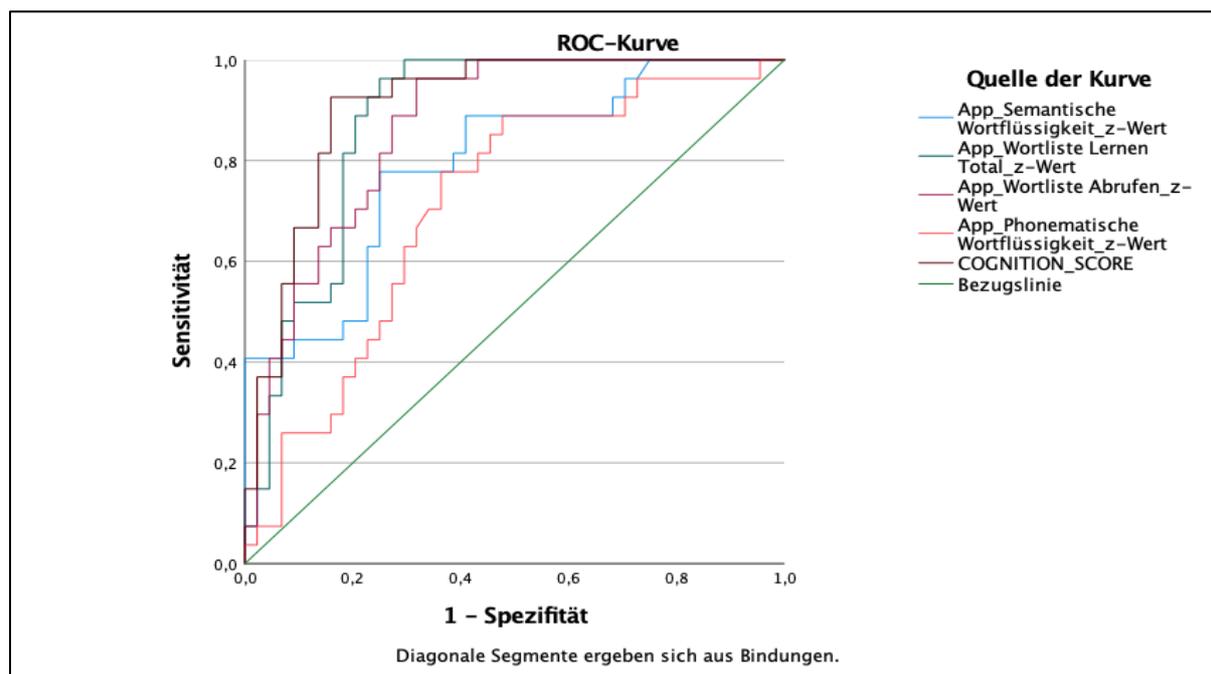


Abbildung 14. Receiver Operating Characteristic (ROC) zur Unterscheidung Gesund vs. AD.

4.5.3 Unterscheidung gesunde Kontrollen vs. MCI/AD

Im dritten Gruppenvergleich konnten erneut sowohl die vier Subtests des CERADs als auch der SB-C zwischen der Kontrollgruppe und MCI oder AD-Patienten unterscheiden, jedoch mit unterschiedlicher Performance. Auch hier war die Leistung des SB-C Scores mit der Leistung

des Wortlisten Lernen Total und des Wortlisten Abruf vergleichbar. Jedoch hatte der Subtest „Wortlisten Lernen Total“ (AUC = 0,73; Sensitivität = 1; 1-Spezifität = 0,52) eine höhere Sensitivität und Spezifität als der SB-C Score (AUC = 0,78; Sensitivität = 0,81; 1-Spezifität = 0,33) (siehe Tabelle 16).

Tabelle 16. Gesund vs. MCI/AD.

	Semantische Wortflüssigkeit	Wortlisten Lernen- Total	Wortlisten Abruf	Phonematische Wortflüssigkeit	SB-C score
AUC	0,712	0,731	0,737	0,642	0,787
p-Wert	0,001	<0,001	<0,001	0,022	<0,001
Sensitivität	0,741	1	0,556	0,788	0,815
1-Spezifität	0,387	0,523	0,198	0,441	0,333

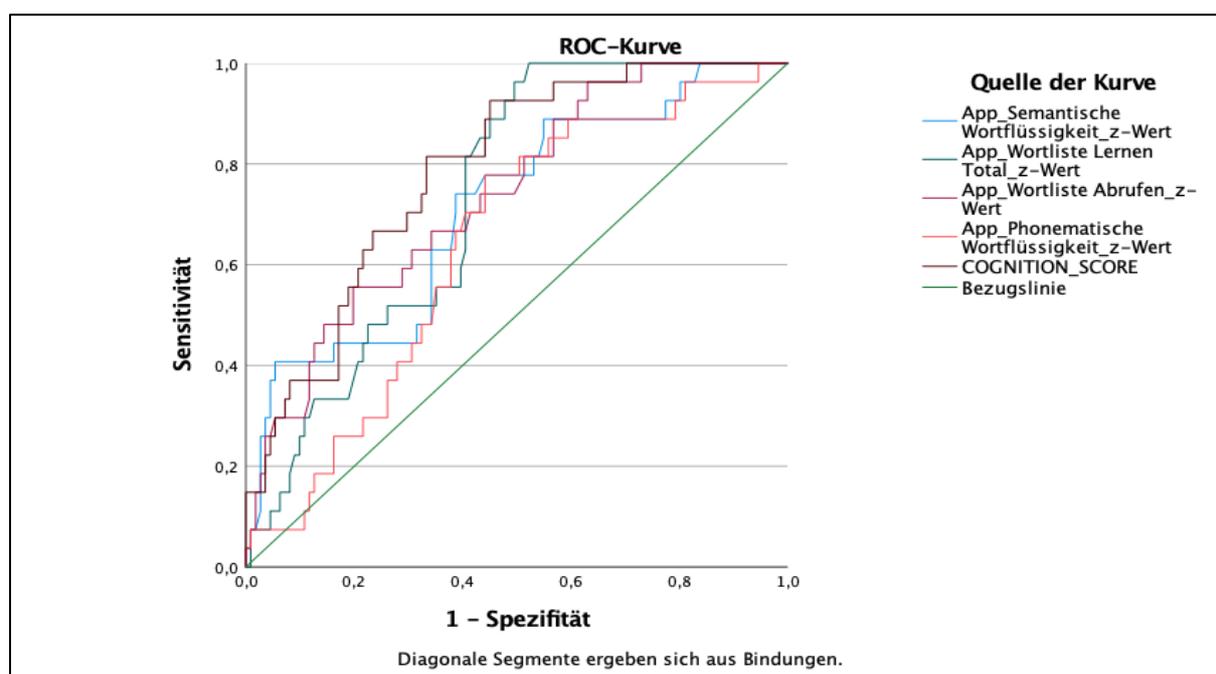


Abbildung 15. Receiver Operating Characteristic (ROC) zur Unterscheidung von Gesund vs. MCI/AD.

4.5.4 Unterscheidung MCI vs. AD

Im vierten Vergleich zwischen den beiden Gruppen (Probanden mit einer MCI und den Probanden mit einer Pathologie AD) konnten sowohl vier Subtests als auch der Sprachbiomarker, statistisch signifikant, zwischen den beiden Probandengruppen unterscheiden. Dabei hatte der Subtest „Wortlisten Lernen Total“ (AUC = 0,73; Sensitivität = 1; 1-Spezifität = 0,52) zwar eine geringere AUC und Spezifität als der SB-C Score

(AUC = 0,78; Sensitivität = 0,81; 1-Spezifität = 0,33), aber wiederum eine höhere Sensitivität. Der SB-C Score war mit einer besseren Sensitivität und Spezifität in seiner Aussagegüte jedoch besser als der Subtest „Wortlisten Abruf“ (siehe Tabelle 14, Abbildung 16).

Tabelle 14. Unterscheidung von MCI vs. AD.

	Semantische Wortflüssigkeit	Wortlisten Lernen- Total	Wortlisten Abruf	Phonematische Wortflüssigkeit	SB-C score
AUC	0,712	0,731	0,737	0,642	0,787
p-Wert	0,001	<0,001	<0,001	0,022	<0,001
Sensitivität	0,741	1	0,556	0,788	0,815
1-Spezifität	0,387	0,523	0,198	0,441	0,333

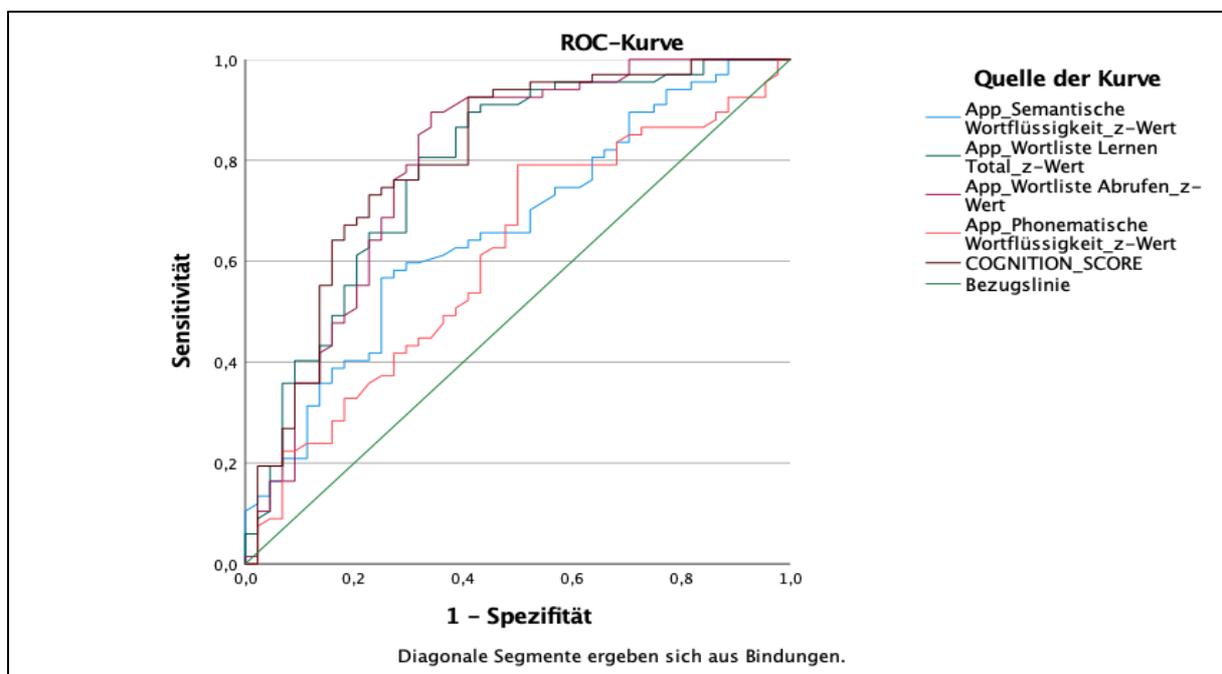


Abbildung 16. Receiver Operating Characteristic (ROC) zur Unterscheidung von MCI vs. AD.

Es lässt sich demnach feststellen, dass der SB-C Score in allen Gruppenvergleichen (HC vs. MCI vs. AD) genauso gut oder teilweise besser zwischen den Gruppen unterscheiden konnte als die verbalen Sprachaufgaben des CERADs (Wortflüssigkeit, verbales Lernen und verbaler Abruf). Beim Vergleich zwischen den gesunden Kontrollprobanden (HC) und Probanden mit kognitiver Beeinträchtigung (MCI/AD) war die Aussagekraft des SB-C gleichwertig mit der des Wortlisten Lernens und Abrufes.

5. Diskussion

Aufgrund einer deutlich gestiegenen Lebenserwartung hat die Prävalenz von Demenzerkrankungen in den letzten Jahrzehnten stark zugenommen. Kognitive Screeningtests eignen sich besonders gut dazu, Personen mit einem Risiko für dementielle Störungsbilder früh zu identifizieren. Traditionell werden hierfür Papier-Bleistift-Tests verwendet, was jedoch zeit- und arbeitsaufwändig ist. Aus diesem Grund werden zunehmend digitale kognitive Screeningverfahren, einschließlich sprachbasierter digitaler Beurteilung, durchgeführt, um die Diagnosestellung sowie die Überwachung des Krankheitsverlaufes zu unterstützen.

Die ELEMENT-Studie zielte in diesem Zusammenhang darauf ab, die diagnostische Genauigkeit eines traditionellen Paper-Pencil-Verfahrens (CERAD) im Rahmen einer neuropsychologischen Demenzabklärung mit einer computerbasierten Version, der Δ elta (Delta)-App, bei Probanden mit kognitiver Beeinträchtigung unterschiedlicher Ätiologie und Schweregrad zu vergleichen. Aus den Sprachdateien, die die App während der kognitiven Testung aufzeichnete, wurden mithilfe einer automatischen Spracherkennung digitale Biomarker extrahiert und deren diagnostische Aussagekraft im Rahmen der Diagnostik neurodegenerativer Erkrankungen überprüft. Um einen direkten Vergleich beider Untersuchungsmethoden an demselben Patientenkollektiv zu ermöglichen, wurde ein randomisiertes, einfach verblindetes, zweiarmiges Studiendesign mit cross-over Design verwendet; es wurden insgesamt 200 Probanden und 27 Kontrollpersonen eingeschlossen. Letztlich hatte die neue Testmodalität der mit einem iPad® elektronisch unterstützten Testung nur einen geringen Einfluss auf die Testergebnisse bei insgesamt repräsentativen Testergebnissen. Hinsichtlich des digitalen Sprachbiomarkers, der aus automatisch extrahierten Sprachdateien während der Testung entwickelt wurde, ließen sich gute bis sehr gute Ergebnisse bei der Unterscheidung der einzelnen klinischen Subgruppen erzielen. Die Klassifikation der Subgruppen durch den SB-C Score war mit der Aussagekraft der sprachbasierten Subtests des CERADs (Wortlisten Lernen, Wortlisten Abruf) vergleichbar, teilweise sogar überlegen.

5.1 Interpretation der erhobenen Testergebnisse

Zunächst erfolgte eine Analyse der Testergebnisse zur Beurteilung von **externer wie interner Validität**. Um die Ergebnisse der Studiengruppe mit einer gesunden Kontrollstichprobe vergleichen zu können, begann ab Dezember 2020 die Rekrutierung der Kontrollpersonen. Über einen Studienflyer wurden sowohl auf dem Klinikgelände sowie im sozialen Umfeld des Studienteams Probanden rekrutiert. Die Kontrollgruppe umfasste 27 Probanden, die dem Studienprotokoll entsprechend nur einmal mittels App getestet wurden. Im direkten Vergleich zur Studiengruppe lag hinsichtlich der Patientencharakteristika ein signifikanter

Altersunterschied vor, denn die Probanden aus der Studiengruppe waren sieben Jahre älter. Der Altersunterschied lässt sich vermutlich darauf zurückführen, dass als potentielle Kontrollpersonen besonders Klinikpersonal sowie Bekannte im häuslichen Umfeld angeworben wurden. Die Erreichbarkeit älterer Personen mithilfe des Flyers gestaltete sich als eher schwierig. Aus der Forschung ist bekannt, dass die Rekrutierung und der Verbleib von älteren Probanden in Forschungsstudien eine Herausforderung darstellt (162), was durch eine geringere Einschluss- sowie höhere Dropout-Quote in dieser Gruppe erklärt werden kann (163). Um den Altersunterschied anzugleichen, wäre es zukünftig sinnvoll, beispielsweise vermehrt bei Hausärzten oder ambulanten Neurologen/Psychiatern zu werben und auch hier die Angehörigen mehr in den Studienprozess mit einzuschließen.

Wie zu erwarten, hatten die Kontrollpersonen bessere **Testergebnisse** als die die Probanden der Studienkohorte. So schnitt die Kontrollgruppe nach z-Standardisierung in allen 18 Subtests des CERADs besser ab, wobei sich in 15 Tests auch ein signifikanter Unterschied ergab. Diese Ergebnisse unterstreichen die Validität der erhobenen Testergebnisse. In einer Studie von Karrasch und Kollegen (2005) zur Untersuchung der kognitiven Leistung von gesunden Kontrollprobanden, Personen mit MC und AD, gemessen mit der CERAD-Testbatterie, waren die Testergebnisse vergleichbar. Auch hier unterschied sich die Kontrollgruppe bei fast allen Subtests von Personen mit einer Alzheimer Demenz und MCI-Patienten außer im Subtest zum Erlernen von Wortlisten (164).

Die Frage, ob die **Testmodalität einen Einfluss auf die Testergebnisse** hat, konnte nur in der Studiengruppe überprüft werden, da nur die Studienpatienten mit beiden Testmodalitäten getestet wurden. Die Testung mittels App hatte in 5 der 18 Subtests des CERADs einen statistisch signifikanten Einfluss auf die Ergebnisse, wobei nach der z-Standardisierung nur in zwei dieser Tests eine Eingruppierung in eine schlechtere Gruppe bzw. Beurteilungskategorie erfolgte. So schnitten Probanden bei Testung mittels App im Untertest „Wortlisten lernen“ signifikant schlechter ab. Der Test umfasst insgesamt drei Durchgänge (Wortliste lernen, Wortlistenabruf, Wortliste wiedererkennen), die separat bewertet werden, sowie einen Gesamtscore („Wortlisten Lernen total“), der die Ergebnisse aller drei Durchgänge zusammenfasst. Ein statistisch signifikanter Unterschied bestand lediglich im ersten Testdurchgang sowie im Gesamtscore. Folglich ist zu diskutieren, ob der Unterschied in den Testergebnissen dadurch zustande gekommen ist, dass die Probanden sich zu Beginn der Testung mittels App an den Modus der Tonaufnahme gewöhnen mussten: erst nach einem Tonsignal erfolgt die Aufnahme sowie die Auswertung der Ergebnisse, anders als im klassischen Paper Pencil Verfahren – hier zeichnete der Untersucher alle Antworten ohne zeitliche Verzögerung durch das Tonsignal auf und ließ diese mit in die Auswertung einfließen. Aus diesem Grund ist die die Bewertung und Validierung psychometrischer Eigenschaften von computergestützten Verfahren sehr wichtig, um das Risiko widersprüchlicher Ergebnisse zu

minimieren und die Gleichwertigkeit zwischen den traditionellen und digitalen Testungen zu gewährleisten (165). In anderen Studien z. B. von Morrison und Kollegen, die die Leistung eines mit dem iPad durchgeführten Wortlisten-Erinnerungstest mit dem standardisierten Rey Auditory Learning Test (RAVLT) an kognitiv beeinträchtigten Probanden und einer gesunden Kontrollgruppe verglichen, waren die Leistungen der Teilnehmer zwischen den beiden Testmodalitäten ebenfalls gleichwertig und genauso sensitiv (166). Auch eine Metaanalyse aus dem Jahr 2018, welche die diagnostische Leistung von computergestützten und Paper Pencil Gedächtnistests bei Patienten mit MCI und einer Demenz verglich, kam zu dem Ergebnis, dass sowohl bei verbalen als auch visuellen Gedächtnistests mit beiden Modalitäten eine vergleichbare diagnostische Leistung erzielt werden kann (167). Es ist also festzuhalten, dass der Einfluss der Testmodalität auf die Testergebnisse in der vorliegenden Arbeit allenfalls marginal war.

Ein weiterer möglicher Confounder, also ein an sich unerwünschter Einflussfaktor auf die Testergebnisse, hätte andererseits auch der **Zeitpunkt der Testung** sein können. Trotz des dreimonatigen Abstandes zwischen erster und zweiter Testung, welcher etwaige Übungs- und Erinnerungseffekte minimieren sollte, schnitt die Studiengruppe bei der Erst- im Vergleich zur Zweittestung in neun Subtests des CERADs signifikant schlechter ab, wobei in nur zwei Subtests nach der z-Standardisierung die Ergebnisse als schlechter zu interpretieren sind. Auch in der Literatur finden werden bei wiederholter Durchführung von neuropsychologischen Tests, z. B. im Rahmen der Überwachung von degenerativen Krankheitsverläufen, eine Verbesserung der Testergebnisse beschrieben (**Übungseffekte**). Die Nichtberücksichtigung solcher Übungseffekte kann zu fehlerhaften diagnostischen Einschätzungen führen. Gemeinsame Moderatoren solcher Effekte können das Alter der Teilnehmer, klinische Diagnosen und die Länge des Test-Retest-Intervalls sein (168). In der Stichprobe von gesunden Erwachsenen ist eine Verbesserung zwischen Test und Re-Test in verbalen Wortlisten Tests prinzipiell zu erwarten und spiegelt vermutlich die zunehmende Gewöhnung und die Vertrautheit mit dem Testverfahren wieder (169). Ein Übungseffekt wurde auch unabhängig von der durchgeführten Testmodalität beobachtet, also unabhängig davon, ob zuerst mit dem iPad oder dem Standardverfahren getestet wurde (166). Tests mit Modifikationen der Antworten zur Vermeidung von Lerneffekten waren im CERAD als auch der computerisierten Version (Delta App) nicht vorgesehen, was als mögliche Erweiterung bzw. Verbesserung diskutiert werden könnte. Generierte Alternativformen können Übungseffekte zwar verringern, aber letztlich auch nicht ausschließen (170). Dies könnte daran liegen, dass dann neben der Erinnerung an bestimmte Aufgaben, auch die Vertrautheit und erlernte Strategien zur erfolgreichen Bewältigung von Aufgaben führen (171).

Des Weiteren lag haben wir überprüft, ob bei verschiedenen kognitiven Einschränkungen bzw. dementiellen Stadien ein **Unterschied in der kognitiven Beeinträchtigung, also ein**

Unterschied hinsichtlich der Testergebnisse im CERAD vorlag. Von den 148 zweimal getesteten Studienteilnehmern hatte fast die Hälfte (48 %) ein mild cognitive impairment (MCI) und 31,1 % eine Demenz vom Alzheimerstyp, die restlichen Patienten hatten andere Demenzformen. Beim Vergleich dieser drei Gruppen war die Häufigkeit an Depressions-Diagnosen signifikant verschieden. Sowohl Depressionen als auch kognitive Beeinträchtigungen sind gleichermaßen Erkrankungen des Alters, die zusätzlich gemeinsam auftreten können. Die exakte Wechselwirkung zwischen den Störungsbildern ist komplex und bislang nicht endgültig erforscht. Depressionen sollen die Manifestation einer Demenz beeinflussen (172), können aber auch als Prodrom einer dementiellen Erkrankung auftreten oder auch eine Reaktion auf den kognitiven Verfall darstellen (173). Die Testergebnisse zwischen Patienten mit MCI, Alzheimer und anderen Demenzen waren in 14 Subtests signifikant verschieden, wobei in 12 Fällen nach z-Standardisierung unterschiedliche Gruppeneinstufungen erfolgten. Insgesamt schnitten die Probanden mit einer Alzheimer Demenz in jedem Item (bis auf die semantische Wortflüssigkeit) am schlechtesten ab. Diese Testergebnisse unterstützten die Hypothese, dass verschiedene Stadien einer kognitiven Einschränkung zu Unterschieden in den Testergebnissen führen. In einer weiteren Studie, die ebenfalls mithilfe der CERAD-Testbatterie die Leistungsfähigkeit gesunder Probanden und Personen mit einer MCI oder einer AD verglich, bestand in vier Subtests ein Unterschied zwischen der MCI und AD-Gruppe, in denen die AD-Gruppe erneut schlechter abschnitt. Fast alle Tests waren empfindlich genug, um beim Vergleich zwischen gesunder Kontrollgruppe und Alzheimer Patienten signifikant zu unterscheiden (164). Diese Ergebnisse stimmen mit unseren Studienergebnissen überein. Fillenbaum und Kollegen untersuchten ebenfalls die Performance im CERAD, hier jedoch bei Personen mit und ohne dementielle Erkrankung. Kongruent mit unseren Ergebnissen bestand ein Unterschied hinsichtlich des Leistungsniveaus, denn die Leistung nahm mit zunehmendem Schweregrad der Demenz ab (174).

Des Weiteren sollte in dieser Studie überprüft werden, ob die delta-App **im klinischen Alltag einsetzbar** ist und das neuropsychologische Testen erleichtert, sowie ob eine valide Erhebung der Testergebnisse möglich ist. Im Rahmen der neuropsychologischen Untersuchungen war die App gut einsetzbar, die Bedienung war problemlos, es kam zu keinen Systemabstürzen. Die Arbeit an sich wurde erleichtert, indem die Antworten des Patienten automatisch aufgezeichnet wurden, somit war ein erneutes Anhören auch nach der Testung zur Bewertung möglich. Die Aufnahme an sich sowie eine automatische Auswertung kann potentiell die Zeit, die das Personal zum Bewerten der Leistung des Probanden benötigt, verkürzen, was als vorteilhaft zu bewerten ist (175). Zusätzlich erleichterte die App das Aufzeichnen von Reaktionszeiten wie z. B. bei den Wortflüssigkeitsaufgaben (semantische und phonematische Wortflüssigkeit) und den Aufgaben zur Verarbeitungsgeschwindigkeit und kognitiven

Flexibilität (Trail Making A & B), da dies automatisch erfolgte. Chin und Kollegen hoben bei einer Testung mit digitalen Geräten ähnliche positive Aspekte hervor: So sei es möglich, nicht nur die Endergebnisse einer Testung zu erfassen, sondern auch Informationen über den Testprozess, einschließlich der Antwortsequenzen und Reaktionszeiten. Insgesamt könnten Tablet- oder Computergestützte Testgeräte die Sensitivität und Durchführbarkeit kognitiver Tests zur Erkennung und Beobachtung kognitiver Veränderungen sogar noch weiter verbessern, was auch größer angelegte kognitive Testungen in klinischen Studien ermöglichen würde (175). Die Probanden nahmen die veränderte Testmodalität gut an, da die Testung von einer geschulten Psychologin durchgeführt wurde und diese zur Unterstützung der Probanden bei der Eingabe der Antworten verfügbar war. Diese positive Akzeptanz lässt sich auch in der weiterführenden Literatur finden. Die Probanden scheinen sich in iPad-basierter Interaktion wohl zu fühlen und sind dazu bereit, zukünftige Testungen wieder im selben Format durchzuführen (176).

Nichtsdestotrotz muss festgehalten werden, dass der Anteil an **Probanden, die aus der Studie ausgeschlossen wurden**, mit 26 % sehr hoch war. Als mögliche Gründe für den Ausschluss sind die geringere kognitive Leistungsfähigkeit der Dropout-Gruppe, eine höhere Rate an Komorbiditäten sowie der zeitliche Abstand zwischen beiden Testzeitpunkten von drei Monaten ebenso wie die andauernde Corona-Pandemie zu nennen. Die unterschiedlichen kognitiven Leistungsniveaus der Studienteilnehmer und derer, die als Dropout-Gruppe ausschieden, bestätigte sich mit Hilfe des statistischen Vergleiches der Leistungen in den neuropsychologischen Subtests des CERADs. Eine Studie von Beishuizen et al. aus dem Jahr 2017 fand ähnliche Ergebnisse bezüglich der kognitiven Einschränkung als mögliche Ursache für den Ausschluss bei vergleichbaren Bedingungen. Neben den schlechteren kognitiven Funktionen bei älteren Probanden waren auch ein höheres Lebensalter, eine Depressionsdiagnose und weitere körperliche Beeinträchtigungen mit einem Abbruch verbunden (177). Die Rate an Dropouts, die mit dem kognitiven Status einer Person zusammenhängen, können das Risiko mit sich bringen, dass es zu einer verzerrten Schätzung des kognitiven Krankheitsverlaufes in der Form einer Unterschätzung von schweren dementiellen Verläufen kommt (178), was es notwendig machen würde, dieses selektive Dropout zu berücksichtigen. In einer Zwischenanalyse der Ergebnisse im Oktober 2021 fiel der hohe Anteil an Dropouts auf, weswegen die Art und Weise der Vereinbarung des zweiten Testtermins sowie die terminliche Erinnerung vom Studienteam geändert wurden. Da Menschen mit einer Demenz häufig Unterstützung bei alltäglichen Aufgaben benötigen, die ein Zeitmanagement erfordern, und es daher notwendig sein kann, sie an Termine zu erinnern (179, 180), erhielten die Probanden nicht wie zuvor den zweiten Termin zum Zeitpunkt der ersten Testung. Ein Terminvorschlag an den Probanden erfolgte vier Wochen vor der zweiten Testung schriftlich durch das Studienteam mit der Bitte um Rückmeldung, ob dieser bei dem

Probanden zeitlich möglich sei. In der rückblickenden Betrachtung wäre es vermutlich von Vorteil gewesen, einen Angehörigen des Probanden mit in die terminliche Koordination mit einzubeziehen, um die kognitiven Defizite des Probanden, die vermutlich auch die Erinnerung an Termine betreffen, auszugleichen. Mit Beginn der Corona Pandemie im Februar 2020 folgte die Ungewissheit, inwieweit die Studie fortgesetzt werden könne, da eine Verunsicherung bei den Studienteilnehmern bestand und die Hygiene- sowie Kontaktbeschränkungsregeln für ambulante Patienten im klinischen Kontext die Fortsetzung der Studie zunächst erheblich erschwerten. Der Ausbruch der Corona Pandemie hatte somit nicht nur Auswirkungen auf das alltägliche Leben, sondern auf klinische Studien. Der Umstand, dass die Aufnahme von Patienten in laufende Studien, sowie der Ablauf klinischer Studien und somit auch die Datenerhebung durch die Corona-Pandemie negativ beeinflusst wurden, ist auch schon andernorts beschrieben worden (181, 182). Letztlich wirkte er sich auch negativ auf die Rekrutierung der Kontrollpersonen aus.

5.2 Entwicklung eines Sprachbasierten Biomarkers

Da die Delta App mit einer automatisierten Spracherkennung während der Durchführung der sprachgebundenen kognitiven Tests Sprachdateien produzierte, wurde im Rahmen dieser Arbeit auch angestrebt, aus diesen Aufnahmen zusätzliche Informationen zu gewinnen: Könnten daraus digitale Biomarker extrahiert werden und können diese beim Screening von MCI und Alzheimer nützlich sein? Hierfür wurden aus der Studiengruppe mit (148 Probanden und 27 Kontrollpersonen) drei Untergruppen extrahiert, die gesunde Kontrollgruppe (healthy controls, HC, n = 27), Probanden mit einer leichten Beeinträchtigung (MCI, n = 71) und Patienten mit einer Alzheimer Demenz (AD, n = 46). Bei den Probanden lagen allerdings nur in 138 Fällen verwendbare Audiofiles vor.

Der ki:elements Sprachbiomarker für Kognition (ki:e SB-C, ki elements, 2022) ist ein zusammengesetzter Score, der aus drei verschiedenen neurokognitiven Subscores (Lernen & Gedächtnis, Exekutive Funktionen, Verarbeitungsgeschwindigkeit) gebildet wird. Diese Scores setzen sich wiederum aus 50 automatisch extrahierten Sprachmerkmalen zusammen. Als Input für diesen SB-C fungierten in unserer Studie die Sprachaufnahmen der CERAD-NP sowie der semantischen Wortflüssigkeit (Semantic Verbal Fluency, SVF). Auf Basis des V3-Rahmens wurde der sprachbasierte Biomarker in den drei Schritten Verifizierung, analytischer und sodann klinischer Validierung (159) beurteilt. Für die *analytische Validierung* wurde der ki:e SB-C mit dem MMSE (bzw. MMST) verglichen, einem Standardmaß für Kognition in der Population für Probanden mit kognitiver Beeinträchtigung. Der MMSE hatte eine starke Korrelation mit dem ki: e SB-C global score, was dafür spricht, dass der digitale Score auch wirklich ein Maß für Kognition erfasst. Im Rahmen der *klinischen Validierung* wurde ein Gruppenvergleich der Biomarker-Scores zwischen den verschiedenen Gruppen (HC, MCI,

AD) durchgeführt, es bestand ein signifikanter Unterschied des ki:e SB-C-Scores zwischen allen drei Gruppen. Es ließ sich also schlussfolgern, dass der ki:e SB-C Score diese klinisch definierten Krankheitsstadien tatsächlich widerspiegelt. Zur Beurteilung des *klinischen Nutzens* erfolgte eine Analyse der diagnostischen Genauigkeit des SB-C hinsichtlich der Unterscheidung zwischen den verschiedenen Gruppen. Der Biomarker Score hatte auch hier gute bis sehr gute diskriminatorische Aussagekraft.

König et al. veröffentlichten in diesem Zusammenhang 2015 eine Studie, in der eine automatische Sprachanalyse zur Identifizierung von MCI-Patienten und Personen mit Alzheimer im Frühstadium eingesetzt wurde (siehe Tabelle 17, S. 68). Die Autoren extrahierten ebenfalls Merkmale aus mehreren kurzen kognitiven Sprachaufgaben (Countdown, Sentence Repeating, Picture description, semantic fluency) und verwendeten die Methode des maschinellen Lernens, um die beste Kombination zu ermitteln. Hier war die Genauigkeit ähnlich wie in unserer Studie (Genauigkeit 71 % für MCI und 89 % für AD). Bei MCI betrug die Genauigkeit 79 % und bei Alzheimer-Patienten 87 %. Darüber hinaus waren König et al. in der Lage, Personen mit MCI von solchen mit Alzheimer mit einer Genauigkeit von 80 % zu unterscheiden, was mit Ergebnissen der hier vorliegenden Studie übereinstimmt (Genauigkeit 76 %) (135). Die bessere Aussagekraft in der Studie von König und Kollegen ist eventuell darauf zurückzuführen, dass andere Testverfahren, wie z. B. der Countdown Test und die Picture description eingesetzt wurden, die für den Grad der Klassifizierung von Bedeutung waren und auch andere Sprachmerkmale zur Klassifikation herangezogen wurden. Die Analyse z. B. von stimmhaften und stimmlosen Segmenten in freier Sprache scheint insgesamt zu einer hohen Klassifikationsgenauigkeit beizutragen.

López-de-Ipina (2018) extrahierten in einer Gruppe von sowohl gesunden Probanden als auch MCI und AD Probanden mithilfe einer automatischen Sprachanalyse verschiedene sprachliche Merkmale aus drei Aufgaben unterschiedlicher sprachlicher Komplexität (animal fluency task, picture description, spontaneous speech; siehe Tabelle 17, S. 68). Deren Klassifikationsgenauigkeit war gut mit den hier vorliegenden Ergebnissen vergleichbar (Genauigkeit MCI 73 % und zwischen 89-95 % für AD). Das beste Gesamtergebnis mit einer Genauigkeit von 95 % erzielte hier eine freie Sprachaufgabe (Spontansprache). Es wird vermutet, dass strukturierte Sprachaufgaben wie z. B. der Boston Naming Test oder auch die verbale Wortflüssigkeit bei Personen mit hohem Bildungsniveau subtile Veränderungen *nicht* erkennen, da der Patient über genügend anderweitige Ressourcen verfügt, dieses Defizit zu kompensieren. Der Anteil gesprochener Wörter oder auch Pausen in freier Konversation könnte möglicherweise empfindlicher in frühen Krankheitsstadien sein (183).

Durch die Kombination verschiedener prosodischer Merkmale im Rahmen eines Demenz-Screenings und freier Sprache konnten **Kato et al.** Probanden mit einer Genauigkeit von 73 % klassifizieren (184). In deren Studie, die sich automatische Sprachanalyse und Maschine

Learning zu Nutzen machte, konnte mit einer Genauigkeit von 78 % zwischen gesunden Kontrollprobanden und Personen mit MCI unterscheiden. In diesem Ansatz von **Toth et al.** wurden ausschließlich freie Sprachaufgaben verwendet. Unter den extrahierten akustischen Parametern (Anzahl & Länge Pausen, Sprech- und Artikulationsraten) waren das Sprechtempo sowie die Anzahl an Pausen die Faktoren, die die signifikantesten Unterschiede erbrachten (185).

Al-Hameed et al. untersuchten 2019 eine Gruppe von verschiedenen neurodegenerativen Erkrankungen, unter anderem MCI und AD, und klassifizierten, anhand von akustischen Merkmalen, mit einer Genauigkeit von 92 % zwischen gesunden und kognitiv beeinträchtigten Probanden. Die Merkmale, die am besten zur Klassifikation beigetragen, waren das Verhältnis von Pausen und Äußerungen der teilnehmenden Personen (132).

Zusammenfassend lässt sich konstatieren, dass die Ergebnisse der hier vorgestellten Studie mit denen anderer Arbeiten gut vergleichbar sind. Der Unterschied, der im Vergleich zu den erwähnten Studien besteht, ist die Nutzung freier Sprache bei der Klassifikation neurodegenerativer Erkrankungen. Wie oben aufgeführt, scheint die Auswertung der Spontansprache die Genauigkeit der Vorhersage leicht zu verbessern. Prinzipiell gibt es zwei Aufgabentypen, die bei der Erforschung der Veränderung in der Sprachproduktion bei Menschen mit Demenz eingesetzt werden: **strukturierte Sprachaufgaben** wie z. B. beim Boston Naming Test oder der verbalen Wortflüssigkeit, bei denen durch einen Stimulus ein Sprachfetzen als Antwort produziert wird. Die Struktur dieser Aufgaben ist sowohl Vor- als auch Nachteil zugleich. Die Leistungen können einerseits besser mit den Ergebnissen anderer Probanden oder auch etablierter Normen verglichen werden, aber die erfassten sprachlichen Fähigkeiten entsprechen andererseits nicht dem alltäglichen Sprachgebrauch. Der zweite Aufgabentyp zur Untersuchung der Veränderung in der Sprachproduktion nutzt die **Spontansprache**, hier werden meist frei gesprochen zusammenhängende Sätze produziert. Provoziert wird die Spontansprache meist durch Aufgaben wie die Bildbeschreibung oder offene Fragen (186). Somit kann die Fähigkeit der teilnehmenden Person in einer realen Kommunikationssituation bewertet werden, was eine große Herausforderung bei neurodegenerativen Störungsbildern darstellt (187). Jedoch ist es schwierig, die Defizite mit anderen Probanden zu vergleichen. In einer Arbeit von Ostrand et al. (2021), die beide Aufgabentypen miteinander kombinierten, hatten die Spontansprachaufgaben eine gute Vorhersagegenauigkeit (42-51%) bezüglich des aktuellen kognitiven Status, aber nur die strukturierten Sprachaufgaben konnten das kognitive Niveau der Nachuntersuchung (29-56% Genauigkeit), die 12 Monate später stattfand, vorhersagen (188). Für zukünftige Studien, die sich mit der Klassifikationsgenauigkeit von automatisierter Spracherkennung und Sprachbiomarkern erkennen, wäre eine Kombination beider Arten von Sprachaufgaben vermutlich sinnvoll.

Um die Ergebnisse der automatischen Sprachanalyse weiter zu verbessern, versuchten andere Studienteams die Analyse mit noch weiteren Methoden zu kombinieren. Fraser und Kollegen (2016) erreichten eine Klassifikationsgenauigkeit von 81 % bei der Unterscheidung gesunder Probanden und Alzheimer-Patienten, indem sie neben den akustischen Merkmalen weitere semantische und syntaktische Merkmale erfassten und in die Analyse mit einbezogen (189). Die Kombination von automatischer Sprachanalyse mit linguistischen Merkmalen, wie sie z. B. von Gosztolya et al. (2019) durchgeführt wurde, verbesserte die Klassifizierung von Kontroll-, MCI- und AD-Patienten von 74-82 % auf 80-86 % (190).

Es lässt sich dennoch festhalten, dass sich auch hier bestätigte, dass die Analyse von Sprachsignalen bei Menschen mit Alzheimer und MCI ein äußerst nützliches Instrument ist, um Veränderungen in der Sprache zu erkennen und sie zur Steigerung der diagnostischen Genauigkeit beim Screening auf Demenz einzusetzen.

Tabelle 17. Vergleich der Studienlage zur Früherkennung neurodegenerativer Erkrankungen durch sprachlicher Biomarker.

	ELEMENT-Studie	König et al. (2015)	López-de-Ipina et al. (2018)	Kato et al. (2018)	Toth et al. (2018)	Al-Hameed et al. (2019)
Sprachliche Aufgaben	Verbale fluency, Verbal learning (CERAD)	Counting backward, sentence repeating, picture description, verbal fluency	Categorial fluency (animal fluency (Kontrollgruppe vs. MCI) picture description and spontaneous speech (Kontrollgruppe vs. AD)	Reply Hasegawa-Demenz-Skala (HDS-R) and spontaneous speech (question-answer-task)	Immediate and delayed recall content of a short film	Conversation (answer of question in examination situation)
Stichprobe	HC (27), MCI (67), AD (44)	HC (15), MCI (23), AD (26)	Drei verschiedene Stichproben: Categorial fluency (187 HC & 38 MCI) Picture description (12 HC & 6 AD) Spontaneous speech (50 HC & 20 AD)	HC (91), MCI (91), AD (91)	HC (36), MCI (48)	HC (15), Neurodegenerative Gruppe (AD (15), aMCI (2), FTD (2), vaskuläre Demenz (1)
Extrahierte Parameter	Semantische und zeitliche Cluster (verbal fluency), primacy & recency Items, serielle Cluster (verbal learning)	Kombination aussagekräftiger Stimmmerkmale aus jeder Aufgabe	Relevante Merkmale aus jeder Aufgabe	Prosodische Merkmale (z. B. Intonation, Sprechrhythmus)	Sprechtempo, Pausen, Länge der Äußerungen	Verschiedene Gruppen von Merkmalen
Vorhersage-Genauigkeit	HC vs. MCI: 71% HC vs. AD: 89% HC vs. MCI/AD: 75% MCI vs. AD: 76%	MCI: 78% AD: 87% MCI vs. AD: 80%	MCI: 73% AD: 89-95% (spontaneous speech)	AD: 74-89% MCI: 70-76%	MCI: 75%	Neurodegenerative Erkrankungen: 81-92%

5.3 Vergleich der Performance vom SB-C mit der konventionellen CERAD Testbatterie

Zuletzt wurde hier der digitale Sprachbiomarker (SB-C) mit den sprachbasierten Subtests der traditionellen Paper-Pencil-Testung des CERADs hinsichtlich dessen Aussagekraft für das Screening einer etwaigen MCI und Alzheimer-Erkrankung verglichen.

In einer Studie, die eine Gruppe von jungen, gesunden Probanden untersuchte, konnten akustische, lexikalische und semantische Sprachmerkmale während der verbalen Produktion extrahiert werden und somit die Aussagekraft des traditionellen Paper-Pencil-Tests erheblich erhöhen (191). Cho und Kollegen haben den zusätzlichen Nutzen einer automatischen Sprachanalyse während einer Wortflüssigkeitsaufgabe bei neurodegenerativen Erkrankungen untersucht. Die automatisierte Sprachanalyse angewendet auf die herkömmliche neuropsychologische Aufgabe konnte auch hier den diagnostischen Wert optimieren, den Krankheitsverlauf überwachen und prognostische Informationen liefern (192). Auch König et al. konnten in ihrer Untersuchung, die die automatische Analyse der semantischen Wortflüssigkeit mit einer manuellen Messung an Probanden mit MCI, Alzheimer und einer gesunden Kontrollgruppe vergleicht, zeigen, dass die automatische Sprachanalyse hoch mit dem traditionellen Verfahren korreliert und äquivalente Ergebnisse liefert (193). Unsere Ergebnisse sprechen dafür, dass der SB-C Score in allen Gruppenvergleichen genauso gut oder teilweise besser zwischen den Gruppen klassifizieren kann im Vergleich zu den verbalen Sprachaufgaben des CERADs (Wortflüssigkeit, verbales Lernen und verbaler Abruf). Beim Vergleich zwischen den gesunden Kontrollprobanden und Probanden mit kognitiver Beeinträchtigung war die Aussagekraft des SB-C gleichwertig mit der des Wortlisten-Lernens und Abrufes.

Zusammenfassend lässt sich feststellen, dass der neu entwickelte SB-C Score nicht nur gut zur Unterscheidung zwischen HC vs. MCI vs. Alzheimer genutzt werden kann, sondern auch hinsichtlich seiner Aussagekraft mit den sprachbasierten Tests des traditionellen Paper-Pencil-Verfahrens (CERAD) vergleichbar ist. Die Ergebnisse unserer Studie könnten neben der Erweiterung der sprachbasierten Aufgaben durch Spontansprachaufgaben auch durch die Auswertung der strukturellen Veränderungen des Gehirns mittels cranialer Bildgebung (MRT), die im Rahmen der Studie erhoben wurden, aber noch nicht ausgewertet sind, verbessert werden. Studien, die als Biomarker z. B. das mit einem MRT gemessene Hippocampus-Volumen (194) oder Amyloid-PET-Tracer (195) verwendeten, hatten so eine Genauigkeit von über 90 % und ermöglichten die Erkennung erster Veränderungen meist Monate oder Jahre vor der Diagnose. Abnormalitäten in strukturellen MRT-Aufnahmen, vor allem Atrophien des medialen Temporallappens und kortikale Ausdünnung in für Alzheimer anfälligen Regionen, sind bereits in der asymptomatischen Phase der Erkrankung nachweisbar und könnten daher auch als zusätzlicher Prädiktor dienen (196).

5.4 Limitationen

Als Limitation dieser Studie ist zu benennen, dass eine recht hohe **Dropout-Rate** von 26 % in der Studiengruppe vorlag, die durch verschiedene Ursachen wie z. B. das defizitäre kognitive Niveau der Teilnehmer, den großen zeitlichen Abstand zwischen den Testzeitpunkt sowie die Corona Pandemie bedingt sein kann. Eine weitere Beeinträchtigung betrifft die **relativ kleine Anzahl an Kontrollprobanden**, welche sowohl durch die Art und Weise der Rekrutierung als auch die Pandemie entstanden sein kann. Zusätzlich ist einschränkend dazu gekommen, dass 10 der 148 **Sprachaufnahmen** in der Gruppe der Studienpatienten wegen schlechter Qualität oder fehlerbehafteter Aufzeichnung nicht mit in die Analyse einfließen konnten. Bei der ELEMENT-Studie handelt es sich außerdem um eine **monozentrische Studie**, eine Ausweitung auf mehrere Zentren wäre sinnvoll, um die Aussagekraft der vorliegenden Ergebnisse zu erhöhen und regionsspezifische Unterschiede zu minimieren. Die Ausweitung der Prüfzentren könnte ebenso dazu beitragen, dass sich die Stichprobe vergrößert und eventuell auch eine **größere Anzahl an Varianten der frontotemporalen Demenz** (z. B. die nicht-flüssige/agrammatische & die semantische Variante der primär progressiven Aphasie) mit in die Analyse einfließen könnten. Das unizentrische Setting mit einer Probandenrekrutierung über eine Gedächtnisambulanz sowie eine fehlende **externe Validierung** könnten zu einem Selektionsbias und somit zu einer Verzerrung der Daten führen. Ein **längeres Follow-up** über mehrere Jahre am selben Patientenkollektiv würde es außerdem ermöglichen, die zeitliche Entwicklung und Veränderung der degenerativen Krankheitsverläufe zu protokollieren und das Potenzial des digitalen Sprachbiomarkers zu überprüfen.

Demgegenüber ist die hochgradig **standardisierte Erhebung der Daten positiv hervorzuheben**, die durch die Nutzung standardisierter neuropsychologischer Verfahren, die neuropsychologisch geschulten Untersucher, den standardisierten Ablauf der Untersuchung sowie die einheitlichen diagnostischen Kriterien (ICD-10), in Hinblick auf die Einordnung der verschiedenen dementiellen Gruppen, gestützt werden. Die Erhebung in einer Gedächtnisambulanz ist ebenso als positiv zu werten, weil die **Mitarbeiter ausreichend Erfahrung im Umgang mit dem Patientenkollektiv** aufweisen und mit den **Erhebungsinstrumenten vertraut** sind. Wichtig ist auch, dass bei der Auswahl der Testpersonen auf das Vorliegen **depressiver Symptomatik** als möglichen Confounder geachtet wurde und Probanden mit einem BDI > 19 (mehr als eine minimale depressive Symptomatik) ausgeschlossen wurden. Eine weitere Stärke ist die **geringe Anzahl an Untersuchern**, die die Daten erfasst, ausgewertet und statistisch analysiert haben.

5.5 Schlussfolgerung und Ausblick

Zusammenfassend belegt die vorliegende Studie, dass die Delta App problemlos funktioniert und im klinischen Alltag erfolgreich eingesetzt werden kann. Durch ihre Nutzung wird das Arbeiten erleichtert, indem sie die Antworten des Probanden aufzeichnet und in Teilen automatisch auswertet, womit die Auswertung der Ergebnisse vereinfacht wird. Zusammen mit dem Team und den ärztlichen Kollegen und Kolleginnen können die Leistungen des Probanden durch die Aufnahmen der App auch nach der Testung beurteilt werden. Dies kann zu einer optimierten Diagnosestellung beitragen. Die erhobenen Testergebnisse waren valide und die Testmodalität (App anstelle Paper-Pencil) hatte allenfalls einen marginalen Einfluss auf die Ergebnisse.

Basierend auf diesen Ergebnissen war es möglich, einen sprachbasierten digitalen Sprachbiomarker zu entwickeln, der eine gut bis sehr gute Performance in der Unterscheidung zwischen klinischen Gruppen (gesund vs. MCI vs. Alzheimer Demenz) bietet. Die Klassifikation der Subgruppen durch den SB-C Score ist mit der Aussagekraft der Subtests des CERADs (Wortlisten Lernen, Wortlisten Abruf) zu vergleichen, teilweise sogar überlegen. Um die Güte des digitalen SB-C Biomarkers noch weiter zu verbessern, wäre eine Integration weiterer sprachbasierter Tests des CERADs z. B. die phonematische Wortflüssigkeit sowie der verbal recall (Wortlisten Abruf), als ein sehr sensibler Prädiktor bei der Identifikation von Alzheimer, zu überlegen. Auch der zusätzliche Einsatz von freien sprachbasierten Aufgaben und dadurch weiteren sprachlichen Parametern (Reaktionszeiten, Pausenlängen, oder ähnliches) wäre denkbar und könnte die Klassifikationsgenauigkeit verbessern. Die Kombination bspw. mit anderen Untersuchungsmodalitäten, wie den im Rahmen dieser Studie durchgeführten Magnetresonanztomographien zur Beurteilung struktureller Veränderungen des Gehirns, steht noch aus und sollte die Vorhersagegüte noch weiter optimieren können.

6. Literaturverzeichnis

1. Langenscheidt Online Wörterbuch: PONS GmbH; [Available from: <https://de.langenscheidt.com/deutsch-latein/>].
2. Information DfMDu. Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme 10. Revision 2020 [Available from: <https://www.dimdi.de/static/de/klassifikationen/icd/icd-10-gm/kode-suche/htmlgm2020/>].
3. Gauthier SR-NPM, J. A. Webster, C. World Alzheimer Report 2021 [
4. Gauthier S, Rosa-Neto P, Morais JA, Webster C. World Alzheimer Report 2021: Journey through the diagnosis of dementia: Alzheimer's Disease International; 2021 [Available from: <https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf>].
5. Foundation AE. Dementia in Europe Yearbook 2019 2019 [Available from: https://www.alzheimer-europe.org/sites/default/files/alzheimer_europe_dementia_in_europe_yearbook_2019.pdf].
6. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, et al. Alzheimer's disease. *Lancet*. 2021;397(10284):1577-90.
7. Wu YT, Beiser AS, Breteler MMB, Fratiglioni L, Helmer C, Hendrie HC, et al. The changing prevalence and incidence of dementia over time - current evidence. *Nat Rev Neurol*. 2017;13(6):327-39.
8. Cao Q, Tan CC, Xu W, Hu H, Cao XP, Dong Q, et al. The Prevalence of Dementia: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 2020;73(3):1157-66.
9. Hasselgren C, Ekbrand H, Halleröd B, Mellqvist Fässberg M, Zettergren A, Johansson L, et al. Sex differences in dementia: on the potentially mediating effects of educational attainment and experiences of psychological distress. *BMC Psychiatry*. 2020;20(1):434.
10. Nebel RA, Aggarwal NT, Barnes LL, Gallagher A, Goldstein JM, Kantarci K, et al. Understanding the impact of sex and gender in Alzheimer's disease: A call to action. *Alzheimers Dement*. 2018;14(9):1171-83.
11. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. *Lancet Neurol*. 2007;6(12):1106-14.
12. Mazure CM, Swendsen J. Sex differences in Alzheimer's disease and other dementias. *Lancet Neurol*. 2016;15(5):451-2.
13. Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of Dementia over Three Decades in the Framingham Heart Study. *N Engl J Med*. 2016;374(6):523-32.
14. Schrijvers EM, Verhaaren BF, Koudstaal PJ, Hofman A, Ikram MA, Breteler MM. Is dementia incidence declining?: Trends in dementia incidence since 1990 in the Rotterdam Study. *Neurology*. 2012;78(19):1456-63.
15. Matthews FE, Stephan BC, Robinson L, Jagger C, Barnes LE, Arthur A, et al. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. *Nat Commun*. 2016;7:11398.
16. Hebert LE, Bienias JL, Aggarwal NT, Wilson RS, Bennett DA, Shah RC, et al. Change in risk of Alzheimer disease over time. *Neurology*. 2010;75(9):786-91.
17. Grasset L, Brayne C, Joly P, Jacqmin-Gadda H, Peres K, Foubert-Samier A, et al. Trends in dementia incidence: Evolution over a 10-year period in France. *Alzheimers Dement*. 2016;12(3):272-80.
18. Langa KM. Is the risk of Alzheimer's disease and dementia declining? *Alzheimers Res Ther*. 2015;7(1):34.
19. Gao S, Burney HN, Callahan CM, Purnell CE, Hendrie HC. Incidence of Dementia and Alzheimer Disease Over Time: A Meta-Analysis. *J Am Geriatr Soc*. 2019;67(7):1361-9.
20. Loy CT, Schofield PR, Turner AM, Kwok JB. Genetics of dementia. *Lancet*. 2014;383(9919):828-40.
21. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. *Lancet*. 2017;390(10113):2673-734.
22. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med*. 2013;29(4):737-52.

23. Wisdom NM, Mignogna J, Collins RL. Variability in Wechsler Adult Intelligence Scale-IV subtest performance across age. *Arch Clin Neuropsychol*. 2012;27(4):389-97.
24. Jorm AF, Jolley D. The incidence of dementia: a meta-analysis. *Neurology*. 1998;51(3):728-33.
25. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology*. 2013;80(19):1778-83.
26. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nature genetics*. 2019;51(3):414-30.
27. van der Lee SJ, Wolters FJ, Ikram MK, Hofman A, Ikram MA, Amin N, et al. The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. *Lancet Neurol*. 2018;17(5):434-44.
28. Genin E, Hannequin D, Wallon D, Slegers K, Hiltunen M, Combarros O, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry*. 2011;16(9):903-7.
29. Lourida I, Hannon E, Littlejohns TJ, Langa KM, Hyppönen E, Kuzma E, et al. Association of Lifestyle and Genetic Risk With Incidence of Dementia. *Jama*. 2019;322(5):430-7.
30. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet*. 2015;385(9984):2255-63.
31. Arenaza-Urquijo EM, Wirth M, Chételat G. Cognitive reserve and lifestyle: moving towards preclinical Alzheimer's disease. *Front Aging Neurosci*. 2015;7:134.
32. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 2020;396(10248):413-46.
33. Torssander J, Erikson R. Stratification and mortality—A comparison of education, class, status, and income. *European Sociological Review*. 2010;26(4):465-74.
34. Lövdén M, Fratiglioni L, Glymour MM, Lindenberger U, Tucker-Drob EM. Education and Cognitive Functioning Across the Life Span. *Psychol Sci Public Interest*. 2020;21(1):6-41.
35. Cha H, Farina MP, Hayward MD. Socioeconomic status across the life course and dementia-status life expectancy among older Americans. *SSM Popul Health*. 2021;15:100921.
36. Institute of Medicine Committee on a National Surveillance System for C, Select Chronic D. The National Academies Collection: Reports funded by National Institutes of Health. A Nationwide Framework for Surveillance of Cardiovascular and Chronic Lung Diseases. Washington (DC): National Academies Press (US) Copyright © 2011, National Academy of Sciences.; 2011.
37. Takeda S, Rakugi H, Morishita R. Roles of vascular risk factors in the pathogenesis of dementia. *Hypertens Res*. 2020;43(3):162-7.
38. Takeda S, Sato N, Ogihara T, Morishita R. The renin-angiotensin system, hypertension and cognitive dysfunction in Alzheimer's disease: new therapeutic potential. *Front Biosci*. 2008;13:2253-65.
39. Walker KA, Power MC, Gottesman RF. Defining the Relationship Between Hypertension, Cognitive Decline, and Dementia: a Review. *Curr Hypertens Rep*. 2017;19(3):24.
40. Hersi M, Irvine B, Gupta P, Gomes J, Birkett N, Krewski D. Risk factors associated with the onset and progression of Alzheimer's disease: A systematic review of the evidence. *Neurotoxicology*. 2017;61:143-87.
41. Evans IE, Martyr A, Collins R, Brayne C, Clare L. Social isolation and cognitive function in later life: A systematic review and meta-analysis. *Journal of Alzheimer's disease*. 2019;70(s1):S119-S44.
42. Kelly M, Duff H, Kelly S, Power J, Brennan S, Lawlor B, et al. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: A systematic review. *Systematic Reviews*, 6 (1), 259. 2017.

43. Alzheimer A. Über eine eigenartige Erkrankung der Hirnrinde. Allgemeine. J Psychiatry and Forensic Medicine-physichish,(Berlin). 1907;64:146-8.
44. Kumar A, Singh A, Ekavali. A review on Alzheimer's disease pathophysiology and its management: an update. Pharmacol Rep. 2015;67(2):195-203.
45. Eratne D, Loi SM, Farrand S, Kelso W, Velakoulis D, Looi JC. Alzheimer's disease: clinical update on epidemiology, pathophysiology and diagnosis. Australas Psychiatry. 2018;26(4):347-57.
46. Baumkötter F, Schmidt N, Vargas C, Schilling S, Weber R, Wagner K, et al. Amyloid precursor protein dimerization and synaptogenic function depend on copper binding to the growth factor-like domain. J Neurosci. 2014;34(33):11159-72.
47. Ghosh AK, Gemma S, Tang J. β -Secretase as a therapeutic target for Alzheimer's disease. Neurotherapeutics. 2008;5(3):399-408.
48. Hur JY. γ -Secretase in Alzheimer's disease. Exp Mol Med. 2022;54(4):433-46.
49. Priller C, Bauer T, Mitteregger G, Krebs B, Kretschmar HA, Herms J. Synapse formation and function is modulated by the amyloid precursor protein. Journal of Neuroscience. 2006;26(27):7212-21.
50. Lane CA, Hardy J, Schott JM. Alzheimer's disease. Eur J Neurol. 2018;25(1):59-70.
51. De-Paula VJ, Radanovic M, Diniz BS, Forlenza OV. Alzheimer's disease. Subcell Biochem. 2012;65:329-52.
52. Rabinovici GD, Karlawish J, Knopman D, Snyder HM, Sperling R, Carrillo MC. Testing and disclosures related to amyloid imaging and Alzheimer's disease: Common questions and fact sheet summary. Alzheimers Dement. 2016;12(4):510-5.
53. 2020 Alzheimer's disease facts and figures. Alzheimers Dement. 2020.
54. Piaceri I, Nacmias B, Sorbi S. Genetics of familial and sporadic Alzheimer's disease. Frontiers in Bioscience-Elite. 2013;5(1):167-77.
55. Long JM, Holtzman DM. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. Cell. 2019;179(2):312-39.
56. Atri A. The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management. Med Clin North Am. 2019;103(2):263-93.
57. Selkoe DJ. Translating cell biology into therapeutic advances in Alzheimer's disease. Nature. 1999;399(6738 Suppl):A23-31.
58. Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg SM, Iadecola C, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2011;42(9):2672-713.
59. Kalra RN. The pathology and pathophysiology of vascular dementia. Neuropharmacology. 2018;134(Pt B):226-39.
60. Skrobot OA, Attems J, Esiri M, Hortobágyi T, Ironside JW, Kalra RN, et al. Vascular cognitive impairment neuropathology guidelines (VCING): the contribution of cerebrovascular pathology to cognitive impairment. Brain. 2016;139(11):2957-69.
61. O'Brien J, Erkinjuntti T, Reisberg B, Roman G, Sawada T, Pantoni L, et al. Vascular cognitive impairment. The Lancet Neurology. 2003;2(2):89-98.
62. Dichgans M, Leys D. Vascular Cognitive Impairment. Circ Res. 2017;120(3):573-91.
63. Hachinski VC, Lassen NA, Marshall J. Multi-infarct dementia: a cause of mental deterioration in the elderly. The Lancet. 1974;304(7874):207-9.
64. Szirmai I, Vastagh I, Szombathelyi É, Kamondi A. Strategic infarcts of the thalamus in vascular dementia. Journal of the neurological sciences. 2002;203:91-7.
65. Kalashnikova L, Gulevskaya T, Kashina E. Disorders of higher mental function due to single infarctions in the thalamus and in the area of the thalamofrontal tracts. Neuroscience and behavioral physiology. 1999;29(4):397-403.
66. Jellinger KA. Morphologic diagnosis of "vascular dementia"—a critical update. Journal of the neurological sciences. 2008;270(1-2):1-12.
67. Wen W, Sachdev PS, Li JJ, Chen X, Anstey KJ. White matter hyperintensities in the forties: their prevalence and topography in an epidemiological sample aged 44–48. Human brain mapping. 2009;30(4):1155-67.

68. De Leeuw F, de Groot JC, Achten E, Oudkerk M, Ramos L, Heijboer R, et al. Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study. The Rotterdam Scan Study. *Journal of Neurology, Neurosurgery & Psychiatry*. 2001;70(1):9-14.
69. Cordonnier C, Al-Shahi Salman R, Wardlaw J. Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting. *Brain*. 2007;130(8):1988-2003.
70. Moulin S, Labreuche J, Bombois S, Rossi C, Boulouis G, Hénon H, et al. Dementia risk after spontaneous intracerebral haemorrhage: a prospective cohort study. *The Lancet Neurology*. 2016;15(8):820-9.
71. Johnston SC, O'Meara ES, Manolio TA, Lefkowitz D, O'Leary DH, Goldstein S, et al. Cognitive impairment and decline are associated with carotid artery disease in patients without clinically evident cerebrovascular disease. *Ann Intern Med*. 2004;140(4):237-47.
72. Balestrini S, Perozzi C, Altamura C, Vernieri F, Luzzi S, Bartolini M, et al. Severe carotid stenosis and impaired cerebral hemodynamics can influence cognitive deterioration. *Neurology*. 2013;80(23):2145-50.
73. Olney NT, Spina S, Miller BL. Frontotemporal Dementia. *Neurol Clin*. 2017;35(2):339-74.
74. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;134(Pt 9):2456-77.
75. Woolley JD, Khan BK, Murthy NK, Miller BL, Rankin KP. The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *J Clin Psychiatry*. 2011;72(2):126-33.
76. Mesulam MM. Primary progressive aphasia--differentiation from Alzheimer's disease. *Annals of neurology*. 1987;22(4):533-4.
77. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;76(11):1006-14.
78. Miller BL, Chang L, Mena I, Boone K, Lesser IM. Progressive right frontotemporal degeneration: clinical, neuropsychological and SPECT characteristics. *Dementia*. 1993;4(3-4):204-13.
79. Seeley WW, Bauer AM, Miller BL, Gorno-Tempini ML, Kramer JH, Weiner M, et al. The natural history of temporal variant frontotemporal dementia. *Neurology*. 2005;64(8):1384-90.
80. Ogar J, Willock S, Baldo J, Wilkins D, Ludy C, Dronkers N. Clinical and anatomical correlates of apraxia of speech. *Brain Lang*. 2006;97(3):343-50.
81. Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, et al. Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol*. 2004;55(3):335-46.
82. Fierini F. Mixed dementia: Neglected clinical entity or nosographic artifice? *J Neurol Sci*. 2020;410:116662.
83. S3-Leitlinien „Demenzen“ 2016: Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN); [Available from: https://www.awmf.org/uploads/tx_szleitlinien/038-013l_S3-Demenzen-2016-07.pdf].
84. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-5)2013.
85. Snowdon DA, Greiner LH, Mortimer JA, Riley KP, Greiner PA, Markesbery WR. Brain infarction and the clinical expression of Alzheimer disease: the Nun Study. *Jama*. 1997;277(10):813-7.
86. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. *The Lancet Neurology*. 2009;8(11):1006-18.
87. Harvey RJ, Skelton-Robinson M, Rossor MN. The prevalence and causes of dementia in people under the age of 65 years. *J Neurol Neurosurg Psychiatry*. 2003;74(9):1206-9.
88. Kelley BJ, Boeve BF, Josephs KA. Young-onset dementia: demographic and etiologic characteristics of 235 patients. *Arch Neurol*. 2008;65(11):1502-8.

89. Chiari A, Vinceti G, Adani G, Tondelli M, Galli C, Fiondella L, et al. Epidemiology of early onset dementia and its clinical presentations in the province of Modena, Italy. *Alzheimers Dement.* 2021;17(1):81-8.
90. Masellis M, Sherborn K, Neto P, Sadovnick DA, Hsiung GY, Black SE, et al. Early-onset dementias: diagnostic and etiological considerations. *Alzheimers Res Ther.* 2013;5(Suppl 1):S7.
91. Draper B, Cations M, White F, Trollor J, Loy C, Brodaty H, et al. Time to diagnosis in young-onset dementia and its determinants: the INSPIRED study. *Int J Geriatr Psychiatry.* 2016;31(11):1217-24.
92. Mendez MF. Early-Onset Alzheimer Disease. *Neurol Clin.* 2017;35(2):263-81.
93. van Vliet D, de Vugt ME, Bakker C, Koopmans RT, Verhey FR. Impact of early onset dementia on caregivers: a review. *Int J Geriatr Psychiatry.* 2010;25(11):1091-100.
94. Mendez MF. The accurate diagnosis of early-onset dementia. *Int J Psychiatry Med.* 2006;36(4):401-12.
95. Bonvicini C, Scassellati C, Benussi L, Di Maria E, Maj C, Ciani M, et al. Next Generation Sequencing Analysis in Early Onset Dementia Patients. *J Alzheimers Dis.* 2019;67(1):243-56.
96. Sakata N, Okumura Y. Job Loss After Diagnosis of Early-Onset Dementia: A Matched Cohort Study. *J Alzheimers Dis.* 2017;60(4):1231-5.
97. Sikes P, Hall M. The impact of parental young onset dementia on children and young people's educational careers. *Br Educ Res J.* 2018;44(4):593-607.
98. Gomperts SN. Lewy Body Dementias: Dementia With Lewy Bodies and Parkinson Disease Dementia. *Continuum (Minneapolis, Minn).* 2016;22(2 Dementia):435-63.
99. Sezgin M, Bilgic B, Tinaz S, Emre M. Parkinson's Disease Dementia and Lewy Body Disease. *Semin Neurol.* 2019;39(2):274-82.
100. McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology.* 2005;65(12):1863-72.
101. Emre M. What causes mental dysfunction in Parkinson's disease? *Mov Disord.* 2003;18 Suppl 6:S63-71.
102. Compta Y, Parkkinen L, O'Sullivan SS, Vandrovcova J, Holton JL, Collins C, et al. Lewy- and Alzheimer-type pathologies in Parkinson's disease dementia: which is more important? *Brain.* 2011;134(Pt 5):1493-505.
103. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of 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):280-92.
104. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimer's & Dementia.* 2016;12(3):292-323.
105. Pike KE, Cavuoto MG, Li L, Wright BJ, Kinsella GJ. Subjective Cognitive Decline: Level of Risk for Future Dementia and Mild Cognitive Impairment, a Meta-Analysis of Longitudinal Studies. *Neuropsychol Rev.* 2021.
106. Petersen RC, Lopez O, Armstrong MJ, Getchius TSD, Ganguli M, Gloss D, et al. Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology.* 2018;90(3):126-35.
107. Liss JL, Seleri Assunção S, Cummings J, Atri A, Geldmacher DS, Candela SF, et al. Practical recommendations for timely, accurate diagnosis of symptomatic Alzheimer's disease (MCI and dementia) in primary care: a review and synthesis. *J Intern Med.* 2021;290(2):310-34.
108. Ganguli M, Snitz B, Saxton J, Chang C, Hall K, Lee C. Outcomes of mild cognitive impairment depend on definition: a population study. *Arch Neurol.*
109. Vos SJ, van Rossum IA, Verhey F, Knol DL, Soinen H, Wahlund LO, et al. Prediction of Alzheimer disease in subjects with amnesic and nonamnesic MCI. *Neurology.* 2013;80(12):1124-32.

110. Brodaty H, Seeher K, Gibson L. Dementia time to death: a systematic literature review on survival time and years of life lost in people with dementia. *Int Psychogeriatr*. 2012;24(7):1034-45.
111. Wimo A, Prince MJ. World Alzheimer Report 2010: the global economic impact of dementia: Alzheimer's Disease International; 2010.
112. Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M. World Alzheimer Report 2015-The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. 2015.
113. Schubert CC, Boustani M, Callahan CM, Perkins AJ, Carney CP, Fox C, et al. Comorbidity profile of dementia patients in primary care: are they sicker? *J Am Geriatr Soc*. 2006;54(1):104-9.
114. Whitmer RA, Karter AJ, Yaffe K, Quesenberry CP, Jr., Selby JV. Hypoglycemic episodes and risk of dementia in older patients with type 2 diabetes mellitus. *Jama*. 2009;301(15):1565-72.
115. Chen Y, Wilson L, Kornak J, Dudley RA, Merrilees J, Bonasera SJ, et al. The costs of dementia subtypes to California Medicare fee-for-service, 2015. *Alzheimers Dement*. 2019;15(7):899-906.
116. Sinclair AJ, Girling AJ, Bayer AJ. Cognitive dysfunction in older subjects with diabetes mellitus: impact on diabetes self-management and use of care services. All Wales Research into Elderly (AWARE) Study. *Diabetes Res Clin Pract*. 2000;50(3):203-12.
117. Tilly J, Reed P, Gould E, Fok A. End-of-life care practice recommendations for assisted living residences and nursing homes serving individuals with dementia. *Alzheimer's Care Today*. 2008;9(2):113-9.
118. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. *N Engl J Med*. 2013;368(14):1326-34.
119. Friedman EM, Shih RA, Langa KM, Hurd MD. US Prevalence And Predictors Of Informal Caregiving For Dementia. *Health Aff (Millwood)*. 2015;34(10):1637-41.
120. Schulz R, Beach SR. Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. *Jama*. 1999;282(23):2215-9.
121. Goren A, Montgomery W, Kahle-Wroblewski K, Nakamura T, Ueda K. Impact of caring for persons with Alzheimer's disease or dementia on caregivers' health outcomes: findings from a community based survey in Japan. *BMC Geriatr*. 2016;16:122.
122. Pinquart M, Sörensen S. Associations of stressors and uplifts of caregiving with caregiver burden and depressive mood: a meta-analysis. *J Gerontol B Psychol Sci Soc Sci*. 2003;58(2):P112-28.
123. Robinson L, Tang E, Taylor J-P. Dementia: timely diagnosis and early intervention. *Bmj*. 2015;350.
124. Michalowsky B, Flessa S, Hertel J, Goetz O, Hoffmann W, Teipel S, et al. Cost of diagnosing dementia in a German memory clinic. *Alzheimers Res Ther*. 2017;9(1):65.
125. Tan CC, Yu JT, Wang HF, Tan MS, Meng XF, Wang C, et al. Efficacy and safety of donepezil, galantamine, rivastigmine, and memantine for the treatment of Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimers Dis*. 2014;41(2):615-31.
126. Rountree SD, Chan W, Pavlik VN, Darby EJ, Siddiqui S, Doody RS. Persistent treatment with cholinesterase inhibitors and/or memantine slows clinical progression of Alzheimer disease. *Alzheimers Res Ther*. 2009;1(2):1-7.
127. Leung CW, Lam TP, Wong KW, Chao VKD. Early detection of dementia: The knowledge and attitudes of primary care physicians in Hong Kong. *Dementia (London)*. 2020;19(3):830-46.
128. Iliffe S, Manthorpe J, Eden A. Sooner or later? Issues in the early diagnosis of dementia in general practice: a qualitative study. *Fam Pract*. 2003;20(4):376-81.
129. Maki Y, Yamaguchi H. Early detection of dementia in the community under a community-based integrated care system. *Geriatr Gerontol Int*. 2014;14 Suppl 2:2-10.
130. The Need for Early Detection and Treatment in Alzheimer's Disease. *EBioMedicine*. 2016;9:1-2.
131. Counts SE, Ikonomic MD, Mercado N, Vega IE, Mufson EJ. Biomarkers for the Early Detection and Progression of Alzheimer's Disease. *Neurotherapeutics*. 2017;14(1):35-53.

132. Al-Hameed S, Benaissa M, Christensen H, Mirheidari B, Blackburn D, Reuber M. A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. *PLoS One*. 2019;14(5):e0217388.
133. Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, et al. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimers Dement*. 2015;11(5):561-78.
134. López-de-Ipiña K, Alonso JB, Travieso CM, Solé-Casals J, Egiraun H, Faundez-Zanuy M, et al. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors (Basel)*. 2013;13(5):6730-45.
135. König A, Satt A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimers Dement (Amst)*. 2015;1(1):112-24.
136. Ostrand R, Gunstad J. Using automatic assessment of speech production to predict current and future cognitive function in older adults. *Journal of Geriatric Psychiatry and Neurology*. 2021;34(5):357-69.
137. Heyman A, Fillenbaum GG. Overview: clinical sites, case material, and special studies. *Neurology*. 1997;49(3 Suppl 3):S2-6.
138. Welsh-Bohmer KA, Mohs RC. Neuropsychological assessment of Alzheimer's disease. *Neurology*. 1997;49(3 Suppl 3):S11-S3.
139. Schmid NS, Ehrensperger MM, Berres M, Beck IR, Monsch AU. The Extension of the German CERAD Neuropsychological Assessment Battery with Tests Assessing Subcortical, Executive and Frontal Functions Improves Accuracy in Dementia Diagnosis. *Dement Geriatr Cogn Dis Extra*. 2014;4(2):322-34.
140. Rosen WG. Verbal fluency in aging and dementia. *Journal of clinical and experimental neuropsychology*. 1980;2(2):135-46.
141. Lezak MD, Howieson DB, Loring DW, Fischer JS. *Neuropsychological assessment*: Oxford University Press, USA; 2004.
142. Eloi JM, Lee J, Pollock EN, Tayim FM, Holcomb MJ, Hirst RB, et al. Boston Naming Test: Lose the Noose. *Arch Clin Neuropsychol*. 2021.
143. Kaplan E, Goodglass H, Weintraub S. Boston naming test. 2001.
144. Cockrell JR, Folstein MF. Mini-mental state examination. *Principles and practice of geriatric psychiatry*. 2002:140-1.
145. Beck IR, Gagneux-Zurbriggen A, Berres M, Taylor KI, Monsch AU. Comparison of verbal episodic memory measures: consortium to establish a registry for Alzheimer's disease-Neuropsychological Assessment Battery (CERAD-NAB) versus California Verbal Learning Test (CVLT). *Arch Clin Neuropsychol*. 2012;27(5):510-9.
146. Saari T, Hallikainen I, Hänninen T, Rätty H, Koivisto A. Relationships between Cognition and Activities of Daily Living in Alzheimer's Disease During a 5-Year Follow-Up: ALSOVA Study. *J Alzheimers Dis*. 2018;64(1):269-79.
147. Mohs RC, Kim Y, Johns CA, Dunn DD, Davis KL. Assessing changes in Alzheimer's disease: Memory and language. 1986.
148. Welsh KA, Butters N, Mohs RC, Beekly D, Edland S, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology*. 1994;44(4):609-.
149. Thalman B, Monsch AU, Schneitter M, Bernasconi F, Aebi C, Camachova-Davet Z, et al. The CERAD neuropsychological assessment battery (CERAD-NAB)—A minimal data set as a common tool for German-speaking Europe. *Neurobiology of Aging*. 2000(21):30.
150. Bortz J, Schuster C. *Statistik für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*: Springer-Verlag; 2011.
151. Fisseni H-J. *Lehrbuch der psychologischen Diagnostik*. Göttingen: Verlag für Psychologie; 1990.
152. Lezak MD. *Neuropsychological Assessment*. New York: Oxford University Press; 1995.
153. Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M. Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. *Archives of general psychiatry*. 2001;58(9):853-8.

154. Lezak MD. *Neuropsychological Assessment* (5th ed.). New York: Oxford University Press; 2012.
155. Biracyaza E, Habimana S, Rusengamihigo D. Psychometric Properties of the Beck Depression Inventory (BDI-II) in Cancer Patients: Cancer Patients from Butaro Ambulatory Cancer Center, Rwanda. *Psychol Res Behav Manag.* 2021;14:665-74.
156. Wintjen L, Petermann F. Beck-depressions-Inventar revision (BDI-II). *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie.* 2010;58(3):243-5.
157. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj Digital Medicine.* 2020;3(1):55.
158. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations. *Digit Biomark.* 2020;4(3):99-108.
159. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj digital Medicine.* 2020;3(1):1-15.
160. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-98.
161. Dorsch F. *Dorsch Psychologisches Wörterbuch*: Huber; 2009.
162. Provencher V, Mortenson WB, Tanguay-Garneau L, Bélanger K, Dagenais M. Challenges and strategies pertaining to recruitment and retention of frail elderly in research studies: a systematic review. *Archives of Gerontology and Geriatrics.* 2014;59(1):18-24.
163. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol.* 2005;58(1):13-9.
164. Karrasch M, Sinervä E, Grönholm P, Rinne J, Laine M. CERAD test performances in amnesic mild cognitive impairment and Alzheimer's disease. *Acta Neurol Scand.* 2005;111(3):172-9.
165. Ruggeri K, Maguire Á, Andrews JL, Martin E, Menon S. Are We There Yet? Exploring the Impact of Translating Cognitive Tests for Dementia Using Mobile Technology in an Aging Population. *Front Aging Neurosci.* 2016;8:21.
166. Morrison RL, Pei H, Novak G, Kaufer DI, Welsh-Bohmer KA, Ruhmel S, et al. A computerized, self-administered test of verbal episodic memory in elderly patients with mild cognitive impairment and healthy participants: A randomized, crossover, validation study. *Alzheimers Dement (Amst).* 2018;10:647-56.
167. Chan JYC, Kwong JSW, Wong A, Kwok TCY, Tsoi KKF. Comparison of Computerized and Paper-and-Pencil Memory Tests in Detection of Mild Cognitive Impairment and Dementia: A Systematic Review and Meta-analysis of Diagnostic Studies. *J Am Med Dir Assoc.* 2018;19(9):748-56.e5.
168. Scharfen J, Jansen K, Holling H. Retest effects in working memory capacity tests: A meta-analysis. *Psychon Bull Rev.* 2018;25(6):2175-99.
169. Benedict RH, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test–Revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist.* 1998;12(1):43-55.
170. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist.* 2012;26(4):543-70.
171. Heilbronner RL, Sweet JJ, Attix DK, Krull KR, Henry GK, Hart RP. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *The Clinical Neuropsychologist.* 2010;24(8):1267-78.
172. Meeks TW, Vahia IV, Lavretsky H, Kulkarni G, Jeste DV. A tune in "a minor" can "b major": a review of epidemiology, illness course, and public health implications of subthreshold depression in older adults. *J Affect Disord.* 2011;129(1-3):126-42.

173. Bennett S, Thomas AJ. Depression and dementia: cause, consequence or coincidence? *Maturitas*. 2014;79(2):184-90.
174. Fillenbaum GG, Peterson B, Welsh-Bohmer KA, Kukull WA, Heyman A. Progression of Alzheimer's disease in black and white patients: the CERAD experience, part XVI. Consortium to Establish a Registry for Alzheimer's Disease. *Neurology*. 1998;51(1):154-8.
175. Chin J, Kim DE, Lee H, Yun J, Lee BH, Park J, et al. A Validation Study of the Inbrain CST: a Tablet Computer-based Cognitive Screening Test for Elderly People with Cognitive Impairment. *J Korean Med Sci*. 2020;35(34):e292.
176. Canini M, Battista P, Della Rosa PA, Catricalà E, Salvatore C, Gilardi MC, et al. Computerized neuropsychological assessment in aging: testing efficacy and clinical ecology of different interfaces. *Computational and mathematical methods in medicine*. 2014;2014.
177. Beishuizen CRL, Coley N, Moll van Charante EP, van Gool WA, Richard E, Andrieu S. Determinants of Dropout and Nonadherence in a Dementia Prevention Randomized Controlled Trial: The Prevention of Dementia by Intensive Vascular Care Trial. *J Am Geriatr Soc*. 2017;65(7):1505-13.
178. Handels R, Jönsson L, Garcia-Ptacek S, Eriksdotter M, Wimo A. Controlling for selective dropout in longitudinal dementia data: application to the SveDem registry. *Alzheimer's & Dementia*. 2020;16(5):789-96.
179. Evans N, Collier L. An exploration of the experience of using calendar reminders for people with dementia and family carers. *Dementia (London)*. 2019;18(5):1912-33.
180. Persson AC, Möller MC, Dahlberg L, Löfgren M, Janeslätt G. Assessing time processing ability and daily time management in persons with dementia: Psychometric properties of three instruments. *Australian Occupational Therapy Journal*. 2022.
181. COVID M. and clinical trials: the Medidata perspective. 2020.
182. Shiely F, Foley J, Stone A, Cobbe E, Browne S, Murphy E, et al. Managing clinical trials during COVID-19: experience from a clinical research facility. *Trials*. 2021;22(1):1-7.
183. Lopez-de-Ipina K, Martinez-de-Lizarduy U, Calvo PM, Mekyska J, Beitia B, Barroso N, et al. Advances on Automatic Speech Analysis for Early Detection of Alzheimer Disease: A Non-linear Multi-task Approach. *Curr Alzheimer Res*. 2018;15(2):139-48.
184. Kato S, Homma A, Sakuma T, Nakamura M. Detection of mild Alzheimer's disease and mild cognitive impairment from elderly speech: Binary discrimination using logistic regression. *Annu Int Conf IEEE Eng Med Biol Soc*. 2015;2015:5569-72.
185. Tóth L, Hoffmann I, Gosztolya G, Vincze V, Szatlóczki G, Bánréti Z, et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. *Current Alzheimer Research*. 2018;15(2):130-8.
186. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJG. Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *Front Psychol*. 2021;12:620251.
187. Murray J, Schneider J, Banerjee S, Mann A. EURO CARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: II--A qualitative analysis of the experience of caregiving. *Int J Geriatr Psychiatry*. 1999;14(8):662-7.
188. Ostrand R, Gunstad J. Using Automatic Assessment of Speech Production to Predict Current and Future Cognitive Function in Older Adults. *J Geriatr Psychiatry Neurol*. 2020;891988720933358.
189. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease*. 2016;49(2):407-22.
190. Gosztolya G, Vincze V, Tóth L, Pákási M, Kálmán J, Hoffmann I. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech & Language*. 2019;53:181-97.
191. Gagliardi G, Kokkinakis D, Duñabeitia JA. Editorial: Digital Linguistic Biomarkers: Beyond Paper and Pencil Tests. *Front Psychol*. 2021;12:752238.
192. Cho S, Nevler N, Parjane N, Cieri C, Liberman M, Grossman M, et al. Automated Analysis of Digitized Letter Fluency Data. *Front Psychol*. 2021;12:654214.
193. König A, Linz N, Troger J, Wolters M, Alexandersson J, Robert P. Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task. *Dement Geriatr Cogn Disord*. 2018;45(3-4):198-209.

194. Chupin M, Gérardin E, Cuingnet R, Boutet C, Lemieux L, Lehericy S, et al. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus*. 2009;19(6):579-87.
195. Morris E, Chalkidou A, Hammers A, Peacock J, Summers J, Keevil S. Diagnostic accuracy of 18F amyloid PET tracers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *European journal of nuclear medicine and molecular imaging*. 2016;43(2):374-85.
196. Tan CC, Yu JT, Tan L. Biomarkers for preclinical Alzheimer's disease. *J Alzheimers Dis*. 2014;42(4):1051-69.

7. Publikation

Die Ergebnisse der vorliegenden Arbeit wurden im Rahmen folgendem **Kongress** präsentiert:

- Vortrag auf dem Kongress der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN): „Früherkennung kognitiver Störungen mittels Sprachanalyse“ (2020)

Publikation:

- Zeuschner H, Mallick E, Tröger J, Linz N, Riemenschneider M. “Screening for different stages of Alzheimer using digital speech biomarkers for Cognition and machine learning based on digital speech-based CERAD subtests”. *Submitted*.

8. Danksagung

Herrn Prof. Dr. Matthias Riemenschneider (Direktor der Klinik für Psychiatrie und Psychotherapie, Universitätsklinikum des Saarlandes) danke ich für die freundliche Überlassung dieses spannenden Promotionsthemas und für seine stete wissenschaftliche Unterstützung sowie die konstruktive Diskussion während der Bearbeitungsphase meiner Dissertation.

Mein besonderer Dank gilt auch Herrn Nicklas Linz und Herrn Dr. Johannes Tröger von ki:elements (Spin-off des DFKI) für die Bereitstellung der Delta App, die Entwicklung des digitalen Sprachbiomarkers sowie die freundliche Unterstützung und Hilfsbereitschaft während der gesamten Studienlaufzeit.

Gerne möchte ich mich beim gesamten Studienteam der Klinik für Psychiatrie und Psychotherapie, insbesondere bei Frau Dr. Geiger und Frau Huber für die exzellente Unterstützung bei der Organisation der Studie, bedanken. Weiter gilt mein besonderer Dank den Psychologen in Ausbildung und meinen Kollegen, die mir bei der Rekrutierung von Studienpatienten und somit der Erhebung der Daten im Rahmen der Gedächtnisambulanz geholfen haben.

Die statistische Auswertung der Daten unterstützte Herr Prof. S. Wagenpfeil vom Institut für Epidemiologie und Biometrie des Universitätsklinikums des Saarlandes.

Mein besonderer Dank gilt meinem Ehemann für seine Geduld, sein Vorbild und seine jederzeit liebevolle Unterstützung, die motivierenden Worte während der Datenerhebung und Verfassung der Dissertation.

Zum Schluss möchte ich meinen Dank an meine Familie, insbesondere an meine Eltern, richten, für ihre außergewöhnliche und stets liebevolle Unterstützung in jeder Lebenslage, denn ohne ihre Hilfe hätte ich das Bisherige nicht erreicht.

9. Lebenslauf

Aus datenschutzrechtlichen Gründen wird der Lebenslauf in der elektronischen Fassung der Dissertation nicht veröffentlicht.

10. Eidesstattliche Erklärung

Erklärung gemäß § 7 Absatz 1 Nr. 2 – Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

1. Prof. Dr. med. M. Riemenschneider: Betreuung der Arbeit

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberaterinnen/Promotionsberater oder anderer Personen) in Anspruch genommen. Außer den Angegebenen hat niemand von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder in ähnlicher Form in einem anderen Verfahren zur Erlangung des Doktorgrades einer anderen Prüfungsbehörde vorgelegt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die Wahrheit gesagt und nichts verschwiegen habe.

Die Bedeutung der eidesstattlichen Erklärung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Erklärung sind mir bekannt.

A handwritten signature in blue ink, appearing to read 'A. Jowes', with a long horizontal line extending to the right.

Saarlouis, 22.02.2023

11. Appendix

11.1 Patienteneinwilligung ELEMENT-Studie

 <p>Forschungsverbund</p>	 <p>Universitätsklinikum des Saarlandes Klinik für Psychiatrie & Psychotherapie</p>	 <p>Saarbrücken</p>
<p>Einwilligung in die Teilnahme an der Studie „Früherkennung kognitiver Störungen mittels Sprachanalyse“</p> <p>Liebe Patientin, lieber Patient, liebe Angehörige, wir sind stolz, die Möglichkeit zu haben, unsere Diagnostik und Therapie durch Forschungsarbeit zu begleiten und zu verbessern.</p> <p>Die vorliegende Studie entsteht in Zusammenarbeit zwischen dem Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Saarbrücken und der Klinik für Psychiatrie und Psychotherapie am Universitätsklinikum des Saarlandes. Unser Interesse liegt am Verlauf unterschiedlicher Hirnleistungen und in der Früherkennung neurologischer Erkrankungsbildern. Ein Schwerpunkt liegt dabei auf der Erfassung spezieller sprachlicher Leistungen, so dass während der Untersuchungen die Sprache aufgezeichnet wird. Unterstützt wird eine nicht-invasive Abbildung der Struktur und der Funktion des Gehirns mittels MRT durchgeführt. Details zu diesem Verfahren können Sie dem beigefügten Informationsblatt entnehmen.</p> <p>Wir suchen nach Möglichkeiten zur besseren Erkennung von Abbauprozessen im Gehirn im Zusammenhang mit sprachlichen Mustern. Diese Querschnittstudie beinhaltet im Fall einer Wiederaufnahme auch die Möglichkeiten einer Verlaufuntersuchung.</p> <p>Für die Teilnahme an der Studie werden bis zu zwei Termine im Bereich Neuropsychologie mehr benötigt, als sich üblicher Weise aus dem Therapiekonzept der Klinik ergibt. In der Studie werden unterschiedliche Testverfahren eingesetzt. Dies sind ausschließlich Verhaltensuntersuchungen und keine so genannten invasiven Verfahren. Alle Daten werden anonymisiert gespeichert und unabhängig vom einzelnen Patienten ausgewertet. Alle ansonsten für das Krankungsbild unserer Patienten erforderlichen Maßnahmen werden wie üblich, unabhängig von der Studie durchgeführt.</p> <p>Ihre Unterschrift benötigen wir aufgrund datenschutzrechtlicher Bestimmungen. Sollten Sie nicht in die Teilnahme einwilligen, so hat dies keine Nachteile für Ihre Versorgung hier im Haus.</p> <p>Wir möchten Sie aber dennoch herzlich bitten, sich an der Studie zu beteiligen. Für mögliche Fragen wenden Sie sich bitte an die Kollegin oder den Kollegen aus dem Bereich Neuropsychologie.</p> <p style="text-align: center;">Vielen Dank für Ihre Mitarbeit. Mit besten Grüßen</p> <div style="display: flex; justify-content: space-between; margin-top: 20px;"> <div style="text-align: center;"> <p>Prof. Dr. M. Riemenschneider Direktor, Klinik für Psychiatrie und Psychotherapie, UKS</p> </div> <div style="text-align: center;"> <p>Dr. Ing. J. Alexandersson DFKI-Intelligente Benutzerschnittstellen</p> </div> </div> <p style="font-size: small; margin-top: 20px;">Studienverantwortliche: Dr. Ing. Alexandersson, Prof. Dr. Riemenschneider</p>		
<p>Name der Patientin / des Patienten <small>(bitte in Druckbuchstaben)</small></p> <p>Ich <small>(bitte in Druckbuchstaben)</small> stimme in meiner Funktion als Betreuer / Bevollmächtigter der Teilnahme an der Studie ELEMENT zu.</p> <p>Datum: _____</p>	<p>Unterschrift: _____</p>	
<p>Studienverantwortliche: Dr. Ing. Alexandersson, Prof. Dr. Riemenschneider</p>		

11.2 Flyer ELEMENT Studie



Lageplan
der Klinik

Anmeldung zur ELEMENT-Studie

Bitte vereinbaren Sie bei Interesse telefonisch einen Termin.

Kontaktadresse

Klinik für Psychiatrie und Psychotherapie
Direktor:
Prof. Dr. med. M. Riemenschneider
Universitätsklinikum des Saarlandes
Gebäude 90
Kirrberger Straße 100
66421 Homburg/Saar

Studienbüro:

Tel.: 06841/16-24430
E-Mail: studien.psychiatrie@uks.eu

Gedächtnisambulanz:

Tel.: 06841/16-24210
E-Mail: psychiatrie.ambulanz@uks.eu

Version 1, Stand: 01/2020



Universitätsklinikum des Saarlandes und Medizinische Fakultät der Universität des Saarlandes

Klinik für Psychiatrie und Psychotherapie
Direktor: Prof. Dr. med. Matthias Riemenschneider

„ELEMENT-Studie“

Früherkennung
kognitiver Störungen mittels
Sprachanalyse



Studieninformation ELEMENT-Studie

Sehr geehrte Interessenten,

wir freuen uns, Ihnen die Studie „ELEMENT“ (Früherkennung kognitiver Störungen mittels Sprachanalyse) nachfolgend näher vorstellen zu dürfen.

Die Studie zielt darauf ab, die Möglichkeiten zur Früherkennung neurodegenerativer Erkrankungen (z.B. Demenzen) zu verbessern.

Das bedeutet, dass diagnostische Verfahren mit einer intelligenten Sprachsoftware danach überprüft werden, ob sie die Möglichkeiten der Früherkennung dementieller Erkrankungen im Vergleich zu üblichen standardisierten Testungen verbessern.

Die Defizite in der Hirnleistungsfähigkeit bei Patienten mit unterschiedlichen kognitiven Einschränkungen bzw. Stadien der Alzheimer-Demenz sollen mit Hilfe einer sprachbasierten kognitiven Testung geprüft werden, um dementielle Entwicklungen früh zu erkennen und behandeln zu können.

Ablauf der Studie

Zu Beginn der Studie erfolgt für die Basisuntersuchung eine Anamnese sowie eine körperliche Untersuchung zur Bestätigung der kognitiven Einschränkungen.

Alle Teilnehmer erhalten zwei kognitive Testungen zu Beginn und zum Ende der Studie sowie ein MRT und ein Abschlussgespräch mit Befundmitteilung.

Während die erste Gruppe von Patienten erst ein standardisiertes kognitives Testverfahren erhält, erfolgt bei der zweiten Gruppe eine sprachbasierte Testung gleichen Inhalts. Nach drei Monaten erfolgt die zweite Testung; die erste Gruppe wird nun mit einer sprachbasierten Testung untersucht und die zweite Gruppe erhält die standardisierte kognitive Testung.

Alle sonstigen für die eingehende Abklärung Ihres Krankheitsbilds erforderlichen Maßnahmen werden unabhängig von der Studie durchgeführt.

Die erhobenen Daten dienen ausschließlich Forschungszwecken.

Alle Daten werden pseudonymisiert gespeichert und ausgewertet.

Sie können jederzeit ohne Angabe von Gründen und ohne Nachteil für Sie von der Studie zurücktreten.

Wer kann an der Studie teilnehmen?

Sie können an dieser Studie teilnehmen, wenn:

- Sie zwischen 50 und 90 Jahre alt sind,
- Ihre Gedächtnis- und Denkfähigkeiten beeinträchtigt sind,
- Sie mobil und gehfähig sind und über ausreichendes Seh- und Hörvermögen verfügen,
- Sie nicht schwer körperlich oder neurologisch erkrankt sind.

Wenn Sie mehr über die Studie und darüber erfahren möchten, was eine Teilnahme mit sich bringt, setzen Sie sich gerne mit uns in Verbindung.

Studienbüro:
Tel.: 06841/16-24430
E-Mail: studien.psychiatrie@uks.eu

Gedächtnisambulanz:
Tel.: 06841/16-24210
E-Mail: psychiatrie.ambulanz@uks.eu

Weitere Informationen finden Sie auf unserer Homepage unter

www.uniklinikum-saarland.de/psychiatrie

11.3 Studieninformationen auf Internetseite des UKS



UKS

Universitätsklinikum des Saarlandes und Medizinische Fakultät der Universität des Saarlandes

Klinik für Psychiatrie und Psychotherapie

Leitung: Prof. Dr. med. Matthias Riemenschneider

Notfall-Informationen +

Blutspende am UKS

Suche

Deutsch Français English

Aktuelles
Klinische Zentren und Einrichtungen
Krankenversorgung
Forschung
Lehre
Wir über uns
Service
Kontakt

Sie befinden sich hier: >> [Startseite](#) >> [Klinische Zentren und Einrichtungen](#) >> [Klinische Zentren und Einrichtungen am UKS](#) >> [Neurologie und Psychiatrie](#) >> [Psychiatrie](#) >> [aktuelle Studien](#)
[Teilnehmer gesucht](#) >> [ELEMENT-Studie Früherkennung kognitiver Störungen mittels Sprachanalyse](#)

aktuelle Studien
 Teilnehmer gesucht

Studieninformationen ELEMENT-Studie

Sehr geehrte Interessenten,

wir freuen uns, Ihnen die Studie „ELEMENT“ (Früherkennung kognitiver Störungen mittels Sprachanalyse) nachfolgend näher vorstellen zu dürfen.

Die Studie zielt darauf ab, die Möglichkeiten zur Früherkennung neurodegenerativer Erkrankungen (z.B. Demenzen) zu verbessern. Das bedeutet, dass diagnostische Verfahren mit einer intelligenten Sprachsoftware danach überprüft werden, ob sie die Möglichkeiten der Früherkennung dementieller Erkrankungen im Vergleich zu üblichen standardisierten Testungen verbessern.

Die Defizite in der Hirnleistungsfähigkeit bei Patienten mit unterschiedlichen kognitiven Einschränkungen bzw. Stadien der Alzheimer-Demenz sollen mit Hilfe einer sprachbasierten kognitiven Testung geprüft werden, um dementielle Entwicklungen früh zu erkennen und behandeln zu können.

Ablauf der Studie

Zu Beginn der Studie erfolgen für die Basisuntersuchung eine Anamnese sowie eine körperliche Untersuchung zur Bestätigung der kognitiven Einschränkungen.

Alle Teilnehmer erhalten zwei kognitive Testungen zu Beginn und zum Ende der Studie sowie ein MRT und ein Abschlussgespräch mit Befundmitteilung.

Während die erste Gruppe von Patienten erst ein standardisiertes kognitives Testverfahren erhält, erfolgt bei der zweiten Gruppe eine sprachbasierte Testung gleichen Inhalts. Nach drei Monaten erfolgt die zweite Testung: die erste Gruppe wird nun mit einer sprachbasierten Testung untersucht und die zweite Gruppe erhält die standardisierte kognitive Testung.

Alle sonstigen für die eingehende Abklärung Ihres Krankheitsbilds erforderlichen Maßnahmen werden unabhängig von der Studie durchgeführt.

Die erhobenen Daten dienen ausschließlich Forschungszwecken.
Alle Daten werden pseudonymisiert gespeichert und ausgewertet.

Sie können jederzeit ohne Angabe von Gründen und ohne Nachteil für Sie von der Studie zurücktreten.

Wer kann an der Studie teilnehmen?

Sie können an dieser Studie teilnehmen, wenn:

- Sie zwischen 50 und 90 Jahre alt sind.
- Ihre Gedächtnis- und Denkfähigkeiten beeinträchtigt sind.
- Sie mobil und gehfähig sind und über ausreichendes Seh- und Hörvermögen verfügen.
- keine wesentlichen körperlichen Grunderkrankungen (z. B. fortgeschrittene Tumorerkrankung, instabile Herzerkrankung), keine wesentlichen neurologischen Erkrankungen (Z. n. schwerem Schädel-Hirn-Trauma, Schlaganfall etc.) oder relevante psychiatrische Begleiterkrankungen (schwere depressive Episode, Schizophrenie) vorliegen.

Wenn Sie mehr über die Studie und darüber erfahren möchten, was eine Teilnahme mit sich bringt, setzen Sie sich gerne mit uns in Verbindung.

Studienbüro:
 Tel.: 06841/16-24430
 E-Mail: studien.psychiatrie@uks.eu
Gedächtnisambulanz:
 Tel.: 06841/16-24210
 E-Mail: psychiatrie.ambulanz@uks.eu

Diesen Text als [pdf](#) downloaden.



Klinik für Psychiatrie und Psychotherapie
 Universitätsklinikum des Saarlandes
Direktor:
 Prof. Dr. med. M. Riemenschneider
 Kirnberger Straße 100
 D-66421 Homburg/Saar

Direktionssekretariat
 Frau I. Freistadt
 Tel.: 06841/16-24202
 Fax: 06841/16-24270
sekretariat.psychiatrie@tj.uks.eu

Ambulanz
 Tel.: 06841/16-24210

Privatambulanz
 Tel.: 06841/16-24202

Für Notfälle außerhalb der Sprechstunde
 Tel.: 06841/16-24100

Anmeldung stationäre Patientenaufnahme
 Tel.: 06841/16-24100

[Startseite der Klinik](#)



Uni Saarland Wissenwertes von A-Z Impressum Datenschutz





Seite am 08.04.2022 besucht.