Brain volumes in late life: gender, hormone treatment and estrogen receptor variants

*Joanne Ryan*a,b, Sylvaine Artero*a, Isabelle Carrière*a, Jacqueline Scali*a, Jerome J. Maller*c, Chantal Meslin*d, Karen Ritchie*a,e, Pierre-Yves Scarabin*f, Marie-Laure Ancelin*a

*aInserm, U1061, Montpellier F-34093; Université Montpellier 1, Montpellier F-34000 France

*bCancer & Disease Epigenetics Group, Murdoch Childrens Research Institute, Flemington Road Parkville 3052, Victoria, Australia

*cMonash Alfred Psychiatry Research Centre, The Alfred & Monash University Central Clinical School, Melbourne 3004, Victoria, Australia

*dCentre for Mental Health Research, Australian National University, Canberra 0200, ACT, Australia

*eFaculty of Medicine, Imperial College, London SW7 2AZ, United Kingdom

*fInserm, U1018, Université Paris Sud 11, UMRS 1018, F-94807, Villejuif, France

Email: joanne.ryan@inserm.fr (JR); sylvaine.artero@inserm.fr (SA); isabelle.carriere@inserm.fr (IC); jacqueline.scali@inserm.fr (JS); jerome.maller@monash.edu (JJM); chantalmeslin@hotmail.com (CM); karen.ritchie@inserm.fr (KR); pierre-yves.scarabin@inserm.fr (PYS); marie-laure.ancelin@inserm.fr (MLA).

*Corresponding author: Dr. Joanne Ryan, Inserm U1061, Nervous System Pathologies: Epidemiological and Clinical Research. Hôpital La Colombière, 39 avenue Charles Flahault, BP 34493, 34093 Montpellier Cedex 5, France. Tel +33 499614565; Fax +33 499614579. joanne.ryan@inserm.fr
ABSTRACT

Structural imaging studies suggest gender differences in brain volumes but whether hormone treatment (HT) can protect against age-related structural changes remains unknown and no prior neuroimaging study has investigated potential interactions between HT and estrogen receptor (ESR) polymorphisms. MRI was used to measure grey and white matter, hippocampal volume, corpus callosum, cerebrospinal fluid (CSF), total intracranial volume (ICV) and white matter lesions (WML) in 582 non-demented older adults. In multivariable analysis when compared to women never using HT, men and women currently on treatment, but not past users, had significantly smaller ratio of grey matter to ICV and increased atrophy (CSF/ICV ratio). Hippocampal and white matter volume, and corpus callosum area were not significantly different across groups. ESR2 variants were not significantly associated with brain measures, but women with the ESR1 rs2234693 C allele had significantly smaller WML. Furthermore this association was modified by HT use. Our results do not support a beneficial effect of HT on brain volumes in older women but suggest the potential involvement of ESR1 in WML.

Keywords

Grey matter, white matter, white matter lesions, hormone treatment, estrogen receptor polymorphisms, gender differences, imaging
1. Introduction

Estrogen is most commonly known as a reproductive hormone, being one of the principal steroid hormones produced by the ovaries but it also plays an important role in the brain, as largely evidenced from animal studies. Estrogen has been implicated in neurogenesis and synaptogenesis during development (McCarthy, 2008), it influences synaptic plasticity, interacts with other neurotrophins (Scharfman and MacLusky, 2008), and is neuroprotective, inducing the survival, differentiation and function of neurons throughout life (Brann et al., 2007, Hojo et al., 2008, Spencer et al., 2008). Through epidemiological studies, estrogen has also been linked to cognitive function (Ryan et al., 2009) and a number of brain-related disorders, such as dementia (Henderson, 2010), depression (Ancelin et al., 2007) and stroke (Scarabin-Carre et al., 2012, Shao et al., 2012). Speculatively, the decline in estrogen levels which accompanies the postmenopausal period in women may also help account for gender-related differences in these disorders. However, studies on the potential beneficial effects of supplementing estrogen in postmenopausal women, through the use of hormone treatment (HT) have produced mixed results (Lisabeth and Bushnell, 2012, Maki et al., 2010, Sherwin and Henry, 2008) with some evidence that estrogen treatment in later life can have detrimental effects (Lisabeth and Bushnell, 2012, Shumaker et al., 2003). Given that the majority of HT preparations are a combination of estrogen and progestagen, it is also possible that progestagen could also influence cognitive function (Hogervorst and Bandelow, 2010).

In terms of structural imaging studies, gender differences in brain volumes have been observed, with a number of studies reporting that men have larger global brain volumes than women and a higher ratio of white matter to whole brain volume (Gur et al., 1999, Passe et al., 1997). Women on the other hand, reportedly have a higher ratio of grey matter (Gur et al., 1999). Loss of brain volume which occurs with normal aging is greater in men overall and tends to begin at an earlier age (Murphy et al., 1996). However after menopause, brain atrophy in women accelerates at a much faster rate (Takeda and Matsuzawa, 1985), suggesting the potential involvement of
estrogen. Other age-related brain changes, such as an increase in white matter lesions (WML), may also be more frequent in older women than men (de Leeuw et al., 2001). A number of structural imaging studies over the last decade have investigated potential associations between HT use and various brain volumes in postmenopausal women (Boccardi et al., 2006, Coker et al., 2010, Cook et al., 2004, Eberling et al., 2003, Erickson et al., 2010, Greenberg et al., 2008, Lord et al., 2010, Sullivan et al., 2005), but the findings have been inconsistent. This likely relates to the fact that the sample sizes have often been very small and failure to adequately control for confounding factors may have influenced the results. It is possible that differences relate to the fact that some studies examined associations with HT overall (Cook et al., 2004, Erickson et al., 2010, Greenberg et al., 2006, Sullivan et al., 2005), while others focused on unopposed estrogen treatment (Boccardi et al., 2006, Eberling et al., 2003, Lord et al., 2010). However one study which examined both, reported that combined estrogen-progestagen HT and estrogen alone were similarly associated with ischemic lesion volume (Coker et al., 2010). The question therefore remains of whether HT can help protect against brain-related structural changes in older women.

Another area which has yet to receive much attention is the potential involvement of estrogen receptors (ESR) in influencing brain volumes. Estrogen exerts its biological effects in large part through intracellular activation of its two principal receptors, estrogen receptor alpha (ESR1) and beta (ESR2), which are located throughout the brain and predominate in limbic-related areas which play an important role in cognitive function (Osterlund and Hurd, 2001). Through binding to its receptor and translocation into the nucleus, this estrogen-receptor complex is able to regulate the expression of hundreds of genes. Gene variants in the ESR could thus modify estrogen signalling and the effect that estrogen has on the brain. This could help explain the mixed findings of previous imaging studies investigating associations between estrogen-containing HT and brain volumes. Indeed, prior work has demonstrated that activation of either ESR1 or ESR2 can mediate estrogen’s beneficial effect on synaptic plasticity (Wilson et al., 2013) and both receptors play a role in preventing neurodegeneration in hippocampal neurons (Zhao et al., 2004). Our own work and that
of others has demonstrated that ESR1 and ESR2 variants are associated with cognitive function and the risk of dementia (Ryan et al., 2013a, Ryan et al., 2013b) (Bertram et al., 2007, Yaffe et al., 2009). Associations with cardiovascular disease and ischemic stroke have also been reported (Li et al., 2012, Rexrode et al., 2007). To date however, only three studies have investigated potential associations between ESR gene variants and brain volumes, one of which was a study of neonates (Knickmeyer et al., 2013). The Rotterdam study reported a significant association between common ESR alpha (ESR1) variants and amygdalar volumes in older women only (90% of whom never used HT) (den Heijer et al., 2004), but no significant association with hippocampal (HC) volume. Other regional brain volumes were not examined. An Italian study of 20 postmenopausal women balanced for HT ever/never use reported that the rs9340799 G allele was associated with smaller grey matter volumes (Boccardi et al., 2008). No subsequent study has attempted to replicate their findings and WML have not been examined. Furthermore, none of these studies examined the potential modifying effect of HT, despite prior studies demonstrating the genetic variation in the ESR could interact with HT to influence a number of health-related outcomes (Almeida and Hutz, 2007, Herrington et al., 2002b, Ryan and Ancelin, 2012, Ryan et al., 2013a). It may be that some women are genetically (in terms of the ESR) more susceptible to the effects of HT on brain volumes.

Using a population-based cohort of non-demented older adults, this study firstly aimed to examine differences in brain volumes (in terms of grey and white matter and WML) and regional brain structures (HC, corpus callosum) between males and females and across females depending on their lifetime use of HT. The second aim of the study was to examine potential associations between ESR gene variants and brain volumes, as well as the potential interaction between these variants and HT. Based on the previous literature, we hypothesise that women who have used HT would have larger grey and/or white matter volumes, as well as larger hippocampal volume and fewer WML than women who never used treatment. We also hypothesised that ESR variants would interact with current HT to modify the association with brain volumes.
2. Methods

2.1 Study population

The data used for this analysis were derived from a general population study of psychiatric disorders among older adults in the French community (the ESPRIT Study) (Ritchie et al., 2004). Eligible participants, who were at least 65 years of age and non-institutionalised, were randomly selected from the electoral rolls in Montpellier between 1999 and 2001. Ethics approval for the ESPRIT study was granted by the Ethics Committee of the University Hospital of Kremlin-Bicêtre (France) and all procedures were carried out with the adequate understanding and written consent of the participants. Of the 1863 participants initially recruited to the ESPRIT Study, 760 participants aged 80 years or less were randomly selected to take part in the imaging study and 690 had complete volumetric data relevant for this analysis. Those participants diagnosed with dementia (n=14) or who were left handed (n=19) were excluded from this analysis. Women who did not respond to the question concerning HT use (n=12) and participants with incomplete genotyping data (n=63) were also excluded. This analysis is thus based on 582 older adults (287 men and 295 women). Participants who took part in the imaging study but were excluded from this analysis, were significantly older (p<0.001), and a higher percentage of excluded women had a history of cardiovascular disease and were currently using more than three medications (p<0.01). Included and excluded participants however, were not significantly different in terms of brain measures, gender, HT use or ESR genotype frequencies.

2.2 Magnetic Resonance Imaging (MRI)

2.2.1 A 1.5T GE Signa Imaging System (General Electric Medical Systems, Wisconsin, USA) was used to acquire a transversal fast multislice double echo T2-weighted 2D axial data, (TR=4400ms, TE1 and TE2= 16ms and 98ms, slice thickness = 4mm, gap= 0.4mm, matrix = 256x256, in-plane resolution = 0.98 x 0.98mm²) that covered the whole brain. T1-weighted volumetric MR imaging was also obtained by using the spoiled gradient echo sequence (TR = 97ms, TE = 4ms) which consisted of
a set of 124 adjacent transverse sections parallel to the anterior commissure-posterior commissure line with a section thickness of 1.5mm (no gap).

2.2.2 Brain volumes (cm$^3$) were determined by segmenting each T1-weighted SPGR image into its component tissue classes, grey matter, white matter and cerebrospinal fluid with SPM5 (Wellcome Department of Cognitive Neurology, London, UK), using the segment.m script developed by Jon Jackson (http://www.fil.ion.ucl.ac.uk/spm/ext/#spm_segment). Subcortical /deep grey / white matter structures were included in the total grey matter and white matter volumes. Other details of the script included no lesion masks, default threshold (0) and separable smoothing kernel = 3x3x3. Total intracranial volume (ICV) was calculated as the sum of grey matter plus white matter plus cerebrospinal fluid volumes (Acosta-Cabronero et al., 2010). All outputs were manually inspected to ensure accurate segmentation and valid data. Brain atrophy was estimated as the ratio of cerebrospinal fluid to ICV.

2.2.3 The corpus callosum (CC) outline was manually traced on the midline sagittal slice of the T1 images using anatomical landmarks in a hierarchical order and the region of interest (ROI) module of AnalyzeTM 9.0 (Brain Imaging Resource, Mayo Clinic, Minnesota, USA) on a Windows XP Professional workstation. The landmarks based on the midline sagittal slice were i) no white matter or only minimal white matter in the cortical mantle surrounding the CC; ii) the inter-thalamic adhesion; and iii) the transparent septum and the visibility of the aqueduct of Sylvius (Anstey et al., 2007).

Hippocampal ROIs were manually outlined on consecutive coronal slices and verified from axial and sagittal orientations (Maller et al., 2007). The anterior tip of the HC until the slice before the opening of the crus of the fornix (CF) was measured as the HC head and body and included the subiculum, CA1–(4) areas, and dentate gyrus (DG), as described previously (Watson et al., 1997), and the HC tail was measured from the slice immediately posterior to that which represented the last slice according to the Watson protocol. The internal structure of the HC tail is the same as in the head and the body, whereby the cornu ammonis has an analogous structure throughout, as does the
gyrus dentatus. From the coronal perspective, measuring the HC until the CF represents the part of the tail which coincides with the coronal section of the pulvinar (which is situated in a supero-medial position). Voluminous choroidal plexuses occupy portions of this region, hence care was taken to exclude them laterally from volumetric estimates. The HC was then followed posteriorly. On initial slices, the tail appears bulgy as an ovoid mass of gray matter on the infero-medial part of the lateral ventricle, and more posteriorly it lies flattened on the superior surface of the parahippocampal gyrus. The tail was outlined until the fasciolar gyrus becomes the subsplenium gyrus curving around the postero-inferior margin of the splenium. The superior border was easy to differentiate from the crus of the fornix. The medial and inferior limits were also easily drawn because of the contrast between grey/white matter. Relevant images from standard atlases were referred to in order to ensure a consistent reference to the boundaries and relevant landmarks for these slices of HC. All HC volumes are expressed in mm$^3$ and CC volumes are expressed in mm$^2$.

CC and HC outlines were traced by two trained researchers blind to the study hypotheses, group assignment and participants’ identity. The reliability of the CC and HC measurements was assessed using a formula to calculate the intra- and inter-class correlations (intra-CC, intra HC and inter-CC, inter HC) that presumes random selection of raters (Shrout and Fleiss, 1979). The two researchers (JM, CM) each retraced five MRI images, which were randomly selected among the images previously traced, and five images which belonged to the group previously traced by the other researcher. Intra-CC was 0.957 for JM and 0.962 for CM. Inter-CC was 0.915. Intra-HC was 0.942 for JM and 0.970 for CM. Inter-HC was 0.939. All these values are well within acceptable limits.

2.2.4 The volume of WML (cm$^3$) was estimated using a semiautomatic method (Brickman et al., 2009, Gurol et al., 2006). Areas of supratentorial WML appearing as hyperintensities were segmented on T2-weighted sequences using MRicro software (Rorden and Brett, 2000). A first layer of ROIs corresponding to WML was created by a semi-automated technique based on intensity thresholding. A second layer of ROIs was then manually outlined on each slide by roughly contouring of all WML. The intersection of the first and second layer was then manually inspected and
automatic total volume of WML obtained, irrespective of underlying cause. Persons with extensive
damage due to stroke were excluded. An experienced reader examined all scans. Another
experienced neurologist examined 80 randomly chosen scans to assess inter-rater reliability. Inter-
rater and intra-rater intra-class correlation coefficients showed good-to-excellent agreement (0.79
and 0.95, respectively).

2.3 Hormone treatment

Current and past use of HT and detailed information relating to the type of treatment was obtained
at baseline. Treatment use was validated by presentation of the prescription or the medication itself
and past users were shown photos to aid with recall. Current users were classified according to the
route of estrogen administration and the type of progestogen (progesterone or progestins). The
duration of current HT and the timing of initiation of first treatment in relation to the menopause
were also examined, with age at menopause being defined as one year without menses.

2.4 Estrogen receptor polymorphisms

Fasting venous blood samples were taken from the participants at baseline. DNA was extracted from
white blood cells (Puregene kit, Qiagen, France) and stored at -80°C. Genotyping was performed by
Kbiosciences (Hoddesdon Herts, UK) using their competitive allele-specific PCR Single-Nucleotide
Polymorphism (SNP) genotyping system (KASPar). The amplified PCR products were analysed by
fluorescence scanning in a BMG labtech Pherastar scanner and the results were interpreted with
their KlusterCaller 1.1 software. The error rate for the KASPar assay system is less than 0.3%. The
most common two ESR1 polymorphisms were examined, rs2234693 and rs9340799 (otherwise
known as Pvull and Xbol), which are located at position 397 and 351 of intron 1 respectively.
Previous studies have reported significant associations between these SNPs and brain-related
disorders, including cognitive impairment and Alzheimer’s disease (www.alzgene.org) (Ryan and
Ancelin, 2012, Sundermann et al., 2010), as well as cholesterol levels and cardiovascular disease (Li
et al., 2012, Rexrode et al., 2007). There is also some evidence that they may be functionally significant (Alonso et al., 2011, Maruyama et al., 2000). Three ESR2 polymorphisms with unknown functional consequences but showing potential causal associations with other hormone-related health outcomes, including cognitive outcomes, were also examined (Rexrode et al., 2007, Ryan and Ancelin, 2012, Ryan et al., 2013b): rs1256049 (position 1082 of exon 5), rs4986938 (position 1730 in the 3’-untranslated region of exon 8) and rs1271572 in the promoter region. While both ESR1 and ESR2 subtypes are expressed throughout the brain, ESR1 is found in higher concentrations in the hypothalamus, the hypothalamic preoptic area and in the amygdala and ESR2 predominates in the hippocampus, entorhinal cortex and thalamus (Osterlund and Hurd, 2001). This localisation may suggest that ESR2 would play a more important role in cognitive function, but research has shown that both receptor subtypes can mediate estrogen’s neuroprotective effect (Zhao et al., 2004).

2.4 Socio-demographic and clinical adjustment variables

Information was gathered on the participant’s age, education level, living status, consumption of alcohol, and smoking status. Regular smoking was defined as smoking at least 10 packs per year and high alcohol consumption was at least 24 grams of alcohol each day. Body mass index (BMI) was calculated as weight (kg) divided by height (metres²). The Mini-International Neuropsychiatry Interview (MINI), a standardized psychiatric examination which has been validated in the general population (Sheehan et al., 1998), was used for the diagnosis of current and past major depressive disorder (MDD), according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria. The Center for Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977) was used for the assessment of depressive symptoms (CES-D ≥16). Low global cognitive function was defined as participants scoring 24 or less on the Mini-Mental State Examination (MMSE) (Folstein et al., 1975).

Information on the health of the participants was obtained through detailed medical questionnaires, a complete inventory of drug use in the preceding month and from fasting blood samples. Participants were classified as having a history of cardiovascular disease if they suffered
from angina pectoris, myocardial infarction, stroke, bradycardia, palpitations or reported cardiovascular surgery. Blood pressure was measured twice in a sitting position using a digital electronic tensiometer OMRON M4, and the average was used in the analyses. Hypertension was defined as resting blood pressure ≥ 160/95 mm Hg or treatment, diabetes as fasting glucose level ≥ 7.0 mmol/l or treatment and hypercholesterolemia as total cholesterol levels ≥ 7.2 mmol/l or treatment. Current use of medications was validated either by presentation of the prescription or the medication itself. In this analysis we considered in particular the use of anticholinergics, or a “large number of medications” (defined as using three or more current medications). Genotyping of the Apoe-ε4 allele was performed at a genotyping facility in Lille, France (http://www.genopole-lille.fr/spip/).

2.5 Statistical analyses

Grey matter and white matter volumes, as well as midsagittal total CC area and total HC volume, were expressed as a ratio of ICV to control or normalize for differences in overall brain size. WML values were transformed by a log_{10} (x+0.01) function given their highly asymmetric distribution and possible null values. All models with WML as the dependent variable were adjusted for total white matter volume included as a covariate. The association between brain volumes and socio-demographic and clinical variables was tested using Pearson’s correlation coefficient and Student’s t-test for continuous and categorical variables, respectively.

Linear regression models were used to determine the association between gender or HT and brain volumes in age-adjusted analysis. Multivariable models were then generated adjusting for age, education level, history of cardiovascular disease, hypertension and number of medications currently used. These variables were chosen as they were most strongly associated with brain volumes in our analysis and were considered to potentially confound the association between brain volumes and HT based on the literature. Adjustment for additional covariates such as low cognitive function, lifetime major depressive disorder, hypercholesterolemia and other variables listed in
Table 1 was also considered in the final multivariable models by adding them consecutively to the models and ensuring that they had no substantial effect on the estimates. Analysis was also repeated by excluding participants with low MMSE scores (MMSE≤24). This was done to ensure that the study findings were not been overly influenced by the inclusion of participants with some degree of cognitive impairment.

Chi-squared tests were used to compare the distribution of ESR genotypes with those predicted under the Hardy-Weinberg equilibrium (Trikalinos et al., 2006) and pair-wise linkage disequilibrium was estimated to determine the correlation between individual ESR1 and ESR2 SNPs. Age and multivariable-adjusted linear regression models were then used to examine the independent association between ESR gene variants and brain volumes, separately in males and females. A first-order interaction between these SNPs and HT use in women was also examined by including a product term in the models. When significant, subsequent analysis was stratified to determine independent group effects. SAS version 9.2 (SAS Institute, Inc., Cary, North Carolina) was used for all of the statistical analysis and the significance level was p<0.05.

3. Results

3.1 Participant’s characteristics

The socio-demographic, lifestyle and health characteristics of the 287 men and 295 women in this study are given in Table 1. Comparing men and women overall, they differed in terms of all of the characteristics listed (p<0.10), with the exception of history of cardiovascular disease (p=0.31), hypertension (p=0.33) and Apoe-ε4 allele carriers (p=0.74). When women were compared according to their use or not of HT, they differed in terms of age, BMI, education, history of cardiovascular disease, hypertension, diabetes and use of anticholinergic medication (p≤0.10).

Table 1 also lists the median and interquartile brain volumes by gender and use of HT, both as absolute volumes (grey and white matter, HC, CC area) and when expressed as a ratio of total ICV. The median percentage of atrophy, measured as the volume of CSF relative to total ICV was 14.8%.
(interquartile range: 13.6 to 16.7%). The total volume of WML ranged from 0.0001 to 126.6 cm³, with 10 males and 16 females having no WML.

Of the women, 173 stated that they had never used HT, 60 were past HT users and 62 women currently used HT. In terms of the type of treatment, the majority of women (58.1%) used transdermal estradiol (12.3% unopposed and 17.5% combined with oral progesterone and 28.3% synthetic progestin), and 15.6% used oral estradiol (10.7% combined with synthetic progestin). Of the remaining women, 6.6% used other forms of HT and 19.7% of women (all past users of HT) did not recall the type of treatment used. The median duration of HT use overall was 10 years (IQR: 4-15 years), while for current users it was 13 years (IQR: 7-18 years) and for past users it was 5 years (IQR: 2-10 years).

3.2 Gender, HT and brain volumes

Comparing brain volumes across the genders, women had a significantly higher ratio of grey matter to ICV compared to men (adjusted linear regression $\beta=0.59$, SE=0.21, $p=0.006$), but there was a non-significant trend for women to have a smaller percentage of white matter to ICV (adjusted linear regression $\beta=-0.31$, SE=0.17, $p=0.07$). Men also had significantly smaller CC area than women (adjusted linear regression $\beta=0.03$, SE=0.007, $p<0.0001$). There was no significant difference between the genders in terms of the volume of HC (adjusted linear regression $\beta=0.09$, SE=0.06, $p=0.12$) or WML (adjusted linear regression $\beta=0.05$, SE=0.07, $p=0.50$). Women were then grouped according to their use of HT and the results of the age- and multivariable-adjusted linear regression analysis for the association with brain volumes are shown in Table 2. The results suggest that both men and women currently using HT had a significantly smaller percentage of grey matter to ICV when compared to women never using HT. In post-hoc pair-wise analysis, men and women currently using HT, also had a significantly smaller percentage of grey matter to ICV compared to past HT users ($p=0.001$ and $p=0.005$ respectively). There was no significant difference in the percentage of grey matter to ICV between past and never HT users ($p=0.22$), nor between women currently using HT
and men (p=0.69). The results were similar in the reversed direction for brain atrophy, with men and women currently using HT having significantly higher atrophy than women never using HT (adjusted linear regression for log-transformed variable $\beta=0.013$, $SE=0.006$, $p=0.04$ and $\beta=0.022$, $SE=0.01$, $p=0.02$ respectively). Men also had significantly higher atrophy than past HT users ($p=0.04$), but among women, the comparison of current and past HT users was not significant ($p=0.31$). All of the significant associations described above remained after further adjustment for the variables listed in Table 1.

In terms of the percentage of white matter or HC volume to ICV, there was no significant main group difference, and CC volume was not significantly different across the groups of women. There were also no significant associations between WML volume and either gender, or according to HT use. When specific characteristics of HT were examined, such as duration of use, type of treatment and age at initiation, we found no significant associations with brain volumes (data not shown).

3.3 ESR polymorphisms and brain volumes

The ESR genotype frequencies in males were as follows: rs2234693 TT=79, CT=153, CC=55; for rs9340799 AA=114, GA=137, GG=36; for rs1271572 GG=91, TG=138, TT=58; for rs4986938 GG=111, GA=132, AA=44; and for rs1256049 GG=264, GA/AA=23. For females the genotype frequencies were: rs2234693 TT=90, CT=148, CC=57; for rs9340799 AA=124, GA=139, GG=32; for rs1271572 GG=89, TG=158, TT=48; for rs4986938 GG=93, GA=155, AA=47; and for rs1256049 GG=269, GA/AA=26. These frequencies were not significantly different from those predicted by the Hardy-Weinberg equilibrium ($p>0.10$ for all tests). The ESR1 SNPs were in strong linkage disequilibrium ($|D'|=0.98$), as were the three ESR2 SNPs ($|D'|>0.90$ for all pairwise comparisons). To maximise the power of subsequent analyses, homozygotes for the variant allele (the smallest group in each case) were combined with the heterozygotes.
Linear regression analyses stratified by gender examined the association between ESR SNPs and brain volumes. Among the men (n=287), no significant associations were identified between the ERS1 or ESR2 variants and the ratio of grey matter, white matter, or HC volume relative to ICV (Supplementary Tables 1 and 2 respectively). However there was a weak trend for men with the TT genotype of ESR1 rs2234693 to have small CC (p=0.07). No significant associations were found with the volume of WML, with similar results in age-adjusted and multivariable-adjusted analyses. Among women, there was also no significant association between either of the ESR1 SNPs and the percentage of grey matter or white matter relative to ICV or in terms of HC or CC volume (Table 3 and Supplementary table 3). However, rs2234693 was significantly associated with the volume of WML. Adjusting for all other variables, this corresponds to a 0.26 difference in the expected geometric mean of the log of WML between women with the C allele and those homozygous for the T allele. Converting back to the original scale, women with the TT genotype have 1.82 times more WML than women CT/CC, when all covariables were held constant. These associations were not modified after further adjustment for covariates listed in Table 1. None of the ESR2 SNPs were significantly associated with brain volumes in women (Table 3 and Supplementary table 3).

We then investigated whether there was any evidence of a gene-environment interaction, whereby these SNPs could interact with HT and modify the association with female brain volumes. For both ESR1 rs2234693 and rs9340799 we found a significant HT interaction on WML (p=0.005 and p=0.006). Stratified linear regression analysis was therefore undertaken (Table 4). The findings for never HT users and current users mimicked those of women overall, whereby the TT genotype of rs2234693 was associated with larger volume of WML (β=0.38 and 0.43 respectively) compared to women with the C allele, although for current users this only reached significance after multivariate analysis. In contrast, for women who were past users of HT, the rs2234693 SNP was not associated with the amount of WML (p=0.48). Similar results were seen for rs9340799, whereby among current HT users the AA genotype was associated with larger WML volume compared to women with the rs9340799 G allele, but there was no significant association among women who never HT. For past
HT users, women homozygous for the A allele, versus those with a G allele, had significantly small WML, however after multivariate adjustment, this association was no longer significant (p=0.17).

4. Discussion

Over the last decade a number of structural studies have investigated associations with HT, however these studies have generally focused on only one specific brain measure, most commonly hippocampal volume and they have been limited in sample size, often comparing groups of less than 20 women. This has limited their ability to consider specific characteristics of HT, to differentiate women who are current and past HT users, and to consider other factors which could potentially have confounded the associations. Prior studies have also failed to take into account genetic factors, in particular variants in the ESR, which could modify the association between HT and brain volumes and men were rarely examined. The results of our relatively large population based study indicate that older men and women currently using HT have a significantly smaller ratio of grey matter to total ICV than women never on treatment or past users of HT. Similar results were found with an approximate measure of brain atrophy in that men and current HT users had significantly more atrophy than never users but the comparison of current and past HT users did not reach significance. This study also reports for the first time that a genetic variant in the ESR1 rs2234693 is associated with the volume of WML in older postmenopausal women, and this variant, as well as another common ESR1 polymorphism rs9340799, can also modify the association between HT and WML.

4.1 HT and brain volumes

Sexual dimorphism in human brain structure has been observed previously (Greenberg et al., 2008, Gur et al., 1999, Lord et al., 2010). These fit with our observations that after taking into account differences in ICV men have significantly smaller grey matter volumes than women, but a trend for larger white matter volumes. However, against the generalised view that estrogen may be beneficial
for cognitive function, as supported by animal experiments and epidemiological studies (Ancelin and Ritchie, 2005, Sherwin and Henry, 2008), we found that women currently using estrogen-containing HT had smaller grey matter volumes and greater atrophy than women who had never used HT. This suggests rather a detrimental effect of postmenopausal HT.

Previous neuroimaging studies in this area report mixed findings. The most frequently studied measure has been HC volume, with some inconsistent findings (Eberling et al., 2003, Greenberg et al., 2006, Raz et al., 2004). A Canadian study reported that 16 women currently using HT had significantly larger right HC volume than 10 past users and 15 women who had never used HT (Lord et al., 2008). However, they also reported that among current users, increased duration of treatment was associated with smaller total HC volume. Another study which examined HC volume in 62 women who had ever used HT and 37 controls, found that no significant difference overall between ever and never users, but the 38 women who had used HT at the menopause had a significantly larger hippocampus (Erickson et al., 2010). A larger study of 213 postmenopausal women however, supports the findings of our study, in that there was no significant difference in HC volume between current, past and never users of HT (Low et al., 2006). The biggest study in this field which involved older women aged 71-89 years from the Women’s Health Initiative Memory Study (WHIMS), actually reported that estrogen or combined treatment was associated with smaller HC and frontal lobe volumes than placebo, although for HC this was of borderline significance (p=0.05) (Resnick et al., 2009). The WHIMS also reported a non-significant trend for HT to be associated with lower total brain volumes (Resnick et al., 2009). A few studies have looked at other brain measures, including grey matter, but equally mixed results have been reported. In direct contrast to the findings of our study, a small study reported that 17 never users of HT had smaller grey matter volumes than 23 ever users, however when separating out ever users, they actually found that the 16 women currently using HT had significantly smaller grey matter than the 7 past users (Boccardi et al., 2006). Another small study reported that postmenopausal women who had used HT since their ovariectomy, had larger grey matter volumes in some brain regions compared to
postmenopausal women never using HT (Robertson et al., 2009). On the other hand, the larger study by Greenberg et al., involving 41 postmenopausal women who used HT and 51 never users of treatment, and which adjusted for ICV, supports the findings of our study in that women using HT had smaller grey matter than never users (Greenberg et al., 2006). The findings of our study in terms of brain atrophy are also supported by two of the largest previous studies which have focused specifically on this brain measure (Coker et al., 2010, Luoto et al., 2000). Together this gives strong support for the notion that current HT use in older women is associated with greater brain atrophy.

Specific characteristics related to HT could help explain the discrepancies in findings between some studies. We found no significant evidence that the type of treatment, duration of use or timing of initiation in relation to the menopause was significantly associated with brain volumes, but the small numbers of women in each subgroup limited the power of this analysis. It has been suggested previously for example, that less than 10 years of HT is associated with greater grey matter volume, but treatment given for more than 10 years was detrimental (Erickson et al., 2007). The median duration of current use in our study was 13 years. The majority of women in our study were using combined treatment consisting of transdermal estradiol and progestagen, which is much higher than that observed in other studies. Interestingly, the study by Greenberger et al., which found similar results to ours, also had a relatively high percentage of women using combined HT (Greenberg et al., 2006). On the other hand, the WHIMS study found no difference in brain measures between women who were given combined or unopposed estrogen treatment (Coker et al., 2010, Resnick et al., 2009). Furthermore, although women in the WHIMS were administered a form of HT (conjugated equine estrogens with or without medroxyprogesterone acetate) which was not used by any women in our study, findings from their study and ours were similar. Other possible explanations for the differing findings may relate to confounding from factors which have not been controlled for in the analyses. Several small studies which reported a positive association between grey matter volume and the use of HT for example, failed to adjust for health-related variables (Boccardi et al., 2006, Cook et al., 2004, Ghidoni et al., 2006, Robertson et al., 2009).
In terms of other brain volumes, a few prior studies have reported no significant association between HT use and white matter volume (Ghidoni et al., 2006, Greenberg et al., 2008, Lord et al., 2010), yet as these studies were small, one could not rule out that the lack of significant association was due to the limited statistical power. Our larger study has confirmed the findings of these studies. Our findings of no association between the extent of WML and HT use are also in accordance with the biggest previous studies in this area (Greenberg et al., 2006, Low et al., 2006), including a randomized controlled trial which found that HT use was not associated with total ischemic lesion volume, or lesions in the white matter (Coker et al., 2010, Greenberg et al., 2008, Low et al., 2006). They did not however, investigate potential associations with ESR variants.

4.2 ESR1 and WML

Despite the widespread distribution of ESR throughout the brain (Osterlund and Hurd, 2001), very few studies have investigated potential associations between ESR variants and brain volumes. While we found no significant association between ESR1 rs2234693 or rs9340799 and the volume of grey matter, white matter, HC or CC area in either men or women, the C allele of rs2234693 and G allele of rs9340799 were associated with significantly smaller volume of WML in women. WML are thought to reflect small-vessel cerebrovascular disease (LadisStudyGroup, 2011, Raz et al., 2007), which increases in occurrence with age (Ylikoski et al., 1995) and have been associated with poorer cognitive performance (Inzitari et al., 2009), mild cognitive impairment (Ritchie et al., 2010) and an increased risk of dementia (Debette and Markus, 2010, Mortamais et al., 2013), particularly in women (Sawada et al., 2000). In line with this, previous studies have also reported significant associations between rs2234693 and cognitive function (Yaffe et al., 2002), the risk of Alzheimer’s disease (Sundermann et al., 2010) and cardiovascular disease in older women (Herrington et al., 2002b, Schuit et al., 2004), however the direction of the association has not been consistently found. We have recently found that in older women specifically, the CC genotype of rs2234693 tended to be associated with lower decline in executive function (Ryan et al., 2013a), which is considered as
amongst the earliest signs of vascular MCI (Howieson et al., 2008, Iachini et al., 2009). The rs2234693 polymorphism may help regulate the expression of ESR1 (Alonso et al., 2011, Maruyama et al., 2000) and alter transcription factor binding (Herrington et al., 2002a) and could thus modify estrogen-mediated signalling. The C allele for example has been associated with increased serum estradiol levels in some studies (Schuit et al., 2005, Sowers et al., 2006), which could provide a mechanism by which it has a beneficial effect on cognitive function, but the exact functional consequence of this variant remains unknown. It could also be in high linkage disequilibrium with other yet undefined functional variants. A previous study examining HC and amygdala volumes found that women (but not men) homozygous for the alternative T allele had smaller amygdala volumes (den Heijer et al., 2004), but no significant association between this ESR1 variant and HC, as we report here. An Italian study of 20 postmenopausal women found a significant association between the G allele of ESR1 rs9340799 and grey matter (Boccardi et al., 2008). ESR1 rs2234693 was not examined, neither were ESR2 variants nor WML. No study to date has investigated whether there is a significant interaction between ESR1 variants and HT on brain volumes.

Previous research provides evidence that there is a high heritability of WML especially in women (Atwood et al., 2004) and our results suggests for the first time that ESR1 may be one of the genetic factors that is implicated. Our findings also indicate a significant gene-environment interaction, whereby past use of HT appeared beneficial for women who were genetically more susceptible to WML in later life. A previous study has found that women with the CC genotype of rs2234693 showed a greater increase in high-density lipoprotein cholesterol response with HT compared to other women (Herrington et al., 2002b), which could support a beneficial effect on reducing the risk of atherosclerosis. The reason for our specific effect of past HT remains unknown, but does not appear to relate to differences in the characteristics of HT treatment. That is, there was no significant difference between current and past HT users in terms of the duration or type of HT or the timing of HT in relation to the menopause and, we controlled for other variables such as age, education level and health-related factors in the analysis. Interestingly, a few prior studies have also
reported beneficial effects of past but not current HT, with less decline in cognitive function (Matthews et al., 1999) and a reduced risk of Alzheimer’s disease (Zandi et al., 2002) compared to never users. Our novel finding warrants further investigation and replication.

4.3 Limitations and strengths

There are some limitations to our study which should be considered when interpreting the results. Although the study population was randomly selected from those living in the community, participants who agreed to participate in the study and those that were included in this analysis were younger and in overall better health. This therefore limits the extent to which these findings can be generalised to the wider community of older adults. Data concerning the use of HT was gathered retrospectively and is subject to recall errors. Current HT users were however encouraged to present medications or prescriptions so that treatment use could be verified and past HT users were shown photos to help aid recall. The sample size, although considered large for this type of study, may have been inadequate to investigate in detail differential associations depending on specific characteristics of HT. In particular, the majority of women in our cohort were using combined HT (estrogen with a progestagen) and thus any associations observed in our analysis may be due to progestagen rather than estrogen per se. As in any genetic association study, bias from population stratification also needs to be considered and French law prohibits the collection of data related to ethnicity. However, the genotype frequencies across analysis sub-groups, is similar to those of Caucasian populations observed previously (Ioannidis et al., 2002). Finally, in terms of brain volumes, we did not investigate volumes within given structures except HC and CC, nor do we have longitudinal structural data, which would assist in determining causality. The strengths of our study are the inclusion of both older men and women and the study size which was larger than the majority of previous studies, thus enabling us to separate women using HT into current and past users. Given the study size and fact that the population was well characterised, has also enabled us to take into consideration a large number of potentially confounding factors. Furthermore, this is
one of the few studies to investigate associations with $ESR$ variants, and the first to look directly for an $ESR$-HT (gene-environment) interaction.

**4.4 Conclusion and perspectives**

This study does not support a beneficial effect of HT on brain volumes and suggests rather that current use may be associated with reduced grey matter volume and higher brain atrophy. Our findings also suggest for the first time that genetic variations in the $ESR1$ may be associated with the risk of WML in older postmenopausal women, which should thus be considered in future analyses. Large prospective studies are needed to see whether these findings can be replicated and to help ascertain causality. Such studies could also examine whether women with certain genotypes have differential changes in brain volumes over time in response to HT.

**5. Acknowledgments**

The ESPRIT project was financed by the regional government of Languedoc-Roussillon and an unconditional grant from Novartis. Part of this project is financed by grants from the Agence Nationale de la Recherche (projects ANR 2007-LVIE-004 and 2007-LVIE-005-01) and France Alzheimer. Joanne Ryan is the holder of an NHMRC Training (Postdoctoral) Fellowship (Overseas Public Health, APP1012735). None of these sponsors had any involvement or say in the study design, data collection, data analysis or in the decision to publish.

**Disclosure statement:** Dr. Ritchie serves on scientific advisory boards for the Biomedical Research Centre, King’s College London, and London and MRC Strategic Steering Committee (Longitudinal Health and Aging Research Unit). All other authors report no actual or potential conflicts of interest.
6. References


Interactive effects of fitness and hormone treatment on brain health in postmenopausal women. Neurobiol Aging. 28, 179-85.


with bone mineral density and fracture risk in women: a meta-analysis. J Bone Miner Res. 17, 2048-60.


Robertson, D., Craig, M., Van Amelsvoort, T., Daly, E., Moore, C., Simmons, A., Whitehead, M., Morris, R. & Murphy, D., 2009 Effects of estrogen therapy on age-related differences in gray matter concentration. Climacteric. 12, 301-9.


