Homocysteine and stroke: evidence on a causal link from mendelian randomisation

Lancet 2005; 365: 224–32

See Comment page 194

Juan P Casas, Leonelo E Bautista, Liam Smeeth, Pankaj Sharma, Aroon D Hingorani**

Centre for Clinical Pharmacology, Department of Medicine, British Heart Foundation Laboratories, Royal Free and University College Medical School, University College London, Rayne Building, 5 University Street, London WC1E 6JJ, UK (J P Casas MD, A D Hingorani FRCP); Population Health Sciences, University of Wisconsin Medical School, Madison, WI, USA (L E Bautista DrPH); Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK (L Smeeth MRCGP); and Hammersmith Hospitals Acute Stroke Unit and Department of Cellular and Molecular Neuroscience, Imperial College London, London, UK (P Sharma PhD)

Correspondence to: Dr Aroon D Hingorani a.hingorani@ucl.ac.uk or Dr Pankaj Sharma psharma@imperial.ac.uk

*These authors contributed equally to the study

Summary

Background Individuals homozygous for the T allele of the *MTHFR* C677T polymorphism have higher plasma homocysteine concentrations (the phenotype) than those with the CC genotype, which, if pathogenetic, should put them at increased risk of stroke. Since this polymorphism is distributed randomly during gamete formation, its association with stroke should not be biased or confounded. We investigated consistency between the expected odds ratio for stroke among TT homozygotes, extrapolated from genotype–phenotype and phenotype–disease studies, and the observed odds ratio from a meta-analysis of genotype–disease association studies.

Methods We searched MEDLINE and EMBASE up to June, 2003, for all relevant studies on the association between homocysteine concentration and the *MTHFR* polymorphism, and until December, 2003, for those on the association between the polymorphism and the risk of stroke. Pooled odds ratios and 95% CI were calculated by random-effects and fixed-effects models. Consistency between expected and observed odds ratios was assessed by interaction test.

Findings 111 studies met the selection criteria. Among 15 635 people without cardiovascular disease, the weighted mean difference in homocysteine concentration between TT and CC homozygotes was 1·93 µmol/L (95% CI 1·38 to 2·47). The expected odds ratio for stroke corresponding to this difference based on previous observational studies was 1·20 (1·10 to 1·31). In our genetic meta-analysis (n=13 928) the odds ratio for stroke was 1·26 (1·14 to 1·40) for TT versus CC homozygotes, similar to the expected odds ratio (p=0·29). Consistency between the odds ratios was preserved in analyses by age-group, ethnic background, and geographical location.

Interpretation The observed increase in risk of stroke among individuals homozygous for the *MTHFR* T allele is close to that predicted from the differences in homocysteine concentration conferred by this variant. This concordance is consistent with a causal relation between homocysteine concentration and stroke.

Introduction

Stroke is the third most common cause of death in more developed countries.[1] About 80% of strokes are thromboembolic (ischaemic) in origin and the remainder are haemorrhagic.[1–3] In the UK, stroke is the largest single cause of severe disability with more than 125 000 incident strokes and about 60 000 deaths due to stroke each year.[2] Because treatments for stroke are limited, the best approach to reduce mortality and morbidity is primary prevention through modification of acquired risk factors (eg, high blood pressure, smoking, diabetes, and atrial fibrillation).[3]

Data from cohort and case-control studies suggest that a raised circulating concentration of homocysteine is associated with a higher risk of stroke.[4–6] However, homocysteine concentration is also related to smoking status, blood pressure, and social class and is higher in people with existing atherosclerosis than in those without.[7,8] Therefore, this relation could be subject to residual confounding, reverse-causality bias, or both.[7,8]

A common functional polymorphism, C677T, in the gene encoding methylenetetrahydrofolate reductase (MTHFR), an enzyme involved in homocysteine metabolism, has been associated with differences in homocysteine concentration.[4,9,10] Since carriage of this variant is subject to the random assortment of maternal and paternal alleles at the time of gamete formation, according to Mendel's second law[11] associations between *MTHFR* genotype and homocysteine concentration or stroke should not be subject to reverse-causality bias and should also be largely free from confounding by other determinants of homocysteine concentration or risk factors for stroke.[9] Moreover, genotype is a fixed characteristic, so there is unlikely to be regression-dilution bias, which results from measurement error and biological variability of the exposure under assessment and leads to underestimation of the association between a risk factor and disease.[12,13] Therefore, if homocysteine increases the risk of stroke, carriage of the *MTHFR* polymorphism that exposes individuals to an increased homocysteine concentration should confer an increased risk of stroke proportional to the difference in homocysteine concentration attributable to variant and to the relative risk observed in non-genetic observational studies.

The investigation of consistency between risk estimates obtained from genotype–disease studies and those from phenotype–disease studies, to provide insight into the nature of the observed associations, has been referred to as "mendelian randomisation".[12,13] This approach has been used to test the nature of the association between homocysteine concentration and coronary heart disease, venous thromboembolism, and stroke.[4,9] However, in the previous genetic analyses of

the *MTHFR* polymorphism and stroke, the number of available studies was small. Also, information on the relative effect of the *MTHFR* polymorphism on homocysteine concentration in different geographical locations, ethnic groups, and age-groups was limited.[4] Demonstration that homocysteine concentration is causally related to the development of stroke would have important implications in primary prevention, because administration of folic acid is known to lower homocysteine concentrations,[14] and a policy of fortification of cereal with folic acid to lower the incidence of neural-tube defects has already been undertaken in North America.

We have carried out two updated and comprehensive meta-analyses. In the first, we investigated the extent to which homocysteine concentrations are determined by the *MTHFR* C677T polymorphism. We also assessed the effect of pre-existing cardiovascular disease, geographical location, ethnic background, and age on this association. In the second meta-analysis, we estimated the odds ratio of stroke conferred by the TT genotype. Our meta-analysis of *MTHFR* C677T polymorphism and homocysteine concentration includes data from 48 more studies than a previous meta-analysis[4] and extends the observations to people without cardiovascular disease. Our meta-analysis of the *MTHFR* C677T polymorphism and stroke includes data from 23 more studies than the previous meta-analysis[4] and assessed the robustness of the results by sensitivity analysis.

Methods

Two electronic databases (MEDLINE and EMBASE) were searched up to June, 2003, for all studies on the association between the *MTHFR* C677T polymorphism and homocysteine concentrations and up to December, 2003, for studies on the association between the *MTHFR* C677T polymorphism and stroke. For the first search, we used the text words, which were also MeSH terms, "polymorphism", "mutation", "genes", and "cardiovascular disease" in combination with "homocysteine". For the second search, the terms used were "stroke", "brain infarction", "cerebral isch(a)emia", "h(a)emorrhagic stroke" and "silent brain infarction" in combination with "genetic", "polymorphism", "mutation", or "genes". Both literature searches were limited to "human" and "English Language". We searched for any additional studies in the references of all identified publications, including previous relevant meta-analyses.[4,15,16]

Selection criteria

For inclusion in the meta-analysis on homocysteine concentrations and the *MTHFR* polymorphism, studies had to have an analytical design (case-control, cohort, or cross-sectional) and had to examine the association between homocysteine concentrations and the polymorphism. Studies were included only if they were published as full-length articles or letters in peer-reviewed journals. For duplicate publications the smaller dataset was excluded.

For inclusion in the meta-analysis on *MTHFR* and stroke, studies had to involve unrelated individuals and to examine the associations between ischaemic or haemorrhagic stroke and the presence of the polymorphism. For the main comparison, only studies published as full-length articles or letters in peer-reviewed journals in English were included. However, to assess the robustness of the association, we did a sensitivity analysis that included silent brain infarction as an outcome, non-full-text papers, and papers published in languages other than English. In all searches, when relevant information was not reported or there was doubt about duplicate publications, we contacted the authors to obtain the required information.

Data extraction

Data for analysis (country of origin, study design, mean age of participants, frequency of genotypes and alleles, homocysteine and folate concentrations, ethnic background, and frequency of cardiovascular risk factors) were extracted and entered into databases by two of us (JPC and PS). The results were compared and disagreements resolved by consensus.

Statistical analysis

We obtained a summary estimate of the effect of raised plasma homocysteine concentrations on risk of stroke from a recently published meta-analysis by Wald and others of eight prospective studies, which included 676 stroke cases, mainly ischaemic in aetiology, in white people, carried out in European and North American countries.[4] In that analysis, a difference of 5 μmol/L in plasma homocysteine was associated with an odds ratio for stroke of 1·59 (95% CI 1·29 to 1·96) adjusted for confounding variables and for regression-dilution bias.[4] We did a meta-analysis to obtain the weighted mean difference in plasma homocysteine concentrations between individuals homozygous for the T allele and those homozygous for the C allele. The weighted mean difference was obtained separately for people with and without known cardiovascular disease (ischaemic heart disease, stroke, or venous thrombosis) and for both groups combined. For these analyses, a random-effects model[17] was used to allow for any heterogeneity across studies. Estimates of the weighted mean difference were obtained for different age-groups, by geographical location, and by ethnic background. We then did a second meta-analysis of all published studies to obtain a summary odds ratio for all strokes for individuals homozygous for the T allele compared with those homozygous for the C allele. Fixed-effects summary odds ratios and 95% CI were calculated by the Mantel-Haenszel method,[18,19] and DerSimonian and Laird's method[17] was used to calculate random-effects summary odds ratios and

their 95% CI. Also, to test the robustness of our findings, we calculated different odds ratios according to outcome (ischaemic stroke confirmed by MRI or CT, haemorrhagic stroke, or silent brain infarction), ethnic background (white and non-white), publication language (English and other), and type of publication (full text or abstract).

We used the DerSimonian and Laird *Q* test[20] to assess the degree of heterogeneity between studies, and funnel plots and Egger's regression asymmetry test to assess small-study bias, of which publication bias is one potential cause.[21] In addition, the influence of individual studies on the summary odds ratio was investigated by re-estimation and plotting of the summary odds ratio in the absence of each study. Meta-regression was used to assess the extent to which different variables explained heterogeneity in the weighted mean difference and in the summary odds ratios.[22] Finally, we used the weighted mean difference in homocysteine concentration by *MTHFR* C677T polymorphism to estimate an expected increase in the risk of stroke assuming that an increase of 5 μmol/L in plasma homocysteine would be associated with an odds ratio for stroke of 1·59 (1·26–1·96) and that this association follows a log-linear relation.[4] The uncertainty surrounding the expected odds ratio is a function of the variability of the weighted mean difference in homocysteine by genotype and the variability of the summary odds ratio from Wald's meta-analysis[4] and cannot be directly calculated. Therefore, to obtain a 95% CI for this odds ratio, we generated a million values from a normal distribution with mean and SD equal to the weighted mean difference and its SE, and a million values from a normal distribution with mean and SD equal to the natural logarithm of the summary odds ratio from Wald's meta-analysis[4] and its SE (calculated from its 95% CI). We used the simulated values to calculate a million estimates of this odds ratio and took the 2·5% and 97·5% centile values of the created empirical distribution as 95% confidence limits. Then the expected odds ratio was compared with the summary odds ratio obtained from the meta-analysis of genetic studies by means of an interaction test.[23] Consistency between the two odds ratios would indicate that the association between plasma homocysteine concentration and stroke seen in non-genetic epidemiological studies is unlikely to be severely affected by residual or reverse-causality bias. Data were analysed by use of the Review Manager software (version 4.2) from the Cochrane Collaboration 2003 and Stata (version 8.0).

Role of the funding source

No funding source had any role in study design; collection, analysis, or interpretation of data; or the writing of the report. All authors had full access to all the data in the study, and all took full responsibility for the decision to submit the paper for publication.

Results

The primary search for studies on homocysteine and the *MTHFR* C677T polymorphism generated 104 potentially relevant studies (see webreferences, numbered w1–w145, at http://image.thelancet.com/extras/03art11437webreferences.pdf) of which 81 met the selection criteria (w1–w81). Of the 23 articles excluded, 16 (w82–w97) did not provide sufficient data for us to calculate the weighted mean difference between the genotypes, and the relevant information could not be obtained from the authors. Two (w98, w99) were discarded as probable duplication, and five (w100–w104) reported only the homocysteine concentrations for TT homozygotes and C-allele carriers but not for the CC genotype alone.

Of the 81 studies included (31 355 individuals), information for the main comparison (TT *vs* CC genotype) was available for 15 635 people without known cardiovascular disease, 6312 with cardiovascular disease, and for 9408 reported only as the combination of individuals with and without cardiovascular disease.

Among individuals without cardiovascular disease (webreferences w1–w41), the weighted mean difference in plasma homocysteine between those homozygous for the T allele and those homozygous for the C allele was 1·93 μmol/L (95% CI 1·38 to 2·47; p<0·0001, figure 1). There was significant inter-study heterogeneity (p for heterogeneity <0·0001). From the variables examined in a meta-regression analysis, the only major source of heterogeneity detected was the mean concentration of serum folate. The weighted mean difference in plasma homocysteine concentrations comparing the TT and CC genotypes was 0·048 μmol/L less for each 1 nmol/L increase in mean serum folate (p=0·035). The weighted mean difference in homocysteine concentration was lower in studies done in North America than in those done in Europe or other continents (table 1). Conversely, the weighted mean serum folate concentration was higher in North American (25·3 nmol/L [7·3 to 43·3]) than in European studies (13·6 nmol/L [11·4 to 15·6]) and those conducted in other continents (15·3 nmol/L [9·0 to 21·6]). No other sources of heterogeneity by age (p=0·94), ethnic background (white *vs* other; p=0·62), smoking status (p=0·41), or sex (p=0·57) were observed in a meta-regression analysis of studies in which these variables were reported.

The distribution of the weighted mean difference in relation to its SD in the funnel plot was symmetrical, and the result of Egger's test was not significant (p=0·72) providing no positive evidence for small-study bias. Visual assessment of a graph of the individual weighted mean difference for each study showed that none of the studies had an undue influence on the overall weighted mean difference.

Among people with cardiovascular disease (w2, w4, w7, w11, w16, w21, w26, w27, w29, w32, w37, w39– w55), the

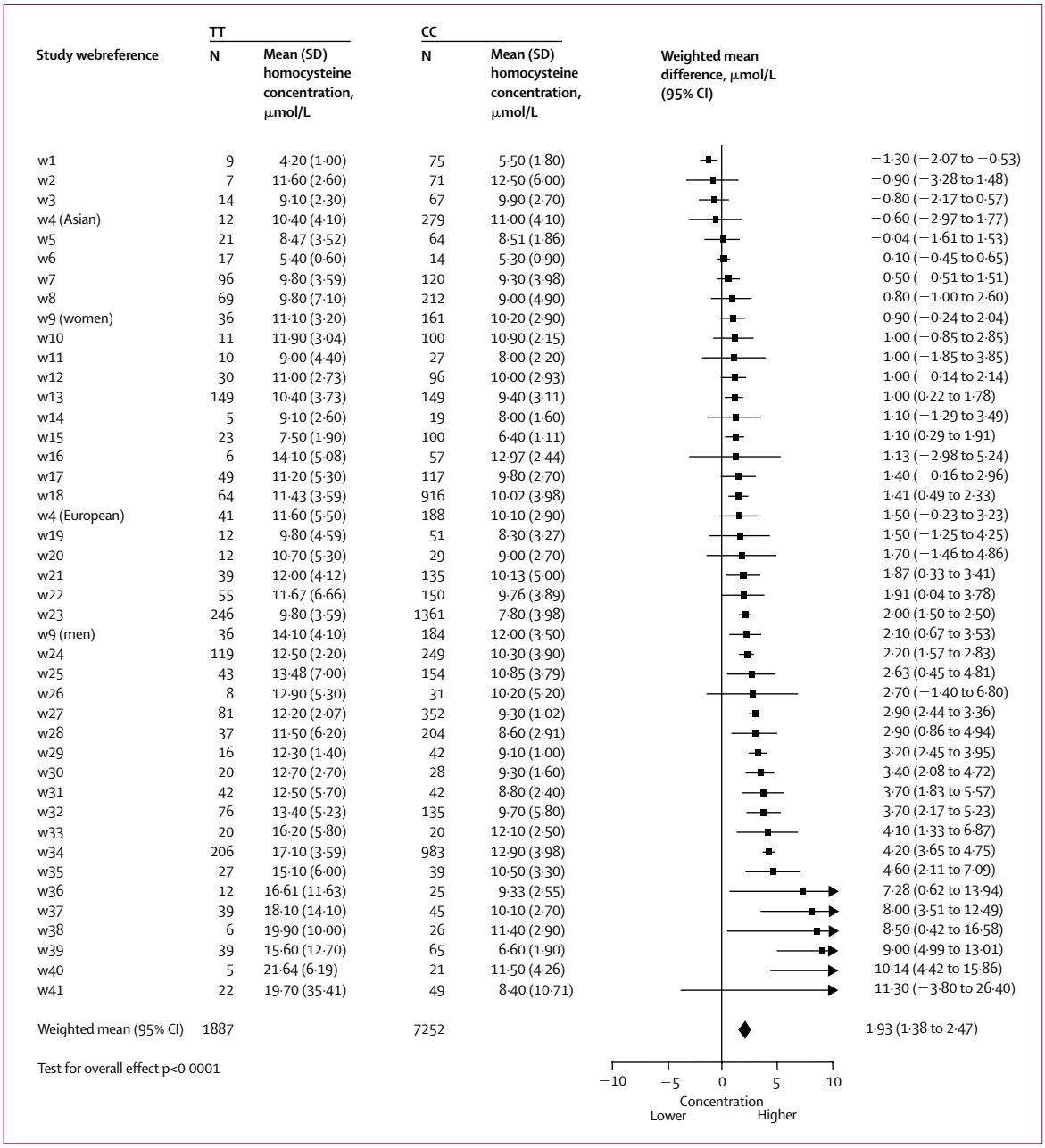

Figure 1: **Weighted mean differences in plasma homocysteine concentration according to the *MTHFR* C677T genotype (TT *vs* CC) among people without known cardiovascular disease**

For webreferences see http://image.thelancet.com/extras/03art11437webreferences.pdf. The total number of people without known cardiovascular disease included in the meta-analysis was 15 635 (TT 1887, CT 6496, CC 7252).

mean homocysteine concentration was 4·35 μmol/L (3·22 to 5·49; p<0·0001) higher for those homozygous for the T allele than for those homozygous for the C allele. A meta-regression analysis indicated that the greater difference in homocysteine by genotype among people with cardiovascular disease was explained partly by lower serum folate concentrations (the crude β coefficient for disease status was 3·04; after adjustment for serum folate it was 2·73).

When the data from all studies, including those that investigated individuals with and without cardiovascular disease without distinction, were combined (w1–w81), the weighted mean difference for homozygotes for the T allele against homozygotes for the C allele was 3·10 μmol/L (2·54 to 3·65; p<0·0001).

According to a previous meta-analysis of prospective studies,[4] a plasma homocysteine concentration higher by 5 μmol/L corresponds to an odds ratio for stroke of 1·59

	Number of participants	Number of studies	Weighted mean difference, μmol/L (95% CI)	p
All studies	9139	43	1·93 (1·38 to 2·47)	<0·0001
Ethnic background				
White	6948	32	1·96 (1·41 to 2·51)	<0·0001
Other	2191	11	1·83 (0·58 to 3·09)	0·004
Mean age				
40 years or older	8575	36	1·96 (1·46 to 2·47)	<0·0001
Less than 40 years	564	7	1·66 (0 to 3·31)	0·05
Geographical location				
Europe	7031	26	2·04 (1·45 to 2·64)	<0·0001
North America	1027	8	0·57 (−0·28 to 1·42)	0·19
Other continents	1081	9	2·97 (1·72 to 4·22)	<0·0001
European and North-American, mean age 40 years or older	7742	29	1·84 (1·29 to 2·39)	<0·0001
White people mean age 40 years or older	6747	29	1·95 (1·39 to 2·51)	<0·0001

Table 1: **Weighted mean difference in plasma homocysteine concentrations according to the *MTHFR* C677T genotype (TT vs CC) among people without cardiovascular disease**

(1·29 to 1·96). A difference in homocysteine concentration of 1·93 μmol/L (1·38 to 2·47) in healthy individuals with the TT genotype would therefore result in an expected odds ratio for stroke of 1·20 (1·10 to 1·31) compared with individuals homozygous for the C allele, if the association between homocysteine and risk of stroke follows a log-linear relation (table 2), and is free from confounding and reverse-causality bias. Expected odds ratios for stroke derived from differences in homocysteine by genotype in all individuals without cardiovascular disease, and also separately by ethnic background, geographical location, and mean age are summarised in table 2.

Of the 49 potentially relevant studies identified in the primary search for the meta-analysis on the *MTHFR* C677T polymorphism and stroke, 30 met the selection criteria (w40, w47, w53, w59, w71, w105–w128). Of the 19 articles excluded for the purpose of the main comparison, three (w129–w131) were published in non-English journals, two (w132, w133) reported the outcome silent brain infarction, and five (w134–w138) were published only as abstracts. Though excluded from the main comparison, these studies were used in a sensitivity analysis. Of the remaining nine publications excluded, five studies (w62, w139–w142) encompassed some duplication. Four (w103, w143–w145) did not report the genotype frequency, and the relevant information could not be obtained from the authors. Of the 30 studies (a total of 6324 cases and 7604 controls) included, 19 were in white people, ten involved Asian participants, and one included both white individuals and people of Afro-Caribbean origin.

	Expected odds ratio for stroke (95% CI)*	Observed genetic odds ratio (95% CI)†	p‡
Main comparison			
All studies	1·20 (1·10 to 1·31)	1·26 (1·14 to 1·40)	0·29
Secondary comparisons			
Studies with mean age 40 years or older	1·20 (1·10 to 1·30)	1·26 (1·14 to 1·40)§	0·30
Europe and North America only	1·17 (1·09 to 1·27)	1·21 (1·02 to 1·43)	0·37
White people only	1·20 (1·10 to 1·31)	1·19 (1·02 to 1·39)	0·39

All odds ratios are based on comparisons of TT *vs* CC genotypes. *Odds ratios were calculated only for individuals without cardiovascular disease. Mean expected odds ratios were calculated with the formula: expected odds ratio=1·59 raised to the power of d/5, where d=weighted mean difference in homocysteine by genotype, on the assumption that a 5 μmol/L increase in homocysteine is associated with an odds ratio for stroke of 1·59.[4] †Odds ratios obtained from the meta-analysis of genotype–disease association studies of *MTHFR* C677T and stroke. ‡For the comparison by use of an interaction test. §The weighted mean age was 58 years for cases and 53 years for controls.

Table 2: **Assessment of consistency between odds ratios derived by extrapolation from phenotype–disease studies and those derived from meta-analysis of genetic studies**

The summary odds ratio, under a fixed-effects model, indicated that individuals with the TT genotype compared with those homozygous for the C allele had an odds ratio for stroke of 1·26 (1·14 to 1·40; p<0·0001; figure 2). There was significant heterogeneity among the results of individual studies (p for heterogeneity=0·034). A sensitivity analysis showed that the study by Morita and colleagues (w47) was the main cause of the heterogeneity. After exclusion of this study, the heterogeneity was no longer significant (p=0·32) but the estimate of the overall effect changed very little and remained significant (odds ratio 1·20 [1·08 to 1·34]; p=0·0006). Similarly, a random-effects model that took into account the variability within and between studies resulted in a similar overall estimate (1·26 [1·07 to 1·47]; p=0·004). A meta-regression analysis showed that ethnic background (white *vs* other, p=0·23) and the presence of risk factors such as age (p=0·38), sex (p=0·46), hypertension (p=0·15), smoking (p=0·43), and diabetes (p=0·10) were not significant sources of heterogeneity in a group of 20 studies (w47, w53, w59, w105, w108, w109, w111, w116–w127) with information on all these variables.

The distribution of the odds ratio in relation to its SD in the funnel plot was symmetrical, and the result of Egger's test was not significant (p=0·32) providing no positive evidence of small-study bias. No individual study had an undue influence on the summary odds ratio. A sensitivity analysis showed a robust association between the *MTHFR* C677T polymorphism and stroke (figure 3). No significant changes in the summary odds ratio were detected after addition of studies published in languages other than English (w129–w131) or in abstract form (w134–w138), or those that used the outcome silent brain infarction (w132, w133). Similarly, no differences in the summary odds ratio were observed when the analysis was restricted according to ethnic background or to studies in which the outcome was solely ischaemic stroke (figure 3). Data from five studies (w106, w110, w111, w118, w129; 611 cases and 2405 controls) for which information was available on the association of the *MTHFR* C677T polymorphism and haemorrhagic stroke (TT *vs* CC) gave a summary odds ratio of 1·16 (0·90 to 1·50; p=0·25), under a fixed-effects model.

Estimates of odds ratios for TT homozygous individuals in comparison with those who had the CC genotype, extrapolated from the genotype effect on homocysteine and the homocysteine–stroke association, were mathematically very similar to those from our genotype–disease meta-analysis (table 2). The p value for interaction for the main comparison of the study, the non-genetic and genetic odds ratios for stroke in TT versus CC individuals, derived from all the data, was 0·29. Therefore, there was no evidence of a significant difference between the expected odds ratio estimated from the increment in the homocysteine concentrations by genotype (*MTHFR* C677T) and that observed from the genotype–disease meta-analysis. Moreover, no significant differences between estimated odds ratios were observed when similar comparisons were done among more homogeneous groups defined by geographical location, ethnic background, or mean age (table 2). In all the comparisons, the expected odds ratios were within the 95% CI of the corresponding observed odds ratio obtained from the genetic association studies.

Discussion

The main finding of these meta-analyses was that the odds ratio for stroke conferred by the *MTHFR* TT genotype was similar to that estimated by use of the homocysteine difference by genotype and homocysteine–stroke odds ratio from phenotype–disease studies. Indeed, after exclusion of the study of Morita and colleagues (w47) which was the cause of much of the heterogeneity in the genotype–stroke analysis, the summary odds ratio for the *MTHFR* TT genotype was 1·20 (95% CI 1·08 to 1·34), identical to the predicted effect of the polymorphism estimated from the studies of genotype–homocysteine differences and homocysteine–stroke risk. Because of the random allocation of genotype in advance of disease development, these results imply that the relation between homocysteine concentration and stroke seen in phenoype–disease studies is not subject to substantial residual confounding or reverse-causality bias. Thus our study provides evidence for a role for homocysteine in stroke pathogenesis.

Our first meta-analysis of studies examining the association between homocysteine concentration and *MTHFR* genotype, involving more than 31 000 people, allowed us to refine the size estimate of the effect of the *MTHFR* C677T polymorphism on plasma homocysteine concentration and to explore potential sources of heterogeneity. The absolute difference in homocysteine concentration conferred by the genotype, though consistent in direction, was greater in people with established atherosclerosis than in individuals who were healthy at the time of measurement. A meta-analysis of individual patients' data had similar findings.[9] In that study, among individuals with coronary heart disease, there was a difference in homocysteine concentration between TT and CC individuals of 2·2 μmol/L; the difference among healthy individuals was smaller, but still significant (1·5 μmol/L).[9] Therefore, our data suggest the presence of an additional effect of disease status on the association between homocysteine

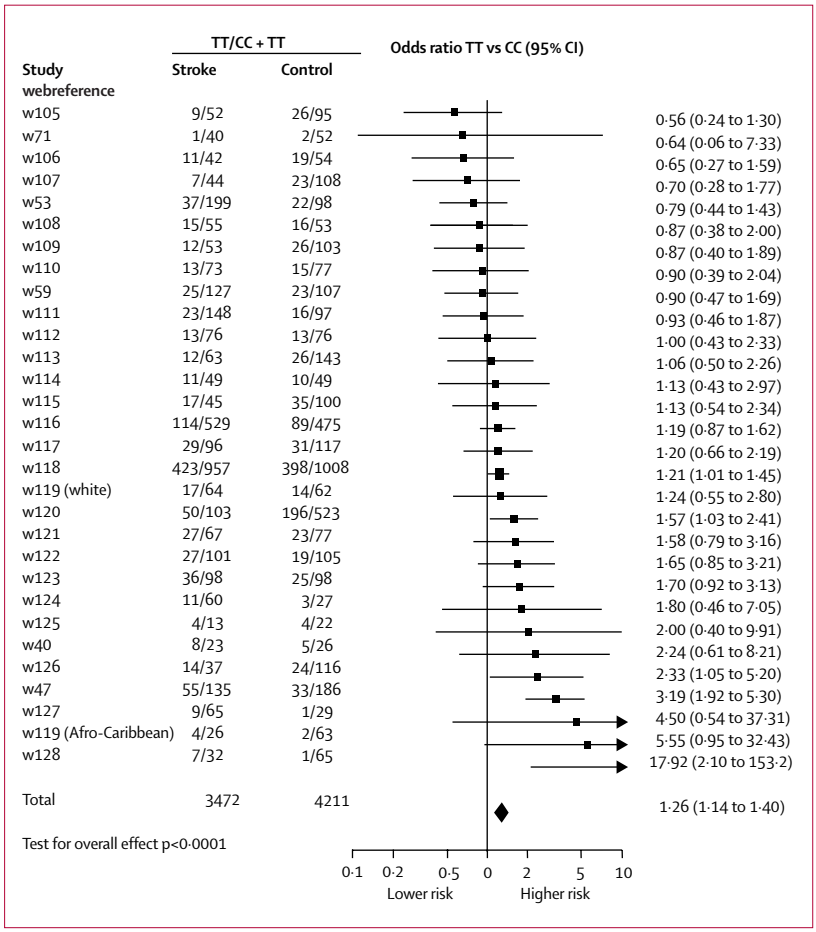

Figure 2: **Odds ratio for stroke in individuals with the TT genotype compared with those homozygous for the C allele of the *MTHFR* C677T polymorphism**

For webreferences see http://image.thelancet.com/extras/03art11437webreferences.pdf. The total number of stroke cases was 6324 (TT 1041, CT 2852, CC 2431) and the total number of controls was 7604 (TT 1140, CT 3393, CC 3071).

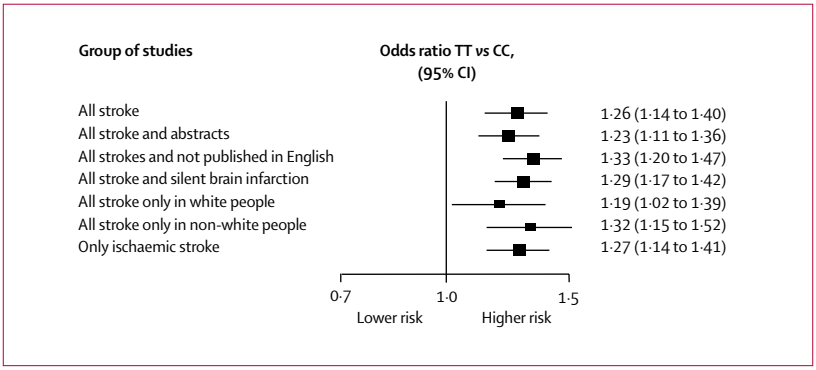

Figure 3: **Sensitivity analysis of the *MTHFR* C677T polymorphism and stroke association**

Summary odds ratio of stroke for TT INDIVIDUALS compared WITH CC INDIVIDUALS by publication type, stroke classification or type, and ethnic group.

concentration and *MTHFR* genotype, partly explained by lower serum folate concentrations in people with atherosclerotic disease. We therefore used estimates from healthy individuals only when calculating the predicted odds ratio for stroke among TT homozygotes. In estimation of an association between genotype and intermediate phenotype, the potential modifier effect of disease status should be examined.[24,25]

The second genetic meta-analysis (*MTHFR* C677T and stroke), which included about 14 000 individuals, allowed us to obtain a precise estimate of the effect of *MTHFR* genotype on stroke risk. Individuals homozygous for the T allele had an odds ratio for stroke of 1·26 (1·14 to 1·40) compared with those homozygous for the C allele. This difference in risk is similar to that expected from the difference in homocysteine concentrations by genotype in healthy individuals (1·20 [1·10 to 1·31]). Moreover, when the comparison was restricted to white individuals or to studies done in Europe or North America, within a similar age range to the population included in the meta-analysis of prospective studies of homocysteine and stroke risk,[4] similar results were obtained.

Clinical studies have shown that supplementation with folic acid and vitamin B12 lowers homocysteine concentrations by about 3 μmol/L.[14,26] If homocysteine is causally associated with an increased risk of stroke, nutritional interventions to lower concentrations might be expected to produce a relative-risk reduction in the incidence of stroke of about 23%.[5] Several randomised clinical trials are currently investigating the effects on cardiovascular outcomes of lowering homocysteine concentration by administration of B vitamins.[27,28] One clinical trial in North America[29] did not detect a beneficial effect of high versus low folate doses in secondary prevention of stroke. However, as suggested by the trial investigators and the results of our study, a larger sample size, longer periods of intervention, and targeting of the intervention to populations with low folate concentrations might all be required to show any potential benefit of this intervention. Such randomised intervention trials of the effect of lowering homocysteine concentrations are important for several reasons: the type of study we have done has several potential limitations; intervention trials are necessary to assess reversibility and could establish precisely the magnitude of any treatment effect; and they could ascertain whether there are any unexpected adverse effects of such therapies.

Despite the association found between the *MTHFR* C677T polymorphism, homocysteine concentration, and stroke risk, the size of the effect is modest compared with those of classic cardiovascular risk factors and does not necessarily provide a rational basis for screening for the polymorphism or for the measurement of homocysteine concentration in isolation, in the prediction of stroke. Whether either of these measurements would add useful predictive information to more established risk prediction tools (eg, Framingham risk equation) will require further investigation.

Our analyses must be interpreted in the context of the limitations of the available data. In the meta-analysis of the *MTHFR* variant and homocysteine concentrations, we found significant heterogeneity. This finding is perhaps not surprising, because expression of the mutation is likely to depend on environmental factors such as folate concentrations, as detected in our meta-regression analysis. This idea is supported by the smaller effect of the *MTHFR* C677T polymorphism on the homocysteine concentration in studies in North America, in which the mean concentrations of serum folate were higher. However, an overview of data from individual patients will be required for more precise quantification of this potential gene–environment interaction.[9,10] For this reason, only the estimates of weighted mean difference derived from a random-effects model were used in this report.

Publication bias or small-study bias was considered as an explanation for the observed associations between the *MTHFR* polymorphism, homocysteine concentration, and stroke risk, but the results obtained from the funnel-plot analysis and Egger's tests did not provide positive evidence for such bias. Although confounding is less likely in analyses of an association of a genotype with disease, some imbalance in the distribution of cardiovascular risk factors by *MTHFR* genotype cannot be totally excluded. However, previous studies that have investigated the effect of the *MTHFR* variant on coronary heart disease have suggested that there was no major systematic confounding from other cardiovascular risk factors (eg, age, sex, hypertension, diabetes, obesity, or alcohol intake).[9,30]

The mendelian randomisation approach used here is a potentially useful tool to assess the nature of the observed associations between putative risk factors and disease. This approach overcomes some potential limitations of observational studies, such as confounding, reverse-causality bias, and regression-dilution bias.[12,13,31,32] However, it also has theoretical limitations, such as the potential for confounding of the association between genotype and intermediate phenotype by linkage disequilibrium with other genes. Similarly, associations of the genotype with environmental exposures that regulate the concentrations of the intermediate phenotype could lead to spurious genotype–disease association. Finally, population stratification (confounding of gene–disease association by ethnicity), multiple disparate (pleiotropic) effects of gene polymorphisms on more than one biological system, or compensatory biological adaptation to its effects on disease risk (canalisation) could distort the observed genotype–disease associations in one or other direction.[12,13,33] There are other practical limitations. Since the estimated relative risk arising from genotype

(eg, *MTHFR* variant) on the intermediate phenotype (eg, homocysteine concentration) will in many cases be small, very large genetic association studies or meta-analysis of smaller studies will be needed to assess the nature of such associations. Unless careful consideration is given to the inclusion or exclusion of studies used in any meta-analyses, and to the potential for publication bias, the comparison of genetic and non-genetic odds ratio estimates could become distorted.

Adequately powered randomised controlled trials of supplementation with folic acid (with or without B vitamins) will be necessary to validate the therapeutic approach of lowering homocysteine concentrations to prevent stroke and other cardiovascular events.[34] The findings of this study and previous analyses of the same type emphasise the potential importance of such trials. The mendelian randomisation approach we and others have used for homocysteine might also be useful in assessment of whether other "novel" risk factors for cardiovascular disease could have aetiological roles.

Contributors

Juan P Casas contributed to protocol design, data extraction, and statistical analysis. Leonelo E Bautista helped with design, statistical analysis, and interpretation of data. Liam Smeeth helped with statistical analysis. Pankaj Sharma contributed to protocol design and data extraction, and Aroon D Hingorani contributed to protocol design and interpretation of data; both these authors contributed equally. All the authors contributed to the writing and revision of the report.

Conflict of interest statement

ADH has received fees for lecturing on features of the assessment and management of cardiovascular-disease risk at meetings organised by a medical conference company with a pharmaceutical sponsor. PS has received honoraria for lecturing in industry-sponsored meetings, industry funding for attending national and international meetings, and research grants from pharmaceutical companies, and has been a paid consultant to the biotech industry. The other authors declare no conflicts of interest.

Acknowledgments

We thank W Lalouschek, T Katsuya, T Ogihara, E Pongracz, A Pezzini, R Grossmann, B Voetsch, J Loscalzo, L L Husemoen, G S Zhang, C Meisel, S Hustad, J Hung, B McQuillan, H Iso, and Y Moriyama for providing relevant information about the genotype frequencies, homocysteine concentrations, and data on cardiovascular risk factors from their studies. LS holds a Medical Research Council Clinician Scientist Fellowship. ADH holds a British Heart Foundation Senior Fellowship.

References

1 WHO. The World Health Report 2002: reducing risks, promoting healthy life. Geneva: WHO, 2002. Also available at http://www.who.int/whr/2002/en/ (accessed Oct 4, 2003).

2 Warlow CP. Epidemiology of stroke. *Lancet* 1998; **352** (suppl 3): SIII1–4.

3 Goldstein LB, Adams R, Becker K, et al. Primary prevention of ischemic stroke: a statement for healthcare professionals from the Stroke Council of the American Heart Association. *Circulation* 2001; **103:** 163–82.

4 Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ* 2002; **325:** 1202–06.

5 Homocysteine Studies Collaboration. Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. *JAMA* 2002; **288:** 2015–22.

6 Bautista LE, Arenas IA, Penuela A, Martinez LX. Total plasma homocysteine level and risk of cardiovascular disease: a meta-analysis of prospective cohort studies. *J Clin Epidemiol* 2002; **55:** 882–87.

7 Brattstrom L, Wilcken DE. Homocysteine and cardiovascular disease: cause or effect? *Am J Clin Nutr* 2000; **72:** 315–23.

8 Ueland PM, Refsum H, Beresford SAA, Vollset SE. The controversy over homocysteine and cardiovascular risk. *Am J Clin Nutr* 2000; **72:** 324–32.

9 Klerk M, Verhoef P, Clarke R, et al. MTHFR 677C→T polymorphism and risk of coronary heart disease: a meta-analysis. *JAMA* 2002; **288:** 2023–31.

10 Brattstrom L, Wilcken DE, Ohrvik J, Brudin L. Common methylenetetrahydrofolate reductase gene mutation leads to hyperhomocysteinemia but not to vascular disease: the result of a meta-analysis. *Circulation* 1998; **98:** 2520–26.

11 Morgan TH. Heredity and sex. New York: Columbia University Press, 1913.

12 Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; **32:** 1–22.

13 Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004; **33:** 30–42.

14 Clarke R, Armitage J. Vitamin supplements and cardiovascular risk: review of the randomized trials of homocysteine-lowering vitamin supplements. *Semin Thromb Hemost* 2000; **26:** 341–48.

15 Kelly PJ, Rosand J, Kistler JP, et al. Homocysteine, MTHFR 677C→T polymorphism, and risk of ischemic stroke: results of a meta-analysis. *Neurology* 2002; **59:** 529–36.

16 Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol* 2004; **61:** 1652–61.

17 DerSimonian R, Laird NM. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7:** 177–88.

18 Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22:** 719–48.

19 Robin JS, Greenland S, Breslow NE. A general estimator of the variance of the Mantel-Haenszel odds ratio. *Am J Epidemiol* 1986; **124:** 719–23.

20 Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in a meta-analysis. In: Egger M, Davey Smith G, Altman DG, eds. Systematic reviews in health care: meta-analysis in context. London: BMJ Publications, 2001: 285–312.

21 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315:** 629–34.

22 Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. *Lancet* 1998; **351:** 123–27.

23 Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003; **326:** 219.

24 Keavney BD, Youngman LD, Palmer A, et al. Large-scale test of hypothesized associations between polymorphisms of lipid-related genes and myocardial infarction in about 5000 cases and 6000 controls. *Circulation* 2000; **102** (suppl II)**:** 852 (abstr).

25 Youngman LD, Keavney BD, Palmer A, et al. Plasma fibrinogen and fibrinogen genotypes in 4865 cases of myocardial infarction and 6002 controls: test of causality by "Mendelian randomization". *Circulation* 2000; **102** (suppl II)**:** 32 (abstr).

26 Wald DS, Bishop L, Wald NJ, et al. Randomized trial of folic acid supplementation and serum homocysteine levels. *Arch Intern Med* 2001; **161:** 695–700.

27 Schnyder G, Roffi M, Flammer Y, Pin R, Hess OM. Effect of homocysteine-lowering therapy with folic acid, vitamin B12, and vitamin B6 on clinical outcome after percutaneous coronary intervention: the Swiss Heart study—a randomized controlled trial. *JAMA* 2002; **288:** 973–79.

28 VITATOPS Trial Study Group. The VITATOPS (Vitamins to Prevent Stroke) Trial: rationale and design of an international, large, simple, randomised trial of homocysteine-lowering multivitamin therapy in patients with recent transient ischaemic attack or stroke. *Cerebrovasc Dis* 2002; **13:** 120–26.

29 Toole JF, Malinow MR, Chambless LE, et al. Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: the Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. *JAMA* 2004; **291:** 565–75.

30 Russo GT, Friso S, Jacques PF, et al. Age and gender affect the relation between methylenetetrahydrofolate reductase C677T genotype and fasting plasma homocysteine concentrations in the Framingham Offspring Study Cohort. *J Nutr* 2003; **133:** 3416–21.

31 Keavney B. Genetic epidemiological studies of coronary heart disease. *Int J Epidemiol* 2002; **31:** 730–36.

32 Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* 2001; **358:** 1356–60.

33 Little J, Khoury MJ. Mendelian randomisation: a new spin or real progress? *Lancet* 2003; **362:** 930–31.

34 Nallamothu BK, Fendrick AM, Rubenfire M, Saint S, Bandekar RR, Omenn GS. Potential clinical and economic effects of homocyst(e)ine lowering. *Arch Intern Med* 2000; **160:** 3406–12.