Clinical validity of medial temporal atrophy as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework

M ten Kate[1], F Barkhof[2,3], M Boccardi[4,5], PJ Visser[1,6], CR Jack Jr[7], KO Lovblad[8], GB Frisoni[4,9], P Scheltens[1], and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers

1. Alzheimer Center & Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, Netherlands
2. Department of Radiology & Nuclear Medicine, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, Netherlands
3. European Society of Neuroradiology (ESNR)
4. Laboratory of Alzheimer's Neuroimaging & Epidemiology, IRCCS Fatebenefratelli, Brescia, Italy.
5. LANVIE - Laboratory of Neuroimaging of Aging, University of Geneva, Switserland
6. Department of Psychiatry & Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, the Netherlands
7. Department of Radiology, Mayo Clinic, Rochester, MN, USA
8. Laboratory of Neuroimaging of Aging, Memory Clinic and LANVIE, University Hospital of Geneva, Geneva, Switzerland
9. Department of Radiology, University Hospital of Geneva, Geneva, Switzerland

Corresponding author:
M. ten Kate
Alzheimer Center & Department of Neurology, VU University Medical Center
PO Box 7056
1007 MB Amsterdam, the Netherlands
e-mail: m.tenkate1@vumc.nl
tel: +31 (0) 20 444 33 37

Manuscript details:
Word count abstract: 162 + Graphical Abstract
Word count main text: 8139
Number of references: 178
Number of tables: 1
Number of figures: 1
Supplementary table: 1

Abstract

Research criteria for Alzheimer's disease recommend the use of biomarkers for diagnosis, but whether biomarkers improve the diagnosis in clinical routine has not been systematically assessed. Aim is to evaluate the evidence for use of medial temporal lobe atrophy (MTA) as biomarker for Alzheimer's disease at the MCI stage in routine clinical practice, with an adapted version of the 5-phase oncology framework for biomarker development. A literature review on visual assessment of MTA and hippocampal volumetry was conducted with other biomarkers addressed in parallel reviews. Ample evidence is available for Phase 1 (rationale for use) and Phase 2 (discriminative ability between diseased and control subjects). Phase 3 (early detection ability) is partly achieved: most evidence is derived from research cohorts or clinical populations with short follow-up but validation in clinical MCI cohorts is required. In phase 4, only the practical feasibility has been addressed for visual rating of MTA. The rest of phase 4 and phase 5 have not yet been addressed.

1. Introduction

Biomarker assessment at the time of clinical dementia is important for differential diagnosis and thereby also for prognosis and potential treatment. New clinical criteria have introduced the use of biomarkers for the diagnosis of Alzheimer's disease (Dubois et al., 2014, 2007; Jack et al., 2011a; McKhann et al., 2011). These criteria also recommend the use of biomarkers for the diagnosis of Alzheimer's disease in non-demented subjects with mild cognitive impairment (MCI) (Albert et al., 2011). Early diagnosis offers an opportunity for early intervention, improved guidance for caregivers and more accurate prognosis.

Several biomarkers for Alzheimer's disease have been developed; however, there is insufficient systematically addressed evidence to implement them for a diagnosis of Alzheimer's disease at the MCI stage in routine clinical practice (Frisoni et al., 2011). To overcome a similar problem in the field of oncology, Pepe and colleagues (2001) suggested to systematize the investigation of cancer biomarkers based on a framework borrowed from drug development. A similar approach may boost the adoption of Alzheimer's disease biomarkers in clinical practice. An effort has recently been launched to adapt the oncology framework to suit the current goal of diagnosis of Alzheimer's disease at the MCI stage, as described in the accompanying summarizing paper (Frisoni et al., in this issue). The present study fulfills a specific part of this wider plan: the analysis of the available evidence for medial temporal lobe atrophy (MTA) on MRI in the context of this framework. The other studies from this effort assessed within the common framework the following biomarkers: episodic memory assessment (Cerami et al., in this issue), cerebrospinal fluid measures (Mattsson et al., in this issue), amyloid PET (Chiotis et al, in this issue), [18F]FDG-PET (Garibotto et al., in this issue), and [123I]-Ioflupane and [123I]-MIBG imaging (Sonni, et al., in this issue).

In a research setting, there are three methods commonly used to assess atrophy of the medial temporal lobe: visual rating, manual volumetry and automated volumetry. In visual rating, atrophy is

assessed on an ordinal scale. The most widely adopted visual rating scale in research, used in more than 100 publications, is the five-point Scheltens scale which was developed over twenty years ago (Scheltens et al., 1992). MTA is visually assessed on coronal images taking into account the width of the choroid fissure, width of temporal horn and height of hippocampus. Although visual rating scales can be broadly applied to a range of imaging acquisition methods and performed by a trained radiologist, it is still only sparsely adopted in clinical practice (Gardeniers et al., 2015). In manual volumetry, hippocampal atrophy is quantified by drawing multiple regions of interest by an experienced rater on adjacent coronal MRI slices, typically 1-3 mm thick. This requires trained raters and is time-consuming. Several automated methods, which estimate volumes of structures by means of a computer algorithm, have been developed. Most of these algorithms involve an automated segmentation and classification of hippocampal tissue. Volume is then calculated as the sum of all voxels classified as hippocampal tissue. These techniques require specialist software and expertise, are also time-consuming and results can vary across scanners and acquisition protocols. Manual and automated volumetry are currently only used in research settings and not yet implemented for clinical use.

In this paper, we will review evidence for the maturity of visual rating of MTA (vMTA) and hippocampal volume (HCV) as biomarkers for Alzheimer's disease at the MCI stage, where evidence is interpreted under the light of an adapted version of the oncology framework (Pepe et al., 2001; Frisoni et al., in this issue).

2. Methods

This review was performed with reference to the oncology framework (Pepe et al., 2001), which was adapted to the field of dementia, specifically to the aim of *performing the diagnosis of sporadic Alzheimer's disease at the MCI stage* (Frisoni et al., in this issue). The lexicon of this framework is extensively described in Frisoni et al., in this issue, and is summarized in this section. Only sporadic, not familial Alzheimer's disease, is considered. The standard reference for diagnosis was Alzheimer's disease neuropathology or the development of incident Alzheimer's dementia at follow-up.

2.1 Glossary

Alzheimer's disease is defined as the presence of Alzheimer's pathology consisting of cerebral amyloid plaques and tangles, supposedly leading to a pattern-specific neurodegeneration (medio-temporal and temporo-parietal distribution). The term is thus independent of the clinical manifestation of the disease.

Alzheimer's dementia is the clinical syndrome featuring acquired and progressive cognitive impairment associated with functional disability as defined by the NINCDS-ADRDA criteria (McKhann et al., 1984). Notably, not all cases of clinically diagnosed Alzheimer's dementia have Alzheimer's disease pathology due to the imperfect accuracy of purely clinical criteria.

Mild cognitive impairment (MCI) is used to indicate the clinical condition between normality and dementia in which patients experience acquired cognitive impairment of greater severity than expected by age, but no functional disability. This population includes cases with Alzheimer's disease (about 50%), cases with other neurodegenerative pathologies (10-15%) and cases without a neurodegenerative disorder (35-40%) (Bennett et al., 2002; Jack et al., 2008b; Jansen et al., 2015; Rowe et al., 2010). The MCI cases with Alzheimer's disease biomarker positivity are defined as *Prodromal Alzheimer's disease* in the clinical criteria by Dubois et al. (2010). The focus of this review is the use of biomarkers in the MCI stage.

Non-Alzheimer neurodegenerative disease includes neurodegenerative disorders that are not primarily due to Alzheimer's pathology. These belong to a large pathological spectrum including hippocampal sclerosis, frontotemporal lobar degeneration (FTLD), progressive supranuclear palsy, corticobasal degeneration, argyrophilic grain disease, Lewy body disease (LBD) and other alpha-synucleinopathies such as multiple system atrophy.

2.2 Conceptual framework

The main phases of the present framework for the development of biomarkers for early diagnosis reflect the phases covered in the original oncology framework by Pepe et al. (2001) and in turn inspired by pharmaceuticals development. The shift of the reference methodological model from the field of oncology to that of dementia implies a shift of aims from screening to diagnosis, as the examined biomarkers are validated for use in the clinical population of MCI (Frisoni et al., in this issue; Boccardi et al., in this issue). The present review assesses the clinical validity of medial temporal lobe atrophy within a translated framework, consisting of five phases with a main aim and various sub-aims. These (sub-)aims and evidence reported are summarized in Table 1, together with the most pertinent references. There are five consecutive phases of development that should be completed before clinical use of the biomarker for prediction of Alzheimer's disease at the MCI stage.

Phase 1: Phase 1 studies are preclinical exploratory studies, in which the aim is to find leads for potential biomarkers by identifying characteristics specific to the disease, based on pathology findings, which could be detected with clinical tests.

Phase 2: Evaluation of the biomarker's ability to discriminate patients with Alzheimer's disease from controls. Ideally, evidence is based on studies in which a diagnosis of Alzheimer's disease is also supported by autopsy findings. Sub-aims of phase 2 focus on defining and optimizing the clinical assay allowing reliable discrimination between patients and controls (sub-aim 1), determining the relation between pathological measurement and biomarker measurement (sub-aim 2) and the assessment of possible differential effects of factors in patients and controls that may influence the

6

thresholds for positivity (sub-aims 3 and 4). Relevant factors can be for example age, apolipoprotein E ε 4 (APOE4) carrier status and educational attainment.

Phase 3: Phase 3 studies consist of prospective longitudinal studies and the main aim is to assess the ability of the biomarker to detect the disease at the MCI stage by evaluating the biomarker's ability to predict the development of incident Alzheimer's dementia at follow-up. In other words, to distinguish MCI progressors from non-progressors. We included studies examining the biomarker at baseline in subjects with MCI and sufficiently long follow-up, ideally over three years. Second main aim is to fine-tune the threshold for positivity. Sub-aims are to assess the impact of covariates on the discriminatory ability of the biomarker in the MCI stage (sub-aim 1), to compare the usefulness of the biomarker in comparison to or in combination with other available biomarkers (sub-aims 2 and 3) and to determine a biomarker testing interval (sub-aim 4).

Phase 4: Main aim of phase 4 is to estimate the accuracy and usefulness of the biomarker-based early diagnosis in real world patients. It consists of prospective diagnostic studies in which the biomarker is used for an early diagnosis of Alzheimer's disease, affecting decision-making regarding patient management and treatment. Sub-aims are to assess the benefit of the biomarker-based early diagnosis (sub-aim 1) and the feasibility of the biomarker assessment (sub-aim 2), and provides preliminary evidence on impact on mortality, morbidity and costs (sub-aim 3), and undetected cases (sub-aim 4).

Phase 5: Phase 5 studies aim to quantify the impact of the biomarker-based early diagnosis on clinically meaningful outcomes and costs. They consist of disease-control studies assessing the reduction in mortality, morbidity and disability allowed by the biomarker-based diagnosis. Sub-aims are to assess cost-effectiveness (sub-aim 1), evaluate compliance in different settings (sub-aim 2) and to compare different biomarker testing protocols (sub-aim 3). However, this phase can be properly carried out only in the context of an effective and accepted treatment available. For the Alzheimer's disease field, only mortality and quality of life may be properly considered within this phase.

2.3 Literature search, article selection and evidence evaluation

References for this review were selected searching the PubMed/Medline database in June 2015. A different search algorithm was used for the aims of the 5-phases, each one comprising an aim-specific key word string and a biomarker specific key-word string: "MTA" OR "medial temporal" OR "hippocamp*". The aim-specific key word strings can be found in supplementary table 1. Only papers published in English were included. References were also selected on the basis of the authors' personal knowledge and by screening references from retrieved articles. When aims were unequivocally achieved, a reference paper or review was selected by the authors. The final selection of articles was based on relevance to topics covered in this review, as judged by the authors.

For all phases, available literature was assessed and used to evaluate whether each aim was considered as Achieved, Partly Achieved, presenting with Preliminary Evidence, or Not Achieved for vMTA and HCV. Results of this assessment are visualized in figure 1. *Fully Achieved:* scientific evidence is available and replicated in representative samples in studies without major methodological faults. *Partly Achieved:* scientific evidence is available but not yet sufficiently replicated, or samples are not representative, or other significant methodological limitations can be found in the available literature. *Preliminary evidence:* only preliminary evidence is available. *Not Achieved:* no evidence was found and no studies are known to be ongoing at the time of the writing of this review.

3. Evidence for clinical validity of medial temporal lobe atrophy on MRI

3.1 Phase 1 – Pilot studies

The aim of phase 1 studies is to find leads for potential biomarkers based on pathological findings. Alzheimer's disease is characterized by extracellular amyloid beta depositions and intraneuronal or extraneuronal neurofibrillary changes, eventually leading to neuronal destruction. Autopsy studies have shown that these changes already start many years before the onset of clinical symptoms with early and prominent neurofibrillary tangles in medial temporal lobe structures (Braak and Braak, 1996, 1991; Delacourte et al., 1999; Duyckaerts et al., 2009). These neuropathological changes in the medial temporal lobe are accompanied by atrophy, which can be visualized on structural MRI (Scheltens et al., 2002).

The first aim can be considered *Fully Achieved* for vMTA and HCV*.*

3.2 Phase 2 – Clinical assay development for clinical disease

The purpose of the second phase is to find a clinical measurement based on the findings from phase 1, which can be easily obtained and sufficiently distinguishes subjects with and without Alzheimer's disease. Secondary aims in this phase are to optimize the procedures for performing the measurement, assess reproducibility, validate the measurement against pathological measurements and assess factors associated with the measurement in controls and diseased subjects.

3.2.1 Phase 2, Primary aim: ability to distinguish patients from controls.

The primary aim of phase 2 is to assess the ability of vMTA and HCV to distinguish patients with Alzheimer's disease from healthy controls.

Many case-control studies have evaluated the use of MTA in differentiating subjects with clinically diagnosed Alzheimer's dementia from healthy controls, which have recently been reviewed (Frisoni et al., 2013). Average specificity of a visual read of MTA to distinguish clinical Alzheimer's dementia from healthy controls is 79% (CI 75-83) with average-good sensitivity of 70% (CI 65-74). In a

pathology verified sample of young Alzheimer's disease subjects (mean age 59), the sensitivity of vMTA was high (92%) with a specificity of 62% (Likeman et al., 2005). Specificities and sensitivities of manual hippocampal volumetry are on average 82% (CI 78-85) and 79% (76-82) respectively (Frisoni et al., 2013). Specificities and sensitivities for automated measurements of HCV are on average 81% (CI 77-85) and 72% (CI 67-77) respectively (Frisoni et al., 2013). The specificty and sensitivity of HCV for detecting Alzheimer's disease were in a similar range in two neuropathological studies: 0.80-0.87% and 0.75-0.82% (Barnes et al., 2006; Gosche et al., 2002).

Although MTA reasonably well distinguishes subjects with dementia from healthy controls, it is less useful in the differential diagnosis of Alzheimer's dementia. Atrophy of the medial temporal lobe and hippocampus is also seen in other neurodegenerative diseases, as well as in vascular dementia (Barnes et al., 2006; Bastos-Leite et al., 2007; de Souza et al., 2013; Galton et al., 2001; Harper et al., 2014; Likeman et al., 2005; van de Pol et al., 2006b).

The primary aim of phase 2 can be considered *Fully Achieved* for vMTA and HCV*.*

3.2.2 Phase 2, Secondary aim 1: optimize procedures for biomarker assessment.

The first secondary aim of phase 2 is to optimize procedures for measuring medial temporal lobe atrophy and to assess the reproducibility of this measurement.

Visual rating of MTA is quick and easy and can be performed by any trained rater, usually a radiologist. In contrast to volumetric methods, vMTA is relatively independent of acquisition protocol. Merely a good quality anatomical scan is required, which is usually a T1-weighted MRI with coronal reconstructions, but can also be a high-resolution CT scan (Wattjes et al., 2009). The Scheltens scale has reasonable inter-rater agreement and reproducibility. A study shortly after the development of the scale reports inter-rater agreements with kappa values of 0.59-0.62 for dichotomized vMTA into present or absent (Scheltens et al., 1995). In more recent publications, the

inter-rater kappa is usually higher, up to 0.90 (Tolboom et al., 2010). This may be due to advances in imaging acquisition techniques with higher field strengths and better display methods or more experience of the raters. Several studies have shown that the reproducibility and accuracy of vMTA is higher when scoring is performed by trained investigators (Boutet et al., 2012; Cavallin et al., 2012b).

Volumetric methods may be influenced by scanner type and acquisition protocol (Huppertz et al., 2010; Jovicich et al., 2009; Nugent et al., 2013; Wonderlick et al., 2009). To increase uniformity of MRI acquisition methods between different sites, the Alzheimer's Disease Neuroimaging Initiative (ADNI), developed a standardized protocol for MRI acquisition (Jack et al., 2008a). Although this sequence has been adopted by some scanner manufacturers, it is not yet widely implemented for clinical use.

Over the years, various methods for manual segmentation of the hippocampus and medial temporal lobe have been developed (Konrad et al., 2009). Due to the complexity of the hippocampal region and different definitions of anatomical landmarks across research groups, manual volumetry has varying reproducibility rates. Recently, effort has been put into the development of a standardized method for manual segmentation of the hippocampus: the EADC-ADNI Harmonized Protocol (HarP) (Boccardi et al., 2015a). Compared to local protocols, using the harmonized protocol for hippocampal segmentation results in higher intra- and interrater agreement (Frisoni et al., 2015). Given the time-consuming and expensive nature of manual outlining, automated volumetry has higher potential to be broadly used in clinical setting. Automated methods approach performance of manual outlining in distinguishing healthy controls, MCI and subjects with Alzheimer's dementia (Frisoni et al., 2013; Shen et al., 2010). Several (semi-)automated algorithms have been developed for the automated measurement of hippocampal volume. Non-commercial, widely used in research, algorithms include FIRST (FMRIB's integrated registration and segmentation tool, FSL) (Patenaude et al., 2011; Smith et al., 2004) and FreeSurfer (Fischl et al., 2002). Several commercially available algorithms have been

developed, such as Assessa® (IXICO) based on the LEAP algorithm (Wolz et al., 2010) and NeuroQuant® (CorTechs Labs). These various methods use different a priori anatomical information on the hippocampus, as well as different computational strategies for volume estimation and therefore also provide different volumes for same subjects (Guadalupe et al., 2014; Yu et al., 2014). Although steps towards the standardization of automated methods have been undertaken (Boccardi et al., 2015b; Jack et al., 2011b; Wolz et al., 2014), their generalizability across acquisition methods and centers is still insufficient for their routine application in clinical practice.

This sub-aim can be considered *Fully Achieved* for vMTA and manual HCV. For automated HCV, this aim is considered *Partly Achieved.*

3.2.3 Phase 2, Secondary aim 2: relationship between pathology and biomarker measurement

Secondary aim 2 of phase 2 consists of determining the relationship between the biomarker measurement and actual pathology.

Medial temporal lobe atrophy assessed on MRI correlates well with neuropathological findings. In post-mortem studies, HCV measured on MRI correlates strongly with histological volume measurements (Bobinski et al., 1999). Furthermore, hippocampal atrophy measurements on MRI are indicative of Braak neurofibrillary tangle stage at autopsy, even in clinically non-demented subjects (Gosche et al., 2002; Jack et al., 2002; Kaur et al., 2014; Whitwell et al., 2008) and HCV on MRI correlates strongly with histopathological measures of neuron count and neurofibrillary tangle density in the hippocampus (Csernansky et al., 2004; Kril et al., 2004). Good correlations between vMTA and Alzheimer pathology at autopsy (Braak staging and plaques and tangles in hippocampus) have been demonstrated by a few studies (Barkhof et al., 2007; Burton et al., 2009). However, high vMTA scores and decreased HCV are not exclusive to Alzheimer pathology (Barkhof et al., 2007; Barnes et al., 2006; Lehmann et al., 2012). Good correlations between Braak stage, neuronal count

and HCV on MRI has also been demonstrated for the HarP manual outlining protocol (Apostolova et al., 2015).

Several studies have compared visual rating of MTA with volumetric methods and found that vMTA ratings correlate well with MRI volumetric measurements (Boutet et al., 2012; Bresciani et al., 2005; Clerx et al., 2013; Ridha et al., 2007; Vermersch et al., 1994; Wahlund et al., 1999).

This sub-aim is considered *Fully Achieved* for vMTA and HCV*.*

3.2.4 Phase 2, Secondary aim 3: impact of covariates on biomarker measurement in healthy controls

Phase 2, secondary aim 3 assesses the impact of covariates on the biomarker level in control subjects.

Several factors that influence medial temporal lobe atrophy in cognitively healthy subjects have been identified: age, apolipoprotein E (APOE) ε4 genotype, vascular risk factors and co-morbid brain disease (mostly psychiatric). In studies examining cognitively healthy elderly subjects, a confounding effect of preclinical Alzheimer's disease cannot be ruled out since only very few studies have examined pathology verified healthy controls.

Numerous studies have demonstrated age-related decline in HCV in cognitively healthy adults (Fjell et al., 2013; Jack et al., 2015; Lockhart and DeCarli, 2014; van de Pol et al., 2006a). Some studies have shown non-linear effects of aging on HCV with increased decline with advancing age (Fjell et al., 2013; Jack et al., 2015; Raz et al., 2005; Walhovd et al., 2011). Age-related decline in HCV is reflected in age-related increase in vMTA scores in cognitively healthy subjects (Cavallin et al., 2012a; Scheltens et al., 1992). Older aged subjects are also more prone to have non-Alzheimer pathology leading to medial temporal lobe atrophy, such as hippocampal sclerosis (Barkhof et al., 2007). Age-related decline in HCV is mediated by a protective effect of higher education (Noble et al., 2012). Other studies have also shown a protective effect of education on HCV (Schreiber et al., 2016).

Various studies report decreased HCV in cognitively healthy subjects with an APOE-ε4 allele compared to APOE-ε4 non-carriers (den Heijer et al., 2012; Taylor et al., 2014). However, this effect is not undisputed with some researchers finding no effect of APOE-ε4 on cross-sectional HCV (Cherbuin et al., 2008; Jack et al., 2015; Okonkwo et al., 2012; Schmidt et al., 1996). A possible explanation for these differences could be the inclusion of different study populations, where some of the studies that failed to find an effect included younger populations.

Studies on associations between vascular risk factors, vascular brain lesions and HCV have yielded inconsistent results (Lockhart and DeCarli, 2014). In the population-based Rotterdam study, researchers found an association between longitudinal hippocampal atrophy rates, vascular white matter lesions and diastolic blood pressure (den Heijer et al., 2012). Another large scale study showed an effect of smoking and diastolic blood pressure on HCV (Janowitz et al., 2014). Others did not find effects of vascular risk factors and vascular white matter hyperintensities on HCV (Gattringer et al., 2012).

Several psychiatric diseases are also associated with decreased HCV in non-demented subjects, such as depression (Brown et al., 2014; Geerlings et al., 2012) and post-traumatic stress disorder (Gurvits et al., 1996).

This sub-aim is considered *Fully Achieved* for vMTA and HCV.

3.2.5 Phase 2, Secondary aim 4: impact of covariates on biomarker measurement in patients with Alzheimer's disease

Phase 2, secondary aim 4 assesses the impact of covariates on the biomarker measurement in subjects with Alzheimer's disease.

Many of the same factors that influence medial temporal lobe atrophy in healthy subjects also apply to subjects with Alzheimer's disease. Additionally, several disease-related factors have been shown to affect the amount of medial temporal lobe atrophy: age-of-onset, disease stage and clinical presentation. A large scale study combining data from the ADNI and AddNeuroMed databases has

shown that in subjects with Alzheimer's dementia, visually rated MTA scores are influenced by age, gender and disease duration (Ferreira et al., 2015). Similarly, lower HCV is found with increasing age and increasing cognitive impairment in patients with Alzheimer's disease (Jack et al., 2012; Peng et al., 2015; Pol et al., 2006a).

Compared to the rest of the brain, medial temporal atrophy is most pronounced at an early disease stage; at later stages other cortical areas also become more affected (McDonald et al., 2009; Scahill et al., 2002; Whitwell et al., 2007). Subjects with early-onset (age ≤ 65 years) Alzheimer's dementia have different atrophy patterns than subjects with late-onset Alzheimer's dementia. Subjects with early-onset Alzheimer's dementia have more pronounced parietal and precuneal atrophy, whereas subjects with late-onset Alzheimer's disease have more prominent medial temporal atrophy (Cavedo et al., 2014; Frisoni et al., 2007; Ishii et al., 2005; Möller et al., 2013). Moreover, early onset-subjects have more frequently a non-memory clinical presentation, which is also associated with relative hippocampal sparing (Koedam et al., 2010; Mendez et al., 2012; Whitwell et al., 2011).

In Alzheimer's disease, carriership of an APOE-ε4 allele is associated with decreased volume of medial temporal lobe structures (Bigler et al., 2000; Boccardi et al., 2004; Geroldi et al., 1999; Hashimoto et al., 2001; Lehtovirta et al., 1995; Manning et al., 2014). This effect may be limited to younger subjects: sensitivity of vMTA to detect early-onset Alzheimer's dementia is high in APOE-ε4 carriers (82%), but only 47% in APOE-ε4 non-carriers (Ferreira et al., 2015), whereas in late-onset the sensitivity was around 80%, regardless of APOE genotype.

Taken together, these studies suggest that medial temporal lobe atrophy may not be useful as a biomarker for early diagnosis of Alzheimer's dementia in young onset patients, especially those without an APOE-ε4 allele or non-memory presentation.

Studies on the effects of vascular pathology on medial temporal lobe atrophy have inconsistent results, with some finding increased medial temporal lobe atrophy associated with white matter hyperintensities (Korf et al., 2005; Leeuw et al., 2006), whereas other studies did not find such a relation (Jang et al., 2013).

This sub-aim is considered *Fully Achieved* for vMTA and HCV*.*

3.3 Phase 3 – Prospective longitudinal repository studies

In phase 3, the ability of medial temporal lobe atrophy to detect subjects with MCI who will progress to Alzheimer's dementia is assessed based on *prospective clinical* studies. Evidence is presented from studies examining medial temporal lobe atrophy as a predictor for clinical decline from the stage of MCI. In this phase, criteria for a positive biomarker test are defined and factors influencing the abilities of the biomarker to detect pre-clinical disease are assessed. Additionally, in this phase different biomarkers are compared and possibly combined for optimal detection.

3.3.1 Phase 3, Primary aim 1: ability of biomarker to predict progression to Alzheimer's dementia

The primary aim of phase 3 is to assess the capacity of the biomarker to detect subjects with MCI who will progress to Alzheimer's dementia.

Compared to cognitively healthy controls, subjects with MCI have reduced HCV in the following order: control > MCI > Alzheimer's dementia (Frisoni et al., 2008; Pennanen et al., 2004; Shen et al., 2010; Shi et al., 2009). A considerable amount of studies support the ability of medial temporal lobe atrophy to predict progression to dementia from the MCI stage. These studies have recently been summarized, resulting in an overall specificity of 75% (CI 67-82) and average sensitivity of 60% (CI 51-68) for the visual rating of MTA on MRI (Frisoni et al., 2013). Most of the studies reported by Frisoni et al. have relatively short clinical follow-up in the order of 1-2 years. deCarli et al. (2007) assessed data from a clinical trial with 3 year longitudinal follow-up in 190 subjects with amnestic MCI and found an average specificity of 98% and sensitivity of 14% for progression to Alzheimer's dementia with a vMTA cut-off score of ≥ 2 (average of left and right) (DeCarli et al., 2007). With a less stringent cut-off (≥1), specificity was 69% and sensitivity 51%. Liu et al. (2013) analyzed MCI subjects from the ADNI cohort with three year follow-up and found specificity of 82% and sensitivity of 32% for the

prediction of progression to Alzheimer's dementia with a vMTA cut-off of ≥ 3 on any side (Liu et al., 2013). Which cut-off scores should be used remains a matter of debate and is elaborated upon in the next section.

Specificities and sensitivities of manual HCV for prediction of progression to dementia at the MCI stage are on average 81% (CI 73-87) and 58% (CI 47-68) respectively (Frisoni et al., 2013). Specificities and sensitivities for automated measurements of HCV are on average 66% (CI 61-71) and 70% (CI 63-76) respectively (Frisoni et al., 2013). Studies in samples with long follow-up (≥ 3 years) have reported similar specificities of 80-87% and sensitivities of 60-67% for manual volumetry (Devanand et al., 2007; Fritzsche et al., 2010) and 50% and 83% for automated HCV (Bakkour et al., 2009).

Medial temporal lobe atrophy performs reasonably well in the prediction of cognitive decline from the MCI stage but is not specific for Alzheimer's disease (Barnes et al., 2006; de Souza et al., 2013; Harper et al., 2014; van de Pol et al., 2006b; Tam et al., 2005). Other imaging markers such as parietal or frontal atrophy may be more useful in differentiating between Alzheimer's disease and other underlying pathologies (Harper et al., 2016, 2015; Koedam et al., 2011; Lehmann et al., 2012; Vemuri et al., 2011).

In MCI subjects who progress to Alzheimer-type dementia, vMTA rating is associated with time to dementia and can therefore also serve as a prognostic marker (van Rossum et al., 2012a). Similar results have been found for volumetric assessments of the hippocampus in MCI (Devanand *et al.*, 2007) and amyloid positive MCI (Jack et al., 2010b; van Rossum et al., 2012b).

Much evidence has already been gathered on the ability of medial temporal lobe atrophy to predict progression to Alzheimer's dementia. However, to fully complete this aim, sensitivities and specificities for progression will also need to be assessed in clinical MCI populations, with longer

follow-up. Therefore, we consider this aim *Partly Achieved* for both vMTA and HCV. Data to fully achieve this aim could be readily available from memory clinic samples with standardized use of MRI (van der Flier et al., 2014).

3.3.2 Phase 3, Primary aim 2: define criteria for positive biomarker test

The second primary aim of phase 3 is to define criteria for a positive biomarker test.

Since there is a considerable influence of age on medial temporal lobe atrophy in both healthy subjects and patients with Alzheimer's dementia, age-related cut-offs for vMTA have been proposed (Barkhof et al., 2007; Ferreira et al., 2015; Pereira et al., 2014; Scheltens et al., 1992). A recent large-scaled study combining data from the ADNI and AddNeuroMed studies examined the effects of APOE-ε4 on vMTA and demonstrated that early-onset Alzheimer's disease subjects without APOE-ε4 have lower vMTA sores than APOE-ε4 carriers (Pereira et al., 2014). This has led to the proposition of age-related cut-offs, stratified by APOE genotype. Using this same dataset, Ferreira and colleagues found the highest performance of vMTA to distinguish Alzheimer's dementia from healthy controls when using the following general age-adjusted cut-offs (average scores): 45-74: ≥ 1.5; 75-84: ≥2 and 85-94: ≥2.5 (sensitivity 80%, specificity 77%). For early-onset (≤65 years) APOE-ε4 non-carriers a vMTA cut-off ≥2 had better performance (Ferreira et al., 2015). These cut-offs will need to be validated for use in prediction of progression from MCI in a prospective clinical setting.

The proposed cut-offs were derived using average vMTA scores of left and right hemisphere. Other researchers have used the highest vMTA score on either side to define abnormality (Scheltens et al, 1992; Geroldi et al., 2006). Pereira et al. (2014) have examined the performance of different vMTA cut-off scores for differentiating patients with Alzheimer's dementia from healthy controls, including cut-offs based on average and highest scores. They found a better performance of age-adjusted average cut-off scores compared to age-adjusted highest cut-off scores, especially in the group of older patients. Another argument for using average scores is the relatively symmetrical distribution

of Alzheimer's pathology, in contrast to some other neurodegenerative diseases (Chan et al., 2001; Boccardi et al., 2003).

For volumetric analysis, several methods to define cut-offs have been used in research settings, depending on the intended use and need for either higher sensitivity or specificity (Bartlett et al., 2012). One strategy that is frequently used involves taking the 95% percentile of a reference population, with the implication that everything below this is abnormal. Another commonly used method is the creation of covariate-corrected Z-scores, also called W-scores (Jack et al., 1997). W-scores represent where a hippocampal volume would fall on the normal distribution of healthy controls, corrected for covariates. W-scores require the availability of a normative data set. Advantages of this method are the possibility to include covariates and the relative robustness of derived cut-offs against the use of different measurement algorithms or acquisition methods.

A big challenge in defining universal cut-offs in automated volumetric analysis is the influence of image acquisition method (such as scanner type and acquisition protocol) on obtained results (Huppertz et al., 2010; Jovicich et al., 2009; Nugent et al., 2013; Wonderlick et al., 2009). Due to the large variability in HCV obtained from different automated methods (Guadalupe et al., 2014; Yu et al., 2014), universal absolute cut-off points cannot be defined. With the development of the HarP (Boccardi et al., 2015a), and of certified labels that may be used to train algorithms for automated segmentation (Boccardi et al., 2015b), results from automated volumetry may become more consistent between methods in the future.

This sub-aim is considered *Partly Achieved* for vMTA and in the stage of *Preliminary Evidence* for HCV.

3.3.3 Phase 3, Secondary aim 1: impact of covariates on biomarker measurement in subjects with MCI

Secondary aim 1 of phase 3 is to explore the impact of covariates on the discriminatory ability of the biomarker to predict progression from the MCI stage.

Few studies have addressed this issue directly. Factors that may impact the ability of medial temporal lobe atrophy to predict Alzheimer's disease at the MCI stage are age, clinical presentation, and APOE genotype. In very old subjects there is age-related hippocampal atrophy in healthy subjects as well as in Alzheimer's dementia patients (van de Pol et al., 2006a), which could affect the discriminatory ability between normal and abnormal. In early-onset Alzheimer's dementia, there is less prominent involvement of the medial temporal lobe, making this biomarker less suitable for young subjects (Frisoni et al., 2007; Möller et al., 2013). Different clinical subtypes of Alzheimer's dementia are associated with specific brain atrophy patterns, which may already be visible in the MCI stage. Compared to amnestic MCI, subjects with non-amnestic MCI have relatively spared medial temporal lobes (Geroldi et al., 2006; Vos et al., 2013), which reduces the sensitivity of medial temporal lobe atrophy for prediction of progression to Alzheimer-type dementia in non-amnestic MCI.

Using ADNI and AddNeuroMed data, it has been shown that APOE-ε4 carriership and early-onset disease before the age of 65 affect the performance of vMTA for prediction of clinical progression from the MCI stage (Pereira et al., 2014). Further replication of these findings in clinical cohorts and extension to volumetric methods are required.

This secondary aim is considered *Partly Achieved* for vMTA and HCV.

3.3.4 Phase 3, Secondary aims 2 & 3: comparison and combination of biomarkers

Secondary aims 2 and 3 of phase 3 are to compare biomarkers and develop algorithms for positivity based on combinations of markers.

Research criteria developed by the National Institute of Aging and the Alzheimer Association (NIA-AA) have incorporated different stages of likelihood of developing Alzheimer's dementia based on

the combination of amyloid markers (either CSF or PET) and injury markers, such as medial temporal

atrophy (Sperling et al., 2011). Neuroimaging and CSF markers are currently widely accepted

biomarkers for Alzheimer's disease in research settings (Hampel et al., 2008). There is, however, not

yet enough knowledge about their use in clinical practice. There is no consensus on which

biomarkers, in what combination and in which order should be used in the work-up of clinical MCI

patients. Various studies have examined the combination of multiple imaging markers for prediction

of decline in patients with MCI, which have recently been summarized (Teipel et al., 2015).

Compared to FDG-PET, measures of hippocampal atrophy seem to be less accurate at predicting

conversion to Alzheimer's dementia from MCI (Brück et al., 2013; Chen et al., 2011; Frisoni et al.,

2013; Shaffer et al., 2013; Yuan et al., 2009) but many studies found highest accuracies for a

combination of biomarkers (Chen et al., 2011; Shaffer et al., 2013). Studies combining information

from amyloid PET and structural MRI have also reported highest accuracies for combinations of both

markers rather than a single biomarker (Jack et al., 2008b; Trzepacz et al., 2014). Multiple studies

have also examined the combined use of CSF and MRI markers for the prediction of clinical

progression in subjects with MCI. Some have found better prediction for CSF and others for MRI but

nearly all of them show added benefit of a combination of both (Bouwman et al., 2007; Eckerström

et al., 2010; Ewers et al., 2012; Galluzzi et al., 2010; Heister et al., 2011; Prestia et al., 2013; Vos et

al., 2012). Only the study from Bouwman et al. assessed vMTA in a clinical cohort, the others

assessed HCV in research settings. Differences in findings between the studies may largely be

explained by use of highly selected samples, different definitions of MCI (or only inclusion of

amnestic MCI) and variations in methods used for deriving the biomarker measures.

Some studies have shown that high vMTA scores and decreased HCV may be better predictors of

time to clinical progression than evidence of amyloid pathology (Jack et al., 2010b; van Rossum et al.,

2012a). Taken together, these studies suggest that different biomarkers provide complementary

information and that combinations of biomarkers may be needed for accurate prediction of clinical

progression to Alzheimer's dementia at the MCI stage. Although recent efforts are focussing on

devising algorithms taking into account all available biomarker data to aid in the prediction of progression to Alzheimer's disease in clinical settings (Rhodius-Meester et al., 2015), more research is still needed to determine which biomarkers should be used, in which order they should be assessed and what to do in the presence of conflicting results.

These sub-aims can be considered *Partly Achieved* for vMTA and HCV.

3.3.5 Phase 3, Secondary aim 4: biomarker testing interval

Secondary aim 4 of phase 3 is to determine a biomarker testing interval.

Progressive hippocampal atrophy is an important imaging finding in Alzheimer's dementia. Compared to cognitively healthy subjects, hippocampal atrophy rates are 2-4 times greater in subjects with Alzheimer's dementia (Barnes et al., 2009; Henneman et al., 2009; Jack et al., 2004; Morra et al., 2008; Schott et al., 2005). For repeated testing to be useful for diagnosis or prognosis in clinical practice, a biomarker should be able to detect changes over short intervals. A visual rating scale is not sensitive enough to detect changes over short term follow-up evaluations in the order of one year (Ridha et al., 2007). Repeated testing may be valuable in the case hippocampal volumetry becomes available for routine clinical application. Hippocampal atrophy rates can be measured reliably over a period of one year and some studies have even show that volumetric atrophy rates can be measured over periods as short as 6 months (Barnes et al., 2008; Holland et al., 2012; Leung et al., 2010; Schuff et al., 2009). Several longitudinal studies have found that increased hippocampal atrophy rates are associated with rapid progression to Alzheimer's dementia in MCI (Leung et al., 2013; Macdonald et al., 2013; Sluimer et al., 2009; Wang et al., 2009) and may perform better in prediction of cognitive decline than baseline hippocampal volumes alone (Henneman et al., 2009; McEvoy et al., 2011).

This sub-aim is not applicable for vMTA. For HCV this sub-aim can be considered at the stage of *Preliminary Evidence*.

3.4 Phase 4 – Prospective diagnostic studies

The primary aim of phase 4 is to determine the operating characteristics of the biomarker in a clinical setting.

Most studies that have examined the use of medial temporal atrophy as a predictor for clinical progression in subjects with MCI have used highly selected populations in terms of MCI subtypes, scan quality and comorbidities and are therefore not generalizable to a memory clinic sample. In order to assess the value of medial temporal atrophy as a biomarker for early diagnosis in MCI, it is important to also examine clinical populations using methods that are feasible for broad implementation. In phase 4 studies, use of the biomarker leads to early diagnosis and the effects on patient management and outcome are assessed. We are not aware of any prospective clinical studies examining the systematic use of medial temporal atrophy for the prediction of Alzheimer's disease at the MCI stage.

3.4.1 Phase 4, Secondary aim 1: characteristics of disease detected by biomarker in early stage

Secondary aim 1 of phase 4 assesses the characteristics of the disease identified in an early stage by the biomarker in a clinical setting, specifically with regard to potential benefit for the patient incurred by early detection.

This aim is *Not Achieved* for vMTA and HCV.

3.4.2 Phase 4, Secondary aim 2: feasibility of biomarker measurement

The secondary aim of phase 4 assesses the practical feasibility of implementing the biomarker measurement in a clinical setting and the compliance of test-positive subjects with work-up and treatment recommendations.

Although not formally assessed, a visual rating of MTA should be feasible in clinical setting. Most

memory clinics already use imaging in the work-up of memory disorders to exclude other (treatable)

pathologies such as tumors, vascular damage as well as give direction to underlying

neurodegenerative pathology such as frontotemporal dementia or progressive supranuclear palsy.

Tertiary memory clinics or centers associated with a research facility often use standardized MRI

protocols and may also adopt structured radiology reporting, including a visual rating of MTA (Boutet

et al., 2012; van der Flier et al., 2014). The application of manual or automated volumetry is still

distant from implementation in daily clinical practice. To be widely implemented in clinical setting,

volumetric analysis should be sufficiently standardized, fully automated and easy to use.

This aim can be considered at the stage of *Preliminary Evidence* for vMTA and *Not Achieved* for HCV*.*

3.4.3 Phase 4, Secondary aim 3: impact on costs and mortality

The aim is to evaluate the effects of biomarker testing on costs and mortality associated with

Alzheimer's disease.

This aim is *Not Achieved* for vMTA and HCV.

3.4.4. Phase 4, Secondary aim 4: monitor undetected cases

This aims includes monitoring disease occurring clinically but not detected by the biomarker testing.

In order words, this subaim assesses how many subjects with MCI show clinical progression in the

absence of medial temporal lobe atrophy at baseline and the clinical trajectories of these subjects.

This aim is *Not Achieved* for vMTA and HCV.

3.5 Phase 5 – Disease control studies

This final phase addresses whether using biomarkers for early diagnosis reduces the burden of

Alzheimer's disease in the general population. There are currently no studies assessing changes in

mortality and morbidity, impact on economic costs or overdiagnosis associated with use of medial temporal lobe atrophy as diagnostic biomarker for Alzheimer's disease at the MCI stage.

This entire phase is considered *Not Achieved* for vMTA and HCV.

4. Conclusions and future perspectives

In this paper we reviewed the evidence for medial temporal lobe atrophy as a biomarker for prediction of Alzheimer's disease at the MCI stage. We performed this review in the context of a wider effort, aiming to accelerate the use of Alzheimer's disease biomarkers at the predementia stage, where differentiating Alzheimer's disease from normal aging and other causes of cognitive impairment is of huge clinical and societal relevance. The effort has borrowed a biomarker validation framework developed for oncology biomarkers (Pepe et al., 2001) and ultimately taking inspiration from the traditional 4-phase drug development framework. Our working group has adapted the oncology framework to the predementia context, highlighting those issues sufficiently investigated and those in need of more research. This will allow funders of biomedical research to prioritize research topics towards the achievement of the ultimate aim of appropriate, effective and efficient use of Alzheimer's disease biomarkers in the clinic.

In this review, we examined the available evidence for use of medial temporal lobe atrophy for prediction of Alzheimer's disease at the MCI stage in the light of this framework. The first phase has been achieved. There is ample evidence to support medial temporal lobe atrophy as a characteristic feature of Alzheimer's disease (phase 1). The second phase, focusing on the ability of the biomarkers to distinguish subjects with Alzheimer's disease from healthy controls has also been achieved for a visual rating of MTA. For HCV, phase 2 has not yet been fully completed, with insufficient progress on the optimization and standardization of measurement algorithms.

There are still some steps to be taken in the third phase. The current evidence does not support the use of vMTA rating or HCV in isolation for the prediction of Alzheimer type dementia at the MCI stage. Accuracy for the prediction of progression to Alzheimer's dementia from the MCI stage has not reached clinically acceptable levels with, on average, sensitivities and specificities below 80% (Frisoni et al., 2013). It should be noted that most studies have been performed on research samples with

short follow-up. Prospective studies on clinical MCI populations with sufficiently long follow-up are lacking in the literature. Studies with short follow-up, in the order of 1-2 years, may underestimate the predictive ability of the biomarker for clinical progression. The predictive ability of medial temporal lobe atrophy is dependent on age, clinical presentation and APOE-ε4 genotype (Ferreira et al., 2015; Pereira et al., 2014; Vos et al., 2013). However, clear guidelines on cut-offs for clinical (sub)populations have not yet been established.

Given the insufficient accuracy of medial temporal lobe atrophy alone for the prediction of Alzheimer's disease at the MCI stage in a clinical setting, phase 4 and phase 5 studies should not be undertaken on this single biomarker but rather focus on assessing the impact of combinations of biomarkers (Fox et al., 2013; Frisoni et al., 2013). Multiple studies have supported evidence for a model in which abnormal amyloid markers are present in early disease; whereas neuronal injury markers, such as medial temporal atrophy, may be more useful in predicting advancing pathology and thereby serve as a prognostic marker, rather than diagnostic marker (Jack et al., 2010a, 2010b, 2011c; van Rossum et al., 2012b). Large size clinic-based studies assessing which combinations of biomarkers should be used, and in which order and what to do in the case of conflicting biomarker results are needed. Related to its potential role as a prognostic marker, HCV may also be valuable as a monitor of disease progression and could be used as a biomarker outcome measure in clinical trials (Drago et al., 2011; Fox and Freeborough, 1997). For the latter, algorithms will need to have very low measurement errors (below 1.5%) to detect effects over a one year follow-up period, as yearly hippocampal atrophy rates are in the order of 2.9-5.6% per year for Alzheimer's disease and 0.3-2.2 in healthy aging (Frisoni et al., 2010), resulting in an average difference of 3% atrophy per year between the groups.

Clinical implementation of a vMTA rating should be achievable, since this is a quick and accessible tool, which can be easily learned with adequate training. As structural imaging is already integrated

in the work-up of patients with dementia in many clinics, it should be practically feasible to implement neuroimaging at the MCI stage as well. Unfortunately, a routine visual assessment of MTA is not yet widely adopted, especially in non-specialized clinics (Gardeniers et al., 2015). Adequate training may have a significant impact on the performance and application of visual rating scales in clinical practice. There might be a role for scientific societies to improve knowledge and know-how on visual ratings through the development of guidelines with reference images and training programs. Hippocampal volumetry using sophisticated analysis methods, rather than a visual read, may ultimately be a more powerful predictor of progression from MCI, but little progress has been made towards the integration of these in clinical work streams (Clerx et al., 2013). The EMA has approved measurements of HCV for use in clinical trials and similar efforts are being undertaken to get approval from the FDA (Hill et al., 2014; Wolz et al., 2014). To be widely adopted in clinical setting, volumetric analysis should be fully automated and easy to use, for example through implementation on the scanner console.

In this paper, we have discussed the assessment of medial temporal lobe atrophy on MRI. Although MRI is the preferred imaging modality in the work-up of dementia, some patients are unable to undergo MRI for various reasons (e.g. claustrophobia, pacemaker, unavailability) or MRI is not included in national guidelines on dementia care (Falahati et al., 2015). In such cases a high resolution CT-scan with multiplanar reconstruction may also be used for the visual assessment of MTA, as has been validated by one study (Wattjes et al., 2009). Future research may focus on further validating the use of vMTA on CT. Due to the lower contrast resolution, CT has only sporadically been used for volumetric measurements (Aguilar et al., 2015).

This review has several limitations. We reviewed the performance of medial temporal lobe atrophy in MCI patients, however, the definition of MCI is not homogeneous across different studies. This issue is addressed by Cerami et al., in this issue. A further limitation of this study is that, notwithstanding

our efforts to be as inclusive as possible, the literature search was not conducted as a formal systematic review. A number of PubMed research strings were proposed centrally for the whole project; however, the literature databases and some selection criteria for included papers were chosen by the authors of each review, who additionally added papers from personal knowledge or other papers reference lists. Second, the original Pepe and colleague's framework was developed to screen cancer in asymptomatic populations and has been further adapted to the early diagnosis of Alzheimer's disease in symptomatic memory clinic patients. Future developments of the field of Alzheimer's disease and in drug development may require and allow to extend this framework to asymptomatic preclinical patients. Therefore, the nature of this whole effort is liable to change in the near future, but still necessary to proceed in a fruitful way to improve clinical practice in the Alzheimer's field.

5. Acknowledgements

Fatebenefratelli, Brescia, Italy) and Agnese Picco (Università di Genova, Genova, Italy) who took care of the logistics of the workshop.

The following scientific societies took part to the Geneva Workshop for the Roadmap of Alzheimer's Biomarkers on December 8-9 2014. Flavio Nobili was delegate from the European Association of Nuclear Medicine (EANM) Neuroimaging Committee. Kaj Blennow was delegate and Chair of the International Federation of Clinical Chemistry and Laboratory Medicine Working Group for CSF proteins (IFCC WG-CSF). Frederik Barkhof was delegate from the European Society of Neuroradiology (ESNR). Stefano Cappa was delegate and Chair of the Federation of European Societies of Neuropsychology (FENS). Urs Mosimann was delegate from the Swiss Federation of Clinical Neuro Societies (SFCNS). The content of this paper represents the opinion of the individual authors and is not necessarily endorsed by the scientific societies which took part to the Geneva Workshop for the Roadmap of Alzheimer's Biomarkers, except the ESNR which formally endorsed it on September 5th 2015.

Disclosure statement

References

Aguilar, C., Edholm, K., Simmons, A., Cavallin, L., Muller, S., Skoog, I., Larsson, E.-M., Axelsson, R., Wahlund, L.-O., Westman, E., 2015. Automated CT-based segmentation and quantification of total intracranial volume. Eur Radiol 25, 3151–3160. doi:10.1007/s00330-015-3747-7

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7, 270–279. doi:10.1016/j.jalz.2011.03.008

Apostolova, L.G., Zarow, C., Biado, K., Hurtz, S., Boccardi, M., Somme, J., Honarpisheh, H., Blanken, A.E., Brook, J., Tung, S., Kraft, E., Lo, D., Ng, D., Alger, J.R., Vinters, H.V., Bocchetta, M., Duvernoy, H., Jack Jr., C.R., Frisoni, G.B., Bartzokis, G., Csernansky, J.G., de Leon, M.J., deToledo-Morrell, L., Killiany, R.J., Lehericy, S., Malykhin, N., Pantel, J., Pruessner, J.C., Soininen, H., Watson, C., 2015. Relationship between hippocampal atrophy and neuropathology markers: A 7T MRI validation study of the EADC-ADNI Harmonized Hippocampal Segmentation Protocol. Alzheimer's & Dementia 11, 139–150. doi:10.1016/j.jalz.2015.01.001

Bakkour, A., Morris, J.C., Dickerson, B.C., 2009. The cortical signature of prodromal AD: regional thinning predicts mild AD dementia. Neurology 72, 1048–1055. doi:10.1212/01.wnl.0000340981.97664.2f

Barkhof, F., Polvikoski, T.M., Straaten, E.C.W. van, Kalaria, R.N., Sulkava, R., Aronen, H.J., Niinistö, L., Rastas, S., Oinas, M., Scheltens, P., Erkinjuntti, T., 2007. The significance of medial temporal lobe atrophy A postmortem MRI study in the very old. Neurology 69, 1521–1527. doi:10.1212/01.wnl.0000277459.83543.99

Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor M, Fox NC, 2006. Measurements of the amygdala and hippocampus in pathologically confirmed alzheimer disease and frontotemporal lobar degeneration. Arch Neurol 63, 1434–1439. doi:10.1001/archneur.63.10.1434

Barnes, J., Bartlett, J.W., van de Pol, L.A., Loy, C.T., Scahill, R.I., Frost, C., Thompson, P., Fox, N.C., 2009. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. Neurobiology of Aging 30, 1711–1723. doi:10.1016/j.neurobiolaging.2008.01.010

Barnes, J., Scahill, R.I., Frost, C., Schott, J.M., Rossor, M.N., Fox, N.C., 2008. Increased hippocampal atrophy rates in AD over 6 months using serial MR imaging. Neurobiology of Aging 29, 1199–1203. doi:10.1016/j.neurobiolaging.2007.02.011

Bartlett, J.W., Frost, C., Mattsson, N., Skillbäck, T., Blennow, K., Zetterberg, H., Schott, J.M., 2012. Determining cut-points for Alzheimer's disease biomarkers: statistical issues, methods and challenges. Biomarkers Med. 6, 391–400. doi:10.2217/bmm.12.49

Bastos-Leite, A.J., van der Flier, W.M., van Straaten, E.C.W., Staekenborg, S.S., Scheltens, P., Barkhof, F., 2007. The Contribution of Medial Temporal Lobe Atrophy and Vascular Pathology to Cognitive Impairment in Vascular Dementia. Stroke 38, 3182–3185. doi:10.1161/STROKEAHA.107.490102

Bennett, D.A., Wilson, R.S., Schneider, J.A., Evans, D.A., Beckett, L.A., Aggarwal, N.T., Barnes, L.L., Fox, J.H., Bach, J., 2002. Natural history of mild cognitive impairment in older persons. Neurology 59, 198–205. doi:10.1212/WNL.59.2.198

Bigler, E.D., Lowry, C.M., Anderson, C.V., Johnson, S.C., Terry, J., Steed, M., 2000. Dementia, Quantitative Neuroimaging, and Apolipoprotein E Genotype. Am J Neuroradiol 21, 1857–1868.

Bobinski, M., de Leon, M.J., Wegiel, J., DeSanti, S., Convit, A., Saint Louis, L.A., Rusinek, H., Wisniewski, H.M., 1999. The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. Neuroscience 95, 721–725. doi:10.1016/S0306-4522(99)00476-5

Boccardi, M., Bocchetta, M., Apostolova, L.G., Barnes, J., Bartzokis, G., Corbetta, G., DeCarli, C., deToledo-Morrell, L., Firbank, M., Ganzola, R., Gerritsen, L., Henneman, W., Killiany, R.J., Malykhin, N., Pasqualetti, P., Pruessner, J.C., Redolfi, A., Robitaille, N., Soininen, H., Tolomeo, D., Wang, L., Watson, C., Wolf, H., Duvernoy, H., Duchesne, S., Jack, C.R., Frisoni, G.B., 2015a. Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance. Alzheimer's & Dementia 11, 126–138. doi:10.1016/j.jalz.2014.02.009

Boccardi, M., Bocchetta, M., Morency, F.C., Collins, D.L., Nishikawa, M., Ganzola, R., Grothe, M.J., Wolf, D., Redolfi, A., Pievani, M., Antelmi, L., Fellgiebel, A., Matsuda, H., Teipel, S., Duchesne, S., Jack, C.R., Frisoni, G.B., EADC-ADNI Working Group on The Harmonized Protocol for Manual Hippocampal Segmentation and for the Alzheimer's Disease Neuroimaging Initiative, 2015b. Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. Alzheimer's & Dementia 11, 175–183. doi:10.1016/j.jalz.2014.12.002

Boccardi, M., Gallo, V., Yutaka, Y., Vineis, P., Padovani, A., Mosimann, U., Giannakopoulos, P., Gold, G., Dubois, B., Jack, C.R., Winblad, B., Frisoni, G.B., Albanese, E., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. The biomarker-based diagnosis of Alzheimer's disease. 2 – Lessons from oncology. Neurobiology of Aging (*in preparation for this issue*)

Boccardi, M., Laakso, M.P., Bresciani, L., Galluzzi, S., Geroldi, C., Beltramello, A., Soininen, H., Frisoni, G.B., 2003. The MRI pattern of frontal and temporal brain atrophy in fronto-temporal dementia. Neurobiology of Aging 24, 95-103. doi:10.1016/S0197-4580(02)00045-3

Boccardi, M., Sabattoli, F., Testa, C., Beltramello, A., Soininen, H., Frisoni, G.B., 2004. APOE and modulation of Alzheimer's and frontotemporal dementia. Neuroscience Letters 356, 167–170. doi:10.1016/j.neulet.2003.11.042

Boutet, C., Chupin, M., Colliot, O., Sarazin, M., Mutlu, G., Drier, A., Pellot, A., Dormont, D., Lehéricy, S., Initiative, A. the A.D.N., 2012. Is radiological evaluation as good as computer-based volumetry to assess hippocampal atrophy in Alzheimer's disease? Neuroradiology 54, 1321–1330. doi:10.1007/s00234-012-1058-0

Bouwman, F.H., Schoonenboom, S.N.M., van der Flier, W.M., van Elk, E.J., Kok, A., Barkhof, F., Blankenstein, M.A., Scheltens, P., 2007. CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment. Neurobiology of Aging 28, 1070–1074. doi:10.1016/j.neurobiolaging.2006.05.006

Braak, H., Braak, E., 1996. Evolution of the neuropathology of Alzheimer's disease. Acta Neurologica Scandinavica 94, 3–12. doi:10.1111/j.1600-0404.1996.tb05866.x

Braak, H., Braak, E., 1991. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 82, 239–259. doi:10.1007/BF00308809

Bresciani, L., Rossi, R., Testa, C., Geroldi, C., Galluzzi, S., Laakso, M.P., Beltramello, A., Soininen, H., Frisoni, G.B., 2005. Visual assessment of medial temporal atrophy on MR films in Alzheimer's disease: comparison with volumetry. Aging Clin Exp Res 17, 8–13. doi:10.1007/BF03337714

Brown, E.S., Hughes, C.W., McColl, R., Peshock, R., King, K.S., Rush, A.J., 2014. Association of depressive symptoms with hippocampal volume in 1936 adults. Neuropsychopharmacology 39, 770–779. doi:10.1038/npp.2013.271

Brück, A., Virta, J.R., Koivunen, J., Koikkalainen, J., Scheinin, N.M., Helenius, H., Någren, K., Helin, S., Parkkola, R., Viitanen, M., Rinne, J.O., 2013. [11C]PIB, [18F]FDG and MR imaging in patients with mild cognitive impairment. Eur J Nucl Med Mol Imaging 40, 1567–1572. doi:10.1007/s00259-013-2478-8

Burton, E.J., Barber, R., Mukaetova-Ladinska, E.B., Robson, J., Perry, R.H., Jaros, E., Kalaria, R.N., O'Brien, J.T., 2009. Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis. Brain 132, 195–203. doi:10.1093/brain/awn298

Cavallin, L., Bronge, L., Zhang, Y., Øksengård, A.-R., Wahlund, L.-O., Fratiglioni, L., Axelsson, R., 2012a. Comparison between visual assessment of MTA and hippocampal volumes in an elderly, non-demented population. Acta Radiol 53, 573–579. doi:10.1258/ar.2012.110664

Cavallin, L., Løken, K., Engedal, K., Øksengård, A.-R., Wahlund, L.-O., Bronge, L., Axelsson, R., 2012b. Overtime reliability of medial temporal lobe atrophy rating in a clinical setting. Acta Radiol 53, 318–323. doi:10.1258/ar.2012.110552

Cavedo, E., Pievani, M., Boccardi, M., Galluzzi, S., Bocchetta, M., Bonetti, M., Thompson, P.M., Frisoni, G.B., 2014. Medial temporal atrophy in early and late-onset Alzheimer's disease. Neurobiology of Aging 35, 2004–2012. doi:10.1016/j.neurobiolaging.2014.03.009

Cerami C., Dubois B., Boccardi M., Monsch A.U., Demonet J.F., Cappa S.F., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. Clinical validity of free and cued wordlist recall as a gateway-biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. Neurobiology of Aging (Submitted, November 23, 2015; NBA 15-927)

Chan, D., Fox, N.C., Scahill, R.I., Crum, W.R., Whitwell, J.L., Leschziner, G., Rossor, A.M., Stevens, J.M., Cipolotti, L., Rossor, M.N., 2001. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Annals of Neurology 49, 433-442. doi:10.1002/ana.92

Chen, K., Ayutyanont, N., Langbaum, J.B.S., Fleisher, A.S., Reschke, C., Lee, W., Liu, X., Bandy, D., Alexander, G.E., Thompson, P.M., Shaw, L., Trojanowski, J.Q., Jack Jr., C.R., Landau, S.M., Foster, N.L., Harvey, D.J., Weiner, M.W., Koeppe, R.A., Jagust, W.J., Reiman, E.M., 2011. Characterizing Alzheimer's disease using a hypometabolic convergence index. NeuroImage 56, 52–60. doi:10.1016/j.neuroimage.2011.01.049

Cherbuin, N., Anstey, K.J., Sachdev, P.S., Maller, J.J., Meslin, C., Mack, H.A., Wen, W., Easteal, S., 2008. Total and Regional Gray Matter Volume Is Not Related to APOE*E4 Status in a Community Sample of Middle-Aged Individuals. J Gerontol A Biol Sci Med Sci 63, 501–504.

Chiotis K., Saint-Aubert L., Boccardi M., Gietl A., Picco A., Varrone A., Garibotto V., Herholz K., Nobili F., Nordberg A., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. Clinical validity of increased cortical uptake of amyloid ligands on PET as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. Neurobiology of Aging (Submitted, November 19, 2015; NBA 15-915)

Clerx, L., van Rossum, I.A., Burns, L., Knol, D.L., Scheltens, P., Verhey, F., Aalten, P., Lapuerta, P., van de Pol, L., van Schijndel, R., de Jong, R., Barkhof, F., Wolz, R., Rueckert, D., Bocchetta, M., Tsolaki, M., Nobili, F., Wahlund, L.-O., Minthon, L., Frölich, L., Hampel, H., Soininen, H., Visser, P.J., 2013. Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. Neurobiology of Aging 34, 2003–2013. doi:10.1016/j.neurobiolaging.2013.02.002

Csernansky, J.G., Hamstra, J., Wang, L., McKeel, D., Price, J.L., Gado, M., Morris, J.C., 2004. Correlations between antemortem hippocampal volume and postmortem neuropathology in AD subjects. Alzheimer Dis Assoc Disord 18, 190–195.

De Souza, L.C., Chupin, M., Bertoux, M., Lehéricy, S., Dubois, B., Lamari, F., Le Ber, I., Bottlaender, M., Colliot, O., Sarazin, M., 2013. Is hippocampal volume a good marker to differentiate Alzheimer's disease from frontotemporal dementia? J Alzheimers Dis 36, 57–66. doi:10.3233/JAD-122293

DeCarli, C., Frisoni, G.B., Clark, C.M., Harvey, D., Grundman, M., Petersen, R.C., Thal, L.J., Jin, S., Jack, C.R., Scheltens, P., Alzheimer's Disease Cooperative Study Group, 2007. Qualitative estimates of medial temporal atrophy as a predictor of progression from mild cognitive impairment to dementia. Arch Neurol 64, 108–115. doi:10.1001/archneur.64.1.108

Delacourte, A., David, J.P., Sergeant, N., Buée, L., Wattez, A., Vermersch, P., Ghozali, F., Fallet-Bianco, C., Pasquier, F., Lebert, F., Petit, H., Menza, C.D., 1999. The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease. Neurology 52, 1158–1158. doi:10.1212/WNL.52.6.1158

Den Heijer, T., van der Lijn, F., Ikram, A., Koudstaal, P.J., van der Lugt, A., Krestin, G.P., Vrooman, H.A., Hofman, A., Niessen, W.J., Breteler, M.M.B., 2012. Vascular risk factors, apolipoprotein E, and hippocampal decline on magnetic resonance imaging over a 10-year follow-up. Alzheimer's & Dementia 8, 417–425. doi:10.1016/j.jalz.2011.07.005

Devanand, D.P., Pradhaban, G., Liu, X., Khandji, A., Santi, S.D., Segal, S., Rusinek, H., Pelton, G.H., Honig, L.S., Mayeux, R., Stern, Y., Tabert, M.H., Leon, M.J. de, 2007. Hippocampal and entorhinal atrophy in mild cognitive impairment Prediction of Alzheimer disease. Neurology 68, 828–836. doi:10.1212/01.wnl.0000256697.20968.d7

Drago, V., Babiloni, C., Bartrés-Faz, D., Caroli, A., Bosch, B., Hensch, T., Didic, M., Klafki, H.-W., Pievani, M., Jovicich, J., Venturi, L., Spitzer, P., Vecchio, F., Schoenknecht, P., Wiltfang, J.,

Redolfi, A., Forloni, G., Blin, O., Irving, E., Davis, C., Hårdemark, H., Frisoni, G.B., 2011. Disease tracking markers for Alzheimer's disease at the prodromal (MCI) stage. J Alzheimers Dis 26 Suppl 3, 159–199. doi:10.3233/JAD-2011-0043

Dubois, B., Feldman, H.H., Jacova, C., Cummings, J.L., DeKosky, S.T., Barberger-Gateau, P., Delacourte, A., Frisoni, G., Fox, N.C., Galasko, D., Gauthier, S., Hampel, H., Jicha, G.A., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Sarazin, M., de Souza, L.C., Stern, Y., Visser, P.J., Scheltens, P., 2010. Revising the definition of Alzheimer's disease: a new lexicon. The Lancet Neurology 9, 1118–1127. doi:10.1016/S1474-4422(10)70223-4

Dubois, B., Feldman, H.H., Jacova, C., DeKosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria. The Lancet Neurology 6, 734–746. doi:10.1016/S1474-4422(07)70178-3

Dubois, B., Feldman, H.H., Jacova, C., Hampel, H., Molinuevo, J.L., Blennow, K., DeKosky, S.T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G.B., Fox, N.C., Galasko, D., Habert, M.-O., Jicha, G.A., Nordberg, A., Pasquier, F., Rabinovici, G., Robert, P., Rowe, C., Salloway, S., Sarazin, M., Epelbaum, S., de Souza, L.C., Vellas, B., Visser, P.J., Schneider, L., Stern, Y., Scheltens, P., Cummings, J.L., 2014. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. The Lancet Neurology 13, 614–629. doi:10.1016/S1474-4422(14)70090-0

Duyckaerts, C., Delatour, B., Potier, M.-C., 2009. Classification and basic pathology of Alzheimer disease. Acta Neuropathol 118, 5–36. doi:10.1007/s00401-009-0532-1

Eckerström, C., Andreasson, U., Olsson, E., Rolstad, S., Blennow, K., Zetterberg, H., Malmgren, H., Edman, A., Wallin, A., 2010. Combination of hippocampal volume and cerebrospinal fluid biomarkers improves predictive value in mild cognitive impairment. Dement Geriatr Cogn Disord 29, 294–300. doi:10.1159/000289814

Ewers, M., Walsh, C., Trojanowski, J.Q., Shaw, L.M., Petersen, R.C., Jack Jr., C.R., Feldman, H.H., Bokde, A.L.W., Alexander, G.E., Scheltens, P., Vellas, B., Dubois, B., Weiner, M., Hampel, H., 2012. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. Neurobiology of Aging 33, 1203–1214.e2. doi:10.1016/j.neurobiolaging.2010.10.019

Falahati, F., Fereshtehnejad, S.-M., Religa, D., Wahlund, L.-O., Westman, E., Eriksdotter, M., 2015. The Use of MRI, CT and Lumbar Puncture in Dementia Diagnostics: Data from the SveDem Registry. Dementia and Geriatric Cognitive Disorders 39, 81–91. doi:10.1159/000366194

Ferreira, D., Cavallin, L., Larsson, E.-M., Muehlboeck, J.-S., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Simmons, A., Wahlund, L.-O., Westman, E., the AddNeuroMed consortium and the Alzheimer's Disease Neuroimaging Initiative, 2015. Practical cut-offs for visual rating scales of medial temporal, frontal and posterior atrophy in Alzheimer's disease and mild cognitive impairment. J Intern Med 278, 277–290. doi:10.1111/joim.12358

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. Neuron 33, 341–355. doi:10.1016/S0896-6273(02)00569-X

Fjell, A.M., Westlye, L.T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Holland, D., Dale, A.M., Walhovd, K.B., Alzheimer Disease Neuroimaging Initiative, 2013. Critical ages in the life course of the adult brain: nonlinear subcortical aging. Neurobiology of Aging 34, 2239–2247. doi:10.1016/j.neurobiolaging.2013.04.006

Fox, C., Lafortune, L., Boustani, M., Dening, T., Rait, G., Brayne, C., 2013. Screening for dementia – is it a no brainer? Int J Clin Pract 67, 1076–1080. doi:10.1111/ijcp.12239

Fox, N.C., Freeborough, P.A., 1997. Brain atrophy progression measured from registered serial MRI: Validation and application to alzheimer's disease. J Magn Reson Imaging 7, 1069–1075. doi:10.1002/jmri.1880070620

Frisoni, G.B., Bocchetta, M., Chételat, G., Rabinovici, G.D., de Leon, M.J., Kaye, J., Reiman, E.M., Scheltens, P., Barkhof, F., Black, S.E., Brooks, D.J., Carrillo, M.C., Fox, N.C., Herholz, K., Nordberg, A., Jack, C.R., Jagust, W.J., Johnson, K.A., Rowe, C.C., Sperling, R.A., Thies, W., Wahlund, L.-O., Weiner, M.W., Pasqualetti, P., Decarli, C., ISTAART's NeuroImaging Professional Interest Area, 2013. Imaging markers for Alzheimer disease: which vs how. Neurology 81, 487–500. doi:10.1212/WNL.0b013e31829d86e8

Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 6, 67–77. doi:10.1038/nrneurol.2009.215

Frisoni, G.B., Hampel, H., O'Brien, J.T., Ritchie, K., Winblad, B., 2011. Revised criteria for Alzheimer's disease: what are the lessons for clinicians? The Lancet Neurology 10, 598–601. doi:10.1016/S1474-4422(11)70126-0

Frisoni, G.B., Henneman, W.J.P., Weiner, M.W., Scheltens, P., Vellas, B., Reynish, E., Hudecova, J., Hampel, H., Burger, K., Blennow, K., Waldemar, G., Johannsen, P., Wahlund, L.-O., Zito, G., Rossini, P.M., Winblad, B., Barkhof, F., 2008. The pilot European Alzheimer's Disease Neuroimaging Initiative of the European Alzheimer's Disease Consortium. Alzheimer's & Dementia 4, 255–264. doi:10.1016/j.jalz.2008.04.009

Frisoni, G.B., Jack, C.R., Bocchetta, M., Bauer, C., Frederiksen, K.S., et al., 2015. The EADC-ADNI Harmonized Protocol for manual hippocampal segmentation on magnetic resonance: Evidence of validity. Alzheimer's & Dementia 11, 111–125. doi:10.1016/j.jalz.2014.05.1756

Frisoni, G.B., Perani D., Bastianello S., Bernardi G., Porteri C., Boccardi M., Cappa S.F., Trabucchi M., Padovani A, 2016. The use of biomarkers for the diagnosis of Alzheimer's disease in clinical practice: the Italian inter-societal roadmap. Neurobiology of Aging, 2016 (Accepted for publication; NBA 15-700)

Frisoni, G.B., Pievani, M., Testa, C., Sabattoli, F., Bresciani, L., Bonetti, M., Beltramello, A., Hayashi, K.M., Toga, A.W., Thompson, P.M., 2007. The topography of grey matter involvement in early and late onset Alzheimer's disease. Brain 130, 720–730. doi:10.1093/brain/awl377

Fritzsche, K.H., Stieltjes, B., Schlindwein, S., van Bruggen, T., Essig, M., Meinzer, H.-P., 2010. Automated MR morphometry to predict Alzheimer's disease in mild cognitive impairment. Int J CARS 5, 623–632. doi:10.1007/s11548-010-0412-0

Galluzzi, S., Geroldi, C., Ghidoni, R., Paghera, B., Amicucci, G., Bonetti, M., Zanetti, O., Cotelli, M., Gennarelli, M., Frisoni, G.B., Group, T.O.M.C.W., 2010. The new Alzheimer's criteria in a naturalistic series of patients with mild cognitive impairment. J Neurol 257, 2004–2014. doi:10.1007/s00415-010-5650-0

Galton, C.J., Gomez-Anson, B., Antoun, N., Scheltens, P., Patterson, K., Graves, M., Sahakian, B.J., Hodges, J.R., 2001. Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. J Neurol Neurosurg Psychiatry 70, 165–173. doi:10.1136/jnnp.70.2.165

Gardeniers, M., Wattjes, M.P., Meulen, E.F.J., Barkhof, F., Bakker, J., 2015. Beeldvormende diagnostiek bij dementie: inventarisatie van de praktijk binnen de Nederlandse geheugenklinieken. Tijdschr Gerontol Geriatr 47, 2–8. doi:10.1007/s12439-015-0154-0

Garibotto V., Herholtz K., Boccardi M., Picco A., Varrone A., Nordberg A., Nobili F., Ratib O., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. Clinical validity of FDG-PET as a biomarker for Alzheimer's disease in the context of a structured 5 –phase development framework. Neurobiology of Aging (Submitted, November 10, 2015; NBA 15-884)

Gattringer, T., Enzinger, C., Ropele, S., Gorani, F., Petrovic, K.E., Schmidt, R., Fazekas, F., 2012. Vascular risk factors, white matter hyperintensities and hippocampal volume in normal elderly individuals. Dement Geriatr Cogn Disord 33, 29–34. doi:10.1159/000336052

Geerlings, M.I., Brickman, A.M., Schupf, N., Devanand, D.P., Luchsinger, J.A., Mayeux, R., Small, S.A., 2012. Depressive symptoms, antidepressant use, and brain volumes on MRI in a population-based cohort of old persons without dementia. J Alzheimers Dis 30, 75–82. doi:10.3233/JAD-2012-112009

Geroldi, C., Pihlajamäki, M., Laakso, M.P., DeCarli, C., Beltramello, A., Bianchetti, A., Soininen, H., Trabucchi, M., Frisoni, G.B., 1999. APOE-ɛ4 is associated with less frontal and more medial temporal lobe atrophy in AD. Neurology 53, 1825–1825. doi:10.1212/WNL.53.8.1825

Geroldi, C., Rossi, R., Calvagna, C., Testa, C., Bresciani, L., Binetti, G., Zanetti, O., Frisoni, G.B., 2006. Medial temporal atrophy but not memory deficit predicts progression to dementia in patients with mild cognitive impairment. Journal of neurology, neurosurgery and psychiatry 77, 1219–1222.

Gosche, K.M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A., 2002. Hippocampal volume as an index of Alzheimer neuropathology Findings from the Nun Study. Neurology 58, 1476–1482. doi:10.1212/WNL.58.10.1476

Guadalupe, T., Zwiers, M.P., Teumer, A., Wittfeld, K., Vasquez, A.A., Hoogman, M., Hagoort, P., Fernandez, G., Buitelaar, J., Hegenscheid, K., Völzke, H., Franke, B., Fisher, S.E., Grabe, H.J., Francks, C., 2014. Measurement and genetics of human subcortical and hippocampal asymmetries in large datasets. Human Brain Mapping 35, 3277–3289. doi:10.1002/hbm.22401

Gurvits, T.V., Shenton, M.E., Hokama, H., Ohta, H., Lasko, N.B., Gilbertson, M.W., Orr, S.P., Kikinis, R., Jolesz, F.A., McCarley, R.W., Pitman, R.K., 1996. Magnetic resonance imaging study of hippocampal volume in chronic, combat-related posttraumatic stress disorder. Biological Psychiatry 40, 1091–1099. doi:10.1016/S0006-3223(96)00229-6

Hampel, H., Bürger, K., Teipel, S.J., Bokde, A.L.W., Zetterberg, H., Blennow, K., 2008. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. Alzheimer's & Dementia 4, 38–48. doi:10.1016/j.jalz.2007.08.006

Harper, L., Barkhof, F., Fox, N.C., Schott, J.M., 2015. Using visual rating to diagnose dementia: a critical evaluation of MRI atrophy scales. J Neurol Neurosurg Psychiatry 86, 1225-33. doi:10.1136/jnnp-2014-310090

Harper, L., Barkhof, F., Scheltens, P., Schott, J.M., Fox, N.C., 2014. An algorithmic approach to structural imaging in dementia. J Neurol Neurosurg Psychiatr 85, 692–698. doi:10.1136/jnnp-2013-306285

Harper, L., Fumagalli, G.G., Barkhof, F., Scheltens, P., O'Brien, J.T., Bouwman, F., Burton, E.J., Rohrer, J.D., Fox, N.C., Ridgway, G.R., Schott, J.M., 2016. MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases. Brain 139, 1211–1225. doi:10.1093/brain/aww005

Hashimoto, M., Yasuda, M., Tanimukai, S., Matsui, M., Hirono, N., Kazui, H., Mori, E., 2001. Apolipoprotein E ɛ4 and the pattern of regional brain atrophy in Alzheimer's disease. Neurology 57, 1461–1466. doi:10.1212/WNL.57.8.1461

Heister, D., Brewer, J.B., Magda, S., Blennow, K., McEvoy, L.K., 2011. Predicting MCI outcome with clinically available MRI and CSF biomarkers. Neurology 77, 1619–1628. doi:10.1212/WNL.0b013e3182343314

Henneman, W.J.P., Sluimer, J.D., Barnes, J., van der Flier, W.M., Sluimer, I.C., Fox, N.C., Scheltens, P., Vrenken, H., Barkhof, F., 2009. Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. Neurology 72, 999–1007. doi:10.1212/01.wnl.0000344568.09360.31

Hill, D.L.G., Schwarz, A.J., Isaac, M., Pani, L., Vamvakas, S., Hemmings, R., Carrillo, M.C., Yu, P., Sun, J., Beckett, L., Boccardi, M., Brewer, J., Brumfield, M., Cantillon, M., Cole, P.E., Fox, N., Frisoni, G.B., Jack, C., Kelleher, T., Luo, F., Novak, G., Maguire, P., Meibach, R., Patterson, P., Bain, L., Sampaio, C., Raunig, D., Soares, H., Suhy, J., Wang, H., Wolz, R., Stephenson, D., 2014. Coalition Against Major Diseases/European Medicines Agency biomarker qualification of

hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease. Alzheimer's & Dementia 10, 421–429.e3. doi:10.1016/j.jalz.2013.07.003

Holland,D., McEvoy, L.C., Dale, A.M., andthe Alzheimer's Disease Neuroimaging Initiative, 2012. Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI. Human Brain Mapping 33, 2586-2602. DOI: 10.1002/hbm.21386

Huppertz, H.-J., Kröll-Seger, J., Klöppel, S., Ganz, R.E., Kassubek, J., 2010. Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. Neuroimage 49, 2216–2224. doi:10.1016/j.neuroimage.2009.10.066

Ishii, K., Kawachi, T., Sasaki, H., Kono, A.K., Fukuda, T., Kojima, Y., Mori, E., 2005. Voxel-based morphometric comparison between early- and late-onset mild Alzheimer's disease and assessment of diagnostic performance of z score images. Am J Neuroradiol 26, 333–340.

Jack, C.R., Albert, M.S., Knopman, D.S., McKhann, G.M., Sperling, R.A., Carrillo, M.C., Thies, B., Phelps, C.H., 2011a. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7, 257–262. doi:10.1016/j.jalz.2011.03.004

Jack, C.R., Barkhof, F., Bernstein, M.A., Cantillon, M., Cole, P.E., DeCarli, C., Dubois, B., Duchesne, S., Fox, N.C., Frisoni, G.B., Hampel, H., Hill, D.L.G., Johnson, K., Mangin, J.-F., Scheltens, P., Schwarz, A.J., Sperling, R., Suhy, J., Thompson, P.M., Weiner, M., Foster, N.L., 2011b. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimer's & Dementia 7, 474–485.e4. doi:10.1016/j.jalz.2011.04.007

Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L.G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Gunnar Krueger, Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008a. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691. doi:10.1002/jmri.21049

Jack, C.R., Dickson, D.W., Parisi, J.E., Xu, Y.C., Cha, R.H., O'Brien, P.C., Edland, S.D., Smith, G.E., Boeve, B.F., Tangalos, E.G., Kokmen, E., Petersen, R.C., 2002. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology 58, 750–757. doi:10.1212/WNL.58.5.750

Jack, C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Trojanowski, J.Q., 2010a. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. The Lancet Neurology 9, 119–128. doi:10.1016/S1474-4422(09)70299-6

Jack, C.R., Lowe, V.J., Senjem, M.L., Weigand, S.D., Kemp, B.J., Shiung, M.M., Knopman, D.S., Boeve, B.F., Klunk, W.E., Mathis, C.A., Petersen, R.C., 2008b. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain 131, 665–680. doi:10.1093/brain/awm336

Jack, C.R., Petersen, R.C., Xu, Y.C., Waring, S.C., O'Brien, P.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 49, 786–794. doi:10.1212/WNL.49.3.786

Jack, C.R., Shiung, M.M., Gunter, J.L., O'Brien, P.C., Weigand, S.D., Knopman, D.S., Boeve, B.F., Ivnik, R.J., Smith, G.E., Cha, R.H., Tangalos, E.G., Petersen, R.C., 2004. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 62, 591–600. doi:10.1212/01.WNL.0000110315.26026.EF

Jack, C.R., Vemuri P, Wiste HJ, et al, 2011c. Evidence for ordering of alzheimer disease biomarkers. Arch Neurol 68, 1526–1535. doi:10.1001/archneurol.2011.183

Jack, C.R., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Lowe, V., Kantarci, K., Bernstein, M.A., Senjem, M.L., Gunter, J.L., Boeve, B.F., Trojanowski, J.Q., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Knopman, D.S., Alzheimer's Disease Neuroimaging Initiative, 2012.

Shapes of the trajectories of 5 major biomarkers of Alzheimer disease. Arch Neurol 69, 856–867. doi:10.1001/archneurol.2011.3405

Jack, C.R., Wiste, H.J., Vemuri, P., Weigand, S.D., Senjem, M.L., Zeng, G., Bernstein, M.A., Gunter, J.L., Pankratz, V.S., Aisen, P.S., Weiner, M.W., Petersen, R.C., Shaw, L.M., Trojanowski, J.Q., Knopman, D.S., Initiative, the A.D.N., 2010b. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. Brain 133, 3336–3348. doi:10.1093/brain/awq277

Jack, C.R., Wiste, H.J., Weigand, S.D., Knopman, D.S., Vemuri, P., Mielke, M.M., Lowe, V., Senjem, M.L., Gunter, J.L., Machulda, M.M., Gregg, B.E., Pankratz, V.S., Rocca, W.A., Petersen, R.C., 2015. Age, Sex, and APOE ε4 Effects on Memory, Brain Structure, and β-Amyloid Across the Adult Life Span. JAMA Neurol 72, 511–519. doi:10.1001/jamaneurol.2014.4821

Jang, J.-W., Kim, S., Na, H.Y., Ahn, S., Lee, S.J., Kwak, K.-H., Lee, M.-A., Hsiung, G.-Y.R., Choi, B.-S., Youn, Y.C., 2013. Effect of white matter hyperintensity on medial temporal lobe atrophy in Alzheimer's disease. Eur Neurol 69, 229–235. doi:10.1159/000345999

Janowitz, D., Schwahn, C., Borchardt, U., Wittfeld, K., Schulz, A., Barnow, S., Biffar, R., Hoffmann, W., Habes, M., Homuth, G., Nauck, M., Hegenscheid, K., Lotze, M., Völzke, H., Freyberger, H.J., Debette, S., Grabe, H.J., 2014. Genetic, psychosocial and clinical factors associated with hippocampal volume in the general population. Transl Psychiatry 4, e465. doi:10.1038/tp.2014.102

Jansen, W., Ossenkoppele, R., Knol, D., et al, 2015. Prevalence of cerebral amyloid pathology in persons without dementia: A meta-analysis. JAMA 313, 1924–1938. doi:10.1001/jama.2015.4668

Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., Pacheco, J., Albert, M., Killiany, R., Blacker, D., Maguire, P., Rosas, D., Makris, N., Gollub, R., Dale, A., Dickerson, B.C., Fischl, B., 2009. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. Neuroimage 46, 177–192. doi:10.1016/j.neuroimage.2009.02.010

Kaur, B., Himali, J.J., Seshadri, S., Beiser, A.S., Au, R., McKee, A.C., Auerbach, S., Wolf, P.A., DeCarli, C.S., 2014. Association between neuropathology and brain volume in the Framingham Heart Study. Alzheimer Dis Assoc Disord 28, 219–225. doi:10.1097/WAD.0000000000000032

Koedam, E.L.G.E., Lauffer, V., van der Vlies, A.E., van der Flier, W.M., Scheltens, P., Pijnenburg, Y.A.L., 2010. Early-Versus Late-Onset Alzheimer's Disease: More than Age Alone. Journal of Alzheimer's Disease 19, 1401–1408. doi:10.3233/JAD-2010-1337

Koedam, E.L.G.E., Lehmann, M., van der Flier, W.M., Scheltens, P., Pijnenburg, Y.A.L., Fox, N., Barkhof, F., Wattjes, M.P., 2011. Visual assessment of posterior atrophy development of a MRI rating scale. Eur Radiol 21, 2618–2625. doi:10.1007/s00330-011-2205-4

Konrad, C., Ukas, T., Nebel, C., Arolt, V., Toga, A.W., Narr, K.L., 2009. Defining the human hippocampus in cerebral magnetic resonance images—An overview of current segmentation protocols. NeuroImage 47, 1185–1195. doi:10.1016/j.neuroimage.2009.05.019

Korf, E.S.C., Scheltens, P., Barkhof, F., de Leeuw, F.-E., 2005. Blood Pressure, White Matter Lesions and Medial Temporal Lobe Atrophy: Closing the Gap between Vascular Pathology and Alzheimer's Disease? Dementia and Geriatric Cognitive Disorders 20, 331–337. doi:10.1159/000088464

Kril, J.J., Hodges, J., Halliday, G., 2004. Relationship between hippocampal volume and CA1 neuron loss in brains of humans with and without Alzheimer's disease. Neuroscience Letters 361, 9–12. doi:10.1016/j.neulet.2004.02.001

Leeuw, F.-E. de, Korf, E., Barkhof, F., Scheltens, P., 2006. White Matter Lesions Are Associated With Progression of Medial Temporal Lobe Atrophy in Alzheimer Disease. Stroke 37, 2248–2252. doi:10.1161/01.STR.0000236555.87674.e1

Lehmann, M., Koedam, E.L.G.E., Barnes, J., Bartlett, J.W., Ryan, N.S., Pijnenburg, Y.A.L., Barkhof, F., Wattjes, M.P., Scheltens, P., Fox, N.C., 2012. Posterior cerebral atrophy in the absence of

medial temporal lobe atrophy in pathologically-confirmed Alzheimer's disease. Neurobiology of Aging 33, 627.e1–627.e12. doi:10.1016/j.neurobiolaging.2011.04.003

Lehtovirta, M., Laakso, M.P., Soininen, H., Helisalmi, S., Mannermaa, A., Helkala, E.-L., Partanen, K., Ryynänen, M., Vainio, P., Hartikainen, P., Riekkinen Sr, P.J., 1995. Volumes of hippocampus, amygdala and frontal lobe in Alzheimer patients with different apolipoprotein E genotypes. Neuroscience 67, 65–72. doi:10.1016/0306-4522(95)00014-A

Leung, K.K., Barnes, J., Ridgway, G.R., Bartlett, J.W., Clarkson, M.J., Macdonald, K., Schuff, N., Fox, N.C., Ourselin, S., 2010. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. NeuroImage 51, 1345–1359. doi:10.1016/j.neuroimage.2010.03.018

Leung, K.K., Bartlett, J.W., Barnes, J., Manning, E.N., Ourselin, S., Fox, N.C., for the Alzheimer's Disease Neuroimaging Initiative, 2013. Cerebral atrophy in mild cognitive impairment and Alzheimer disease Rates and acceleration. Neurology 80, 648–654. doi:10.1212/WNL.0b013e318281ccd3

Likeman, M., Anderson, V.M., Stevens, J.M., Waldman, A.D., Godbolt, A.K., Frost, C., Rossor, M.N., Fox, N.C., 2005. Visual assessment of atrophy on magnetic resonance imaging in the diagnosis of pathologically confirmed young-onset dementias. Arch. Neurol. 62, 1410–1415. doi:10.1001/archneur.62.9.1410

Liu, Y., Mattila, J., Ruiz, M.Á.M., Paajanen, T., Koikkalainen, J., van Gils, M., Herukka, S.-K., Waldemar, G., Lötjönen, J., Soininen, H., for The Alzheimer's Disease Neuroimaging Initiative, 2013. Predicting AD Conversion: Comparison between Prodromal AD Guidelines and Computer Assisted PredictAD Tool. PLoS ONE 8, e55246. doi:10.1371/journal.pone.0055246

Lockhart, S.N., DeCarli, C., 2014. Structural Imaging Measures of Brain Aging. Neuropsychol Rev 24, 271–289. doi:10.1007/s11065-014-9268-3

Macdonald, K.E., Bartlett, J.W., Leung, K.K., Ourselin, S., Barnes, J., ADNI investigators, 2013. The value of hippocampal and temporal horn volumes and rates of change in predicting future conversion to AD. Alzheimer Dis Assoc Disord 27, 168–173. doi:10.1097/WAD.0b013e318260a79a

Manning, E.N., Barnes, J., Cash, D.M., Bartlett, J.W., Leung, K.K., Ourselin, S., Nick C. Fox1 for the Alzheimer's Disease NeuroImaging Initiative, 2014. APOE ε4 Is Associated with Disproportionate Progressive Hippocampal Atrophy in AD. PLoS ONE 9, e97608. doi:10.1371/journal.pone.0097608

Mattsson N., Lönnborg A., Boccardi M., Blennow K., Hansson O., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. Clinical validity of Aβ42, tau, and phospho-tau in the cerebrospinal fluid as biomarkers for Alzheimer's disease in the context of a structured 5-phase development framework. Neurobiology of Aging (Accepted for publication; NBA 15-765)

McDonald, C.R., McEvoy, L.K., Gharapetian, L., Fennema-Notestine, C., Hagler, D.J., Holland, D., Koyama, A., Brewer, J.B., Dale, A.M., 2009. Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. Neurology 73, 457–465. doi:10.1212/WNL.0b013e3181b16431

McEvoy, L.K., Holland, D., Hagler, D.J., Fennema-Notestine, C., Brewer, J.B., Dale, A.M., 2011. Mild Cognitive Impairment: Baseline and Longitudinal Structural MR Imaging Measures Improve Predictive Prognosis. Radiology 259, 834–843. doi:10.1148/radiol.11101975

McKhann G.M., 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, 939–939. doi:10.1212/WNL.34.7.939

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's

Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7, 263–269. doi:10.1016/j.jalz.2011.03.005

Mendez, M.F., Lee, A.S., Joshi, A., Shapira, J.S., 2012. Nonamnestic Presentations of Early-Onset Alzheimer's Disease. Am J Alzheimers Dis Other Demen 27, 413–420. doi:10.1177/1533317512454711

Möller, C., Vrenken, H., Jiskoot, L., Versteeg, A., Barkhof, F., Scheltens, P., van der Flier, W.M., 2013. Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. Neurobiology of Aging 34, 2014–2022. doi:10.1016/j.neurobiolaging.2013.02.013

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Avedissian, C., Madsen, S.K., Parikshak, N., Hua, X., Toga, A.W., Jack Jr., C.R., Weiner, M.W., Thompson, P.M., 2008. Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. NeuroImage 43, 59–68. doi:10.1016/j.neuroimage.2008.07.003

Noble, K.G., Grieve, S.M., Korgaonkar, M.S., Engelhardt, L.E., Griffith, E.Y., Williams, L.M., Brickman, A.M., 2012. Hippocampal volume varies with educational attainment across the life-span. Front Hum Neurosci 6. doi:10.3389/fnhum.2012.00307

Nugent, A.C., Luckenbaugh, D.A., Wood, S.E., Bogers, W., Zarate, C.A., Drevets, W.C., 2013. Automated subcortical segmentation using FIRST: Test-retest reliability, interscanner reliability, and comparison to manual segmentation: Reliability of Automated Segmentation Using FIRST. Human Brain Mapping 34, 2313–2329. doi:10.1002/hbm.22068

Okonkwo, O.C., Xu, G., Dowling, N.M., Bendlin, B.B., LaRue, A., Hermann, B.P., Koscik, R., Jonaitis, E., Rowley, H.A., Carlsson, C.M., Asthana, S., Sager, M.A., Johnson, S.C., 2012. Family history of Alzheimer disease predicts hippocampal atrophy in healthy middle-aged adults. Neurology 78, 1769–1776. doi:10.1212/WNL.0b013e3182583047

Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56, 907–922. doi:10.1016/j.neuroimage.2011.02.046

Peng, G.-P., Feng, Z., He, F.-P., Chen, Z.-Q., Liu, X.-Y., Liu, P., Luo, B.-Y., 2015. Correlation of hippocampal volume and cognitive performances in patients with either mild cognitive impairment or Alzheimer's disease. CNS Neurosci Ther 21, 15–22. doi:10.1111/cns.12317

Pennanen, C., Kivipelto, M., Tuomainen, S., Hartikainen, P., Hänninen, T., Laakso, M.P., Hallikainen, M., Vanhanen, M., Nissinen, A., Helkala, E.-L., Vainio, P., Vanninen, R., Partanen, K., Soininen, H., 2004. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. Neurobiology of Aging 25, 303–310. doi:10.1016/S0197-4580(03)00084-8

Pepe, M., Etzioni, R., Feng, Z., Potter, J.D., Thompson, M.L., Thornquist, M., Winget, M., Yasui, Y., 2001. Phases of Biomarker Development for Early Detection of Cancer. J Natl Cancer Inst 93, 1054–1061.

Pereira, J.B., Cavallin, L., Spulber, G., Aguilar, C., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Spenger, C., Aarsland, D., Lovestone, S., Simmons, A., Wahlund, L.-O., Westman, E., AddNeuroMed consortium and for the Alzheimer's Disease Neuroimaging Initiative, 2014. Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs. J Intern Med 275, 317–330. doi:10.1111/joim.12148

Prestia, A., Caroli, A., Herholz, K., Reiman, E., Chen, K., Jagust, W.J., Frisoni, G.B., Translational Outpatient Memory Clinic Working Group, Alzheimer's Disease Neuroimaging Initiative, 2013. Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series. Alzheimer's & Dementia 9, 677–686. doi:10.1016/j.jalz.2012.09.016

Raz, N., Lindenberger, U., Rodrigue, K.M., Kennedy, K.M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., Acker, J.D., 2005. Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. Cerebral Cortex 15, 1676–1689. doi:10.1093/cercor/bhi044

Rhodius-Meester, H.F.M., Koikkalainen, J., Mattila, J., Teunissen, C.E., Barkhof, F., Lemstra, A.W., Scheltens, P., Lötjönen, J., van der Flier, W.M., 2015. Integrating Biomarkers for Underlying

Alzheimer's Disease in Mild Cognitive Impairment in Daily Practice: Comparison of a Clinical Decision Support System with Individual Biomarkers. Journal of Alzheimer's Disease 50, 261–270. doi:10.3233/JAD-150548

Ridha, B.H., Barnes, J., van de Pol, L.A., Schott, J.M., Boyes, R.G., Siddique, M.M., Rossor, M.N., Scheltens, P., Fox, N.C., 2007. Application of Automated Medial Temporal Lobe Atrophy Scale to Alzheimer Disease. Archives of Neurology 64, 849. doi:10.1001/archneur.64.6.849

Rowe, C.C., Ellis, K.A., Rimajova, M., Bourgeat, P., Pike, K.E., Jones, G., Fripp, J., Tochon-Danguy, H., Morandeau, L., O'Keefe, G., Price, R., Raniga, P., Robins, P., Acosta, O., Lenzo, N., Szoeke, C., Salvado, O., Head, R., Martins, R., Masters, C.L., Ames, D., Villemagne, V.L., 2010. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. Neurobiology of Aging, Alzheimer's Disease Neuroimaging Initiative (ADNI) Studies 31, 1275–1283. doi:10.1016/j.neurobiolaging.2010.04.007

Scahill, R.I., Schott, J.M., Stevens, J.M., Rossor, M.N., Fox, N.C., 2002. Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. PNAS 99, 4703–4707. doi:10.1073/pnas.052587399

Scheltens, P., Fox, N., Barkhof, F., De Carli, C., 2002. Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. The Lancet Neurology 1, 13–21. doi:10.1016/S1474-4422(02)00002-9

Scheltens, P., Launer, L.J., Barkhof, F., Weinstein, H.C., van Gool, W.A., 1995. Visual assessment of medial temporal lobe atrophy on magnetic resonance imaging: Interobserver reliability. J Neurol 242, 557–560. doi:10.1007/BF00868807

Scheltens, P., Leys, D., Barkhof, F., Huglo, D., Weinstein, H.C., Vermersch, P., Kuiper, M., Steinling, M., Wolters, E.C., Valk, J., 1992. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry 55, 967–972. doi:10.1136/jnnp.55.10.967

Schmidt, H., Schmidt, R., Fazekas, F., Semmler, J., Kapeller, P., Reinhart, B., Kostner, G.M., 1996. Apolipoprotein E4 allele in the normal elderly: neuropsychologic and brain MRI correlates. Clinical Genetics 50, 293–299. doi:10.1111/j.1399-0004.1996.tb02377.x

Schott, J.M., Price, S.L., Frost, C., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2005. Measuring atrophy in Alzheimer disease A serial MRI study over 6 and 12 months. Neurology 65, 119–124. doi:10.1212/01.wnl.0000167542.89697.0f

Schreiber, S., Vogel, J., Schwimmer, H.D., Marks, S.M., Schreiber, F., Jagust, W., 2016. Impact of lifestyle dimensions on brain pathology and cognition. Neurobiology of Aging 40, 164–172. doi:10.1016/j.neurobiolaging.2016.01.012

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack, C.R., Weiner, M.W., 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain 132, 1067–1077. doi:10.1093/brain/awp007

Shaffer, J.L., Petrella, J.R., Sheldon, F.C., Choudhury, K.R., Calhoun, V.D., Coleman, R.E., Doraiswamy, P.M., 2013. Predicting Cognitive Decline in Subjects at Risk for Alzheimer Disease by Using Combined Cerebrospinal Fluid, MR Imaging, and PET Biomarkers. Radiology 266, 583–591. doi:10.1148/radiol.12120010

Shen, L., Saykin, A.J., Kim, S., Firpi, H.A., West, J.D., Risacher, S.L., McDonald, B.C., McHugh, T.L., Wishart, H.A., Flashman, L.A., 2010. Comparison of manual and automated determination of hippocampal volumes in MCI and early AD. Brain Imaging Behav 4, 86–95. doi:10.1007/s11682-010-9088-x

Shi, F., Liu, B., Zhou, Y., Yu, C., Jiang, T., 2009. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: Meta-analyses of MRI studies. Hippocampus 19, 1055–1064. doi:10.1002/hipo.20573

Sluimer, J.D., van der Flier, W.M., Karas, G.B., van Schijndel, R., Barnes, J., Boyes, R.G., Cover, K.S., Olabarriaga, S.D., Fox, N.C., Scheltens, P., Vrenken, H., Barkhof, F., 2009. Accelerating

regional atrophy rates in the progression from normal aging to Alzheimer's disease. European Radiology 19, 2826. doi: 10.1007/s00330-009-1512-5

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, Mathematics in Brain Imaging 23, Supplement 1, S208–S219. doi:10.1016/j.neuroimage.2004.07.051

Sonni I., Ratib O., Boccardi M., Picco A., Herholz K., Nobili F., Varrone A., and the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers, 2016. Clinical validity of presynaptic dopaminergic imaging with ^{123}I-ioflupane and noradrenergic imaging with ^{123}I-MIBG SPECT in the differential diagnosis between Alzheimer's disease and Dementia with Lewy bodies in the context of a structured 5–phase development framework. Neurobiology of Aging (Submitted, September 29, 2015; NBA 15-758)

Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack Jr., C.R., Kaye, J., Montine, T.J., Park, D.C., Reiman, E.M., Rowe, C.C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M.C., Thies, B., Morrison-Bogorad, M., Wagster, M.V., Phelps, C.H., 2011. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7, 280–292. doi:10.1016/j.jalz.2011.03.003

Tam, C.W.C., Burton, E.J., McKeith, I.G., Burn, D.J., O'Brien, J.T., 2005. Temporal lobe atrophy on MRI in Parkinson disease with dementia A comparison with Alzheimer disease and dementia with Lewy bodies. Neurology 64, 861–865. doi:10.1212/01.WNL.0000153070.82309.D4

Taylor, J.L., Scanlon, B.K., Farrell, M., Hernandez, B., Adamson, M.M., Ashford, J.W., Noda, A., Murphy, G.M., Weiner, M.W., 2014. APOE-epsilon4 and aging of medial temporal lobe gray matter in healthy adults older than 50 years. Neurobiology of Aging 35, 2479–2485. doi:10.1016/j.neurobiolaging.2014.05.011

Teipel, S., Drzezga, A., Grothe, M.J., Barthel, H., Chételat, G., Schuff, N., Skudlarski, P., Cavedo, E., Frisoni, G.B., Hoffmann, W., Thyrian, J.R., Fox, C., Minoshima, S., Sabri, O., Fellgiebel, A., 2015. Multimodal imaging in Alzheimer's disease: validity and usefulness for early detection. The Lancet Neurology. doi:10.1016/S1474-4422(15)00093-9

Tolboom, N., van der Flier, W.M., Boverhoff, J., Yaqub, M., Wattjes, M.P., Raijmakers, P.G., Barkhof, F., Scheltens, P., Herholz, K., Lammertsma, A.A., van Berckel, B.N.M., 2010. Molecular imaging in the diagnosis of Alzheimer's disease: visual assessment of [11C]PIB and [18F]FDDNP PET images. J Neurol Neurosurg Psychiatry 81, 882–884. doi:10.1136/jnnp.2009.194779

Trzepacz, P.T., Yu, P., Sun, J., Schuh, K., Case, M., Witte, M.M., Hochstetler, H., Hake, A., 2014. Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia. Neurobiology of Aging 35, 143–151. doi:10.1016/j.neurobiolaging.2013.06.018

Van de Pol, L.A., Hensel, A., Barkhof, F., Gertz, H.J., Scheltens, P., van der Flier, W.M., 2006a. Hippocampal atrophy in Alzheimer disease: Age matters. Neurology 66, 236–238. doi:10.1212/01.wnl.0000194240.47892.4d

Van de Pol, L.A., Hensel, A., van der Flier, W.M., Visser, P.J., Pijnenburg, Y. a. L., Barkhof, F., Gertz, H.J., Scheltens, P., 2006b. Hippocampal atrophy on MRI in frontotemporal lobar degeneration and Alzheimer's disease. J Neurol Neurosurg Psychiatry 77, 439–442. doi:10.1136/jnnp.2005.075341

Van der Flier, W.M., Pijnenburg, Y.A.L., Prins, N., Lemstra, A.W., Bouwman, F.H., Teunissen, C.E., van Berckel, B.N.M., Stam, C.J., Barkhof, F., Visser, P.J., van Egmond, E., Scheltens, P., 2014. Optimizing Patient Care and Research: The Amsterdam Dementia Cohort. Journal of Alzheimer's Disease 41, 313–327. doi:10.3233/JAD-132306

Van Rossum, I.A., Visser, P.J., Knol, D.L., van der Flier, W.M., Teunissen, C.E., Barkhof, F., Blankenstein, M.A., Scheltens, P., 2012a. Injury Markers but not Amyloid Markers are

Associated with Rapid Progression from Mild Cognitive Impairment to Dementia in Alzheimer's Disease. Journal of Alzheimer's Disease 29, 319–327. doi:10.3233/JAD-2011-111694

Van Rossum, I.A., Vos, S.J.B., Burns, L., Knol, D.L., Scheltens, P., Soininen, H., Wahlund, L.-O., Hampel, H., Tsolaki, M., Minthon, L., L'Italien, G., van der Flier, W.M., Teunissen, C.E., Blennow, K., Barkhof, F., Rueckert, D., Wolz, R., Verhey, F., Visser, P.J., 2012b. Injury markers predict time to dementia in subjects with MCI and amyloid pathology. Neurology 79, 1809–1816. doi:10.1212/WNL.0b013e3182704056

Vemuri, P., Simon, G., Kantarci, K., Whitwell, J.L., Senjem, M.L., Przybelski, S.A., Gunter, J.L., Josephs, K.A., Knopman, D.S., Boeve, B.F., Ferman, T.J., Dickson, D.W., Parisi, J.E., Petersen, R.C., Jack Jr., C.R., 2011. Antemortem differential diagnosis of dementia pathology using structural MRI: Differential-STAND. NeuroImage 55, 522–531. doi:10.1016/j.neuroimage.2010.12.073

Vermersch, P., Leys, D., Scheltens, P., Barkhof, F., 1994. Visual rating of hippocampal atrophy: correlation with volumetry. J Neurol Neurosurg Psychiatry 57, 1015.

Vos, S., van Rossum, I., Burns, L., Knol, D., Scheltens, P., Soininen, H., Wahlund, L.-O., Hampel, H., Tsolaki, M., Minthon, L., Handels, R., L'Italien, G., van der Flier, W., Aalten, P., Teunissen, C., Barkhof, F., Blennow, K., Wolz, R., Rueckert, D., Verhey, F., Visser, P.J., 2012. Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI. Neurobiology of Aging 33, 2272–2281. doi:10.1016/j.neurobiolaging.2011.12.017

Vos, S., van Rossum, I.A., Verhey, F., Knol, D.L., Soininen, H., Wahlund, L.-O., Hampel, H., Tsolaki, M., Minthon, L., Frisoni, G.B., Froelich, L., Nobili, F.,van der Flier, W., Blennow, K., Wolz, R., Scheltens, P., Visser, P.J., 2013. Prediction of Alzheimer disease in subjects with amnestic and nonamnestic MCI. Neurology 80, 1124–1132. doi:10.1212/WNL.0b013e318288690c

Wahlund, L.-O., Julin, P., Lindqvist, J., Scheltens, P., 1999. Visual assessment of medial temporal lobe atrophy in demented and healthy control subjects: correlation with volumetry. Psychiatry Research: Neuroimaging 90, 193–199. doi:10.1016/S0925-4927(99)00016-5

Wang, P.-N, Liu, H.-C., Lirng, J.-F., Lin, K.N., Wu, Z.A., 2009. Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. Psychiatry Research: Neuroimaging 171, 221–231. doi:10.1016/j.pscychresns.2008.05.002

Walhovd, K.B., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D.H., Greve, D.N., Fischl, B., Dale, A.M., Fjell, A.M., 2011. Consistent neuroanatomical age-related volume differences across multiple samples. Neurobiology of Aging 32, 916–932. doi:10.1016/j.neurobiolaging.2009.05.013

Wattjes, M.P., Henneman, W.J.P., van der Flier, W.M., de Vries, O., Träber, F., Geurts, J.J.G., Scheltens, P., Vrenken, H., Barkhof, F., 2009. Diagnostic Imaging of Patients in a Memory Clinic: Comparison of MR Imaging and 64–Detector Row CT. Radiology 253, 174–183. doi:10.1148/radiol.2531082262

Whitwell, J.L., Clifford, R.J., Przybelski, S.A., Parisi, J.E., Senjem, M.L., Boeve, B.F., Knopman, D.S., Petersen, R.C., Dickson, D.W., Josephs, K.A., 2011. Temporoparietal atrophy: a marker of AD pathology independent of clinical diagnosis. Neurobiology of Aging 32, 1531–1541. doi:10.1016/j.neurobiolaging.2009.10.012

Whitwell, J.L., Josephs, K.A., Murray, M.E., Kantarci, K., Przybelski, S.A., Weigand, S.D., Vemuri, P., Senjem, M.L., Parisi, J.E., Knopman, D.S., Boeve, B.F., Petersen, R.C., Dickson, D.W., Jack, C.R., 2008. MRI correlates of neurofibrillary tangle pathology at autopsy A voxel-based morphometry study. Neurology 71, 743–749. doi:10.1212/01.wnl.0000324924.91351.7d

Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack, C.R., 2007. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain 130, 1777–1786. doi:10.1093/brain/awm112

Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., 2010. LEAP: Learning embeddings for atlas propagation. NeuroImage 49, 1316–1325. doi:10.1016/j.neuroimage.2009.09.069

Wolz, R., Schwarz, A.J., Yu, P., Cole, P.E., Rueckert, D., Jack Jr., C.R., Raunig, D., Hill, D., 2014. Robustness of automated hippocampal volumetry across magnetic resonance field strengths and repeat images. Alzheimer's & Dementia 10, 430–438.e2. doi:10.1016/j.jalz.2013.09.014

Wonderlick, J.S., Ziegler, D.A., Hosseini-Varnamkhasti, P., Locascio, J.J., Bakkour, A., van der Kouwe, A., Triantafyllou, C., Corkin, S., Dickerson, B.C., 2009. Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. Neuroimage 44, 1324–1333. doi:10.1016/j.neuroimage.2008.10.037

Yu, P., Sun, J., Wolz, R., Stephenson, D., Brewer, J., Fox, N.C., Cole, P.E., Jack, C.R., Hill, D.L.G., Schwarz, A.J., 2014. Operationalizing hippocampal volume as an enrichment biomarker for amnestic MCI trials: effect of algorithm, test-retest variability and cut-point on trial cost, duration and sample size. Neurobiology of Aging 35, 808–818. doi:10.1016/j.neurobiolaging.2013.09.039

Yuan, Y., Gu, Z.-X., Wei, W.-S., 2009. Fluorodeoxyglucose–Positron-Emission Tomography, Single-Photon Emission Tomography, and Structural MR Imaging for Prediction of Rapid Conversion to Alzheimer Disease in Patients with Mild Cognitive Impairment: A Meta-Analysis. Am J Neuroradiol 30, 404–410. doi:10.3174/ajnr.A1357

Phase Design	General aim	Specific aim		Progress	Evidence in Alzheimer's disease	Important references
Phase 1—Preclinical Exploratory Studies	Identify the rational of the biomarker, based on pathology	Primary:	To identify leads for potentially useful biomarkers and prioritize identified leads.	Fully achieved	Pathological studies show early medial temporal lobe involvement with neuronal loss in hippocampus.	Braak and Braak, 1991; Braak and Braak, 1996; Delacourte et al., 1999; Duyckaerts et al., 2009
Phase 2—Clinical Assay Development for Clinical Disease	Define the ability of the biomarker to discriminate patients from controls	Primary	To estimate TPR and FPR or ROC curve for the assay and to assess its ability to distinguish subjects with and without disease	Fully achieved	vMTA and HCV separate clinical Alzheimer's dementia patients from cognitively healthy subjects with good sensitivity and specificity. Few studies have examined pathologically verified samples. vMTA and HCV are less useful in differential diagnosis of dementia patients.	Frisoni et al., 2013; Likeman et al., 2005; Barnes et al., 2006; Gosche et al., 2002; Bastos-Leite et al., 2007; Galton et al., 2001; Haper et al., 2014; van de Pol et al., 2006b
		Secondary 1:	To optimize procedures for performing the assay and to assess the reproducibility of the assay within and between laboratories.	Fully achieved for vMTA; Partly achieved for HCV	vMTA has good reproducibility in trained raters. Standardized method of collecting MRI (ADNI protocol) and manual volumetry (EADC-ADNI harmonized segmentation protocol) are sparsely implemented in memory clinics. Much variability between different automated HCV measurement algorithms.	Scheltens et al., 1995; Tolboom et al., 2010; Boutet et al., 2012; Cavallin et al. 2012b; Jack et al., 2008a; Boccardi et al., 2015a; Guadalupe et al., 2014; Yu et al., 2014
		Secondary 2:	To determine the relationship between biomarker tissue measurements made on tissue (phase 1) and the biomarker measurements made on the noninvasive clinical specimen (phase 2)	Fully achieved	Good correlation between hippocampal size on MR and histological measurements and severity of neurodegenerative changes on pathology.	Bobinski et al., 1999; Gosche et al., 2002; Jack et al. 2002; Kaur et al., 2014; Whitwell et al., 2008; Csernansky et al., 2004; Kril et al., 2004; Barkhof et al., 2007; Burton et al., 2009; Apostolova et al., 2015
		Secondary 3:	To assess factors associated with biomarker status or level in control subjects. If such factors affect the biomarker, thresholds for test positivity may need to be defined separately for target subpopulations.	Fully achieved	Well-known age-associated HCV loss (higher vMTA scores). Amount of atrophy also dependent on APOE-ε4 carriership, vascular pathology, education; impact of these latter variables only relevant for volumetry, not for visual rating.	Fjell et al., 2013; Jack et al., 2015; Lockhart and DeCarli, 2014; den Heijer et al., 2012; Taylor et al., 2014; Cherbuin et al., 2008; Janowitz et al., 2014; Gattringer et al., 2012

		Secondary 4:	To assess factors associated with biomarker status or level in diseased subjects—in particular, disease characteristics.	Fully achieved	Similar as for healthy subjects: influence of age, APOE-ε4, vascular pathology. Additionally, clinical presentation is relevant: subjects with early onset and primary non-memory presentation have relatively spared medial temporal lobes.	Ferreira et al., 2015; Jack et al., 2012; Peng et al., 2015; Cavedo et al., 2014; Möller et al., 2013; Ishii et al., 2005; Koedam et al., 2010; Mendez et al., 2012; Whitwell et al. 2011; Bigler et al., 2000; Boccardi et al., 2004; Geroldi et al., 1999; Hashimoto et al., 2001; Lehtovirta et al., 1995; Manning et al., 2014
Phase 3—Prospective Longitudinal Repository Studies	Define the ability of the biomarker to detect the disease in its early phase	Primary 1:	To evaluate the capacity of the biomarker to detect the earliest disease stages	Partly achieved	No evidence in clinical MCI cohorts with long follow-up. In research populations or clinical populations with shorter follow-up, there is a reasonably good specificity but lower sensitivity to predict clinical progression in subjects with MCI.	Frisoni et al., 2013; deCarli et al., 2007; Liu et al., 2013; Devanand et al., 2007; Fritzsche et al., 2010; Bakkour et al., 2009
		Primary 2:	To define criteria for a biomarker positive test in preparation for phase 4.	Partly achieved for vMTA; Preliminary Evidence for HCV	Age and APOE-ε4 related cut-offs for vMTA based on discrimination of controls from Alzheimer's dementia. Validation needed in clinical MCI cohorts. No universal cut-offs for volumetry; substantial variability between acquisition protocols and measurement algorithms.	Barkhof et al., 2007; Ferreira et al., 2015; Pereira et al., 2014; Jack et al, 1997
		Secondary 1:	To explore the impact of covariates on the discriminatory abilities of the biomarker before clinical diagnosis.	Partly achieved	Impact of age, APOE-ε4 genotype and clinical presentation.	Pereira et al., 2014; Geroldi et al., 2006; Vos et al., 2013
		Secondary 2:	To compare markers with a view to selecting those that are most promising	Partly Achieved	Various studies on association of two or more core biomarkers; usually best predictive value for combination of amyloid marker with an injury marker.	Teipel et al., 2015; Brück et al., 2013; Chen et al., 2011; Shaffer et al., 2013; Yuan et al., 2009; Jack et al., 2008b; Trzepacz et al., 2014; Bouwman et al., 2007; Eckerström et al., 2010; Ewers et al., 2012; Galluzzi et al., 2010; Heister et al., 2011; Prestia et al., 2013; Vos et al., 2012; Rhodius-Meester et al., 2015
		Secondary 3:	To develop algorithms for positivity based on combinations of markers.			

		Secondary 4:	To determine a biomarker testing interval for phase 4 if repeated testing is of interest.	Not applicable for vMTA; Preliminary evidence for HCV	Not applicable for vMTA. Preliminary evidence on added value of hippocampal atrophy rates for predicting clinical progression to Alzheimer's dementia in patients with MCI.	Ridha et al., 2007; Leung et al., 2010, 2013; Macdonald et al., 2013; Henneman et al., 2009; McEvoy et al., 2011; Holland et al., 2012; Sluimer et al., 2009; Wang et al., 2009
Phase 4—Prospective Diagnostic Studies	Quantify the biomarker accuracy and usefulness in patients diagnosed and treated based on the BM	Primary:	To determine the operating characteristics of the biomarker-based test in a relevant population by determining the detection rate and the false referral rate. Studies at this stage involve testing people and lead to diagnosis and treatment.	Not achieved		
		Secondary 1:	To describe the characteristics of disease detected by the biomarker test—in particular, with regard to the potential benefit incurred by early detection.	Not achieved		
		Secondary 2:	To assess the practical feasibility of implementing the diagnostic program and compliance of test-positive subjects with work-up and treatment recommendations.	Preliminary evidence for vMTA; Not achieved for HCV	Imaging integrated in standard work-up for dementia patients in most memory clinics. vMTA and quantitative assessment not yet widely implemented.	Boutet et al., 2012; van der Flier et al., 2014
		Secondary 3:	To make preliminary assessments of the effects of biomarker testing on costs and mortality associated with the disease.	Not achieved		
		Secondary 4:	To monitor disease occurring clinically but not detected by the biomarker testing protocol.	Not achieved		
Phase 5—Disease Control Studies	Quantify the impact of the biomarker-based diagnosis on clinically meaningful outcomes and	Primary:	To estimate the reductions in disease-associated mortality, morbidity, and disability afforded by biomarker testing.	Not achieved		
		Secondary 1:	To obtain information about the costs of biomarker testing and treatment and the cost per life saved or per quality-adjusted life year	Not achieved		

| | costs | Secondary 2: | To evaluate compliance with testing and work-up in a diverse range of settings. | Not achieved | | |
| | | Secondary 3: | To compare different biomarker testing protocols and/or to compare different approaches to treating test positive subjects in regard to effects on mortality and costs. | Not achieved | | |

Table 1: Available or required evidence indicating full, partial or lack of achievement of the phases adapted from the oncology framework (Pepe et al., 2001) for visual rating of medial temporal lobe atrophy (vMTA) and hippocampal volume (HCV).

Figure 1: Synopsis of the maturity of a visual rating of medial temporal lobe atrophy (upper panel) and hippocampal volumetry (lower panel) as borrowed from the oncology framework (Pepe et al., 2001). AD: Alzheimer's disease; HC: healthy controls; HCV: hippocampal volume; MCI: mild cognitive impairment; vMTA: visual rating of medial temporal lobe atrophy

Supplementary table 1:

Phase	Aim	Aim-specific key words string
Phase 1: <u>preclinical exploratory studies</u>	<u>**Primary aim:**</u> To identify leads for potentially useful biomarkers and prioritize identified leads.	NO STRINGS USED
Phase 2: <u>clinical assay development for clinical disease</u>	<u>**Primary aim:**</u> To estimate TPR and FPR or ROC curve for the assay and to assess its ability to distinguish subjects with and without disease.	("accuracy" OR "sensitivity" OR "specificity" OR "ROC" OR "predictive value") AND (a) AND (b) AND (d).
	<u>**Secondary aim 1:**</u> To optimize procedures for performing the assay and to assess the reproducibility of the assay within and between laboratories.	("standardization" OR "visual" OR "measure" OR "assessment" OR "reading" OR "quantification" AND ("reproducibility" OR "reliability" OR "agreement") AND (a) AND (d)
	<u>**Secondary aim 2:**</u> To determine the relationship between biomarker tissue measurements made on tissue (phase 1) and the biomarker measurements made on the noninvasive clinical specimen (phase 2).	("autopsy" OR "autoptic" OR "patholog*" OR "neuropatholog*" OR "istopathol*") AND MRI AND (a) AND (d)
	<u>**Secondary aim 3:**</u> To assess factors (e.g. sex, age, etc.), associated with biomarker status or level in control subjects. If such factors affect the biomarker, thresholds for test positivity may need to be defined separately for target subpopulations.	- ("APOE" OR "Apolipoprotein E") AND (b) AND (d) - ("vascular risk factors" OR "white matter hyperintensities) AND (b) AND (d)
	<u>**Secondary aim 4:**</u> To assess factors associated with biomarker status or level in diseased subjects—in particular, disease characteristics.	- ("APOE" OR "Apolipoprotein E") AND (a) AND (d) - ("vascular risk factors" OR "white matter hyperintensities) AND (a) AND (d) - ("early-onset" OR "late-onset") AND (a) AND (d)
Phase 3: <u>Prospective repository studies</u>	<u>**Primary aim:**</u> To evaluate the capacity of biomarkers to detect pre-clinical disease and define criteria for a positive biomarker test in preparation for phase 4.	- ("follow-up" OR "followup" OR "conversion" OR "progression" OR "decline" OR "predict*") AND MRI AND ("visual" OR "rating") AND (c) AND (d) - ("cut-off*" OR "cut-point") AND (d)
	<u>**Secondary aim 1:**</u>To explore the impact of covariates on the discriminatory abilities of the biomarker before clinical diagnosis.	- ("APOE" OR "Apolipoprotein E") AND (c) AND (d) - ("amnestic" OR "non-amnestic") AND (c) AND (d)
	<u>**Secondary aim 2:**</u> To compare markers with a view to selecting those that are most promising.	("follow-up" OR "followup" OR "conversion" OR "progression" OR "decline" OR "predict*") AND ("combinat*" OR "associat*" OR "compar*") AND MRI AND (a) AND (c)
	<u>**Secondary aim 3:**</u>To develop algorithms for positivity based on combinations of markers.	("follow-up" OR "followup" OR "conversion" OR "progression" OR "decline" OR "predict*") AND ("combinat*" OR "associat*" OR "compar*") AND MRI AND (a) AND (c)

	Secondary aim 4: To determine a biomarker testing interval for phase 4 if repeated testing is of interest.	("atrophy rates") AND (a) AND (c) AND (d)
Phase 4: <u>Prospective Diagnostic Studies</u>	**Primary aim:** To determine the operating characteristics of the biomarker-based test in a relevant population by determining the detection rate and the false referral rate. Studies at this stage involve testing people and lead to diagnosis and treatment.	("diagnosis" OR "treatment") AND (a) AND (c) AND (d)
	Secondary aim 1: To describe the characteristics of disease detected by the biomarker test—in particular, with regard to the potential benefit incurred by early detection.	("clinical diagnosis" OR "treatment" OR "memory clinic") AND ("benefits" OR "outcome" OR "improve*") AND (a) AND (c) AND (d)
	Secondary aim 2: To assess the practical feasibility of implementing the case finding program and compliance of test-positive subjects with work-up and treatment recommendations.	("clinical diagnosis" OR "treatment" OR "memory clinic") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life") AND (a) AND (d)
	Secondary aim 3: To make preliminary assessments of the effects of biomarker testing on costs and mortality associated with the disease.	("clinical diagnosis" OR "treatment" OR "memory clinic") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life" OR "cost*") AND (a) AND (d)
	Secondary aim 4: To monitor disease occurring clinically but not detected by the biomarker testing protocol.	("clinical diagnosis" OR "memory clinic" OR "criteria" OR "recommendation*") AND ("accuracy" OR "sensitivity" OR "specificity" OR "ROC" OR "predictive value" OR "concordance" OR "confirm" OR "negative detection rate" OR "negative referral rate" OR "false negative rate") AND (a) AND (d)
Phase 5: <u>Disease Control Studies</u>	**Primary aim:** To estimate the reductions in disease-associated mortality, morbidity, and disability afforded by biomarker testing.	("diagnosis" OR "detection") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life" OR "financial impact" OR "cost*" OR "effectiveness") AND (a) AND (d)
	Secondary aim 1: To obtain information about the costs of biomarker testing and treatment and the cost per life saved or per quality-adjusted life year	("diagnosis" OR "detection") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life" OR "outcome*") AND ("financial impact" OR "cost*" OR "effectiveness") AND (a) AND (d)
	Secondary aim 2: To evaluate compliance with testing and work-up in a diverse range of settings.	("diagnosis" OR "treatment") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life") AND ("primary care" OR "second level" OR "third level") AND "memory clinic" AND "cost*" AND (a) AND (d)
	Secondary aim 3: To compare different biomarker testing protocols and/or to compare different approaches to treating test positive subjects in regard to effects on mortality and costs.	"diagnosis" OR "treatment") AND ("protocol" OR "recommendation*" OR "criteria") AND ("benefit*" OR "compliance" OR "mortality" OR "morbidity" OR "QoL" OR "quality of life") AND ("financial impact" OR "cost*" OR "effectiveness") AND (a) AND (d)

(a)   ("Alzheimer*")
(b)   ("Healthy controls" OR "Cognitively normal" OR "controls" OR "normal")
(c)   ("MCI" OR "mild cognitive impairment" OR "prodromal")
(d)   ("MTA" OR "medial temporal" OR "hippocamp*")