**We would like to thank the reviewers for all their valuable comments, which we believe have contributed to improve the manuscript. Please, find below our point-by-point responses (in blue ink). All changes are tracked in the revised manuscript.**

**Reviewing: 1**

The manuscript "The interactive effect of demographic and clinical factors on hippocampal volume: a multi-cohort study on 1958 cognitively normal individuals" by Ferreira et al. reports the results of an extremely interesting multicenter investigation aiming at assessing both difference and reliability of the most used database in structural hippocampal research.

The analyzed material is extremely large, resulting in accurate and reliable analyses. Results identify pros and cons of the various image sets in an appropriate and critical way.

The manuscript is well written and pretty clear in the aims, methods, results and discussion.

Minor comments

1.1. The last sentence of the Discussion at page 16 (Nonetheless….) is not supported by the present data and it is debatable: I would remove it from the text.

We agree. We have removed the sentence as suggested.

1.2. The authors should add a sentence/paragraph into the Discussion or Conclusion (point (1) at page 17) suggesting, according to the present findings, a strategy the investigators should put on in choosing the reference database depending on the material and the aim of their own study (i.e. which parameters/similartities have to be privileged and which ones left behind).

Thanks to the reviewer for this very nice suggestion. We have included the following sentence (p.16 first paragraph):
"The strategy for recruiting individuals will depend on the study aims and the materials to be used. Regarding hippocampal atrophy as imaging biomarker, our findings show that age and global brain atrophy are the most important factors to consider, but gender, education, MMSE, and TIV should also be taken into account".

**Reviewing: 2**

The authors have performed a nice study comparing different cohorts of healthy controls from both selected and population based cohorts. This is an important study and it addresses a lot of important issues. There are also some limitations that needs to be worked on before it is ready for publication. My comments are attached in a separate file.

2.1. In the abstract the aim of the paper is stated: 1) To evaluate the interactive effect

of key demographic and clinical factors on hippocampal volume. 2) Investigate how comparable the control groups from three multi centre cohort studies are with five single-center population based cohorts. In the results section it is summarized which variables that could explain variability in hippocampal volume, none of which were suprizing or not known from previous work. The authors also found that two of the cohort studies differed from population based studies. Not so surprising either since the recruitement is quite different. Further down in the conclusion the authors state that this may have important implications for future research in translating imaging biomarkers to the general population. I agree that a main challenge is to validate predementia markers when the variability in normal populations is so large, but again. Do the authors suggest a better way to recruit control groups for clinical cohort studies?

As the reviewer points out, most of the variables that explained variability in hippocampal volume are known from previous work. This is nonetheless the first time that all these variables are tested in an interactive way in the same cohort. The advantage of this is that we could investigate interactions between these variables. Other advantage is that we ascertained which variables have a stronger influence on hippocampal volume once they all are considered simultaneously.

Yes, another finding is that the multi-center cohorts differed from the population-based studies. As the reviewer mentions, this is not surprising, but this was not enough documented in previous studies. The current study helps to identify what confounding variables should be taken into account when extrapolating results from these multi-center studies to other cohorts, which may be of relevance for many researchers. We have indicated this in the revised manuscript as suggested by the reviewers in comment 1.2 and 2.6 (p.16 first paragraph).

What the reviewer points out about the way to recruit control groups for clinical cohort studies is crucial. The answer is not straightforward and this will depend on the aims of such studies. The problem in our opinion is that many researchers have been basing their own research on findings coming from ADNI. However, this is very dangerous since as we demonstrate here, extrapolation needs to be taken with great care. We have commented on this in the conclusion. Please see also responses to comments 1.2 and 2.6.

2.2. Introduction: I think the introduction is too long, and personally I think that there is too much focus on AD, when the focus of this paper is hippocampal volume in healthy controls and populations. If this was shortened, it would be easier to read.

Thanks to the reviewer for this suggestion. We have shortened the introduction and lessen the focus on AD.

2.3. On page 6 the authors write that ADNI controls resemble clinical populations. I think it is unclear and do not understand what exactly is meant by this. It needs to be rephrased.

Yes, we agree. We have removed that sentence to avoid confusion, also aiming to shorten the introduction as suggested above.

2.4. In the primary aim I think the paper is good and they have done a good job in studying the effects of several demographic and clinical variable on hippocampal volume. In their secondary aim, however I think it is not so good. The main reason is that they do not at all consider the fact that these different studies all rely on MRI from many different centers and skanner. Especially the multi-centre studies where eg more than 30 centres have recruited subjects to ADNI. There are also publications that the variation between scanners is as high as 10 % for some of the volumes in the ADNI MRI study. Introducing two more of these multicentre cohorts, what effect does this have on their results? It is of course interesting to se how much these differences can be minimized with clinical and demographic covariates, but how much variation is due to scanners and protocol differences? How generalizable are studies like this? In the five population based studies there were only one centre per study, which minimizes the difference based on MRI equipment. We have seen that choice of MRI sequence can give higher volumes in Freesurfer than a different sequence of the same brain in the same person and on the same scanner. Therefore systematic differences can be introduced through such choices. I would like the authors to include analyses for this and also discuss this, given their secondary aim for this paper.

Thanks to the reviewer for this comment. A huge effort was done in ADNI in order to harmonize the different scanners and being able to obtain as similar as possible images across centers. Also, most of the other cohorts included in the current study followed ADNI protocols. We are aware about variation between scanners in volumetric estimations. What the reviewer mentions here is correct and actually adds to the reasons for being careful when translating results from ADNI to other cohorts. This is indeed the main message of the current study.

Nonetheless, despite that 10% variation from some volumes in the ADNI MRI study, hippocampal volume estimations are rather stable as compared to other regions in the brain. Thanks to this, it is possible to combine hippocampal volumes from different software and field strengths (not only scanners) as demonstrated in previous studies (Briellmann et al., 2001; Whitwell et al., 2012; Hibar et al., 2015). This is one of the reasons why we selected hippocampal volume instead of any other neuroimaging variable, among other reasons.

Unfortunately, we cannot know how much of the variance is due to scanners and protocol differences. We have now recognized this as a limitation (p.15 line 8). To note, this affects only those analyses where the different cohorts were combined or compared to each other, but not those analyses carried out for the separate samples. This is mentioned in the limitations as well (p.14 last line).

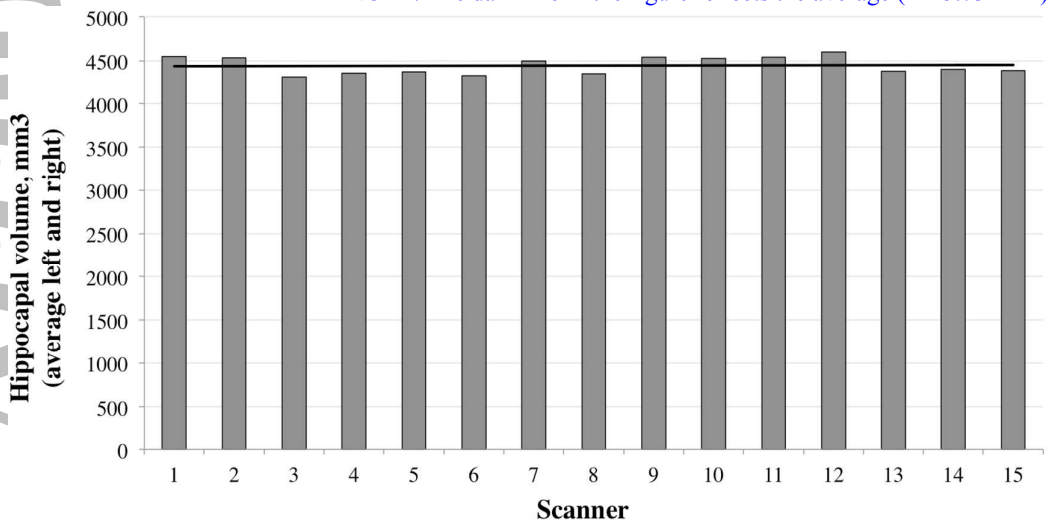
We cannot compare volumetric estimations across the different ADNI centers as suggested by the reviewer. However, we have done that using data from ICINET, which was available at our group. ICINET is a multi-center collaboration and include data from several types of scanners (GE, Philips and Siemens) as well as different field strength (1.5T and 3T) across Sweden. **The same participant was scanned in all these scanners and results are shown in the figure below.** As it can be seen, the percentage difference between the scanners providing the largest and smallest hippocampal volume is only 6.1%. To note, scanners and T1-weighted sequences are not harmonized in ICINET. Hence, smaller variation in hippocampal volume would

be expected across centers in ADNI, where strict harmonization was performed and only 1.5T data was used. As mentioned above, we know of studies reporting high variation from some volumes in the ADNI MRI study (e.g. Kruggel et al. 2010 Neuroimage 49:2123-33), but hippocampal volume estimations are rather stable as compared to other regions in the brain.

Finally, we wanted to highlight that when looking at Figure 1 in the manuscript, one can see that variation in ADNI is as large as in any other single-center sample and, indeed, variation is larger in the single-center samples of GENIC, BioFINDER and SNACK. Further, the difference between GENIC (largest hippocampal volume) and BRC (smallest hippocampal volume) is of 17.5%. This value clearly exceeds the mentioned 10% difference and our reported 6.1% difference for ICINET. Thus, although differences in scanners/protocols between centers in this study might explain part of the variance in hippocampal volume, other factors also explain such differences. Our results indicate that the considered demographic and clinical variables are more important in explaining differences between cohorts than different scanners/protocols. Certainly, between-cohorts differences were dramatically reduced by controlling for these demographic and clinical variables (Figure 4 in the manuscript) by reducing $\eta^2_{par}$ = 19% to $\eta^2_{par}$ = 2%. So, we could say that variance potentially attributable to differences in scanners/protocols is part of statistical noise in our study. Of note, the large sample used allowed us to reduce this noise to the minimum. So, part of the residual differences in Figure 4 ($\eta^2_{par}$ = 2%) might well correspond to differences in scanners/protocols and possibly some other confounding variables.

All these points are now reflected in the limitations section (end of p.14).

NOTE: The dark line in the figure reflects the average (4445.73 $mm^3$)




2.5. Discussion: In general I would like it to be shortended and more to the point. I think that the authors repeat themselves several times eg with stating that the interactive effect of demographic and clinical factors have not been previously

investigated. I have not counted, but I think this is a sign that the authors need to rewrite and shorten their discussion to make it more to the point and less repetitive.

We have shortened the discussion. Further, we have put special attention to repetitive sentences and removed them from the current version of the manuscript.

2.6. The conclusion section is long, and can also be shortened without a problem. In the conclusion I think the authors should give some indication on how they think future studies could be planned instead of repeating points from their discussion about how these things have been done so far.

Thank you for this good suggestion. We have shortened the conclusion and given the indications requested by the reviewer (see conclusion).

**TITLE PAGE**

# The interactive effect of demographic and clinical factors on hippocampal volume: a multi-cohort study on 1958 cognitively normal individuals

Daniel Ferreira [1], Oskar Hansson [2], José Barroso [3], Yaiza Molina [3], Alejandra Machado [3], Juan Andrés Hernández-Cabrera [3], J-Sebastian Muehlboeck [1], Erik Stomrud [2], Katarina Nägga [2], Olof Lindberg [1,2], David Ames [4], Grégoria Kalpouzos [5], Laura Fratiglioni [5,6], Lars Bäckman [5,6], Caroline Graff [7,8], Patrizia Mecocci [9], Bruno Vellas [10], Magda Tsolaki [11], Iwona Kłoszewska [12], Hilkka Soininen [13], Simon Lovestone [14], Håkan Ahlström [15], Lars Lind [16], Elna-Marie Larsson [15], Lars-Olof Wahlund [1], Andrew Simmons [1,17,18,19], and Eric Westman [1], for the AddNeuroMed consortium, for the Alzheimer's Disease Neuroimaging Initiative (ADNI)#, and for the Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL) research group.

**Affiliations:**

[1] Division of Clinical Geriatrics, Centre for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, 14186 Stockholm, Sweden; [2] Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden; [3] Department of Clinical Psychology, Psychobiology and Methodology, University of La Laguna, 38071 La Laguna, Spain; [4] National Ageing Research Institute, Royal Melbourne Hospital, 03050 Victoria, Australia; Academic Unit for Psychiatry of Old Age, Department of Psychiatry, The University of Melbourne, St. Vincent's Aged Psychiatry Service, St George's Hospital, 03065 Victoria, Australia; [5] Aging Research Centre (ARC), Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, 11330 Stockholm, Sweden; [6] Stockholm Gerontology Research Centre, 11330 Stockholm, Sweden; [7] Division of Neurogeriatrics, Centre for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, 14186 Stockholm, Sweden; [8] Department of Geriatric Medicine, Karolinska University Hospital Huddinge, 14186 Stockholm, Sweden; [9] Institute of Gerontology and Geriatrics, University of Perugia, 06100 Perugia, Italy; [10] INSERM U 558, University of Toulouse, 31024 Toulouse, France; [11] 3rd Department of Neurology, Aristoteleion Panepistimeion Thessalonikis, 54124 Thessaloniki, Greece; [12] Medical University of Lodz, 92216 Lodz, Poland; [13] University of Eastern Finland and Kuopio University Hospital, 70211 Kuopio, Finland; [14] Department of Psychiatry, Warneford Hospital, University of Oxford, OX37JX Oxford, UK; [15] Department of Surgical Sciences, Radiology, Uppsala University, 75185 Uppsala, Sweden; [16] Department of Medical Sciences, Uppsala University, 75185 Uppsala, Sweden; [17] NIHR Biomedical Research Centre for Mental Health, SE58AF London, UK; [18] NIHR Biomedical Research Unit for Dementia, SE58AF London, UK; [19] Institute of Psychiatry, King's College London, SE58AF London, United Kingdom.

# Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). 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 can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

**Running headline:** interactive effects on hippocampal volume

**Count:** 14 text pages, 4 figures, and 4 tables

**Correspondence:** Daniel Ferreira, PhD. Division of Clinical Geriatrics; Centre for Alzheimer Research; Department of Neurobiology, Care Sciences and Society; Karolinska Institutet. Novum, Plan 5, 14186 Stockholm (Sweden). Email: daniel.ferreira.padilla@ki.se. Telephone: +46720128047. Fax: +46858585470.

**Keywords:**

Alzheimer's disease, multi-cohort, aging, magnetic resonance imaging, hippocampal volume

RRID:SCR_003819, RRID:SCR_003007, RRID:SCR_001847



**ABSTRACT**

**Background:** Alzheimer's disease is characterized by hippocampal atrophy. Other factors also influence the hippocampal volume, but their interactive effect has not been investigated before in cognitively healthy individuals. To evaluate the interactive effect of key demographic and clinical factors on hippocampal volume, in contrast to previous studies frequently investigating these factors in a separate manner. To investigate how comparable the control groups from ADNI, AIBL, and AddNeuroMed are with five population-based cohorts.

**Methods:** 1958 participants were included (100 AddNeuroMed, 226 ADNI, 155 AIBL, 59 BRC, 295 GENIC, 279 BioFiNDER, 398 PIVUS, and 446 SNAC-K). ANOVA and random forest were used for testing between-cohort differences in demographic-clinical variables. Multiple regression was used to study the influence of demographic-clinical variables on hippocampal volume. ANCOVA was used to analyze whether between-cohort differences in demographic-clinical variables explained between-cohort differences in hippocampal volume.

**Results:** Age and global brain atrophy were the most important variables in explaining variability in hippocampal volume. These variables were not only important themselves but also in interaction with gender, education, MMSE, and total intracranial volume. AddNeuroMed, ADNI and AIBL differed from the population-based cohorts in several demographic-clinical variables that had a significant effect on hippocampal volume.

**Conclusion:** Variability in hippocampal volume in individuals with normal cognition is high. Differences that previously tended to be related to disease mechanisms could also be partly explained by demographic and clinical factors independent from the disease. Further, cognitively normal individuals especially from ADNI and AIBL are not representative of the general population. These findings may have important implications for future research and clinical trials, translating imaging biomarkers to the general population, and validating current diagnostic criteria for Alzheimer's disease and predementia stages.

## INTRODUCTION

Hippocampal atrophy has become a well-established biomarker of Alzheimer's disease (AD) (Frisoni et al., 2010; Morris et al., 2014). Factors such as age, gender, education, global brain atrophy, intracranial volume, and APOE ε4 genotype are known to influence hippocampal volume (Yamaguchi et al., 2002; Morra et al., 2009; Crivello et al., 2010; Raz et al., 2010; Striepens et al., 2011; Noble et al., 2012; Janowitz et al., 2014; Yuefeng et al., 2014; Jiang et al., 2014; Shpanskaya et al., 2014; Voevodskaya et al., 2014). However, other studies have failed to show an effect of these factors on hippocampal volume (Yamaguchi et al., 2002; Morra et al., 2009; Striepens et al., 2011; Janowitz et al., 2014; Jiang et al., 2014; Shpanskaya et al., 2014; Yuefeng et al., 2014). An explanation for these conflicting results is that all previous studies have focused on one or a few of these factors at a time. The interactive influence of all these factors on hippocampal volume in the same study and sample has not previously been investigated.

A substantial number of the publications on hippocampal atrophy comes from three large multi-center cohorts, i.e. ADNI (Mueller et al., 2005), AIBL (Ellis et al., 2010), and AddNeuroMed (Lovestone et al., 2009). However, specific recruiting procedures led to the collection of highly selected samples. In consequence, these cohorts have a higher prevalence of individuals with family history of dementia, participants are younger, more educated and have better global cognitive status than that reported in the general population (Whitwell et al., 2012; Brodaty et al., 2014). Nonetheless, studies comparing ADNI, AIBL, and AddNeuroMed versus population-based samples are scarce and have mainly focused on the MCI and AD groups. Therefore, it is still unclear to what extent control groups from ADNI, AIBL and AddNeuroMed are representative of the general population and whether results are generalizable.

The first aim of this study was to investigate the simultaneous effect of several demographic and clinical factors on hippocampal volume in healthy individuals. To that end, we combined eight large-scale international cohorts, leading to the largest sample to date in a study of this kind. The second aim was to investigate how comparable the control groups from ADNI, AIBL, and AddNeuroMed are with population-based cohorts. To ascertain this is critical and may have important implications since most of the results coming from ADNI, AIBL and AddNeuroMed directly depend on the characteristics of the control group. A specific question was whether between-cohort differences in hippocampal volume could be successfully minimized by statistical control of key demographic and clinical factors.

## MATERIALS AND METHODS

### Participants

This study includes a total of 1958 cognitively normal individuals from the following eight large-scale international cohorts: AddNeuromed (http://www.innomed-addneuromed.com/, RRID:SCR_003819), ADNI (http://adni.loni.usc.edu/, RRID:SCR_003007), and AIBL (https://aibl.csiro.au/) (multi-center cohorts), and BioFINDER, BRC, GENIC, PIVUS, and SNAC-K (single-center population-based cohorts). Cohorts' characteristics and eligibility criteria are displayed in Tables 1 and 2. Approval was obtained from local ethics committees. Data collection was carried out in accordance with relevant regulations at each center and participants gave written consent in accordance with the Declaration of Helsinki.

### Demographic and clinical variables

Age, gender, education and handedness were selected as demographic variables. Clinical information included the mini-mental state examination (MMSE), clinical dementia rating

(CDR) scale, and several instruments for assessing depressive symptomatology and functional activity (Table 3). Subjective memory complaints were operationalized as detailed in Table 2. Cerebrospinal fluid (CSF) levels of $A\beta_{42}$, total tau (T-tau), and phosphorylated tau (p-tau) were also measured in ADNI and BioFINDER (Supplementary Table 1a).

**Magnetic Resonance Imaging**

High-resolution 3D T1-weighted sequences were acquired in all the cohorts. MRI scanner and acquisition parameters are detailed in Supplementary Table 1b. Image processing was performed with FreeSurfer 5.1.0 (http://surfer.nmr.mgh.harvard.edu/, RRID:SCR_001847) using TheHiveDB database system (Muehlboeck et al., 2014). FreeSurfer provides measurements of cortical and subcortical volumes, as well as an estimation of the total intracranial volume (TIV) (Fischl et al., 2002; Desikan et al., 2006). Left and right hippocampal volumes were summed together. The Brain volume (BV)/CSF index was also calculated as a proxy of global brain atrophy using the following formula: Brain volume (BV)/CSF index = (total grey matter volume + total white matter volume) / total CSF volume. This index correlates with clinical measures, CSF biomarkers and cognition, and has been proposed for staging individuals according to the degree of global brain atrophy and for monitoring disease progression (Orellana et al., 2016). Lower values of the BV/CSF index denote more atrophy.

**Statistical analyses**

One-way independent ANOVA and ANCOVA were performed to test between-cohort differences. All p-values (two-sided) were adjusted using Bonferroni correction for multiple comparisons both across dependent variables and post-hoc paired comparisons. The Spearman's rank correlation was used to investigate relationships between variables. Multiple linear regression was performed to analyze the influence of demographic and clinical variables on

hippocampal volume. Random forest analysis and dominance analysis were performed to investigate the importance of the demographic and clinical variables in explaining differences between cohorts as well as variability in hippocampal volume, respectively (Breiman, 2001; Liaw and Wiener, 2002; Grömping, 2007). Importance is reported as *i* and reflects the relative error in classification when a predictor is excluded from the model (in random forest analysis) (Breiman, 2001; Liaw and Wiener, 2002), and the relative percentage of the variance of the regression model explained by a given predictor (in dominance analysis, multiple linear regression) (Grömping, 2007). The statistical design used for each of the analyses performed is detailed in Supplementary Table 2. Effect sizes are reported as partial eta squared ($\eta^2_{par}$) and standardized beta (β). Results were considered significant when $p \leq 0.05$.

## RESULTS

### Between-cohort differences in demographic and clinical variables

Significant between-cohort differences were found in all the studied variables (Table 3). Age, memory complaints, depressive symptoms and the BV/CSF index showed the greatest effect sizes ($\eta^2_{par} \geq 0.30$). Random forest analysis demonstrated that age was the most important variable in explaining differences between cohorts (*i*=526.3), followed by education level (*i*=190.4), TIV (*i*=148.3), the BV/CSF index (*i*=143.6), MMSE (*i*=77.3), and Gender (*i*=64.4). ADNI and AddNeuroMed recruited significantly older samples as compared with AIBL and most of the population-based cohorts. Both ADNI and AIBL recruited highly educated individuals, while AddNeuroMed was comparable to the population-based cohorts. BRC and GENIC showed the smallest TIV values while ADNI, BioFINDER and PIVUS showed the largest TIV values. Finally, the BV/CSF index had higher values (i.e. less atrophy) in the cohorts AIBL, GENIC and

SNAC-K. Lower values of the BV/CSF index (i.e. more atrophy) were observed in AddNeuroMed and PIVUS.

**Between-cohort differences in hippocampal volume**

Hippocampal volume was significantly larger in AIBL than in AddNeuroMed and ADNI, and was comparable with that displayed by GENIC and SNAC-K. BRC and PIVUS were the cohorts with smallest hippocampal volume, and BioFINDER was in between (Table 3, Fig. 1).

**Effect of demographic and clinical variables on hippocampal volume**

Multiple regression analyses performed for the whole sample showed that age and the BV/CSF index were the most important variables in explaining variability in hippocampal volume (Table 4a). Age and the BV/CSF index correlated with each other ($r=-0.575$; $p<0.001$). Moreover, both variables were not only important themselves but also in interaction with gender, education, MMSE, and TIV (Table 4b). Fig. 2 shows the interaction between age and the BV/CSF index (unstandardized beta = 0.975; $t_{(1897)}=11.307$; $p<0.001$): smaller hippocampal volume was associated with lower BV/CSF index (i.e. more atrophy) in the older participants, but not in the younger ones. We then repeated the same multiple regression model for each separate cohort, also including depressive symptoms as a predictor. The association between hippocampal volume and the different demographic and clinical variables was modulated by the cohort factor (Fig. 3). Patterns of association were largely the same across cohorts. Age, the BV/CSF index and TIV showed the largest standardized regression coefficients (absolute ß > 0.20), although education, MMSE, and depressive symptoms also showed significant associations especially in the population-based cohorts (absolute ß = 0.08 – 0.12). Finally, we performed new multiple regression models for each separate cohort, but this time including all the available demographic and clinical variables (predictors included on each model are specified in Supplementary Table 2

"Extended model"). Results indicated similar patterns as above but three new variables showed significant association with hippocampal volume. Higher scores in the CDR, presence of the APOE ε4 allele, and higher levels of CSF T-tau were significantly associated with smaller hippocampal volume (ß = -0.15; ß = -0.17; ß = -0.18, respectively).

**Between-cohort differences in hippocampal volume are largely explained by between-cohort differences in demographic and clinical variables**

An ANCOVA was performed to test between-cohort differences in hippocampal volume when accounting for the effect of age, gender, education, MMSE, TIV, and the BV/CSF index. Results showed a dramatic reduction in the effect size from 19% in the original ANOVA for between-cohort differences in hippocampal volume ($F_{(7, 1904)}$= 63.188; p<0.001) (Table 3) to 2% in this ANCOVA ($F_{(7, 1898)}$= 4.429; p<0.001). Moreover, most post-hoc comparisons became non-significant. AIBL and SNAC-K displayed the largest hippocampal volumes, significantly different from those found in BioFINDER and PIVUS. The effect sizes of the covariates showed that TIV ($\eta^2_{par}$=15%), the BV/CSF index ($\eta^2_{par}$=13%), and age ($\eta^2_{par}$=9%), had the strongest confounding effect. Fig. 4 shows how original between-cohort differences in hippocampal volume (blue line) are attenuated when controlling for the above-mentioned covariates (orange line).

**DISCUSSION**

In this study, age and global brain atrophy (i.e. BV/CSF index) were the most important variables in explaining variability in hippocampal volume, and were not only important themselves but also in interaction with gender, education, MMSE, and TIV. AddNeuroMed, ADNI and AIBL differed from population-based cohorts in key demographic and clinical

variables that were found to largely explain between-cohort differences in hippocampal volumes. Below we discuss several important implications of the findings as well as considerations for generalization of results from these highly selected samples to the general population.

Age was the most important variable in explaining differences between cohorts, followed by education level, TIV, and the BV/CSF index. Participants in ADNI and AddNeuroMed were older in comparison with AIBL and most of the population-based cohorts. This result is different from previous studies where the control groups from ADNI and AIBL were younger than those from the population-based cohorts of the Mayo Clinic Study of Aging (MCSA) and the Sydney Memory and Aging Study (MAS), respectively (Whitwell et al., 2012; Brodaty et al., 2014). Of note, because of the recent interest in studying the preclinical stage of AD (Sperling et al., 2011), as well as in conducting aging research from a lifespan perspective (Walhovd et al., 2014), especially focused on middle-age populations (Ferreira et al., 2015), there is a clear trend for contemporary aging studies to include younger cohorts than those of AddNeuroMed and ADNI.

As previously reported, education levels were found to be considerably higher in ADNI and AIBL compared to AddNeuroMed and most of the population-based cohorts (Whitwell et al., 2012; Brodaty et al., 2014). Education is one of the most frequently used proxies of cognitive reserve (Stern, 2009). Therefore, this finding has important implications since extensive evidence exist about the impact of cognitive reserve in both cognition and brain structure. It has previously been suggested that a large proportion of the ADNI controls could be on the path to AD dementia, although cognitive reserve mechanisms may have protected them from cognitive decline (Whitwell et al., 2012).

Regarding TIV, the cohorts of BRC and GENIC showed the smallest TIV values while ADNI, AIBL, BioFINDER and PIVUS showed the largest. Whitwell *et al.* (2012) (Whitwell et al.,

2012) did not find significant differences in TIV between healthy individuals in ADNI and MCSA. Larger TIV has previously been related to higher brain reserve and protection against AD pathology (Stern, 2009; Whitwell, 2010). The fact that individuals in the ADNI cohort have larger TIV in combination with higher level of education is thus of interest. A recent study showed that larger TIV attenuated the impact of brain atrophy on clinical disease progression in the MCI patients from ADNI (Guo et al., 2013).

The cohorts AIBL, GENIC and SNAC-K had less brain atrophy (i.e. higher BV/CSF index) compared to AddNeuroMed and PIVUS. We are not aware of previous studies comparing the control groups from AddNeuroMed, ADNI or AIBL versus population-based cohorts in terms of global brain atrophy. When looking at other imaging markers, previous studies have reported higher rates of hippocampal atrophy (Whitwell et al., 2012), reduced cortical volume (Nettiksimmons et al., 2010), and unusually high amyloid load (Jagust et al., 2009) in the ADNI cohort. In AIBL, GENIC and SNAC-K, less global brain atrophy (i.e. higher BV/CSF index) could be explained by the fact that these cohorts have the youngest participants. Low educational level, relatively old participants, and high prevalence of mild depression (40%) may explain why the subjects in AddNeuroMed and PIVUS had a lower BV/CSF index. Vascular comorbidity and vascular risk factors were also frequent in PIVUS (Lind et al., 2005). All these factors have previously been associated with brain atrophy (Raz et al., 2010; Noble et al., 2012; Janowitz et al., 2014; Jiang et al., 2014; Shpanskaya et al., 2014; Yuefeng et al., 2014).

We also found between-cohort differences in gender distribution, MMSE, CDR, presence of subjective memory complaints, depressive symptomatology and APOE ε4 distribution. The effect of these and the previously discussed factors of age, education, TIV and global brain atrophy on hippocampal volume is widely known. However, limited research has focused on

cognitively normal individuals (Yamaguchi et al., 2002; Morra et al., 2009; Crivello et al., 2010; Raz et al., 2010; Striepens et al., 2011; Noble et al., 2012; Janowitz et al., 2014; Yuefeng et al., 2014; Jiang et al., 2014; Shpanskaya et al., 2014; Voevodskaya et al., 2014), and negative results have also been reported (Yamaguchi et al., 2002; Morra et al., 2009; Striepens et al., 2011; Janowitz et al., 2014; Jiang et al., 2014; Shpanskaya et al., 2014; Yuefeng et al., 2014). An explanation for these conflicting results is that all previous studies have focused on one or a few of the factors at a time. Therefore, some of the missing factors could be exerting an unobserved confounding effect and partially drive the results.

Increasing evidence shows that some of these factors play a relevant role in disease progression (e.g. cognitive reserve (Sperling et al., 2011)), magnifying variability on rates of hippocampal decline along the stages of MCI and AD. Therefore, the influence of several key demographic and clinical factors on brain structure and disease progression may add something valuable to the explanation of different subtypes in AD.

The main contribution of the present study is the demonstration that once all these factors are simultaneously considered, age and global brain atrophy are the most important factors in explaining variability in hippocampal volume. These variables were not only important in themselves but also in interaction with gender, education, MMSE, and TIV. The fact that global brain atrophy (i.e. the BV/CSF index) strongly correlated with age indicates that reduced hippocampal volume in cognitively normal individuals seems to be primarily explained by a process of global brain atrophy, presumably age-related. This would be true only for the older individuals. The interaction obtained suggests that reduced hippocampal volume in younger individuals could be indicative of either preclinical pathological changes related to a certain neurodegenerative disease or simply premorbid small hippocampal volume.

A previous study also found differences between ADNI controls and a population-based cohort in key demographic and clinical variables (Whitwell et al., 2012). In that study, hippocampal volume was larger in the ADNI controls than in those from the MCSA cohort and these differences were no longer significant after matching the cohorts by age, gender, education, APOE ε4 genotype and MMSE. To the best of our knowledge, the control groups from AIBL and AddNeuroMed have not been previously compared with population-based cohorts in terms of hippocampal volume. Our results together with recent research (Whitwell et al., 2012; Brodaty et al., 2014) demonstrate that control groups from ADNI, AIBL and AddNeuroMed may not be representative of the general population.

This study has several strengths: (1) the use of the largest sample to date in a study of this kind (N=1958); (2) the inclusion and comparison of the three currently most widely used multi-center cohorts in dementia imaging research (i.e. AddNeuroMed, ADNI and AIBL); and (3) the interactive evaluation of several demographic and clinical variables associated with hippocampal volume, in contrast to virtually all previous studies frequently investigating these factors in a separate manner.

Some limitations should also be considered. Some variables were missing for some of the cohorts, limiting the consideration of several clinical variables when performing analyses at the whole sample level. This was addressed by running analyses for each separate cohort including all the available variables. Further, imaging data from different centers with different MRI equipment and sequences were used. This affects only those analyses where the different cohorts were combined or compared to each other, but not those analyses carried out for the separate samples. Moreover, most of the MRI sequences were designed to be comparable with the ADNI

protocol. Some factors could still have some influence in the imaging measurements, especially differences in field strength. Nonetheless, excellent agreement between hippocampal volumes measured across different field strengths has been previously demonstrated for FreeSurfer (Briellmann et al., 2001; Whitwell et al., 2012). Other studies have also compared cohorts where MRI data were acquired in different centers, equipment, and field strengths (Whitwell et al., 2012; Hibar et al., 2015). Still, we cannot know how much of the variance in hippocampal volume is due to scanners and protocol differences in this study. Finally, several life-style factors such as smoking and cardiometabolic factors have been previously identified as determinants of hippocampal volume in cognitively normal individuals (Janowitz et al., 2014) but were not considered in this study.

## CONCLUSION

This study may have important implications for the use of hippocampal atrophy as a biomarker. The results highlight the large variability in hippocampal volume during the cognitively normal stage. This is important when trying to disentangle disease mechanisms from the effect of several demographic and clinical factors. Another important conclusion is that the samples of AddNeuroMed, ADNI and AIBL are not representative of the general population. Both conclusions must be taken into account when (1) designing research where a clinical group is recruited from a specialized center and compared with a control group from the general population; (2) designing future clinical trials, which are often based on highly selected populations; (3) translating imaging biomarkers to the general population; and (4) applying the new diagnostic criteria for AD and predementia stages, in which imaging biomarkers are important for increasing certainty about the underlying disease (Albert et al., 2011; McKhann et al., 2011; Sperling et al., 2011). The key factor here is how well knowledge can be translated

from clinical settings to the general population, and vice versa. The strategy for recruiting individuals will depend on the study aims and the materials to be used. Regarding hippocampal atrophy as a biomarker, our findings show that age and global brain atrophy are the most important factors to be considered, but gender, education, MMSE, and TIV should also be taken into account.

## REFERENCES


Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7:270–279.

Breiman L. 2001. Random forests. Mach Learn 45:5–32.

Briellmann RS, Syngeniotis A, Jackson GD. 2001. Comparison of Hippocampal Volumetry at 1.5 Tesla and at 3 Tesla. Epilepsia 42:1021–1024.

Brodaty H, Mothakunnel A, de Vel-Palumbo M, Ames D, Ellis KA, Reppermund S, Kochan NA, Savage G, Trollor JN, Crawford J, Sachdev PS. 2014. Influence of population versus convenience sampling on sample characteristics in studies of cognitive aging. Ann Epidemiol 24:63–71.

Crivello F, Lemaître H, Dufouil C, Grassiot B, Delcroix N, Tzourio-Mazoyer N, Tzourio C, Mazoyer B. 2010. Effects of ApoE-epsilon4 allele load and age on the rates of grey matter and hippocampal volumes loss in a longitudinal cohort of 1186 healthy elderly persons. Neuroimage 53:1064–1069.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany R. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968 – 980.

Ellis KA, Rowe CC, Villemagne VL, Martins RN, Masters CL, Salvado O, Szoeke C, Ames D. 2010. Addressing population aging and Alzheimer's disease through the Australian imaging biomarkers and lifestyle study: collaboration with the Alzheimer's Disease Neuroimaging Initiative. Alzheimers Dement 6:291–296.

Ferencz B, Laukka EJ, Lövdén M, Kalpouzos G, Keller L, Graff C, Wahlund LO, Fratiglioni L, Bäckman L. 2013. The influence of APOE and TOMM40 polymorphisms on hippocampal volume and episodic memory in old age. Front Hum Neurosci 7:198.

Ferreira D, Correia R, Nieto A, Machado A, Molina Y, Barroso J. 2015. Cognitive decline before the age of 50 can be detected with sensitive cognitive measures. Psicothema 27:216–222.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, Van Der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355.

Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM. 2010. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 6:67–77.

Grömping U. 2007. Estimators of relative importance in linear regression based on variance decomposition. Am Stat 61:139–147.

Guo L-H, Alexopoulos P, Wagenpfeil S, Kurz A, Perneczky R. 2013. Brain size and the compensation of Alzheimer's disease symptoms: a longitudinal cohort study. Alzheimers Dement 9:580–586.

Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivières S, Jahanshad N, Toro R, Wittfeld K, Abramovic L, Andersson M, Aribisala BS, Armstrong NJ, Bernard M, Bohlken MM, Boks MP, Bralten J, Brown AA, Mallar Chakravarty M, Chen Q, Ching CRK, Cuellar-Partida G, den Braber A, Giddaluru S, Goldman AL, Grimm O, Guadalupe T, Hass J, Woldehawariat G, Holmes AJ, Hoogman M, Janowitz D, Jia T, Kim S, Klein M, Kraemer B, Lee PH, Olde Loohuis LM, Luciano M, Macare C, Mather KA, Mattheisen M, Milaneschi Y, Nho K, Papmeyer M, Ramasamy A, Risacher SL, Roiz-Santiañez R, Rose EJ, Salami A, Sämann PG, Schmaal L, Schork AJ, Shin J, Strike LT, Teumer A, van Donkelaar MMJ, van Eijk KR, Walters RK, Westlye LT, Whelan CD, Winkler AM, Zwiers MP, Alhusaini S, Athanasiu L, Ehrlich S, Hakobjan MMH, Hartberg CB, Haukvik UK, Heister AJGAM, Hoehn D, Kasperaviciute D, Liewald DCM, Lopez LM, Makkinje RRR, Matarin M, Naber MAM, Reese McKay D, Needham M, Nugent AC, Pütz B, Royle NA, Shen L, Sprooten E, Trabzuni D, van der Marel SSL, van Hulzen KJE, Walton E, Wolf C, Almasy L, Ames D, Arepalli S, Assareh AA, Bastin ME, Brodaty H, Bulayeva KB, Carless MA, Cichon S, Corvin A, et al. 2015. Common genetic variants influence human subcortical brain structures. Nature 520: 224–229.

Jagust WJ, Landau SM, Shaw LM, Trojanowski JQ, Koeppe RA, Reiman EM, Foster NL, Petersen RC, Weiner MW, Price JC, Mathis CA. 2009. Relationships between biomarkers in aging and dementia. Neurology 73:1193–1199.

Janowitz D, Schwahn C, Borchardt U, Wittfeld K, Schulz A, Barnow S, Biffar R, Hoffmann W, Habes M, Homuth G, Nauck M, Hegenscheid K, Lotze M, Völzke H, Freyberger HJ, Debette S, Grabe HJ. 2014. Genetic, psychosocial and clinical factors associated with hippocampal volume in the general population. Transl Psychiatry 4:e465.

Jiang J, Sachdev P, Lipnicki DM, Zhang H, Liu T, Zhu W, Suo C, Zhuang L, Crawford J, Reppermund S, Trollor J, Brodaty H, Wen W. 2014. A longitudinal study of brain atrophy over two years in community-dwelling older individuals. Neuroimage 86:203–211.

Liaw A, Wiener M. 2002. Classification and regression by random forest. R news 2:18–22.

Lind L, Fors N, Hall J, Marttala K, Stenborg A. 2005. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. Arterioscler Thromb Vasc Biol 25:2368–2375.

Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, Spenger C, Tsolaki M, Vellas B, Wahlund L-O, Ward M. 2009. AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Ann N Y Acad Sci 1180:36–46.

Magnil M, Janmarker L, Gunnarsson R, Björkelund C. 2013. Course, risk factors, and prognostic factors in elderly primary care patients with mild depression: a two-year observational study. Scand J Prim Health Care 31:20–25.

McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. 2011. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7:263–269.

Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Toga AW, Jack CR, Schuff N, Weiner MW, Thompson PM. 2009. Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. Neuroimage 45:S3-S15.

Morris JC, Blennow K, Froelich L, Nordberg A, Soininen H, Waldemar G, Wahlund L-O, Dubois B. 2014. Harmonized diagnostic criteria for Alzheimer's disease: recommendations. J Intern Med 275:204–213.

Muehlboeck J-S, Westman E, Simmons A. 2014. TheHiveDB image data management and analysis

framework. Front Neuroinform 7:49.

Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L. 2005. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement 1:55–66.

Nettiksimmons J, Harvey D, Brewer J, Carmichael O, DeCarli C, Jack CJ, Petersen R, Shaw L, Trojanowski J, Weiner M, Beckett L. 2010. Subtypes based on CSF and MRI markers in normal elderly predict cognitive decline. Neurobiol Aging 31:1419–1428.

Noble KG, Grieve SM, Korgaonkar MS, Engelhardt LE, Griffith EY, Williams LM, Brickman AM. 2012. Hippocampal volume varies with educational attainment across the life-span. Front Hum Neurosci 6:307.

Orellana C, Ferreira D, Muehlboeck J-S, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Soininen H, Lovestone S, Simmons A, Wahlund L-O, Westman E. 2016. Measuring global brain atrophy with the brain volume / cerebrospinal fluid index: normative values, cut-offs and clinical associations. Neurodegener Dis 16:77–86.

Qiu C, Zhang Y, Bronge L, Herlitz A, Aspelin P, Bäckman L, Fratiglioni L, Wahlund LO. 2012. Medial temporal lobe is vulnerable to vascular risk factors in men: a population-based study. Eur J Neurol 19:876–883.

Raz N, Ghisletta P, Rodrigue KM, Kennedy KM, Lindenberger U. 2010. Trajectories of brain aging in middle-aged and older adults: regional and individual differences. Neuroimage 51:501–511.

Shpanskaya K, Choudhury K, Hostage CJ, Murphy K, Petrella J, Doraiswamy P. 2014. Educational attainment and hippocampal atrophy in the Alzheimer's disease neuroimaging initiative cohort. J Neuroradiol 41:350–357.

Sperling R, Aisen P, Beckett L, Bennett D, Craft S, Fagan A, Iwatsubo T, Jack CJ, Kaye J, Montine T, Park D, Reiman E, Rowe C, Siemers E, Stern Y, Yaffe K, Carrillo M, Thies B, Morrison-Bogorad M, Wagster M, Phelps C. 2011. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7:280–292.

Stern Y. 2009. Cognitive reserve. Neuropsychologia 47:2015–2028.

Striepens N, Scheef L, Wind A, Meiberth D, Popp J, Spottke A, Kölsch H, Wagner M, Jessen F. 2011. Interaction effects of subjective memory impairment and ApoE4 genotype on episodic memory and hippocampal volume. Psychol Med 41:1997–2006.

Svanborg P, Ekselius L. 2003. Self-assessment of DSM-IV criteria for major depression in psychiatric out- and inpatients. Nord J Psychiatry 57:291–296.

Voevodskaya O, Simmons A, Nordenskjöld R, Kullberg J, Ahlström H, Lind L, Wahlund L-O, Larsson E-M, Westman E. 2014. The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease. Front Aging Neurosci 6:264.

Walhovd KB, Fjell AM, Espeseth T. 2014. Cognitive decline and brain pathology in aging--need for a dimensional, lifespan and systems vulnerability view. Scand J Psychol 55:244–254.

Wancata J, Alexandrowicz R, Marquart B, Weiss M, Friedrich F. 2006. The criterion validity of the geriatric depression scale: a systematic review. Acta Psychiatr Scand 114:398–410.

Wang R, Fratiglioni L, Laveskog A, Kalpouzos G, Ehrenkrona CH, Zhang Y, Bronge L, Wahlund LO, Bäckman L, Qiu C. 2014. Do cardiovascular risk factors explain the link between white matter hyperintensities and brain volumes in old age? A population-based study. Eur J Neurol 21:1076–1082.

Whitwell JL. 2010. The protective role of brain size in Alzheimer's disease. Expert Rev Neurother 10:1799–1801.

Whitwell JL, Wiste HJ, Weigand SD, Rocca WA, Knopman DS, Roberts RO, Boeve BF, Petersen RC, Jack CR. 2012. Comparison of imaging biomarkers in the Alzheimer Disease Neuroimaging

Initiative and the Mayo Clinic Study of Aging. Arch Neurol 69:614–622.

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, Nordberg A, Bäckman L, Albert M, Almkvist O, Arai H, Basun H, Blennow K, de Leon M, DeCarli C, Erkinjuntti T, Giacobini E, Graff C, Hardy J, Jack C, Jorm A, Ritchie K, van Duijn C, Visser P, Petersen RC. 2004. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 256:240-246.

Yamaguchi S, Meguro K, Shimada M, Ishizaki J, Yamadori A, Sekita Y. 2002. Five-year retrospective changes in hippocampal atrophy and cognitive screening test performances in very mild Alzheimer's disease: the Tajiri Project. Neuroradiology 44:43–48.

Yuefeng L, Jinchuan Y, Dongqing W, Meifang S, Yan Z, Xiaolan Z, Ping J, Ruigen Y, Liang Z. 2014. Magnetic resonance study of the structure and function of the hippocampus and amygdala in patients with depression. Chin Med J 127:3610–3615.

Zhang Y, Qiu C, Lindberg O, Bronge L, Aspelin P, Bäckman L, Fratiglioni L, Wahlund LO. 2010. Acceleration of hippocampal atrophy in a non-demented elderly population: the SNAC-K study. Int Psychogeriatr 22:14–25.

Zigmond AS, Snaith RP. 1983. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 67:361–370.

## ACKNOWLEDGEMENTS

The authors have no competing financial interests.

The authors thank Swedish Brain Power; the Strategic Research Programme in Neuroscience at Karolinska Institutet (StratNeuro); the Strategic Research Programme in Neuroscience at Lund University (MultiPark); European Research Council; Swedish Research Council; the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet as well as between Region Skåne and Lund University; Demensförbundet; Karolinska Institutet Forskningsstiftelser; Sigurd och Elsa Goljes Minne; Stiftelsen för Ålderssjukdomar vid Karolinska Institutet; Gamla Tjänarinnor; and Loo och Hans Ostermans Stiftelse.

Some results presented in this article have been partially published in Neurodegener Dis 2015, 15 (suppl 1): 183, as an abstract of an Oral Communication presented at the 12th International Conference on Alzheimer's & Parkinson's Diseases (ADPD), Nice (France), March 2014, being also awarded with the Junior Faculty Award.

*ADDNEUROMED:* AddNeuromed is supported by InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework programme priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics and Biotechnology for Health.

*ADNI:* ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and several generous contributions from the following: Alzheimer's Association, Alzheimer's Drug Discovery Foundation, BioClinica Inc., Biogen Idec Inc., Bristol-Myers Squibb Co., Eisai, Inc., Elan Pharmaceuticals, Inc., Eli Lilly and Co., F. Hoffmann-La Roche Ltd and its affiliated company Genentech Inc., GE

Healthcare, Innogenetics NV, IXICO Ltd, Janssen Alzheimer Immunotherapy Research & Development LLC, Johnson & Johnson Pharmaceutical Research & Development LLC, Medpace Inc., Merck & Co. Inc. Meso Scale Diagnostics LLC, NeuroRx Research, Novartis Pharmaceuticals Corp., Pfizer Inc., Piramal Imaging, Servier, Synarc Inc. and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (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. Data collection and sharing for in ADNI were funded by the National Institutes of Health (NIH) (Grant U01 AG024904). 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.

*AIBL:* The AIBL Research Group thanks all the participants who took part in the AIBL study and the clinicians who referred participants. The AIBL study (www.AIBL.csiro.au) is a collaboration between CSIRO, Edith Cowan University (ECU), The Florey Institute of Neuroscience and Mental Health (FINMH), National Ageing Research Institute (NARI) and Austin Health. It also involves support from CogState Ltd., Hollywood Private Hospital, and Sir Charles Gairdner Hospital. The study received funding support from CSIRO, the Science and Industry Endowment Fund (www.SIEF.org.au), NHMRC and Dementia Collaborative Research Centres (DCRC), Alzheimer's Australia (AA), Alzheimer's Association and the McCusker Alzheimer's Research Foundation.

*BRC:* Data acquisition and analysis were supported by funds from National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health and the NIHR Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, Kings College London. Co-authors A.S. and S.L. were also supported by funds from these institutions.

*GENIC:* Thanks to *Fundación Cajacanarias* (Canary Islands, Spain); *Servicio de Resonancia Magnética para Investigaciones Biomédicas del SEGAI* (University of La Laguna, Spain); Dr. Diaz-Flores Varela (Hospital Universitario de Canarias) for his collaboration in clinical inspection of magnetic resonance images; and Dr. Antonio Rodríguez for providing access to participants and helpful assistance. Data used in preparation of this article is part of the GENIC-database (Group of Neuropsychological Studies of the Canary Islands, University of La Laguna, Spain. Principal investigator: Professor José Barroso). The following collaborators contributed to the GENIC-database but did not participate in analysis or writing of this report (in alphabetic order by family name): Nira Cedrés, Rut Correia, Patricia Díaz, Nerea Figueroa, Iván Galtier, Lisset González, Teodoro González, Zaira González, Cathaysa Hernández, Edith Hernández, Nira Jiménez, Judith López, Antonieta Nieto, María Sabucedo, Elena Sirumal, Marta Suárez, Manuel Urbano, Pedro Velasco.

*BioFINDER:* The study was supported by the European Research Council, the Swedish Research Council, the Strategic Research Programme in Neuroscience at Lund University (MultiPark), the Crafoord Foundation, the Swedish Brain Foundation, The Swedish Alzheimer foundation, the Torsten Söderberg Foundation at the Royal

Swedish Academy of Sciences, and the regional agreement on medical training and clinical research (ALF) between Region Skåne and Lund University.

*PIVUS:* PIVUS was supported by AstraZeneca R&D Mölndal and Uppsala University Hospital.

*SNAC-K:* The Swedish National study on Aging and Care, SNAC (www.snac.org) is financially supported by the Ministry of Health and Social Affairs, Sweden, the participating County Councils and Municipalities, and the Swedish Research Council. We are grateful to the SNAC-K participants and to our colleagues in the SNAC-K group for their collaboration in data collection and management.



## FIGURE LEGENDS

**Figure 1. Between-cohort differences in hippocampal volume.** Hippocampal volume was calculated by summing left and right sides and values are expressed in cubic millimeters. Values represent median and confidence intervals. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen.

**Figure 2. Effect of demographic and clinical variables on hippocampal volume (whole sample): significant interaction between age and the BV/CSF index** (Multiple linear regression. Interaction between the effects of global brain atrophy (i.e. BV/CSF index) and age on hippocampal volume: unstandardized beta = 0.975; $t_{(1897)}$=11.307; two-sided p<0.001. N = 1912). Hippocampal volume was calculated by summing left and right sides. The BV/CSF index

was studied as a proxy of global brain atrophy. Two groups were created by separating the upper bound of age (old) versus the lower bound of age (young). Y-axis represents raw hippocampal volume in $mm^3$ and x-axis represent mean centered values of the BV/CSF index.

**Figure 3. Effect of demographic and clinical variables on hippocampal volume (separately by cohorts).** The figure schematizes the results from the multiple regression models (backwards) performed separately for the eight study cohorts. Only predictors remaining in the final models are displayed (criterion for excluding predictors from the models: two-sided $p<0.10$). Predictors included in the original models as well as sample size can be consulted in Supplementary Table 2. P-values of the primary regression models were adjusted using Bonferroni correction for multiple comparisons. Gender (0 female; 1 male); Education Level (0 low; 1 high). High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education. Depression was measured with GDS (AddNeuroMed, ADNI, AIBL, BRC, and GENIC), HADS (BioFINDER), and MADRS (SNAC-K). The BV/CSF index was studied as a proxy of global brain atrophy. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; MMSE = mini-mental state examination; TIV = total intracranial volume.

**Figure 4. Attenuation of between-cohort differences in hippocampal volume when controlling for demographic and clinical variables** (ANOVA: $F_{(7,\ 1904)}= 63.188$; two-sided $p<0.001$; $\eta^2_{par}= 19\%$; ANCOVA: $F_{(7,\ 1898)}= 4.429$; two-sided $p<0.001$; $\eta^2_{par}= 2\%$). The original

ANOVA (N = 1958) was repeated for the same sample size than the ANCOVA (N = 1912) to allow perfect comparability of the results. Hippocampal volume was calculated by summing left and right sides and values are expressed in cubic millimeters. Covariates included in the ANCOVA model are age, gender (0 female; 1 male); education level (0 low; 1 high), MMSE, TIV and the BV/CSF index. High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education. The BV/CSF index was studied as a proxy of global brain atrophy. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; MMSE = mini-mental state examination; TIV = total intracranial volume.

Table 1. Description of cohorts' characteristics

| | AddNeuroMed | ADNI | AIBL | BRC | GENIC | BioFINDER | PIVUS | SNAC-K |
|---|---|---|---|---|---|---|---|---|
| **Full name** | - | Alzheimer's Disease Neuroimaging Initiative | Australian Imaging Biomarkers and Lifestyle study | King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort | Group of Neuropsychological Studies from the Canary Islands | Swedish BioFINDER study | Prospective Investigation of the Vasculature in Uppsala Seniors | Swedish National Study on Aging and Care in Kungsholmen |
| **Type** | Multi-centre Case-control (AD, MCI, CTRL) | Multi-centre Case-control (AD, MCI, CTRL) | Multi-centre Case-control (AD, MCI, CTRL) | Single-centre Population-based specifically including AD, MCI and CTRL groups | Single-centre Population-based specifically including MCI and CTRL groups | Single-centre Population-based | Single-centre Population-based | Single-centre Population-based |
| **Country** | Finland, France, Greece, Italy, Poland, and UK | USA and Canada | Australia | UK (London) | Spain (Canary Islands) | Sweden (Malmö) | Sweden (Uppsala) | Sweden (Stockholm) |
| **Design** | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal | Prospective longitudinal |
| **Period** | 2004-present | 2003-present | 2006-present | 2007-present | 2005-present | 2009-present | 2001-present | 2001-present |
| **Brief project description** | Cross European study, part of the InnoMed | Launched by the National Institute on Aging, the National | Launched by the Australian CSIRO and a number of | BRC is a neuroimaging study which was designed | Study from the University of La Laguna, aimed to | Study from Lund University, aimed to investigate the | Study from the University of Uppsala, aimed to | Carried out by the Stockholm Gerontology |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (Innovative Medicines in Europe), funded by the European Union (FP6 and FP7), as well as members of the EFPIA. Designed to find biomarkers or tests for AD | Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and non-profit organizations. Established to develop standardized imaging techniques and biomarker procedures in AD research | leading Australian research organizations. A study to discover which biomarkers, cognitive characteristics, and health and lifestyle factors determine subsequent development of symptomatic AD | to establish imaging markers for the earlier detection and diagnosis of AD. Data was collected at the Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London and South London and Maudsley NHS Foundation Trust. London, UK | investigate the cognitive and imaging profile associated to normal aging from the middle-age adulthood to the old age. Individuals living in the Canary Islands | preclinical stages of Alzheimer's disease and other dementia disorders. Further aims are to study midlife risk factors for future development of amyloid, tau and vascular pathologies | investigate the predictive power of different measurements of endothelial function and arterial compliance in a random sample of 1000 normal elderly individuals living in the community of Uppsala | Research Centre, the ARC at Karolinska Institutet and the Stockholm University, to detect the influence of lifetime genetic, environmental and biological factors on medical, psychological and social health in late adulthood. Individuals living in the area of Kungsholmen/Essin geöarna (Stockholm) |
| **Recruitment procedure** | Non-related members of the patient's families, caregiver's relatives and social centres | Advertisements in newspapers | Medial appeal and through participants' treating physicians | Non-related members of the patient's families, caregivers, relatives and social centres | Relatives and acquaintances of the research staff and students from the University of La | Random sampling from the cardiovascular cohort of the population-based | Random sample of 1000 subjects aged 70 years old at baseline and chosen from the register of | Random epidemiological sample |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | for the elderly or GP surgeries | | | for the elderly, GP surgeries or advertisements in newspapers | Laguna, personnel from local schools, and through participants' GPs | Malmö Diet and Cancer Study conducted in the city of Malmö, Sweden | community living | |
| **Data obtained from** | Eric Westman and Andrew Simmons are part of AddNeuroMed | https://ida.loni.usc.edu, PI Michael M. Weiner | https://ida.loni.usc.edu and https://aibl.csiro.au (EoI), PI David Ames | Personal contact to PI Simon Lovestone and Andy Simmons | Personal contact to PI José Barroso | Personal contact to PI Oskar Hansson | Personal contact to PI Lars Lind | Personal contact to PI Laura Fratiglioni |
| **Key references, other sources** | Lovestone et al. (2009) www.innomed-addneuromed.com | Mueller et al. (2005) www.adni-info.org | Ellis et al. (2009) www.aibl.csiro.au | - | Ferreira et al. (2015) | www.biofinder.se | Lind et al. (2005) www.medsci.uu.se/pivus | Zhang et al. (2010) www.snac-k.se |

AD = Alzheimer's disease; MCI = mild cognitive impairment; CTRL = control; UK = United Kingdom; FP = Framework Programme; EFPIA = European Federation for Pharmaceutical Industries and Associations; USA = United States of America; PI = principal investigator; CSIRO = Commonwealth Scientific Industrial and Research Organization; EoI = Expression of Interest; ARC = Aging Research Centre.

Table 2. Eligibility criteria

| | AddNeuroMed | ADNI | AIBL | BRC | GENIC | BioFINDER | PIVUS | SNAC-K |
|---|---|---|---|---|---|---|---|---|
| Inclusion Criteria | | | | | | | | |
| Age | ≥ 65 years | 55 – 90 years | ≥ 60 years | ≥ 60 years | ≥ 35 years | ≥ 60 years | All individuals are 70 years old at baseline (Data at five years follow-up when MRI is collected is included in the current study) | ≥ 60 years |
| MMSE | 24 – 30 | 24 – 30 (exceptions for subjects with less than 8 years of education) | 24 – 30 | 24 – 30 | 24 – 30 | 28-30 (at screening visit) | - | 24 – 30 [1] |
| CDR | 0 | 0 | 0 or 0.5 | 0 | - | 0 | - | - |
| Depression | GDS ≤ 5 | GDS ≤ 5 | GDS ≤ 5 | GDS ≤ 5 | No clinical diagnosis of depression, no antidepressant drugs | No current clinical major depressive episode | - | No clinical diagnosis of depression (DSM-IV, ICD-10) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cognition | Preserved cognitive and functional status | Normal education-adjusted performance in Logical Memory II (WMS-R) | Generally, normal education-adjusted performance in Logical Memory II (WMS-R), although some participants demonstrating failure were further investigated and included | Preserved cognitive and functional status | Preserved cognitive and functional status | Preserved cognitive and functional status | Normal elderly individuals | Non-demented individuals |
| Subjective complaints | A percentage of the participants expressed subjective concern about their memory function or thinking capacity. Subjective memory complaints were elicited by the question: "Do you have problems with memory or | No memory complaints aside from those common to other normal subjects | A percentage of the participants expressed subjective concern about their memory function. Subjective memory complaints were elicited by the question: "Do you have difficulties with your memory?" | A percentage of the participants expressed subjective concern about their memory function or thinking capacity. Subjective memory complaints were elicited by the question: "Do you have problems with memory or | A percentage of the participants expressed subjective concern about their cognitive functions. Subjective memory complaints were elicited by the question: "Do you have difficulties with your memory?" | No memory complaints aside from those common to other normal subjects. Subjective memory complaints were elicited by the question: "Do you have problems with memory or thinking?" | - | A percentage of the participants expressed subjective concern about their memory function. Subjective memory complaints were elicited by the question: "Do you think your memory has got worse?" |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | thinking?" | | | thinking?" | | | | |
| Functional activity | Absence of significant impairment in activities of daily living | Absence of significant impairment in activities of daily living | Intact social and occupational functioning | Absence of significant impairment in activities of daily living | FAQ ≤5 | Absence of significant impairment in activities of daily living | Normal elderly subjects | Non-demented, non-institutionalized and non-disabled |
| Medication | Stable medication | Stable medication | - | Stable medication | - | Stable medication | - | - |
| Health | Good general health | Good general health | - | Good general health | Good general health | Good general health | Vascular comorbidity and vascular risk factors are not excluded | Healthy elderly people |
| Exclusion criteria | | | | | | | | |
| Cognitive impairment | Dementia (DSM-IV criteria), MCI | Dementia (DSM-IV criteria), MCI | Dementia (DSM-IV, and ICD-10 criteria), MCI (Winblad et al., 2004) | Dementia (DSM-IV criteria), MCI | Dementia (DSM-IV criteria), MCI (Winblad et al., 2004) | Dementia (DSM-IV, and ICD-10 criteria), MCI (Winblad et al., 2004) | - | Dementia (DSM-IV criteria) |
| Disease | Significant neurological or psychiatric illness, significant unstable systemic illness or | Significant neurological or psychiatric illness, significant unstable systemic illness or | Significant neurological or psychiatric illness, cancer (except basal cell skin carcinoma), | Significant neurological or psychiatric illness, significant unstable systemic illness or | Significant neurological or psychiatric illness, significant systemic illness or organ | Significant neurological or psychiatric illness, significant unstable systemic illness or | - | Significant neurological or psychiatric illness, significant systemic illness or organ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | organ failure | organ failure | symptomatic stroke, and uncontrolled diabetes | organ failure | failure | organ failure | | failure [2] |
| Medication | - | Current use of psychoactive medications or warfarin | - | - | Current use of psychoactive medications | Use of anti-psychotic or sedative medications | - | - |
| Substance abuse | Alcohol or substance misuse | History of alcohol or substance abuse or dependence | Current regular alcohol use exceeding two standard drinks per day for women or four per day for men | Alcohol or substance misuse | History of alcohol or substance abuse or dependence | Current alcohol or substance misuse | - | Current alcohol or substance misuse [2] |

ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; MMSE = Mini-Mental State Examination; CDR = Clinical Dementia Rating; GDS = Geriatric Depression Scale; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke - the Alzheimer's Disease and Related Disorders Association; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders-Fourth edition; MCI = mild cognitive impairment; WMS-R = Wechsler Memory Scale-Revised edition; FAQ = Functional Activity Questionnaire; TIA = transitory ischemic attack.

[1] The SNAC-K is a population-based study that recruited healthy elderly people without dementia living in the area of Kungsholmen/Essingeöarna (Stockholm). Some participants turned out to have MMSE scores below 24. For the purposes of the current study, selection criteria based on previous SNAC-K studies (Qiu et al., 2012; Ferencz et al., 2013; Wang et al., 2014) were applied.

[2] The SNAC-K is a population-based study that recruited healthy elderly people without dementia living in the area of Kungsholmen/Essingeöarna (Stockholm). Some participants turned out to have medical conditions associated with cognitive impairment. For the purposes of the current study, selection criteria based on previous SNAC-K studies (Ferencz et al., 2013) were applied. In particular, the following participants were excluded because of psychiatric disorder (n = 27), neurological disease (n = 11), MMSE < 24 (n = 2), alcohol dependence syndrome (n = 3). Moreover, 49 further participants were excluded due to suboptimal MRI image quality.

Table 3. Demographic variables, clinical variables and hippocampal volume

| | AddNeuroMed | ADNI | AIBL | BRC | GENIC | BioFINDER | PIVUS | SNAC-K | p | $\eta^2_{par}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample size, n | 100 | 226 | 155 | 59 | 295 | 279 | 398 | 446 | - | - |
| Age, mean (SD) | 74.2 (5.6) | 76.0 (5.1) | 72.8 (7.1) | 77.2 (5.8) | 54.2 (10.7) | 73.2 (5.1) | 75.0 (0.0) | 70.2 (8.9) | <0.001 | 0.52 |
| Age, range | 61 – 88 | 60 – 90 | 60 – 88 | 67 – 91 | 35 – 83 | 65 – 87 | 75 | 60 – 96 | - | - |
| Gender, % female | 57% | 48% | 52% | 59% | 54% | 61% | 47% | 40% | 0.002 | 0.01 |
| Education, years | 10.1 (4.5) | 16.1 (2.8) | - | 13.0 (3.4) | - | 11.9 (3.7) | - | 12.5 (4.5) | <0.001 | 0.17 |
| Education, % high level [a] | 52% | 98% | 93% | 88% | 58% | 72% | 43% | 73% | <0.001 | 0.15 |
| MMSE, mean (SD) | 29.0 (1.3) | 29.1 (1.0) | 28.7 (1.2) | 29.1 (1.1) | 28.8 (1.3) | 29.0 (0.9) | 28.7 (1.4) | 29.1 (1.0) | <0.001 | 0.03 |
| MMSE, range | 25 – 30 | 25 – 30 | 25 – 30 | 26 – 30 | 24 – 30 | 27 – 30 | 21 – 30 | 25 – 30 | - | - |
| Memory complaints, % | 25% | - | 56% | 44% | 27% | 0% | - | 72% | <0.001 | 0.32 |
| CDR, score (%) | 0 (98%). 0.5 (2%) | 0 (100%) | 0 (94%). 0.5 (6%) | 0 (100%) | - | 0 (100%) | - | 0 (0%) | <0.001 | 0.04 |
| Activities of daily living | - | ***FAQ:*** 0.1 (0.6) | - | - | ***FAQ:*** 0.4 (0.8) [c] | - | - | - | <0.001 | 0.03 |
| | - | - | - | - | - | - | - | ***KATZ:*** 0.0 (0.0) | - | - |
| Depressive symptomatology | ***GDS:*** 5.5 (1.1) | ***GDS:*** 0.9 (1.2) | ***GDS:*** 1.0 (1.3) | ***GDS:*** 1.7 (2.1) | ***GDS:*** 2.4 (2.2) [d] | - | - | - | <0.001 | 0.43 |
| | - | - | ***HADS:*** 2.6 (2.2) | - | - | ***HADS:*** 2.0 (2.3) | - | - | 0.008 | 0.02 |
| | - | - | - | - | - | ***CSDD:*** 1.0 (2.3) | - | - | - | - |
| | - | - | - | - | - | - | - | ***MADRS:*** 1.8 (2.6) | | |
| Mild depression, % [b] | 40% | 0% | 4% [e] | 7% | 10% | 3% [e] | - | 1% | <0.001 | 0.17 |
| APOE ε4, % carriers | 33% | 27% | 42% | 33% | - | 29% | - | 26% | 0.010 | 0.01 |
| TIV, $dm^3$ | 1.48 (0.14) | 1.53 (0.17) | 1.53 (0.16) | 1.46 (0.19) | 1.46 (0.16) | 1.56 (0.16) | 1.55 (0.16) | 1.52 (0.25) | <0.001 | 0.03 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Global brain atrophy | 23.5 (11.1) | 24.7 (11.5) | 28.2 (12.8) | 25.0 (10.6) | 47.3 (19.2) | 24.5 (10.7) | 21.2 (8.5) | 31.6 (14.1) | <0.001 | 0.30 |
| Hippocampal volume, $mm^3$ | 7013 (865) | 7118 (888) | 7492 (890) | 6803 (903) | 8244 (1036) | 7133 (945) | 6929 (764) | 7504 (968) | <0.001 | 0.19 |

ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort ; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; SD = standard deviation; MMSE = mini-mental state examination; CDR = clinical dementia rating; ADL = activities of daily living; FAQ = functional activities questionnaire; KATZ = Katz Index of independence in activities of daily living; GDS = geriatric depression scale; HADS = hospital anxiety and depression scale; CSDD = Cornell scale for depression in dementia; MADRS = Montgomery-Asberg depression rating scale; APOE ε4 = apolipoprotein E allele ε4; TIV = total intracranial volume; $dm^3$ = cubic decimetres; $mm^3$ = cubic millimetres. The BV/CSF index was studied as a proxy of global brain atrophy. Hippocampal volume was calculated by summing left and right sides. Bonferroni correction for multiple comparisons (p-value ≤ 0.003). [a] Years of education was dichotomized in high and low level to allow comparability between cohorts providing years of education and those only providing levels of education. High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education; [b] Clinical cut-offs used for determining mild depression are the ones originally proposed in the scales: GDS ≥6 (Wancata et al., 2006); HADS ≥8 (Zigmond and Snaith, 1983), and MADRS ≥13 (Svanborg and Ekselius, 2003; Magnil et al., 2013); [c] FAQ in GENIC (11 items) was converted to the same scale than FAQ in ADNI (10 items) (conversion factor = 0.909); [d] GDS values for 158 individuals were estimated from BDI by applying the following formula: BDI_to_GDS = (BDI z score * SD of GDS) + mean of GDS. SD and mean of GDS were calculated from the same population (N=134); [e] Percentage obtained from the HADS scale.

Table 4. Effect of demographic and clinical variables on hippocampal volume (whole sample)

**(a) Importance**

**Multiple regression model:**

$R^2$= 53%; $F_{(6, 1905)}$=355.747; p<0.001

| Predictors (X) | β | p | *i* |
|---|---|---|---|
| Age | -0.363 | <0.001 | 22% |
| BV/CSF index | 0.384 | <0.001 | 18% |
| TIV | 0.369 | <0.001 | 9% |
| Gender | 0.056 | 0.004 | 3% |
| Education level | 0.058 | <0.001 | 1% |
| MMSE | 0.055 | <0.001 | 1% |

**(b) Interactions**

**Multiple regression model:**

$R^2$= 57%; $F_{(14, 1897)}$=180.710; p<0.001

| Predictors (X) | Unstandardized beta | T value | p |
|---|---|---|---|
| Age | -42.083 | -20.135 | <0.001 |
| BV/CSF index | 32.785 | 21.597 | <0.001 |
| TIV | 0.002 | 20.417 | <0.001 |
| Gender | 99.756 | 2.663 | 0.008 |
| Education level | 71.363 | 2.090 | 0.037 |
| MMSE | 34.486 | 2.543 | 0.011 |
| BV/CSF index * Age | 0.975 | 11.307 | <0.001 |
| BV/CSF index * Gender | 9.189 | 4.329 | <0.001 |
| BV/CSF index * Education level | -6.735 | -2.246 | 0.025 |
| Age * MMSE | 3.802 | 2.921 | 0.004 |
| Age * Education level | -13.147 | -2.967 | 0.003 |
| Gender * Education level | 192.786 | 2.398 | 0.017 |
| Gender * TIV | 0.001 | 2.056 | 0.040 |
| Education level * TIV | -0.001 | -2.715 | 0.007 |

*i* = importance from dominance analysis in multiple linear regression. It reflects the relative percentage of the variance of the regression model explained by a given predictor; TIV = total intracranial volume; MMSE = Mini-Mental State Examination; Gender (0 female; 1 male); Education Level (0 low; 1 high). High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education. All possible interactions among age, the BV/CSF index, TIV, gender, education level and MMSE were tested. Only significant (p<0.05) predictors and interactions are presented in the table. Predictors included in the models can be consulted at Supplementary Table 4. The BV/CSF index was studied as a proxy of global brain atrophy.

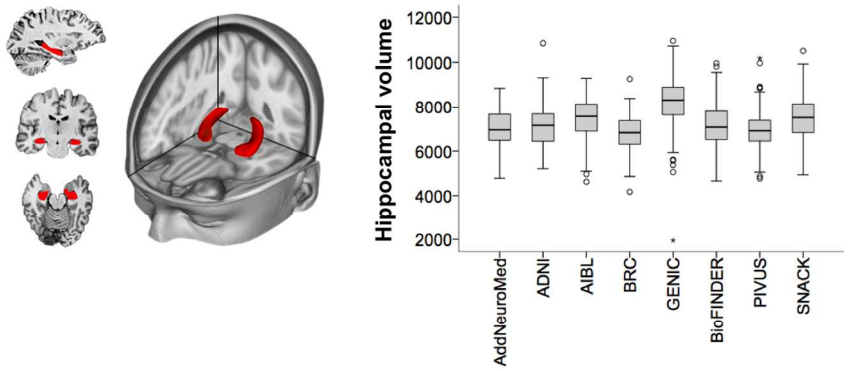



Figure 1. Between-cohort differences in hippocampal volume. Hippocampal volume was calculated by summing left and right sides and values are expressed in cubic millimeters. Values represent median and confidence intervals. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen.

529x396mm (72 x 72 DPI)

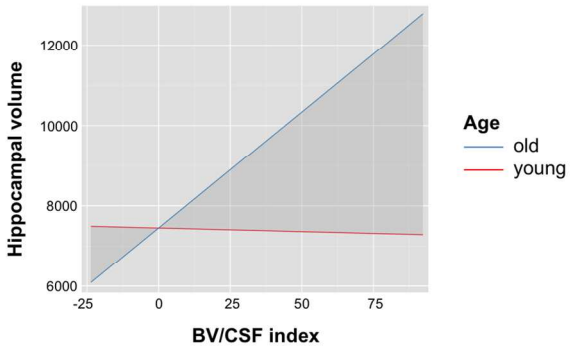



Figure 2. Effect of demographic and clinical variables on hippocampal volume (whole sample): significant interaction between age and the BV/CSF index (Multiple linear regression. Interaction between the effects of global brain atrophy (i.e. BV/CSF index) and age on hippocampal volume: unstandardized beta = 0.975; t(1897)=11.307; two-sided p<0.001. N = 1912). Hippocampal volume was calculated by summing left and right sides. The BV/CSF index was studied as a proxy of global brain atrophy. Two groups were created by separating the upper bound of age (old) versus the lower bound of age (young). Y-axis represents raw hippocampal volume in mm3 and x-axis represent mean centered values of the BV/CSF index.

529x396mm (72 x 72 DPI)

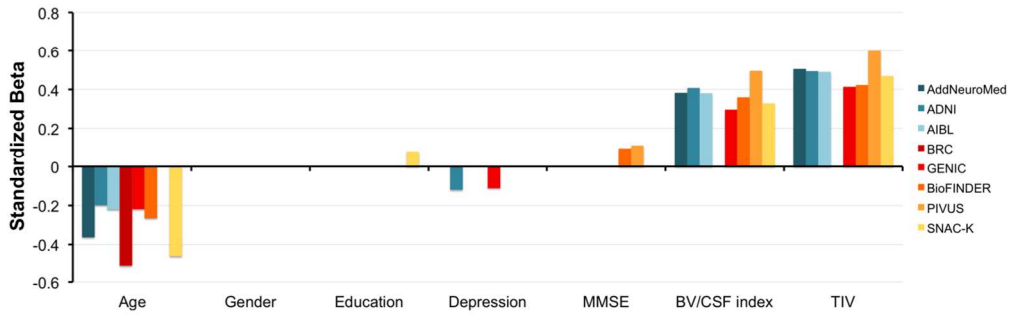



Figure 3. Effect of demographic and clinical variables on hippocampal volume (separately by cohorts). The figure schematizes the results from the multiple regression models (backwards) performed separately for the eight study cohorts. Only predictors remaining in the final models are displayed (criterion for excluding predictors from the models: two-sided $p<0.10$). Predictors included in the original models as well as sample size can be consulted in Supplementary Table 2. P-values of the primary regression models were adjusted using Bonferroni correction for multiple comparisons. Gender (0 female; 1 male); Education Level (0 low; 1 high). High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education. Depression was measured with GDS (AddNeuroMed, ADNI, AIBL, BRC, and GENIC), HADS (BioFINDER), and MADRS (SNAC-K). The BV/CSF index was studied as a proxy of global brain atrophy. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; MMSE = mini-mental state examination; TIV = total intracranial volume.

529x396mm (72 x 72 DPI)

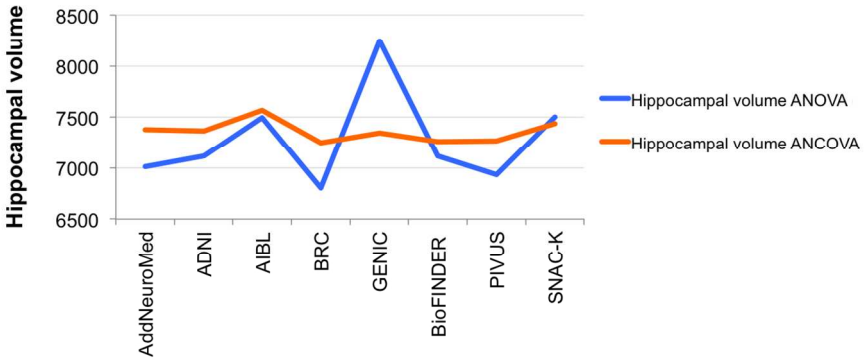



Figure 4. Attenuation of between-cohort differences in hippocampal volume when controlling for demographic and clinical variables (ANOVA: $F(7, 1904)= 63.188$; two-sided $p<0.001$; $\eta 2par= 19\%$; ANCOVA: $F(7, 1898)= 4.429$; two-sided $p<0.001$; $\eta 2par= 2\%$). The original ANOVA (N = 1958) was repeated for the same sample size than the ANCOVA (N = 1912) to allow perfect comparability of the results. Hippocampal volume was calculated by summing left and right sides and values are expressed in cubic millimeters. Covariates included in the ANCOVA model are age, gender (0 female; 1 male); education level (0 low; 1 high), MMSE, TIV and the BV/CSF index. High level of education corresponds with 9 or more years of education, while low level of education corresponds with less than 9 years of education. The BV/CSF index was studied as a proxy of global brain atrophy. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging Biomarkers and Lifestyle study; BRC = King's Health Partners Biomedical Research Centre for Mental Health Dementia Cohort; GENIC = Group of Neuropsychological Studies from the Canary Islands; PIVUS = Prospective Investigation of the Vasculature in Uppsala Seniors; SNAC-K = Swedish National Study on Aging and Care in Kungsholmen; MMSE = mini-mental state examination; TIV = total intracranial volume.

529x396mm (72 x 72 DPI)

**Author/s:**
Ferreira, D; Hansson, O; Barroso, J; Molina, Y; Machado, A; Andres Hernandez-Cabrera, J; Muehlboeck, J-S; Stomrud, E; Nagga, K; Lindberg, O; Ames, D; Kalpouzos, G; Fratiglioni, L; Backman, L; Graff, C; Mecocci, P; Vellas, B; Tsolaki, M; Kloszewska, I; Soininen, H; Lovestone, S; Ahlstrom, H; Lind, L; Larsson, E-M; Wahlund, L-O; Simmons, A; Westman, E

**Title:**
The interactive effect of demographic and clinical factors on hippocampal volume: A multicohort study on 1958 cognitively normal individuals

**Date:**
2017-06-01

**Citation:**
Ferreira, D., Hansson, O., Barroso, J., Molina, Y., Machado, A., Andres Hernandez-Cabrera, J., Muehlboeck, J. -S., Stomrud, E., Nagga, K., Lindberg, O., Ames, D., Kalpouzos, G., Fratiglioni, L., Backman, L., Graff, C., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I. ,... Westman, E. (2017). The interactive effect of demographic and clinical factors on hippocampal volume: A multicohort study on 1958 cognitively normal individuals. HIPPOCAMPUS, 27 (6), pp.653-667. https://doi.org/10.1002/hipo.22721.