

**ORIGINAL
RESEARCH**

F.L. Giesel
H.K. Hahn
P.A. Thomann
E. Widjaja
E. Wignall
H. von Tengge-Kobligh
J. Pantel
P.D. Griffiths
H.O. Peitgen
J. Schroder
M. Essig

Temporal Horn Index and Volume of Medial Temporal Lobe Atrophy Using a New Semiautomated Method for Rapid and Precise Assessment

BACKGROUND AND PURPOSE: Quantitative markers of Alzheimer disease (AD), particularly in the early stages, are needed for clinical assessment and monitoring. We have evaluated a novel method to segment and visualize the ventricular system and obtain volumetric measures thereof. The temporal horn volume (THV) and index in patients with mild cognitive impairment (MCI) and in those with AD were evaluated.

METHODS: High-resolution T1-weighted volume imaging was performed in 52 subjects (21 patients with MCI, 10 with AD, and 21 healthy control subjects). An interactive watershed transformation and semiautomated histogram analysis were implemented to produce segmented THV and temporal horn indices (THI) (ratio of THV to lateral ventricular volume).

RESULTS: Cerebral ventricular and temporal horn size could be semiautomatically quantified from all 52 datasets. The method was fast and rater-independent. Qualitative ventricular inspections using surface rendering shading could uncover atrophic process with enlargement of the whole and especially temporal horn volume. Both THV and THI of patients with AD were significantly larger than those of patients with MCI or control subjects ($P < .005$). There was no significant difference in THV and THI between patients with MCI or control subjects ($P > .05$). There was a significant correlation between the neuropsychologic performance and both THI and THV across groups ($P < .01$).

CONCLUSION: THV and THI could be used as markers of AD in the clinical environment and are expected to be helpful in monitoring therapeutic intervention.

Alzheimer disease (AD) is the most common form of dementia among the elderly, with a prevalence of approximately 5% in people older than 65 years. Despite increased knowledge of the pathophysiologic processes underlying AD, the diagnosis relies largely on subjective clinical judgments (behavioral or cognitive).^{1,2} The estimated accuracy of using this approach is approximately 85%³ compared with histologic data, which are considered the reference standard. Earlier studies⁴⁻⁶ showed that structural changes in the medial temporal lobe (hippocampus, parahippocampal gyrus, entorhinal cortex) might serve as indirect markers for AD. Further studies have shown that atrophy of the hippocampal formation is a pathologic hallmark of AD, with a diagnostic accuracy of more than 90%.⁷⁻⁹

No fast and robust method for direct hippocampal volumetric quantification is available, and structural changes may not be detected reliably by visual inspection alone until the advanced stages of the disease. Therefore, structural imaging techniques using whole-brain volumetric quantification and voxel compression subtraction were developed to quantify volumetric changes in the brain. Current segmentation methods are either time consuming or unreliable. Direct hippocampal segmentation is still a challenging task, in that parts

of its border are poorly differentiated in typical images.¹⁰ Even highly advanced segmentation methods based on deformable shape models and atlases are reported to require at least 20 to 30 minutes per dataset for manual landmark placement.^{10,11} Therefore, manual sampling of medial temporal lobe structures still constitutes the “gold standard.”¹²

Because the temporal horns of the cerebral ventricles are adjacent to the hippocampal formation, we proposed the temporal horn volume (THV) and the temporal horn index (THI = THV/lateral ventricular volume) as indirect and sensitive regional measures for hippocampal and parahippocampal atrophy in a preliminary report.¹³ This article puts special emphasis on the validation of the THV and THI as surrogate markers of volume loss in the hippocampal formation. These measurements were correlated with the neuropsychologic outcome of 31 patients and 21 control subjects.

Methods

Subjects

Subjects were recruited from the section for geriatric psychiatry at the University of Heidelberg as well as from participants of a population-based, interdisciplinary, longitudinal study of aging in the Heidelberg area. The clinical evaluation of all subjects included ascertainment of personal and family history as well as physical, neurologic, and neuropsychologic examination. Those with a history of ischemic heart disease, cancer, and cerebrovascular risk factors (hypertension and diabetes) were excluded. Neuropsychologic performance was assessed with the use of an extensive test battery that has been described in detail.¹⁴ In brief, neuropsychologic assessment comprised attention and concentration, primary and secondary memory, speed of processing, visuospatial orientation, and word fluency. To compare

Received September 7, 2005; accepted after revision November 3.

From the German Cancer Research Center (F.L.G., H.v.T.-K., M.E.), Heidelberg, Germany; MeVis—Center for Medical Diagnostic System and Visualization (H.K.H., H.O.P.), Bremen, Germany; Section for Geriatric Psychiatry (P.A.T., J.S.), University of Heidelberg, Germany; Section of Academic Radiology (F.L.G., E.Wid., E.Wig., P.D.G.), University of Sheffield, United Kingdom; and Department of Psychiatry (J.P.), University of Frankfurt, Germany.

Address correspondence to Frederik Lars Giesel, MD, Deutsches Krebsforschungszentrum (DKFZ), Department of Radiology, INF 280, 69120 Heidelberg, Germany; e-mail: f.giesel@dkfz.de

the degree of cognitive impairment among the 3 study groups the global deterioration scale (GDS)¹⁵ was performed on each subject.

Fifty-two subjects, including 21 control subjects, 21 patients with mild cognitive impairment (MCI), and 10 patients with AD were recruited in this study. The mean age of the control subjects was 66.5 years (range, 65–67). Mild cognitive impairment was defined by the Levy criteria of aging-associated cognitive decline.¹⁶ The mean age of those with MCI was 66.6 years (range, 65–67). Mild to moderate AD was defined by the National Institute of Neurological and Communication Disorders and Stroke-Alzheimer Disease and Related Disorders Association criteria.¹ The mean age of those with AD was 65.1 years (range, 54–74). Evaluation of all patient data was performed under a research protocol reviewed and approved by the Institutional Review Board of the University of Heidelberg.

Imaging and Data Analysis

All examinations were performed on a 1.5T clinical MR scanner (Magnetom Vision; Siemens Medical Solutions, Erlangen, Germany). A T1-weighted coronal high-resolution volume acquisition was performed (repetition time [TR], 4 ms; flip angle, 13°; field of view [FOV], 250 mm; matrix, 256 × 256; 128 contiguous sections of 1.8 mm) and supplemented with axial T2-axial weighted and fluid-attenuated inversion recovery (FLAIR) images. The acquired images were anonymized by the removal of DICOM (standard for Digital Imaging and Communication in Medicine) header information. The data were assessed by a blinded reader.

The T1-weighted volumetric datasets were analyzed with a combination of fast 3D marker based segmentation and histogram analysis. All of the image analysis methods have been integrated as modules into the research and development platform MeVisLab (MeVis, Bremen, Germany) dedicated to rapid application prototyping in medical image analysis and visualization.¹⁷

Image Segmentation and Temporal Horn Volume Estimation

Segmentation of lateral ventricles including frontal horn, trigone, posterior horn, and temporal horn (TH) was performed applying an interactive watershed transform (IWT)¹⁸ to the original images. This requires 2 to 6 anatomic landmarks, depending on the image quality and width of the temporal horns, to segment the ventricles and was performed twice for every dataset, once with and once without the temporal horns, resulting in 2 volume estimates for each side, which were named LVV (lateral ventricular volume) and LVV*, respectively. The temporal horn index was defined as $THI = THV/LVV$, whereas temporal horn volumes were calculated by subtraction $THV = LVV - LVV^*$. Because of the thin shape and small volume of temporal horns, we found this subtraction more robust to partial volume averaging than a direct TH segmentation.¹⁹ The posterior TH limit was defined by a plane through 3 distinct anatomic landmarks, which were interactively identified on 3 synchronized orthogonal viewers (Fig 1). The landmarks were the anterior borders of the left and right trigone between frontal and temporal horn and the posterior tip of the inferior colliculus.

A semiautomated analysis of regional histograms was applied to both segmentations (each containing 2 objects, left and right lateral ventricle) to robustly estimate LVV(L/R) and LVV*(L/R).¹⁷ The analysis was based on a trimodal Gaussian model (CSF, white matter, and gray matter) that explicitly included partial volume terms, so-called Mixed Gaussians. Assuming uniformly distributed partial volume effects, Mixed Gaussians take the form of plateau curves. The model is

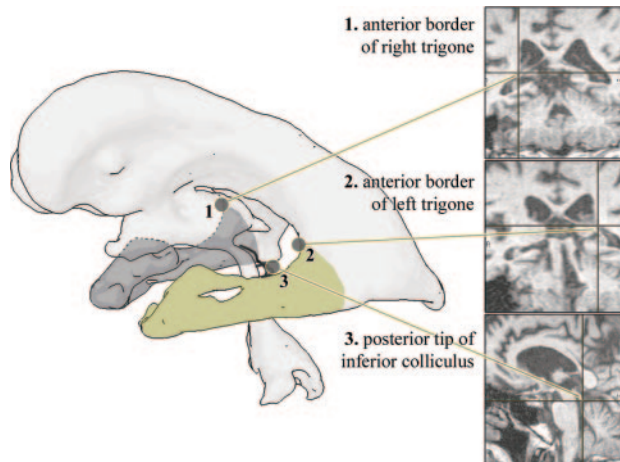


Fig 1. Posterior limit of temporal horns defined by a cut plane through 3 distinct anatomic landmarks that can be defined reliably and independently of patient positioning or head rotation, respectively.

fitted to the histogram data by minimizing squared errors. The reproducibility of this method has previously been evaluated on healthy volunteers.¹⁹ Mean interexamination SD values were as low as 0.07 mL for THV and 0.56% for THI.

Statistics

The age of the 3 different groups was compared to explore potential age differences between groups by using Mann-Whitney *U* test. The right and left LVV, THV, and THI values of patients with AD, subjects with MCI, and control subjects were compared with the use of analysis of variance, followed by Mann-Whitney *U* test. Pearson correlation was calculated between the morphometric values and the neuropsychologic measures. Statistical analysis was performed using SPSS version 11 (SPSS, Chicago, Ill).

Results

In all 52 cases, the ventricles were successfully quantified and visualized. A representation of the ventricles from each group is shown in Fig 2. The 3 groups did not differ significantly in age and sex ($P > .05$). The values for right and left LVV, THV, and THI of the control subjects, subjects with MCI, and patients with AD are summarized in Table 1. There was no significant difference between right, left, and total LVV, THV, and THI of subjects with MCI and control subjects ($P > .05$). The right, left, and total LVV, THV, and THI of patients with AD were significantly larger compared with control subjects ($P < .05$) and with subjects with MCI ($P < .05$).

Comparison of the GDS scores as measures for global cognitive functioning revealed significant differences between healthy control subjects (mean, 1.29; SD 0.46), subjects with MCI (mean, 1.71; SD 0.46) and patients with AD (mean, 3.90; SD 0.88). Moreover, the scores achieved in the reproduction item of the selective reminding test differed significantly among the groups (control subjects: mean, 7.06, SD 0.68; subjects with MCI: mean, 6.50, SD 0.69; and patients with AD: mean, 3.86, SD 1.11). Statistical analysis of a putative correlation between morphometric values and neuropsychologic measures showed a significant negative correlation between the THI and the declarative reminding test ($P < .01$) as well as between THV and the declarative reminding test ($P < .01$)

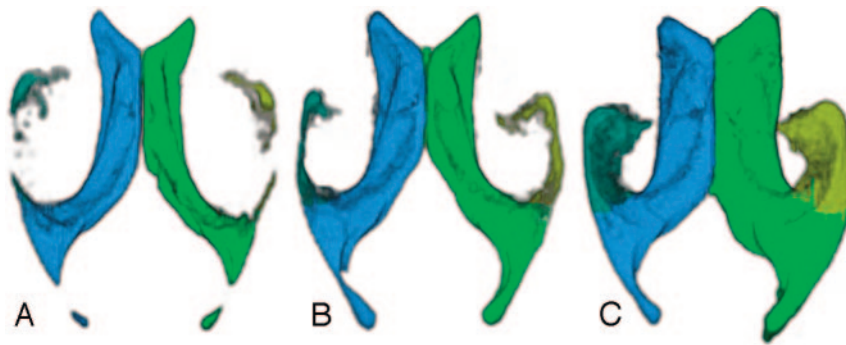


Fig 2. 3D appearance of the ventricles in a control subject (A), subject with MCI (B), and a patient with AD (C). The temporal horn of the lateral ventricle is larger in the patient with AD compared with the control subject or patient with MCI.

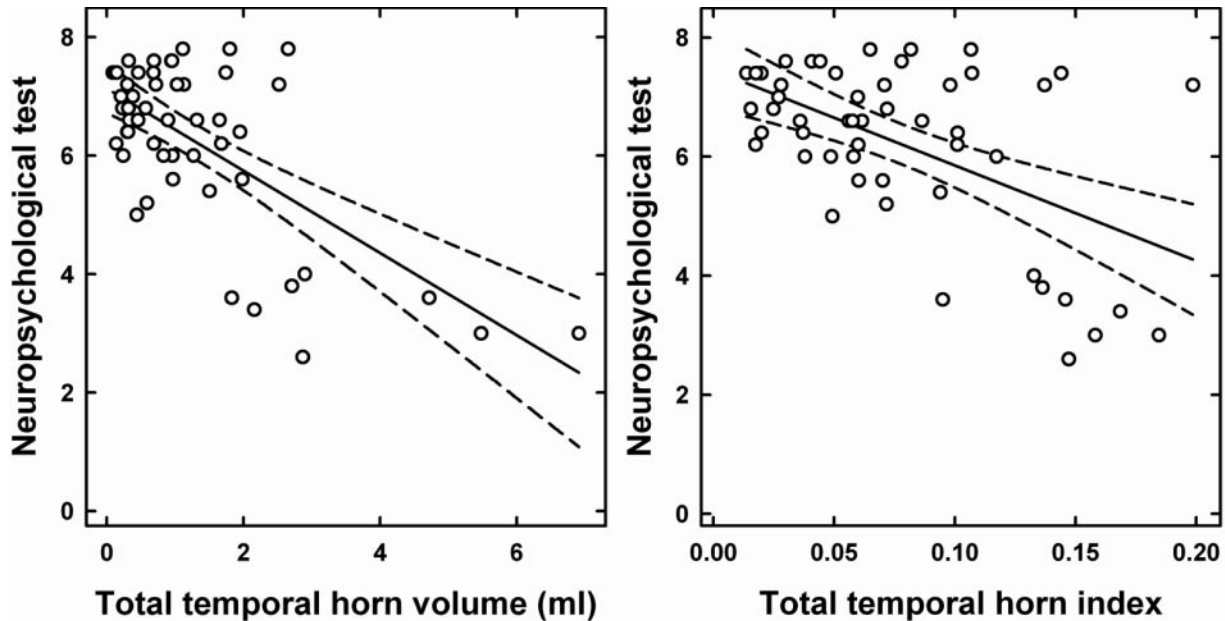


Fig 3. Significant negative correlation between the THV/THI and neuropsychologic test ($P < .01$). $\text{THI} = \text{Neuropsychologic test} = 7.45 \pm 15.75 \times \text{tthi}$ $R^2 = 0.29$; $\text{THV} = \text{Neuropsychologic test} = 7.12 \pm 0.67 \times \text{tthv}$ $R^2 = 0.44$.

(Fig 3). In addition, an individual analysis of the AD subgroup showed a significant correlation of THV and THI and the declarative reminding test [correlation coefficient $\kappa = 0.65$; $P < .05$].

Discussion

Current consensus in the scientific and clinical community has emphasized the need for early detection and the need to develop sensitive objective markers that may serve as adjuncts to current clinical and neuropsychologic tests. Such objective markers of AD should detect the disease at its early stages and could also be used to monitor therapeutic intervention. Direct and focused imaging assessment of regional atrophy of medial temporal lobe (MTL) seem most promising and enable the discrimination of AD from healthy control subjects with the use of MR imaging.^{20,21}

A variety of measurement-based techniques have been used to try to discriminate between patients with AD from those with other disorders, such as vascular dementia or frontotemporal dementia. Manual/semiautomated segmentation of the hippocampus takes approximately 20–30 minutes, depending on user experience.²² Jack et al²³ have shown that the rate of temporal horn enlargement is similar to that of hippocampal atrophy, indicating that the temporal horn mea-

surements could be used in place of hippocampal measurements. Earlier studies attempted to define sentinel changes that will allow the use of linear measurements of the temporal horn to support clinical decisionmaking.^{24,25} These studies have yielded variable results, with sensitivities ranging from 33% to 93% and specificity of approximately 95%.^{24,26-29} These differences are likely to reflect different methodologic approaches to assess medial temporal lobe; this emphasizes the necessity of a reliable and economic approach to quantify brain atrophy. Most previous investigators used manual tracing techniques, however, which are time-consuming and rater-dependent.¹² In contrast, we describe a fast and robust ventricular quantification method of evaluating the THV. This technique takes 3 minutes to evaluate the THV and therefore lends itself well to application in busy clinical settings.

As shown in a previous study,¹⁷ cerebral ventricular segmentation and volumetry based on the IWT and histogram analysis are largely independent of the landmark positions. The robustness and reproducibility of THV and THI has been evaluated by our group.¹⁷ We found that both the intraobserver and interobserver variations are very small.¹⁸ Measuring THV by subtraction rather than directly has a 2-fold motivation. First, the reproducibility is better for THV than for LVV, reflecting the fact that variations of LVV and LVV* are

Right, left, and total left LVV, THV, and THI of the controls and MCI and AD patients

	Controls (SD)	MCI (SD)	AD (SD)
Mean left LVV (mL)	14.42 (7.02)	12.78 (5.36)	25.64 (16.08)
Mean right LVV (mL)	13.24 (5.61)	11.38 (5.11)	22.80 (9.47)
Mean total LVV (mL)	27.66 (12.47)	24.17 (10.07)	48.44 (24.25)
Mean left THV (mL)	0.46 (0.36)	0.40 (0.36)	1.65 (0.84)
Mean right THV (mL)	0.43 (0.36)	0.46 (0.36)	1.74 (1.22)
Mean total THV (mL)	0.89 (0.68)	0.86 (0.68)	3.39 (1.80)
Mean left THI	0.03 (0.02)	0.03 (0.02)	0.07 (0.02)
Mean right THI	0.03 (0.02)	0.04 (0.02)	0.07 (0.03)
Mean total THI	0.06 (0.03)	0.07 (0.04)	0.14 (0.04)

Note:—LVV indicates lateral ventricular volume; THV, temporal horn volume; THI, temporal horn index (THV/LVV ratio); MCI, mildly cognitive impaired; AD, Alzheimer disease.

positively correlated. Second, regional histograms of solely the tiny temporal horns are extremely sparse. In comparison, the histograms for the 2 larger objects yield a higher robustness and reproducibility. From the user perspective, however, directly segmenting the temporal horns or excluding them from a previous lateral ventricle segmentation is equivalent.

Even though it is not fully automated, our method has shown an excellent intraobserver and interobserver as well as interexamination reproducibility.^{15,19} It is important to note that even fully automated image analysis methods that are perfectly reproducible on a given dataset by definition may fail to provide reproducible measurements on repeated examinations. The only critical interactive steps in our method are landmark definitions that an experienced user can perform rapidly and reliably, whereas an automated method might fail in case of a noisy image or heavily deformed anatomy. We thus conclude that our method combines interactive with automated processing steps in a sensible fashion to maximize the precision and robustness of the derived results.

In a study of 192 subjects with probable AD who underwent 2 MR imaging examinations at an interval of 1 year, correlations between image-based volumetric change and change in behavioral and cognitive measures were found to be even greater for the temporal horn than for the hippocampus.³⁰ This further supports the suitability of THV/THI as a reliable biomarker in AD. Jack et al³⁰ demonstrated the technical feasibility of using structural MR imaging measures as a surrogate end point of disease progression in therapeutic trials, resulting in markedly lower estimated sample size requirements for clinical trials. Our method of evaluating THV and THI could be used in therapeutic trials to monitor disease progression because it is semiautomated and quick to perform.

The definition of MCI is a subject of controversy. Some believe that MCI represents a very mild form of AD and that virtually all patients with MCI may have the pathologic changes of AD.³¹ Others use the term to refer to a more heterogeneous group of patients, with some proportion of subjects remaining stable or even “reverting” back to normal and others progressing to develop additional deficits and meeting the criteria for AD. The criteria for identification and classification of subjects with MCI are evolving. Although certain criteria have been drawn up to help identify those with MCI,³² there remains controversy on which memory test(s) to use, what cutoffs to apply to determine impairment, and how to

deal with fluctuations in performance. Because of this diverse definition of MCI, reports on the association of MCI with medial temporal lobe atrophy have varied.^{19,33–36} In this study, we have found no significant difference in the THV and THI between subjects with MCI and control subjects. However, we have found a significant correlation between the THI and THV and memory performance. This may reflect the heterogeneity of the patient population within the MCI group.

Conclusion

We report the use of a semiautomated technique in assessing THV and THI in subjects with MCI and patients with AD. In this study, we found that the THV and THI of patients with AD were significantly larger than that of control subjects or subjects with MCI, and there was no significant difference in the THV and THI of subjects with MCI and control subjects. This method of assessing for changes in THV and THI is rapid, objective, and independent of user experience and is therefore suitable for application in the clinical setting as well as in therapeutic trials.

References

- McKhann G, Drachman D, Folstein M, et al. **Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease.** *Neurology* 1984;34:939–44
- Kazee AM, Eskin TA, Lapham LW, et al. **Clinicopathologic correlates in Alzheimer disease: assessment of clinical and pathologic diagnostic criteria.** *Alzheimer Dis Assoc Disord* 1993;7:152–64
- Tierney MC, Fisher RH, Lewis AJ, et al. **The NINCDS-ADRDA Work Group criteria for the clinical diagnosis of probable Alzheimer's disease: a clinicopathologic study of 57 cases.** *Neurology* 1988;38:359–64
- Seab JP, Jagust WJ, Wong ST, et al. **Quantitative NMR measurements of hippocampal atrophy in Alzheimer's disease.** *Magn Reson Med* 1988;8:200–08
- Kesslak JP, Nalcioglu O, Cotman CW. **Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease.** *Neurology* 1991;41:51–54
- Jack CR Jr, Petersen RC, O'Brien PC, et al. **MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease.** *Neurology* 1992;42:183–88
- de Leon MJ, George AE, Golomb J, et al. **Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease.** *Neurobiol Aging* 1997;18:1–11
- Nagy Z, Hindley NJ, Braak H, et al. **The progression of Alzheimer's disease from limbic regions to the neocortex: clinical, radiological and pathological relationships.** *Dement Geriatr Cogn Disord* 1999;10:115–20
- Smith AD, Jobst KA. **Use of structural imaging to study the progression of Alzheimer's disease.** *Br Med Bull* 1996;52:575–86
- Shen D, Moffat S, Resnick SM, et al. **Measuring size and shape of the hippocampus in MR images using a deformable shape model.** *Neuroimage* 2002;15:422–34
- Hogan RE, Mark KE, Wang L, et al. **Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients.** *Radiology* 2000;216:291–97
- Geuze E, Vermetten E, Bremner JD. **MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed.** *Mol Psychiatry* 2005;10:147–59
- Hahn HK, Rexilius J, Schlüter M. **Fast and robust quantification of parahippocampal atrophy via temporal horn index.** In: Tolxdorff T, Braun J, Handels H, et al, eds. *Bildverarbeitung für die Medizin 2004: Algorithmen—Systeme—Anwendungen.* Berlin: Springer-Verlag; 2004:371–75. Available at: <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-116/p371.pdf>
- Schroder J, Kratz B, Pantel J, et al. **Prevalence of mild cognitive impairment in an elderly community sample.** *J Neural Transm Suppl* 1998;54:51–59
- Reisberg B, Ferris SH, de Leon MJ, et al. **The Global Deterioration Scale for assessment of primary degenerative dementia.** *Am J Psychiatry* 1982;139:1136–39
- Levy R. **Aging-associated cognitive decline.** Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization. *Int Psychogeriatr* 1994;6:63–68
- Hahn HK, Millar WS, Klinghammer O, et al. **A reliable and efficient method for cerebral ventricular volumetry in pediatric neuroimaging.** *Methods Inf Med* 2004;43:376–82

18. Hahn HK, Peitgen HO. IWT—Interactive watershed transform: a hierarchical method for efficient interactive and automated segmentation of multidimensional gray-scale images. *Proc SPIE* 2003;5032:643–53.
19. Parnetti L, Lowenthal DT, Presciutti O, et al. 1H-MRS, MRI-based hippocampal volumetry, and 99mTc-HMPAO-SPECT in normal aging, age-associated memory impairment, and probable Alzheimer's disease. *J Am Geriatr Soc* 1996;44:133–38
20. Juottonen K, Laakso MP, Partanen K, et al. Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease. *AJNR Am J Neuroradiol* 1999;20:139–44
21. 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;49:786–94
22. Thacker NA, Varma AR, Bathgate D, et al. Dementing disorders: volumetric measurement of cerebrospinal fluid to distinguish normal from pathologic findings—feasibility study. *Radiology* 2002;224:278–85
23. Jack CR Jr, Petersen RC, Xu Y, et al. Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology* 1998;51:993–99
24. Frisoni GB, Beltramello A, Weiss C, et al. Linear measures of atrophy in mild Alzheimer disease. *AJNR Am J Neuroradiol* 1996;17:913–23
25. Dahlbeck JW, McCluney KW, Yeakley JW, et al. The interuncal distance: a new MR measurement for the hippocampal atrophy of Alzheimer disease. *AJNR Am J Neuroradiol* 1991;12:931–32
26. Soininen H, Puranen M, Riekkinen PJ. Computed tomography findings in senile dementia and normal aging. *J Neurol Neurosurg Psychiatry* 1982;45:50–54
27. Scheltens P, Leys D, Barkhof F, et al. Atrophy of medial temporal lobes on MRI in “probable” Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* 1992;55:967–72
28. Erkinjuntti T, Lee DH, Gao F, et al. Temporal lobe atrophy on magnetic resonance imaging in the diagnosis of early Alzheimer's disease. *Arch Neurol* 1993;50:305–10
29. Frisoni GB, Geroldi C, Beltramello A, et al. Radial width of the temporal horn: a sensitive measure in Alzheimer disease. *AJNR Am J Neuroradiol* 2002;23:35–47
30. Jack CR Jr, Slomkowski M, Gracon S, et al. MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. *Neurology* 2003;60:253–60
31. Morris JC, Storandt M, McKeel DW Jr, et al. Cerebral amyloid deposition and diffuse plaques in “normal” aging: Evidence for presymptomatic and very mild Alzheimer's disease. *Neurology* 1996;46:707–19
32. Petrella JR, Coleman RE, Doraiswamy PM. Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. *Radiology* 2003;226:315–36
33. Pantel J, Kratz B, Essig M, et al. Parahippocampal volume deficits in subjects with aging-associated cognitive decline. *Am J Psychiatry* 2003;160:379–82
34. Soininen HS, Partanen K, Pitkanen A, et al. Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: correlation to visual and verbal memory. *Neurology* 1994;44:1660–68
35. Laakso MP, Partanen K, Lehtovirta M, et al. MRI of amygdala fails to diagnose early Alzheimer's disease. *Neuroreport* 1995;6:2414–18
36. Laakso MP, Soininen H, Partanen K, et al. MRI of the hippocampus in Alzheimer's disease: sensitivity, specificity, and analysis of the incorrectly classified subjects. *Neurobiol Aging* 1998;19:23–31