Alzheimer's & Dementia 10 (2014) 421–429

Featured Articles

Coalition Against Major Diseases/European Medicines Agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease

Derek L. G. Hill[a], Adam J. Schwarz[b], Maria Isaac[c], Luca Pani[c], Spiros Vamvakas[c], Robert Hemmings[c], Maria C. Carrillo[d], Peng Yu[b], Jia Sun[b,e], Laurel Beckett[f], Marina Boccardi[g], James Brewer[h], Martha Brumfield[i], Marc Cantillon[j], Patricia E. Cole[b], Nick Fox[k], Giovanni B. Frisoni[g], Clifford Jack[l], Thomas Kelleher[m], Feng Luo[m], Gerald Novak[n], Paul Maguire[o], Richard Meibach[p], Patricia Patterson[q], Lisa Bain[r], Cristina Sampaio[s], David Raunig[t], Holly Soares[m], Joyce Suhy[u], Huanli Wang[f], Robin Wolz[a,v], Diane Stephenson[i,*]

[a]*IXICO Ltd., London, UK*
[b]*Eli Lilly and Company, Indianapolis, IN, USA*
[c]*European Medicines Agency, London, UK*
[d]*Alzheimer's Association, Chicago, IL, USA*
[e]*The University of Texas School of Public Health, Houston, TX, USA*
[f]*University of California, Davis, CA, USA*
[g]*IRCCS Fatebenefratelli, Brescia, Italy*
[h]*University of California, San Diego, CA, USA*
[i]*Coalition Against Major Diseases, Critical Path Institute, Tucson, AZ, USA*
[j]*Wellnessmanagments.com, Livingston, NJ, USA*
[k]*UCL Institute of Neurology, London, UK*
[l]*Mayo Clinic, Rochester, MN, USA*
[m]*Bristol Myers Squibb, Wallingford, CT, USA*
[n]*Janssen Pharmaceutical Research and Development, Titusville, NJ, USA*
[o]*Novartis, Basel, Switzerland*
[p]*Novartis, NJ, USA*
[q]*AstraZeneca, Wilmington, DE, USA*
[r]*Independent science writer, Elverson, PA, USA*
[s]*CHDI Foundation Inc., Princeton, NJ, USA*
[t]*ICON Medical Imaging, Warrington, PA, USA*
[u]*SYNARC Inc., Newark, CA, USA*
[v]*Department of Computing, Imperial College London, London, UK*

Abstract

Background: Regulatory qualification of a biomarker for a defined context of use provides scientifically robust assurances to sponsors and regulators that accelerate appropriate adoption of biomarkers into drug development.

Methods: The Coalition Against Major Diseases submitted a dossier to the Scientific Advice Working Party of the European Medicines Agency requesting a qualification opinion on the use of hippocampal volume as a biomarker for enriching clinical trials in subjects with mild cognitive impairment, incorporating a scientific rationale, a literature review and a de novo analysis of Alzheimer's Disease Neuroimaging Initiative data.

The views expressed in this article are the personal views of the authors and may not be understood nor quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties, or any of the national agencies.

*Corresponding author. Tel.: +520-382-1405.
E-mail address: DStephenson@c-path.org

http://dx.doi.org/10.1016/j.jalz.2013.07.003

Results: The literature review and de novo analysis were consistent with the proposed context of use, and the Committee for Medicinal Products for Human Use released an opinion in November 2011.
Conclusions: We summarize the scientific rationale and the data that supported the first qualification of an imaging biomarker by the European Medicines Agency.

Keywords: Alzheimer's disease; Hippocampal volume; Mild cognitive impairment; Alzheimer's Disease Neuroimaging Initiative

1. Introduction

Decreased hippocampal volume (HCV) is one of the best established biomarkers used in research studies to stage the progression of Alzheimer's disease (AD) pathology in the brain of patients across the spectrum of the disease [1,2]. A supporting body of literature over approximately 20 years indicates that changes in HCV are most rapid around the onset of dementia [1,3], and there is substantial evidence that reductions in HCV occur at prodromal phases before the development of clinical dementia [2]. It is therefore considered that HCV represents a biomarker that could be used to enrich clinical trials with individuals who are not yet clinically demented but are likely to progress rapidly.

Scientific assessment of the potential for use of biomarkers in clinical trials can be advanced in a structured fashion through the process of biomarker qualification, a process recently introduced by regulatory agencies including the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA). Regulatory qualification of a biomarker for a defined context of use provides scientifically robust assurances to sponsors and regulators that accelerate appropriate adoption of biomarkers into drug development and clinical practice. Such assurances saves time and money by removing the burden of proof on each individual sponsor to provide data to regulatory agencies on biomarker performance and validation.

In the European Union, the EMA, based in London, is the central regulatory agency that reviews new medicinal products. The evaluation is the responsibility of the Committee for Medicinal Products for Human Use (CHMP), which established a Scientific Advice Working Party (SAWP) as one of its supporting Committees to provide scientific advice to applicants.

In additional to the SAWP providing independent expert advice to sponsors seeking marketing authorization, it also runs the qualification of novel methodologies procedure [4], established by the EMA in 2008, which can result in one of two possible outcomes: (*i*) CHMP qualification *advice* based on the evaluation of the scientific rationale and on preliminary data submitted, relevant to the development of future protocols and methods for further method development toward qualification; and (*ii*) CHMP qualification *opinion* based on the assessment of submitted data, relevant to the acceptability of a specific use of the proposed method (e.g., use of a novel methodology or an imaging method) in a research and development context (nonclinical or clinical studies). After publication of a draft qualification opinion, the CHMP evaluation is open to scientific scrutiny and public comment to ensure that adopted opinions are broadly accepted within the community.

The Coalition Against Major Diseases (CAMD) is one of seven precompetitive consortia of the Critical Path Institute created to deliver on the US FDA's Critical Path Initiative [5] to accelerate the development of therapies for AD and Parkinson's disease by generating the best methods and tools for evaluating drug efficacy, expediting clinical trials, and streamlining review by regulatory agencies [6].

In April 2011, CAMD submitted a dossier to the SAWP requesting a qualification opinion on the use of HCV as a biomarker for enrichment in AD trials in the predementia or prodromal phase. SAWP responded with a list of discussion points and questions in May 2011. CAMD submitted a formal written response to several of the questions and then met with SAWP representatives during a face-to-face meeting in June 2011 to respond to the remaining questions. At this meeting, SAWP posed several additional questions and, in August 2011, CAMD submitted a formal written response to these questions. In September 2011, SAWP approved and CHMP adopted the qualification opinion on the use of HCV as a candidate biomarker for AD for release for public consultation. The consultation period ended on November 1, 2011, and the opinion was adopted by CHMP on November 17, 2011 [7].

This article summarizes the content of the data submitted to the EMA, the discussion process between CAMD and the EMA to address outstanding questions and concerns of the EMA, and the resulting qualification opinion.

2. Rationale for seeking qualification of HCV

Advances in the understanding of AD pathophysiology show that the onset of pathology begins decades before the onset of clinical symptoms [8]. Early treatment of AD is thought to offer the best opportunity for effective intervention [9]. For this to be demonstrated, clinical trials must be performed using participants affected during an early phase of the disease process (e.g., predementia). Clinical criteria exist for a prodromal disease stage defined as amnestic mild cognitive impairment (MCI), characterized by objective memory deficits but the absence of frank

dementia [10]. However, clinically defined MCI represents a heterogeneous group, with some people remaining stable for many years, some reverting to clinical normality, some progressing to other types of dementia, and only 10% to 15% per year progressing to AD dementia [11]. Thus, it is challenging to identify people who have a diagnosis of MCI who are most likely to progress to AD dementia based on clinical criteria alone. Because hippocampal atrophy accelerates during the MCI stage of AD, it is thought to represent a "proximity marker," or staging tool, to help identify people with MCI at increased risk of imminent clinical decline [12]. Recent guidelines have been advanced that propose that accurate diagnosis of MCI requires the use of both cognitive tests in combination with biomarkers [13–15]. It has been proposed that a single measurement of HCV from a structural magnetic resonance (MR) image in predemented individuals with episodic memory deficit can be used to select an "enriched" cohort of patients with MCI more likely to progress to AD dementia during the course of an AD clinical trial.

In previous clinical trials, in which enrollment was predicated on a diagnosis of MCI based on cognitive function alone (i.e., without assessing biomarkers), the expected rate of conversion was, in general, not estimated accurately [16,17]. As a result, protocol amendments were necessary to increase the sample size and/or increase duration of the trial—in some cases, up to 4 years—resulting in unacceptably high costs, long trials, and unnecessary exposure to treatment [11]. Enriching trials with participants more likely to undergo rapid clinical deterioration would allow for increased statistical power and smaller sample sizes in MCI trials [18]. More important, MR images are widely available globally for implementation in international trials and, because an MR image is performed invariably at baseline for radiological screening, the addition of a quantitative HCV measurement can be a cost-effective addition to the trial procedures, imposing no additional burden on the patient. Recent advances in automated methods for segmentation and volumetry have catalyzed more cost-effective and efficient implementation of structural imaging in clinical trials (e.g., [19]).

The topography of brain atrophy in AD mirrors that of neurofibrillary pathology [20–22]. Atrophy begins, and is ultimately most severe, in the medial temporal lobe, particularly the entorhinal cortex and hippocampus. Rates of change in several volumetric measures, including whole-brain and hippocampal atrophy, correlate closely with changes in cognitive performance [23]. In a meta-analysis, medial temporal lobe atrophy is estimated to have 73% sensitivity and 81% specificity for predicting whether patients with amnestic MCI will progress to dementia [24].

The Dubois research criteria for the diagnosis of AD [14] specify that, to meet the criteria for probable AD, an affected individual must satisfy core clinical criteria and at least one or more of the supportive biomarker criteria. Volumetric magnetic resonance imaging (vMRI) measures that satisfy this specification include measures that indicate volume loss of hippocampus, entorhinal cortex, or amygdala using qualitative visual scoring or quantitative volumetry of regions of interest. Indeed, at the MCI stage, magnetic resonance imaging (MRI) methods appear to provide more sensitive prediction of clinical progression than cerebrospinal fluid (CSF) or cognitive testing alone [25]. Moreover, rates of change in several volumetric measures correlate closely with cognitive decline [23]. Thus, structural MRI has gained increasing acceptance in clinical settings as a sensitive and powerful marker of neurodegeneration and consequent cognitive decline.

3. Methods

3.1. Literature review

A systematic review of the literature was conducted following established methods. MEDLINE (via PubMed) and EMBASE searches identified studies published in English between January 1, 1995, and March 23, 2011, that enrolled elderly participants with predementia or MCI. Only longitudinal studies with at least 18 months of follow-up data were included, and studies must have included baseline quantitative data on HCV as well as diagnostic measures (sensitivity, specificity, area under the curve [AUC], and so on) or hazard ratio and odds ratio for vMRI volumes in predicting progression from MCI to AD.

Twenty-seven papers met these criteria. Because the heterogeneity in methodology used by various studies precluded direct quantitative comparisons of results, CAMD used a Cochrane vote-counting analysis [26] of the literature relating to the use of HCV as a biomarker of AD. In addition, a sensitivity analysis was performed to address the potential risk of bias arising from heterogeneity.

3.2. De novo analysis

To supplement the published literature, and to address EMA concerns about possible publication and selection bias in the literature, CAMD also performed a de novo analysis on HCV measures from MR images acquired in the Alzheimer's Disease Neuroimaging Initiative (ADNI), a natural history study with data acquisition and quality control closely approximating those used in clinical trials. The primary purpose of this analysis was to assess the hypothesis that HCV could discriminate accurately patients at high risk of short-term progression from MCI to AD. A secondary aim was to compare the predictive performance of different HCV algorithms in this respect.

A Cox regression analysis was performed with age, gender, Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-Cog) score, apolipoprotein E status, and intracranial volume included as covariates in the model. The ADNI

subjects were split randomly at enrollment into 10 subgroups by the CAMD–ADNI statistics core team, with four subgroups designated as a training set and the remaining subgroups as a test set. Prediction performance was calculated with respect to the receiver–operating characteristics (ROC) for predicting conversion to clinical dementia within 2 years. To determine whether the method used to quantify HCV impacts significantly the prediction of conversion from MCI to AD, data from four separate algorithms for measuring HCV were analyzed—namely, FreeSurfer [27,28], NeuroQuant [29,30], Learning Embeddings for Atlas Propagation [31], and Hippocampus Multi-Atlas Propagation and Segmentation [32,33]. All calculations used the bilateral (average of left and right) HCV from each subject. At the time these calculations were performed, HCV measures were available from 319 (FreeSurfer), 322 (NeuroQuant), 322 (Learning Embeddings for Atlas Propagation), and 289 (Hippocampus Multi-Atlas Propagation and Segmentation) MCI subjects, with 2 years of clinical follow-up. These data sets were split in a ratio of approximately 45:55 for training and testing, respectively. ROCs were generated and the AUC was used as a performance metric for comparing the algorithms.

The SAWP requested additional information comparing the performance of CSF and HCV in predicting conversion to AD. CAMD therefore conducted a de novo analysis of other biomarkers used for enrichment of prodromal AD clinical trial populations. In this analysis, the area under the ROC curve for the CSF biomarkers (amyloid-β_{1-42}, phosphorylated tau, total tau, and their linear combination) corresponding to a conversion to clinical dementia within 2 years was compared with that found using HCV measures.

3.3. Clinical trial enrichment scenario: Worked example

To illustrate how enrichment would work in practice, a worked example was produced using the distribution of HCVs in the ADNI-1 healthy control cohort as a normative reference sample and the ADNI amnestic MCI cohort as a mock screening population. This made clear that a prerequisite step is the selection of a volumetric analysis algorithm to be used both to select the cut point and to analyze the subsequent screening images. In the example, FreeSurfer was selected for this purpose. HCVs in the reference sample were corrected for age and intracranial volume using linear regression, yielding a distribution of adjusted HCVs from which a cut point suitable for use as an inclusion criterion in screening could be defined. For the purposes of the example, the adjusted HCV corresponding to the 10th percentile of the reference distribution (approximately 1.3 standard deviations below the mean) was chosen as the cut point. These regression coefficients were then used to correct each individual HCV measure from the MCI cohort (also calculated using FreeSurfer) for its corresponding age and ICV values. Each subject within the MCI "screening" population was then included in the enriched group only if their adjusted HCV was less than the cut point. The fraction of subjects who progressed to clinical dementia, along with the annual change of two commonly used clinical instruments (ADAS-Cog13 and Mini-Mental State Examination), were calculated from the enriched population and compared with values calculated using the entire (unenriched) screening population.

4. Results

4.1. Results of literature review

Of the 27 studies included in this meta-analysis [12,33–50,e1–e9], all (with the exception of two of the studies [40,e8]) supported the context of use despite variations in magnet strength, acquisition protocol for HCV reconstruction, clinical definition of MCI, participants' medical and medication history, and study size. Of the 25 supportive studies, 13 reported Cox proportional hazard ratios (range, 0.21–15.8), six reported sensitivities (range, 50–90.9), and five reported specificities (range, 61.9–90%). Two studies [49,e2] reported that the association between HCV and conversion was no longer significant after adjustment for age, sex, and intracranial volume. The two nonsupportive studies showed no significant difference in baseline HCVs between participants who progressed to AD and those who did not, although one of these studies was considered too small to draw meaningful conclusions [40] and the other reported that some participants may have been in an earlier stage of disease than the MCI category of major interest for the current context of use [e8].

As a result of this literature review, CAMD concluded that there is substantial and consistent evidence to support the use of an HCV measurement taken at a single time point as an appropriate measure of risk of progression to AD for subject inclusion in an MCI clinical trial.

A summary of the results of the literature review is included in Table E1.

4.2. Results of the de novo analysis

The de novo ROC analysis compared the sensitivity vs. specificity curves for predicting clinical conversion from amnestic MCI to AD dementia within 2 years by four different HCV algorithms (Fig. 1, Table 1). Prediction performance was very similar for all four algorithms, with area under the ROC curve values ranging from 0.73 to 0.76. These results suggest that the HCV quantification method (the one heterogeneous variable across the studies analyzed here de novo) does not impact the utility of HCV as an enrichment biomarker and that four major image analysis algorithms being used in the field all support the proposed context of use. The AUC values found in the de novo analysis were also similar to those (range, 0.60–0.77) that were found in the literature review (Table 2). Three of these four articles, [38,42,e10] used manual tracing and the fourth [e9] used automated FreeSurfer software. Taken together, CAMD concluded that these similar results, despite different HCV quantification approaches, supported the context of use.

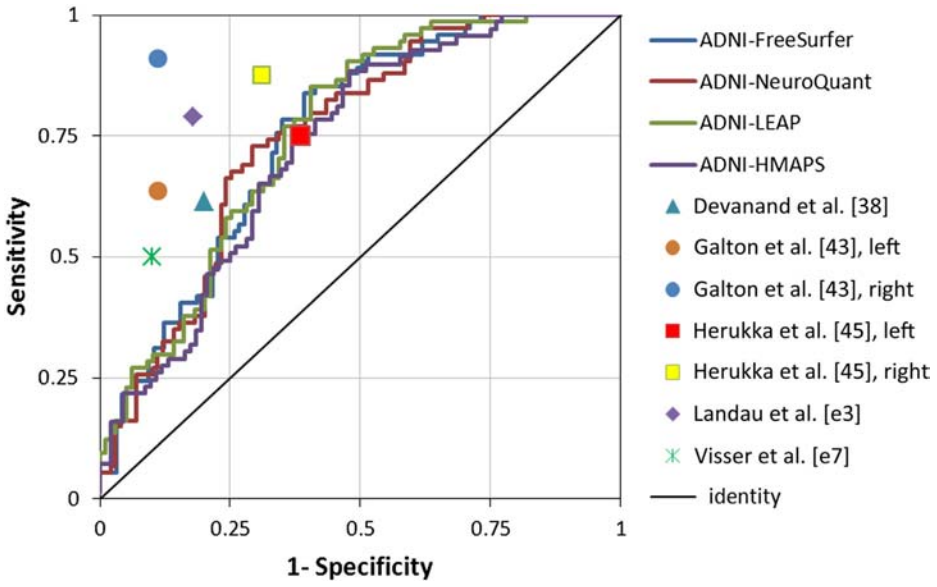

Fig. 1. Receiver–operating characteristic curves for four hippocampal volume HCV algorithms (FreeSurfer, NeuroQuant, LEAP, and HMAPS) applied to Alzheimer's Disease Neuroimaging Initiative (ADNI)-1 data, along with point sensitivity vs. specificity results reported in the Coalition Against Major Diseases literature review [38,43,45,e3,e7]. LEAP, Learning Embeddings for Atlas Propagation; HMAPS, Hippocampus Multi-Atlas Propagation and Segmentation.

Comparison of the AUC values of baseline vMRI of the hippocampus and CSF analytes (amyloid-β_{1-42}, phosphorylated tau, total tau, single and combined) indicate that vMRI in this data set is as good as, or slightly better at, predicting conversion to AD than CSF analytes—specifically, a greater ROC AUC was observed for MRI than CSF biomarkers in this comparison (Fig. 2).

4.3. Clinical trial enrichment scenario: Worked example

The worked example based on data from ADNI-1 illustrated how the biomarker could be used in practice. Using a cut point corresponding to the 10th percentile of the adjusted HCVs in the reference sample and excluding individuals with HCVs above this cut point (45.7% of this screening population), the enriched sample enrolled in the hypothetical trial had a 2-year rate of conversion to AD dementia of 57.2%, a 35% increase over the 2-year conversion rate obtained in the entire MCI sample (i.e., in the absence of MRI-based screening [42.3%]). The average annual cognitive changes in the enriched population were greater than in the nonenriched sample (ADAS-Cog13, 0.106/year vs. 0.091/year; Mini-Mental State Examination, −0.048 vs. −0.026/year).

Fig. 3 provides a general process that is recommended that sponsors follow when applying HCV-based enrichment.

Table 1
Results of Coalition Against Major Diseases' *de novo* analysis. The AUC for four different hippocampal volume quantification algorithms applied to ADNI-1 data indicate the prediction by MRI hippocampal volume of clinical conversion to Alzheimer's dementia within two years.

Algorithm	Training, n	Testing, n	AUC based on clinical conversion
LEAP	149	173	0.7565
NeuroQuant	149	173	0.7516
FreeSurfer	148	171	0.7536
HMAPS	128	161	0.7290

Abbreviations: AUC, area under the receiver–operating characteristic curves; LEAP, Learning Embeddings for Atlas Propagation; HMAPS, Hippocampus Multi-Atlas Propagation and Segmentation.

5. The CHMP qualification opinion

The final qualification opinion issued by the CHMP is as follows:

"Low hippocampal volume, as measured by MRI and considered as a dichotomized variable (low volume or not), appears to help enriching recruitment into clinical trials aimed at studying drugs potentially slowing the progress/conversion to AD dementia of the included subjects. Low hippocampal volume might be considered a marker of

Table 2
AUC values reported in the Coalition Against Major Diseases literature review

Study	n	AUC based on clinical conversion
Bakkour et al. [e9]	49	0.65
Devanand et al. [38]	139	0.77
Fleisher et al. [e10]	129	0.60
Galluzzi et al. [42]	90	0.73

Abbreviation: AUC, area under the receiver–operating characteristic curves.

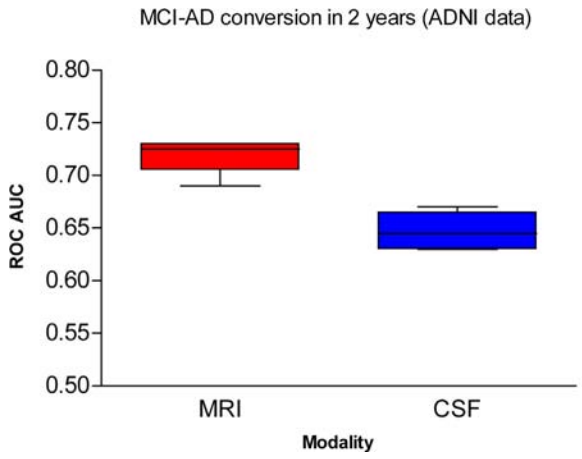

Fig. 2. Comparison of receiver–operating characteristics (ROC) area-under-the-curve (AUC) values obtained from Alzheimer's Disease Neuroimaging Initiative-1 data for the prediction of conversion from mild cognitive impairment to clinical Alzheimer's disease dementia using hippocampal volume (data shown across four algorithms are included in Fig. 1 and Table 1; range, 0.69–0.73) and cerebrospinal fluid (CSF) analytes (data shown across four analyte combinations: amyloid-β_{1-42} alone, phosphorylated tau alone, total tau alone, and the linear combination of all three; range, 0.63–0.67). MRI, magnetic resonance imaging.

progression to dementia in subjects with cognitive deficit compatible with predementia stage of AD (Dubois 2007). However, neither the actual value of low hippocampal volume to accurately predict rate of such progression to dementia in the referred subjects nor the relative value of other biomarkers have been reported.

"As currently planned in the current opinion subjects might be included in the studies based on clinical criteria and low hippocampal volume biomarker (if positive). The CHMP has given a previous positive opinion in the predementia stage of Alzheimer's disease: cerebrospinal fluid-related biomarkers for drugs affecting amyloid burden. This may lead first to a heterogeneous population and, moreover, it will not be possible to explore the relationship among them. Although not required form a regulatory perspective, the concomitant assessment of the two biomarkers in predementia stage of AD would be of great value.

"The process of measurement of low hippocampal volume is also complex. To obtain reliable results implies the standardization of all steps (at least imaging acquisition protocol, imaging reconstruction/analysis methods, timing to conversion, etc). International guidelines have been produced. These guidelines must be enforced."

6. Discussion

The vast majority of the published study results support individually the position that the presence of hippocampal atrophy identified in participants diagnosed with MCI who progressed to AD dementia sooner and more reliably than those with larger HCVs. The de novo analysis also supported CAMD's proposal that smaller HCVs are associated with more rapid clinical decline.

In implementing HCV as a biomarker as part of trial eligibility, participants with episodic memory deficits (the core diagnostic criteria of Dubois [14]) would receive an MRI HCV measurement as part of the trial screening process. This enrichment is expected to result in a population with a steeper and more homogeneous clinical trajectory, enabling a trial sponsor to run more efficient clinical trials in amnestic MCI populations with reduced subject numbers and increased power. However, any such enrichment strategy will result in an increased number of screen

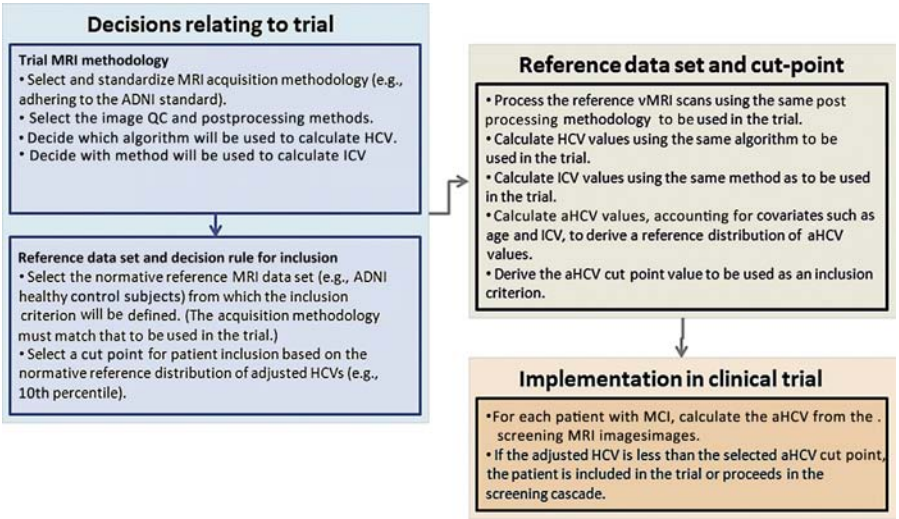

Fig. 3. Flow diagram indicating an operational "algorithm" for the use of hippocampal volume (HCV) in a clinical trial. MRI, magnetic resonance imaging; ADNI, Alzheimer's Disease Neuroimaging Initiative; QC, quality control; HCV, hippocampal volume; ICV, intracranial volume, vMRI, volumetric MRI; aHCV, adjusted HCV; MCI, mild cognitive impairment.

failures, an important practical consideration in the context of clinical trials [e11], and this should be modeled by the sponsor in estimating the overall time and cost of performing a trial.

CAMD acknowledged that HCV is not specific for AD. Substantial literature supports the evidence that numerous factors and comorbidities can result in reduced HCV [e12–e16]. Yet, the goal of implementing HCV as a prognostic biomarker is to augment the current inclusion criteria by enrolling a subset of the participants corresponding to those at greater risk of imminent clinical decline. The greater ROC AUC calculated for MRI HCV compared with CSF biomarkers confirms the practical value of HCV. Clinical exclusion criteria should be used to minimize the number of subjects enrolled who might have low HCV resulting from other conditions.

HCV-based enrichment is not dependent on the mechanism of the investigational compound. As a result, defining the use of the biomarker in the context of any single compound or mechanism (e.g., the amyloid hypothesis) would restrict unnecessarily its application, utility, and generalizability for sponsors developing therapeutics for AD. Furthermore, the data collected and presented here are independent of any investigational intervention.

The EMA recommended that standardization and international harmonized guidelines be followed when implementing the biomarker in clinical trials. Best practice includes the use of acquisition and quality control methods consistent with those used in ADNI [e17], and the use of centralized analysis using a single, validated measurement procedure. The gold standard for hippocampal volumetry is manual segmentation by a human expert, and an internationally harmonized protocol is currently under development that will be finalized by late 2013 [e18], and will also provide reference data for retraining automatic algorithms according to these harmonized guidelines.

Hippocampal volume measures might be used alone for enrichment, or in combination with other biomarkers such as CSF or amyloid positron emission tomographic (PET) imaging. Collection of CSF by lumbar puncture is considered to be more invasive and less widely available than MRI; and PET imaging is more expensive and less widely available than MRI. Thus, HCV represents a potentially more feasible tool for characterizing and selecting participants for global clinical trials. Since the qualification opinion was published, further work on comparisons of HCV and CSF and combining the validations has been performed [e19–e22]. It has been reported that a combination of CSF and HCV shows increased predictive value in defining people that have MCI resulting from AD compared with each biomarker alone [e3,e16,e20,e22]. The EMA recommended that multiple biomarkers be assessed concurrently in future MCI research. The practical implications of biomarker-based enrichment (either alone or in combination), including the impact of increased screen failure rates vs. the need to enroll fewer subjects, are beginning to be elucidated [e11,e21,e22], but remain to be fully tested in clinical trial settings.

Notably, the qualification of HCV as a biomarker for trial enrichment in AD clinical trials represents the first clinical imaging biomarker qualified by regulators for implementation in clinical trials. Subsequently, the EMA has also qualified amyloid PET neuroimaging for trial enrichment in predementia AD clinical trials [e23], and an amyloid imaging PET tracer has also been approved by both the FDA and the EMA as a medical device to enable detection of the presence of amyloid. Regulatory acceptance and recommendations for implementation of these biomarkers in clinical trials will compress the timelines and increase the chances of success in MCI clinical trials.

Acknowledgments

The Critical Path Institute's Coalition Against Major Diseases is supported by grant no. U01FD003865 from the US Food and Drug Association and by Science Foundation Arizona under grant no. SRG 0335-08.

RESEARCH IN CONTEXT

1. Systematic review: During the process of development of the (CAMD) biomarker qualification initiative, we performed a comprehensive literature review encompassing January 1995 through March 2011.
2. Interpretation: The findings and evaluation of the results from the literature and de novo analyses carried out were designed to support the proposed application of the biomarker in clinical trials of drug candidates targeting predementia stages of Alzheimer's disease.
3. Future directions: The impact of biomarker qualification is to gain efficiency in drug development by implementing the HV biomarker as a way to enrich clinical trials at the early stages of the disease. The focus was on automated image analysis algorithms designed to measure hippocampal volume. Advances in newer algorithms will lead to further refinement of this specific qualification. Furthermore, future applications and recommendations will include combinations of biomarkers to refine further the patient populations to greater homogeneity.

References

[1] Jack CR Jr., Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol 2010;9:119–28.

[2] Jack CR Jr. Alliance for aging research AD biomarkers work group: structural MRI. Neurobiol Aging 2011;32:S48–57.
[3] Jedynak BM, Lang A, Liu B, Katz E, Zhang Y, Wyman BT, et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's Disease Neuroimaging Initiative cohort. Neuroimage 2012;63:1478–86.
[4] European Medicines Agency. Qualification of novel methodologies for drug development: guidance to applicants. Accessed November, 21, 2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004201.pdf. 2012.
[5] Food and Drug Administration. Innovation or stagnation: challenges and opportunity on the critical path to new medical products. Accessed February 26, 2013. Available from: http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm. 2004.
[6] Romero K, Corrigan B, Neville J, Kopko S, Cantillon M. Striving for an integrated drug development process for neurodegeneration: the Coalition Against Major Diseases. Neurodegen Dis Manage 2011;1:1–7.
[7] European Medicines Agency. Qualification opinion of low hippocampal volume (atrophy) by MRI for use in regulatory clinical trials: in pre-dementia stage of Alzheimer's disease. Accessed November, 21, 2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2011/10/WC500116264.pdf. 2011.
[8] Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement 2012;8:S1–68.
[9] Sperling RA, Jack CR Jr., Aisen PS. Testing the right target and right drug at the right stage. Sci Transl Med 2011;3:111–33.
[10] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999;56:303–8.
[11] Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. Arch Neurol 2009;66:1447–55.
[12] Wang PN, Liu HC, Lirng JF, Lin KN, Wu ZA. Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. Psychiatry Res 2009;171:221–31.
[13] Albert MS, Dekosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. 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.
[14] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007;6:734–46.
[15] Visser PJ, Vos S, van Rossum I, Scheltens P. Comparison of International Working Group criteria and National Institute on Aging–Alzheimer's Association criteria for Alzheimer's disease. Alzheimers Dement 2012;8:560–3.
[16] Jelic V, Kivipelto M, Winblad B. Clinical trials in mild cognitive impairment: lessons for the future. J Neurol Neurosurg Psychiatry 2006;77:429–38.
[17] Petersen RC. Mild cognitive impairment clinical trials. Nat Rev Drug Discov 2003;2:646–53.
[18] Schott JM, Bartlett JW, Barnes J, Leung KK, Ourselin S, Fox NC. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. Neurobiol Aging 2010;31:1452–62.
[19] Nestor SM, Gibson E, Gao FQ, Kiss A, Black SE. A direct morphometric comparison of five labeling protocols for multi-atlas driven automatic segmentation of the hippocampus in Alzheimer's disease. Neuroimage 2012;66C:50–70.
[20] Whitwell JL, Jack CR Jr., Parisi JE, Knopman DS, Boeve BF, Petersen RC, et al. Rates of cerebral atrophy differ in different degenerative pathologies. Brain 2007;130:1148–58.
[21] Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. Acta Neuropathol 1991;82:239–59.
[22] Whitwell JL, Josephs KA, Murray ME, Kantarci K, Przybelski SA, Weigand SD, et al. MRI correlates of neurofibrillary tangle pathology at autopsy: a voxel-based morphometry study. Neurology 2008;71:743–9.
[23] Frisoni GB, Fox NC, Jack CR Jr., Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 2010;6:67–77.
[24] Schmand B, Huizenga HM, van Gool WA. Meta-analysis of CSF and MRI biomarkers for detecting preclinical Alzheimer's disease. Psychol Med 40:135–45.
[25] Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology 2009;73:294–301.
[26] Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0, updated March 2011. Available from: http://www.cochrane-handbook.org edition.
[27] Freesurfer. Available from: http://surfer.nmr.mgh.harvard.edu/.
[28] Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 1999;9:179–94.
[29] NeuroQuant. Available from: http://www.cortechs.net/products/neuroquant.php.
[30] Brewer JB. Fully-automated volumetric MRI with normative ranges: translation to clinical practice. Behav Neurol 2009;21:21–8.
[31] Wolz R, Aljabar P, Hajnal JV, Hammers A, Rueckert D. LEAP: Learning Embeddings for Atlas Propagation. Neuroimage 2010;49:1316–25.
[32] Barnes J, Foster J, Boyes RG, Pepple T, Moore EK, Schott JM, et al. A comparison of methods for the automated calculation of volumes and atrophy rates in the hippocampus. Neuroimage 2008;40:1655–71.
[33] Leung KK, Barnes J, Ridgway GR, Bartlett JW, Clarkson MJ, Macdonald K, et al. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. Neuroimage 2010;51:1345–59.
[34] Auriacombe S, Helmer C, Amieva H, Berr C, Dubois B, Dartigues JF. Validity of the free and cued selective reminding test in predicting dementia: the 3C study. Neurology 2010;74:1760–7.
[35] Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. Neurobiol Aging 2000;21:19–26.
[36] Desikan RS, Cabral HJ, Fischl B, Guttmann CR, Blacker D, Hyman BT, et al. Temporoparietal MR imaging measures of atrophy in subjects with mild cognitive impairment that predict subsequent diagnosis of Alzheimer disease. AJNR Am J Neuroradiol 2009;30:532–8.
[37] Desikan RS, Fischl B, Cabral HJ, Kemper TL, Guttmann CR, Blacker D, et al. MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. Neurology 2008;71:819–25.
[38] Devanand DP, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. Neurology 2007;68:828–36.
[39] Eckerstrom C, Olsson E, Borga M, Ekholm S, Ribbelin S, Rolstad S, et al. Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: the Goteborg MCI study. J Neurol Sci 2008;272:48–59.
[40] Fellgiebel A, Dellani PR, Greverus D, Scheurich A, Stoeter P, Muller MJ. Predicting conversion to dementia in mild cognitive impairment by volumetric and diffusivity measurements of the hippocampus. Psychiatry Res 2006;146:283–7.
[41] Fleisher AS, Sun S, Taylor C, Ward CP, Gamst AC, Petersen RC, et al. Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment. Neurology 2008;70:191–9.
[42] Galluzzi S, Geroldi C, Ghidoni R, Paghera B, Amicucci G, Bonetti M, et al. The new Alzheimer's criteria in a naturalistic series of patients with mild cognitive impairment. J Neurol 2010;257:2004–14.

[43] Galton CJ, Erzinclioglu S, Sahakian BJ, Antoun N, Hodges JR. A comparison of the Addenbrooke's Cognitive Examination (ACE), conventional neuropsychological assessment, and simple MRI-based medial temporal lobe evaluation in the early diagnosis of Alzheimer's disease. Cogn Behav Neurol 2005;18:144–50.
[44] Henneman WJ, Sluimer JD, Barnes J, van der Flier WM, Sluimer IC, Fox NC, et al. Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. Neurology 2009;72:999–1007.
[45] Herukka SK, Pennanen C, Soininen H, Pirttila T. CSF Abeta42, tau and phosphorylated tau correlate with medial temporal lobe atrophy. J Alzheimers Dis 2008;14:51–7.
[46] Jack CR Jr., Petersen RC, Grundman M, Jin S, Gamst A, Ward CP, et al. Longitudinal MRI findings from the vitamin E and donepezil treatment study for MCI. Neurobiol Aging 2008;29:1285–95.
[47] Jack CR Jr., Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 2000;55:484–9.
[48] Jack CR Jr., Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology 1999;52:1397–403.
[49] Jack CR Jr., Shiung MM, Weigand SD, O'Brien PC, Gunter JL, Boeve BF, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 2005;65:1227–31.
[50] Jack CR Jr., Wiste HJ, Vemuri P, Weigand SD, Senjem ML, Zeng G, et al. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. Brain 2010;133:3336–48.

Further readings

[e1] Kantarci K, Petersen RC, Boeve BF, Knopman DS, Weigand SD, O'Brien PC, et al. DWI predicts future progression to Alzheimer disease in amnestic mild cognitive impairment. Neurology 2005; 64:902–4.

[e2] Killiany RJ, Hyman BT, Gomez-Isla T, Moss MB, Kikinis R, Jolesz F, et al. MRI measures of entorhinal cortex vs hippocampus in preclinical AD. Neurology 2002;58:1188–96.

[e3] Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, et al. Comparing predictors of conversion and decline in mild cognitive impairment. Neurology 2010;75:230–8.

[e4] Stoub TR, Rogalski EJ, Leurgans S, Bennett DA, deToledo-Morrell L. Rate of entorhinal and hippocampal atrophy in incipient and mild AD: relation to memory function. Neurobiol Aging 2010; 31:1089–98.

[e5] Tapiola T, Pennanen C, Tapiola M, Tervo S, Kivipelto M, Hanninen T, et al. MRI of hippocampus and entorhinal cortex in mild cognitive impairment: a follow-up study. Neurobiol Aging 2008;29:31–8.

[e6] Visser PJ, Scheltens P, Verhey FR, Schmand B, Launer LJ, Jolles J, et al. Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. J Neurol 1999;246:477–85.

[e7] Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J Neurol Neurosurg Psychiatry 2002; 72:491–7.

[e8] Whitwell JL, Shiung MM, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, et al. MRI patterns of atrophy associated with progression to AD in amnestic mild cognitive impairment. Neurology 2008;70:512–20.

[e9] Bakkour A, Morris JC, Dickerson BC. The cortical signature of prodromal AD: regional thinning predicts mild AD dementia. Neurology 2009;72:1048–55.

[e10] Fleisher AS, Raman R, Siemers ER, Becerra L, Clark CM, Dean RA, et al. Phase 2 safety trial targeting amyloid beta production with a gamma-secretase inhibitor in Alzheimer disease. Arch Neurol 2008;65:1031–8.

[e11] Lorenzi M, Donohue M, Paternico D, Scarpazza C, Ostrowitzki S, Blin O, et al. Enrichment through biomarkers in clinical trials of Alzheimer's drugs in patients with mild cognitive impairment. Neurobiol Aging 2010;31:1443–51.

[e12] Debette S, Seshadri S, Beiser A, Au R, Himali JJ, Palumbo C, et al. Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. Neurology 2011;77:461–8.

[e13] den Heijer T, Launer LJ, Prins ND, van Dijk EJ, Vermeer SE, Hofman A, et al. Association between blood pressure, white matter lesions, and atrophy of the medial temporal lobe. Neurology 2005; 64:263–7.

[e14] Korf ES, White LR, Scheltens P, Launer LJ. Midlife blood pressure and the risk of hippocampal atrophy: the Honolulu Asia Aging Study. Hypertension 2004;44:29–34.

[e15] Strassburger TL, Lee HC, Daly EM, Szczepanik J, Krasuski JS, Mentis MJ, et al. Interactive effects of age and hypertension on volumes of brain structures. Stroke 1997;28:1410–7.

[e16] Walhovd KB, Fjell AM, Brewer J, McEvoy LK, Fennema-Notestine C, Hagler DJ Jr., et al. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. AJNR Am J Neuroradiol 2010; 31:347–54.

[e17] Jack CR Jr., Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008;27:685–91.

[e18] Boccardi M, Ganzola R, Bocchetta M, Pievani M, Redolfi A, Bartzokis G, et al. Survey of protocols for the manual segmentation of the hippocampus: preparatory steps toward a joint EADC–ADNI harmonized protocol. J Alzheimers Dis 2011;26:61–75.

[e19] Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's Disease Neuroimaging Initiative. Arch Gen Psychiatry 2011;68:961–9.

[e20] Stricker NH, Dodge HH, Dowling NM, Han SD, Erosheva EA, Jagust WJ. CSF biomarker associations with change in hippocampal volume and precuneus thickness: implications for the Alzheimer's pathological cascade. Brain Imaging Behav 2012.

[e21] Vos S, van Rossum I, Burns L, Knol D, Scheltens P, Soininen H, et al. Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI. Neurobiol Aging 2012;33:2272–81.

[e22] Yu P, Dean RA, Hall SD, Qi Y, Sethuraman G, Willis BA, et al. Enriching amnestic mild cognitive impairment populations for clinical trials: optimal combination of biomarkers to predict conversion to dementia. J Alzheimers Dis 2012;32:373–85.

[e23] European Medicines Agency. Qualification opinion of Alzheimer's disease novel methodologies/biomarkers for PET amyloid imaging (positive/negative) as a biomarker for enrichment, for use in regulatory clinical trials in predementia Alzheimer's disease. Accessed December 28, 2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2012/04/WC500125018.pdf. 2012.

Table E1
Longitudinal studies of hippocampal atrophy and progression to Alzheimer's disease

Study	Follow-up, mo (range)	Type of subjects	Sample size, n	Converting to AD, n	Stable MCI, n	Comparison	HR, OR (95% CI)	*P* value	Sensitivity, % (95% CI)	Specificity, % (95% CI)	AUC (95% CI)	Summary
Bakkour et al. [e9]	32.4	QAD (CDR = 0.5)	49	20	29	QAD-AD vs. QAD			83	50	0.65	Entorhinal volume was a better predictor than HCV and other brain regions for predicting progression to AD.
Convit et al. [35]	38.4	Normal or MCI	46	14	32							Baseline measures of HC showed declining subjects had 11.3% of reduction in the HC compared with nondecliners.
Desikan et al. [36]	60	MCI	129	44	85	MCI-S vs. MCI-AD	Crude HR, 0.64 (0.47–0.87); adjusted HR, 0.73 (0.51–1.04)	.005, .08				Automated MRI measurement of HCV is a statistically significant predictor of progression from MCI to AD.
Desikan et al. [37]	108	MCI	47	25	22							HC and other brain regions had greater rates of atrophy for converters vs. nonconverters using post hoc comparisons.
Devanand et al. [38]	36	MCI	139	31	102	MCI-AD vs. MCI-S	3.62 (1.93–6.80), 2.89 (1.52–5.51)	<.0001, .019	61.3	80 (fixed)	0.77	In the total subject sample, Cox proportional hazards models, age stratified, with intracranial volume as a covariate, showed smaller volume of the HC (risk ratio per 1 mL volume reduction, 3.62; 95% CI, 1.93–6.80, $P < .0001$) was associated with the hazard of conversion to AD.
Eckerstrom et al. [39]	24	MCI	42	13	21							Total HC, right HC, and left HC volumes were significantly smaller in AD converters vs. nonconverters.
Fellgiebel et al. [40]	18	aMCI	13	6	7							Baseline HC volumes were not significantly different between stable patients with MCI and converters to AD. There was no significant association between reduced HC size and conversion rate in patients with MCI.
Fleisher et al. e10	36	aMCI	129	53	76	MCI-S vs. MCI-AD					60.4% (52–68.8)	HC and ventricular volumes were associated with progression from aMCI to AD.
Galluzzi et al. [42]	24.0 ± 13.9 (SD)	MCI	90	24	51	MCI-NC vs. MCI-AD					AUC, 0.73; 95% CI, 0.57–0.89	MCI-AD showed higher prevalence of MT atrophy relative to MCI-NC (55% vs. 18%, $P = .002$ on post hoc comparison). All the patients with MCI with MT atrophy converted to AD, but only 48% of those without MT atrophy (log-rank test, $P = .0007$). The accuracy of MT atrophy in discriminating MCI-AD from MCI-NC.
Galton et al. [2005]	24	Questionable dementia (memory complaints yielded a CDR = 0.5)	31	11	18	Nonconverters vs. AD converters			Left HC, 63.6; right HC, 90.9	Left HC, 88.9; right HC, 88.9		AD converters had greater HC atrophy compared with nonconverters. Both right and left HC scores were significant predictors of outcome between converter and nonconverter groups.
Henneman et al. [44]	21.6	MCI	44	23	16	MCI-AD vs. MCI-S	HC volume, 7.4 (2.4–23.0); HC atrophy rate, 3.9 (1.6–9.9)	<.05				Regional measures of hippocampal atrophy are the strongest predictors of progression to AD using Cox proportional hazards models.
Herukka et al. [45]	40.6–57.2	MCI	21	8	13	MCI-AD vs. MCI-S	Right HC, 15.8 (1.4–174.2)		Left HC, 75; right HC, 87.5	Left HC, 61.5; right HC, 69		Subjects with MCI progressing to AD had lower volumes in all MRI measures compared with subjects with stable MCI.
Jack et al. [50]	19	MCI	218	89	129	MCI-S vs. MCI-AD	HR, 2.6 (1.8–3.8); 25% vs. 75%	.001				HC atrophy indicates how far along the neurodegenerative path one is, and hence how close to progressing to dementia.
Jack et al. [46]	36		131	52	79							Annual percentage changes were greater in subjects who converted to AD than nonconverters for each of four brain atrophy rate measures.
Jack et al [49]	22.8	aMCI	72	39	33	MCI-AD vs. MCI-S	HC volume, 1.51 (1.1–2.0); HC APC, 1.13 (0.8–1.5)	.002				Adjusted baseline HC volume and whole-brain volumes were significantly associated with conversion from MCI to AD.
Jack et al. [47]	34.8 (24–48)	MCI	43	18	25							Baseline HC volumes of MCI decliners were significantly more atrophic than in those who were stable.
Jack et al. [48]	32.6	MCI	80	27	53	MCI-S vs. MCI-AD	0.69	.015				In univariate analyses, HCV was a statistically significant predictor of the risk of progression to AD.
Kantarci et al. [e1]	36.4	aMCI	21	12	9	MCI-AD vs. MCI-S	2.5 (1.0–6.2)	.02				Higher HC ADC values ($P = .002$) and lower HC W scores (greater atrophy) ($P = .02$) in subjects with aMCI at baseline are associated with a higher relative risk of progression to AD.
Killiany et al. [e2]	36	QAD (CDR = 0.5)	94	21	73	CDR 0.5-AD vs. CDR 0.5	1.5 (1.0–2.31)	<.05	NS	NS		Subjects showing greatest HC atrophy were 1.5 times more likely to develop AD on follow-up for each quartile of decreasing volume (OR, 1.5; CI, 1.0–2.31; $\chi^2 = 3.79$; $P < .05$).

(*Continued*)

Table E1
Longitudinal studies of hippocampal atrophy and progression to Alzheimer's disease (*Continued*)

Study	Follow-up, mo (range)	Type of subjects	Sample size, n	Converting to AD, n	Stable MCI, n	Comparison	HR, OR (95% CI)	*P* value	Sensitivity, % (95% CI)	Specificity, % (95% CI)	AUC (95% CI)	Summary
Landau et al. [e3]	22.8	MCI	85	28	57	MCI-AD vs. MCI-S	2.49 (1.02–5.96)	.04	79	82		Subjects categorized as AD on FDG-PET, HCV, phosphorylated tau 181p, phosphorylated tau 181p/A 1-42, AVLT, and, marginally, *APOE* had a higher risk of converting than non-AD subjects on each measure.
Leung et al. [33]	22.8	MCI	335	123	204							Comparison of MCI subgroups (reverters, stable, and converters) showed HCVs were lower and rates higher in converters compared with stable and reverter groups.
Stoub et al. [e4]	60	aMCI	29	11	18							Subjects with aMCI converting to AD had smaller entorhinal and HCVs at baseline compared with stable control subjects.
Tapiola et al. [e5]	34 (10–54)	MCI	60	9	47	MCI-S vs. MCI-AD	Left HC, 0.739 (0.55–1.00); right HC, 0.668 (0.49–0.91); total HC, 0.815 (0.69–0.97)	.05, <.01, <.05				In all MCI subjects, Cox regression analysis showed baseline volumes of the right HC predicted the progression of MCI to dementia during the follow-up. When subjects with non-AD dementia were excluded, the right, left, and total HCVs significantly predicted the progression to AD.
Visser et al. [e6]	36	MCI	13	9	4	MCI-S vs. MCI-AD	HC OR, 0.21 (0.05–0.99)	.02				Memory dysfunction is a better predictor of AD than the volumes of the HC or the parahippocampal gyrus or the medial temporal lobe atrophy score.
Visser et al. [e7]	22.8 (12–36)	MCI	30	7	23				50	90		HCV at baseline was statistically significantly associated with a diagnosis of AD at follow-up. Trend analyses and logistic regression showed that HCV was a better predictor of outcome than the volume of the parahippocampal gyrus.
Wang et al. [12]	21.9 (10.7—32.8)	aMCI	58	19	39	MCI-AD vs. MCI-S	Left HC volume HR, 0.38 (0.10–0.88)	.03	76.2			HC volume predicted MCI progression to AD.
Whitwell et al. [e8]	44	aMCI	63	42	21							HC volume showed no significant differences between the aMCI–stable group and the a MCI–progressors group.

Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment; HR, hazard ratio; OR, odds ratio; CI, confidence interval; AUC, area under the curve; QAD, questionable AD; CDR, clinical dementia rating scale; HCV, hippocampal volume; HC, hippocampus; MCI-S, MCI stable; aMCI, amnestic MCI; SD, standard deviation; MCI-NC, MCI–no conversion; MT, medial temporal lobe; FDG-PET, fluorodeoxyglucose–positron emission tomography; *APOE*, apolipoprotein E.