MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Subjective memory complaints versus caregivers'-reported memory deficits and their predictive value in conversion to dementia“

verfasst von / submitted by

Natalia Kusto, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie UG2002
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Claus Lamm

Acknowledgement

I would like to thank first of all Ass.-Prof. Priv.-Doz. Mag. Dr. Johann Lehrner for his supervision and constructive feedback on this master's thesis that was a result of a very educational internship in AKH. I would also like to express my gratitude to Univ.-Prof. Dr. Claus Lamm for his official supervision, support and feedback on my thesis.

Furthermore, I would like to thank my parents and grandparents for their constant support and faith in me. Big thank you goes also to my chosen family and friends, especially Raffael and Olcia for being there for me and listening to all of the updates. I thank Maria Chiara for many Skype calls, that made studying more productive and enjoyable. Katalin, thank you for pushing.

Table of Contents

1. Introduction

According to a statistic conducted by World Health Organization, around 50 million people worldwide are diagnosed with dementia, and each year there are almost 10 million newly diagnosed cases (“Dementia,” 2019). In Austria, nearly 100 thousand people live with dementia (Wancata et al., 2001) and according to estimates the number will more than double until year 2050.

The World Alzheimer report 2011 states that the majority of the dementia cases have not been formally diagnosed and therefore do not receive any support in form of care or treatment (Prince et al., 2018). As the illness progresses, the cognitive impairment reduces the ability to cope with more complex daily tasks. In the later stages of dementia those difficulties increase to the extent, that external support of a caregiver is necessary (Prince et al., 2013*).* To reduce the negative impact of the illness both on the individual and on the society level, an effective early diagnosis and subsequent early interventions are of crucial importance to reduce the gap between early dementia onsets and received treatment (Prince et al., 2018*).*

A majority of the population over 70 years old perceives their cognitive decline (Jessen et al., 2010). One of the first concerns repeatedly reported by older patients are subjective memory complaints (SMC). The link between SMC and aging has been studied for many years. In the review of Jonker et al. (2000), the reported prevalence rates of SMC reached up to over 50%. Furthermore, research shows that patients reporting SMC are at higher risk for future conversion to dementia (Jessen et al., 2020a). SMC, therefore, should be assessed and taken into account during the neuropsychological diagnostic of cognitive state in elder patients.

There are multiple tools allowing to determine the SMC and their severity. The focus of the current study was on the Forgetfulness Assessment Inventory (FAI), a questionnaire that assesses SMC, and its clinical value in predicting the development of dementia at the early stages of cognitive decline. Our objective was also to evaluate whether the standardized z-values are a better predictor of the conversion to dementia than an unstandardized raw score, as other studies

showed discrepancies while compering raw-score and z-score values (Bondi et al., 2003). Furtherly, although SMC do not necessarily need to be confirmed by a caregiver, we were interested in answering the question if her/his assessment of patients' memory performance can be used to predict the conversion to dementia reliably.

The following chapter will tackle the crucial theoretical concepts such as subjective memory complaints, conversion to dementia, and how to assess a reliable change in a clinical context. The theoretical chapter will be followed by a detailed description of the methodology, characterizing our sampling, study design, and neuropsychological tests crucial for stated research interest. Then, the obtained results of our longitudinal study will be reported. Lastly, both the results and the limitations of the study will be critically discussed.

2. Theory

In this theoretical part of the thesis we define the concepts of subjective memory complaints, describe a model of conversion to dementia, and characterize methods determining reliable change in a clinical setting.

2.1. Subjective memory complaints (SMC)

The subjective memory complaints (SMC) can be defined as a person's perceived memory performance that is independent from their objective memory performance. In the literature the SMC are also synonymously often referred to as subjective memory decline (SMD) or subjective memory impairment (SMI) (Edmonds et al., 2014). While reviewing the literature, we will use the term SMC regardless of the term used in the original study for clarity reasons.

SMC are often first symptoms of cognitive decline noticed by elderly people. However, the research findings on the link between SMC and objective memory impairment are not consistent (Reid & MacLullich, 2006). As mentioned before, the complaints about daily forgetfulness are often made by individuals with no direct impairment on objective neuropsychological testing (Edmonds et al., 2014; Jonker et al., 2000). On the other hand, not all patients with an objective cognitive impairment experience their forgetfulness and thus show SMC (Lenehan et al., 2012).

Several studies report, that SMC have been associated with other variables, such as depressive symptoms, rather than objective memory decline (Lehrner et al., 2014; Reid & MacLullich, 2006). A link between SMC and education level has also been described (Jonker et al., 2000). In the former, a positive correlation was reported, indicating that SMC are experienced to a higher extent while patients reach higher scores on depression scales, whereas the latter showed a negative correlation linking higher levels of education to less reported SMC. Demographic variables such as age and gender has been associated with SMC as well. For older patients, the relationship between expressed SMC and measured objective

cognitive decline was stronger compared to younger individuals (Crumley et al., 2014). According to Pérès et al. (2011), women experiencing SMC are at higher risk of conversion to dementia in the long time period.

Having considered the number of confounding variables while assessing the predictive value of SMC on conversion to dementia, researchers advise on taking those into consideration during the diagnostic process (Molinuevo et al. 2017).

2.2. Conversion to dementia

While describing the link between the SMC and objective memory performance, we tackled the next theoretical concept of the thesis being the conversion to dementia. To define conversion to dementia, a three-stage mode of dementia development introduced by Jessen et al. (2010) will be discussed first.

2.2.1. Stage 1: Subjective cognitive decline

The first stage of the model is subjective cognitive decline (SCD). The Subjective Cognitive Decline Initiative (SCD-I) Working Group aimed to unify the terminology used in research and developed a detailed concept with diagnostic criteria for SCD (Jessen et al., 2014). SCD is often accompanied by subjective memory complaints yet is not limited only to the domain of memory. Other cognitive domains such as executive functions or language may be affected as well. The prevalence rates of SCD differ highly. A study of Jessen et al. (2010) reported that between 50% and 80% of people over 70 years old report cognitive decline.

According to Jessen et al. (2014), SCD can be linked to other variables such as aging, personality, psychiatric disorders, medication, etc. Among patients actively seeking help at memory clinics, those showing higher levels of neuroticism tend to report SCD more often (Goldberg et al., 2020).

Moreover, SCD has been linked to Alzheimer's disease (AD) through neuroimaging methods and biomarkers being volume loss of the grey matter (Striepens et al., 2010), cerebral hypometabolism (Mosconi et al., 2008) and amyloid deposition (Perrotin, 2012). Lista et al. (2015) in a review article reported possible uses of magnetic resonance imaging and positron emission tomography in early stages dementia diagnosis. Currently however, the use of biomarkers alone has not a sufficient predictive value to be used for predicting conversion to dementia (Jessen, 2020).

2.2.2. Stage 2: Mild cognitive impairment

Second stage in the model proposed by Jessen et al. (2010) is the mild cognitive impairment (MCI). A meta-analysis of conversion rates from SCD to MCI reported a 27% rate (Mitchell et al., 2014). As a diagnostic entity MCI was firstly described by Petersen (1999, 2004) and is often considered being the transitional stage to dementia. Patients with MCI may experience similar symptoms to dementia, yet the severity levels of their cognitive impairment are not high enough to be diagnosed with dementia.

First set of diagnostic criteria proposed at the Mayo Alzheimer's Disease Research Center (Petersen, 2004) were:

1. subjective memory complaints, when possible confirmed by a caregiver;
2. objective memory impairment corrected for age;
3. good general cognition corrected for age;
4. ability to function independently in a daily life;
5. no dementia.

Those criteria have been later updated to include not only memory deficits but cognitive deficits in other domains as well. MCI can therefore be further divided into two subtypes being amnestic and non-amnestic MCI, (aMCI and naMCI respectively) (Petersen, 2004). Amnestic MCI (aMCI) is diagnosed when patients' memory is significantly impaired, yet the impairment does not meet criteria

necessary for dementia diagnosis. Other domains like language or executive function are not affected. Nonamnestic MCI (naMCI) is diagnosed when there is an impairment in other domains mentioned before, but there is no memory impairment (Petersen, 2011).

2.2.3. Stage 3: Dementia

The third and last stage of the dementia development model of Jessen et al. (2010) is the diagnosis of dementia itself. Dementia is a syndrome caused by brain disfunction with impaired higher cortical functions such as for instance memory, orientation, learning and language. It is often accompanied by changes in emotional control, social behavior and motivation (DIMDI, 2013). In comparison to MCI, decline due to dementia has to be severe enough to restrict patient's functioning in her or his daily life.

The criteria of Alzheimer's Disease diagnosis had changed over time. As the current study is of retrospective nature, we will firstly introduce criteria used during the data collection period and then the most current ones.

According to NINCDS-ADRDA Work Group (McKhann et al., 1984), the criteria for diagnosis of AD are divided into required, probable and possible ones. All required and probable criteria have to be met in order to get the Alzheimer's Disease Definite diagnosis. Those criteria are:

1. "histopathologic evidence obtained from a biopsy or autopsy;
2. dementia established by clinical examination and documented by the Mini-Mental Test, Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests;
3. deficits in ≥2 areas of cognition;
4. progressive worsening of memory and other cognitive functions;
5. no disturbance of consciousness;
6. onset at age >40 to <90 years;

7. absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition" (McKhann et al., 1984).

The possible criteria include "dementia syndrome, in the absence of other neurologic, psychiatric, or systemic disorders sufficient to cause dementia, and presence of variations in onset, in presentation, or in clinical course; presence of a second systemic or brain disorder sufficient to produce dementia but is not considered to be the cause of the dementia; a single, gradually progressive severe cognitive deficit is identified in the absence of other identifiable cause" (McKhann et al., 1984).

The term dementia was recently substituted with a new one – the major neurocognitive disorder (major NCD) in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013). However, both terms can be used interchangeably. According to DSM-5, several cognitive domains can be affected in dementia: attention, executive function, learning and memory, social cognition, motor function and language. There are also multiple subtypes of the disease, based on its causes, for example dementia due to Alzheimer's Disease, dementia due to Parkinson's Disease, dementia due to traumatic brain injury, etc. (Sachev et al., 2014).

The current criteria for dementia diagnosis are memory impairment and additional decline in cognitive abilities, such as:

- speaking coherently or understanding spoken or written language;
- recognizing or identifying objects (intact sensory function assumed);
- performing motor activities (intact motor and sensory function and understanding of the task assumed);
- thinking abstractly, making accurate judgments, and planning and carrying out complex tasks.

The presented diagnostic criteria of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-11; World Health Organization, 2019) are very similar to criteria of DSM-5.

2.3. Subjective memory complaints and conversion to dementia

The research on the subjective memory complaints and their value in predicting conversion to dementia is not conclusive. A longitudinal study of Wang et al. (2004) showed no significant difference at baseline in SMC between patients that got diagnosed with dementia at the follow-up and patients that did not. Wang et al. (2004) concluded therefore, that self-reported SMC are not useful at predicting conversion to dementia. Another study showed however, that patients experiencing SMC are indeed at higher risk for eventual conversion dementia at the further stage of their lives (Reisberg et al., 2010). A meta-analysis showed, that patients diagnosed with SCD are twice as likely to develop dementia compared to patients with no reported SCD and reported a 14% conversion rate from SCD to dementia (Mitchell et al., 2014). Jessen et al. (2014a) as well linked reported SMC to higher risk of conversion to dementia for patients with SCD at baseline.

As mentioned before, the time interval between the baseline and follow up differs strongly between studies investigating the link between SMC and dementia. So do conversion rates mentioned in the literature. However, the progression rates tend to be higher in the first two years of the follow-up and then decline (Mitchell & Shiri-Feshki, 2008).

A longitudinal study of Wang et al. (2004) showed a 15% rate of patients with SMC at baseline converted to dementia, whereas a study of Silva et al. (2016) showed that within 2 years 36.6% of non-demented patients with SMC converted to dementia. In the meta-analysis of Mitchell et al. (2014), 6.6% patients experiencing SMC converted to MCI and 2.3% of patients to dementia within one year. Looking

at the time period of four years, the percentages raised to 24.4% and 10.9% respectively. Patients without SMC had a 4.6% chance of progressing to dementia. When exploring the differences between converters and non-converters regarding SMC at baseline, Lehrner et al. (2016) found no differences between the groups. Correlations between SMC and other variables were also not consistent.

The predictive value of biomarkers in progression from SCD to MCI and further to AD, showed that at highest risk are patients with higher abnormalities in cerebrospinal fluid markers such as tau protein and amyloid (Wolfsbruber et al., 2017).

2.4. Reliability change assessment

To assess change in repeated testing situations, apart from introducing a control group, finding the right retest interval and deciding on similar or same form of the follow-up testing, there are different statistical methods that can be introduced to assess for change. The first method, Simple Discrepancy Score, measures the difference between the second and first testing point ($X_2 – X_1$) (Duff et al., 2012). Another method, known as the Standard Deviation Index, takes the standard deviation of the test score at the first testing point (SD_1) creating a z-score. The z-score of ±1.645 indicates a reliable change, as it corresponds to a 90% confidence level (Hinton-Bayre, 2010). Another method introduced by Jacobson & Truax (1992), the Reliable Change Index (RCI), was created in order to assess reliable change in psychotherapy research. Even though RCI by Jacobson & Truax (1992) includes the test-retest reliability, neither the practice effects nor the score variability at the follow-up are included.

In the following part, we will describe further three models for reliable change commonly used in the literature.

2.4.1. Chelune et al.

Because as opposed to psychological conditions, the cognitive symptoms often can be influenced by repeated testing, Chelune et al. (1993) proposed an extension to the RC model by Jacobson & Truax (1992) including practice effects. They assessed reliable change using the following formula:

$$RC = (X_2 - X_1) - (M_2 - M_1) / SE$$

$M_2 - M_1$ is a difference score between the means of the follow-up and baseline measurement in the healthy control group.

Both methods assessing reliable change described shortly above, can be criticized for not including patients' individual variables into the formula. To tackle that criticism and to predict the score reached at the follow-up, regression-based methods can be used. Depending on which variables are used to predict the follow-up score, the standardized regression based formula, referred to as SRB, can be either simple if using only the score from the first testing to predict the score at the follow-up, or complex if we include other variables in the regression as well. We chose two regression-based RC models that will be described in the next section.

2.4.2. McSweeny et al.

McSweeny et al. (1993) proposed a regression-based method that incorporated in its formula both, the relevant demographical variables such as age and education level, and the interval between the two testing points. Their RCI can be calculated with the formula presented below. SEE stands for the standard error of the estimate from the regression equation.

$$RCI_{SRB} = (Y - Y') / SEE$$

$$SEE = SD_2 \sqrt{(1- \text{test-retest-reliability}_2)}$$

SD_2 is the control group retest standard deviation. The formula to predict the estimated score at the follow-up (Y') was:

$$Y' = bX + c$$

$$b = r * (S_1 / S_2)$$

$$b' = r * (var_2 / var_1)$$

$$c = M_2 - b\, M_1$$

The bX stands for the slope of least squares regression line of retest score on baseline score, and c for the constant. S_2 is the retest variance of the control group, S_1 is the control group variance at baseline.

2.4.3. Maassen et al.

According to several studies (Maassen et al., 2006), the estimates for the main elements of RCI_{SRB} can be also obtained from descriptive data. Both SRB and RCI_{SRB} can be estimated by means and standard deviations from both testing points. The proposed formulas for the estimates are:

$$Y' = b'X + c'$$

$$b' = S_2 / S_1$$

$$c' = (M_2 – [b'M_1])$$

$$SEE' = \sqrt{[(SD_2^2 + SD_1^2)(1 – r_{12})]}$$

b' is the slope (beta coefficient) and c' is the intercept of the regression model. S_2 is the retest variance of the control group, S_1 is the control group variance of the test at baseline. SEE' stands for the standard error of the estimate.

2.5. Study aim

Aim of the present retrospective quasi-experimental longitudinal study is to assess the prognostic value of subjective memory complaints in conversion to dementia. We will compare measures of SMC at baseline of patients being diagnosed with either subjective cognitive deficits (SCD) or mild cognitive impairment, both amnestic and non-amnestic (aMCI, naMCI), with measures of SMC of patients that converted to dementia (AD) at the second testing. We will also conduct the same analysis for the caregivers' rating of patients' memory performance to test the hypothesis that the informants' ratings of memory performance do not have a better predictive value then self-ratings of SMC in regard to conversion to dementia. Moreover, we will also compare the predictive value for both SMC raw scores and standardized z-scores in terms of conversion to dementia. Additionally, we will explore the correlations between SMC and other variables. Finally, we aim to describe the differences between selected methods assessing reliable change.

2.6. Research questions

The master thesis aims to answer following research questions and to test hypotheses stated below.

Research question 1: Can the measurement of subjective memory complaints (SMC) predict the conversion to dementia?

> Hypothesis 1 (H1): Subjective memory complaints cannot predict conversion to dementia.

Research question 2: Can the caregivers' ratings of patients' measurement of memory performance predict the conversion to dementia?

Hypothesis 2 (H2): Patients' memory performance rated by their caregivers cannot predict conversion to dementia.

Research question 3: Can a z-score value of subjective memory complaints based on patient's raw-score predict patients' conversion to dementia?

Hypothesis 3 (H3): A z-score cannot predict conversion to dementia.

Research question 4: Is there a correlation between subjective memory complaints and both demographical variables such as sex, age and education, and NTBV, BDI and WST scores?

To address research question 4 we conduct an exploratory analysis of correlations between FAI score and variables of interested listed in the research question 4 to verify possible correlations between variables.

Research question 5: Is there a difference between methods for assessing reliable change for subjective memory complaints in patients who convert to dementia and who did not convert?

To answer the research question 5 we will descriptively compare reliable change methods.

3. Methods

3.1. Participants and data collection

The present study was part of a research project conducted at the Medical University of Vienna “The Vienna Conversion to Dementia Study” and was approved by the Ethics Committee of the Medical University of Vienna (EK-2231/2019). The data used for the study was collected at the Department of Neurology between June 2001 and October 2018 as part of the neuropsychological assessment with an approval from the Ethics Committee of the Medical University of Vienna (EK-174/2008).

The clinical sample consisted of patients that consulted the outpatient clinic by their own initiative or were referred by a doctor after having reported experiencing memory deficits. In the study we included patients that met following six criteria:

1. Minimum 50 years old;
2. Mini Mental State Examination (MMSE) score higher than 23;
3. No neurological diseases such as stroke, traumatic brain injury, multiple sclerosis or epilepsy;
4. No current major psychiatric diagnosis according to ICD-10 (Dilling et al., 2008) with exclusion of subdepression (Lehrner et al., 2014);
5. No diagnosis of dementia already at baseline (Saß et al., 2003);
6. No language, motor, auditory or visual deficits, or other medical conditions that would influence patient’s cognitive performance.

The six inclusion criteria were chosen in order to control for the possible cognitive deterioration affecting patients’ cognitive performance, and therefore interfering with the aim of the study (Stephan et al., 2011). Patients showing depressive symptoms were included in further analysis as depressive symptoms were shown to be common in the elderly patients (Lehrner et al., 2014).

The healthy control group (CG) was chosen based on the same inclusion criteria listed above except the second criterion that was changed to the MMSE score ≥ 27.

If patients were accompanied by their caregivers, the caregivers were asked to take part in the assessment as well. Each participant included in the study signed a written consent.

3.2. Testing procedure

3.2.1. Mini Mental State Examination (MMSE)

After an assessment during anamnesis to screen for exclusion criteria, we used the Mini Mental State Examination (MMSE) to assess cognitive deficits (Folstein et al, 1975). As stated before, only patients with the MMSE score higher than 23 points underwent further neuropsychological tests and were included into present study. Other tests will be described in the section below. The whole neuropsychological testing took approximately 45 to 90 minutes. Other neurocognitive tests used for our purpose were described in the following part of the thesis.

3.2.2. The Wortschatztest (WST)

To measure patient's verbal comprehension, we used a standardized vocabulary test, The Wortschatztest (WST) (Schmidt & Metzler, 1992). Each item consists of six words, of which only one is a real word. Patient's WST score can be further converted to the Intelligence Quotient (IQ) score.

3.2.3. Beck Depressions-Inventory (BDI-II)

To test patients' depressive symptoms in the last two weeks, we used revised version of Beck Depressions-Inventory II (Hautzinger et al. 2006). BDI-II consists of 21 items. Each question can be rated on a four-point Likert scale (from 0 to maximum 3 points per question). A BDI-II score higher than 13 is an indication for mild depressive symptoms.

3.2.4. Neurological Test Battery Vienna (NTBV)

The Neurological Test Battery Vienna (NTBV), designed to detect dementia in a clinical setting, was introduced to measure patients' cognitive functioning. The NTBV consists of tests assessing domains such as attention, language, memory and executive functions (Lehrner et al., 2005). The single subtests of NTBV categorized per each domain are presented below.

Domain 1: Attention

- AKT: The Alters-Konzentrationstest (AKT) (Gatterer, 1990)
- Digit Symbol Test, a subtest of the German WAIS-R (Tewes, 1994)
- the c.I Symbol Test, a symbol counting subtest from the C.I. test (Lehrl & Fischer, 1997)
- TMTB: The Trail Making Test B (Reitan, 1979)
- TMT B - TMT A: the score difference of the Trail Making Test A and B.

Domain 2: Language

- SVF *(Goodglass & Kaplan, 1983)*
- BNT: The modified Boston Naming Test (Morris, et al., 1989)

Domain 3: Memory

- Verbal Selective Reminding Test (VSRT) with its subtests: Immediate Recall, Total Recall, Delayed Recall and Recognition (Lehrner et al., 2006)

Domain 4: Executive functions

- TMT A: The Trail Making Test A (Reitan, 1979)

- The Five-Point-Test (Regard et al., 1982)
- The Maze Test (Oswald & Fleischmann, 1997)
- The Stroop Test from the NAI Battery (Oswald & Fleischmann, 1997)
- The Interference Test from the C.I. (Lehrl & Fischer, 1997)
- PVF (Goodglass & Kaplan, 1983)

Cronbach alpha of NTBV ranges between .87 and .89 (Lehrner et al., 2007), and its test-retest reliability lies between *r* =.69 and *r* =.94 (Lehrner et al., 2007; Macher 2013; Hitzl, 2015 as cited in Lehrner et al. 2016).

For each subtest score, a corresponding z-value was calculated. In the current study patients with a MMSE score under 24 were given a short version of NTBV. Those patients were not included in the analysis.

3.2.5. The Forgetfulness Assessment Inventory (FAI)

The Forgetfulness Assessment Inventory (FAI) was designed to measure subjective memory complaints in daily life situations. The FAI consists of 16 questions related to subjective memory performance within past four weeks. The patient is asked to rate single items on a 5-point Likert scale from 1 (never) to 5 (very often). The questions address areas of a daily life such as remembering faces, names, birthdays and recent tv or radio content. The FAI-raw-score is an arithmetic mean of items' sum divided by the number of answered questions. The higher FAI-score means more subjective memory complaints. The FAI result can be further converted into a standardized z-score. The FAI has a good internal consistency with a Cronbach alpha of 0.93 (Lehrner et al., 2014).

The same version of the Forgetfulness Assessment Inventory was handed in to the caregivers.

3.3. Study design

To address our research questions, we conducted a retrospective quasi-experimental longitudinal study. As a result of anamnesis, clinical interviews and described above neuropsychological testing at the first testing point (TP 1), the experimental group (EG) was divided into four subgroups based on the received diagnosis – into subjective cognitive decline (SCD), nonamnestic mild cognitive impairment (naMCI), amnestic mild cognitive impairment (aMCI) and Alzheimer's disease (AD) group.

The SCD diagnosis was given when the patient experienced subjective memory complains and hers/his z-scores in every domain tested were greater than -1.5 SD. (Lehrner et al., 2014) The MCI was diagnosed using Petersen's criteria (Petersen, 2004), when the patient reached a z-score of -1.5 SD below the norm corrected for age, sex and education level (Lehrner et al., 2014). Alzheimer's Disease (AD) was diagnosed using the criteria of NINCDS-ADRDA Working Group (McKhann et al., 1984).

The independent variable (IV) assessing the subjective memory complaints was the Forgetfulness Assessment Inventory (FAI) score obtained at the TP 1 as part of the neuropsychological testing. In addition to the patients' FAI raw scores, we calculated their z-values as well. In case of patients being accompanied by their caregivers, the caregivers' rating of patients' memory performance in the last four weeks was also assessed using FAI to answer our research questions.

The second testing point (TP 2) took place 12 to 48 months after TP 1 and followed the same procedure.

Conversion to dementia was defined as receiving the Alzheimer's Disease diagnosis at TP 2 and was our dichotomous dependent variable (DV) coded with 0 or 1, indicating not converting or converting to dementia respectively. In the present thesis we only included patients that got diagnosed with dementia due to Alzheimer's disease.

3.4. Statistical analysis

Statistical analysis of the collected data was performed using SPSS Statistics for Windows (Version 26.0). The significance level was set to $\alpha = .05$. The results will be referred to as significant when $p \leq .05$ and as highly significant when $p \leq .001$ (Bortz & Schuster, 2010).

The data was firstly checked for plausibility in order to tract missing or incorrect values and filled in (missing diagnosis) or corrected (incorrectly inputted value). The binary variable *conversion to dementia* (0/1) was created based on the diagnosis received at the follow-up testing.

In the first place, the demographic and clinical data were descriptively analyzed and checked for normality of the distribution. As we assumed that the single subtests of NTBV correlate within the test battery, we tested for multicollinearity. To extract the highest loading subtest per each factor, we conducted a factor analysis of the NTBV test scores.

In order to calculate z-scores from FAI raw-scores, a GAMLSS model was fitted to data from 504 healthy controls collected at the Medical University of Vienna between 2001 and 2018. The model simultaneously fits P-splines (non-parametric penalized smoothing splines) to the mean, variance and skewness parameters of a left-censored BCCG (Box-Cox Cole and Green) distribution. Left-censoring was used to account for a distributional spike at the lowest FAI value of 1 which demonstrates that younger patients would possibly perform even better than best on the FAI scale. A Generalized Akaike Information Criterion was used with penalization 4.5 for selecting degrees of freedom within each P-spline, and with penalization 3.0 for selecting model terms. Worm plots and Q-statistics were used to check model assumptions. A selection among left-censored versions of normal, log-normal and BCCG distribution was done using AIC.

In the next step we conducted three logistic regressions for patients and their caregivers with the aim to assess the predictive power of subjective memory complaints in conversion to dementia. The binary logistic regression allows to

predict a dichotomous variable with a categorical or continuous one. The assumptions for the binary logistic regression are the linearity of logit (linear relation between the independent variable and log odds) and no multicollinearity (predictors should not correlate highly with each other) (Gollwitzer et al., 2013). As opposed to interpreting the results of the linear regression, a linear interpretation of the results of the logistic regression is not possible. The sign (+/-) of the regression coefficients allows to interpret the probability of the outcome. A positive regression coefficient means a higher probability for a positive outcome when scoring higher on a variable. A negative regression coefficient indicates a lower probability for a positive outcome when getting a higher value (Backhaus et al., 2016). The Odds Ratios (OR), obtained during the logistic regression analysis, are also an indicator of the probability for the outcome. OR > 1 means a higher probability compared with the reference group, OR < 1 a lower probability. Yet it is important to underline that the interpretation of the odds should be made with caution, as the probabilities are relative.

After having tested the assumptions of the logistic regression described above, we conducted three binary logistic regressions for patients and their caregivers with the aim to assess the predictive power of subjective memory complaints in conversion to dementia. As we were interested whether there is a difference between the predictive power of raw-scores and calculated z-values, our independent variables for the logistic regressions were 1. *patients' FAI raw score (*FAI*)*, 2. *patients' FAI z-value*. To answer our research question, the independent variable for the third logistic regression was 3. *caregivers' FAI raw score (FAI*_Proxy*)*. The dichotomous dependent variable for each logistic regression was patients' conversion to dementia. Age, sex, education level and subtests of NTBV extracted in the factor analysis were included as controlled variables. To interpret the goodness-of-fit if the model, we used the Nagelkerke R-Squared. The higher value of the estimate indicates a better model fit. When Nagelkerke R-Squared > .05, the model can be described as good.

Furthermore, a Receiver Operator Characteristics analysis (ROC) was conducted in order to predict the predictive validity of subjective memory complaints in conversion to dementia, both for the FAI raw-scores and FAI z-scores of patients'

self-ratings and FAI raw-scores of caregivers' ratings with the conversion to dementia as positive condition. We calculated: Youden's index using the formula J = Sensitivity + Specificity − 1 (Youden, 1950), sensitivity, specificity, positive predictive value (PPW), negative predictive value (NPV) and both positive and negative likelihood ratios (LR+, LR-) (Jaeschke et al. 1994). The test sensitivity of the test describes the probability of correct diagnosis of dementia, whereas the test specificity shows the probability of correctly diagnosis a healthy subject. PPV is an estimate of the probability of dementia diagnosis and receiving a positive test result (diagnosis of dementia). NPV describes the probability of being healthy and receiving a negative test result (not getting the dementia diagnosis) at the same time. LR+ measures the odds of having dementia and getting a positive test result in comparison to a person that does not have the disease and tests positive. LR- measures the odds of having dementia and getting a negative test result in comparison to a person that does not have the disease and tests negative. To interpret the calculated likelihoods, we used probabilities proposed by Jaeschke et al. (1994):

- convincing diagnostic evidence LR+ > 10, LR- < 0.1
- high diagnostic evidence LR+ 5-10 LR- 0.1.-0.2
- weak diagnostic evidence LR+ 2-5 LR- 0.2-0.5
- scarce diagnostic evidence LR+ 1-2 LR- 0.5-1

The ROC was furtherly evaluated in regard to its prognostic based on the area under the curve (AUC) (Bortz & Döring, 2007). Looking at AUC enables to evaluate how accurate the test is in discriminating between patients that have converted to dementia and those who have not (Gollwitzer et al., 2013). A perfect test is represented by an AUC of 1, a poor test can be characterized by an AUC of .5, as it means the test predictions are random and the test use is therefor of low value (Fawcett, 2006). A AUC of 0 means that the test falsely classified all subjects (Hajian-Tilaki, 2013).

The next step of our analysis was to calculate total group Spearman correlation between subjective memory complaints and selected variables. We included both demographical variables such as sex, age, education, and tests measuring

objective memory performance (NTBV). We also looked at the intelligence level (WST) and depressive symptoms (BDI-II).

Finally, in the exploratory part of our study we calculated and descriptively characterized the Reliable Change Index (RCI). We wanted to assess the accuracy of different methods for reliable change in the context of neuropsychological testing.

The FAI baseline scores of each patient (X_1) was converted into a z-score using the formula presented below. The z-values were then put in ascending order.

$$(X_1 – M_1) / (SD_1 - X_1)$$

Afterwards, we calculated mean (M), standard deviation (SD) and variance of FAI for both testing points (M_1, M_2, SD_1, SD_2, S_1, S_2) for the control group and used the formula presented below to calculate the RCI by Chelune et al. (1993). In order to calculate the standard error of difference (SED), SD_1 and test-retest-reliability of the healthy control group were used.

Furthermore, we compared the three models assessing change by Chelune et al. (1993), McSweeny et al. (1993) and Maassen et al. (2006) for all clinical diagnostic groups and for converters and non-converters to dementia. The calculations were made with the *Reliable Change Calculator 1.0* provided by Dr. Hinton-Bayre. Due to our control group size < 50, we considered change to be significant at a t value of ± 1.684 instead of the z-score of ± 1.645 following the suggestions of Hinton-Bayre (2010).

4. Results

Statistical analysis was performed using SPSS Statistics (Version 26.0). Unless stated otherwise, the significance level was set to $\alpha = 0.05$.

4.1. Descriptive analysis

We conducted our descriptive analysis for the whole sample, and for each diagnostic group, being the healthy controls, SCD, naMCI and aMCI. The Kolmogorov-Smirnov test was used to test for normal distribution.

Our total sample size consisted of 299 participants, 153 women (51.2%) and 146 men (48.8%). Due to the distribution characteristics, we described most of the demographical and clinical variables with the median and the interquartile range (IQR). The median age of participants was 67 (IQR = 61-73). The education level in our sample displayed a median of 12 (IQR = 8-16). At baseline point, 38 patients (12.7%) were diagnosed with SCD, 215 with the MCI (71.9%), of which 94 (31.4%) received the diagnosis of aMCI and 121 (40.5%) of naMCI. Our control group consisted of 46 subjects (15.4%). For more details see Table 1.

Table 1

Demographic Characteristics of the Sample at Baseline

Baseline	ALL	HC	SCD	naMCI	aMCI
N	299	46	38	121	94
Sex (w/m)	153/146	16/30	17/21	72/49	48/46
Age	67 (61-73)	61 (58.75-62)	65 (59.75-73)	69 (63-74.5)	69. 5 (62-75)
Education	12 (8-16)	12 (8-18)	12 (9.75-16)	11 (8-15)	12 (8-16)
MMSE	28 (27-28)	29 (28-30)	29 (27.75-30)	28 (27-29)	28 (26-29)
WST	110 (101-122)	116 (107-122)	118 (107-125)	110 (99-118)	110 (99-121)
BDI-II	8 (3-13)	2 (1-5.25)	8 (5.75-13)	10 (6-14.75)	8 (4-13.25)

Note. Median, (Interquartile Range); Sex (w/m), Number of women/men; HC, Healthy Control Group; SCD, Subjective Cognitive Decline; aMCI, amnestic Mild Cognitive Impairment; naMCI, non-amnestic Mild Cognitive Impairment; MMSE, Mini mental State Examination; WST, Wortschatztest; BDI, Beck Depression Inventory.

The FAI questionnaire assessing SMC was completed by 299 patients and 50 caregivers at baseline, and 259 patients and 45 caregivers at follow-up. Differences in means for converters and non-converters are displayed in Table 2.

Table 2

The mean FAI Scores and FAI z-scores for Self-Rating and Caregivers'-Rating at Baseline and Follow-up

Baseline	**Self** N = 299		**Proxy** N = 50	
	Non-Converters	**Converters**	**Non-Converters**	**Converters**
	N=270	N=29	N=37	N=13
FAI	2.82	3.07	2.88	3.31
FAI z-score	-0.68	-0.98	-	-
Follow-up	**Self** N = 259		**Proxy** N = 45	
	Non-Converters	**Converters**	**Non-Converters**	**Converters**
	N=245	N=14	N=28	N=17
FAI	2.76	3.07	3.10	3.60
FAI_z	-.57	-.39	-	-

Note. Mean; FAI, the Forgetfulness Assessment Inventory; FAI_z, the Forgetfulness Assessment Inventory z-scores; Self, the Forgetfulness Assessment Inventory Self-Rating; Proxy, the Forgetfulness Assessment Inventory Caregiver's Rating.

At the follow-up, 57 patients (19.1%) were diagnosed with SCD, 167 with the MCI (55.9%), of which 78 (26.1%) received the diagnosis of aMCI and 89 (29.8%) of naMCI. 29 patients converted to dementia at the follow-up. Therefore, the conversion rate in our study was 9.7%. The detailed conversion rates and the number of converted patients for each clinical diagnosis at baseline and follow-up are shown in Table 3.

Table 3

Rates of conversion to dementia

		T2					
		HC	SCD	naMCI	aMCI	AD	TOTAL
T1	HC	46 (100)	-	-	-	-	46
	SCD	-	22 (57.9)	9 (23.7)	7 (18.4)	-	38
	aMCI	-	8 (8.5)	18 (19.1)	41 (43.6)	27 (28.7)	94
	naMCI	-	27 (22.3)	62 (51.2)	30 (24.8)	2 (1.7)	121
TOTAL		46	57	89	78	29	299

Note. N (Percentage); HC, Healthy Control Group; SCD, Subjective Cognitive Decline; aMCI, amnestic Mild Cognitive Impairment; naMCI, non-amnestic Mild Cognitive Impairment; AD, Alzheimer´s Disease; T1, Baseline; T2, Follow-up.

The characteristics of converters and non-converters are displayed in Table 4. The median age of non-converters was 66 (60-73) and of converters 73 (67-78). The follow-up testing in our sample took place after 24 months (15-36), whereas the interval's median for non-converters was 25 (15.75-37) and for converters 29 (14-32).

Table 4

Characteristics of Non-Converters and Converters at Baseline and Follow-up

	Non-Converters N=270		**Converters** N=29	
	Baseline	Follow-up	Baseline	Follow-up
Sex (w/m)	138/132	-	15/14	-
Age	66 (60-73)		73 (67-78)	
Education	12 (8-16)	-	12 (8-16)	-
Interval	-	25 (15.75-37)	-	29 (14-32)
MMSE	29 (27-29)	28 (27-29)	26 (25-27)	24 (21-25)
WST	110 (101-121)	111 (101-122)	110 (99.6-122)	109 (94.5-124.75)
BDI-II	8 (3-13)	8 (3-14)	6 (4-10)	6.5 (4.75-9.25)
FAI_Self	2.82*	2.76*	3.07*	3.10*
FAI_z	-.68*	-.98*	-.98*	-.39*
FAI_Proxy	2.88*	2.98*	3.31*	3.60*

*Note. Median, (Interquartile Range), Mean *; Sex (w/m), Number of women/men; HC, Healthy Control Group; SCD, Subjective Cognitive Decline; aMCI, amnestic Mild Cognitive Impairment; naMCI, non-amnestic Mild Cognitive Impairment; MMSE, Mini mental State Examination;* WST, Wortschatztest; BDI, Beck Depression Inventory; FAI_Self, the Forgetfulness Assessment Inventory Self Rating; FAI_z, the Forgetfulness Assessment Inventory z-scores; FAI_Proxy, the Forgetfulness Assessment inventory Caregiver's Rating.

4.2. Prediction of conversion to dementia

To answer the research questions 1 and 2, whether subjective memory complaints measured with the FAI score or the caregiver's assessment of patient's condition (FAI_Proxy) can be used for predicting the conversion to dementia at the second testing point, we conducted the following analyses. In order to include some subtests of NTBV as predictors at the later stage of analysis, we conducted a factor analysis of NTBV.

As the subtests of NTBV test battery are strongly correlated with each other, in order to meet the criterium of no multicollinearity, we calculated the variance inflation factor (VIF) for the subtests of NTBV. We also included other variables of

interest for the current study: age, sex, education, FAI, BDI and WST. The VIF method allows to show whether predictors have a linear relationship with each other. The cut-off point was set to VIF > 10 indicating a strong linear relationship (Field, 2013). As three of NTBV subtests (Stroop Test Interference, Stroop Test Words Stroop Test Color) had the VIF greater than 10, we conducted a principle component analysis (PCA) to reduce the number of predictors. The obtained Kaiser-Meyer-Olkin criterion was .839, indicating a good fit to perform PCA (Field, 2013). Moreover, the Barlett-Test was highly significant (p <.001), meaning that items correlate highly enough with each other to perform a PCA. In order to select the number of components, we first looked at the eigenvalues of each component. According to Kaiser's criterion, the components with an eigenvalue >1 can be extracted. However, two out of six factors had their eigenvalue almost equal 1. Therefore, in the next step we had analyzed the scree plot and decided on extracting four factors explaining 63.63% of variance. Table 5 shows the structure matrix for single subtests of NTBV. The subtest loading the highest for each factor are shown in bold.

Table 5

Structure Matrix for subtests of NTBV

	Component			
	1	2	3	4
TMT-B	**-,803**	-,343	,520	,364
Digit-Symbol Test	,772	,464	-,561	,113
AKT Sum	,765	,389	-,595	,004
5 Point Test Correct	,760	,321	-,446	,131
Maze Total/Time	,746	,245	-,319	-,129
AKT Time	-,735	-,372	,640	,030
Maze Time	-,734	-,233	,350	,191
Interference Time (c.I.)	-,732	-,400	,695	-,221
Interference Total / Time (c.I.)	,724	,423	-,664	,226
SVF Correct	,693	,552	-,435	,097
TMT-B – TMT-A	-,688	-,259	,439	,427
TMT-A	-,665	-,377	,446	,003
Symbols counting (c.I.)	-,604	-,192	,315	-,151
PVF Correct	,560	,305	-,358	,292
BNT	,371	,300	-,092	-,217
VSRT Learning Performance	,441	**,930**	-,374	-,039
VSRT Delayed Recall	,396	,904	-,408	-,111
VSRT Word Span	,383	,821	-,249	,024
VSRT Recognition	,134	,690	-,238	,020
Stroop Test Time Words	-,580	-,433	**,957**	-,054
Stroop Test Total/Time	,508	,352	-,918	,106
Stroop Test Interference	-,437	-,355	,904	,030
Stroop Test Time Colors	-,609	-,404	,644	-,215
5 Points Test Perseverations	-,004	-,046	-,052	**,717**

Note. Extraction Method Oblimin Rotation; NTBV, the Neurological Test Battery Vienna; TMT-B, Trail Making Test Version B; AKT Sum, Age Concentration Test Sum; AKT Time, Age Concentration Test Time; c.I., Cerebral Insufficiency Test; SVF, Semantic Verbal Fluency; TMT-B – TMT-A, Interference; TMT-A, Trail Making Test Version A; PVF, phonematic Verbal Fluency; BNT, Boston Naming Test; VSRT, Verbal Selective Reminding Test.

As predictors for further analysis, we included therefore only those subtests of NTBV that loaded the highest on each factor extracted. Those predictors were TMT-B, VSRT learning performance, Stroop Test Time Words and 5-Points-Perservations, covering cognitive domains such as divided attention, memory, executive functions and problem solving respectively.

After having extracted the predictors, we tested the assumptions for the logistic regression. Multicollinearity criterium (all VIF values were smaller than 5) and linearity of logit were both met. Afterwards we conducted logistic regressions to answer research questions 1-3. First, due to the exploratory aim of the study the enter method was used in order to identify the significant predictors within our model. Then, we conducted the analysis using the backward stepwise elimination method with significant predictors only in order to extract the ones with highest predictive value and to reduce possibility of the Type II error (Field, 2013). The predictors were subjective memory complaints, followed by age, sex, education and 4 subtests of NTBV chosen in the process of factor analysis.

The results of the binary logistic regression with the conversion to dementia as dependent variable for patient's rating of SMC, the FAI, are shown in Table 6. Significant results were found for the two subtests of NTBV, the VSRT Learning Performance ($p < .001$) and Stroop Test Time Words ($p = .03$).

Table 6

Logistic regression for FAI, the Enter method

							95% C.I.for Exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
FAI	-,525	,399	1,728	1	,189	,592	,271	1,294
Age	-,034	,036	,864	1	,353	,967	,900	1,038
Sex	,832	,593	1,970	1	,160	2,298	,719	7,346
Education	,100	,070	2,026	1	,155	1,105	,963	1,267
TMT-B	,007	,005	2,079	1	,149	1,007	,997	1,017
VSRT LP	-,202	,041	23,685	1	,000	,817	,753	,886
Stroop Test Time Words	,036	,017	4,530	1	,033	1,037	1,003	1,072
5 Points Test Perseverations	,034	,050	,461	1	,497	1,034	,938	1,141
Constant	4,910	3,739	1,724	1	,189	135,662		

Note. R_2 *= .250 (Cox & Snell), .534 (Nagelkerke); Model* $\chi 2$ *= 84.168, p<.001;* FAI, the Forgetfulness Assessment Inventory Self Rating; TMT-B, Trail Making Test Version B; VSRT LP, Verbal Selective Reminding Test Learning Performance.

Table 7 shows the results of the logistic regression using the backwards stepwise elimination method.

Table 7

Logistic Regression, the backward stepwise elimination for VSRT LP and Stroop Test Time Words

							95% C.I.for Exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
VSRT LP	-,170	,036	22,917	1	,000	,844	,787	,904
Stroop Test Time Words	,039	,013	8,767	1	,003	1,040	1,013	1,067
Constant	2,279	1,601	2,025	1	,155	9,763		

Note. R_2 *= .222 (Cox & Snell), .475 (Nagelkerke); Model* $\chi 2$ *= 73.647, p<.001;* VSRT LP, Verbal Selective Reminding Test Learning Performance.

Significant results were found for the two subtests of NTBV, the VSRT Learning Performance ($p < .001$) and Stroop Test Time Words ($p = .003$). With each point reached on VSRT Learning Performance the Odds decrease by 0.84 [0.78; 0.90]. Each second longer on the Stroop Test Time Words, increased the odds by 1.041 [1.01; 1.07]. As we did not find significant results for FAI, we confirmed our H1, that SMC measured with FAI cannot predict conversion to dementia.

The results of the binary logistic regression with the conversion to dementia as dependent variable for 50 caregivers' ratings are shown in Table 8. The FAI_Proxy did not show a significant predictive value in prediction of conversion to dementia.

Table 8

Logistic regression for FAI_Proxy, the Enter method

							95% C.I.for Exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
FAI_Proxy	,212	,714	,088	1	,767	1,236	,305	5,006
Age	-,042	,082	,264	1	,608	,959	,816	1,126
Sex	1,179	1,118	1,113	1	,291	3,253	,364	29,089
Education	-,121	,153	,633	1	,426	,886	,657	1,194
TMT-B	-,001	,009	,009	1	,925	,999	,983	1,016
VSRT LP	-,214	,094	5,128	1	,024	,808	,671	,972
Stroop Test Time Words	,030	,029	1,073	1	,300	1,031	,973	1,092
5 Points Test Perseverations	,023	,168	,020	1	,889	1,024	,737	1,422
Constant	8,096	9,254	,765	1	,382	3280,983		

Note. R_2 *= .374 (Cox & Snell), .548 (Nagelkerke); Model* $\chi 2 = 23.389$, *p=.003;* FAI_Proxy, the Forgetfulness Assessment Inventory Caregiver's Rating; TMT-B, Trail Making Test Version B; VSRT LP, Verbal Selective Reminding Test Learning Performance.

Table 9

Logistic Regression for FAI_Proxy at Step 8

							95% C.I.for Exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
VSRT LP	-,174	,057	9,321	1	,002	,840	,751	,940
Constant	5,214	1,942	7,209	1	,007	18,758		

Note. R2 = .286 (Cox & Snell), .419 (Nagelkerke); Model χ2 = 16.830, p<.001; FAI_Proxy, the Forgetfulness Assessment Inventory Caregiver's Rating; VSRT LP, Verbal Selective Reminding Test Learning Performance.

As can be seen in Table 8 and Table 9, only one predictor, the VSRT Learning Performance, was significant (p = .002). Scoring one point less in the VSRT Learning Performance subtest decreased the odds of conversion to dementia by 0.84 [0.75; 0.94]. Hence, we confirmed our second hypothesis (H2), stating that the caregivers' rating of patients' memory performance assessed with FAI cannot predict the conversion to dementia.

For the results of the logistic regression for prediction of conversion to dementia using the z-scores of FAI, see Table 10.

Table 10

Logistic regression for FAI_z, the Enter method

							95% C.I.for Exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
FAI_z	,292	,256	1,305	1	,253	1,339	,811	2,211
Age	-,037	,037	,998	1	,318	,964	,897	1,036
Sex	,773	,590	1,716	1	,190	2,166	,682	6,881
Education	,105	,070	2,234	1	,135	1,110	,968	1,274
TMT-B	,007	,005	2,170	1	,141	1,007	,998	1,017
VSRT LP	-,199	,041	23,471	1	,000	,819	,756	,888
Stroop Test Time Words	,035	,017	4,366	1	,037	1,036	1,002	1,071
5 Points Test Perseverations	,032	,050	,402	1	,526	1,032	,936	1,139
Constant	3,725	3,484	1,143	1	,285	41,453		

Note. R_2 *= .249 (Cox & Snell), .531 (Nagelkerke); Model* $\chi 2 = 83.709$, $p<.001$; FAI_z, the Forgetfulness Assessment Inventory z scores; TMT-B, Trail Making Test Version B; VSRT LP, Verbal Selective Reminding Test Learning Performance.

We did not find a significant result for FAI_z as our predictor of conversion to dementia. Therefore, we confirmed our H3, that SMC measured with FAI z-score cannot predict conversion to dementia.

Table 11

Logistic Regression for FAI z-scores at Step 5

							95% C.I.for EXP(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
VSRT LP	-,179	,036	25,025	1	,000	,836	,780	,897
Stroop Test Time Words	,040	,014	8,823	1	,003	1,041	1,014	1,069
Constant	-2,144	1,604	1,786	1	,181	8,534		

Note. R2 = .231 (Cox & Snell), .494 (Nagelkerke); Model $\chi 2 = 76,847$, $p<.001$; FAI_z, the Forgetfulness Assessment Inventory z-scores; VSRT LP, Verbal Selective Reminding Test Learning Performance.

Significant results were found for two objective measures of memory performance, the VSRT Learning Performance ($p < .001$) and Stroop Test Time Words ($p = .003$). With each point reached on VSRT Learning Performance the Odds decrease by 0.84 [0.78; 0.90]. Each second longer on the Stroop Test Time Words, increased the odds by 1.04 [1.02; 1.07].

Next, we conducted the Receiver-Operating-Characteristics (ROC) analysis of FAI to evaluate the predictive power of the questionnaire. We investigated its predictive value for the self- and caregiver-rating of memory performance, as well as for standardized z-scores calculated on the base of self-rating.

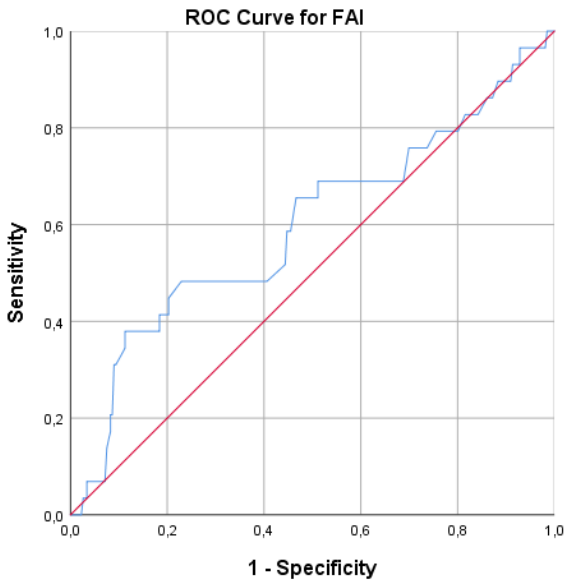

Figure 1. ROC Curve for FAI and conversion to AD

The ROC Curve for the FAI self-rating (FAI) is presented in the Figure 1. The conversion to dementia was set as the positive condition. The AUC for the FAI was .60 [.48, .72], and the chosen Cut-off score was 3.159. With that Cut-off score, FAI showed sensitivity of .48 [.30, .68], specificity of .67 [.61, .73], PPV of .14 [.09, .20], NPV of .92 [.89, .95], LR+ of 1.45 [0.97, 2.22] and LR- of 0.78 [0.54, 1.11]. The results of the ROC analysis are shown in Table 12.

Table 12

AUC Analysis for FAI

			Asymptotic 95% Confidence Interval	
Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
,599	,062	,081	,476	,721

Note. FAI, the Forgetfulness Assessment Inventory Self Rating,

The AUC for caregiver's rating, FAI_Proxy, was .65 [.46, .85]. Using the Cut-off score of 3.565, sensitivity of .54 [.25, .81], specificity of .81 [.65, .92], PPV of .5 [.30, .70], NPV of .83 [.73, .90], LR+ of 2.84 [1.23, 6.56] and LR- of 0.57 [0.31, 1.04] was reached. The ROC Curve for caregiver' assessment of patients' cognitive decline can be seen in Figure 2.

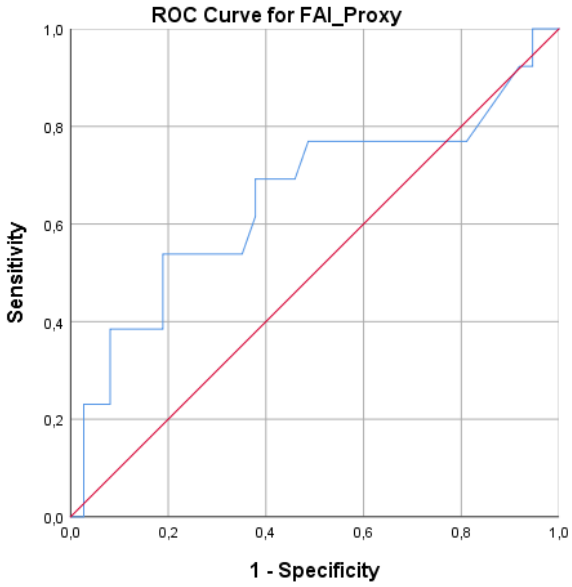

Figure 2. ROC Curve for FAI_Proxy and conversion to AD

The results of the ROC analysis for FAI_Proxy are shown in Table 13.

Table 13

AUC Analysis for FAI_Proxy

			Asymptotic 95% Confidence Interval	
Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
,653	,099	,104	,459	,847

Note. FAI_Proxy, the Forgetfulness Assessment Inventory Caregiver's Rating.

The analysis of AUC of FAI z-scores (FAI_z) followed. As for the raw scores of FAI and FAI_Proxy, the conversion to dementia was set as the positive condition. The ROC Curve for the FAI z-scores is presented in Figure 3.

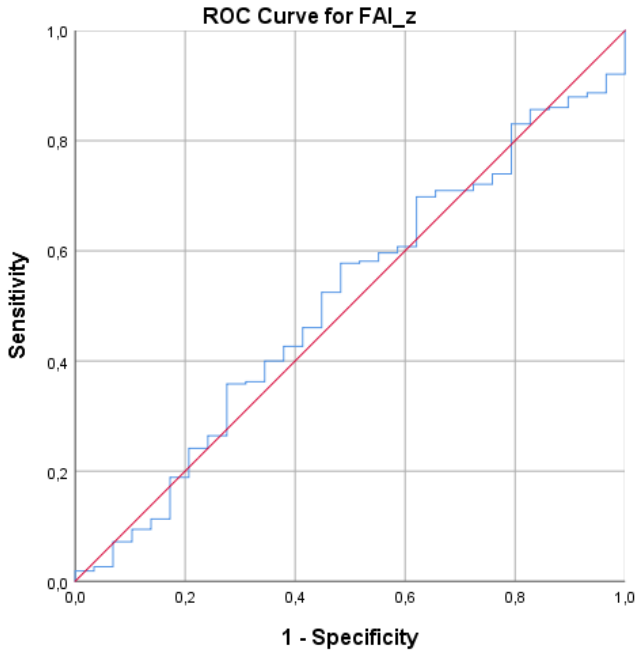

Figure 3. ROC Curve for FAI z-scores and conversion to AD

The AUC for the z-scores of subjective memory complaints assessed with FAI was .58 [.45, .70]. The Cut-off score was -0.627. Based on that, sensitivity reached .45 [.27, .64], specificity .56 [.50, .62], PPV .11 [.07, .15], NPV of .90 [.86, .93], LR+ of 1.02 [0.66, 1.56], and LR- of 0.99 [0.70, 1.39]. The results of the ROC analysis for FAI z-scores are shown in Table 14. Having implemented the criteria of Jaeschke et al. (1994), we found scarce diagnostic evidence for the questionnaire.

Table 14

AUC Analysis for FAI_z

			Asymptotic 95% Confidence Interval	
Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
,508	,056	,891	,398	,617

Note. FAI_z, the Forgetfulness Assessment Inventory z-scores.

The results of AUC analysis for FAI_Self, FAI_Proxy and FAI_z show poor discriminative power of the test to predict the conversion to dementia. The summary of the ROC analysis of FAI, FAI_Proxy and FAI_z can be found in Table 15.

Table 15

Results of ROC analysis with calculated sensitivity, specificity, PPV, PPV, LR+ and LR- at the chosen Cut-off scores for FAI, FAI_Proxy and FAI_z

	Cut-off	SE	SP	PPV	NPV	LR+	LR-	AUC
FAI	3.159	.48 [.30, .68]	.67 [.61, .73]	.14 [.09, .20]	.92 [.89, .95]	1.45 [0.97, 2.22]	0.78 [0.54, 1.11]	.60 [.48, .72]
FAI_Proxy	3.565	.54 [.25, .81]	.81 [.65, .92]	.50 [.30, .70]	.83 [.73, .90]	2.84 [1.23, 6.56]	0.57 [0.31, 1.04]	.653 [,459; ,847]
FAI_z	-0.627	.44 [.25, .64]	.56 [.50, .63]	.11 [.07, .15],	. 90 [.86, .93]	1.02 [0.66, 1.56]	0.99 [0.70, 1.39]	.508 [.398, .617]

Note. FAI, the Forgetfulness Assessment Inventory; FAI_Proxy, the Forgetfulness Assessment Inventory Caregivers' Rating; FAI_z, the Forgetfulness Assessment Inventory z-scores; *SE* Sensitivity; *SP* Specificity; *PPV,* Positive predicted value; *NPV,* Negative predicted value; *LR+,* Positive likelihood ratio; *LR-,* Negative likelihood ratio; *AUC,* Area under curve.

4.3. Correlation analysis

To see to what extent FAI, FAI_Proxy and FAI_z correlate with demographical variables and objective cognitive performance measured by NTBV, we conducted Spearman correlation analyses, as our data was not normally distributed. Included variables were age, sex, education, subsets of NTBV, BDI and WST. For interpreting the correlation coefficients Cohen's criteria will be used, where $r > .1$ indicates a small effect, $r > .3$ a moderate effect, and $r > .5$ a strong effect (Cohen, 1988). The results of the correlation analysis are presented in Table 16.

Firstly, we looked how the FAI, FAI_Proxy and FAI_z correlate with each other at baseline and follow-up. FAI and FAI_z correlated strongly with each other at baseline, as the z-score is calculated based on the raw-score. The FAI baseline score correlated significantly with the FAI follow-up score and the FAI_z at the follow-up. The effect size for the later decreased however to small. FAI_Proxy did

not correlate significantly with the FAI and FAI_z neither at baseline nor follow-up. The caregivers' rating at baseline showed a correlation of moderate size with the rating at follow-up.

For patients' SMC (FAI) at baseline, we found 12 significant positive correlations. Higher SMC correlated with the higher age of the participant and more depressive symptoms. For the NTBV subsets, we also found a positive correlation for FAI with the Interference Time, TMTB, Concentration I, Symbols Counting, TMTA, Stroop Test Time Words and Stroop Test Time Colors. Moreover, we found 12 significant negative correlations for NTBV subtests, indicating, that for Interference Total/Time, ACT Time, ACT Sum, Digit-Symbol Test, SVF Correct, PVF Correct, whole VSRT, SVF Correct, Stroop Test Total/Time, 5 Point Test Correct and BNT, the lower objective memory performance correlated with higher subjective memory complaints.

At the follow-up, patients' SMC correlated significantly with many subtests of NTBV. Apart from positive correlations with Interference Time, TMTB, Concentration I, TMTA and the Stroop Test Interference Stroop Test Time Words and Stroop Test Time Colors observed already at baseline, we found positive correlations for the interference between TMT-B and TMT-A (TMT-B –TMT-A). Along with the negative correlations at baseline for subtests of NTBV being Interference Total/Time, ACT Sum, Digit-Symbol Test, SVF Correct, PVF Correct, whole VSRT, SVF Correct, Stroop Test Total/Time and 5 Point Test Correct, we observed 2 additional ones for BNT. The higher BDI-II score correlated with higher SMC. Moreover, patients' lower WST score indicating a lower IQ, correlated with higher SMC.

Table 16

Spearman's Correlation Coefficients (r_s) between FAI, FAI_z, FAI_Proxy, demographical variables and objective performance measures at Baseline and Follow-up

	Baseline			Follow-up		
	Self	**Proxy**	**z**	**Self**	**Proxy**	**z**
FAI	1.000	-.072	-.211**	.656**	.135	-.282**
FAI_Proxy	-.072	1.000	. 219	.149	.471**	.147
FAI_z	-.264**	,219	1.000	-.211**	-.020	.642**
Sex	-.029	-.036	-.030	.091	-.063	-.024
Age	.155**	.123	-.085	.240**	.021	-.176*
Education	-.042	-.079	-.018	-.146*	.210	-.043
WST	-.066	-.217	,020	-.125*	.091	-.018
BDI-II	.336**	.023	-,107	.393**	-.193	-.185**
Interference Time (c.I.)	.211**	.209	-.082	.217**	.097	-.165*
Interference Total/Time (c.I.)	-.213**	-.179	.078	-.220*	-.063	.152*
TMT-B	.170**	.120	-.131*	.190**	.009	-.106
AKT Time	.141*	.044	-.010	.248**	-.028	-.066
AKT Sum	-.151**	-.077	.023	-.254**	.023	.067
Symbols Counting (c.I.)	.108	.129	-.084	.108	.340*	-.087
TMT-B – TMT-A	.140*	.067	-.114	.152*	.030	-.088
Digit-Symbol Test	-.196**	-.070	.045	-.273**	-.001	.028
TMT-A	.166**	.123	-.072	.192**	-.009	-.108
PVF Correct	-.167**	.018	,073	-.228**	.252	.158*
VSRT Learning Performance	-.247**	-.347*	.147*	-.253**	-.181	.116
VSRT Delayed Recall	-.217**	-.193	.123*	-.200**	-.172	.094
VSRT Word Span	-.160**	-.181	.195**	-.236**	-.129	.140*
VSRT Recognition	-.180**	-.188	.114	-.215**	-.276	.103
SVF Correct	-.293**	.016	.143*	-.160*	-.038	.169**
Stroop Test Interference	.112	-.039	-.006	.352**	-.031	-.057
Stroop Test Time Words	.196**	-.006	-.031	.164**	.000	-.103
Stroop Test Time Colors	.231**	.149	-.029	.170**	.040	-.101
Stroop Test Total/Time	-.151**	-.031	-.025	-.158**	-.017	.037
Maze Time	.100	-.033	-.018	.109	.023	-.036
Maze Time Total/Time	-.079	-.016	.005	-.099	-.051	.036
5 Point Test Correct	-.223**	-.013	.025	-.329**	.066	.001
5 Point Test Perseverations	-.091	.067	-.038	-.062	.200	.034
BNT	-.139*	-.008	.115*	-.145*	-.021	.102

Note. FAI, the Forgetfulness Assessment Inventory Self Rating; FAI_Proxy, the Forgetfulness Assessment Inventory Caregiver's Rating; FAI_z, the Forgetfulness Assessment Inventory z-scores; WST, Wortschatztest,

Standardized Vocabulary Test; BDI, Beck Depression Inventory; c.I., Cerebral Insufficiency Test; TMT-B, Trail Making Test Version B; AKT Time, Age Concentration Test Time; AKT Sum, Age Concentration Test Sum; TMT-B – TMT-A, Interference, Difference between scores of TMT-B and TMT-A; TMT-A, Trail Making Test Version A; PVF, phonematic Verbal Fluency; VSRT, Verbal Selective Reminding Test; SVF, Semantic Verbal Fluency; significance level: *p<.05. **p<.01, (two-tailed).

For FAI_Proxy, the caregiver's assessment of patient's memory performance we found only one significant negative correlation for the NTBV subtest VSRT Learning Performance at baseline. At the follow-up, the VSRT Learning Performance no longer showed a moderate negative correlation. In addition, a higher score at Symbols Counting correlated positively with higher caregivers' rating of patients' performance.

The FAI_z correlated negatively with BDI-II at both testing points, suggesting that patients complaining of SMC often report depressive symptoms. We found significant correlations with objective memory tests at baseline for subtest of NTBV such as TMT-B, VSRT Word Span, VSRT Delayed Recall, VSRT Learning Performance SVF Correct and BNT. At the follow-up, we found significant correlations for Interference Time, Interference Total/Time, PVF Correct, WSRT Word Span and SVF Correct. For more details, see Table 16.

4.4. Reliable Change Analysis

Reliable change was analyzed to examine the changes in performance across time for each diagnostic group (SCD, aMCI, naMCI) that are not due to cognitive impairments. We calculated the mean, variance, standard deviation for 45 healthy controls (HC) and the confidence interval for both baseline and follow-up scores. At baseline the HC group mean was 2.27, standard deviation was .66. At the follow-up the HC mean was 2.15 and standard deviation .61. We assessed the test-retest-reliability. The correlation coefficient was .624. The scores of the HC group are displayed in Table 17.

Table 17

FAI characteristics for the healthy control, SCD, aMCI and naMCI group at baseline and follow-up

		Sample size	Baseline		Follow-up		Direction of dp*	r_{xy}
			Mean	SD	Mean	SD		
HC	FAI	45	2.27	0.66	2.15	0.61	$S_y<S_x$	.423
	FAI_z	45	0.03	1.04	.28	0.97	$S_y<S_x$	.421
SCD	FAI	37	2.72	0.69	2.69	0.71	$S_y>S_x$	.471
	FAI_z	37	-0.80	1.24	-0.73	1.19	$S_y<S_x$	.782
naMCI	FAI	119	2.94	0.74	2.98	0.67	$S_y<S_x$	.623
	FAI_z	119	-0.76	1.13	-0.70	1.01	$S_y<S_x$	.505
aMCI	FAI	94	3.07	0.75	2.90	0.72	$S_y<S_x$	.692
	FAI_z	94	-.94	1.12	-0.76	1.17	$S_y>S_x$	.632

Note. FAI, Forgetfulness Assessment Inventory; SD_x, standard deviation at baseline; SD_y, standard deviation at follow-up; dp, differential practice; r, the test-retest reliability coefficient; HC, Healthy Control Group, SCD, Subjective Cognitive Decline; naMCI, nonamnestic Mild Cognitive Impairment; aMCI, amnestic Mild Cognitive Impairment.

Due to a sample size of the control group smaller than 50, we used the t value instead of a z-score to interpret the results. At 90% confidence level the t value considered as reliable change is ± 1.684.

Before presenting the results of different methods assessing reliable change, it is important to underline, that the higher the FAI score of the patient is, the more severely she or he experiences the subjective memory decline, therefore an improvement in SCD is indicated by lower FAI score.

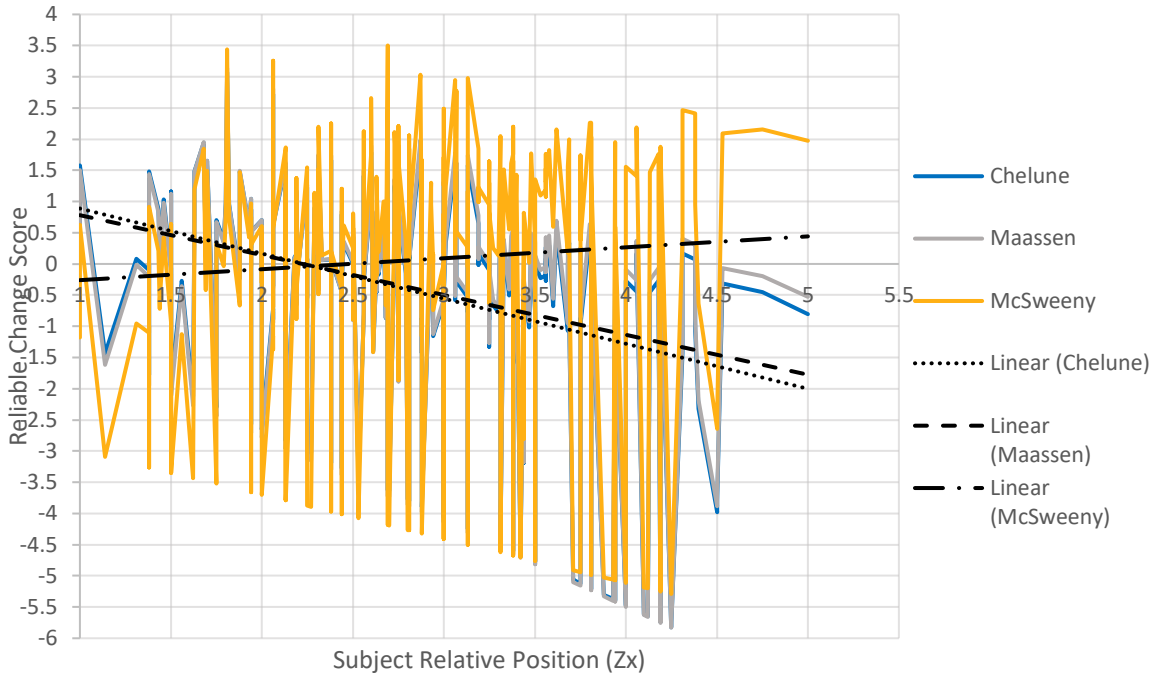

Figure 4. Reliable Change line plots for Subjective Memory Complaints (FAI)

The Figure 4. presents the responsiveness of the model McSweeny et al. (1993), Chelune et al. (1993) and Maasen et al. (2006).

Table 18 shows the classifications of patients in the diagnostic groups and the differences between the methods for determining reliable change.

Table 18

Comparison of methods assessing reliable change between different diagnostic groups for FAI

	SCD N = 37			naMCI N = 118			aMCI N = 94		
	↑	→	↓	↑	→	↓	↑	→	↓
Chelune	3 (8.1)	33 (89.2)	1 (2.7)	20 (17.0)	97 (82.2)	1 (0.8)	15 (16.0)	76 (80.8)	3 (3.2)
McSweeny	3 (8.1)	27 (73.0)	7 (18.9)	20 (17.0)	89 (76.4)	9 (7.6)	16 (17.0)	58 (61.7)	20 (21.3)
Maassen	3 (8.1)	33 (89.2)	1 (2.7)	20 (17.0)	96 (81.3)	2 (1.7)	16 (17.0)	74 (78.7)	4 (4.3)

Note. N (Percentage); SCD, Subjective Cognitive Decline; naMCI, nonamnestic Mild Cognitive Impairment; aMCI, amnestic Mild Cognitive Impairment.

Comparing the reliable change in the different diagnostic groups, SCD group showed a decline between 2.7% and 18.9%, aMCI group deteriorated in 3.2% to 21.3%, the naMCI group in 0.8% to 7.6% of cases. All models showed a similar pattern in detection of reliable change. The regression-based model by McSweeny et al. (1993) was the most responsive to negative change.

When comparing the non-converters and converters, the regression-based model by McSweeny et al. (1993) was most responsive to negative change.

Table 19

Comparison of methods assessing reliable change between converters and non-converters for FAI

	Non-converters N = 265			Converters N = 29		
	↑	→	↓	↑	→	↓
Chelune	33 (12.5)	226 (85.3)	6 (2.2)	6 (20.7)	23 (79.3)	0 (0.0)
McSweeny	36 (13.6)	191 (72.1)	38 (14.3)	6 (20.7)	17 (58.6)	6 (20.7)
Maassen	34 (12.8)	221 (83.4)	10 (3.8)	6 (20.7)	23 (79.3)	0 (0.0)

Note. N (Percentage).

For non-converters, a decline was observed in 2.2% to 14.3% cases. Between 0.0% and 20.7% of converters deteriorated over time, as can be seen in Table 19.

For FAI_z we also looked at the responsiveness of models of McSweeny et al. (1993), Chelune et al. (1993) and Maasen et al. (2006). The results are presented in the Figure 5. The selected models show a very similar pattern in detection of both positive and negative reliable change.

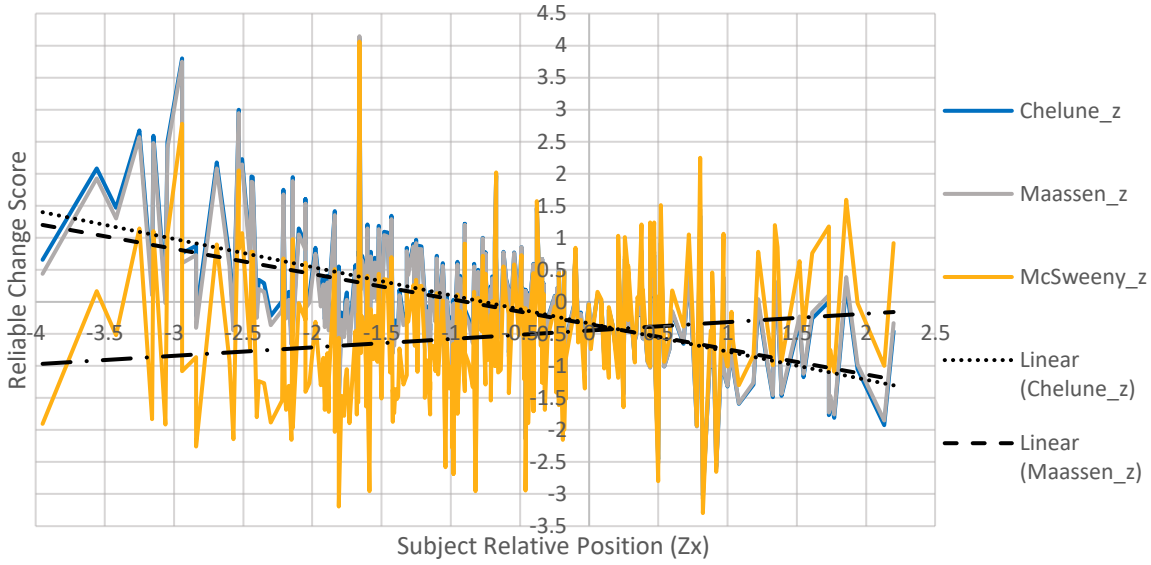

Figure 5. Reliable Change line plots for FAI_z

The Table 20 shows the classifications of patients for the diagnostic groups of SCD, naMCI and aMCI according to their decline, stability or improvement classified with the three methods determining reliable change. The reliable change for the SCD group showed a decline between 2.7% and 21.6%, aMCI group deteriorated in 1.1% to 14.9%, the naMCI group in 5.1% to 9.3.1% of cases. All models showed a similar pattern in detection of reliable change. However, the regression-based model by McSweeny et al. (1993) was the most responsive to negative change, and less responsive to positive change for the aMCI and naMCI.

Table 20

Comparison of methods assessing reliable change between different diagnostic groups for FAI_z

	SCD N = 37			naMCI N = 118			aMCI N = 94		
	↑	→	↓	↑	→	↓	↑	→	↓
Chelune	1 (2.7)	35 (94.6)	1 (2.7)	7 (5.9)	105 (89.0)	6 (5.1)	5 (5.3)	88 (93.6)	1 (1.1)
McSweeny	0 (0.0)	29 (78.4)	8 (21.6)	2 (1.7)	105 (89.0)	11 (9.3)	1 (1.1)	79 (84.0)	14 (14.9)
Maassen	1 (2.7)	35 (94.6)	1 (2.7)	7 (5.9)	105 (89.0)	6 (5.1)	4 (4.3)	89 (94.6)	1 (1.1)

Note. N (Percentage); SCD, Subjective Cognitive Decline; naMCI, nonamnestic Mild Cognitive Impairment; aMCI, amnestic Mild Cognitive Impairment.

As for the FAI, we also looked at the differences for FAI_z between non-converters and converters. For non-converters, the most responsive model to negative change was the regression-based model by McSweeny et al. (1993). For the converters group, the model of Maasen et al. (2006) showed a slightly higher responsiveness to decline. From 3.0% to 11.3% non-converters and from 3.5% to 13.7% converters deteriorated over time, as can be seen in Table 21.

Table 21

Comparison of methods assessing reliable change between converters and non-converters for FAI_z

	Non-converters N = 265			**Converters** N = 29		
	↑	→	↓	↑	→	↓
Chelune	12 (4.5)	245 (92.5)	8 (3.0)	2 (6.9)	26 (89.6)	1 (3.5)
McSweeny	4 (1.5)	231 (87.2)	30 (11.3)	0 (0.0)	26 (89.6)	3 (10.4)
Maassen	12 (4.5)	245 (92.5)	8 (3.0)	1 (3.5)	24 (82.8)	4 (13.7)

Note. N (Percentage).

5. Discussion

The present study investigated the prognostic value of subjective memory complaints assessed with the Forgetfulness Assessment Inventory in prediction of conversion to dementia in a clinical context in patients with SCD and MCI. With the progressing aging of society, successful prediction of dementia conversion has a high clinical relevance, as it allows a timely start of possible therapy or an implementation of cognitive training, which had showed a positive influence on well-being in patients experiencing SMC (Pereira-Morales et al., 2018).

We found support for our hypothesis that neither subjective memory complaints (raw score and z-score) nor the caregivers' assessment of patients' memory performance measured with the FAI can predict conversion to dementia. The FAI showed a poor discrimination value for both patients' and caregivers' rating (AUC of .60 and .65 respectively). A similar (AUC of .62) predictive validity was found in another study of Lehrner et al. (2016). The predictive value was even lower for the FAI z-score (AUC of .43). However, in order to measure the predictive validity of SMC, assessing the complaints with a questionnaire with a higher predictive validity or collecting a higher number of caregivers' evaluations would be desirable in the future in order to reach a more conclusive result.

The exploratory analysis of correlations between SMC and other variables showed that older patients report SMC both at baseline and follow-up. Patients with more depressive symptoms showed higher FAI scores, linking the SMC to depression, as previously had been reported in other studies (Lehrner et al., 2014; Reid & MacLullich, 2006). In contrast to the study of Jonker et al. (2000), we did not observe a negative correlation between SMC and patients' length of education at baseline, only at follow-up. However, the effect size was small. SMC correlated with many objective measures of memory deficits as well both at baseline and follow-up, which is in line with researchers stating that SMC are a reflection of patient's cognitive decline (Jessen et al., 2020a) and that they are frequently

present already at first stage of SCD (Jessen et al., 2014). Interestingly, looking at the correlations between the caregivers' assessment and patients' performance, we found only one correlation with an objective measure of patients' cognitive state at baseline and two at follow-up. A small sample size of the caregivers' group might had been an influence factor here. Nonetheless, most correlations coefficients found in our study were according to Cohen's criteria (1988) small to moderate.

As in our study we were looking at change in repeated testing situation, in order to assessing change in the neuropsychological context of cognitive decline, we lastly analyzed and compared three methods by Chelune et al. (1993), McSweeny et al. (1993) and Maassen et al. (2006). In general, our study showed a superiority of regression-based models in assessing reliable change as postulated in previous studies (Hinton-Bayre, 2010). In our study, the most responsive model detecting reliable change was the model by McSweeny et al. (1993). For patients with Parkinson's Disease, the model by Maassen et al. (2006) showed the highest responsiveness both to positive and negative change (Großbichler, 2020). The difference in responsiveness may be explained due to a higher variance observed at baseline than at follow-up and a test-retest reliability < .7 in our study (Hinton-Bayre, 2016).

The conversion rates in our study reached 9.7% (29 out of 299 patients). The observed rates are similar to those reported in prior studies (Lehrner et al., 2016; Silva et al., 2016). In general, observed conversion rates support the three-stage model of dementia progression, with SCD followed by MCI and dementia diagnosis (Jessen et al., 2010, 2014). None of the SCD patients converted to dementia. Only patients diagnosed with MCI, (2 patients: naMCI → AD; 27 patients: aMCI → AD), converted to dementia at the follow-up. Those numbers are in line with the findings previously reported, suggesting that the aMCI patients are at highest risk of conversion to dementia (Lehrner et al., 2005). In our sample 35 cases had regressed from MCI to SCD (27 patients: naMCI → SCD; 8 patients aMCI → SCD). Possible reasons for that regression can be low motivation level during the baseline testing or other confounding variables.

5.1. Limitations

We would like to point out some limitations of the present study. First of all, the sampling and exclusion criteria implemented in the current study can be seen as a possible limitation. A preselection took already place before the testing, as all subjects were either referred to the clinic to take part in the neuropsychological testing by a doctor or underwent the testing voluntarily. Therefore, our study sample is clinical, not probabilistic, and the results cannot be generalized to the general population. The memory clinic setting shows a higher risk of conversion to dementia for patients with SCD compared to recruitment in community setting (Slot et al., 2019; Rodríguez-Gómez et al., 2015).

Secondly, the whole neuropsychological testing took approximately two to three hours, which can be considered another limitation. Some patients reported being tired or seemed disinterested close to the end of the examination. As the FAI was administered after the NTBV test battery, which requires high level of attention and concentration from the patients, the patients could have possibly been less motivated to reflect on their memory performance. The testing situation was a stressful event for some patients as well, which might have influenced the test results.

Moreover, one of the study aims was to assess, whether a caregivers' assessment of patients' memory performance can be a better predictor than self-reported subjective memory complaints. Yet due to the small sample size of the caregivers that filled in the FAI, the statistical power of our study suffered. Future studies could aim to collect more caregivers' assessments to be able to answer the research question more reliably.

The FAI score is calculated as the average score of answered questions. There are 16 questions all together, it was however possible for patients or caregivers to omit some questions when completing the paper-pencil version of the FAI. The scores might have therefore been influenced through the missing answers. It would

be interesting to analyze, what items were more likely omitted and whether it could have influenced the results.

5.2. Conclusion and future prospects

In conclusion, our study showed that SMC assessed with the FAI are not a reliable predictor for conversion to dementia, neither being assessed by the patient nor by their caregiver. A z-score based on the FAI score showed even a smaller predictive value.

In case of patients complaining of SMC, a broad and individually adapted diagnostic process should be implemented in order to identify possible medical conditions underlying the first stage of dementia, the SCD. Especially, while the pharmacological treatments are still not efficient and require further testing (Jessen, 2020). Also a progressive value of biomarkers in combination with screening methods as MMSE and other objective neuropsychological tests in prediction of conversion to dementia from MCI should be considered (Frölich et al., 2017).

As Mitchell and Shiri-Feshki (2008) showed that the risk of conversion to dementia decreases over time, it is crucial to suggest a follow-up examination to patients consulting their cognitive state and reporting memory deficits. For patients reporting SMC diagnosed with SCD, Wolfsgruber et al. (2016) suggested certain need for differentiation between stable and unstable SCD. Repeatedly expressed memory complaints and consistent SCD diagnosis was namely associated with increased risk for conversion to dementia.

Although subjective memory complaints are commonly experienced by patients, they have a lower predictive power compared to objective memory complaints while predicting conversion to dementia. However, patients at early dementia stages often express SMC. As SCD is a reliable predictor of further cognitive decline and have been linked to an increased risk of conversion to both MCI and

dementia (Jessen et al., 2020), SMC should not be ignored or underrated during the diagnostic process.

6. References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

World Health Organization (2019). International statistical classification of diseases and related health problems (11th ed.). https://icd.who.int/

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2016). *Multivariate Analysemethoden*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bondi, M. W., Houston, W. S., Salmon, D. P., Corey-Bloom, J., Katzman, R., Thal, L. J., & Delis, D. C. (2003). Neuropsychological deficits associated with Alzheimer's disease in the very-old: Discrepancies in raw vs. standardized scores. *Journal of the International Neuropsychological Society*, *9*(5), 783–795. https://doi.org/10.1017/S1355617703950119

Bortz, J., & Döring, N. (2007). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler: Limitierte Sonderausgabe.* Springer-Verlag.

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Springer.

Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, *7*(1), 41.

Cohen, J. (1988). Statistical power analysis for the social sciences.

Crumley, J. J., Stetler, C. A., & Horhota, M. (2014). Examining the relationship between subjective and objective memory performance in older adults: A meta-analysis. *Psychology and Aging*, *29*(2), 250.

Dementia, (19.09.2019). Retrieved on 29.04.2020 from https://www.who.int/news-room/fact-sheets/detail/dementia

Dilling, H., Mombour, W., Schmidt, M. H., Schulte-Markwort, E., & Remschmidt, H. (2010). ICD-10 Kapitel V (F). Klinisch-diagnostische Leitlinien. *Internationale Klassifikation psychischer Störungen*, *3*.

DIMDI (Hrsg.). (2013). Systematisches Verzeichnis Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision. http://www.icd-code.de/icd/code/ICD-10-GM-2013.html

Duff, K., Callister, C., Dennett, K., & Tometich, D. (2012). Practice effects: A unique cognitive variable. *Clinical Neuropsychologist*, *26*(7), 1117–1127. https://doi.org/10.1080/13854046.2012.722685

Edmonds, E. C., Delano-Wood, L., Galasko, D. R., Salmon, D. P., & Bondi, M. W. (2014). Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. *Journal of the International Neuropsychological Society*, *20*(8), 836–847. https://doi.org/10.1017/S135561771400068X

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Field, A. (2013). Discovering statistics using SPSS (3. Edition). *NY: Sage Publications*.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Frölich, L., Peters, O., Lewczuk, P., Gruber, O., Teipel, S. J., Gertz, H. J., ... & Hüll, M. (2017). Incremental value of biomarker combinations to predict progression of

mild cognitive impairment to Alzheimer's dementia. *Alzheimer's research & therapy*, *9*(1), 84.

Gatterer, G. (1990). Alters-Konzentrations-Test (AKT) [Age Concentration Test]. *Göttingen: Hogrefe*.

Gollwitzer, M., Eid, M., & Schmitt, M. (2013). *Statistik und Forschungsmethoden*. Beltz Verlagsgruppe.

Goldberg, S. M., Lopez, O. L., Cohen, A. D., Klunk, W. E., Aizenstein, H. A., Mizuno, A., & Snitz, B. E. (2020). The roles of study setting, response bias, and personality in subjective memory complaints of cognitively normal older adults. *International Psychogeriatrics*, 1-12.

Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.

Großbichler, A. (2020). *Psychosocial Outcome after Deep Brain Stimulation of patients with Parkinson's Disease – a retrospective study* (Master's thesis, Medical University of Vienna).

Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, *4*(2), 627–635.

Hautzinger, M., Keller, F., & Kühner, C. (2006). BDI-II Beck-Depressions-Inventar. *Auflage Harcourt Test Services, Frankfurt/Main*.

Hinton-Bayre, A. D. (2010). Deriving reliable change statistics from test-retest normative data: Comparison of models and mathematical expressions. *Archives of Clinical Neuropsychology*, *25*(3), 244–256. https://doi.org/10.1093/arclin/acq008

Hinton-Bayre, A. D. (2016). Clarifying discrepancies in responsiveness between reliable change indices. *Archives of Clinical Neuropsychology*, *31*(7), 754-768.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp.

Jacobson, N. S., & Truax, P. (1992). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research.

Jaeschke, R., Guyatt, G. H., Sackett, D. L., Guyatt, G., Bass, E., Brill-Edwards, P., ... & Haynes, B. (1994). Users' Guides to the Medical Literature: III. How to Use an Article About a Diagnostic Test B. What Are the Results and Will They Help Me in Caring for My Patients?. *Jama*, *271*(9), 703-707.

Jessen, F., Wiese, B., Bachmann, C., Eifflaender-Gorfer, S., Haller, F., Kölsch, H., Luck, T., Mösch, E., Van Den Bussche, H., Wagner, M., Wollny, A., Zimmermann, T., Pentzek, M., Riedel-Heller, S. G., Romberg, H. P., Weyerer, S., Kaduszkiewicz, H., Maier, W., & Bickel, H. (2010). Prediction of dementia by subjective memory impairment effects of severity and temporal association with cognitive impairment. *Archives of General Psychiatry*, *67*(4), 414–422. https://doi.org/10.1001/archgenpsychiatry.2010.30

Jessen, F., Amariglio, R. E., Van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., Van Der Flier, W. M., Glodzik, L., Van Harten, A. C., De Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., … Wagner, M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's and Dementia*, *10*(6), 844–852. https://doi.org/10.1016/j.jalz.2014.01.001

Jessen, F., Wolfsgruber, S., Wiese, B., Bickel, H., Mösch, E., Kaduszkiewicz, H., ... & Weyerer, S. (2014a). AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. *Alzheimer's & Dementia*, *10*(1), 76-83.

Jessen, F. (2020). Pharmacological prevention of cognitive decline and dementia. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*.

Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., ... & Sikkes, S. A. (2020a). The characterisation of subjective cognitive decline. *The Lancet Neurology*.

Jonker, C., Jonker, M. I., & Schmand, B. (2000). Are memory complaints predictive for dementia? A review of clinical and population-based studies. *International Journal of Geriatric Psychiatry*, *15*(11), 983–991. https://doi.org/10.1002/1099-1166(200011)15:11<983::AID-GPS238>3.0.CO;2-5

Lehrl, S., & Fischer, B. (1997). Kurztest für cerebrale Insuffizienz. *Göttingen: Hogrefe*.

Lehrner, J., Gufler, R., Guttmann, G., Maly, J., Gleiß, A., Auff, E., & Dal-Bianco, P. (2005). Annual conversion to Alzheimer disease among patients with memory complaints attending an outpatient memory clinic: the influence of amnestic mild cognitive impairment and the predictive value of neuropsychological testing. *Wiener Klinische Wochenschrift*, *117*(18), 629-635.

Lehrner, J., Gleiß, A., Maly, J., Auff, E., & Dal-Bianco, P. (2006). The Verbal Selective Reminding Test (VSRT). *NEUROPSYCHIATRIE-MUNCHEN-DEISENHOFEN-*, *20*(3), 204.

Lehrner, J., Moser, D., Klug, S., Glei, A., Auff, E., Dal-Bianco, P., & Pusswald, G. (2014). Subjective memory complaints, depressive symptoms and cognition in patients attending a memory outpatient clinic. *International Psychogeriatrics*, *26*(3), 463–473. https://doi.org/10.1017/S1041610213002263

Lehrner, Johann, Bodendorfer, B., Lamm, C., Moser, D., Dal-Bianco, P., Auff, E., & Pusswald, G. (2016). Subjektive Gedächtnisbeschwerden bei Patienten mit subjektiven kognitiven Beschwerden und Patienten mit leichter kognitiver Beeinträchtigung und ihre Relevanz für Früherkennung von Demenz. *Zeitschrift Fur Neuropsychologie*, *27*(2), 85–93. https://doi.org/10.1024/1016-264X/a000175

Lenehan, M. E., Klekociuk, S. Z., & Summers, M. J. (2012). Absence of a relationship between subjective memory complaint and objective memory impairment in mild cognitive impairment (MCI): Is it time to abandon subjective memory complaint as an MCI diagnostic criterion? *International Psychogeriatrics*, *24*(9), 1505–1514. https://doi.org/10.1017/S1041610212000695

Lista, S., Molinuevo, J. L., Cavedo, E., Rami, L., Amouyel, P., Teipel, S. J., ... & Zetterberg, H. (2015). Evolving evidence for the value of neuroimaging methods and biological markers in subjects categorized with subjective cognitive decline. *Journal of Alzheimer's Disease*, *48*(s1), S171-S191.

Maassen, G. H., Bossema, E. R., & Brand, N. (2006). Reliable change assessment with practice effects in sport concussion research: A comment on Hinton-Bayre. *British journal of sports medicine*, *40*(10), 829-833.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*(7), 939-939.

McSweeny, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). “T scores for change”: An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist*, *7*(3), 300-312.

Mitchell, A. J., & Shiri-Feshki, M. (2008). Temporal trends in the long term risk of progression of mild cognitive impairment: a pooled analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(12), 1386-1391.

Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(6), 439–451. https://doi.org/10.1111/acps.12336

Molinuevo, J. L., Rabin, L. A., Amariglio, R., Buckley, R., Dubois, B., Ellis, K. A., ... & Reisberg, B. (2017). Implementation of subjective cognitive decline criteria in research studies. *Alzheimer's & Dementia*, *13*(3), 296-311.

Morris, J. C., Mohs, R. C., & Rogers, H. (1989). Consortium To Establish a Registry for Alzheimer's Disease (CERAD) Clinical and Neuropsychological. *Psychopharmacology bulletin*, *24*(3-4), 641.

Mosconi, L., De Santi, S., Brys, M., Tsui, W. H., Pirraglia, E., Glodzik-Sobanska, L., Rich, K. E., Switalski, R., Mehta, P. D., Pratico, D., Zinkowski, R., Blennow, K., & de Leon, M. J. (2008). Hypometabolism and Altered Cerebrospinal Fluid Markers in Normal Apolipoprotein E E4 Carriers with Subjective Memory Complaints. *Biological Psychiatry*, *63*(6), 609–618. https://doi.org/10.1016/j.biopsych.2007.05.030

Oswald, W. D., & Fleischmann, U. M. (1997). Nürnberger-Alters-Inventar (4. Aufl.).

Perrotin, A. (2012). Subjective Cognition and Amyloid Deposition Imaging. *Archives of Neurology*, *69*(2), 223. https://doi.org/10.1001/archneurol.2011.666

Pereira-Morales, A. J., Cruz-Salinas, A. F., Aponte, J., & Pereira-Manrique, F. (2018). Efficacy of a computer-based cognitive training program in older people with subjective memory complaints: a randomized study. *International Journal of Neuroscience*, *128*(1), 1-9.

Pérès, K., Helmer, C., Amieva, H., Matharan, F., Carcaillon, L., Jacqmin-Gadda, H., ... & Dartigues, J. F. (2011). Gender differences in the prodromal signs of dementia: Memory complaint and IADL-restriction. A prospective population-based cohort. *Journal of Alzheimer's Disease*, *27*(1), 39-47.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Petersen, R. C. (2011). Clinical practice. Mild cognitive impairment. *The New England journal of medicine*, *364*(23), 2227.

Prince, M. J., Prina, M., & Guerchet, M. (2013). *World Alzheimer report 2013: Journey of caring: an analysis of long-term care for dementia*. Alzheimer's Disease International.

Prince, M., Bryce, R., & Ferri, C. (2018). World Alzheimer Report 2011: The benefits of early diagnosis and intervention.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and motor skills*, *55*(3), 839-844.

Reid, L. M., & MacLullich, A. M. J. (2006). Subjective memory complaints and cognitive impairment in older people. *Dementia and Geriatric Cognitive Disorders*, *22*(5–6), 471–485. https://doi.org/10.1159/000096295

Reisberg, B., Jamil, I. A., Khan, S., Monteiro, I., Torossian, C., Ferris, S., Sabbagh, M., Gauthier, S., Auer, S., Shulman, M. B., Kluger, A., Franssen, E., & Wegiel, J. (2010). Staging Dementia. *Principles and Practice of Geriatric Psychiatry: Third Edition*, 162–169. https://doi.org/10.1002/9780470669600.ch31

Reitan, R. M. (1979). *Trail making test: TMT*. Testzentrale.

Rodríguez-Gómez, O., Abdelnour, C., Jessen, F., Valero, S., & Boada, M. (2015, September 24). Influence of Sampling and Recruitment Methods in Studies of Subjective Cognitive Decline. *Journal of Alzheimer's Disease*. IOS Press. https://doi.org/10.3233/JAD-150189

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature Reviews Neurology*, *10*(11), 634.

Saß, H., Wittchen, H. U., Zaudig, M., & Houben, I. (2003). Diagnostische Kriterien des diagnostischen und statistischen Manuals psychischer Störungen. *DSMIV-TR. Hogrefe, Göttingen*.

Schmidt, K. H., & Metzler, P. (1992). *Wortschatztest*. Beltz.

Silva, M. R., Moser, D., Pflüger, M., Pusswald, G., Stögmann, E., Dal-Bianco, P., Auff, E., & Lehrner, J. (2016). Subjektive Gedächtnisbeschwerden und Awareness bei Patienten mit leichter kognitiver Beeinträchtigung und Patienten mit Alzheimer Demenz. *Neuropsychiatrie*, *30*(2), 103–112. https://doi.org/10.1007/s40211-016-0185-y

Slot, R. E., Sikkes, S. A., Berkhof, J., Brodaty, H., Buckley, R., Cavedo, E., ... & Lista, S. (2019). Subjective cognitive decline and rates of incident Alzheimer's disease and non–Alzheimer's disease dementia. *Alzheimer's & Dementia*, *15*(3), 465-476.

Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., Suliman, H., Wagner, M., Schild, H. H., & Jessen, F. (2010). Volume loss of the medial temporal lobe structures in subjective memory impairment. *Dementia and Geriatric Cognitive Disorders*, *29*(1), 75–81. https://doi.org/10.1159/000264630

Tewes, U. (Ed.). (1994). *Hamburg-Wechsler Intelligenztest für Erwachsene: HAWIE-R; Revision 1991*. Huber.

Wancata, J., Kaup, B., & Krautgartner, M. (2001). Projections in the incidence of dementia in Austria for the years 1951 to 2050. *Wiener klinische Wochenschrift*, *113*(5-6), 172-180.

Wang, L., Van Belle, G., Crane, P. K., Kukull, W. A., Bowen, J. D., McCormick, W. C., & Larson, E. B. (2004). Subjective memory deterioration and future dementia in people aged 65 and older. *Journal of the American Geriatrics Society*, *52*(12), 2045-2051.

Wolfsgruber, S., Kleineidam, L., Wagner, M., Mösch, E., Bickel, H., Lühmann, D., ... & Brettschneider, C. (2016). Differential risk of incident Alzheimer's disease dementia in stable versus unstable patterns of subjective cognitive decline. *Journal of Alzheimer's Disease*, *54*(3), 1135-1146.

Wolfsgruber, S., Polcher, A., Koppara, A., Kleineidam, L., Frölich, L., Peters, O., ... & Kornhuber, J. (2017). Cerebrospinal fluid biomarkers and clinical progression in patients with subjective cognitive decline and mild cognitive impairment. *Journal of Alzheimer's Disease*, *58*(3), 939-950.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32-35.

List of Figures

List of Tables

Appendix

Abstract English

Background: Subjective memory complaints (SMC) are commonly reported by patients at early stages of Alzheimer's Disease (AD). Therefore, the current study aimed to determine the predictive value of SMC in prediction of conversion to AD. Additionally, we analyzed the correlation between SMC and objective memory performance, and methods assessing reliable change in a neuropsychological context.

Methods: 299 patients including 46 controls underwent a neuropsychological examination. The SMC of patients were assessed with the Forgetfulness Assessment Inventory (FAI). Using the FAI, caregivers rated patients' memory performance within last two weeks as well. The follow-up examination took place after 12 to 48 months.

Results: At the follow-up examination, the conversion rate to AD was 9.7%. Neither SMC reported by patients nor their caregivers' assessment of patients' memory performance were a significant predictor of conversion to AD. Compared to raw scores, z-scores showed even lower predictive value predicting dementia. SMC correlated with the age, depressive symptoms and certain objective cognitive tests. The most responsive method for assessing reliable change was regression-based model of McSweeney.

Conclusion: SMC have a lower predictive power compared to objective memory complaints while predicting the conversion to AD.

Abstract Deutsch

Grundlagen: Subjektive Gedächtnisbeschwerden werden häufig von PatientInnen ausgesprochen, die sich in der Anfangsphase der Alzheimer‘s Demenz befinden. Daher, Ziel unserer Studie war, die prädiktive Validität der subjektiven Gedächtnisbeschwerden für die Vorhersage der Konversion zu Alzheimer‘s Demenz zu evaluieren. Zusätzlich wurden Korrelationen zwischen subjektiven Gedächtnisbeschwerden und objektiven Gedächtnisleistung analysiert und Modelle zur Einschätzung reliabler Veränderungen im neuropsychologischen Kontext deskriptiv beschrieben.

Methodik: An der retrospektiven neuropsychologischen Untersuchung haben 299 PatientInnen gesamt 46 Kontrollen teilgenommen. Die subjektiven Gedächtnisbeschwerden und deren Einschätzung durch Bezugspersonen wurden mithilfe der Forgetfulness Assessment Inventory erhoben. Die Follow-up Untersuchung wurde zwischen 12 und 48 Monaten durchgeführt.

Ergebnisse: Bei der Follow-up Untersuchung wurde die Konversionsrate zu Alzheimer‘s Demenz von 9.7% beobachtet. Weder subjektive Gedächtnisbeschwerden noch von der Bezugsperson eingeschätzte kognitive Leistung der PatientInnen können die Konversion zu Alzheimer‘s Demenz hervorsagen. Die z-Werte hatten mit Vergleich zu den FAI-Rohwerte eine noch niedrigere prädiktive Validität. Die subjektiven Gedächtnisbeschwerden korrelierten mit Alter, depressiven Symptomen und objektiven kognitiven Tests. Die Analyse der Methoden zur Erfassung reliablen Veränderungen ergab, dass die regressionsbasiere Methode von McSweeny die mit der höchsten Responsivität auf Veränderung war.

Konklusion: Subjektive Gedächtnisbeschwerden haben eine niedrigere prädiktive Validität bei der Vorhersage der Konversion zu Alzheimer‘s Demenz.