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


IXICO Ltd., London, UK
Eli Lilly and Company, Indianapolis, IN, USA
European Medicines Agency, London, UK
Alzheimer’s Association, Chicago, IL, USA
The University of Texas School of Public Health, Houston, TX, USA
University of California, Davis, CA, USA
IRCCS Fatebenefratelli, Brescia, Italy
University of California, San Diego, CA, USA
Coalition Against Major Diseases, Critical Path Institute, Tucson, AZ, USA
Wellnessmanagements.com, Livingston, NJ, USA
UCL Institute of Neurology, London, UK
Mayo Clinic, Rochester, MN, USA
Bristol Myers Squibb, Wallingford, CT, USA
Janssen Pharmaceutical Research and Development, Titusville, NJ, USA
Novartis, Basel, Switzerland
Novartis, NJ, USA
AstraZeneca, Wilmington, DE, USA
Independent science writer, Elverson, PA, USA
CHDI Foundation Inc., Princeton, NJ, USA
ICON Medical Imaging, Warrington, PA, USA
SYNARC Inc., Newark, CA, USA
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.
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 metanalysis, 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
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-β₁₋₄₂, 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. The 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.
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 biomarkers—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.

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 1

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Training, n</th>
<th>Testing, n</th>
<th>AUC based on clinical conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEAP</td>
<td>149</td>
<td>173</td>
<td>0.7565</td>
</tr>
<tr>
<td>NeuroQuant</td>
<td>149</td>
<td>173</td>
<td>0.7516</td>
</tr>
<tr>
<td>FreeSurfer</td>
<td>148</td>
<td>171</td>
<td>0.7536</td>
</tr>
<tr>
<td>HMAPS</td>
<td>128</td>
<td>161</td>
<td>0.7290</td>
</tr>
</tbody>
</table>

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


Further readings


[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.


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

<table>
<thead>
<tr>
<th>Study</th>
<th>Follow-up, mo (range)</th>
<th>Type of subjects</th>
<th>Sample size, n</th>
<th>Converting to AD, n</th>
<th>Stable MCI, n</th>
<th>Comparison</th>
<th>HR, OR (95% CI)</th>
<th>P value</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>AUC (95% CI)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bakkour et al. [e9]</td>
<td>32.4</td>
<td>QAD (CDR 0.5)</td>
<td>49</td>
<td>20</td>
<td>29</td>
<td>QAD-AD vs. QAD</td>
<td>85</td>
<td>50</td>
<td>0.65</td>
<td>Entorhinal volume was a better predictor than HCV and other brain regions for predicting progression to AD. Baseline measures of HCV showed declining subjects had 11.3% of reduction in the HCV compared with nondecliners.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convit et al. [35]</td>
<td>38.4</td>
<td>Normal or MCI</td>
<td>46</td>
<td>14</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Automated MRI measurement of HCV is a statistically significant predictor of progression from MCI to AD.</td>
<td></td>
</tr>
<tr>
<td>Desikan et al. [36]</td>
<td>60</td>
<td>MCI</td>
<td>129</td>
<td>44</td>
<td>85</td>
<td>MCI-S vs. MCI-AD</td>
<td>Crude HR, 0.64 (0.47–0.87); adjusted HR, 0.73 (0.51–1.04)</td>
<td>.005, .08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desikan et al. [37]</td>
<td>108</td>
<td>MCI</td>
<td>47</td>
<td>25</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Devanand et al. [38]</td>
<td>36</td>
<td>MCI</td>
<td>139</td>
<td>31</td>
<td>102</td>
<td>MCI-AD vs. MCI-S</td>
<td>3.62 (1.93–6.80); 2.69 (1.52–5.51)</td>
<td>&lt;.0001, .019</td>
<td>61.3</td>
<td>80 (fixed)</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Eckerstrom et al. [39]</td>
<td>24</td>
<td>MCI</td>
<td>42</td>
<td>13</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleisher et al. [40]</td>
<td>18</td>
<td>aMCI</td>
<td>13</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleisher et al. [41]</td>
<td>36</td>
<td>aMCI</td>
<td>129</td>
<td>53</td>
<td>76</td>
<td>MCI-S vs. MCI-AD</td>
<td>0.73 (0.48–1.12); 0.72 (0.47–1.09)</td>
<td>.10, .06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galluzzi et al. [42]</td>
<td>24.0 ± 13.9 (SD)</td>
<td>MCI</td>
<td>90</td>
<td>24</td>
<td>51</td>
<td>MCI-NC vs. MCI-AD</td>
<td>3.91 (1.6–9.9); HC atrophy rate, 3.9 (1.6–9.9)</td>
<td>0.0007, 0.0007</td>
<td>50.6%</td>
<td>57.2%</td>
<td>0.73; 95% CI, 0.37–0.97</td>
<td></td>
</tr>
<tr>
<td>Galton et al. [2005]</td>
<td>24</td>
<td>Questionable dementia (memory complaints yielded a CDR 0.5)</td>
<td>31</td>
<td>11</td>
<td>18</td>
<td>Nonconverters vs. AD converters</td>
<td>Left HC, 63.6; right HC, 90.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henneman et al. [44]</td>
<td>21.6</td>
<td>MCI</td>
<td>44</td>
<td>23</td>
<td>16</td>
<td>MCI-AD vs. MCI-S</td>
<td>HC volume, 7.4 (2.4–23.0); HC atrophy rate, 3.9 (1.6–9.9)</td>
<td>&lt;.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herukka et al. [45]</td>
<td>40.6–57.2</td>
<td>MCI</td>
<td>21</td>
<td>8</td>
<td>13</td>
<td>MCI-AD vs. MCI-S</td>
<td>Right HC, 15.8 (1.4–174.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jack et al. [50]</td>
<td>19</td>
<td>MCI</td>
<td>218</td>
<td>89</td>
<td>129</td>
<td>MCI-S vs. MCI-AD</td>
<td>HR, 2.6 (1.8–3.8); 25% vs. 75%</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jack et al. [46]</td>
<td>36</td>
<td>MCI</td>
<td>131</td>
<td>52</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jack et al. [49]</td>
<td>22.8</td>
<td>aMCI</td>
<td>72</td>
<td>39</td>
<td>33</td>
<td>MCI-AD vs. MCI-S</td>
<td>HC volume, 1.51 (1.1–2.0); HC APC, 1.13 (0.8–1.5)</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jack et al. [47]</td>
<td>34.8 (24–48)</td>
<td>MCI</td>
<td>43</td>
<td>18</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jack et al. [48]</td>
<td>32.6</td>
<td>MCI</td>
<td>80</td>
<td>27</td>
<td>53</td>
<td>MCI-S vs. MCI-AD</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kantarcı et al. [61]</td>
<td>36.4</td>
<td>aMCI</td>
<td>21</td>
<td>12</td>
<td>9</td>
<td>MCI-AD vs. MCI-S</td>
<td>2.5 (1.0–6.2)</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiliaany et al. [62]</td>
<td>36</td>
<td>QAD (CDR 0.5)</td>
<td>94</td>
<td>21</td>
<td>73</td>
<td>CDR 0.5-AD vs. CDR 0.5</td>
<td>1.5 (1.0–2.31)</td>
<td>&lt;.05</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Study</th>
<th>Follow-up, mo (range)</th>
<th>Type of subjects</th>
<th>Sample size, n</th>
<th>Converting to Stable</th>
<th>Boyd, OR (95% CI)</th>
<th>P value</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>AUC (95% CI)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landau et al. [e3]</td>
<td>22.8</td>
<td>MCI</td>
<td>85</td>
<td>28</td>
<td>57</td>
<td>2.49 (1.02–5.96)</td>
<td>.04</td>
<td>79</td>
<td>82</td>
<td>.02</td>
</tr>
<tr>
<td>Leung et al. [33]</td>
<td>22.8</td>
<td>MCI</td>
<td>335</td>
<td>123</td>
<td>204</td>
<td>.02</td>
<td>79</td>
<td>82</td>
<td>.02</td>
<td>Comparison of MCI subgroups (reverters, stable, and converters) showed HCVs were lower and rates higher in converters compared with stable and reverter groups.</td>
</tr>
<tr>
<td>Stoub et al. [e4]</td>
<td>60</td>
<td>aMCI</td>
<td>29</td>
<td>11</td>
<td>18</td>
<td>.03</td>
<td>79</td>
<td>82</td>
<td>.03</td>
<td>Subjects with aMCI converting to AD had smaller entorhinal and HCV volumes at baseline compared with stable control subjects.</td>
</tr>
<tr>
<td>Tapiola et al. [e5]</td>
<td>34 (10–54)</td>
<td>MCI</td>
<td>60</td>
<td>9</td>
<td>47</td>
<td>.05 &lt; .01, &lt; .05</td>
<td>.05</td>
<td>79</td>
<td>82</td>
<td>.05</td>
</tr>
<tr>
<td>Visser et al. [e6]</td>
<td>36</td>
<td>MCI</td>
<td>13</td>
<td>9</td>
<td>4</td>
<td>.02</td>
<td>79</td>
<td>82</td>
<td>.02</td>
<td>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.</td>
</tr>
<tr>
<td>Visser et al. [e7]</td>
<td>22.8 (12–36)</td>
<td>MCI</td>
<td>30</td>
<td>7</td>
<td>23</td>
<td>.03</td>
<td>79</td>
<td>82</td>
<td>.03</td>
<td>HCV volume 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.</td>
</tr>
<tr>
<td>Wang et al. [12]</td>
<td>21.9 (10.7–32.8)</td>
<td>aMCI</td>
<td>58</td>
<td>19</td>
<td>39</td>
<td>.03</td>
<td>76.2</td>
<td></td>
<td></td>
<td>HC volume predicted MCI progression to AD.</td>
</tr>
<tr>
<td>Whitwell et al. [e8]</td>
<td>44</td>
<td>aMCI</td>
<td>63</td>
<td>42</td>
<td>21</td>
<td>.03</td>
<td>76.2</td>
<td></td>
<td></td>
<td>HC volume showed no significant differences between the aMCI-stable group and the aMCI-progressors group.</td>
</tr>
</tbody>
</table>

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.