Neurobiology of Aging 52 (2017) 153–166

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Roadmap to Alzheimer's Biomarkers in the Clinic

Clinical validity of delayed recall tests as a gateway biomarker for Alzheimer's disease in the context of a structured 5-phase development framework

Chiara Cerami [a,b,c,*], Bruno Dubois [d], Marina Boccardi [e,f], Andreas U. Monsch [g], Jean Francois Demonet [h], Stefano F. Cappa [b,i], for the Geneva Task Force for the Roadmap of Alzheimer's Biomarkers

[a] Vita-Salute San Raffaele University, Milan, Italy
[b] Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy
[c] Clinical Neuroscience Department, San Raffaele Turro Hospital, Milan, Italy
[d] Dementia Research Center and Department of Neurology, Salpêtrière University Hospital, Paris University, Paris, France
[e] Laboratory of Neuroimaging & Alzheimer's Epidemiology, IRCCS Centro S. Giovanni di Dio Fatebenefratelli, Brescia, Italy
[f] LANVIE (Laboratory of Neuroimaging of Aging) – Department of Psychiatry, University of Geneva, Geneva, Switzerland
[g] Memory Clinic, University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland
[h] Leenaards Memory Centre – CHUV, Lausanne University, Lausanne, Switzerland
[i] NeTS Center, Istituto Universitario di Studi Superiori, Pavia, Italy

ARTICLE INFO

Article history:
Received 23 November 2015
Received in revised form 5 March 2016
Accepted 22 March 2016

Keywords:
Episodic memory
Delayed recall
Cued recall tasks
Free word list tests
Biomarker development
Early diagnosis
Biomarker-based diagnosis
Alzheimer's disease
Mild cognitive impairment
5-Phases

ABSTRACT

Although Alzheimer's disease criteria promote the use of biomarkers, their maturity in clinical routine still needs to be assessed. In the light of the oncology framework, we conducted a literature review on measures used to assess delayed recall impairment due to medial temporal lobe dysfunction (i.e., free and cued word list recall tests). Ample evidence is available for phases 1 (rationale for use), 2 (discriminative ability), and 3 (early detection ability) for many of the tests in routine use. Evidence about phase 4 (performance in real world) and phase 5 (quantify impact and costs) is yet to come. Administration procedures have been standardized and cutoff scores are well validated in large Alzheimer's disease and mild cognitive impaired series. Some aspects (e.g., different task formats), however, hamper the comparability of results among different populations and the reproducibility between laboratories. No definite guideline for their use can thus be proposed at the moment. Accordingly, the maturity of such markers is not yet sufficient and requires future investigation to promote the proper use of memory measures in clinical settings.

1. Introduction

The correct identification of Alzheimer's disease (AD) represents a challenge for clinicians especially at the predementia stages. Recent developments in this area of research are specifically devoted to support the early identification of AD pathology in vivo and to the application of reliable biomarkers of disease in clinical settings. The need of more accurate early and differential diagnosis, indeed, prompted the development of new research criteria supporting the use of biomarkers in order to recognize AD in prodromal or even preclinical stages (Albert et al., 2011; Dubois et al., 2007, 2010, 2014; Jack et al., 2011; McKhann et al., 2011; Sperling et al., 2011).

Since the introduction of new clinical criteria, AD research has been mainly focused on the application of distinctive topographic (e.g., 18F-fluorodeoxyglucose positron emission tomography [FDG-PET]; magnetic resonance imaging [MRI]) and pathophysiological (e.g., amyloid-PET or cerebrospinal fluid [CSF]) biomarkers in clinical research. Several methodological problems have emerged about their implementation in clinical routine. Neither a definite diagnostic algorithm nor clear quantitative measures for the use of biomarkers in patients suspected for AD have been clearly outlined.

* Corresponding author at: Division of Neuroscience, Università Vita-Salute San Raffaele and San Raffaele Scientific Institute, Via Olgettina, 60, 20132 Milan, Italy. Tel.: +39 0226435760; fax: +39 0226435738.
E-mail address: cerami.chiara@hsr.it (C. Cerami).

0197-4580/$ – see front matter © 2016 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2016.03.034

Thus, different biomarkers and measurement tools are used by researchers according to their availability in community-based and clinical-based studies, obviously resulting in heterogeneous findings.

To overcome similar problems in the field of oncology, Pepe et al. (2001) suggested systematizing the investigation of cancer biomarkers on the basis of the methodology used for pharmacological investigation. Since a formal structure to guide the process of AD biomarker development was lacking so far, an effort has recently been launched to adopt the previously mentioned oncology model to effectively systematize the available scientific evidence for the use of biomarkers in AD diagnosis, with the aim to promote rigor in their application to clinical settings. The present study focuses on the analysis of the maturity of the assessment of episodic memory by means of delayed recall tasks in the framework of this model.

Whether cognitive testing can properly be considered as a marker of disease is an open question. The concept is fully compatible with a broad definition, such as "…a characteristic which can be objectively measured and evaluated as an indicator of a physiological as well as a pathological process or pharmacological response to a therapeutic intervention" (Jain, 2012). In the field of dementia, however, neuropsychological testing is generally considered separately from imaging and CSF biomarkers (e.g., Ewers et al., 2012). Within the MCI and/or prodromal AD context, neuropsychological testing is usually considered as a sort of "gatekeeper" for the application of biomarkers, as the presence of objective impairment is required by the diagnostic criteria to separate these conditions from subjective complaints. In any case, given the central role of cognitive assessment, it is surprising that studies assessing the sensitivity and specificity of neuropsychological tests for diagnosis of AD or the predictive value of the progression from mild cognitive impairment (MCI) to AD are relatively scant and heterogeneous in methodology. In particular, the presence of an early and significant objective deficit of memory and learning has been considered as the main criterion supporting the diagnosis of typical AD condition for decades (APA, 2000; McKhann et al., 1984; Diagnostic and Statistical Manual of Mental Disorders [DSM-IV]), and an impaired memory performance in comparison to a healthy control group is considered as the best cognitive predictor of the development of future AD (Elias et al., 2000; Sarazin et al., 2007; Small et al., 2000). According to Braak and Braak staging of AD (Braak and Braak, 1991), the earliest neuropathological changes involve the entorhinal cortex and the hippocampal structures, disconnecting the Papez circuit and selectively affecting the ability to consolidate new information. This results in an impaired performance on delayed recall memory tasks (Squire et al., 2004).

Moreover, in the context of recent diagnostic criteria for the very early and/or prodromal stages of AD, the performance in specific memory tests has gained a special status. The presence of impaired memory performance on objective testing is required for the definition of MCI (hence "amnesic MCI") or prodromal AD status. If a subject presents with subjective memory complaints, in the absence of objective memory dysfunction, he or she is defined as having a "subjective memory impairment", an inconsistently defined construct (Abdulrab and Heun, 2008), which in clinical practice does not usually lead to further investigation but only to reassurance and long-term follow-up (Berrios et al., 2000; Jessen et al., 2014).

Since cognitive assessment remains a critical component of diagnosis in clinical and research settings, it is vital to determine the capacity of specific memory measures in detecting early disease changes and predicting disease progression, in order to recommend tests having the greatest predictive accuracy. The tasks assessing memory ability may differ with respect to the modality of stimulus presentation, the testing procedure, the structure of the to-be-remembered information or the presence of facilitators to improve encoding and recall (cued paradigms).

Many memory measures are used in clinical and research settings. The most common are measures of delayed free recall of word lists. A variety of standardized verbal learning tasks, such as the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1941), the California Verbal Learning Test (CVLT; Delis et al., 1987, 2000; CVLT-II, Delis et al., 2000), and the Hopkins Verbal Learning Test (Brandt and Benedict, 2001), are commonly used for clinical diagnosis and disease monitoring of AD dementia and mild cognitive impaired (MCI) patients. In addition, some word list tasks, often quicker to administer, that is, with fewer words to learn (10-word list) and less learning trials (2 or 3), are part of neuropsychological batteries, such as the Wechsler Memory battery (http://www.pearsonclinical.com/), Alzheimer's Disease Assessment Scale-cognition (ADAS-cog) (Mohs et al., 1983), Consortium to establish a registry for Alzheimer's disease (CERAD) (Morris et al., 1988), or Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). Direct measurement of the number of items recalled at the learning trials (i.e., the immediate recall) or after a time delay (i.e., the delayed recall), as well as the difference between immediate and delayed recall (i.e., the savings) are the main measures obtained from the word list free recall tasks used to evaluate patient performances.

Other neuropsychological tests for the assessment of verbal long-term memory are logical memory (short story recall) and associative learning tasks (e.g., in the Wechsler Memory battery). Both immediate and delayed scores are obtained from story recall tests. In this case, the processing of a coherent stream of information typically benefits from intrinsic semantic organization of the material, while the ability to self-generate organizational strategies is required for free recall of word list tasks—with the exception of the CVLT (Randolph et al., 1994). Other tests assess nonverbal memory, such as, the delayed recall of Rey figure or the Cambridge Neuropsychological Test Automated Battery (http://www.cambridgecognition.com/), testing the ability to form and remember associations between the attributes of an experience. These tasks are sensitive to the functional integrity of the medial temporal lobe (Eichenbaum and Cohen, 2001; Meltzer and Constable, 2005), as performance largely depends on the ability to bind and encode arbitrary information.

Though the amnestic syndrome is typical of the onset of AD, impairments on delayed recall tasks may also be present in non-AD neurodegenerative diseases (e.g., behavioral variant of frontotemporal dementia; Hornberger et al., 2010) as well as in other conditions (e.g., vascular MCI, depression), characterized by cognitive deficits that can affect the learning phase or the encoding and recall processes (Dickerson and Eichenbaum, 2010). In AD, the amnestic profile is typically characterized by poor learning and rapid forgetting over relatively short periods, reflecting damage to the hippocampal structures (Squire et al., 2004). To accurately detect memory impairment of the hippocampal type, the design of the test used to assess memory ability, and in particular of the learning phase of the task, is crucial. Free recall is dependent upon intact attentional processing, registration, and retrieval mechanisms. An effective retrieval of the to-be-remembered information can be better achieved with an "encoding specificity", based on the correspondence of the semantic cue during encoding and retrieval. The use of this learning technique produces efficient results in healthy subjects (Ivnik et al., 1997). In agreement to this evidence, the specific neuropsychological features of the memory impairment of AD has been stressed in the International Working Group research diagnostic criteria (Dubois et al., 2010, 2014) and defined as follows: "objective evidence of significantly impaired episodic memory on testing, generally consisting of a recall deficit that does not improve significantly with cueing or recognition testing after