Cortical Gray Matter Atrophy in Healthy Aging Cannot Be Explained by Undetected Incipient Cognitive Disorders: A Comment on Burgmans et al. (2009)

Anders M. Fjell, Lars T. Westlye, Thomas Espeseth, and Ivar Reinvang
University of Oslo

Kristine B. Walhovd
University of Oslo

The Alzheimer’s Disease Neuroimaging Initiative

Burgmans, van Boxtel, Vuurman, et al. (2009) published an interesting study titled “The Prevalence of Cortical Gray Matter Atrophy May Be Overestimated in the Healthy Aging Brain” on how subclinical cognitive disorders may affect correlations between age and cortical volume. Correlations between cortical gray matter volume and age were found in 30 elderly with cognitive decline after 6 years, but not in 28 elderly without cognitive decline. This study is important, and demonstrates that preclinical cognitive disorders may affect cortical brain volumes before being detectable by neuropsychological tests. However, we are not convinced by the conclusions: “. . . gray matter atrophy . . . is to a lesser extent associated with the healthy aging process, but more likely with brain processes underlying significant cognitive decline” (p. 547) and “. . . cortical gray matter atrophy in the aging brain may be overestimated in a large number of studies on healthy aging” (p. 547). We analyzed the cross-sectional MR data (n = 1,037) as well as longitudinal data from a sample of very well-screened elderly followed by cognitive testing for 2 years. In the cross-sectional data, the correlations between age and brain volumes were generally not much reduced when the upper age limit was lowered. This would not be expected if age-related incipient cognitive disorders caused the correlations given that the incidence of cognitive decline increased with age. Longitudinally, 1-year atrophy was identified in all tested regions. It is likely that cortical brain atrophy is manifested in cognitively normal elderly without subclinical cognitive disorders.

Keywords: aging, atrophy, cerebral cortex, hippocampus, cognition

Supplemental materials: http://dx.doi.org/10.1037/a0018827.supp

Anders M. Dale and Dominic Holland
University of California, San Diego

Anders M. Fjell, Lars T. Westlye, Thomas Espeseth, Ivar Reinvang, and Kristine B. Walhovd, Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway; Anders M. Dale, Multimodal Imaging Laboratory, Department of Radiology, and Department of Neurosciences, University of California, San Diego; Dominic Holland, Department of Neurosciences, University of California, San Diego.

Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators is available in the supplemental materials at: http://dx.doi.org/10.1037/a0018827.supp

Anders M. Dale is a founder and holds equity in CorTechs Labs., Inc., and also serves on the scientific advisory board. The terms of this arrangement have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies.

This research was funded by the following grants: Norwegian Research Council Grants 177404/W50 to Kristine B. Walhovd, 175066/D15 to Anders M. Fjell, 154313/V50 to Ivar Reinvang, and 177458/V50 to Thomas Espeseth; University of Oslo to Kristine B. Walhovd and Anders M. Fjell; National Institutes of Health Grants R01-NS39581, R37-AG11230, and R01-RR13609; National Center for Research Resources (NCRR) Grants P41-RR14075 and R01 RR16594-01A1; NCRR BIRN Morphometric Project Grants BIRN002 and U24 RR021382; National Institute for Biomedical Imaging and Bioengineering Grants R01 EB001550 and R01EB006758; National Institute for Neurological Disorders and Stroke Grant R01 NS052585-01; as well as the Mental Illness and Neuroscience Discovery (MIND) Institute, part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health (NIH) through NIH Roadmap for Medical Research Grant U54 EB005149. Additional support was provided by the Autism & Dyslexia Project funded by the Ellison Medical Foundation. We thank the developers of the OASIS (Open Access Series of Imaging Studies) database for access to MRI data constituting Samples 4 and 5 of the present work. According to Marcus et al. (2007), the OASIS database is supported by NIH Grants P50 AG05681, P01 AG03991, P20 MH071616, RR14075, RR 16594, and BIRN002; the Alzheimer’s Association; the James S. McDonnell Foundation; the MIND Institute; and the Howard Hughes Medical Institute. The longitudinal data collection and sharing were funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI, Principal Investigator: Michael Weiner; NIH Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and through generous contributions from the following: Pfizer Inc., Wyeth Research, Bristol-Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Merck & Co. Inc., AstraZeneca AB, Novartis Pharmaceuticals Corporation, Alzheimer’s Association, Eisai Global Clinical Development, Elan Corporation PLC, Forest Laboratories, and the Institute for the Study of Aging, with participation from the U.S. Food and Drug Administration. Industry partnerships are coordinated through the Foundation for the National Institutes of Health. 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 of Neuroimaging at the University of California, Los Angeles.

Correspondence concerning this article should be addressed to Anders M. Fjell, Department of Psychology, University of Oslo, Norway. P.h. 1094 Blindern, 0317 Oslo, Norway. E-mail: andersmf@psykologi.uio.no
MRI can be used to predict development of age-related cognitive disorders (Burgmans, van Boxtel, Smeets, et al., 2009; Driscoll et al., 2009; Fischl et al., 2002), and elderly with mild cognitive impairment (MCI) have reduced volume or thickness of the cerebral cortex and a range of subcortical structures (de Leon et al., 2006; Fennema-Notestine et al., 2009; Jack et al., 2008). MR findings also show consistent cortical and subcortical atrophy in healthy elderly (Fjell et al., 2009; Raz et al., 2005; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003). A possible problem in all studies of healthy aging is that even well-screened samples of healthy persons may include participants with preclinical cognitive disorders. Burgmans, van Boxtel, Vuurman, et al. (2009) overcame this problem by using cognitive follow-up data obtained 6 years after MR scanning. In the group of cognitively stable elderly, age did not correlate with cortical or hippocampal volume. The authors suggest that amount of cortical gray matter atrophy in healthy aging may be overestimated in previous studies, and that gray matter atrophy in hippocampus, parahippocampal, and frontal brain areas to a lesser extent is associated with the aging process.

We find this conclusion surprising. Cross-sectional studies have shown that changes in cortical thickness and volume follow a mainly linear trajectory throughout the adult life span (Allen, Bruss, Brown, & Damasio, 2005; Fjell et al., 2009; Walhovd et al., 2009). If preclinical age-related cognitive disorders are causing a substantial part of the age correlations, we would expect a reduction of brain volumes in the latter part of the life span only. We believe that the study by Burgmans, van Boxtel, Vuurman, et al. (2009) is well conducted and important, but we think that more evidence is needed to support their conclusion. First, the sample was small, with 28 and 30 participants in two groups. Second, as acknowledged by the authors, only cross-sectional data were presented.

In the present report, we present two sets of analyses. First, we reanalyzed data from several previously published cross-sectional MR studies of aging (Espeseth et al., 2006; Fjell et al., 2009; Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Raz et al., 2005; Walhovd et al., 2005, 2009; Westlye et al., 2009), yielding a total sample of 1,037. As the prevalence of age-related cognitive disorders increases with higher age, the proportion of preclinical cognitive disorders will likely increase with the age of the sample if the cognitive screening has not been sufficiently thorough, as implied by Burgmans, van Boxtel, Vuurman, et al. (2009). Thus, by restricting the upper limit of the age span, we expected less contamination by preclinical cognitive disorders. If Burgmans, van Boxtel, Vuurman, et al. are correct in their conclusion, one should find that the correlations between age and brain structure will decrease and eventually vanish as the upper limit of the age span is lowered.

Second, we have recently shown substantial longitudinal cortical reductions over only 1 year in healthy elderly from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) sample (Fjell et al., 2009). The ADNI database presently contains extensive cognitive follow-up data for 2 years. Here, we reanalyzed 1-year longitudinal MR data from normal controls from ADNI, excluding all participants showing any signs of cognitive decline 2 years after the first scanning.

**Cross-Sectional Analysis**

**Sample.** The sample was drawn from the multisample study of aging, coordinated from the Center for the Study of Human Cognition, Department of Psychology, University of Oslo, and consisted of several independent samples from Norway and the United States, with a total sample of 1,037 participants (642 women, 395 men) and an age range of 18 to 94 years (M = 47.0 years, SD = 19.9). All participants were well screened for diseases and history of neurological conditions and for cognitive deficits or dementia by standardized tests. Details of each sample, screening criteria, and key publications are presented in online supplemental Table 1.

**MR acquisition and analysis.** All scans were obtained from 1.5 T magnets from two different manufacturers (Siemens, Erlangen, Germany; General Electric [GE], Milwaukee, WI) and from five different models (Siemens Avanto, Symphony, Sonata, Vision; GE Signa). All participants in each sample were scanned on the same scanner. T1-weighted sequences were acquired (3-D magnetization prepared gradient-echo for Siemens; 3-D spoiled gradient recalled pulse sequences for GE). Multiple scans were acquired for 846 of the participants within the same scanning session and averaged to increase the signal-to-noise ratio. The details of the sequences are presented in online supplemental Table 2.

All data sets were processed and analyzed with FreeSurfer 4.01 (http://surfer.nmr.mgh.harvard.edu/) at the Neuroimaging Analysis Lab, Center for the Study of Human Cognition, University of Oslo, with the additional use of computing resources from the TITAN grid cluster (http://hpc.uio.no/index.php/TITAN) run by the Research Computing Services Group at USIT, University of Oslo. The procedure is described in detail elsewhere (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993) and in online supplemental material 3.

**Longitudinal Analysis**

**Sample.** The longitudinal raw data were obtained from the ADNI database (www.loni.ucla.edu/ADNI; Principal Investigator Michael W. Weiner, VA Medical Center and University of California–San Francisco; for more information, see http://www.adni-info.org). 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. ADNI eligibility criteria are described at http://www.adni-info.org/index.php?option=com_content&task=view&id=9&Itemid=43. Briefly, participants were 55–90 years of age, had an informant able to provide an independent evaluation of functioning, and spoke either English or Spanish. All participants were willing and able to undergo test procedures, including neuroimaging, and agreed to longitudinal follow-up. Users of specific psychoactive medications were excluded. General inclusion and exclusion criteria were as follows for normal subjects: Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) scores between 24 and 30 (inclusive), Clinical Dementia Rating (CDR; Morris, 1993) of zero, nondepressed, non-MCI, and nondemented. In addition, CDR sum of boxes (CDR–SB) was calculated as a measure of clinical
functioning (does not need to be zero). The subject pool was further restricted to those subjects for whom adequate processed and quality checked MR data were available by February 2009. One hundred eighteen healthy elderly had been tested at Visits 1 (baseline), 3 (1 year), and 5 (2 years), and had MR of sufficient quality. MR change was measured from Visit 1 to Visit 3 (1-year change). To be included in the analyses, CDR–SD and MMSE needed to remain identical or improve over 2 years. In addition, raw scores on each of the following seven neuropsychological tests had to remain at 90% of initial score or higher: Auditory Verbal Learning Test (AVLT) learning (hits – false alarms), AVLT 30-min delayed recall (hits – false alarms; Rey, 1941), digit span (sum of forward and backward), clock copying (Goodglass & Kaplan, 1983), digit symbol substitution test (Wechsler, 1981), and logical memory test immediate recall and logical memory test delayed recall (Wechsler, 1987). This means, for instance, that a participant who scored 10 points on the AVLT delayed recall test at baseline and 8 points after 2 years would be excluded, whereas a participant who scored 9 points after 2 years could be included. After applying these very stringent exclusion criteria, 21 participants were included in the longitudinal analyses (mean age at baseline = 76.8 years, range = 63.0–90.2 years; 6 women and 15 men).

**MR acquisition and analysis.** All scans used for the present analyses were from 1.5 T scanners. Data were collected across a variety of scanners with protocols individualized for each scanner, as defined at http://www.loni.ucla.edu/ADNI/Research/Cores/index.shtml. Raw DICOM MRI scans (including two T1-weighted volumes per case) were downloaded from the public ADNI site (http://www.loni.ucla.edu/ADNI/Data/index.shtml) and processed as described elsewhere (Fennema-Notestine et al., 2009). A recently developed procedure for measuring longitudinal change was used (Holland et al., 2009) based on nonlinear registration of images and linear elasticity (see online supplemental material 3). The method enables very precise registration involving large or subtle deformations, even at small spatial scales with low boundary contrast. The procedure produces an estimate of the percentage cortical volume loss at each vertex and within each region of interest (ROI). The method has been validated in model studies of complex spherical-shell geometries with low contrast and noise, where a prescribed volume change is numerically estimated to accuracies of within 0.5% (Holland et al., 2009). It is beyond the scope of the present article to compare the sensitivity of this method with the manual tracings used by Burgmans, van Boxtel, Vuroman, et al. (2009).

### Statistical Analyses

Intracellular volume was regressed out from all volume measures before they were entered into statistical analyses. In the cross-sectional data, age was correlated with each of the ROIs in the age ranges 18–95, 18–80, 18–70, and 18–60 years, and the changes in the Pearson correlation coefficient were inspected. Hippocampal volume was not included as an ROI in the cross-sectional analyses because this structure is known to follow a highly nonlinear trajectory over the adult life span, and the present correlation approach would then be inappropriate. For the rest of the ROIs, degree of nonlinearity was tested by multiple regressions with brain volume as dependent and age and square of age as simultaneous predictors. One-sample t tests were used for the longitudinal data to test whether significant atrophy was seen over 1 year (from Visit 1 to Visit 3).

### Results

#### Cross-Sectional Results

Cross-sectional age correlations are presented in Table 1. All ROIs correlated significantly (p < .01) with age for all age spans, except for parahippocampal gyrus when the sample was restricted to participants under 60 years. None of the correlation coefficients for the ROIs except parahippocampal gyrus were notably lower when the upper limit of the sample was lowered. For parahippocampal volume, however, a steady reduction in correlation strength was observed, from −.29 for the full sample, through −.21 and −.10, and finally .02 for the sample under 60 years.

We also tested for nonlinearity by running multiple regressions with brain volume as dependent and age and square of age as simultaneous predictors. In one case, square of age was not significant (inferior prefrontal); in four cases, the term was positive (anterior and posterior cingulate, orbital, and dorsolateral prefrontal), indicating a U-shaped relationship; and in one case (parahippocampal gyrus), the quadratic term was negative, indicating an inverse U-shaped relationship.

#### Longitudinal Results

One-sample t tests were used to test whether 1-year atrophy in the 21 cognitively very stable participants was significantly different from zero. Significant reduction was observed in all ROIs, with annual change in the range −0.31% (parahippocampus) to −0.98% (hippocampus; see Table 2).

<table>
<thead>
<tr>
<th>Cortical region</th>
<th>&lt;95 years</th>
<th>&lt;80 years</th>
<th>&lt;70 years</th>
<th>&lt;60 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior cingulate</td>
<td>−.24**</td>
<td>−.25**</td>
<td>−.27**</td>
<td>−.33**</td>
</tr>
<tr>
<td>Dorso/medial prefrontal</td>
<td>−.67**</td>
<td>−.67**</td>
<td>−.66**</td>
<td>−.67**</td>
</tr>
<tr>
<td>Inferior prefrontal</td>
<td>−.65**</td>
<td>−.63**</td>
<td>−.56**</td>
<td>−.47**</td>
</tr>
<tr>
<td>Orbital prefrontal</td>
<td>−.60**</td>
<td>−.58**</td>
<td>−.54**</td>
<td>−.53**</td>
</tr>
<tr>
<td>Parahippocampal</td>
<td>−.29**</td>
<td>−.21**</td>
<td>−.10**</td>
<td>.02</td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>−.52**</td>
<td>−.49**</td>
<td>−.47**</td>
<td>−.43**</td>
</tr>
</tbody>
</table>

*p < .01. **p < 10^-10.*

---

**Table 1**

**Pearson Correlations of Cortical Atrophy With Age With Different Upper Age Limits**
Discussion

The results of Burgmans, van Boxtel, Smeets, et al. (2009) are interesting because they indicate that MR can be used to identify differences between groups with no versus very mild or incipient cognitive disorders. This confirms the potential value of MR morphometry in detecting and monitoring age-related cognitive conditions. However, we do not find adequate support for the authors’ conclusion that substantial cortical gray matter changes may not be associated with the healthy aging process, and that age-related atrophy typically has been overestimated in previous studies, which generally lack follow-up data on cognitive function. We believe that two lines of evidence speak against this interpretation.

First, because cognitive disorders are increasingly common with older age, it is reasonable to assume that the prevalence of incipient cognitive problems will be higher among participants in their 80s than participants in their 70s. Still, previous research has tended to report that cortical gray matter atrophy follows a largely linear pattern during the adult life span (Allen et al., 2005; Fjell et al., 2009; Wallhovd et al., 2009). This would be a surprising finding if preclinical cognitive disorders explained a substantial proportion of the atrophy among the elderly. In the present report, we analyzed data from more than 1,000 participants, drawn from previously published cross-sectional studies of healthy aging (Espeseth et al., 2006; Fjell et al., 2009; Fotenos et al., 2005; Raz et al., 2005; Wallhovd et al., 2009; Westlye et al., 2009). For all cortical ROIs except the parahippocampal gyrus, the age correlations remained relatively stable when the upper age limit was systematically lowered. For all ROIs except one, the relationships between age and volume were best described as linear or as following a U-shaped curve. We believe this indicates that the effects were not caused by inclusion of participants with preclinical cognitive disorders.

The samples included in the analyses were all well screened at baseline, but no follow-up cognitive data had been used to exclude possible participants with declining cognitive functions. The samples are representative of MR studies of healthy aging, indicating that the previously reported age differences in brain morphometry to a lesser degree can be attributed to incipient cognitive disorders. If the screening procedures in these samples had been insufficient to exclude preclinical conditions, the proportion of participants with declining cognitive function would obviously be too low to affect the age relationships notably. One possible exception from this pattern was parahippocampal cortex, where the correlation coefficients steadily were reduced as the sample became younger. The medial temporal lobe is often found to be vulnerable to early manifestations of MCI and Alzheimer’s disease (AD; Dickerson et al., 2008; Du et al., 2007), even though entorhinal cortex is often found to be affected earlier in the disease than parahippocampal cortex. As AD is the most common cause of cognitive decline among the elderly, it is possible that an increasing fraction of the participants included may have incipient AD among the oldest participants, and that this affected the age correlation seen here. Fotenos, Mintun, Snyder, Morris, and Buckner (2008) showed that nondemented elderly with higher socioeconomic status had smaller brain volumes and higher rates of longitudinal brain atrophy than elderly with lower socioeconomic status. Thus, participants with higher socioeconomic status and normal neuropsychological functioning may still have incipient AD, but this is impossible to detect without follow-up data. Although it is highly speculative, this may contribute to the observed age differences in parahippocampal cortex. To allow detection of age effects in parahippocampal cortex in healthy elderly without incipient cognitive disorders, it may be necessary to employ a more sensitive design than the usual cross-sectional studies, that is, longitudinal investigations.

Second, Burgmans, van Boxtel, Vuurman, Smeets, Gronenschild, Uylings, & Jolles (2009) reported cross-sectional results, implying that the findings pertain to age-related differences, not atrophy per se. As the authors acknowledge, longitudinal data are needed to confirm their conclusions. To address this issue, we analyzed longitudinal data from the ADNI. For these data sets, cognitive follow-up data were available for 2 years only, not 6 years after scanning as in Burgmans et al. Still, we applied very strict exclusion criteria, ensuring no decline in CDR–SB or MMSE scores, and no more than 10% decline on any of seven other neuropsychological tests over 2 years. This led to inclusion of less than 18% of the initial sample of participants classified as normal elderly at baseline. Rates of atrophy were calculated over only 1 year, and cognitive data were available for 1 year after the last scan. Thus, even though it would be beneficial to have cognitive follow-up data for longer time periods, we believe it is very unlikely that any of these 21 participants had incipient cognitive disorders. Still, in all the ROIs tested, significant atrophy over 1 year was found, on the order of 0.31% to 0.98%. Both hippocampal and parahippocampal volumes declined significantly. Thus, even in a group of superstable elderly, significant longitudinal atrophy was observed over as short a time interval as 1 year. The magnitude of atrophy may seem small, but over a decade, the median ROI would show more than 5.4% decline, and hippocampus would shrink about 10.8%. If such changes are happening throughout large parts of the adult life span, the percentage reduction will be substantial. Further studies are needed to better disentangle the implications of such atrophy for clinical neuroscience.

In a recent study, we found that cerebrospinal fluid levels of amyloid and tau proteins, both of which are sensitive to AD, were related to brain atrophy also in healthy elderly (Fjell et al., in press), but that the pattern of brain changes did not resemble that seen in AD (Fjell et al., 2009). We need more knowledge about which factors mediate brain atrophy in healthy elderly and what consequences the changes may have for cognitive function. An interesting hypothesis is also that high-functioning persons may show less age-related brain decline because cognitive activity may...
increase cortical thickness (Draganski et al., 2006). Experimental manipulations are needed to address this issue.

In conclusion, we believe that evidence from both cross-sectional and longitudinal data indicates that reduction in brain volumes is a part of the normal aging process, and not necessarily related to pathological processes underlying cognitive decline. The differences between the present results and findings of Burgmans, van Boxtel, Vuurman, et al. (2009) may be explained by limited power due to small sample size in combination with a cross-sectional design. Longitudinal designs have much higher sensitivity than cross-sectional designs. Combined with small sample size, the study may not have had sufficient statistical power to detect changes in healthy aging, but only age-related differences associated with pathology or rapid cognitive decline. Much more power and longitudinal data are needed to “prove the null hypothesis” that there are no age-related brain changes in healthy aging.

References


Received September 16, 2009
Revision received November 30, 2009
Accepted December 8, 2009