Mild Cognitive Impairment: Baseline and Longitudinal Structural MR Imaging Measures Improve Predictive Prognosis[1]

Linda K. McEvoy, PhD
Dominic Holland, PhD
Donald J. Hagler, Jr, PhD
Christine Fennema-Notestine, PhD
James B. Brewer, MD, PhD
Anders M. Dale, PhD
For the Alzheimer's Disease Neuroimaging Initiative

Purpose: To assess whether single-time-point and longitudinal volumetric magnetic resonance (MR) imaging measures provide predictive prognostic information in patients with amnestic mild cognitive impairment (MCI).

Materials and Methods: This study was conducted with institutional review board approval and in compliance with HIPAA regulations. Written informed consent was obtained from all participants or the participants' legal guardians. Cross-validated discriminant analyses of MR imaging measures were performed to differentiate 164 Alzheimer disease (AD) cases from 203 healthy control cases. Separate analyses were performed by using data from MR images obtained at one time point or by combining single-time-point measures with 1-year change measures. Resulting discriminant functions were applied to 317 MCI cases to derive individual patient risk scores. Risk of conversion to AD was estimated as a continuous function of risk score percentile. Kaplan-Meier survival curves were computed for risk score quartiles. Odds ratios (ORs) for the conversion to AD were computed between the highest and lowest quartile scores.

Results: Individualized risk estimates from baseline MR examinations indicated that the 1-year risk of conversion to AD ranged from 3% to 40% (average group risk, 17%; OR, 7.2 for highest vs lowest score quartiles). Including measures of 1-year change in global and regional volumes significantly improved risk estimates ($P = 001$), with the risk of conversion to AD in the subsequent year ranging from 3% to 69% (average group risk, 27%; OR, 12.0 for highest vs lowest score quartiles).

Conclusion: Relative to the risk of conversion to AD conferred by the clinical diagnosis of MCI alone, MR imaging measures yield substantially more informative patient-specific risk estimates. Such predictive prognostic information will be critical if disease-modifying therapies become available.

Supplemental material: *http://radiology.rsna.org/lookup/suppl/doi:10.1148/radiol.11101975/-/DC1*

[1] From the Departments of Radiology (L.K.M., D.J.H., C.F., J.B.B., A.M.D.), Neurosciences (D.H., J.B.B., A.M.D.), and Psychiatry (C.F.), University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093. Received September 30, 1010; revision requested November 22; revision received January 5, 2011; accepted January 19; final version accepted January 27. Supported by the National Institutes of Health (grants R01 AG031224, K01 AG029218, K02 NS067427, U01 AG024904, P30 AG010129, K01 AG030514). A complete list of the organizations that provided support for this study is given in the Acknowledgments. **Address correspondence to** L.K.M. (e-mail: *lkmcevoy@ucsd.edu*).

The data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database *(http://www.loni.ucla.edu/ADNI)*. As such, the investigators in ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis of data in or the writing of this report. A complete list of the ADNI investigators is available at *http://www.loni.ucla.edu/ADNI/Data/ADNI_Authorship_List.pdf*.

Although there currently are no disease-modifying therapies approved for use in individuals with Alzheimer disease (AD), numerous potential treatments are under development (1). There remains hope that one or more of these potential treatments will be effective at altering the course of AD in the coming years (1). Treatment is likely to be most beneficial if it is applied early, before the onset of dementia (2,3). Since these therapies may be associated with substantial adverse effects, including brain edema and hemorrhage (4), the ability to identify an individual's risk of developing AD will be critical for clinical decision making (5).

Amnestic mild cognitive impairment (MCI) is associated with an elevated risk of AD, with affected individuals having an annual conversion rate of 15%–20% per year compared with a conversion rate of 1%–2% for the general population (6,7). Although the majority of individuals with MCI eventually develop AD, some develop other clinical conditions, others remain stable for long periods, and still others revert to normal cognitive status (8).

Biomarkers consistent with AD may help to improve the predictive prognosis in individuals with MCI. Structural magnetic resonance (MR) imaging is sensitive to neurodegeneration in cases of AD and MCI, depicting widespread atrophy in cortical association areas with prominent involvement of mesial temporal regions (9–11). The magnitude and pattern of atrophy in MCI are predictive of AD (12–16), as is the rate of whole-brain and regional atrophy progression (12,17–20). The purpose of this study was to assess whether single-time-point and longitudinal volumetric MR imaging measures provide predictive prognostic information in patients with MCI.

Advances in Knowledge

- A quantitative estimate of the risk of conversion to Alzheimer disease (AD) over the short term for a patient with mild cognitive impairment (MCI) can be derived from a summary numeric value that reflects the degree to which the individual's MR images depict the regional atrophy pattern associated with AD.
- Inclusion of measures of change over a 1-year period derived from serial MR examinations facilitates substantial improvement in risk prediction compared with the risk prediction achieved by using MR images obtained at a single time point.

Materials and Methods

One author (A.M.D.) is a founder of and holds equity interest in CorTechs Labs (La Jolla, Calif) and serves on the scientific advisory board of this company. The terms of this arrangement were reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. The spouse of another author (L.K.M.) is president of CorTechs Labs. The authors had control of the data and information submitted for publication.

In this prospective study, we analyzed MR imaging data that were collected and processed between August 26, 2005, and October 14, 2010, and clinical diagnostic follow-up data that were available November 17, 2010. This study was conducted with institutional review board approval and in compliance with Health Insurance Portability and Accountability Act regulations. Written informed consent was obtained from all participants or the participants' legal guardians.

Description of ADNI

The data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database *(www.loni.ucla.edu/ADNI)*. The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and nonprofit organizations. The primary goal of ADNI is to test whether serial MR imaging, positron emission tomography (PET), biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians in developing new treatments and monitoring the effectiveness of these therapies and to help lessen the time and costs of clinical trials.

ADNI is the result of the efforts of many coinvestigators from a broad range of academic institutions and private corporations; the principal investigator is Michael Weiner of the University of California, San Francisco. Participating subjects were recruited from more than 50 sites across the United States

Implication for Patient Care

- Volumetric analysis of structural MR imaging data can be used to quantify the 1-year risk of converting to AD in individual patients with amnestic MCI; the ability to determine an individual's risk of developing AD will be critical if potentially aggressive disease-modifying treatments become available.

Published online before print
10.1148/radiol.11101975

Radiology 2011; 259:834–843

Abbreviations:
AD = Alzheimer disease
ADNI = Alzheimer's Disease Neuroimaging Initiative
MCI = mild cognitive impairment
QDA = quadratic discriminant analysis

Author contributions:
Guarantors of integrity of entire study, L.K.M., J.B.B., A.M.D.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, L.K.M., C.F., J.B.B., A.M.D.; clinical studies, J.B.B.; statistical analysis, L.K.M., J.B.B., A.M.D.; and manuscript editing, all authors

Funding:
This research was supported by the NIH (grants R01 AG031224, K01 AG029218, K02 NS067427, U01 AG024904, P30 AG010129, K01 AG030514).

Potential conflicts of interest are listed at the end of this article.

See also the article by Chiang et al in this issue.

and Canada. For ADNI, 229 cognitively healthy individuals, 398 subjects with MCI, and 192 subjects with mild AD were recruited. The healthy control subjects and participants with MCI are being followed up for 3 years, with follow-up visits at 6, 12, 18 (for subjects with MCI only), 24, and 36 months. The participants with AD are being followed up for 2 years, with follow-up visits at 6, 12, and 24 months. (Visit *http://www.adni-info.org/* for additional information.)

Participants

ADNI eligibility criteria have been described elsewhere (7). Briefly, subjects are 55–90 years of age and have a study partner who is able to provide an independent evaluation of their cognitive function. The control subjects have Mini-Mental State Examination (MMSE) scores of 24–30 and a clinical dementia rating of 0. The subjects with MCI have MMSE scores of 24–30, subjective memory complaints, objective memory loss measured by using education-adjusted Wechsler Memory Scale Logical Memory II test scores, a clinical dementia rating of 0.5, preserved ability to perform daily living activities, and absence of dementia. The subjects with AD have MMSE scores of 20–26 and a clinical dementia rating of 0.5 or 1.0 and meet the criteria for probable AD of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (21).

MR imaging data that were available as of October 14, 2010, were analyzed. Of the 229 healthy control subjects participating in ADNI, 14 (6.1%) individuals whose cognitive status progressed to MCI or AD over the course of the study and 12 (5.2%) whose baseline MR images did not meet local quality control criteria were excluded, leaving 203 (88.6%) healthy control subjects. Of the 398 subjects with MCI in ADNI, 43 (10.8%) were excluded because their baseline MR images did not meet local quality control criteria and an additional 38 (9.5%) were excluded because they did not have 1-year clinical follow-up data, leaving 317 (79.6%) subjects with MCI. Of the 192 subjects with AD in ADNI, 28 (14.6%) were excluded because their baseline MR images did not meet local quality control criteria, leaving 164 (85.4%) subjects with AD in the analysis.

Demographic Data

Parameter	Healthy Control Subject Group ($n = 203$)	MCI Group ($n = 317$)	Mild AD Group ($n = 164$)
Subject age (y)	76.04 ± 5.09	74.74 ± 7.33	74.94 ± 7.60
No. of male subjects*†	105 (52)	199 (63)	82 (50)
Years of education‡	16.02 ± 2.89	15.76 ± 2.96	14.76 ± 3.18
MMSE score‡	29.13 ± 1.03	27.05 ± 1.74	23.32 ± 2.00
CDR-SB score†	0.03 ± 0.12	1.57 ± 0.88	4.23 ± 1.59
APOE ε4 allele*‡	53 (26)	173 (55)	110 (67)

Note.—Unless otherwise noted, data are mean values ± standard deviations. APOE = apolipoprotein E, MMSE = Mini-Mental State Examination, CDR-SB = clinical dementia rating scale sum of boxes.

* Data are numbers of subjects, with percentages in parentheses.

† $P < .01$ for comparison of the groups.

‡ $P < .001$ for comparison of the groups.

Group demographic data are presented in the Table. The subject groups did not differ significantly in age ($F_{2,682} = 2.53$, $P = .081$). The subjects with AD had less education ($F_{2,682} = 8.70$, $P < .001$) than did the healthy control subjects and the subjects with MCI, and the MCI group contained the highest proportion of men ($\chi^2 = 9.99$, $P = .007$). One-year follow-up MR imaging data that met local quality control criteria were available for 149 (73.4%) healthy control subjects, 104 (63.4%) subjects with AD, and 199 (62.8%) subjects with MCI. For the longitudinal analyses, the MCI group was further reduced to 170 (53.6%) participants who had 2-year clinical diagnostic data and whose cognitive status had not declined to AD within the first year of the study.

Clinical diagnostic follow-up data were available for 315 (99.4%) of the participants with MCI at 6 months, for all 317 (100%) participants at 12 months, for 281 (88.6%) participants at 18 months, for 258 (81.4%) participants at 24 months, and for 188 (59.3%) participants at 36 months. Conversion to AD was determined according to the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association. Operational criteria are provided in Appendix E1 (online).

Procedures

Raw Digital Imaging and Communications in Medicine MR images (including two T1-weighted image volumes per subject per visit) were downloaded from the public ADNI Web site *(http://www.loni.ucla.edu/ADNI/Data/index.shtml)*. These data were collected by using a variety of MR units, with imaging protocols individualized for each unit (22). (Visit *http://www.loni.ucla.edu/ADNI/Research/Cores/index.shtml*.) Locally, the raw MR imaging data were reviewed for quality, automatically corrected for spatial distortion due to gradient nonlinearity (23) and B_1 field inhomogeneity (24), and registered, and the two image volumes in each subject were averaged to improve the signal-to-noise ratio. Volumetric segmentation (25) and cortical surface reconstruction (26–28) and parcellation (29,30), based on a locally optimized version of the FreeSurfer software package (version 3.02; developed at the University of California, San Diego, CorTechs Labs, and the Martinos Center for Biomedical Imaging, Boston, Mass), were used to quantify the baseline thicknesses and volumes of brain regions, as described in detail elsewhere (11,19). Local optimization of this software included improved corrections for spatial distortions induced by gradient nonlinearities (23) and B_1 field inhomogeneity (24), achieved by using algorithms updated to reflect the MR unit upgrades at the different data collection sites. The image

correction procedures performed by the ADNI investigators have not been updated for recent MR unit changes and thus may not provide as precise a correction for site-specific distortion effects and signal inhomogeneities. These corrections improve performance in automated segmentation and parcellation procedures, enabling sensitive detection of subtle effects (23). Volumetric data were corrected for differences in head size (31) by dividing each measurement by the estimated total intracranial volume, which was computed as the sum of the volumes of all four ventricles, the white matter, and the gray matter, with the brainstem excluded.

Longitudinal changes in brain structure measures were quantified by using a method developed at the Multimodal Imaging Laboratory, University of California, San Diego, called quantitative anatomic regional change analysis, or QUARC (19). The MR images of each subject at each follow-up time point were corrected for spatial distortion due to gradient nonlinearity and were rigid-body aligned, and the two images obtained at that time point were averaged. The averaged follow-up image was registered to the subject's baseline image by using a 12-parameter affine registration and then intensity normalized to the baseline image by using an iterative procedure. A deformation field was then calculated from the nonlinear registration, as described by Holland et al (19), and used to align the images at the subvoxel level, resulting in a one-to-one correspondence between each vertex on the baseline and follow-up images. Unlike some longitudinal analysis methods that have been shown to yield significant overestimations of the changes owing to bias in serial registration (32), the longitudinal registration procedure used in this investigation is free of bias. Subcortical segmentation and cortical parcellation labels from the baseline image were used to extract an average volume change for each region of interest. A trained technician who had 18 months experience and was supervised by image analysis experts (D.H., A.M.D., 5 and 20 years of experience, respectively) performed a visual quality control in the volume change field, with blinding to the diagnosis, to exclude cases with registration degradation due to artifacts (eg, patient motion, imaging unit changes over time [19]).

Statistical Analyses

Analyses of covariance were used to compare continuous measures between the groups, and χ^2 statistics were used to compare categorical measures between the groups. Group comparisons of clinical measures, such as the clinical dementia rating scale sum of boxes score and the Mini-Mental State Examination score, included age, sex, and education as covariates. Prior to the analyses, the effects of sex and age were regressed from all brain measures.

To determine the sensitivity and specificity for discriminating between AD and healthy control subject data, ridge-regularized quadratic discriminant analyses (QDAs) were performed by using the regularization parameter that provided the best generalization performance based on leave-one-out cross validation. QDA was implemented in MatLab (release 2010b; MathWorks, Natick, Mass) and was used in this study because data from the healthy control subjects and subjects with AD showed different covariance structures. With QDA, one assumes that the distribution of feature vectors in each class is a multivariate Gaussian distribution; however, unlike the case with linear discriminant analysis, the within-class covariance structure may differ. This allows the decision surface to be quadratic in form. When the covariance matrices differ and can be accurately estimated, as was the case here, QDA facilitates better classification performance than does linear discriminant analyses (33). Linear support vector machine models are another commonly used approach to multivariate classification and have been applied to voxel-wise data to distinguish subjects with AD from healthy control subjects (34–37). However, the resulting linear decision surface is suboptimal when the covariance of features differs between groups. This necessitates a linear surface decision that is suboptimal when the variance of features differs between groups. Furthermore, the large number of features relative to the number of training samples may result in overfitting of the data. Finally, a limited number of discrete regions of interest that are known to be affected early in AD may have greater face validity as a biomarker of disease than a complex pattern of voxels across the cortex, where the signal intensity of one voxel is indicative of AD but that of a neighboring voxel is not.

In our investigation, the input features were restricted to two nonspecific measures of brain atrophy—the whole-brain volume and the volume of the inferior lateral ventricle (temporal horn of the lateral ventricle)—and to measures in eight regions previously found to best facilitate the differentiation of AD data from healthy control subject data (15): hippocampus, entorhinal cortex, middle and superior temporal gyri, bank of the superior temporal sulcus, medial and lateral orbitofrontal cortices, and isthmus cingulate (Fig 1). For each region, left and right hemisphere values were averaged.

These regions were chosen on the basis of findings in a prior analysis involving a subset of the healthy control subjects and subjects with AD included in the current study (211 of 367 subjects), and there was a small possibility that this would cause an optimistic bias in the estimation of sensitivity and specificity for differentiating healthy control subjects from subjects with AD. To verify that this did not occur, the classification accuracy achieved by using the entire group was compared with that achieved by using the subset of subjects with AD and healthy control subjects who were not involved in the prior feature selection analysis (n = 156).

Separate analyses were performed to determine classification accuracy by using baseline MR imaging measures alone and by using both baseline measures and measures of the 1-year rate of change. Receiver operating characteristic curves were computed from the resulting QDA scores by using leave-one-out cross validation. Areas under the receiver operating characteristic curve were compared by using the method of Hanley and McNeil (38).

QDA classifiers that were trained by using data from all subjects in the

Figure 1

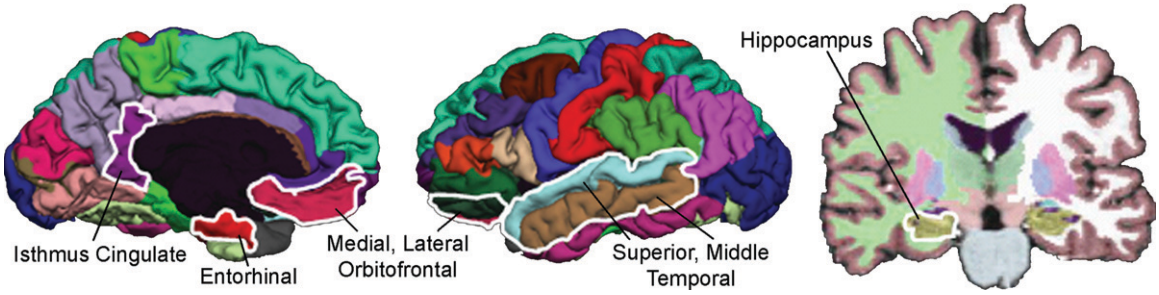

Figure 1: Regions examined in QDA are outlined (in white) on medial (left) and lateral (middle) views of reconstructed parcellated left-hemisphere cortical surface and on segmented coronal view (right) of brain. Average thickness or volume across left- and right-hemisphere regions was used. Examined regions included isthmus cingulate, entorhinal cortex, medial and lateral orbitofrontal cortices, superior and middle temporal gyri, bank of superior temporal sulcus (not visible), and hippocampus. These regions were chosen on the basis of their utility for discrimination of AD data from healthy control subject data in a prior study involving a subset of the subjects with AD and healthy control subjects evaluated in the current study (211 [57.5%] of 367 subjects) (28). Whole-brain and inferior lateral ventricle volumes (not shown) also were included in the analysis.

combined group of healthy control subjects and subjects with AD were applied to the MCI group data to obtain a QDA score for each subject. The QDA score is the log of the a posteriori probability of belonging to the AD group, based on the estimated mean and covariance matrix for each group. The QDA score derived from the baseline data is the atrophy score, and the QDA score derived from the baseline and 1-year follow-up measures is the atrophy progression score. Kaplan-Meier survival curves were estimated for atrophy score and atrophy progression score quartiles. Odds ratios and corresponding 95% confidence intervals of the risk of converting to AD as a function of highest versus lowest score quartile were computed for atrophy and atrophy progression scores. The risk of conversion from MCI to AD was estimated as a continuous function of atrophy score and atrophy progression score percentile by using a binomial smoothing spline-fitting procedure (39) implemented in MatLab and was applied to the binary variable of conversion at 1 year for each subject. The improvement in cognitive status prediction that resulted from the inclusion of atrophy progression measures derived at follow-up MR imaging was considered the integrated discrimination improvement (40). The integrated discrimination improvement is essentially the difference between the improvement in average sensitivity and the increase in the average false-positive rate facilitated by the addition of a new variable into the discriminative model (40).

Results

Demonstrating the sensitivity of the MR measures to neurodegeneration in AD, Figure 2 shows plots of brain region volumes (as percentages of estimated total intracranial volume) and thicknesses and the changes in volume and thickness across all evaluatable follow-up visits for four of the brain regions evaluated. Figure 3 shows scatterplots of baseline volume and thickness versus annual percentage change in these measures, after the effects of age and sex were regressed out.

The QDA classifier applied to the baseline MR measures enabled the differentiation of subjects with AD from healthy control subjects with high cross-validated sensitivity (85%, 140 of 164 subjects with AD) and specificity (93%, 189 of 203 healthy control subjects). Similarly, high classification accuracy was achieved for the 156 subjects who were not included in the prior published analysis (83% sensitivity [69 of 83 subjects with AD], 93% specificity [68 of 73 healthy control subjects]). Incorporating the 1-year change measures improved cross-validated sensitivity (91%, 95 of 104 subjects with AD) and specificity (91%, 133 of 147 healthy control subjects) to result in significantly larger areas under the receiver operating characteristic curve (0.967 vs 0.936, $P < .01$) (Figs E1 and E2 [online]).

Applying the QDA classifiers, which were trained with data from all healthy control subjects and subjects with AD, to the data from the subjects with MCI yielded atrophy and atrophy progression scores for each patient with MCI. Kaplan-Meier survival curves of atrophy score quartiles for the MCI group are shown in Figure 4a. Individuals with atrophy scores in the highest quartile had a significantly higher risk of conversion to AD than did those with scores in the lowest quartile (odds ratio, 7.2; 95% confidence interval: 3.45, 14.96). The individualized risk estimates obtained by fitting the 1-year conversion risk as a continuous function of atrophy score percentile are shown in Figure 4b. Risks of converting to AD ranged from 3% to 40%. (The straight horizontal line in Fig 4b indicates the average conversion risk for the MCI group [17%].) Relative to this average risk, an approximately fivefold decrease in risk was observed in the subjects with scores in the lowest percentile, whereas those with scores in the highest percentile had a greater than twofold increase in risk.

Kaplan-Meier survival curves for atrophy progression score quartiles are shown in Figure 5a. This analysis was

Figure 2

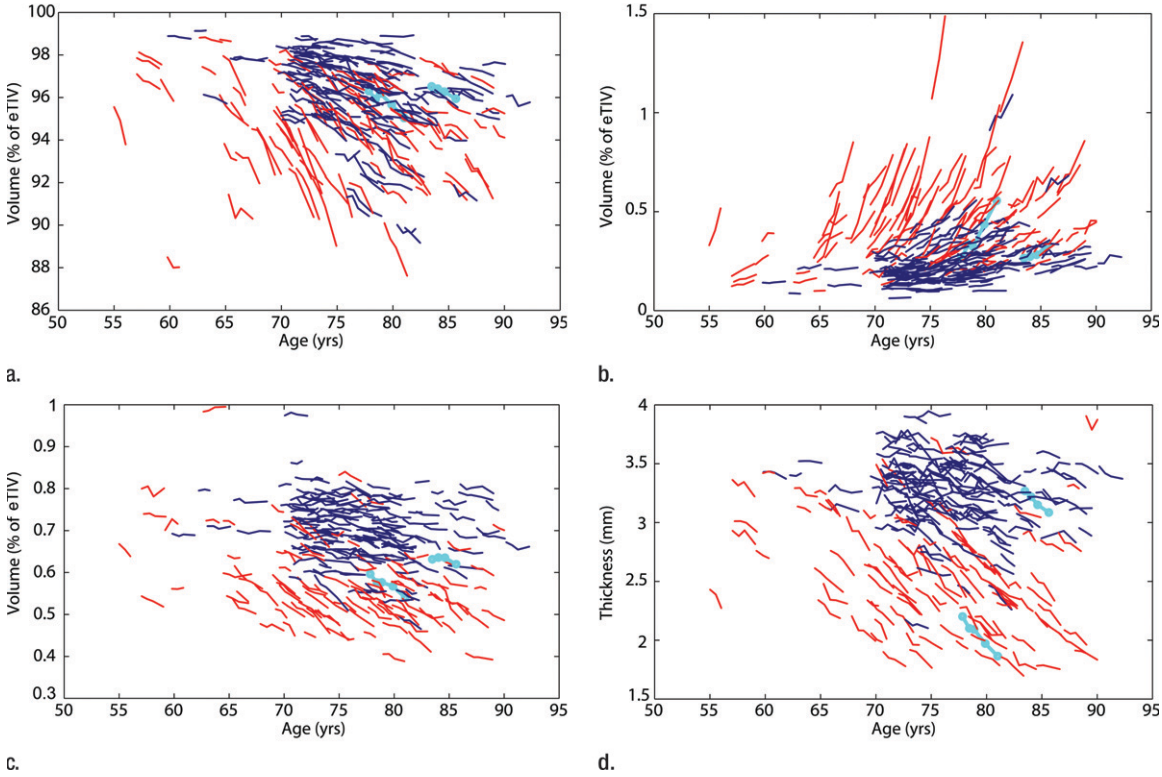

Figure 2: Spaghetti plots illustrate volumes (as percentages of estimated total intracranial volume *[eTIV]*) and volume changes across all evaluatable follow-up visits as functions of age for **(a)** whole brain, **(b)** inferior lateral ventricle, and **(c)** hippocampus. **(d)** For entorhinal cortex, thicknesses and changes in thickness are shown. Healthy control subject data are in dark blue; AD data are in red. Data from the two healthy control subjects whose cognitive status changed to AD by the time of these analyses are in light blue. Although these individuals were excluded from the healthy control subject group before the analyses, their data are shown here as evidence of the sensitivity of these structural MR measures to AD atrophy in individuals in a clinically unrecognized prodromal AD state.

restricted to the patients with MCI whose cognitive status had not changed to AD by the 1-year visit and who had valid 1-year follow-up MR imaging data and 2-year clinical diagnostic data. The risk of conversion to AD increased with increasing atrophy progression score. Individuals with atrophy progression scores in the highest quartile had a significantly higher risk of developing AD within the next year than did those with scores in the lowest quartile (odds ratio, 12.0; 95% confidence interval: 4.1, 35.2). Individualized risk estimates, shown in Figure 5b, ranged from 3% to 69%. Relative to the average group risk (27%), a greater than sixfold decrease in risk of conversion within the next year was observed in those subjects with scores in the lowest percentile, whereas those with scores in the highest percentile had a greater than twofold increase in risk.

To determine the value of including measures of longitudinal change in addition to the atrophy measures from a single-time-point MR imaging examination, individualized risk estimates were derived from the atrophy scores for thickness and volume measures calculated at the 1-year follow-up MR examination. The thickness and volume measures from the baseline MR examination were multiplied by the 1-year percentage change value and added to the baseline volume and thickness measures to obtain thickness and volume values for 1-year follow-up MR imaging. With use of the atrophy scores from 1-year MR imaging, the risks of conversion to AD during the subsequent year (ie, year 2 of the study) ranged from 3% to 43% (Fig 5c), whereas the risks based on the atrophy progression scores (calculated by using baseline MR and 1-year change measures) ranged from 3% to 69%. The significant integrated discrimination improvement score (0.081; z score, 3.24; P = .001) demonstrated that the inclusion of rate-of-change measures significantly improved predictive ability relative to that afforded by using single-time-point measures.

Discussion

The results of this study demonstrate that patient-specific estimates of the risk of conversion from MCI to AD can be derived from quantitative measures of brain atrophy obtained from both single-time-point and serial MR imaging

Figure 3

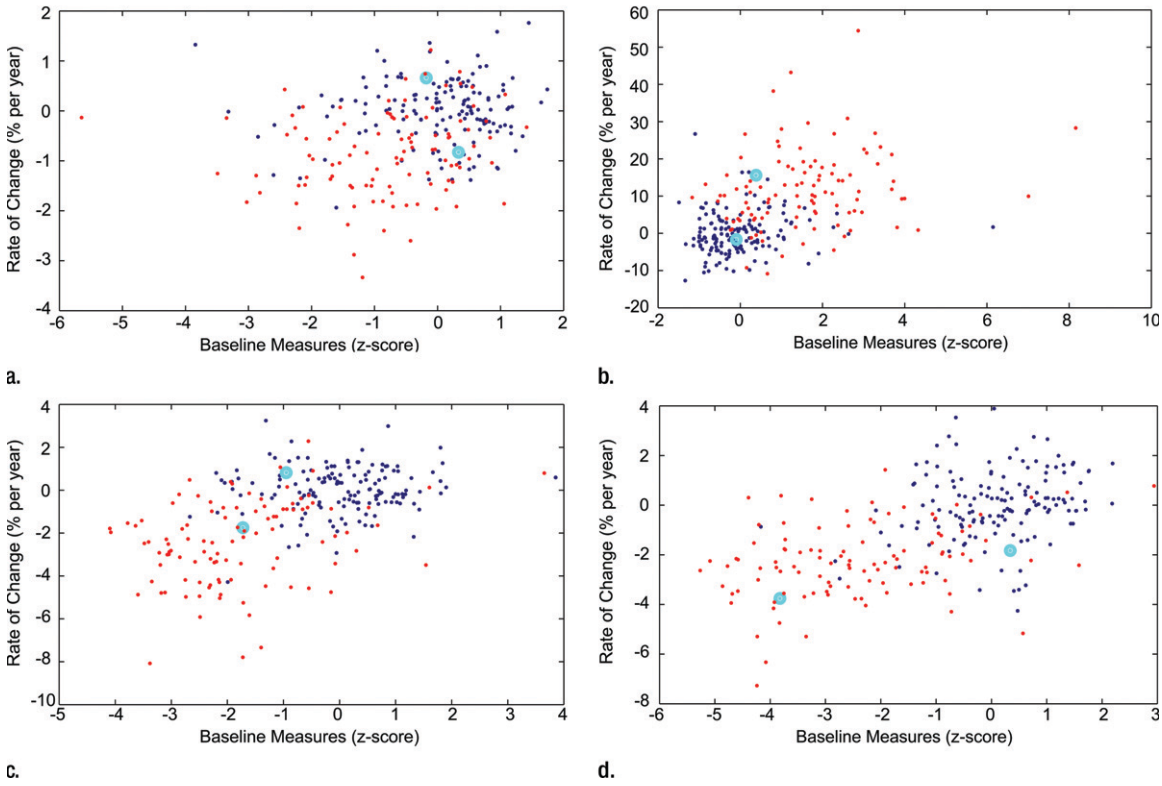

Figure 3: Scatterplots of baseline measures versus annual percentage changes for **(a)** whole brain, **(b)** inferior lateral ventricle, **(c)** hippocampus, and **(d)** entorhinal cortex after effects of age and sex are regressed out. Dark blue dots represent healthy control subject data, and red dots represent AD data. Light blue circles represent data for two healthy control subjects whose cognitive status changed to AD. As evident from these data, the informative value of baseline and longitudinal change measures differs for different brain regions. For example, healthy control subjects and subjects with AD show substantial overlap in whole-brain and inferior ventricle volumes at baseline, but subjects with AD show larger annual percentage changes than do healthy control subjects. Greater group variance in baseline and longitudinal change measures are observed in entorhinal cortex and to a lesser extent in hippocampus. For all regions, baseline measures and percentage changes differ significantly between healthy control subject and AD groups ($P < .001$ for all comparisons). See Table E1 (online) for statistical comparisons.

examinations. Use of these measures substantially improves risk predictions compared with risk predictions based on the clinical MCI diagnosis alone. An individual who receives a diagnosis of amnestic MCI based on the commonly accepted clinical criteria used in ADNI has a 15%–20% risk of developing AD within 1 year (6,7). However, use of structural MR imaging data from a baseline examination considerably improves risk prediction, enabling the identification of individuals who show only slightly elevated risk relative to elderly subjects without cognitive impairment (3% per year) and individuals with a substantially higher risk (up to 40% per year). Information regarding the rate of atrophy progression over a 1-year period enables even greater risk discrimination compared with the information gleaned from a single MR examination, enabling the identification of individuals who are at very high risk (69%) of developing AD within the next year.

This study builds upon a large body of literature addressing the sensitivity of structural MR imaging to neurodegeneration in AD and the capability of multivariate models based on structural MR imaging data in the prediction of clinical decline in subjects with MCI (16,35,41). The survival curves from baseline MR examinations reported herein show greater risk differentiation than do those reported by Vemuri et al (16). An explanation for this could be the measurements of the anatomies of individual subjects in the current study rather than the measurements that Vemuri et al (16) obtained by using the atlas- and voxel-based morphometric method, and the potentially better generalization performance of the discriminant model due to the use of a small number of predefined features rather than a large number of voxels across the imaging volume (33).

The finding that the inclusion of information regarding the rate of atrophy progression identified at serial MR imaging improves risk prediction was not unexpected, as prior studies have shown atrophy rate to have higher sensitivity than baseline measures for predicting cognitive decline (12,42,43). When brain measure information is obtained at only a

Figure 4

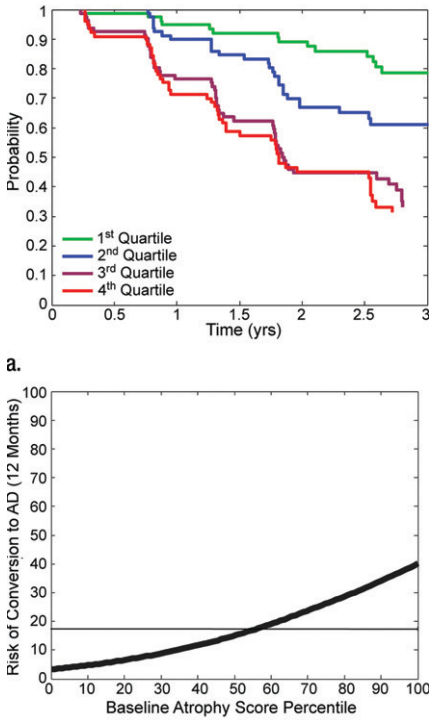

Figure 4: **(a)** Kaplan-Meier survival curves of increasing atrophy score quartiles for patients with MCI. Probabilities of remaining free of AD as functions of time are shown for atrophy scores derived from baseline MR measures. Probabilities at 0–25th (1st quartile), 26th–50th (2nd quartile), 51st–75th (3rd quartile), and 76th–100th (4th quartile) percentiles are shown. Log rank statistics for each quartile versus next higher quartile are as follows: 1st versus 2nd quartile: $\chi^2 = 5.16$, $P = .023$; 2nd versus 3rd quartile: $\chi^2 = 12.49$, $P < .001$; and 3rd versus 4th quartile: $\chi^2 = 0.127$, $P = .72$. **(b)** Thick curved line on graph illustrates risk of conversion to AD within 1 year as a continuous function of baseline atrophy score percentile, estimated by using binomial smoothing spline-fitting procedure. Straight horizontal line indicates average group risk (17%). A new patient's atrophy score can be converted to a percentile score on the basis of the distribution of scores in the current study sample. The new individual's risk estimate can then be read from this graph.

Figure 5

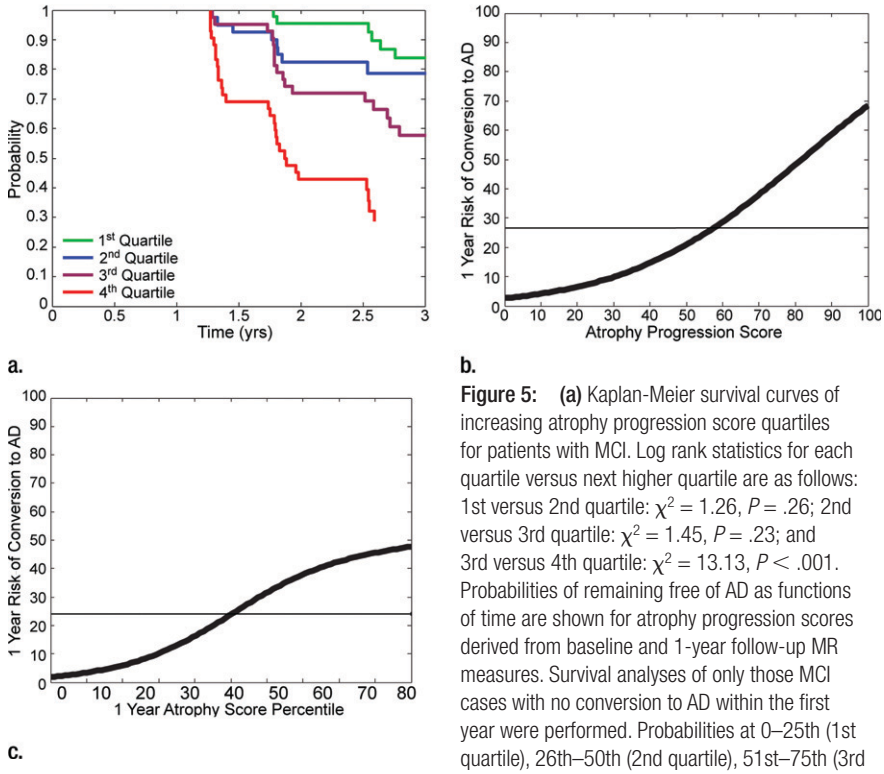

Figure 5: **(a)** Kaplan-Meier survival curves of increasing atrophy progression score quartiles for patients with MCI. Log rank statistics for each quartile versus next higher quartile are as follows: 1st versus 2nd quartile: $\chi^2 = 1.26$, $P = .26$; 2nd versus 3rd quartile: $\chi^2 = 1.45$, $P = .23$; and 3rd versus 4th quartile: $\chi^2 = 13.13$, $P < .001$. Probabilities of remaining free of AD as functions of time are shown for atrophy progression scores derived from baseline and 1-year follow-up MR measures. Survival analyses of only those MCI cases with no conversion to AD within the first year were performed. Probabilities at 0–25th (1st quartile), 26th–50th (2nd quartile), 51st–75th (3rd quartile), and 76th–100th (4th quartile) percentiles are shown. **(b)** Thick curved line on graph illustrates risk of conversion to AD within subsequent year as a continuous function of atrophy progression score percentile. **(c)** Thick curved line on graph illustrates risk of conversion to AD within subsequent year as a continuous function of atrophy score percentile, derived by using the 1-year MR imaging data as the baseline. Comparison of results in **b** and **c** illustrate the additional risk modification information provided by including rate-of-change measures. Atrophy measures from a single time point provide valuable risk information, enabling the identification of individuals at much lower (3%) and much higher (43%) risk of developing AD compared with average group risk of 27% (straight horizontal line in **c**). However, information derived from follow-up MR imaging enables even greater separation of risk, with the 1-year risk of conversion ranging from a risk similar to that of non–cognitive impaired elderly individuals (3%) to 69% (**b**).

single time point, differences in regional thicknesses and volumes among individuals can mask disease-related atrophy, as was the case in a single healthy control subject in the current study whose cognitive status changed to AD. This individual had baseline values well within the normal range on each of four MR images collected over a 3-year period. However, the rate of change over time was substantially accelerated relative to that observed in the healthy individuals who remained stable and was particularly evident in the entorhinal cortex.

In addition to structural MR imaging parameters, several other potential biomarkers may aid in the early diagnosis of AD; these include decreased glucose metabolism at fluorine 18 fluorodeoxyglucose (FDG) PET, evidence of amyloid-β pathology from cerebrospinal fluid or PET imaging of brain amyloid deposition, and evidence of neurodegeneration from cerebrospinal fluid levels of tau proteins. Among these biomarkers, MR parameters seem to be the best suited for predictive prognosis in cases of MCI: Although hypometabolism at FDG PET correlates with clinical symptoms and is used to predict conversion to AD (44–46), PET is not as widely available and is more expensive compared with MR imaging, and it involves exposure to ionizing radiation. Furthermore, MR imaging measures have been found to be just as sensitive for detecting mild and prodromal AD as PET measures (47), and the neurodegeneration evident at MR imaging may precede the hypometabolism detected with FDG PET (48).

Biomarkers of amyloid pathology have high sensitivity for detecting AD (49), but a substantial number of healthy individuals have test results that are positive for amyloid pathology. It remains unclear when, or even whether, such individuals will develop AD. Thus, biomarkers of amyloid pathology in patients with MCI do not necessarily indicate a high risk of imminent conversion to AD. Increased levels of cerebrospinal fluid tau proteins are associated with neurodegeneration and when present alone or in combination with cerebrospinal fluid amyloid levels are predictive of clinical decline in subjects with MCI (50). However, cerebrospinal fluid biomarkers have been found to be less sensitive than MR imaging measures in the prediction of clinical decline and conversion to AD (51,52).

MR imaging is a promising adjunct to the clinical diagnosis in cases of AD. As observed in a recent review (53), barriers to the routine clinical use of structural MR imaging measures, such as lack of standardized image acquisition protocols, spatial distortion and motion artifacts, labor-intensive image analyses, and lack of normative values, are being overcome with the aid of large multisite clinical trials such as ADNI. Much effort was expended in ADNI to develop image acquisition protocols that were suitable for automated image analyses and standardized across multiple sites and vendor platforms (22). Efforts of ADNI (22) and the Morphometry Biomedical Informatics Research Network (23) have facilitated the development of automated methods for high-throughput semiautomated image analysis that are robust to intersite variation and sensitive to longitudinal change and that yield individual-specific thickness and volume measures (11,19). The ADNI also provides the data needed to compute normative values. These advancements have paved the way for the translation of research results on the sensitivity and predictive value of MR imaging biomarkers for detecting and monitoring AD to clinical practice (53).

This study had several limitations. First, we do not have histopathologic verification of the diagnosis of AD in any patient. Second, the estimates of risk for conversion to AD were obtained in a large group of patients with a diagnosis of MCI based on objective criteria determined by means of consensus judgment among leading experts in the field. Thus, the subject sample was not representative of the general clinical population, which includes diverse groups of patients with memory problems. This limitation may have also been a strength, however, in that the patient population was highly selected to represent individuals who are most likely to have prodromal AD, with the exclusion of those with other potential causes of memory impairment. Thus, the current results may be underestimations of the utility of MR imaging in general clinical practice, where patients with less uniform causes of memory impairment are encountered. Nevertheless, validation studies involving more representative clinical populations are warranted.

Acknowledgments: We thank Matthew Erhart, BS, Alain Koyama, BS, Robin Jennings, BS, Michele Perry, MS, Christopher Pung, BA, and Elaine Wu, BS, for downloading and preprocessing the ADNI MR imaging data. Data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca, Bayer Schering Pharma, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Company, Medpace, Merck and Company, Novartis, Pfizer, F. Hoffman-La Roche, Schering-Plough, Synarc, Wyeth, and nonprofit partners the Alzheimer's Association and the Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health *(http://www.fnih.org/)*. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514 and the Dana Foundation.

Disclosures of Potential Conflicts of Interest: L.K.M. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: none to disclose. Other relationships: spouse is president of CorTechs Labs. **D.H.** No potential conflicts of interest to disclose. **D.J.H.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: is a consultant for CorTechs Labs. Other relationships: none to disclose. **C.F.** No potential conflicts of interest to disclose. **J.B.B.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: received grants or grants are pending from General Electric Medical Foundation and Janssen Alzheimer's Immunotherapy. Other relationships: none to disclose. **A.M.D.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: is founder and holds equity in CorTechs Labs; serves on scientific advisory board of CorTechs Labs. Other relationships: none to disclose.

References

1. Rafii MS, Aisen PS. Recent developments in Alzheimer's disease therapeutics. BMC Med 2009;7:7.
2. Cummings JL, Doody R, Clark C. Disease-modifying therapies for Alzheimer disease: challenges to early intervention. Neurology 2007;69(16):1622–1634.
3. Aisen PS. Commentary on "a roadmap for the prevention of dementia II: Leon Thal Symposium 2008"—facilitating Alzheimer's disease drug development in the United States. Alzheimers Dement 2009;5(2):125–127.
4. Black RS, Sperling RA, Safirstein B, et al. A single ascending dose study of bapineuzumab in patients with Alzheimer disease. Alzheimer Dis Assoc Disord 2010;24(2):198–203.
5. McEvoy LK, Brewer JB. Quantitative structural MRI for early detection of Alzheimer's disease. Expert Rev Neurother 2010;10(11):1675–1688.
6. Petersen RC, Thomas RG, Grundman M, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. N Engl J Med 2005;352(23):2379–2388.
7. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. Neurology 2010;74(3):201–209.
8. Larrieu S, Letenneur L, Orgogozo JM, et al. Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. Neurology 2002;59(10):1594–1599.
9. Jack CR Jr, Petersen RC, Xu YC, et al. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 1997;49(3):786–794.
10. Dickerson BC, Feczko E, Augustinack JC, et al. Differential effects of aging and Alzheimer's disease on medial temporal lobe cortical thickness and surface area. Neurobiol Aging 2009;30(3):432–440.
11. Fennema-Notestine C, Hagler DJ Jr, McEvoy LK, et al. Structural MRI biomarkers for

preclinical and mild Alzheimer's disease. Hum Brain Mapp 2009;30(10):3238–3253.
12. Jack CR Jr, Shiung MM, Weigand SD, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 2005;65(8):1227–1231.
13. Whitwell JL, Shiung MM, Przybelski SA, et al. MRI patterns of atrophy associated with progression to AD in amnestic mild cognitive impairment. Neurology 2008;70(7):512–520.
14. Bakkour A, Morris JC, Dickerson BC. The cortical signature of prodromal AD: regional thinning predicts mild AD dementia. Neurology 2009;72(12):1048–1055.
15. McEvoy LK, Fennema-Notestine C, Roddey JC, et al. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. Radiology 2009;251(1):195–205.
16. Vemuri P, Wiste HJ, Weigand SD, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology 2009;73(4):294–301.
17. Fox NC, Cousens S, Scahill R, Harvey RJ, Rossor MN. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. Arch Neurol 2000;57(3):339–344.
18. Thompson PM, Hayashi KM, de Zubicaray G, et al. Dynamics of gray matter loss in Alzheimer's disease. J Neurosci 2003;23(3):994–1005.
19. Holland D, Brewer JB, Hagler DJ, Fennema-Notestine C, Dale AM; Alzheimer's Disease Neuroimaging Initiative. Subregional neuroanatomical change as a biomarker for Alzheimer's disease. Proc Natl Acad Sci U S A 2009;106(49):20954–20959.
20. McDonald CR, McEvoy LK, Gharapetian L, et al. Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. Neurology 2009;73(6):457–465.
21. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34(7):939–944.
22. Jack CR Jr, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008;27(4):685–691.
23. Jovicich J, Czanner S, Greve D, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage 2006;30(2):436–443.
24. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 1998;17(1):87–97.
25. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33(3):341–355.
26. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 1999;9(2):179–194.
27. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 1999;9(2):195–207.
28. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci U S A 2000;97(20):11050–11055.
29. Fischl B, van der Kouwe A, Destrieux C, et al. Automatically parcellating the human cerebral cortex. Cereb Cortex 2004;14(1):11–22.
30. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31(3):968–980.
31. Buckner RL, Head D, Parker J, et al. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. Neuroimage 2004;23(2):724–738.
32. Thompson WK, Holland D. Bias in tensor based morphometry stat-ROI measures may result in unrealistic power estimates. Neuroimage (in press).
33. Duda RO, Hart PE, Strok DG. Pattern classification. New York, NY: Wiley, 2000.
34. Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging 2008;29(4):514–523.
35. Fan Y, Batmanghelich N, Clark CM, Davatzikos C; Alzheimer's Disease Neuroimaging Initiative. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. Neuroimage 2008;39(4):1731–1743.
36. Klöppel S, Stonnington CM, Chu C, et al. Automatic classification of MR scans in Alzheimer's disease. Brain 2008;131(Pt 3):681–689.
37. Vemuri P, Gunter JL, Senjem ML, et al. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. Neuroimage 2008;39(3):1186–1197.
38. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839–843.
39. Gu C. Smoothing spline ANOVA models. New York, NY: Springer-Verlag, 2002.
40. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27(2):157–172; discussion 207–212.
41. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. Neuroimage 2009;44(4):1415–1422.
42. Rusinek H, De Santi S, Frid D, et al. Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. Radiology 2003;229(3):691–696.
43. Risacher SL, Shen L, West JD, et al. Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. Neurobiol Aging 2010;31(8):1401–1418.
44. Mosconi L. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease: FDG-PET studies in MCI and AD. Eur J Nucl Med Mol Imaging 2005;32(4):486–510.
45. Herholz K, Carter SF, Jones M. Positron emission tomography imaging in dementia. Br J Radiol 2007;80(Spec No 2):S160–S167.
46. Landau SM, Harvey D, Madison CM, et al. Comparing predictors of conversion and decline in mild cognitive impairment. Neurology 2010;75(3):230–238.
47. Karow DS, McEvoy LK, Fennema-Notestine C, et al. Relative capability of MR imaging and FDG PET to depict changes associated with prodromal and early Alzheimer disease. Radiology 2010;256(3):932–942.
48. Villain N, Fouquet M, Baron JC, et al. Sequential relationships between grey matter and white matter atrophy and brain metabolic abnormalities in early Alzheimer's disease. Brain 2010;133(11):3301–3314.
49. Rabinovici GD, Jagust WJ. Amyloid imaging in aging and dementia: testing the amyloid hypothesis in vivo. Behav Neurol 2009;21(1):117–128.
50. Mattsson N, Zetterberg H, Hansson O, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA 2009;302(4):385–393.
51. Fjell AM, Walhovd KB, Fennema-Notestine C, et al. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. J Neurosci 2010;30(6):2088–2101.
52. Vemuri P, Wiste HJ, Weigand SD, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: diagnostic discrimination and cognitive correlations. Neurology 2009;73(4):287–293.
53. Brewer JB. Fully-automated volumetric MRI with normative ranges: translation to clinical practice. Behav Neurol 2009;21(1):21–28.