**RESEARCH PAPER**

**Candidate-gene analysis of white matter hyperintensities on neuroimaging**

Theresa Tran, Ioana Cotlarciuc, Sunaina Yadav, Nazeeha Hasan, Paul Bentley, Christopher Levi, Bradford B Worrall, James F Meschia, Natalia Rost, Pankaj Sharma

**ABSTRACT**

**Background** White matter hyperintensities (WMH) are a common radiographic finding and may be a useful endophenotype for small vessel diseases. Given high heritability of WMH, we hypothesised that certain genotypes may predispose individuals to these lesions and consequently, to an increased risk of stroke, dementia and death. We performed a meta-analysis of studies investigating candidate genes and WMH to elucidate the genetic susceptibility to WMH and tested associated variants in a new independent WMH cohort. We assessed a causal relationship of WMH to methylene tetrahydrofolate reductase (MTHFR).

**Methods** Database searches through March 2014 were undertaken and studies investigating candidate genes in WMH were assessed. Associated variants were tested in a new independent ischaemic cohort of 1202 WMH patients. Mendelian randomization was undertaken to assess a causal relationship between WMH and MTHFR.

**Results** We identified 43 case-control studies interrogating eight polymorphisms in seven genes covering 6,314 WMH cases and 15,461 controls. Fixed-effects meta-analysis found that the C-allele containing genotypes of the aldosterone synthase CYP11B2 T(−344)C gene polymorphism were associated with a decreased risk of WMH (OR=0.61; 95% CI, 0.44 to 0.84; p=0.003). Using mendelian randomisation the association among MTHFR C677T, homocysteine levels and WMH, approached, but did not reach, significance (expected OR=1.75; 95% CI, 0.90–3.41; observed OR=1.68; 95% CI, 0.97–2.94). Neither CYP11B2 T(−344)C nor MTHFR C677T were significantly associated when tested in a new independent cohort of 1202 patients with WMH.

**Conclusions** There is a genetic basis to WMH but anonymous genome wide and exome studies are more likely to provide novel loci of interest.

**INTRODUCTION**

White matter hyperintensities (WMH) are defined as diffuse white matter abnormalities detected on T2-weighted or fluid attenuated inversion recovery (FLAIR) MRI and appearing as regions of low attenuation on brain CT scans. They are a common radiological finding, particularly in older individuals, with a reported prevalence of up to 95%. These lesions, of presumed vascular origin, may represent an endophenotype for small cerebral vessel diseases such as stroke and dementia; thus, WMH could be used in early diagnosis and guided management of these conditions. WMH has consistently been associated with increasing age and hypertension, as well as smoking, previous stroke or TIA, and elevated homocysteine (Hcy) levels. Studies have reported high heritability estimates ranging between 55% and 71%, implying a significant genetic component to WMH development.

To date, a single WMH GWAS has been published by the CHARGE consortium that identified six single nucleotide polymorphisms (SNPs) in a novel locus on chromosome 17q25 associated with WMH burden in stroke-free participants. The most significantly associated SNP on 17q25 was rs3744028 with a reported p value of $1.0 \times 10^{-9}$ after adjustment for hypertension. This association between the 17q25 locus and WMH has recently been tested in a cohort of ischaemic patients with stroke, where it has replicated in association with WMH volume but not lacunar stroke status. The latter may suggest that these two diseases have distinct pathogeneses of cerebral microangiopathy.

Several statistically underpowered small candidate gene studies on WMH have been published, but the results remain invalid due to low power. By consolidating data from these smaller studies, a literature-based meta-analysis is considered to be the next best way to increase power and find a true genetic risk association. We conducted a comprehensive meta-analysis of all case-control studies investigating candidate genes in WMH, tested our findings in a new independent WMH cohort and sought to identify a causal relationship with methylene tetrahydrofolate reductase (MTHFR).

**METHODS**

A comprehensive search strategy in electronic databases (PubMed, Google Scholar, Embase) was undertaken using a range of search terms for WMH (leukoaraiosis, white matter hyperintensities, white matter lesions, white matter disease, age-related white matter changes, homocysteine, hyperhomocysteinemia) in combination with the...
Boolean operator AND/OR (genetics, genotype, genes or polymorphism). Further searches were conducted for each gene identified, using specific gene names combined with the WMH search terms. Additional studies were found by hand searching reference lists of relevant papers. For duplicate papers, the largest cohort was selected. A variety of different methods were utilised to quantify WMH levels or volumes, but in general visual rating scales were more commonly used than automated programmes. Most papers reported genotype frequencies within the Hardy–Weinberg equilibrium. Our study complied with PRISMA guidelines.

**Study selection**

Inclusion criteria were: (1) case–control studies where WMH was reported as a grade on a standardised scale or as a volume, (2) WMH was objectively confirmed by MRI or CT brain scans, (3) genotype frequency was reported for WMH cases and controls. Studies were excluded if they did not explicitly distinguish WMH from other brain lesions such as lacunes and microinfarcts. For the Mendelian randomisation part of this study, additional selection was based on plasma Hcy levels for cases and controls in the studies of participants of European descent reporting SDs associated with mean Hcy levels.

**Data extraction**

Some studies quantified WMH grade on a standardised scale and presented this as dichotomous data, so for each genotype the number of participants in the highest and lowest WMH grade groups was extracted. Where studies subdivided WMH into two categories: deep (or subcortical) WMH and periventricular WMH, without providing data for WMH as a whole, data for the deep WMH was analysed. In cases where the scale cut-off could be chosen, the upper grade group included Fazekas scale 2 or 3 or equivalent. For continuous WMH grade data, the mean grade and SD for each genotype were taken, and for studies presenting WMH data as a volume, we extracted the mean volume and SD for each genotype.

**Data analysis for meta-analysis**

Data were analysed using Review Manager V5.2. Using a Mantel–Haenszel statistical method, a pooled OR and 95% CI were calculated for each SNP-WMH association. Statistical significance was set at p<0.05. Where significant heterogeneity was detected, a random-effects analysis model was utilised in order to account for this interstudy heterogeneity. In all other cases, a fixed effects model was used. Heterogeneity was assessed with an F test for each meta-analysis (significance set at p<0.10) and an iterative analysis was performed where significant heterogeneity was found. Publication bias was assessed with Funnel plots and by performing Egger’s regression analysis (two-tailed tests) using Comprehensive Meta-Analysis V2.0 (CMA).

**Data analysis for Mendelian randomisation**

Mendelian randomisation allows the testing of a causal effect of observed data in the presence of confounding factors. Review Manager was used to calculate an OR and 95% CI for MTHFR-WMH grade association using the TT vs CC model so as to be in keeping with the model used by Casas et al15 who report a weighted mean difference in the Hcy level between TT and CC-genotype to be 1.93 μmol/L in their meta-analysis. A pooled mean difference in Hcy levels between cases with WMH and controls was calculated and then converted into a corresponding OR of WMH for that specific increase in Hcy level using CMA software. The expected OR was then calculated using the following formula:

Expected OR = $X^{Y/z}$

where $X=$ the OR of risk of WMH for a Z μmol/L increase in plasma Hcy levels.

And Y=the mean difference in the Hcy level (μmol/L) between TT and CC-genotype participants.

To calculate the 95% CI for the expected OR, we took the natural log of this number to determine the logged OR. The 95% CI for this logged OR is calculated by taking 1.96×SE on either side of this logged OR. The SE is taken as the square root of the sum of the reciprocals of the number of cases and controls. The exponential function in Excel was used to convert the upper and lower CI limits into the 95% CI limits for the original OR.

**Replication of the associated genetic variants**

Any associated genetic variants were tested in a cohort of 1202 ischaemic stroke cases of European ancestry with genome-wide genotyping available (‘MGH,’ ‘IGS,’ ‘ASGS,’ ‘SWISS’ cohorts) that was previously used to replicate chromosome 17q25 locus association with WMH.15 In this cohort, WMH volume (WMHv) was measured using a previously validated, semiautomated volumetric method (facilitated by MRicro, University of Nottingham School of Psychology, Nottingham, UK; http://www.mricro.com). For this analysis, MRI scans obtained from a 1.5 T scanner were converted from DICOM into Analyse format, and the contiguous, supratentorial axial T2 FLAIR sequences were cross-referenced with diffusion-weighted images (DWI) and examined to exclude hyperintensities that represent oedema, acute ischaemia or chronic infarcts. WMHs were manually outlined as regions of interest, and their intersections with automatically derived intensity thresholds were manually examined and touched up by a trained reader. The total WMHv was calculated by doubling the measurement from the hemisphere unaffected by stroke, or by adding bilateral WMHv values in participants with infratentorial DWI lesions. To control for variation in head size, the intracranial area (ICA) was calculated from two middle sagittal T1 slices and used to normalise WMHv by multiplying it by the individual-to-the-mean ICA ratio.20 21 Specific study characteristics and genotyping information of this cohort are previously described.15

**RESULTS**

Our initial search identified 1248 studies with 20 additional records from references of relevant articles. Removal of duplicates and matches to our predefined inclusion and exclusion criteria resulted in 43 studies which had available data for meta-analysis interrogating nine polymorphisms in 10 different genes ([figure 1](#figure1)). The majority used MRI to assess WMH in (mostly Caucasian) participants aged between 60 and 80 years. Most of the genes studied were involved in the renin-angiotensin-aldosterone system (ACE, angiotensinogen (AGT), angiotensin II receptor 1, aldosterone synthase). The other identified genes had roles in Hcy levels (MTHFR), antiatherosclerosis (paraoxonase 1) and cholesterol regulation (apolipoprotein E). Other gene polymorphisms studied were brain-derived neurotrophic factor/Val66Met22 and nitric oxide synthase 3/G894T23 24 but some studies had yet to be replicated25 and others did not have data available for meta-analysis24 (table 1).
MTHFR677 cytosine/thymine (TT vs CT/CC)

Six studies investigated the association between the MTHFR C677T polymorphism (TT vs CT/CC) and WMH (n=4002).\textsuperscript{25-30} Five studies (n=2988) assessed the genotype difference between the lower and upper WMH grade groups\textsuperscript{25,26,28-30} (figure 2 A) and one study (n=1014) compared the mean WMH volume between different genotypes\textsuperscript{27} (figure 2 B). A fixed-effects meta-analysis demonstrated a trend of increased risk burden of WMH with MTHFR TT compared to the CT/CC genotype (OR=1.19; 95% CI 0.95 to 1.51; p=0.14; I²=28%; p=0.24).

Aldosterone synthase CYP11B2 T(-344)C (CC/CT vs TT)

Two studies (n=1153) evaluated the association between the aldosterone synthase CYP11B2 T(-344)C polymorphism and dichotomous graded WMH, and the fixed-effects meta-analysis demonstrated that the C-allele-containing genotypes were at a reduced risk of white matter lesions (OR=0.98; 95% CI 0.87 to 1.11; p=0.78; I²=5%; p=0.39). Within participants with WMH, there was no significant predominance of the e4 allele-containing genotypes (WMH grade, continuous data, pooled standardised mean difference=0.29; 95% CI –0.03 to 0.61; p=0.07; I²=0%; p=0.87; WMH volume, pooled standardised mean difference=0.06; 95% CI –0.02 to 0.14; p=0.14; I²=47%; p=0.09).

Table 1  Summary table demonstrating each gene, polymorphism, model used, number of cases and controls and resulting OR, CI and p values

<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism</th>
<th>Model used</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>95% CI</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apolipoprotein E4</td>
<td>E4 carriers vs non-carriers</td>
<td></td>
<td>2614</td>
<td>5008</td>
<td>0.98</td>
<td>0.87 to 1.11</td>
<td>0.78</td>
</tr>
<tr>
<td>Apolipoprotein E2</td>
<td>E2 carriers vs non-carriers</td>
<td></td>
<td>248</td>
<td>569</td>
<td>1.42</td>
<td>0.46 to 4.43</td>
<td>0.54</td>
</tr>
<tr>
<td>ACE</td>
<td>Insertion/deletion</td>
<td></td>
<td>756</td>
<td>1365</td>
<td>1.46</td>
<td>0.92 to 2.31</td>
<td>0.11</td>
</tr>
<tr>
<td>MTHFR C677T</td>
<td>TT vs CT/CC</td>
<td></td>
<td>800</td>
<td>2188</td>
<td>1.19</td>
<td>0.95 to 1.51</td>
<td>0.14</td>
</tr>
<tr>
<td>Angiotensigen M235T</td>
<td>TT/MT vs MM</td>
<td></td>
<td>328</td>
<td>806</td>
<td>1.12</td>
<td>0.84 to 1.50</td>
<td>0.44</td>
</tr>
<tr>
<td>Angiotensigen II receptor 1</td>
<td>CC vs AC/AA</td>
<td></td>
<td>105</td>
<td>354</td>
<td>1.23</td>
<td>0.59 to 2.54</td>
<td>0.58</td>
</tr>
<tr>
<td>Aldosterone synthase</td>
<td>Ti-344C vs TT</td>
<td></td>
<td>197</td>
<td>956</td>
<td>0.61</td>
<td>0.44 to 0.84</td>
<td>0.003</td>
</tr>
<tr>
<td>Paraoxonase L55M</td>
<td>LL/LM vs MM</td>
<td></td>
<td>77</td>
<td>266</td>
<td>1.42</td>
<td>0.61 to 3.28</td>
<td>0.41</td>
</tr>
</tbody>
</table>

MTHFR, methylene tetrahydrofolate reductase.
Ethnic differences in plasma Hcy levels are well documented with East Asians consistently reported to have significantly lower Hcy levels compared to Caucasians.67–70 (table 2). Given these ethnic disparities, we considered it appropriate to exclude studies of East Asian (ie, Japanese, Korean) participants from the Mendelian randomisation, as the majority of studies were conducted in participants of European descent. The remaining four studies covering 745 Caucasian participants were meta-analysed and a comparison of WMH cases vs controls found a pooled mean difference in plasma Hcy levels of 3.71 μmol/L (95% CI 2.79 to 4.63; p<0.00001; I²=0%) (figure 3). CMA V.2.0 software was used to calculate the corresponding pooled OR of risk of WMH for this mean difference in Hcy levels using a fixed effects analysis model, OR=2.93 (95% CI 2.18 to 3.94).

In a meta-analysis of 42 studies, we had previously examined the effect of MTHFR on plasma Hcy levels in healthy participants (n=15 635) and reported the weighted mean difference in the Hcy level between TT and CC-genotype to be 1.93 μmol/L (95% CI 1.38 to 2.47; p<0.0001).17 Using these three pieces of data, the expected OR was calculated using the following formula:

\[
\text{Expected OR} = \frac{2.93}{1.93} = 1.57
\]

where:
- 2.93 is the OR of risk of WMH for a 3.709 μmol/L increase in plasma Hcy levels,
- 1.93 is the mean difference in the Hcy level (μmol/L) between TT and CC-genotype participants.

### Table 2 Ethnic differences in plasma homocysteine level between East Asian and Caucasian participants

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>n</th>
<th>Mean Hcy± SD (μmol/L)</th>
<th>p Value</th>
<th>Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anand et al68</td>
<td>Europeans</td>
<td>326</td>
<td>10.0±3.8</td>
<td>0.02</td>
<td>Chinese had significantly lower Hcy levels.</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>317</td>
<td>9.2±3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carmel et al69</td>
<td>White</td>
<td>237</td>
<td>14.8*</td>
<td>&lt;0.05</td>
<td>Whites had higher Hcy concentrations than Asian-Americans.</td>
</tr>
<tr>
<td></td>
<td>Asian-Americans</td>
<td>68</td>
<td>12.8*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senaratne et al70</td>
<td>Caucasians</td>
<td>106</td>
<td>10.8±0.6</td>
<td>&lt;0.001</td>
<td>East Asians had significantly lower plasma Hcy compared to Caucasians.</td>
</tr>
<tr>
<td></td>
<td>East Asians (Chinese, Japanese)</td>
<td>17</td>
<td>7.6±0.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*SD not reported by study.

Hcy, total homocysteine.
To calculate the 95% CI for the expected OR of 1.75, we took the natural log of 1.75 to get the logged OR of 0.56. The 95% CI for this logged OR was −0.108 to 1.227 and was calculated by taking $1.96 \times SE$ on either side of 0.56. The SE was 0.34, which was calculated as the square root of the sum of the reciprocals of the number of cases and controls. Using the exponential function in excel, these limits were converted into the 95% CI limits for the original OR of 1.75 giving EXP (−0.108) $=0.897$ to EXP (1.227)=3.412. From our meta-analysis of two studies investigating the association between MTHFR (TT vs CC) and WMH grade, the observed OR for WMH was 1.68 (95% CI 0.97 to 2.94, $p=0.07$). Despite the results not reaching statistical significance, the similarity between the expected and observed ORs supports a likely causal relationship between MTHFR and WMH.

**Replication of associated genetic variants**

We examined the association between CYP11B2T(-344)C and MTHFR C677T polymorphisms and WMH quantified on the MRI using a validated, semiautomated volumetric protocol in an independent cohort of 1202 ischaemic stroke cases with WMH. There was no association between either polymorphism and the WMH volume in this cohort (CYP11B2T(-344)C, $P=0.5755$; MTHFR C677T, $P=0.68$).

**DISCUSSION**

In this largest study to date of candidate genes in WMH burden, we interrogated 6253 WMH cases and 15 239 controls for eight polymorphisms in seven genes (APO E/e4 and e2, ACE insertion/deletion, MTHFR C677T, AGT M235T, AGTR1 A1166C, CYP11B2 T344C, PON1 L55M). Our analysis demonstrated a likely genetic effect for ischaemic white matter disease, with an apparent inverse association between CYP11B2 and WMH. A trend for positive association between MTHFR and WMH severity could not be completely interrogated, given the relatively small sample size of the available studies as compared to other well-powered genetic studies on stroke.

The MTHFR gene is involved in plasma Hcy levels and may contribute to endothelial dysfunction, which is one of the suggested mechanisms behind WMH. While the association between the MTHFR TT-genotype and WMH fell shy of statistical significance, the totality of the data suggested a trend for association. The Mendelian randomisation approach allowed us to investigate any potential relationship between MTHFR and WMH in more depth and evaluate for potential causality. The specific strengths of this method are that confounding factors are equally distributed among genotypes, which facilitates causality to be tested in their presence, whereas measurement error bias, reverse causality and selection biases are largely overcome. Using this approach yielded similar values for the expected and observed OR, and there was considerable overlap of their 95% CIs. Given that the two values are derived from meta-analyses of the different study types (genetic association vs observational)—either of which is prone to a different source of bias—might be suggestive of a causal association between Hcy and WMH burden. However, this analysis was insufficiently powered, and future studies using an adequate sample size may prove more conclusive.

A number of study limitations need to be documented. Publication bias is always a concern in meta-analysis. However, funnel plots were produced for each gene-WMH association and Egger’s regression analysis (two-tailed test) was performed to assess publication bias. Given that the majority of included studies reported non-significant results, substantial publication bias is considered unlikely, although it can never be completely excluded. Further, of the 78 studies investigating the association of candidate genes and WMH, just under half did not have usable data for meta-analysis. It may be that the authors of these papers did not consider the data to be interesting enough to report (selective outcome reporting). The vast majority of these studies found no association and their inclusion would have strengthened our finding of no relationship between WMH and any of the studied gene polymorphisms. Some studies reported data according to a genetic model they had chosen rather than reporting event rates for each genotype separately, which limited our ability to incorporate their studies into other genetic models. There was considerable variation in WMH measurement methods used between studies, which introduces methodological heterogeneity. Assessing WMH using visual rating scales can be subjective and observer dependent, although most papers reported good inter-rater agreement. The inclusion of CT and MRI studies adds another source of intersubject heterogeneity. CT has been shown to be less sensitive at detecting WMH and its use may result in an underestimation of the true WMH load within those studies. However, removal of these studies did not lead to a substantial change in the pooled OR; thus, we considered it appropriate to include them. Finally, results could be confounded by failure to adjust for age, intracranial volume and vascular risk factors in all studies. Additionally, consideration ought to be given in analyses to the known WMH risk factors since genes may be exerting their effect through these factors. The variable disease status of the study populations could have introduced heterogeneity into our analysis. For example, six studies in our APO E4 analysis were conducted in participants with probable or pathologically confirmed Alzheimer’s disease. Combining these studies with those of asymptomatic participants could have confounded our results. Finally, a number of covariates which we are not able to assess because of a lack of complete data sets may influence our final results.

Despite undertaking, to the best of our knowledge, the largest meta-analysis to date along with studying a new independent WMH cohort, the genetics of this condition remains unclear. Future genetic studies not using an a priori hypothesis may shed further light on this field.

**Author affiliations**

1 Institute of Cardiovascular Research, Royal Holloway University of London (ICR2UL) and Ashford & St Peters NHS Foundation Trust, London, UK
2 Imperial College Cerebrovascular Research Unit (ICCRU), Imperial College London, London, UK
3 Department of Neurology, John Hunter Hospital, Hunter Medical Research Institute, Newcastle, New South Wales, Australia
4 Departments of Neurology and Public Health Sciences, University of Virginia Health System, Charlottesville, Virginia, USA

To calculate the 95% CI for the expected OR of 1.75, we took the natural log of 1.75 to get the logged OR of 0.56. The 95% CI for this logged OR was −0.108 to 1.227 and was calculated by taking $1.96 \times SE$ on either side of 0.56. The SE was 0.34, which was calculated as the square root of the sum of the reciprocals of the number of cases and controls. Using the exponential function in excel, these limits were converted into the 95% CI limits for the original OR of 1.75 giving EXP (−0.108) $=0.897$ to EXP (1.227)=3.412. From our meta-analysis of two studies investigating the association between MTHFR (TT vs CC) and WMH grade, the observed OR for WMH was 1.68 (95% CI 0.97 to 2.94, $p=0.07$). Despite the results not reaching statistical significance, the similarity between the expected and observed ORs supports a likely causal relationship between MTHFR and WMH.
Cerebrovascular disease

5Department of Neurology, Mayo Clinic, Jacksonville, Florida, USA
6Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

Acknowledgements PS is a former Department of Health (UK) Senior Fellow. PB holds a Department of Health (UK) Senior Fellowship.

Contributors PS conceived the work. TT undertook the primary search. NR undertook the replication. All authors provided data and contributed to the final draft of the manuscript.

Funding NH is funded by a MRC (UK) PhD studentship and SY through a British Council/UK-India Education Research Initiative (UKIERI) awarded to PS.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES
Cerebrovascular disease


