A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline

Tom den Heijer,1,2,* Fedde van der Lijn,4,5,* Peter J. Koudstaal,1 Albert Hofman,2 Aad van der Lugt,4 Gabriel P. Krestin,4 Wiro J. Niessen4,5,6 and Monique M. B. Breteler2

1 Department of Neurology, Erasmus MC, Rotterdam, The Netherlands
2 Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands
3 Department of Neurology, Sint Franciscus Gasthuis, Rotterdam, The Netherlands
4 Department of Radiology, Erasmus MC, Rotterdam, The Netherlands
5 Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
6 Imaging Science and Technology, Faculty of Applied Sciences, Delft University, Delft, The Netherlands
7 Department of Forensic Molecular Biology, Erasmus MC, Rotterdam, The Netherlands

*These authors contributed equally to this work.

Correspondence to: Dr Tom den Heijer,
Department of Neurology,
Sint Franciscus Gasthuis,
PO Box 10900,
3004 BA, Rotterdam, The Netherlands
E-mail: t.denheijer@sfg.nl

Hippocampal atrophy is frequently observed on magnetic resonance images from patients with Alzheimer’s disease and persons with mild cognitive impairment. Even in asymptomatic elderly, a small hippocampal volume on magnetic resonance imaging is a risk factor for developing Alzheimer’s disease. However, not everyone with a small hippocampus develops dementia. With the increased interest in the use of sequential magnetic resonance images as potential surrogate biomarkers of the disease process, it has also been shown that the rate of hippocampal atrophy is higher in persons with Alzheimer’s disease compared to those with mild cognitive impairment and the healthy elderly. Whether a higher rate of hippocampal atrophy also predicts Alzheimer’s disease or subtle cognitive decline in non-demented elderly is unknown. We examine these associations in a group of 518 elderly (age 60–90 years, 50% female), taken from the population-based Rotterdam Scan Study. A magnetic resonance imaging examination was performed at baseline in 1995–96 that was repeated in 1999–2000 (in 244 persons) and in 2006 (in 185 persons). Using automated segmentation procedures, we assessed hippocampal volumes on all magnetic resonance imaging scans. All persons were free of dementia at baseline and followed over time for cognitive decline and incident dementia. Persons had four repeated neuropsychological tests at the research centre over a 10-year period. We also continuously monitored the medical records of all 518 participants for incident dementia. During a total follow-up of 4360 person-years, (mean 8.4, range 0.1–11.3), 50 people developed incident dementia (36 had Alzheimer’s disease). We found an increased risk to develop incident dementia per standard deviation faster rate of decline in hippocampal volume [left hippocampus 1.6 (95% confidence interval 1.2–2.3, right hippocampus 1.6 (95% confidence interval 1.2–2.1)]. Furthermore, decline in hippocampal volume predicted onset of clinical dementia when corrected for baseline hippocampal volume. In people who remained free of dementia during the whole follow-up period, we found that decline in hippocampal volume paralleled, and preceded, specific decline in delayed word recall. No associations were found in this sample between rate of hippocampal atrophy, Mini Mental...
Introduction

One of the major challenges for research into Alzheimer’s disease is to identify people in the earliest phase of the disease so that they may enter clinical trials (Sonnen et al., 2008). Persons with so-called mild cognitive impairment (Petersen et al., 1999) have subjective complaints and a high conversion risk to develop clinical Alzheimer’s disease (Palmer et al., 2008b). However, half of the patients with Alzheimer’s disease have never reported subjective memory complaints before diagnosis and would therefore not come to the attention of medical care or specialized memory clinics (Palmer et al., 2008a). Therefore, objective biomarkers for the onset of the disease process, irrespective of complaints or cognitive symptoms, are necessary. Moreover, association studies of early biomarkers and genetic and environmental factors could give insight in the pathogenesis and aetiology of the disease. With the increased use of MRI-based biomarkers, attention has focused on the medial temporal lobe, as this region is clearly affected by specific Alzheimer neuropathology early on in the disease course (Braak and Braak, 1997). A large number of studies found smaller hippocampal volumes on MRI in patients with Alzheimer’s disease or mild cognitive impairment, compared with healthy controls (Jack et al., 1992, 1997; Convit et al., 1995; Fox et al., 1996). We have previously shown that even in elderly without cognitive symptoms or complaints, a small hippocampal volume on MRI predicts Alzheimer’s disease (den Heijer et al., 2006). However, we also showed that a large portion of people with smaller hippocampal volumes on MRI do not develop dementia. To improve prediction of Alzheimer’s disease or cognitive decline further, follow-up brain imaging may distinguish persons with a small, yet stable volume, from those with a declining volume due to a neurodegenerative process. Longitudinal MRI scanning of the hippocampus has been performed previously in a few studies. In a set of young patients with familial Alzheimer’s disease, hippocampal volume change was found to be an earlier and better predictor compared with a single volume measurement (Ridha et al., 2006). In the elderly, rates of hippocampal atrophy on MRI were found to be higher in cases with Alzheimer’s disease and mild cognitive impairment compared with controls (Du et al., 2004; Jack et al., 2004). However, in another follow-up study of 3 years among 27 elderly patients with Alzheimer’s disease, longitudinal MRI hippocampal data did not improve diagnostic accuracy over a single volume measurement (Laakso et al., 2000). In the current study, we examine whether decline in hippocampal volume on MRI is associated with cognitive decline and incident clinical dementia, as measured in a large population study (the Rotterdam Scan Study) during a 10-year follow-up.

Materials and methods

Setting and participants

The Rotterdam Scan Study is a large population-based cohort study among non-demented elderly in the Netherlands with baseline examinations from 1995 to 1996 (Breteler, 2000). For details on selection criteria and differences between participants and non-participants, we refer to de Leeuw et al. (1999). In 1995–1996, 518 non-demented elderly (age 60–90, 50% female) underwent, among other examinations, 3D brain MRI scanning and cognitive testing. After these baseline examinations, there were four different examination rounds with cognitive testing and two follow-up brain MRI scans within a time frame of 10 years (most recent examination in 2006). Figure 1 shows the time frame of the study cohort and the number of participants at each examination round. All participants gave written informed consent after complete description of the study. The medical ethics committee of Erasmus MC approved the study protocol.

Magnetic resonance imaging sequences

At examinations in 1995–96, in 1999–2000 and in 2006, the whole brain was imaged using a 1.5 T MRI unit. The MRI sequence of the first two MRI examinations was a custom made 3D Half-Fourier Acquisition Single-Shot Turbo Spin Echo (HASTE) sequence (inversion time 440 ms, repetition time 2800 ms, 128 contiguous sagittal slices of 1.25 mm, acquisition matrix 192 × 256, field of view 256 × 256 mm). Two HASTE modules were sequentially acquired after the inversion pulse (effective echo time of 29 ms and 440 ms) of which the first was used for volumetric assessments of the hippocampus (den Heijer et al., 2003). Due to the availability of newer MRI techniques and a new MR scanner, the third examination (in 2006), was performed with a 3D T1-weighted sequence (inversion time 400 ms, repetition time 14.8 ms, time to echo 2.8 ms, 96 axial slices of 1.6 mm interpolated to 192 slices of 0.8 mm, acquisition matrix 416 × 256, field of view 250 × 250 mm). At baseline, we performed manual segmentations of the hippocampus (den Heijer et al., 2003). However, due to the labour-intensive nature of such measurements and the risk of errors in reproducibility, we developed an automatic method to segment the hippocampus on all three sequential MRI scans.

Automated segmentation of the hippocampus

The hippocampus was segmented using an automated method described previously (van der Lijn et al., 2008). The two most important components of this method are a statistical intensity model and a spatial probability map. The intensity model describes the typical intensities of the hippocampus and the background. The spatial probability

Keywords: hippocampus; Alzheimer; MRI; memory; population

Abbreviations: HASTE = half-fourier acquisition single-shot turbo spin echo
The T1-weighted images were segmented with the intention for all three time points are shown in Fig. 2. The HASTE images were segmented according to the baseline hippocampus assessments. Example results of the automated segmentations were visually inspected by a single rater (TdH) who forming its spatial map. The results of all automated hippocampal examinations. The spatial probability map was again created by registering the HASTE image and hipppocampus. In these instances, the segmentation was manually corrected using non-rigidly registering the same 20 labelled images to the unlabelled target image, deforming the manual segmentations and averaging them. The first follow-up was acquired with the same MRI sequence as the baseline examination. Consequently, the same intensity model could be used for these images. For the scans at first follow-up, the spatial probability map was obtained by first non-rigidly registering the baseline to the follow-up image and subsequently deforming the baseline probability map. For the second follow-up the baseline intensity model could not be used, since those images were obtained with a T1-weighted sequence. Therefore, we created a new training set by manually segmenting the hippocampus in 18 examinations that were acquired with the same scanner and sequence as the second follow-up examinations. The spatial probability map was again created by registering the baseline image to the images at second follow-up and deforming its spatial map. The results of all automated hippocampal segmentations were visually inspected by a single rater (TdH) who was blinded for cognitive status. In a number of cases the segmented area included the entorhinal cortex or extended to the collateral sulcus. In these instances, the segmentation was manually corrected using FSLView (http://www.fmrib.ox.ac.uk/fsl/) and volumes were recalculated. Manual correction was necessary in 212 (11%) of all 1984 hippocampus assessments. Example results of the automated segmentation for all three time points are shown in Fig. 2.

To estimate the effect of the different MRI sequences on the volume measurements, eight elderly subjects were scanned with the 3D HASTE sequence and 3D T1-weighted sequence within a 3-month period. The HASTE images were segmented according to the baseline procedure. The T1-weighted images were segmented with the intensity model of the second follow-up. For these latter images, the probability map was created by registering the HASTE image and deforming its probability map. The hippocampal volumes derived from both MRI sequences within this short follow-up of 3 months should be approximately identical. There was indeed a strong correlation between the two hippocampal volume measurements [Pearson r = 0.97 (P < 0.001)]. However, the mean total hippocampus volume measured in the HASTE images was 6.13 ± 0.98 ml versus 5.28 ± 0.98 ml in the T1 images. Because of this systematic under-segmentation in the T1-weighted images, we could not infer absolute volume decline in hippocampal volume over the 10-year follow-up. Ranking of subjects according to rate of hippocampal decline was however possible. Therefore, we transformed all hippocampal volumes to Z-scores (individual volume–population mean/standard deviation) at each time point. By definition, the average Z-scores at each point time were zero with a standard deviation of one. If a person declined in hippocampal volume more rapidly than its counterpart he or she would decline in Z-score over time.

### Decline in hippocampal volume

The decline in hippocampal volume was modelled using a linear random-effects model. This approach uses all available hippocampal data, and accounts for within-person correlation over time, which results in increased statistical power for estimating effects (Diggle et al., 1994). We used PROC Mixed models (Statistical SAS 9.1, PROC MIXED; SAS Institute, Cary, NC, USA) to model hippocampal volume decline. Using the population with at least one repeated hippocampal volume measurement, we used hippocampal volumes at baseline and follow-up as outcome variable, and follow-up time from baseline as independent variable. The estimated fixed effect and the individual random effects were added to obtain estimated slopes of the individual Z-score declines in hippocampal volume.

### Incident dementia

None of the 518 participants of the cohort had dementia at baseline. We followed the cohort for incident dementia with a strict protocol (Ruitenberg et al., 2001; Vermeer et al., 2003; den Heijer et al., 2006). Briefly, participants were cognitively screened at follow-up visits (1997–99, 1999–2000, 2004–05, 2006) with the Mini Mental State Examination and the Geriatric Mental Schedule. When screened positive, they were assessed with the Cambridge Examination for Mental Disorders of the Elderly interview. Participants who were then thought to have dementia were examined by a neurologist and underwent additional neuropsychological testing by a neuropsychologist. The number of participants that could be examined in person at

### Figure 1

Time frame of the study cohort (n = 518 at baseline in 1995–96). Outlined in dark boxes are examination rounds which included MRI scanning. Number of participants at those rounds are given for the MRI part and neuropsychological examinations. At each examination round, in dashed boxes are given the number of participants that had died before we could invite them for participation (cumulative over time), and the number that refused to visit the research centre, or had a contraindication (CI) for MRI scanning.
the follow-up visits was 436 in 1997–99, 366 in 1999–2000, 287 in 2004–05 and 192 in 2006. To avoid missing incident dementia cases among those people who did not come to the research centre, we also continuously monitored the medical records of all 518 participants at the general practitioners’ offices to obtain information on diagnosed dementia and other major morbidity or death. This was done as part of the strict follow-up protocol of the Rotterdam Study (Ott et al., 1998). Research assistants regularly checked the medical records of all participants, and medical discharge letters of memory clinics, neuropsychologists, and the Regional Institute for Ambulatory Mental Health Care were given to the research physician. Follow-up of medical record information until January 1, 2006 was complete for >99% of the cohort. A diagnosis of dementia and subtype was made by a panel that consisted of a neurologist, neuropsychologist and research physician with the use of standard international criteria (McKhann et al., 1984; Román et al., 1993). The onset of dementia was defined as the date on which clinical symptoms allowed the diagnosis of dementia. Duration of follow-up for each participant was calculated from baseline examination until death, diagnosis of dementia or the end of follow-up, whichever came first.

Cognitive decline

In addition to the short cognitive screening for dementia with Mini Mental State Examination and Geriatric Mental Schedule, persons underwent extensive neuropsychological testing (Prins et al., 2005). In short, we assessed memory function by means of a 15-word verbal learning test. The sum of words recalled at three trials was used to define immediate recall. After having done other cognitive tests in 15 min, a delayed recall phase was introduced. We assessed executive function with the Stroop test (part three interference), and the Letter–Digit Substitution Task. All tests were done at the research centre during the follow-up examination rounds, except in 1997–99 at which time only Mini Mental State Examination screening was done. For each person, we therefore had a maximum of five Mini Mental State Examination scores (including baseline) and four other neuropsychological test scores over the 10-year follow-up period. To determine cognitive decline in any of these tests, we used a linear random-effects model similar to that used for hippocampal volume decline. After excluding persons with incident dementia (as they could have extreme cognitive tests results or unreliable data), we used PROC Mixed models (Statistical SAS 9.1, PROC MIXED; SAS Institute, Cary, NC, USA) to model cognitive decline. Cognitive decliners in each separate neuropsychological test were defined as having an individual rate of decline one standard deviation faster than the average cognitive decline.

Confounders

The following variables were used as potential confounders: age at baseline, sex, educational level, total grey matter volume and white matter lesion volume. The latter two volumes were measured on the 3D MRI scans using an automated segmentation procedure volume (Ikram et al., 2008). Hippocampal volumes were calculated as percentages of the total intracranial volume (in millilitres). This was measured using an automated method on the baseline MRI (Ikram et al., 2008).

Data analysis

We used Cox’s proportional hazards models to quantify the association between hippocampal volume (baseline and decline) and risk of incident dementia. Hazard ratios of dementia were calculated per standard deviation decrease of baseline Z-score or standard deviation of Z-score of hippocampal decline. For the analyses on decline in hippocampal volume, we used only incident dementia cases with a baseline and a first follow-up scan before the clinical diagnosis was made. First, we investigated overall dementia then separately incident Alzheimer dementia. ANCOVA was used to compare the means of MRI and cognitive variables between persons with and without dementia. We used logistic regression to quantify the association between hippocampal volume (baseline and decline) and cognitive decline. To assess whether hippocampal decline could precede cognitive decline we also separately investigated the decline in hippocampal volume from baseline to first follow-up with cognitive decline after the first follow-up MRI as dependent variable. Finally, we used ANCOVA to compare the means of all MRI volumes and individual cognitive scores between the cognitive decliners and non-decliners.

Results

Baseline characteristics and characteristics at time of follow-up MRI scanning are shown in Table 1. From the 518 persons at baseline, 128 had three MRI scans, 173 had two MRI scans and 217 had a baseline MRI scan only. Average time between the first and the last MRI scan was 10.4 years (range 9.7–11.1). Persons who did not have any of the two follow-up MRI examinations after baseline (n=217) were in general older (at baseline 4.5 years, \(P<0.001\)), had a lower Mini Mental State Examination (−0.5 points, \(P<0.001\)), had a lower baseline hippocampal Z-score (−0.22, \(P=0.02\) both for left and right hippocampal volume), but were similar with respect to sex distribution compared with persons who had at least one follow-up MRI.
(n = 501). The absolute decline in left hippocampal volume from the baseline to first follow-up scan was 0.52% per year and in right hippocampal volume 0.51% per year. As expected due to the systematic under-segmentation of the hippocampus on the third MRI scan, the absolute volume decline from the first to second follow-up MRI was higher (1.7% per year for the left hippocampus and 1.6% per year for the right hippocampus).

During a follow-up period of 4360 person-years, (mean 8.4, range 0.1–11.3), 50 persons developed incident dementia (36 had Alzheimer’s disease). Of these 50 persons, 21 were diagnosed based on the in-person screening and 29 were diagnosed based on medical information. Similar to findings previously reported in this dataset, baseline hippocampal volumes were associated with risk of dementia [age, sex and education adjusted HR for standard deviation decrease in left hippocampus 2.3 (95% CI 1.7–3.1) and for the right hippocampus 2.0 (1.5–2.6)]. There was no difference in effect sizes of the association between the hippocampus and Alzheimer’s disease or non-Alzheimer’s dementia (data not shown).

Of the 50 persons with incident dementia, 13 had two MRI scans before their clinical diagnosis. We found that a one SD faster decline in hippocampal volume was associated with a higher risk to develop dementia [age, sex and education adjusted HR for standard deviation faster decline in hippocampal volume from first to second follow-up scan was 1.39 (95% CI 0.96–2.00) for the left hippocampus and 1.55 (95% CI 1.12–2.15)]. After adjustment for baseline hippocampal volumes, the risk associated with decline in hippocampal volume remained. Baseline total grey matter volume was not significantly associated with decline in hippocampal volume (Pearson correlation coefficient −0.07, P = 0.31 for decline in left hippocampal volume and −0.08, P = 0.21 for decline in right hippocampal volume). More extensive white matter lesions at baseline were associated with faster decline in hippocampal volume (Pearson correlation coefficient −0.16, P = 0.02 for decline in left hippocampal volume and −0.16, P = 0.01 for decline in right hippocampal volume). Adjusting the associations between decline in hippocampal volume and risk of dementia for baseline grey matter volume and white matter lesions did not change the relations (data not shown).

Figure 3 shows the Z-scores of the hippocampus relative to the specific clinical diagnosis of Alzheimer’s dementia. The majority of persons with incident Alzheimer’s dementia [24 of the 36 (67%) for the left hippocampus and 27 of the 36 (75%) for the right hippocampus] had a Z-score of hippocampal volume below population average (i.e. <0) years before clinical diagnosis.

Table 2 shows the average absolute Z-scores of hippocampal and grey matter volume, and white matter lesions at baseline and first MRI follow-up in persons with and without incident dementia. The cognitive test scores at baseline and first follow-up of the persons with incident dementia are shown in Table 3.

We then investigated the association between decline in hippocampal volume and cognitive decline in the cohort of persons who remained dementia free (Table 4). There were 414 subjects without incident dementia who had at least one repeated cognitive test after baseline. Of this group, 283 subjects had one or more follow-up hippocampal volume measurements. Those who declined faster in hippocampal volume had a statistically significant faster decline in delayed memory recall. We similarly found that a standard deviation faster decline in hippocampal volume from first to second scan predicted the risk to develop decline in delayed memory after the second scan [odds ratio for left hippocampus 1.39 (95% CI 0.96–2.00) and for the right hippocampus 1.55 (95% CI 1.02–2.35)]. We found a borderline statistically significant association between decline in right hippocampal volume and risk to decline in Letter–Digit Substitution Task [odds ratio 1.34 (95% CI 0.98–1.82), P = 0.06)]. Additional adjustment for baseline total grey matter volume and white matter lesions did not change any of the associations. We repeated the analyses for persons who had all three MRI scans available (n = 128) with Z-scores based on these subjects only. Similar associations were found, though with lower statistical significance, between Z-score hippocampal decline and decline in delayed recall [odds ratio per
standard deviation decline in Z-score 1.54 (95% CI 0.98–2.42, \( P = 0.06 \)) for the left hippocampus and 1.35 (95% CI 0.85–2.14, \( P = 0.19 \)) for the right hippocampus).

When we focused on the cognitive profiles of subjects who had a decline in delayed recall, we found that they performed worse at any time in several other cognitive tests including Mini Mental State Examination (Table 5). They had specifically lower hippocampal volumes, but no total grey matter volume reduction (Table 6). They also had increased white matter lesion load at the second follow-up MRI.

### Discussion

In this large population-based cohort study, we found that decline in hippocampal volume on MRI was associated with an increased risk to develop dementia or cognitive decline, particularly decline in delayed recall.

Before discussing these findings, several limitations of our study need to be addressed. First, although we had a large cohort, few people developed dementia. At baseline, there was a natural selection of relatively healthy elderly willing to undergo MRI and...
### Table 3  Cognitive characteristics of persons with incident dementia versus those without incident dementia

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mini Mental State Examination</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident dementia (n = 241)</td>
<td>28.0 (1.8)</td>
<td>27.8 (1.8)</td>
<td>−0.06 (0.57)</td>
</tr>
<tr>
<td>Incident dementia (n = 13)</td>
<td>26.7 (2.6)†</td>
<td>26.0 (2.8)†</td>
<td>−0.23 (0.71)</td>
</tr>
<tr>
<td><strong>Immediate word recall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident dementia (n = 241)</td>
<td>20.8 (5.3)</td>
<td>24.9 (4.8)</td>
<td>1.2 (1.4)</td>
</tr>
<tr>
<td>Incident dementia (n = 13)</td>
<td>18.7 (4.3)</td>
<td>18.0 (5.8)†</td>
<td>−0.2 (1.6)†</td>
</tr>
<tr>
<td><strong>Delayed word recall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident dementia (n = 241)</td>
<td>6.5 (2.7)</td>
<td>7.8 (3.0)</td>
<td>0.4 (0.8)</td>
</tr>
<tr>
<td>Incident dementia (n = 13)</td>
<td>4.6 (2.3)†</td>
<td>4.4 (2.3)†</td>
<td>−0.1 (0.6)†</td>
</tr>
<tr>
<td><strong>Stroop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident dementia (n = 241)</td>
<td>58.2 (24.9)</td>
<td>53.3 (20.6)</td>
<td>−1.2 (6.1)</td>
</tr>
<tr>
<td>Incident dementia (n = 13)</td>
<td>57.4 (18.7)</td>
<td>56.0 (23.3)</td>
<td>−0.4 (3.1)</td>
</tr>
<tr>
<td><strong>Letter–Digit Substitution Task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No incident dementia (n = 241)</td>
<td>27.8 (6.6)</td>
<td>27.1 (7.0)</td>
<td>−0.2 (1.3)</td>
</tr>
<tr>
<td>Incident dementia (n = 13)</td>
<td>26.0 (7.4)</td>
<td>20.5 (9.3)</td>
<td>−1.6 (1.2)</td>
</tr>
</tbody>
</table>

Only persons with incident dementia were included that were diagnosed after the first follow-up MRI. Persons with incident dementia between baseline and first follow-up (n = 3) were excluded as well as persons who had no first follow-up MRI. Adjusted for age, sex and educational level.

†P < 0.05 compared with no incident dementia.

### Table 4 Association between decline in hippocampal volume and cognitive decline in persons without incident dementia

<table>
<thead>
<tr>
<th></th>
<th>MMSE</th>
<th>Word recall</th>
<th>Delayed recall</th>
<th>Stroop</th>
<th>LDST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hippocampus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>1.26 (0.91–1.73)</td>
<td>1.10 (0.77–1.56)</td>
<td>1.40 (1.00–1.97)</td>
<td>1.07 (0.72–1.59)</td>
<td>1.37 (0.96–1.96)</td>
</tr>
<tr>
<td>Right</td>
<td>1.24 (0.90–1.72)</td>
<td>1.11 (0.77–1.61)</td>
<td>1.66 (1.18–2.33)</td>
<td>1.28 (0.87–1.89)</td>
<td>1.35 (0.94–1.94)</td>
</tr>
</tbody>
</table>

Numbers are odds ratios (95% confidence interval) per SD decrease in baseline hippocampal volume and hippocampal volume decline. Adjustments were made for age, sex, educational level and total intracranial volume. n = 283.

a Cognitive decline was defined as having a rate of decline one SD faster than average.

MMSE = Mini Mental State Examination; LDST = Letter–Digit Substitution Task.

### Table 5 Cognitive characteristics of persons without decline in delayed recall versus those with decline in delayed recall

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mini Mental State Examination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-decliners (n = 343)</td>
<td>28.2 (1.7)</td>
<td>27.9 (1.7)</td>
<td>28.0 (1.6)</td>
<td>27.9 (1.6)</td>
<td>27.0 (2.5)</td>
</tr>
<tr>
<td>Decliners (n = 71)</td>
<td>27.1 (2.0)†</td>
<td>27.1 (1.9)†</td>
<td>27.0 (2.4)†</td>
<td>26.7 (2.9)†</td>
<td>23.9 (4.1)†</td>
</tr>
<tr>
<td><strong>Immediate word recall (total number of words)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-decliners (n = 343)</td>
<td>21.4 (4.9)</td>
<td>–</td>
<td>25.3 (5.3)</td>
<td>21.7 (7.8)</td>
<td>23.5 (6.1)</td>
</tr>
<tr>
<td>Decliners (n = 71)</td>
<td>16.1 (3.7)†</td>
<td>–</td>
<td>19.7 (3.8)†</td>
<td>15.0 (4.5)†</td>
<td>17.7 (5.1)†</td>
</tr>
<tr>
<td><strong>Delayed word recall (total number of words)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-decliners (n = 343)</td>
<td>6.8 (2.2)</td>
<td>–</td>
<td>8.2 (2.6)</td>
<td>7.1 (2.6)</td>
<td>7.6 (3.0)</td>
</tr>
<tr>
<td>Decliners (n = 71)</td>
<td>3.2 (1.2)†</td>
<td>–</td>
<td>4.5 (1.8)†</td>
<td>3.7 (1.5)†</td>
<td>3.1 (1.8)†</td>
</tr>
<tr>
<td><strong>Stroop (total seconds)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-decliners (n = 343)</td>
<td>56.8 (21.3)</td>
<td>–</td>
<td>52.8 (20.8)</td>
<td>57.2 (21.9)</td>
<td>61.0 (25.3)</td>
</tr>
<tr>
<td>Decliners (n = 71)</td>
<td>61.4 (29.6)</td>
<td>–</td>
<td>60.0 (28.9)†</td>
<td>66.0 (33.1)†</td>
<td>69.8 (30.8)</td>
</tr>
<tr>
<td><strong>LDST (total number of correctly substituted)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-decliners (n = 343)</td>
<td>28.2 (6.8)</td>
<td>–</td>
<td>28.0 (6.8)</td>
<td>26.7 (7.2)</td>
<td>26.8 (7.4)</td>
</tr>
<tr>
<td>Decliners (n = 71)</td>
<td>25.7 (6.5)†</td>
<td>–</td>
<td>23.9 (6.9)†</td>
<td>24.4 (6.8)</td>
<td>23.2 (6.5)†</td>
</tr>
</tbody>
</table>

Cognitive decline in delayed recall was defined as having a decline one SD faster than average. Persons with incident dementia were excluded from these analyses.

*P < 0.05 for comparison between decliners and non-decliners.

other examinations. However, if anything, this healthy participator bias will lead to inconclusive findings and not to the present significant results. Second, within those who developed incident dementia, only 13 had a repeated MRI scan before clinical diagnosis. On the other hand, the advantage of our population approach, compared with studies in clinical settings, is that we did include patients when they were in a very early phase of the dementia. Also, the large amount of neuropsychological data in subjects who remained dementia free allowed us to examine subtle cognitive decline in association with (change in) hippocampal volumes. Third, we have a thorough medical follow-up of all participants including those who refused in-person examination, we cannot fully exclude the possibility that we miss some subjects with mild dementia who have no contact with the medical system. However, as we included these subjects in the ‘no dementia’ group, this would only reduce the possibility of finding an association. Fourth, at invitation for follow-up MRI scans those who were older were less likely to participate. These older subjects had the smallest hippocampal volumes at baseline. What we were able to visualize as hippocampal decline was probably an underestimation of the true decline that we could have measured if we had the opportunity to have MRI images of everyone including those who died in between MRI examinations. When computing the rate of decline in hippocampus between the baseline and first follow-up MRI scan, we found a decline of 0.5% per year that is comparable with other studies (Barnes et al., 2008). Finally, during the follow-up of the MRI, techniques inevitably had changed and improved and in 2006 we used a different 3D MRI sequence shortly after each other. However, absolute hippocampal volumes were systematically underestimated in the last MRI sequence. We used Z-scores to describe a subject’s hippocampal volume in the distribution at each point in time, therefore ranking of subjects according to severity of decline in Z-scores can be used.

Previous studies have shown that the hippocampus on MRI is atrophied in patients with Alzheimer’s disease (Fox et al., 1996; Horn et al., 1996), patients with mild cognitive impairment (Du et al., 2001; Devanand et al., 2007), and in cognitively healthy elderly subjects destined to develop Alzheimer’s disease (den Heijer et al., 2006). We found that the majority of patients who developed dementia had a smaller baseline hippocampal volume years before their clinical diagnosis than people who remained free of dementia. This is in line with clinical studies showing that ~80–90% of established patients with Alzheimer’s disease have a small hippocampal volume (Ridha et al., 2007; Colliot et al., 2008). Pathological validation studies have shown that hippocampal atrophy on MRI correlates with the specific Alzheimer’s disease neuropathology (Bobinski et al., 2000; Gosche et al., 2002). A longitudinal study in patients who carried an autosomal dominant mutation for Alzheimer’s disease found that a decline in hippocampal volume could be detected 5 years before the clinical diagnosis (Ridha et al., 2006). In the elderly, follow-up studies in patients with Alzheimer’s disease and mild cognitive impairment have shown an approximately two to four times faster rate of decline in hippocampal volume than in healthy controls (Jack et al., 2000; Laakso et al., 2000; Wang et al., 2003; Morra et al., 2008). In these studies, MRI scans were done when clinical diagnosis had already been made. We showed that decline in hippocampus is strongly associated with risk to develop dementia.

Although decline in hippocampal volume in addition to a single measurement was predictive of Alzheimer’s disease, it was not always observed in subjects with incident Alzheimer’s disease, suggesting that within a clinical diagnosis of Alzheimer’s disease heterogeneity exists in its pathological substrate. Furthermore, hippocampal atrophy on MRI has also been described in patients with other dementia types such as frontotemporal dementia (Barnes et al., 2006), and vascular dementia (Kril et al., 2002). Combining hippocampal volume measures with MRI volumes of other regions, such as the thalamus (de Jong et al., 2008) or changes assessed with other imaging techniques, such as diffusion...
tensor imaging or PET, might further improve identification of presymptomatic persons with Alzheimer's disease.

In addition to the analyses on clinical dementia, we also showed that in subjects who remained free of dementia over 10 years, a decline in hippocampal volume paralleled, or could even precede subtle cognitive decline, particularly in delayed memory function. No association was found with executive function tasks, although a borderline statistically significant association was found between hippocampal decline and decline in performance on the Letter-Digit Substitution Task. In this task however, not only executive function is tested, as remembering letters connected with digits may improve performing the task. In a smaller setting (Stoub et al., 2008) and a sample mixed with Alzheimer's disease, mild cognitive impairment and healthy controls (Mungas et al., 2005), rate of hippocampal decline was similarly found to be associated with memory function. These findings support the notion that subtle delayed memory decline with hippocampal volume decline can be observed long before a clinical diagnosis of dementia is made. We showed that persons who had a decline in delayed memory had similarly low test scores on executive function and a faster decline in Mini Mental State Examination score, suggesting that they are likely to develop dementia.

However, not all subjects with hippocampal decline or cognitive decline will eventually develop dementia. This could be either due to the fact they die before developing significant cognitive decline, or that hippocampal volume decline can be part of healthy ageing. A pathological study suggested differences in neuronal loss within the specific subregions of the hippocampus between normal ageing and Alzheimer’s disease (West et al., 1994). With recently introduced high resolution 7 T MRI scanners, subregion imaging of the hippocampus in vivo has become feasible, potentially allowing discrimination of normal ageing from Alzheimer’s disease (Thomas et al., 2008).

**Funding**


**References**


