**The Biomarker-based Diagnosis of Alzheimer’s Disease. 2 – Lessons From Oncology**

Marina Boccardi1,2, Valentina Gallo3, Yutaka Yasui4,5, Paolo Vineis6, Alessandro Padovani7, Urs Mosimann8,9, Panteleimon Giannakopoulos10, Gabriel Gold11, Bruno Dubois12, Clifford R. Jack Jr13, Bengt Winblad14, Giovanni B. Frisoni1,2,15, Emiliano Albanese16, for the Geneva Task Force for the Roadmap of Alzheimer’s Biomarkers17

1LANVIE – Laboratory of Neuroimaging of Aging, University of Geneva, Geneva, Switzerland

2Laboratory of Neuroimaging and Alzheimer Epidemiology, IRCCS S.Giovanni di Dio-Fatebenefratelli, Brescia, Italy

3Queen Mary, University of London, Barts and the London School of Medicine, Centre of Primary Care and Public Health, Blizard Institute, London, UK

4School of Public Health, University of Alberta, Alberta, Canada

5Department of Epidemiology & Cancer Control, St. Jude Children’s Research Hospital, Memphis, TN, USA

6School of Public Health, Imperial College London, London, UK

7Neurology Unit, Dept. Clinical and Experimental Sciences, University Health and Wealth of Brescia

8Gerontechnology & Rehabilitation Group, University of Bern, Bern, Switzerland

9University Hospital of Old Age Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland

10Department of Psychiatry, Faculty of Medicine, University of Geneva, Geneva, Switzerland

11Department of Internal Medicine, Rehabilitation and Geriatrics, University Hospitals and University of Geneva, 3 chemin du Pont-Bochet, 1226, Thônex, Switzerland

12Dementia Research Center and Department of Neurology; Salpêtrière University Hospital, Paris University, Paris, France

13Department of Radiology, Mayo Clinic, Rochester, MN, USA

14Karolinska Institutet, Dept NVS, Center for Alzheimer Research, Div of Neurogeriatrics, Huddinge, Sweden

15Memory Clinic, University Hospitals and University of Geneva, Switzerland

16Department of Psychiatry, WHO Collaborating Centre, University of Geneva, Switzerland

17http://centroalzheimer.it/public/MB/BM-Roadmap/The\_Geneva\_AD\_Biomarker\_Roadmap\_Task\_Force.docx

Abstract

Biomarkers for the diagnosis of Alzheimer’s disease (AD) are not yet validated for use in clinical settings. We aim to provide a methodological framework for their systematic validation, by reference to that developed for oncology biomarkers. As for this discipline, the steps for the systematic validation of AD biomarkers need to target analytical validity, clinical validity and clinical utility. However, the premises are different from oncology: the nature of disease (neurodegeneration *versus* cancer), the purpose (improve diagnosis in clinically affected *versus* screening preclinical individuals) and the target population (mild cognitive impairment patients referring to memory clinics *versus* general population) lead to important differences, influencing both the design of validation studies and the use of selected biomarkers. This framework is applied within a wider initiative to assess the current available evidence on the clinical validity of biomarkers for AD, for the final aim to identify gaps and research priorities, and to inform coordinated research efforts boosting AD biomarkers research.

Keywords: biomarkers, dementia, Alzheimer, early diagnosis, validation, methodology.

1. *From a pathological to a clinico/biological approach to the diagnosis of AD*

The aim of this paper is to define a methodological framework for the validation of biomarkers for diagnosing Alzheimer’s Disease (AD) in people referred to memory clinics or other specialist outpatient service, and meeting current diagnostic criteria of mild cognitive impairment (MCI) (Albert et al., 2011), which includes an important proportion of patients in the prodromal phase of AD (Dubois et al, 2007). Although the definite diagnosis of AD may be posed only after pathological confirmation, possible or probable AD can be diagnosed assessing the clinical features listed in widely accepted sets of diagnostic criteria (i.e., McKhann et al., 1984; 2011; the 10th edition of the International Classification of Diseases, ICD-10 (WHO - World Health Organization 1992), and the US Diagnostic and Statistical Manual of Mental Disorders, DSM-5 (APA - American Psychiatric Association 2013)). In the last decade evidence accumulated that the accuracy of the *in vivo* diagnosis can be improved using assays indicating the presence of the key pathological hallmarks of AD. These assays of *in vivo* biological or molecular characteristics of the pathological process underlying AD may be used as clinical biomarkers because they are associated with disease status or progression, and may capture biological responses to pharmacological or non-pharmacological interventions.

The presence of the main neuropathological AD changes, namely extra- (amyloid) and intra- (tau) cellular lesions, synaptic dysfunction and neuronal death may be identified during the long prodromal phase that precedes the clinical onset of AD. Methods of detection include direct evidence of brain amyloidosis and tau deposits from amyloid (Klunk *et al*., 2004) and tau ligands uptake at PET imaging (Villemagne *et al*., 2014; Johnson et al., 2016; Scholl et al., 2016; Schwarz et al., 2016; Brier et al., 2016), or indirect evidence such as the altered concentrations of the Abeta42 and tau proteins in CSF specimens (Blennow, *et al*., 1995, Iqbal and Grundke-Iqbal, 1997). The downstream synaptic dysfunction and loss of brain integrity may be identified using functional (FDG-PET, e.g. temporo-parietal hypometabolism; - de Leon, *et al*., 1983, McGeer, *et al*., 1986), and structural neuroimaging (MRI). In particular, medial temporal atrophy, assessed either visually (Scheltens *et al*., 2002) or quantitatively (Boccardi et al., 2015) has great prognostic value and automated quantification is now accessible to physicians through different services (Tanpitukpongse et al., 2017). Finally, other assays tapping pathophysiological processes (namely, degeneration of the dopaminergic nigrostriatal pathway with 123IMIBG scintigraphy - Treglia and Cason, 2012, and myocardial postganglionic sympathetic dysfunction with 123I-Ioflupane SPECT - Papathanasiou *et al*., 2012) may be used to exclude non-AD degenerative disorders (e.g., Lewy body dementia). The possibility to detect or exclude pathophysiological processes typical of AD has additional value considering that differential diagnosis is further complicated by the atypical presentations of AD (Dubois et al., 2010; McKhann et al., 2011).

The contribution of these AD biomarkers to improve the accuracy of the clinical diagnosis depends on the demonstration of their analytical validity, that is their ability to detect the key pathological hallmarks of AD and the correlated brain damage and dysfunctions. The available empirical evidence on the analytical and clinical validity of the aforementioned AD biomarkers is presented in the six reviews reported in this issue (Cerami *et al*.; Garibotto et al.; Sonni *et al*.; TenKate *et al*.; Mattsson *et al*.; Chiotis *et al*., 2017, in this issue). The availability of *in vivo* measures of AD pathology of proven analytical validity has the transformative potential to provide a diagnosis of AD based on a clinico-biological rather than clinico-pathological basis. Moreover, because the neuropathology underlying AD accumulates gradually over several decades and the insidious onset of the disease reflects the long induction and latency periods (Jack et al., 2013), the possibility to accurately measure AD-related brain changes *in vivo* can substantially contribute to the detection of AD at the preclinical stage when future curative treatments might be more efficacious.

* 1. *Relevance of early diagnosis*

Currently, the ability of biomarkers alone to predict the clinical expression of AD in not yet symptomatic individuals is not known (Sperling *et al*., 2011). Together with the lack of disease-modifying treatments, the use of biomarkers for population screening is currently not justified in terms of costs and benefits. On the other hand, improving diagnosis in people with the highest probability of having the disease, at the earliest possible time, is an aim that can and should be reasonably pursued in the short term. This can be done by targeting outpatient clinical settings, where people with a high probability of having AD seek medical advice often before the disease has impacted on autonomous living. The possibility to provide an early diagnosis in this context critically depends on two main factors. First, the affected people or their carers must *express their concern*. This can be negatively impacted by low awareness, limited understanding and stigma (WHO Dementia Report 2012) particularly in low- and middle-income countries (Albanese *et al.,* 2011). Secondly, the health system must *respond* *adequately*: such responsiveness can be improved using a biomarker-informed diagnostic approach. Consistent with this context, our effort focuses on the use of biomarkers for detecting AD in *symptomatic,* including mildly symptomatic individuals who (or whose caregivers) express concern for their decline. Indeed this scenario is frequent and has become the norm rather than the exception in high-income countries memory clinics. Providing an early diagnosis in symptomatic individuals seeking help would respond to an expressed, not proactively elicited concern, and would maximize benefits from the available knowledge and resources as to date.

We maintain that early diagnosis is potentially beneficial even in the absence of a disease-modifying drug. First, it provides an explanation for mild symptoms and perceived changes in cognitive function that some people recognize as problematic. Second, early diagnosis would maximize the possibility that interventions that are proven to ameliorate the quality of life of patients and families can be started as early as possible. In fact, early interventions may be beneficial in people at risk (Kivipelto et al., 2013; Ngandu *et al*., 2015;Andrieu et al., 2011; de Souto Barreto et al., 2016; Cesari et al., 2015), and may favor compression of cognitive morbidity (Langa *et al.,* 2008; Petersen et al., 2005), contributing to delaying disability with considerable savings on direct (e.g., institutionalization) and indirect (e.g. informal caregiving) costs. Moreover, past failure of AD randomized controlled trials (Winblad *et al*., 2016) was indeed ascribed, as a possible cause, to the lack of accurate recruitment procedures in confirming the presence of the brain damage that the drug under study was designed to treat or reduce (Mangialasche *et al*., 2010). Nowadays, biomarker-based diagnosis is being used to inform better study designs, particularly for the selection and inclusion of participants for experimental studies testing potential beneficial effects of interventions targeting specific disease mechanisms. On the clinicaltrials.gov web-site, a search using the string “Alzheimer AND anti-amyloid AND phase 3” (January 10, 2017) sorted out 3 non-overlapping studies, of which one (Solanezumab) includes amyloid-positivity at PET or CSF as an inclusion criterion. The same search using “anti-tau” found one study, testing NPT088, which is actually in Phase 1 but does use amyloid-PET positivity as an inclusion criterion. This use of biomarkers will definitely increase the power of clinical trials, however, their accuracy will critically influence the appropriateness of subjects selection and, thus, the final power of these studies.

*1.2 Objective*

The aim of this paper is to define a methodological framework for the validation of AD diagnostic biomarkers in people referred to a memory clinic or other specialist outpatient service and who meet current MCI diagnostic criteria (Albert *et al*., 2011). The ultimate goal is to improve early diagnosis of AD in memory clinics and to operationalize biomarker-based diagnostic criteria originally conceived for research (Dubois *et al.,* 2014; Albert *et al.,* 2011).

*2.0 Methods*

*2.1 Context of use*

Research on AD biomarkers needs to be conducted within a conceptual framework that specifies the *purpose* of their use, the *nature* of the disease, and the *population* in which these tests would be used (Zimmern, 2009).

First, as illustrated in the background, the *purpose* of using biomarkers in this context is to determine whether the impairment of clinically diagnosed MCI in patients accessing memory clinics is due to AD. The salient feature of the *nature* of the target disease is its insidious onset due to slowly progressive, well-characterized neurodegenerative processes that begin long before the overt clinical onset of AD. Still, the lack of disease modifying therapies, and the non–systematic link between biomarkers positivity and disease expression strongly discourage screening programs in the general population to identify cases proactively (Wimo and Winblad, 2015). As a consequence, our *population* of interest is the MCI population that is non-proactively screened but who is referred to a memory clinic or other third-level specialist health service, by a general practitioner or through another primary level entry point. People with subjective cognitive complaints, also accessing memory clinics, are not considered in this manuscript since their probability of having AD is lower than that of MCI.

*2.2 A methodological framework for setting systematic validation studies*

Three key aspects need to be separately and sequentially considered for biomarker evaluation: analytical validity, clinical validity and clinical utility. These are sequentially ordered, as the lack of more basic evidence would not justify the efforts for studies meant to progressively test the validity of the biomarker for implementation into clinical use.

In the following paragraphs we define and describe these evaluation phases, consistent with how these were originally featured in the field of genomics (Kroese *et al*., 2004), and later detailed for diagnostic molecular biomarkers (Zimmer *et al*., 2009).

*2.2.1 Analytical validity*

The terms biomarker and diagnostic test will be deliberately interchangeably used in this context (Zimmern, 2009). A clear distinction is needed, however, between a *test* (or *biomarker*) and an *assay*. The assay is the method used to ascertain whether a certain biological or molecular characteristic is present and/or to quantify its presence. The test (or biomarker) is the use of this quantitative measure with a predefined purpose, for a specific disease and in a given population. Analytical validity (i.e. accuracy) is demonstrated with respect to an existing standard, and is also present when the assay provides measurements with sufficient precision (i.e. reliability), that is consistent over time and in different contexts or circumstances. Although analytical validity may be seen as a mere technical step, it is a fundamental pre-requisite to determine the clinical validity and utility of a test, and it allows defining standards of use.

*2.2.2 Clinical validity*

The clinical validity of a test (or a biomarker) is its ability to detect the presence of a sign that is clearly separated from normality on the one hand, and from ‘adjacent’ signs (or proxies for diseases) on the other hand. Although this relies on its ability to measure the biological or molecular features that are specific to the disease, clinical validity is in essence the extent to which a test does measure what it sets out to measure rather than the mere underlying bio-molecular phenomena. Therefore, various aspects are implicated in clinical validity, from face and criterion validity, to construct and predictive validity. These validity facets may not need to be separately demonstrated, but rather explored jointly to determine the various aspects of the clinical validity of a test. It is noteworthy that the process used to explore a test’s construct validity is iterative in a manner comparable to the process of formulation-confirmation of hypotheses and theories in observational studies, because the construct that is intended to be measured is itself under investigation and can, or should, be progressively redefined on the basis of the empirical evidence. Therefore, epidemiological studies on broadly representative samples of the general population are needed to demonstrate the biomarker-disease association. In turn, observational research of biomarker-disease association must be informed by our progressively refined and improved understanding of the disease itself. This is particularly true in the case of AD, because, as discussed above, there is a marked, largely not understood, inter-individual variability in clinical expression amongst people with comparable brain damage.

Once the biomarker-disease association is established and understood, standard tests to determine the customary validity measures (i.e., sensitivity and specificity) should be conducted to formally explore how the test performs in practice. From these, positive and negative predictive values can be calculated under an appropriate study design to finally inform the interpretation of tests results (i.e., biomarker positivity) at an individual level and from a clinical perspective. It must be noted that the test properties can only be established with respect to an existing gold standard measure of reference.

*2.2.3 Clinical utility*

The clinical utility of a test is a function of the clinical implications of the results. Because of the inexorable progressive nature of AD and the absence of disease-modifying treatments, the main clinical implication at present is still limited. However, in the presence of brain damage the clinical expression of dementia not only varies between individuals but may also be modulated by factors that enhance or reduce cognitive reserve. As such, the purpose of the test is of paramount importance to establish its clinical utility, which can potentially be achieved even though the disease (i.e., MCI due to AD) is not yet fully understood.

 *2.3 Similar initiatives: the case of oncology*

In 2001, oncologists developed a 5-phases systematic framework for the validation of biomarkers, which has profoundly inspired the present work (Pepe *et al*., 2001). These phases map onto the consecutive steps required by the aforementioned *analytical validity*, *clinical validity* and *clinical utility*. However, relevant differences between the fields of oncology and AD should be considered. The *purpose* of the original biomarker framework consisted of validating and refining the performance of screening programs that could reliably detect the preclinical phases of *cancer* in the *general population*, while we focus our attention on early diagnosis of patients who already seek advice for cognitive symptoms.

The 5 phases framework for the validation of oncology biomarkers are the following (Pepe *et al*., 2001).

**Phase 1: “**Preclinical studies”. These are pilot studies to identify characteristics unique to tumor tissue that might lead to ideas for clinical tests for detecting cancer. Examples are the measurement of methylation signatures in cancer tissue and normal tissue from the same patients, and the identification of methylation changes in candidate genes (e.g., typically in relation to breast cancer (see for example Feinberg *et al*., 1983 as an early paper; and Cho *et al*., 2010). These observations in humans have been complemented by work in animals.

**Phase 2**: “Clinical Assay Development for Clinical Disease”. This phase is aimed to show that the clinical assay can distinguish subjects with cancer from those without cancer, in order to be considered promising for screening. This also includes the definition itself of an effective assay to be collected based on standard procedures throughout laboratories. Case-control studies to further develop the hypotheses generated in Phase 1 are developed, e.g., formal study designs comparing methylation levels of cancerous and non-cancerous tissue in the same people or specific methylation signatures of white blood cells within case control studies.

**Phase 3**: “Retrospective Longitudinal Repository Studies” provide evidence regarding the capacity of the biomarker to detect preclinical disease, by analyzing preclinical samples, e.g. through nested case-control studies in cohorts (as a recent example on breast cancer: van Veldhoven *et al*., 2015). Samples from cancer patients collected before disease onset at recruitment are compared with matched controls in the cohort studies. The methylation levels before disease onset are compared with those of people not developing cancer.

**Phase 4**: “Prospective Screening Studies” are performed in clinical or screening settings, and full diagnostic procedures are applied to positive individuals. Prospective nested designs provide demonstration of predictive value of epigenetic changes. For example, in lung cancer the same CpGs found to be associated with smoking (notably AHRR) are predictive of lung cancer prospectively, through statistical methylation analysis (Fasanelli *et al*, 2015). No similar examples are available yet for other cancers.

**Phase 5**: the “Cancer-Control Studies” phase addresses whether screening reduces the burden of cancer on the population. Overall societal benefit, compliance with (or actual implementation of) the screening program, and cost-effectiveness are assessed within this phase. These phases can be illustrated through the examples of several recent advancements on the development of biomarkers in oncology, e.g. in genomics (inherited gene variants and acquired mutational spectra) and in epigenomics - including DNA methylation and miRNA. The final aim is to apply the new markers, including epigenetic approaches, to complement early detection (e.g., CT scans for lung cancer Boeri *et al*., 2011; Sozzi et al, 2014, or ongoing trials involving miRNA markers, and driver mutations in premalignant lung lesions (Izumchenko et al, 2015). Ideally, these data will inform a diagnostic algorithm including current system for detecting cancer and can be used to predict cancer onset with potentially a greater sensitivity compared to current imaging methods.

Note that phase 1 and 2 encompass analytical validity, phase 3 and 4 clinical validity, and phase 5 includes studies assessing clinical utility.

*2.3.1 Impact of this framework on the oncology field*

The major impact of the use of this common framework in the field of oncology has been the stimulation of a concerted set of actions and research activities that has favored a harmonized, goal-oriented research agenda. A number of screening tools have been developed, validated and tested, some of which are currently in use to assist in the early detection of cancers in the general or in at-risk populations. The framework has also clarified the reciprocal importance of the five phases and the need to clearly separate the existing evidence for analytical validity from that of clinical validity. Hence, the framework reduced duplicated efforts and allowed to use systematically the large body of empirical evidence already available at the time to design studies more rapidly and fill the identified gaps.

*3. A 5-phases methodological framework for AD biomarkers diagnosis in the MCI population.*

The definition of an analogous framework in the field of AD is thus equally expected to facilitate the systematization of the existing empirical evidence, facilitating the evaluation of evidence on the validity of AD biomarkers as implemented within the Geneva Biomarker Roadmap Initiative (http://centroalzheimer.it/public/MB/BM-Roadmap/The\_Geneva\_AD\_Biomarker\_Roadmap\_Task\_Force.docx ) (Cerami *et al*.; Ten Kate et al, *et al*.; Sonni *et al*.; Garibotto *et al*.; Mattsson *et al*.; Chiotis *et al*., 2017, in this issue). We translated the described 5 phases from the oncology framework to the AD context as follows.

*3.1 Phase 1: Pilot studies*

*3.1.1 Phase 1 Aims*. Phase 1 aims to identify and prioritize promising biomarkers for the diagnosis of AD. The properties of healthy and diseased target tissue guide the recognition of the unique features that separate healthy and diseased subjects (morphological abnormalities or different expression of proteins).

*3.1.2 Phase 1 Design.* For each biomarker, Phase 1 is needed to estimate to what extent it distinguishes between cases and controls. This will be done calculating sensitivity, specificity, true positive ratio, false positive ratio for dichotomous measures and receiver operating characteristics (ROC) curves for continuous ones.

*3.1.3 Phase 1 Target Population*. The target sample includes specimens (brain or CSF tissue or cerebral function) from any stage of pathologically confirmed AD patients and healthy controls (or structural and functional comparison between affected/non affected tissues in the same patient).

*3.1.4 Phase 1 Gold standard*. The gold standard for phase 1 is pathology. Each assay has a different pathological gold standard. For instance, PET/CSF amyloid assays are specific to cortical neuritic plaques; MR hippocampal atrophy biomarkers reflect neuronal loss; FDG-PET temporoparietal hypometabolism biomarker is related to synaptic and neuronal loss; CSF total tau biomarkers to neuronal injury and phosphor-tau to neurofibrillary tangles formation (Zetterberg, 2017).

*3.1.5 Phase 1 Outcomes*. The outcome measures for Phase 1 are values of the assay in “target” versus healthy tissues (i.e. non-affected tissue from the same subject or same tissue from non-diseased subjects). The sequential spread of Alzheimer’s pathology through brain regions (Abeta, Tau, and neurodegeneration) is well known. Thus, in order to demonstrate their analytical validity, assays of Abeta, Tau, and neurodegeneration need to be consistent with this temporal and regional progression as a function of Thal’s (Thal *et al*., 2002) and Braak’s neuropathological stages (Braak & Braak, 1991). Atypical presentations of AD (Dubois et al., 2010, McKhann et al., 2011), as well as atypical atrophy (Caso et al., 2015; Buyn et al., 2015) and progression patterns (Buyn et al., 2015) should also be taken into account.

*3.2 Phase 2: Clinical Essay Development for Clinical Disease*

*3.2.1 Phase 2 Aims.* This phase aims to define non-invasive biomarkers that successfully distinguish between dementia cases and controls. True positive rate and false positive rate, or receiving operating characteristics curves for the essay will be estimated. A secondary aim of this phase is to optimize procedures for reproducing assays, to assess the consistency of the measurements taken from the non-invasive assay and from tissue, and to assess how covariates affect the status or level of the biomarker taken on the essay.

*3.2.2 Phase 2 Design*. Case-control studies, with cases defined by AD diagnosis and matched/unmatched controls.

*3.2.3 Phase 2 Target Population*. In the absence of pathological confirmation, we included clinically overt AD patients, and cognitively healthy persons. Note that comparisons of AD cases and controls from different source populations or using different study procedures do not qualify as a proper Phase 2 study. What mainly differentiates phase 1 from phase 2 is the use of surrogate markers in vivo, and the application of more stringent epidemiological criteria for the case-control study design. Every effort should be made to minimize selection and recall bias, and confounders should be carefully accounted for.

*3.2.4 Phase 2 Gold standard*. In the absence of a neuropathological diagnosis of AD, the clinical diagnosis of AD dementia is used as definition of caseness in this phase.

*3.2.5 Phase 2 Outcomes*. The examined outcome is the ability of biomarkers to discriminate cases from controls. The caveat of increased heterogeneity given by the fact that a clinical rather than neuropathological definition of the disease is used needs to be stressed here (Kovacs *et al*., 2013).

*3.3 Phase 3: Prospective Longitudinal Repository Studies*

*3.3.1 Phase 3 Aims*. This phase is designed to assess the ability of the biomarkers to detect the disease at the earliest possible phase, and ultimately sets out to demonstrate the predictive validity of the clinical test(s) in population-based representative samples. As defined in sections 1.1 and 2.1 (“*population of interest”*), although AD may now be detected even earlier, we limit our effort to the MCI population. Besides the mentioned convenience considerations, this is also consistent with the current clinical criteria (Albert et al., 2011; Sperling et al., 2011; Dubois et al., 2014). In oncology, the ability of biomarkers in the earliest phase has been more often assessed with case-control studies nested in cohorts with biological samples collected at recruitment. Very few studies would have this design in the field of MCI and AD dementia; as a consequence, by extension we include here all studies that have collected biomarkers in MCI populations and have collected follow-up data on their conversion to AD. In the oncology field, phase 3 is based on nested case-control studies of longitudinal cohorts where the disease status (cancer/no cancer) is known at time 0, blood samples were collected at time 0-n (5-10-20 years before), and the biomarker is measured *a posteriori* on the blood (hence the original notation of “retrospective longitudinal repository studies”). In the absence of assays repositories taken in asymptomatic subjects and stored prospectively, the adaptation to the field of AD of Phase 3 requires a prospective design, where time 0 is when patients have MCI, the biomarkers are measured, and disease status (i.e. incident dementia) is ascertained at time 0+n (generally after 1 to 5 years). Thus, the primary aims consist of evaluating the ability of the biomarker to detect the disease *in situ* in MCI, as the earliest possible phase of AD, and to define criteria for test positivity that accurately distinguish MCI due to AD from other diseases that may cause or explain the observed MCI phenotype. Secondary aims consist of the formal exploration of the impact of relevant covariates in the discriminatory ability of the biomarker, comparisons across the performances of different biomarkers to operate a progressive selection in the sake of parsimony, and finally the development and subsequent validation of sets of diagnostic criteria based on algorithms combining biomarkers for optimal performance. Since these studies are aimed to fine-tune the parameters of the biomarker, assays can be taken from everyday clinical context, however the biomarker is assessed only for experimental purposes and is not used to help diagnosis nor treatment of these patients. An example for this kind of study is CSF Abeta42 and T-tau assessment by Eckerstrom et al., 2000, where the same assays were used to define normality cut-offs based on the performance of Abeta42 and T-tau in separating converter and non-converter MCI subjects within the sample itself: both baseline and follow-up diagnoses were thus defined independently on CSF assays, and follow-up diagnosis served as the reference standard for the study.

*3.3.2 Phase 3 Design*. The study design in the Framework Phase 3 in AD is a prospective, cohort study. Data on the biomarkers and the clinical status of MCI are collected at baseline, the outcome being incident dementia (i.e. progression to AD). Note that Phase 3 studies are purely observational, that is the biomarkers results at baseline are not used for clinical diagnosis, prognosis, or treatment. At least 2 independent well designed and adequately powered studies may be required to guarantee replicability and generalizability of findings.

*3.3.3 Phase 3 Target Population*. The examined population consists of MCI subjects who are representative of those who meet the MCI criteria in the general population. The operationalization of the MCI diagnosis at this stage may not be specified and various neuropsychological profiles may be considered appropriate to best reflect the range of possible presentations at this stage.

*3.3.4 Phase 3 Gold standard*. Analogous to previous phases, Alzheimer’s pathology is the ideal gold standard, which would in essence confirm that the MCI phenotype is occurring in the presence of the AD neuropathology and can be ascribed to this. However, as said, the paucity of data, and limited number of autopsies make incident dementia or progression of cognitive impairment a necessary proxy of the ideal gold standard at this stage, acceptable in virtue of the progressive nature of the disease. Statistical power in these studies will depend not only on the sample size but also on the length of the follow-up periods. For instance, in our reviews (Garibotto *et al*.; Sonni *et al*.; Cerami *et al*.; Ten Kate *et al*.; Chiotis *et al*.; Mattsson *et al*., 2017, in this issue) we have considered studies with 3 years follow-up consistently with studies on MCI conversion in clinical samples (Landau *et al*., 2010).

*3.3.5 Phase 3 Outcome*. The outcome is the ability of biomarkers in predicting AD in people who meet MCI diagnostic criteria at baseline. Thus, the progression of MCI to non-AD dementia subtypes should also be clearly distinguished at the follow-up. This requires that differential diagnosis is performed in a valid and standardized manner, and according to existing diagnostic criteria (i.e. DSM-5 ‘Major Neurocognitive Disorder due to AD’).

*3.4 Phase 4: Prospective Diagnostic Studies*

*3.4.1 Phase 4 Aims*. The primary aim of this phase is to determine the operating characteristics of the biomarker in a representative population, by determining the true and false positive referral rates. Similar to studies that investigate interventions and treatment, Phase 4 is somewhat concerned with the exploration of effectiveness rather than efficacy. Secondary aims include the collection and interpretation of data about the feasibility and scalability of the diagnostic protocol, including the acceptability of the diagnostic procedures required to conduct the clinical tests, and the attached costs. The latter may vary as the clinical validity and utility of the biomarker is demonstrated and the indication of the use of biomarker in the clinical routine is established, its use consequently expanding to ‘spend to save’ thanks to the benefits of an improved early diagnosis. A final aim is the continuous and extended monitoring of disease occurrence in these representative samples in those whose diagnostic tests are negative in the face of the MCI phenotype.

This phase may appear similar to Phase 3. Indeed, in the case of AD both phase 3 and phase 4 studies target clinical MCI patients. However, unlike Phase 3, in Phase 4 patients are actually diagnosed (i.e. classified) and treated on the basis of the test (i.e. biomarker) results, which are also the object of the investigation in this phase. These studies are performed in a naturalistic context to safeguard the ecological conditions of the research, and the external validity of results. The internal validity may be somewhat lower because of less stringent sampling and selection procedures and greater measurement errors. This would be necessary to guarantee that MCI patients in Phase 4 are truly representative of everyday memory clinic patients, who may also have comorbidities, and for whom a lower compliance to the study protocol may be expected, compared to those examined in Phase 3 studies. Examples of Phase 4 studies are those currently assessing the incremental value of amyloid imaging. In these studies (Boccardi et al., JAMA Neurol 2016; Zwan et a., 2014), MCI cases from memory clinics undergo amyloid-PET at the end of the traditional diagnostic work-up. Physicians record diagnosis, diagnostic confidence and treatment plan at the end of the traditional work-up, and again after receiving amyloid-PET scan results. The latter diagnosis and treatment are those finally given to patients. Different from Phase 3 studies, participants also take a benefit from the exam that is being assessed to fulfill Phase 4 aims.

*3.4.2 Phase 4 Design*. Studies fulfilling Phase 4 aims are prospective cohort studies conducted in large and representative samples drawn from the target population.

*3.4.3 Phase 4 Target Population*. Same as in Phase 3. However, less stringent selection criteria may be used in Phase 4 studies due to the inevitably higher amount of comorbidity of patients representative of memory clinics populations, and the likely involvement of multiple study centers. These should recruit participants without operating exclusions possibly introducing unwanted bias, and reducing the external validity of results.

*3.4.4 Phase 4 Gold standard*. Same as in Phase 3, that is the clinical diagnosis of AD at follow-up.

*3.4.5 Phase 4 Outcome*. The outcomes assessed in Phase 4 studies are the proportion of cases correctly diagnosed, accounting for relevant baseline correlates (including age, sex and severity), and measures of compliance with the study protocol. Preliminary estimates of disease-associated morbidity, quality of life, and costs are also explored to evaluate the cost-effectiveness and scalability of the diagnostic procedures once these are translated into the clinical routine.

*3.5 Phase 5: Disease-Control Studies.*

*3.5.1 Phase 5 Aims*. Phase 5 studies aim to address whether the biomarker-based diagnosis is able to reduce the burden of Alzheimer’s dementia in the MCI population. Even if the examined biomarkers were able to detect the disease earlier, and at the MCI stage, there are several reasons why they may not have an overall benefit for the diagnosed population or the individual patient. Phase 5 corresponds conceptually to Phase 4 or post-marketing surveillance studies in the field of drug development. Thus, Phase 5 examines the benefits of the biomarker-based diagnostic program for the population of interest, and the factors that might impact on these benefits.

*3.5.2 Phase 5 Design*. A surveillance system should be devised to guarantee systematic, routine observational assessment of accepted practice, and indicators need to be considered on a local basis (and depending on the health and other monitoring systems already in place), accounting for their importance, and the feasibility of data collection and the limitations of the data themselves. Prospective studies conducted using observational routinely collected health data may be used in combination with purposely-designed cohort studies (Frisoni *et al*., in this issue).

*3.5.3 Phase 5 Target Population*. Same as Phase 4. Note that the sample size will plausibly be very large (range of thousands) implying that both academic and non-academic memory clinics (e.g. secondary referral centers) are involved in these monitoring steps.

*3.5.4 Phase 5 Outcome*. Assessed outcomes consist of mortality (or survival after the AD diagnosis), co–morbidity occurrence, detection and control, quality of life of patients and their caregivers, care needs and disability, direct and indirect costs of care (including diagnostic procedures and clinical follow-ups). Savings due to proper use of biomarkers may be reflected also in reduced costs due to possibly delayed clinically relevant transitions requiring increased time for caregiving, up to institutionalization. A proper computation of costs should be taken into account also to adjust the output of aims from previous phases, like the definition of an optimal algorithm (e.g., priority may be given to exams providing the same information with similar reliability but lower costs, as in the case of amyloidosis investigated using CSF samples or PET scan). Standard dissemination to memory clinics, harmonization of reports to patients, and access to user-friendly, cost effective and standard services to clinicians should also be systematically implemented within this phase.

*4. Discussion*

In this paper we described an adapted methodological framework aimed at systematizing the current research efforts and programs in the field of AD diagnostic biomarkers validation in the MCI population. The ultimate goal of the framework consists of the translation of current research diagnostic criteria that have introduced the use of AD biomarkers into clinical criteria to be formally adopted in memory clinics or other specialist health services beyond research uses. Further, the exact delineation of each and all phases of the framework and its adaptation to the field of AD research was combined with a series of up to date reviews to allow the identification of gaps in the evidence (Cerami *et al*.; Garibotto *et al*.; Sonni *et al*.; Ten Kate *et al*.; Chiotis *et al*.; Mattsson *et al*., in this issue), as well as the identification of the ethical implications of the current use of biomarkers (Porteri *et al*., in this issue). These gaps should be filled before completing downstream studies on the applicability of AD biomarkers in clinical practice. Collecting the adequate evidence of validity of biomarkers will eventually allow not only an improved and earlier diagnosis, but also a cost-effective refund strategy for biomarker-based diagnostic procedures.

To the best of our knowledge, and as evident from the results of the mentioned reviews on the development of biomarkers based on this methodological framework, no similar methods have explicitly guided the research on the clinical validity of AD biomarkers, and the available evidence appears to be the results of relatively uncoordinated efforts. On the other hand, our initiative is inspired and similar to that launched by oncologists in 2001 (Pepe *et al*, 2001). Our effort differs importantly from that of oncologists because the methodological framework should be applied to diagnosis and not to disease screening, and because of the different basic premises, i.e., nature of the disease, purpose and population defining the context of use. Another major difference between oncology and AD is that histological confirmation of the disease diagnosis is feasible *ex vivo* for oncology, but not for AD. Therefore the oncology analogy is only partly applicable to AD and the AD field will have to take a different approach to obtain reference standards in clinical and community-based samples. Of course the limited availability of neuropathological confirmation in AD makes this field prone to a greater extent of misclassification error of the outcome. We underline that, in our initiative, we did not address the fact that possible heterogeneities on the use of stains, antibodies, and on structural/morphological assessments may be a source of uncontrolled variability that can potentially influence downstream validation phases. Indeed they may require adequate systematization and standardization for the definition of the gold-standard itself, that anyway is not the specific topic of our effort. On the other hand, the possibility may be taken into account to perform studies where a strong reference-standard is provided by clinical diagnosis not only confirmed at follow-up but also confirmed by strong pathophysiological biomarkers, for example by prove of amyloid and tau positivity using PET scans. Although these exams are very promising, they are at the same time the object of our validation efforts nowadays, and proposing this kind of reference-standard at the moment may fault our methodological framework with recursive reasoning. However, it is conceivable that the fulfilment of proper validation of some biomarkers may in the near future supply an adequate biomarker-based gold-standard, that will greatly improve the current reference of clinical diagnosis while replacing pathology satisfactorily.

The defined framework denotes a logical sequence in the kind of studies to be performed for proper validation of biomarkers. Performance of downstream studies in the lack of “upstream” evidence might include unwanted noise or bias. An example is the detrimental variability of performance of FDG-PET in the early detection of AD (Smailagic et al., 2015), that can be attributed to the lack of a standard method for scan reading and of a standard threshold for positivity (Garibotto et al., in this issue).

The framework proposed here is of course open to dynamic and reciprocal interaction with the whole body of research currently performed. For example, large studies including the investigation of the natural history of AD (see as ongoing examples EPAD - <http://ep-ad.org/> - and AMYPAD - <http://www.amypad.eu/> ) can highlight and guide research for investigating the prognostic features of the disease detected by the biomarker (Phase 4, SA 1) and may, in turn, be influenced by our framework to structure the aims and sequence of investigations within the study. Moreover, our increasing knowledge of AD is pointing out that its different components may be heterogeneously present in different individuals, or appear with different sequence in time (Jack et al., 2013; Jack et al., 2016). Our framework refers to AD as a unitary pathophysiological process defined by presence of both amyloid and tau pathology. However, we believe that the proper validation of each individual biomarker with respect to the specific phenomenon that it is meant to measure guarantees its effectiveness independently on our future conceptualization of AD as a possibly more heterogeneous disorder. Indeed, this happened in the field of oncology, where the proper development of biomarkers (following the framework that inspired this initiative) indeed led to the current models of precision medicine that are proving increasingly effective in cancer therapy.

Since we recommend that dementia and MCI screening in the general population should be discouraged (Wimo and Winblad, 2015), we have intentionally focused on clinical samples. Nevertheless, future applications of the framework do not exclude epidemiological studies conducted in representative population-based samples of community dwelling healthy asymptomatic older adults. This is the most desirable final endpoint, especially considering future availability of disease-modifying drugs. The fact that the pathological hallmarks defining AD may be detected long before its clinical expression, and that many clinical trials failed possibly due to late interventions pose a strong rationale for pursuing a future, medium term goal of using biomarkers for preclinical diagnosis of AD. To achieve this end, the practical steps aimed to implement the present framework to the field of preclinical AD would be analogous to those that we have undertaken for MCI, i.e.: I) perform surveys of available literature on asymptomatic individuals, II) outline research priorities, and III) set new studies following the sequential steps outlined by this framework. Also in this case, attention should be paid to accurately distinguish the population of interest, as different population would convey different *a priori* probability of AD, which would bias results about biomarkers efficiency. For example, population at risk for AD (for being carriers of the ApoE4 allele, or due to recurrent cases in the family), or preclinical AD (for being carriers of dominant mutation causing AD), or SCI subjects (Jessen et al., 2014) may all heterogeneously differ from the general population. Thus the performance of biomarker test aimed to population screening should be examined in real population-based cohort and may not benefit of data already collected using samples from asymptomatic subjects conveying specific risk for AD, that are frequently investigated nowadays.

The approach proposed in this manuscript guarantees that the use of AD biomarkers will not lead to over-diagnosis (Moynihan *et al*., 2012) and will not divert precious resources that should prioritize health and social care aimed at improving the quality of life of MCI patients who may develop dementia, and their families.

*5. Glossary*

Alzheimer’s dementia. The clinical syndrome featuring acquired and progressive cognitive impairment associated with functional disability as defined by the NINCDS-ADRDA criteria (McKhann *et al*., 1984). Sixty-five to 80% of cases have Alzheimer’s pathology (plaques and tangles) and have Alzheimer’s disease. The diagnosis of Alzheimer’s disease dementia can be achieved in vivo by demonstrating positivity to biomarkers of Alzheimer’s pathology (decreased Abeta42 and increased tau and phospho-tau in the CSF, and increased uptake of amyloid and tau ligands on PET) (McKhann *et al*., 2011).

Alzheimer’s disease. Alzheimer’s pathology (plaques and tangles) isolated or associated with clinical symptoms (Dubois *et al*., 2016).

Alzheimer’s pathology. Extracellular senile plaques and intraneuronal neurofibrillary tangles, made of mainly beta-amyloid and hyperphosphorilated tau protein, respectively. Neurodegenerative changes are usually associated, co-localizing with tangles. The three pathological variants are recognized of typical (limbic and neocortical), limbic predominant, and hippocampal sparing (Murray *et al*., 2011).

Biomarker development. The process of discovery, analytical validation, clinical validation, and demonstration of clinical utility. A structured framework has been developed for oncology biomarker development made of the 5 phases of: (i) preclinical exploratory studies; (ii) clinical assay development for overt disease; (iii) prospective retrospective longitudinal repository studies; (iv) prospective case finding studies, and (v) disease control studies (Pepe at al., 2001).

Mild Cognitive Impairment. Syndrome featuring primary cognitive impairment but no disability. About half have Alzheimer’s pathology, 40% have no neurodegenerative disease (normal ageing), and 10% have neurodegenerative disease other than Alzheimer’s (hippocampal sclerosis, tau only dementia, primary age-related tauopathy, frontotemporal degeneration, pure cortical Lewy body disease, etc.) (Bennett *et al*., 2012; Rowe *et al*., 2010; Jack *et al*., 2008).

Roadmap. Objective-oriented, structured, and efficient action plan. In science & technology also called “strategic research agenda”

*6. Acknowledgements*

The Geneva Task Force for the Roadmap of Alzheimer’s Biomarkers includes the participants to a workshop held in Geneva on December 8-9, 2014. The P.I. of the Geneva Roadmap effort is Giovanni B Frisoni, with Bengt Winblad and Clifford R Jack Jr as co-PIs. The task force includes experts in biomarker development from the oncology community; experts on diagnostic AD biomarkers from Europe; representatives of pertinent scientific societies (Federation of European Societies of Neuropsychology - FENS, European Neurological Society of Neuroradiology - ENSNR, International Foundation of Clinical Chemistry and Laboratory Medicine - IFCCLM, European Association of Nuclear Medicine - EANM, and Swiss Federation of Clinical Neuro Societies - SFCNS); representatives of patient advocates, bioethicists and regulatory agencies, and early career researchers. The Geneva Task Force has been endorsed by the EADC – European Alzheimer’s Disease Consortium.

The workshop was funded thanks to a competitive grant by the Swiss National Science Foundation (“Early diagnosis of Alzheimer's disease with biomarkers: Now despite no cure, or later «only if»? International Exploratory Workshops - IZ32Z0\_157953) and unrestricted grants from: Alzheimer Forum Switzerland, Association pour la Recherche sur Alzheimer, Genève; Piramal, Eli Lilly & Company, General Electric, Guerbet, TEVA Pharma; Academie Suisse de Sciences Médicales, Vifor Pharma Switzerland., Novartis, Siemens, and IXICO. The Alzheimer's Association hosted the first follow-up meeting of the initiative at the 2015 AAIC congress in Washington. We acknowledge the help from Margherita Mauri and Daria Gennaro (IRCCS 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 from 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.

P Giannakopoulos' participation was supported by the Swiss National Foundation (SNF -3200B0-1161193 and SPUM 33CM30-124111). CR Jack Jr’s participation was supported by the National Institutes of Health (NIH - R01-AG011378, RO1-AG041851, U01-AG06786, U01-AG024904, R01-AG37551, R01-AG043392, R01-NS092625) and the Alexander Family Alzheimer's Disease Research.

**Conflict of interest**

CR Jack has provided consulting services for Eli Lilly and owns stock in Johnson and Johnson.

The other authors have no conflict of interest relative to the content of this paper.

**References**

Albanese E, Liu Z, Acosta D, Guerra M, Huang Y, Jacob KS, Jimenez-Velazquez IZ, Llibre Rodriguez JJ, Salas A, Sosa AL, Uwakwe R, Williams JD, Borges G, Jotheeswaran AT, Klibanski MG, McCrone P, Ferri CP, Prince MJ. Equity in the delivery of community healthcare to older people: findings from 10/66 Dementia Research Group cross-sectional surveys in Latin America, China, India and Nigeria. BMC Health Serv Res 2011;11:153.

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. 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. Alzheimers Dement 2011;7:270-9.

Andrieu S, Aboderin I, Baeyens JP, et al. IAGG workshop: health promotion program on prevention of late onset dementia. J Nutr Health Aging, 2011;15:562–575.

Blennow K, Wallin A, Agren H, Spenger C, Siegfried J, Vanmechelen E. Tau protein in cerebrospinal fluid: a biochemical marker for axonal degeneration in Alzheimer disease? Mol Chem Neuropathol 1995;26:231-45.

Boccardi M, Bocchetta M, Apostolova LG, Barnes J, Bartzokis G, Corbetta G, DeCarli C, deToledo-Morrell L, Firbank M, Ganzola R, Gerritsen L, Henneman W, Killiany RJ, Malykhin N, Pasqualetti P, Pruessner JC, Redolfi A, Robitaille N, Soininen H, Tolomeo D, Wang L, Watson C, Wolf H, Duvernoy H, Duchesne S, Jack CR Jr, Frisoni GB; EADC-ADNI Working Group on the Harmonized Protocol for Manual Hippocampal Segmentation. Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance. Alzheimers Dement 2015;11:126-38.

Boccardi M, Altomare D, Ferrari C, Festari C, Guerra UP, Paghera B, Pizzocaro C, Lussignoli G, Geroldi C, Zanetti O, Cotelli MS, Turla M, Borroni B, Rozzini L, Mirabile D, Defanti C, Gennuso M, Prelle A, Gentile S, Morandi A, Vollaro S, Dalla Volta G, Bianchetti A, Conti MZ, Cappuccio M, Carbone P, Bellandi D, Abruzzi L, Bettoni L, Villani D, Raimondi MC, Lanari A, Ciccone A, Facchi E, Di Fazio I, Rozzini R, Boffelli S, Manzoni L, Salvi GP, Cavaliere S, Belotti G, Avanzi S, Pasqualetti P, Muscio C, Padovani A, Frisoni GB, for the INDIA-FBP working group. The incremental diagnostic value of 18F-Florbetapir imaging in real-life memory clinic patients with cognitive impairment: the INDIA-FBP study. JAMA Neurology 2016 (in press).

Boeri, M., Verri, C., Conte, D., Roz, L., Modena, P., Facchinetti, F., Sozzi, G. MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. Proc Natl Acad Sci USA 2011;108:3713-8.

Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 1991;82:239-59.

Brier MR, Gordon B, Friedrichsen K, McCarthy J, Stern A, Christensen J, et al. Tau and Aβ imaging, CSF measures, and cognition in Alzheimer’s disease. Sci Transl Med 2016;8:338-66.

Byun MS, Kim SE, Park J, Yi D, Choe YM, Sohn BK, Choi HJ, Baek H, Han JY, Woo JI, Lee DY; Alzheimer’s Disease Neuroimaging Initiative. Heterogeneity of Regional Brain Atrophy Patterns Associated with Distinct Progression Rates in Alzheimer's Disease. PLoS One. 2015;10(11):e0142756.

Caso F, Agosta F, Mattavelli D, Migliaccio R, Canu E, Magnani G, Marcone A, Copetti M, Falautano M, Comi G, Falini A, Filippi M. White Matter Degeneration in Atypical Alzheimer Disease. Radiology. 2015;277(1):162-72.

Cerami C, Dubois B, Boccardi M, Monsch AU, Demonet JF, Cappa SF. 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 . Neurobiol Aging 2017 (in this issue).

Cesari M, Vellas B, Hsu FC, Newman AB, Doss H, King AC, Manini TM, Church T, Gill TM, Miller ME, Pahor M; LIFE Study Group. A physical activity intervention to treat the frailty syndrome in older persons-results from the LIFE-P study. J Gerontol A Biol Sci Med Sci 2015;70:216-22.

Chiotis K, Saint-Aubert L, Boccardi M, Gietl A, Picco A, Varrone A, et al. 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 Neurobiol Aging 2017 (in this issue).

Cho YH, Yazici H, Wu HC, Terry MB, Gonzalez K, Qu M, et al. Aberrant promoter hypermethylation and genomic hypomethylation in tumor, adjacent normal tissues and blood from breast cancer patients. Anticancer Res 2010;30:2489–96

de Leon, M.J., George, A.E., Ferris, S.H., Rosenbloom, S., Christman, D.R., Gentes, C.I., Reisberg, B., Kricheff, II, Wolf, A.P. Regional correlation of PET and CT in senile dementia of the Alzheimer type. AJNR Am J Neuroradiol 1983;4:553-6.

de Souto Barreto P, Denormandie P, Lepage B, Armaingaud D, Rapp T, Chauvin P, Vellas B, Rolland Y. Effects of a long-term exercise programme on functional ability in people with dementia living in nursing homes: Research protocol of the LEDEN study, a cluster randomised controlled trial. Contemp Clin Trials. 2016;47:289-95.

Dubois B, Feldman HH, Jacova C, Dekosky ST, 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 PJ, Scheltens P. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007;6:734-46.

Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ, Scheltens P. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurol. 2010;9(11):1118-27.

Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, DeKosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert MO, Jicha GA, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 2014;13:614-29.

Eckerstrom C, Andreasson U, Olsson E, Rolstad S, Blennow K, Zetterberg H, Malmgren H, Edman A, Wallin A. Combination of hippocampal volume and cerebrospinal fluid biomarkers improves predictive value in mild cognitive impairment. Dementia and geriatric cognitive disorders 2010;29:294-300.

Fasanelli F, Baglietto L, Ponzi E, Guida F, Campanella G, Johansson M, Grankvist K, Johansson M, Assumma MB, Naccarati A, Chadeau-Hyam M, Ala U, Faltus C, Kaaks R, Risch A, De Stavola B, Hodge A, Giles GG, Southey MC, Relton CL, Haycock PC, Lund E, Polidoro S, Sandanger TM, Severi G, Vineis P. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. Nat Commun 2015;6:10192.

Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 1983;301:89–92

Garibotto V, Herholtz K, Boccardi M, Picco A, Varrone A, Nordberg A, et al. Clinical validity of FDG-PET as a biomarker for Alzheimer’s disease in the context of a structured 5 –phase development framework. Neurobiol Aging 2017 (in this issue).

Frisoni GB, D Perani, S Bastianello, G Bernardi, C Porteri, M Boccardi, SF Cappa, M Trabucchi, A Padovani. Biomarkers for the diagnosis of Alzheimer’s disease in clinical practice: the Italian inter-societal roadmap. Neurobiol Aging 2017 (in this issue).

Iqbal K, Grundke-Iqbal I. Elevated levels of tau and ubiquitin in brain and cerebrospinal fluid in Alzheimer's disease. Int Psychogeriatr. 1997;9 Suppl 1:289-96; discussion 317-21.

Izumchenko E, Chang X, Brait M, Fertig E, Kagohara LT, Bedi A, Marchionni L, Agrawal N, Ravi R, Jones S, Hoque MO, Westra WH, Sidransky D. Targeted sequencing reveals clonal genetic changes in the progression of early lung neoplasms and paired circulating DNA. Nat Commun 2015;6:8258.

Jessen F, Amariglio RE, van Boxtel M, Breteler M, Ceccaldi M, Chételat G, Dubois B, Dufouil C, Ellis KA, van der Flier WM, Glodzik L, van Harten AC, de Leon MJ, McHugh P, Mielke MM, Molinuevo JL, Mosconi L, Osorio RS, Perrotin A, Petersen RC, Rabin LA, Rami L, Reisberg B, Rentz DM, Sachdev PS, de la Sayette V, Saykin AJ, Scheltens P, Shulman MB, Slavin MJ, Sperling RA, Stewart R, Uspenskaya O, Vellas B, Visser PJ, Wagner M; Subjective Cognitive Decline Initiative (SCD-I) Working Group. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers Dement. 2014;10:844-52.

Kivipelto M, Solomon A, Ahtiluoto S, et al. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress. Alzheimers Dement 2013; 9:657–65.

Jack CR Jr, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 2013;12:207-16.

Jack CR Jr, Wiste HJ, Weigand SD, Knopman DS, Lowe V, Vemuri P, Mielke MM, Jones DT, Senjem ML, Gunter JL, Gregg BE, Pankratz VS, Petersen RC. Amyloid-first and neurodegeneration-first profiles characterize incident amyloid PET positivity. Neurology. 2013;81:1732-40.

Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, Hampel H, Jagust WJ, Johnson KA, Knopman DS, Petersen RC, Scheltens P, Sperling RA, Dubois B. A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology. 2016;87(5):539-47.

Johnson KA, Shultz A, Betensky RA, Becker JA, Sepulcre J, Rentz DM, et al. Tau positron emission tomographic imaging in aging and early Alzheimer's disease. Ann Neurol 2016;79:110-9.

Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Bergstrom M, Savitcheva I, Huang GF, Estrada S, Ausen B, Debnath ML, Barletta J, Price JC, Sandell J, Lopresti BJ, Wall A, Koivisto P, Antoni G, Mathis CA, Langstrom B. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Annals of neurology 2004;55:306-19.

Kovacs GG, Milenkovic I, Wöhrer A, Höftberger R, Gelpi E, Haberler C, Hönigschnabl S, Reiner-Concin A, Heinzl H, Jungwirth S, Krampla W, Fischer P, Budka H. Non-Alzheimer neurodegenerative pathologies and their combinations are more frequent than commonly believed in the elderly brain: a community-based autopsy series. Acta Neuropathol 2013;126:365-84.

Kroese M, Zimmern R, Sanderson S. Genetic tests and their evaluation: Can we answer the key questions? Genet Med 2004;6:475–80.

Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, Petersen RC, Shaw LM, Trojanowski JQ, Jack CR Jr, Weiner MW, Jagust WJ; Alzheimer's Disease Neuroimaging Initiative. Comparing predictors of conversion and decline in mild cognitive impairment. Neurology. 2010;75:230-8.

Langa KM1, Larson EB, Karlawish JH, Cutler DM, Kabeto MU, Kim SY, Rosen AB. Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity? Alzheimers Dement. 2008;4:134-44.

Mangialasche F, Solomon A, Winblad B, Mecocci P, Kivipelto M. Alzheimer's disease: clinical trials and drug development. Lancet Neurol. 2010;9:702-16.

Mattsson N, Lönnborg A, Boccardi M, Blennow K Hansson O. 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, 2017 (in this issue).

McGeer PL, Kamo H, Harrop R, McGeer EG, Martin WR, Pate BD, Li DK. Comparison of PET, MRI, and CT with pathology in a proven case of Alzheimer's disease. Neurology 1986;36:1569-74.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939-44.

McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:263-9.

Moynihan R. Preventing overdiagnosis: the myth, the music, and the medical meeting. BMJ. 2015;350:h1370.

Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, Bäckman L, Hänninen T, Jula A, Laatikainen T, Lindström J, Mangialasche F, Paajanen T, Pajala S, Peltonen M, Rauramaa R, Stigsdotter-Neely A, Strandberg T, Tuomilehto J, Soininen H, Kivipelto M. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet. 2015;385:2255-63.

Papathanasiou ND, Boutsiadis A, Dickson J, Bomanji JB. Diagnostic accuracy of 123I-FP-CIT (DaTSCAN) in dementia with Lewy bodies: a meta-analysis of published studies. Parkinsonism Relat Disord 2012;18:225–9.

Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93:1054-61.

Petersen RC, Thomas RG, Grundman M, Bennett D, Doody R, Ferris S, Galasko D, Jin S, Kaye J, Levey A, Pfeiffer E, Sano M, van Dyck CH, Thal LJ; Alzheimer's Disease Cooperative Study Group. Vitamin E and donepezil for the treatment of mild cognitive impairment. N Engl J Med 2005;352:2379-88.

Porteri C, Albanese E, Scerri C, Carrillo MC, Snyder HM, Martensson B, Baker M, Giacobini E, Boccardi M, Winblad B, Frisoni GB, Hurst S, and the Geneva Task Force for the Roadmap of Alzheimer’s Biomarkers. The Biomarker-based Diagnosis of Alzheimer’s Disease. 1 – Ethical and Societal Issues. Neurobiology of Aging (2017; in this issue,)

Scheltens P, Fox N, Barkhof F, De Carli C. Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. The Lancet Neurology 2002;1:13–21.

Scholl M, Lockhart SN, Schonhaut DR, O'Neil JP, Janabi M, Ossenkoppele R, Baker SL, Vogel JW, Faria J, Schwimmer HD, Rabinovici GD, Jagust WJ. PET Imaging of tau deposition in the aging human brain. Neuron 2016;89:971-82.

Schwarz AJ, Yu P, Miller BB, Shcherbinin S, Dickson J, Navitsky M, Joshi AD, Devous MD Sr, Mintun MS. Regional profiles of the candidate tau PET ligand 18F-AV-1451 recapitulate key features of Braak histopathological stages. Brain 2016; 139:1539-50.

Smailagic N, Vacante M, Hyde C, Martin S, Ukoumunne O, Sachpekidis C. ¹⁸F-FDG PET for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2015;1:CD010632.

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. Clinical validity of presynaptic dopaminergic imaging with 123I-ioflupane and noradrenergic imaging with 123I-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 (2017, in this issue).

Sozzi G, Boeri M, Rossi M, Verri C, Suatoni P, Bravi F, Roz L, Conte D, Grassi M, Sverzellati N, Marchiano A, Negri E, La Vecchia C, Pastorino U. Clinical utility of a plasma-based miRNA signature classifier within computed tomography lung cancer screening: a correlative MILD trial study. J Clin Oncol 2014;32:768-73.

Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:280-92.

Tanpitukpongse TP1, Mazurowski MA1, Ikhena J1, Petrella JR2. Predictive Utility of Marketed Volumetric Software Tools in Subjects at Risk for Alzheimer Disease: Do Regions Outside the Hippocampus Matter? AJNR Am J Neuroradiol. 2017.

Ten Kate M, Barkhof F, Boccardi M, Visser PJ, Lovblad KO, Jack CR, et al., and the Geneva Task Force for the Roadmap of Alzheimer’s Biomarkers. Clinical validity of medial temporal atrophy as a biomarker for Alzheimer’s disease in the context of a structured 5–phase development framework Neurobiol Aging 2017 (in this issue).

Thal DR, Rüb U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology. 2002;58:1791-800.

Treglia G, Cason E. Meta-analysis on MIBG scintigraphy in differential diagnosis between Parkinson’s disease and neurodegenerative parkinsonism. Parkinsonism Relat Disord 2012;18:805; author reply 806.

van Veldhoven K, Polidoro S, Baglietto L, Severi G, Sacerdote C, Panico S, Mattiello A, Palli D, Masala G, Krogh V, Agnoli C, Tumino R, Frasca G, Flower K, Curry E, Orr N, Tomczyk K, Jones ME, Ashworth A, Swerdlow A, Chadeau-Hyam M, Lund E, Garcia-Closas M, Sandanger TM, Flanagan JM, Vineis P. Epigenome-wide association study reveals decreased average methylation levels years before breast cancer diagnosis. Clin Epigenetics 2015;7:67.

Villemagne VL, Furumoto S, Fodero-Tavoletti MT, Mulligan RS, Hodges J, Harada R, et al. In vivo evaluation of a novel tau imaging tracer for Alzheimer's disease. Eur J Nucl Med Mol Imaging 2014;41:816-26.

WHO Dementia report 2012. World Health Organization. Dementia: a public health priority. 2012, Geneva, Switzerland. <http://apps.who.int/iris/bitstream/10665/75263/1/9789241564458_eng.pdf?ua=1>

Wimo A, Winblad B. Brain health: a primary health care viewpoint. J Am Med Dir Assoc 2015;16:720-1.

Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, Cedazo-Minguez A, Dubois B, Edvardsson D, Feldman H, Fratiglioni L, Frisoni GB, Gauthier S, Georges J, Graff C, Iqbal K, Jessen F, Johansson G, Jönsson L, Kivipelto M, Knapp M, Mangialasche F, Melis R, Nordberg A, Rikkert MO, Qiu C, Sakmar TP, Scheltens P, Schneider LS, Sperling R, Tjernberg LO, Waldemar G, Wimo A, Zetterberg H. Defeating Alzheimer's disease and other dementias: a priority for European science and society. Lancet Neurol 2016;15:455-532.

Zetterberg H. Review: Tau in biofluids - relation to pathology, imaging and clinical features. Neuropathol Appl Neurobiol. 2017. doi: 10.1111/nan.12378.

Zimmern RL. Testing challenges: evaluation of novel diagnostics and molecular biomarkers. Clin Med (Lond) 2009;9:68-73.

Zwan MD, Bouwman FH, Van der Flier WM, Lammertsma A, Van Berckel B, Scheltens P. Diagnostic value of amyloid imaging in early onset dementia. Alzheimers Dement 2014;10:P14.