

Causal Relationship of Susceptibility Genes to Ischemic Stroke: Comparison to Ischemic Heart Disease and Biochemical Determinants

Paul Bentley^{1*}, George Peck¹, Liam Smeeth², John Whittaker², Pankaj Sharma¹

1 Imperial College Cerebrovascular Research Unit, Clinical Neurosciences, Charing Cross Hospital, Imperial College London, London, United Kingdom, **2** Department of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, United Kingdom

Abstract

Interrelationships between genetic and biochemical factors underlying ischemic stroke and ischemic heart disease are poorly understood. We: 1) undertook the most comprehensive meta-analysis of genetic polymorphisms in ischemic stroke to date; 2) compared genetic determinants of ischemic stroke with those of ischemic heart disease, and 3) compared effect sizes of gene-stroke associations with those predicted from independent biochemical data using a mendelian randomization strategy. Electronic databases were searched up to January 2009. We identified: 1) 187 ischemic stroke studies (37,481 cases; 95,322 controls) interrogating 43 polymorphisms in 29 genes; 2) 13 meta-analyses testing equivalent polymorphisms in ischemic heart disease; and 3) for the top five gene-stroke associations, 146 studies (65,703 subjects) describing equivalent gene-biochemical relationships, and 28 studies (46,928 subjects) describing biochemical-stroke relationships. Meta-analyses demonstrated positive associations with ischemic stroke for factor V Leiden Gln506, ACE I/D, MTHFR C677T, prothrombin G20210A, PAI-1 5G allele and glycoprotein IIIa Leu33Pro polymorphisms (ORs: 1.11 – 1.60). Most genetic associations show congruent levels of risk comparing ischemic stroke with ischemic heart disease, but three genes—glycoprotein IIIa, PAI-1 and angiotensinogen—show significant dissociations. The magnitudes of stroke risk observed for factor V Leiden, ACE, MTHFR and prothrombin, but not PAI-1, polymorphisms, are consistent with risks associated with equivalent changes in activated protein C resistance, ACE activity, homocysteine, prothrombin, and PAI-1 levels, respectively. Our results demonstrate causal relationships for four of the most robust genes associated with stroke while also showing that PAI-1 4G/5G polymorphism influences cardiovascular risk via a mechanism not simply related to plasma levels of PAI-1 (or tPA) alone.

Citation: Bentley P, Peck G, Smeeth L, Whittaker J, Sharma P (2010) Causal Relationship of Susceptibility Genes to Ischemic Stroke: Comparison to Ischemic Heart Disease and Biochemical Determinants. PLoS ONE 5(2): e9136. doi:10.1371/journal.pone.0009136

Editor: Katrina Gwinn, Baylor College of Medicine, United States of America

Received: October 7, 2009; **Accepted:** January 4, 2010; **Published:** February 9, 2010

Copyright: © 2010 Bentley et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: PB & PS are supported by Senior Clinical Fellowships from Dept. of Health (UK). LS is supported by a Senior Clinical Fellowship from the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: p.bentley@imperial.ac.uk

Introduction

Stroke is one of the leading causes of death, disability, and health finance cost in both developed and developing world countries [1]. Understanding the genetic contributions to ischemic stroke is important not only so as to explain, or predict, the minority of cases that occur in the absence of well-established risk factors, such as smoking, hypertension and diabetes [2], but also to account for wide variability of stroke incidence within individuals who do harbour these common, acquired risk-factors [3]. Moreover, appreciating the biochemical basis for risk-associated genes can motivate novel therapeutic strategies, including pharmacogenomics [4].

As the cumulative number of studies reporting positive genetic associations with stroke increases, the main challenges are deciding which associations are reliable and robust, and then deciphering the role of putative gene effects in terms of causation [5]. The present study attempts to address these issues by firstly, presenting the most comprehensive meta-analysis to date of all candidate genetic polymorphisms associated with ischemic stroke. Secondly, we relate these observed gene effect sizes with those predicted from pathophysiologically-related studies.

Since many of the candidate genes tested for an association with ischemic stroke have originated from studies in ischemic heart disease, and given overlapping pathophysiologies of these two diseases [6], it is meaningful to investigate whether specific genetic polymorphisms associate with clinical arteriopathic syndromes in general, e.g. due to a tendency to stiffen arteries [7], or whether certain genes exert organ-specific effects [8–10]. Furthermore, where positive associations do exist between genes and stroke it is critical to validate whether these effects are consistent with the risks attributed to their putative biological intermediates. For example, if a stroke-associated gene is also associated with a prothrombotic tendency, then does the degree of thrombophilia imparted by the genotype-in-question associate with a similar degree of risk of stroke, using independent data sets? We attempted to answer this question for all robust positive gene associations using a method based upon mendelian randomization [11].

Results

Ischemic Stroke Candidate Gene Meta-Analysis

We identified 187 candidate genetic polymorphism case-control studies (References S1), incorporating 37,481 ischemic stroke cases

and 95,322 controls that fulfilled the inclusion criteria. Between them, 43 polymorphisms were interrogated in 29 genes, with the mean number of studies per candidate polymorphism being 6.6 (95% CIs 4.4 – 8.8). For 23 out of the 43 candidate polymorphisms (53%), the combined studies comprised >1000 cases (and >1000 controls) in aggregate. It is these that are focused on in the rest of the results. Note that these represent 16 out of 29 candidate *genes*, because for several genes more than one polymorphism was tested; PDE 4D (6 SNPs), angiotensinogen (2 SNPs) and HFE (2 SNPs) being those for which there are >1000 pooled cases. Eighteen polymorphisms (42% of the total; representing 11 genes) were investigated by at least one study comprising a total sample size of >1000.

Of the 23 genetic polymorphisms candidates tested in >1000 cases, six polymorphisms in six genes were found to show an overall significant effect, with no significant between-study heterogeneity (Figures 1–7). These were, in order of case-numbers: factor V Leiden Gln506, angiotensin converting enzyme (ACE) I/D, methylene tetrahydrofolate reductase (MTHFR) C677T, prothrombin G20210A, plasminogen activator inhibitor-1 5G allele and glycoprotein IIIa Leu33Pro. The summary ORs for

these genes ranged from 1.15 (95% CI: 1.06 – 1.25) for ACE I/D, to 1.60 (95% CI: 1.28 – 2.00) for prothrombin G20210A. The corresponding population attributable risks for the genes listed above are, respectively: 1.8%, 3.9%, 3.1%, 1.9%, 11.2% and 5.8% (total: 27.5%). The remainder 17 polymorphisms that were tested in >1000 pooled cases failed to demonstrate association with ischemic stroke (Figure 7; Figures S1). Within this group, ten polymorphisms showed between-study heterogeneity ($p < 0.05$).

Of the 20 candidate polymorphisms that were tested in <1000 cases, four were found to show positive associations: factor VII R353Q (793 cases), protein Z G79A (741 cases), glycoprotein 1b-alpha Met-Thr (564 cases), and intercellular adhesion molecule-1 E469K (356 cases).

For all of the above associations there was no publication bias towards smaller studies, as indicated by non-significant ($p > 0.1$) regression intercepts in Egger funnel-plots of effect size against sample variance [12].

In order to evaluate the relative efficiency of the candidate gene method over time, we divided studies according to whether our meta-analysis either identified, or failed to identify, a significant association, and plotted pooled case numbers for each of the

Factor V Leiden (Arg506Gln) Gln+ vs ArgArg

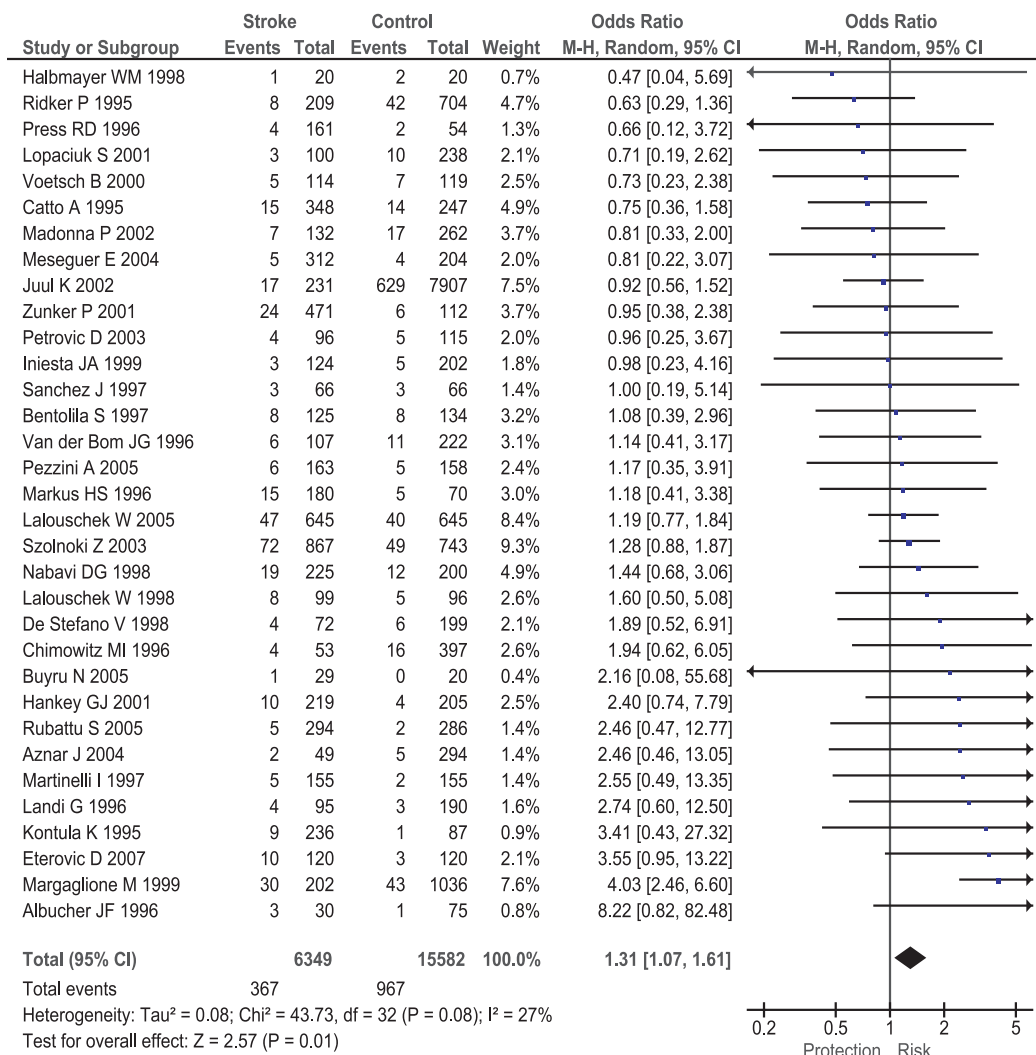


Figure 1. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: Factor V Leiden Arg506Gln.

doi:10.1371/journal.pone.0009136.g001

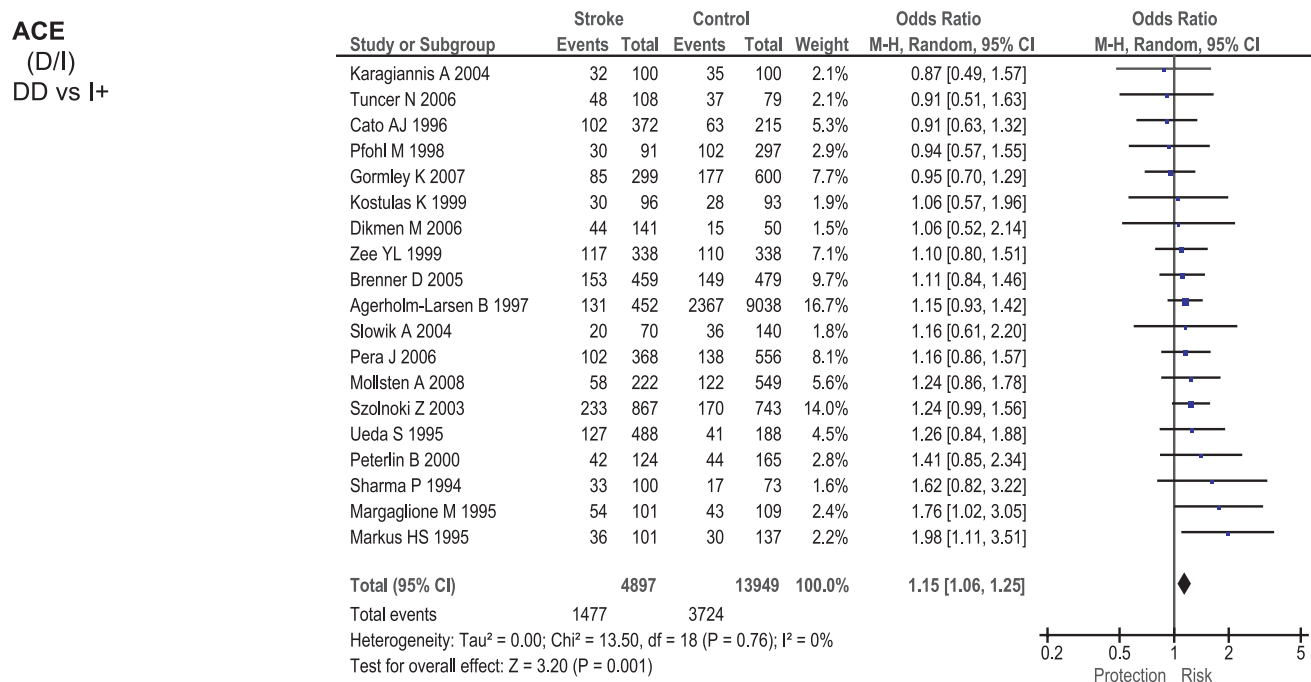


Figure 2. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: ACE D/I.
doi:10.1371/journal.pone.0009136.g002

largest polymorphisms against publication year (Figure 8 A and B). This shows that during the first decade of published studies (1993 – 2003) candidate polymorphisms were predominantly those found to be associated with stroke (according to our meta-analysis), whereas more recently (2004+), an increasing number of studied cases are for polymorphisms that show no association after pooling. Such declining success of candidate-gene studies is also seen by plotting the probability that cases were tested for polymorphisms that were found after meta-analysis to be associated, rather than unassociated, over time (Figure 8 C). We note that there were no correlations between OR magnitude, or study size, and publication date, when analyzing polymorphisms individually or grouping as a whole (all $p < 0.1$).

Comparison with Ischemic Heart Disease

Meta-analyses in myocardial infarction and/or ischemic heart disease were found [13–22] for thirteen of the genetic polymorphisms tested in ischemic stroke for which there were >1000 pooled cases. Comparing odds ratios for each genetic polymorphism between ischemic cardiac disease and ischemic stroke identified four profiles (Figure 9):

- 1) Polymorphisms associated with risk of *both* ischemic heart disease and stroke: factor V Leiden Gln506, ACE I/D, MTHFR C677T, prothrombin G20210A. The 95% confidence intervals (CIs) for ORs of ischemic stroke and cardiac disease overlapped for all of these gene variants.
- 2) Polymorphisms associated with risk of *either* ischemic stroke or myocardial ischemia but not the other disease type – i.e. dissociations: glycoprotein IIIa Leu33Pro conferring a risk for stroke, but not ischemic heart disease; PAI-1 4G- versus-5G associated positively with cardiac disease, but negatively with stroke, and angiotensinogen M235T posing a risk for coronary stenosis, but a trend for protection against stroke. The 95% CIs for risk for all three polymorphisms were non-

overlapping comparing ischemic stroke with ischemic heart disease, although restricting cardiac datasets to those in which Caucasian ethnicity was definitely specified results in marginal overlap for glycoprotein IIIa Leu33Pro (ischemic heart disease OR: 0.99 – 1.09) and angiotensinogen M235T (coronary stenosis OR: 1.0 – 1.16).

- 3) Polymorphisms associated with myocardial ischemia but not ischemic stroke, but where confidence intervals clearly overlapped. In these cases, we calculated the minimum case number required for 90% power (and $\alpha = 0.05$) assuming the same effect size as in ischemic heart disease. This showed that in all cases an inadequate sample size in stroke may account for the lack of significance here (Figure 9; third section, final column).
- 4) Polymorphisms not associated with either ischemic stroke or heart disease yet tested in >1000 cases of each disease viz. interleukin-6 G174C [21] and HFE C282Y [22].

Comparison with Biochemically-Predicted Risk

For each of the stroke-associated genetic polymorphisms identified by our primary meta-analysis we searched for: 1) studies providing differential measurements of a biochemical marker – i.e. an intermediate phenotype, IP – related to the genotypic variants of interest, among subjects *without* cardiovascular disease, and 2) studies providing an estimate of ischemic stroke risk based upon incremental change in the same IP. Such studies were available for the following polymorphism – IP pairs: factor V Leiden – activated protein C resistance, ACE D/I – ACE activity, MTHFR – homocysteine levels, prothrombin G20210A – prothrombin levels, and PAI-1 4G/5G – PAI-1 levels.

The number of studies providing differential IP levels according to genotype ranged between 3 and 70, incorporating between 843 and 46,743 healthy subjects, for each genotype of interest. We performed meta-analyses that pooled these biomarker level differences for each

**MTHFR
(C677T)
TT vs C+**

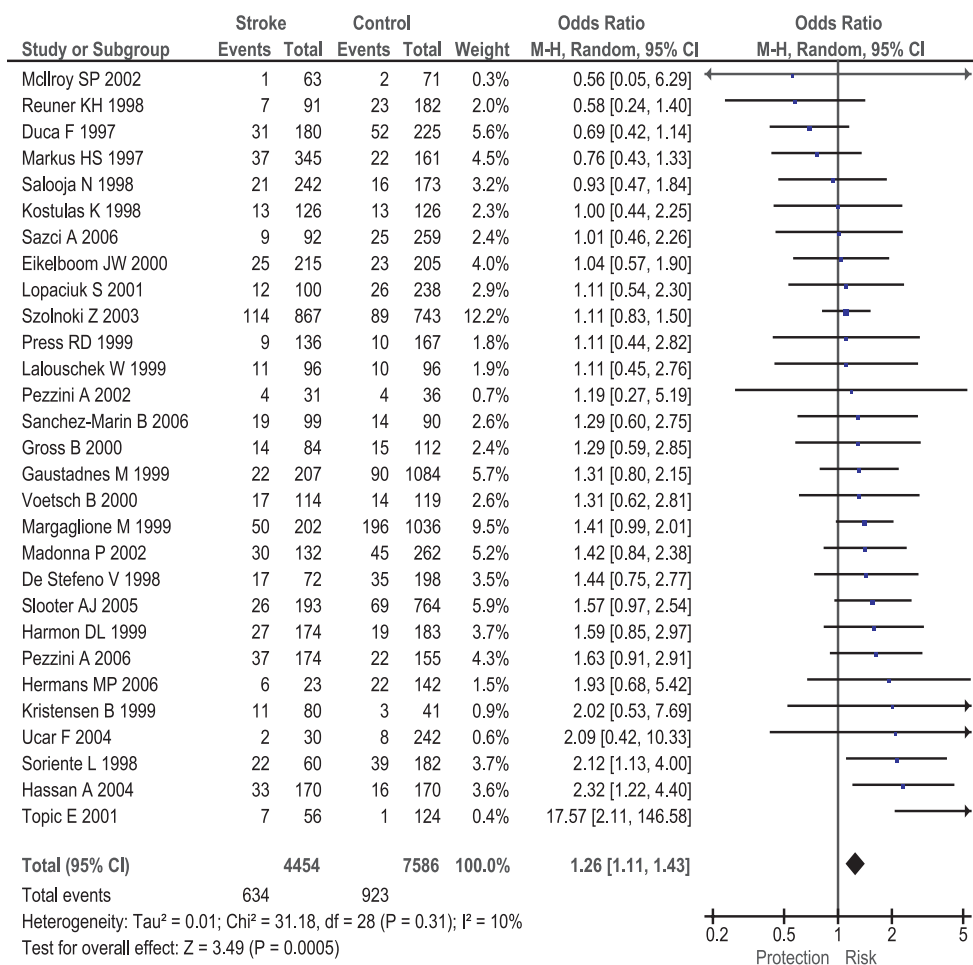


Figure 3. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: MTHFR C677T.
doi:10.1371/journal.pone.0009136.g003

of the six genotype – IP pairs (Figures 10 – 15), with the weighted mean differences, and 95% CIs, summarised in Figure 16 (2nd column).

The number of studies providing estimates of ischemic stroke risk relative to incremental changes in the above IPs ranged between 1 and 9, incorporating between 245 and 1688 cases, and 459 and 30574 controls. Where more than one study existed that provided an estimate of IP – stroke risk, we performed a meta-analysis that combined these ORs (Figures 10 and 13 – 15; lower forest plots), whilst first adjusting each value to the pooled estimate of biomarker level change for the genotype comparison of interest (see above), assuming a log-linear relationship. In the case of ACE activity and homocysteine levels, a single estimate of biomarker – stroke risk was extracted from a single study [23], and previous meta-analysis [24], respectively, and then scaled in the same way. The results of these calculations for expected risk of the genotype comparisons of interest (Figure 16, 4th column) were then compared with the observed risk for each of the same genotype contrasts, based upon data used in the first part of this paper (5th column).

As an example, Factor V Leiden 506 mutation ArgGln, relative to wild-type ArgArg, is associated with a weighted mean average of 0.74 APTT ratio decrease in the activated protein C resistance test in healthy subjects (Figure 10, upper). Independent data from ischemic stroke case – control studies suggest that a 0.74 decrease

in APTT ratio using the same test is associated with a risk odds-ratio of 1.31 (1.11 – 1.54) (Figure 10, lower). This estimated OR is very close to the observed OR of 1.30 (1.11 – 1.51) noted from our earlier meta-analysis (Figure 1).

For each of the gene – biomarker pairs (Figure 16), the mean level of expected risk fell close to, and certainly within the 95% confidence intervals, the observed risk, with the exception of PAI-1. For this latter pairing, the expected risk of the genotype comparison 5G5G vs 4G4G based upon PAI-1 levels was significantly lower than that observed. This arose because whilst PAI-1 levels are found to be higher in stroke than in controls, the 5G5G genotype – that itself is associated with elevated stroke risk relative to 4G4G – was associated with a ~25% reduction in PAI-1 levels relative to 4G4G. This result was unaffected by whether PAI-1 levels were derived from antigen concentration or enzyme activity data. We also estimated expected stroke risk on the basis of tPA levels, and its relationship with the PAI-1 4G/5G genotype. The 5G5G relative to 4G4G genotype was associated with a slight, albeit non-significant, increase in tPA levels, which itself is independently associated with stroke. However, the expected stroke risk using tPA data was also less than that observed.

Discussion

The current study takes a meta-analytic approach to identify the totality of candidate genetic polymorphisms reliably associated

Prothrombin
(G20210A)
A+ vs GG

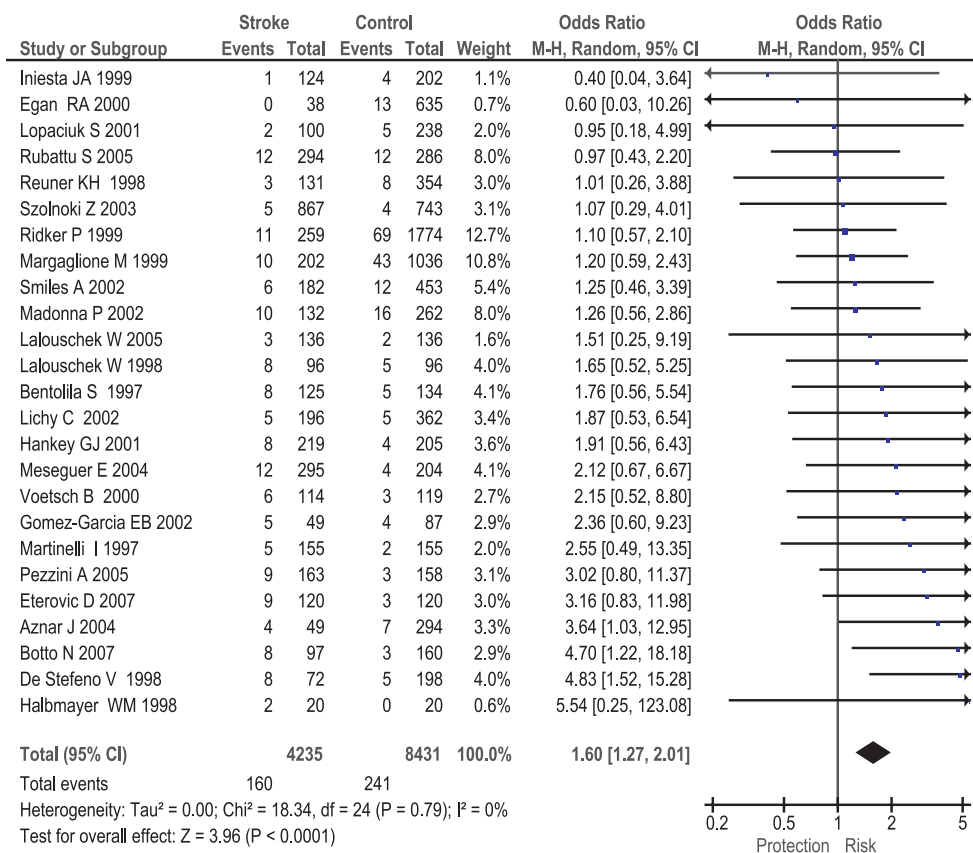


Figure 4. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: prothrombin G20210A.

doi:10.1371/journal.pone.0009136.g004

with ischemic stroke. Furthermore, we show how these results compare with the same polymorphisms in ischemic cardiac disease, and test their causal relationship with likely biochemical intermediaries. These points are discussed in turn.

Meta-Analysis of Genetic Polymorphisms in Ischemic Stroke

Since our last comprehensive meta-analysis of gene effects in ischemic stroke [25], the size of the pooled study sample from

PAI-1
(4G / 5G)
5G vs 4G Allele

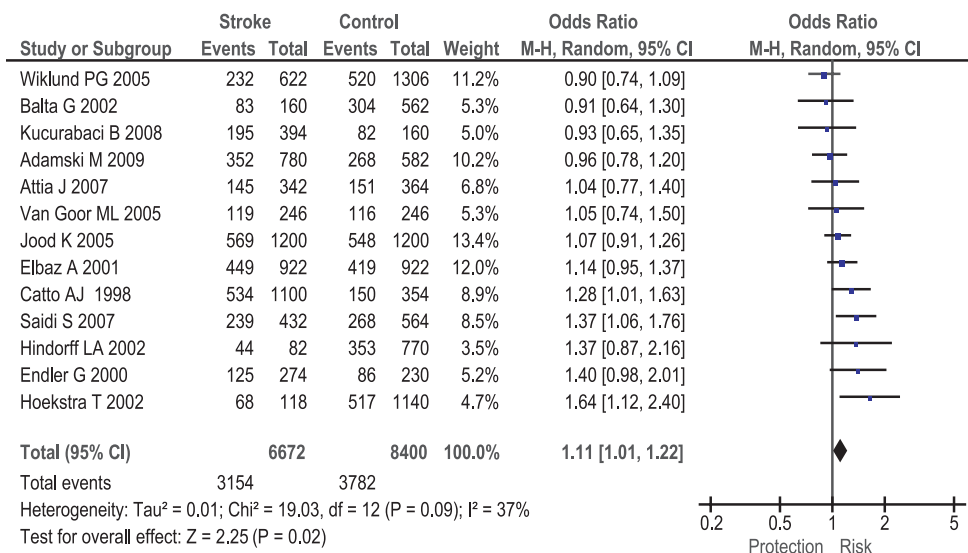


Figure 5. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: PAI 5G allele.

doi:10.1371/journal.pone.0009136.g005

**Glycoprotein IIIa
(Leu-to-Pro)
Pro+ vs LeuLeu**

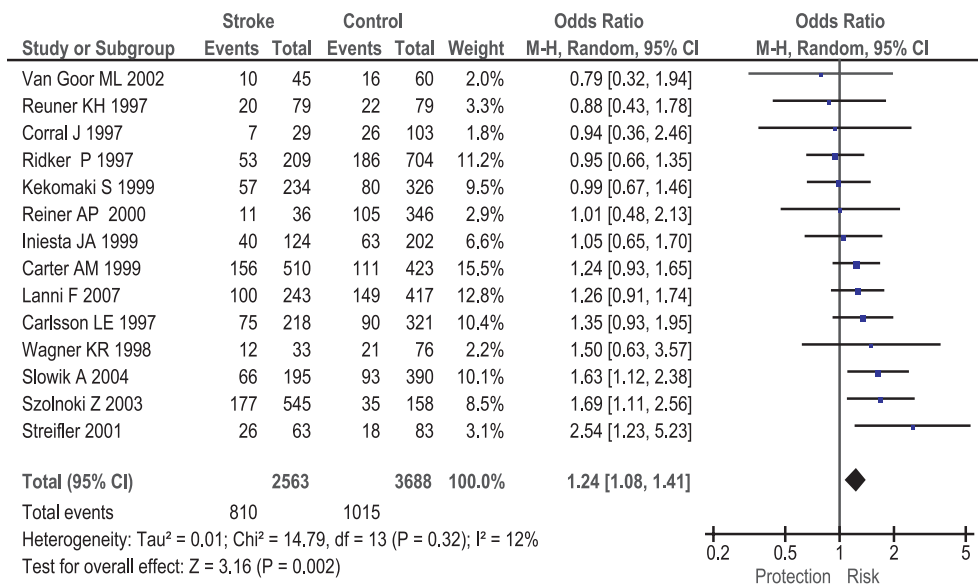


Figure 6. Forest plots showing positive associations of ischemic stroke with the following genetic polymorphisms: glycoprotein IIIa Leu33Pro.

doi:10.1371/journal.pone.0009136.g006

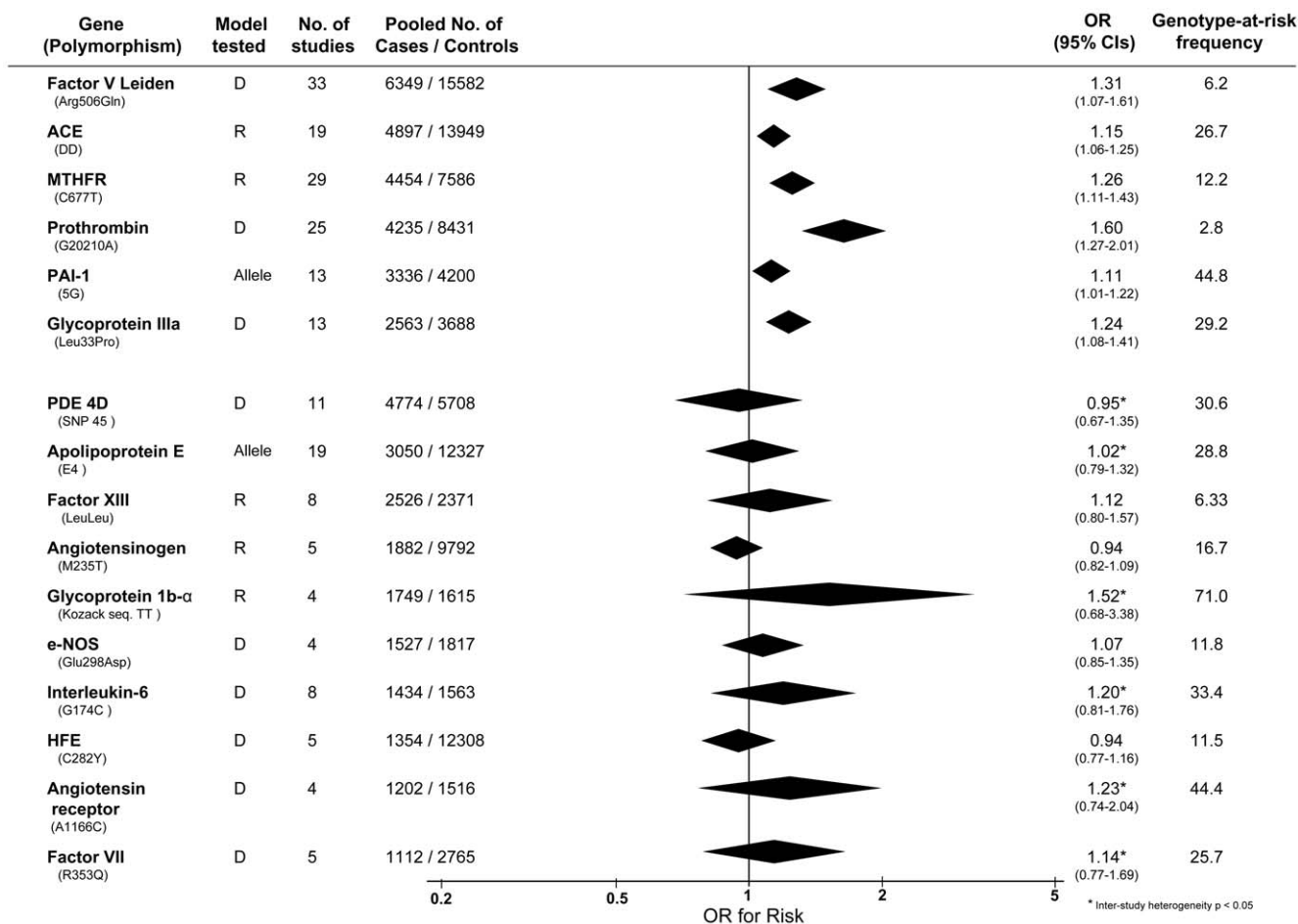


Figure 7. Summary of meta-analyses testing associations of candidate genetic polymorphisms with ischemic stroke. Table reports genetic model tested (D – dominant; R – recessive); numbers of studies; numbers of pooled cases and controls; ORs with 95% confidence intervals, and at-risk genotype frequency.

doi:10.1371/journal.pone.0009136.g007

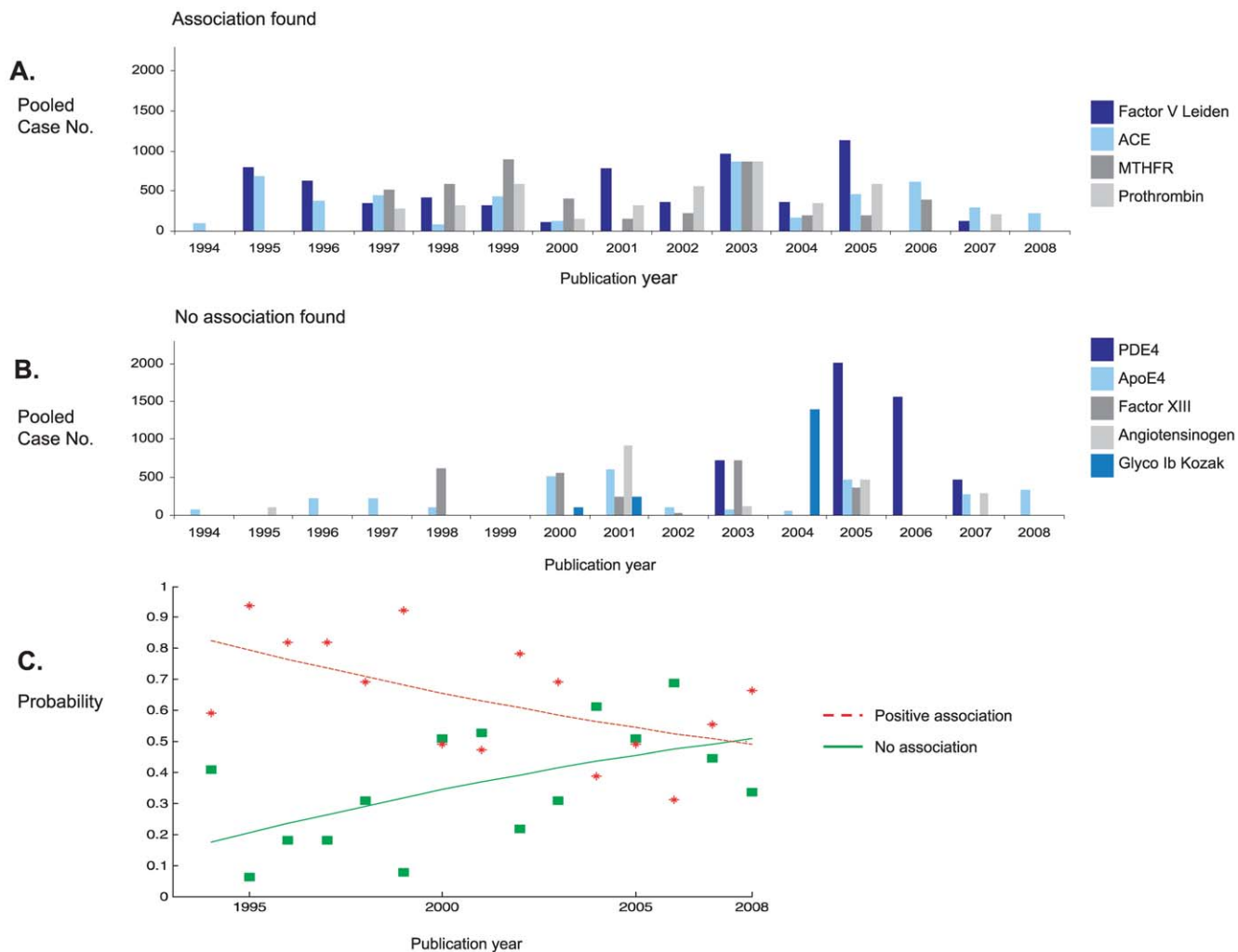


Figure 8. Trends in success of candidate gene approach. **A:** Numbers of pooled cases published over time testing for polymorphisms positively associated with stroke according to the present meta-analyses. **B:** As for A, but for polymorphisms found to show no stroke association according to the present meta-analysis. **C:** Changes in time of probability that cases were tested for polymorphism subsequently found to show association (red) or no association (green) with stroke.
doi:10.1371/journal.pone.0009136.g008

which we were able to extract genetic polymorphism frequency data has more than doubled (from ~18000 to ~37000 cases), as has the number of genetic polymorphisms in which more than 1000 cases have been tested for each (from 8 to 23). In order to increase reliability [5] we confined our results to gene variants studied in over 1000 patients, and to studies where ischemic stroke was confirmed radiologically, and that were based predominantly in Caucasian adults. We discounted studies where there was publication bias, and applied random-effects models to allow for between-study heterogeneity. After applying these criteria we found positive associations with polymorphisms in the following six genes: Factor V Leiden, ACE, MTHFR, prothrombin, PAI-1, and glycoprotein-III.

As well as identifying two further genetic associations not identified reliably in our earlier meta-analysis [25], our study supports the validity of the meta-analysis technique by finding very similar effect sizes for the four positive associations identified previously in the face of an increase in patient numbers by approximately 50%. There was no reduction in effect size comparing the current results with those from five years earlier [25] as is sometimes observed [26]. A difference in our result for

the PAI-1 polymorphism relative to a recent meta-analysis [27] can be explained by our more rigorous inclusion criteria that restricted data to Caucasian-predominant populations [28], and confirmed cerebral infarcts [29].

Whilst the effect size of each positive gene association was small (odds ratios of 1.11 to 1.60), the overall contribution that genetic factors make towards stroke is likely to be relatively large given the frequency of these risk variants in the general population (from 3 to 45% each). The sum of the population attributable risks across all the gene associations identified here was ~30%. These results are in keeping with models of common complex diseases in which relatively small numbers of common polymorphisms, each with only small hazard ratios, can account for large proportions of population attributable risk [30]. By contrast, certain well-described single-gene mutations may confer a high relative risk of stroke, e.g. CADASIL, but contribute very little to overall stroke occurrence by virtue of their rarity.

It is likely that future discovery of disease-associated genes will rest increasingly with genome-wide association studies (GWAS), rather than candidate-gene strategies [9]. For example, the recent finding from genome-wide searches that polymorphisms on chromosome 9p

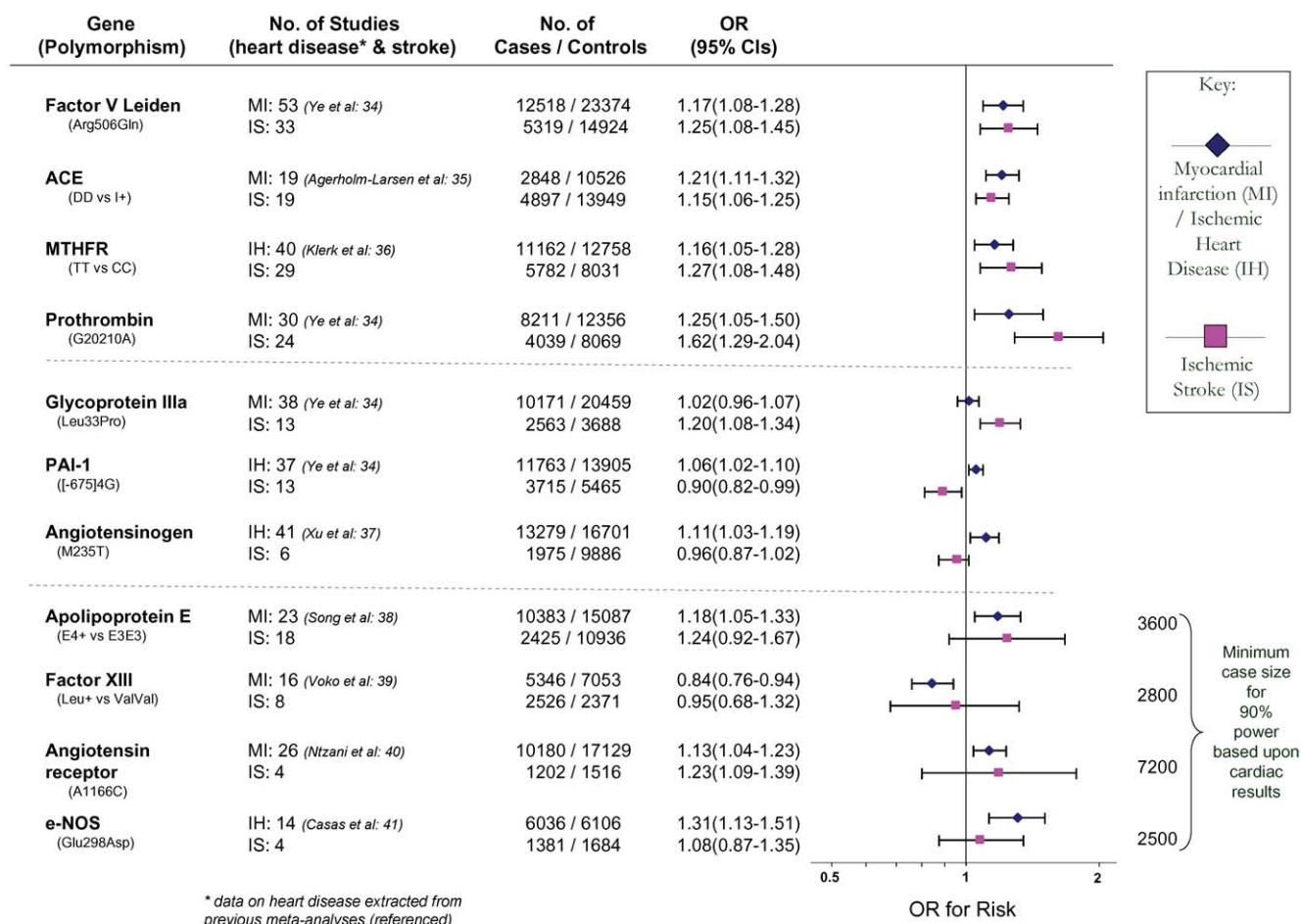


Figure 9. Risk of genetic polymorphisms compared between ischemic stroke (IS) and ischemic heart disease (IH) or myocardial infarction (MI). Contrasts represent per-allele effects except where indicated. Polymorphisms are grouped into those showing consistent associations for both types of disease; those showing dissociated effects; and those showing a significant risk of heart disease but not stroke. For the latter the minimum sample numbers estimated to achieve 90% power based upon each polymorphism's effect size for ischemic heart disease are shown.

doi:10.1371/journal.pone.0009136.g009

are associated with myocardial infarction, has prompted preliminary testing of the same polymorphisms in ischemic stroke [31,32]. In fact, we found that an increasing proportion of recruited cases to candidate-gene studies over time are testing genetic polymorphisms that, after pooling over large numbers, fail to show associations. This is unlikely to reflect a publication bias towards positive results amongst earlier publications since we found no association between effect magnitude, or sample size, and publication date; moreover, the summary ORs for candidate polymorphisms have not changed significantly over five years [25]. In the last two years the overall number of candidate-gene studies, and recruited patients, has fallen that may reflect increasing inefficiency in selecting candidate genes and increased employment of GWAS strategies.

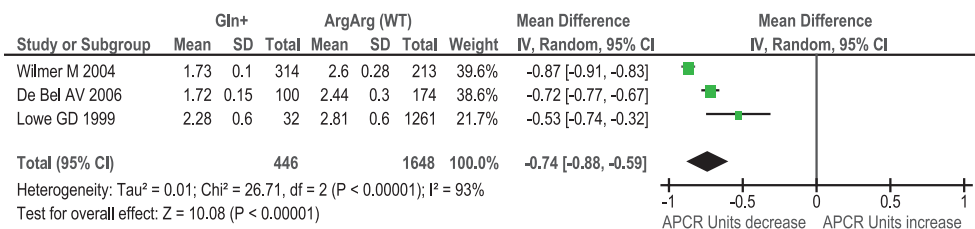
Comparison of Genetic Effects with Ischemic Heart Disease

Given overlapping pathological substrates for ischemic heart disease and ischemic stroke, as well as a common set of acquired risk factors and treatment strategies [9], we expected a broadly similar set, and relative strength, of genetic associations. For the majority of polymorphisms reported here this is what was found.

Polymorphisms were found to impose similar degrees of risk for ischemic heart disease and stroke, or were found to show no associations for both. In some cases, a similar risk was found for both disease types, but significance was only achieved for ischemic heart disease, probably due to inadequate power of pooled stroke studies.

Three genetic polymorphisms tested in both ischemic heart disease and ischemic stroke showed differing effects between diseases that did not seem to arise from inadequate power in either group. We are cautious in our interpretation of these apparent disease 'dissociations' as the two sets of meta-analyses are not strictly controlled for confounders such as age, sex, or co-morbidities, and case ascertainment methods differ, as are likely to be the thresholds at which coronary versus cerebrovascular ischemia manifest themselves clinically. Nevertheless, both our meta-analyses in stroke, and those cited in heart disease, are based predominantly in non-selective Caucasian adult populations; required both clinical and confirmatory diagnostic procedures; and checked for publication bias. Furthermore, in the cases of PAI-1 and angiotensinogen polymorphisms, it is difficult to account for how the same genotype could produce *opposite* effects on risk comparing the two disease types on the basis of confounding or bias alone.

Factor V Leiden – Activated Protein C Resistance ArgArg vs ArgGln



Activated Protein C Resistance – Stroke per 0.74 APCR decrease

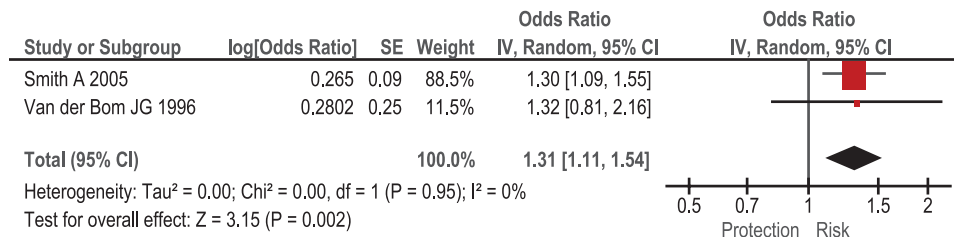


Figure 10. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: Factor V Leiden and activated Protein C resistance ratio. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.
doi:10.1371/journal.pone.0009136.g010

Further studies, ideally within the same populations, will be needed to confirm whether these three gene-disease dissociations are real, or whether they reflect methodological differences. The fact that these dissociations relate to proteins in three different physiological systems, viz. platelets (glycoprotein IIIa), clotting (PAI-1) and blood pressure (angiotensinogen) - for which there also

polymorphisms showing concordant disease effects, e.g. glycoprotein 1b-alpha, factor V Leiden and ACE, respectively – suggests that differences in risk-profile between ischemic heart disease and stroke cannot be attributed simply to a single pathophysiology e.g. hypertension. Differences between ischemic stroke and heart disease in terms of these three pathophysiology has support from

ACE – ACE activity DD vs II

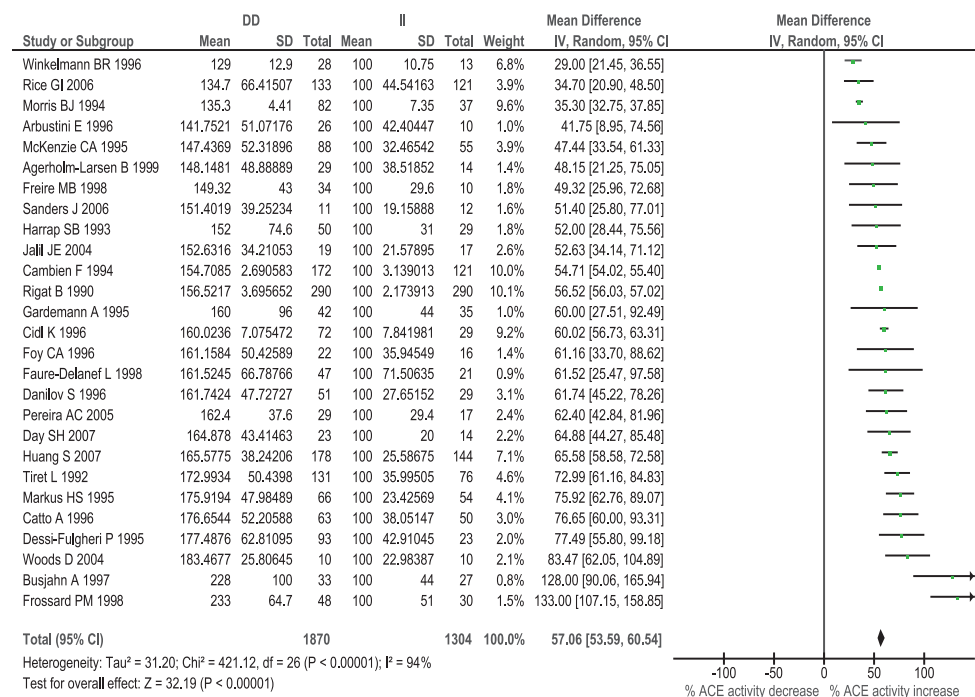


Figure 11. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: ACE D/I and ACE activity. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.
doi:10.1371/journal.pone.0009136.g011

MTHFR – homocysteine levels TT vs CC

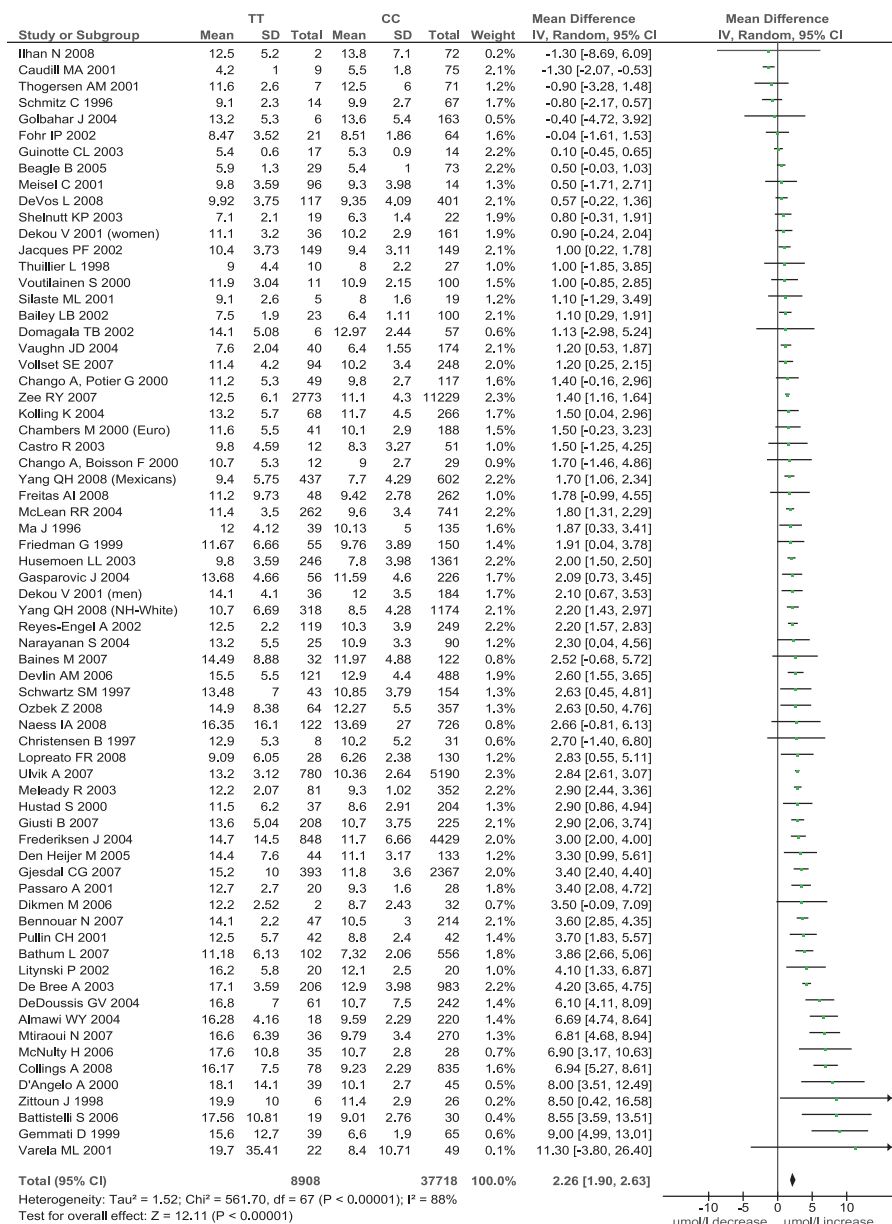


Figure 12. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: MTHFR and homocysteine levels. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.

doi:10.1371/journal.pone.0009136.g012

other sources. For example, ischemic stroke relative to ischemic heart disease has a stronger relationship with hypertension [33,34], whilst the protective profile of anti-platelet and thrombolytic drugs differs between these two diseases [9,35]. Genetic influences on ischemia may also differ according to vessel size [10].

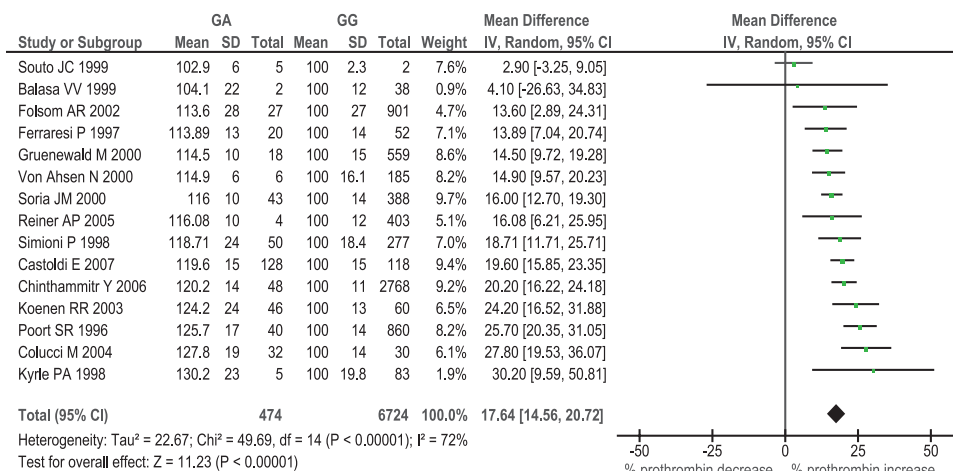
Comparison of Genetic Effects with Biochemical Markers of Risk

For each of the positive gene associations with ischemic stroke that we identified, we performed further analyses using separate data to establish whether the putative biochemical intermediaries of these gene variants are associated with equivalent quantitative levels of risk. The method used here was based upon mendelian randomization in which one starts with an observational

association between an environmental (e.g. biochemical) factor and disease, and then secondarily investigates whether a concordant level of risk occurs for a genotype that simulates the environmental factor - thereby making a stronger case for the factor being causative of the disease [36]. In the current paper, we start with positive gene-stroke associations and then subsequently interrogate independent biochemical data with the expectation of finding concordant levels of risk.

For the four strongest positive gene-stroke associations— factor V Leiden, MTHFR, ACE and prothrombin - concordance between observed risk and that predicted from their associated biochemical changes was found. Moreover, the mean levels of predicted and observed risks for these genotypes lay very close to each other. Importantly therefore, we show here, for the first time, that the

Prothrombin gene - Prothrombin levels GA vs GG



Prothrombin levels – Stroke per 17.64 % increase

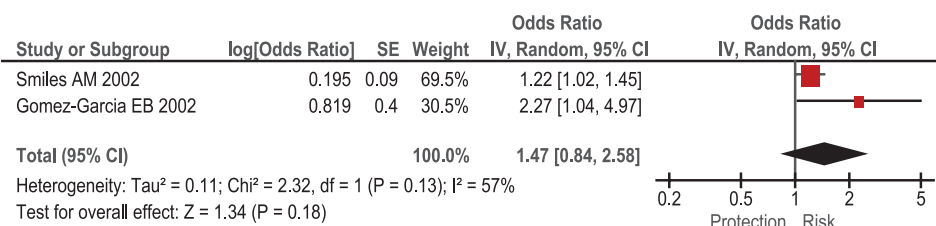


Figure 13. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: Prothrombin G20210A and prothrombin levels. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.
doi:10.1371/journal.pone.0009136.g013

four genetic variants most reliably associated with ischemic stroke are also associated with biochemical changes that themselves are related to equivalent levels of risk. This concordance both validates our original gene – stroke positive associations, and furthermore, suggests that the risk imparted by each genotype variant is the direct consequences of each gene's understood biochemical actions. The fact that each of these genes exerts biochemical or haematological changes that are measurable systemically (i.e. from venous plasma samples), rather than being specifically cerebrovasculally based, is in keeping with our other finding that these four gene variants exert similar levels of risk on ischemic heart disease as on ischemic stroke. In the case of the concordant relationship between MTHFR genotype and homocysteine in their separate associations with stroke risk, we have replicated our earlier findings in the face of more than a fourfold increase in meta-analysis size [11].

In contrast to concordant gene-biochemical risk estimates observed for the four largest genetic associations, we observed a discordant gene – biochemical relationship for the PAI-1 4G/5G polymorphism. Specifically, the variant 5G5G, relative to 4G4G, was associated with an elevated risk of ischemic stroke, but *decreased* PAI levels. However, in separate case-control studies, stroke - as well as atherothrombosis and ischemic cardiac disease [37,38] - is associated with *increased* PAI-1 levels. Indeed we saw earlier how in ischemic heart disease, it is the 4G allele - associated with higher PAI-1 levels – that is associated with risk. The expected risk of stroke of the 4G/5G PAI polymorphism was also less than that observed using tPA levels, that are strongly influenced by PAI-1 levels.

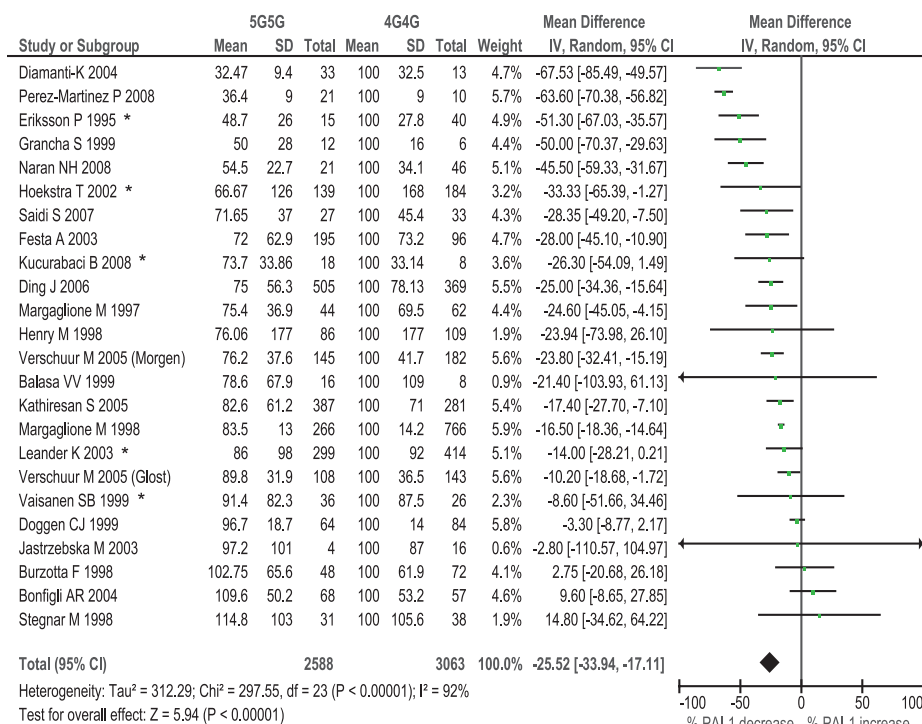
There are several possible explanations for the apparent PAI-1 gene-biochemical paradox in the case of stroke. Firstly, it is possible

that the PAI-1 5G allele association with stroke is false, e.g. because of reporting bias. However, it is unclear why such a false association should emerge in the opposite direction to that expected from ischemic cardiac disease and PAI-1 level data, and Egger's regression test argues against publication bias here [12]. Secondly, the finding that 5G polymorphism depresses PAI-1 levels over the course of the subject's life may have different pathophysiological implications, e.g. by predisposing to stroke, than the same depression of PAI-1 levels found at a single point in later life, when this is found to be protective [36]. A third possibility is that the PAI-1 genetic 4G/5G polymorphism is associated with a brain-specific factor that influences stroke more strongly than does its actions on PAI-1 levels. This may occur because of linkage disequilibrium with an as-yet unidentified gene [39], or because of pleiotropy [36]. For example, while raised PAI-1 levels in plasma may raise thrombotic risk, raised PAI-1 levels in carotid vessel wall may serve to stabilise atheromatous plaques [40]. If the PAI 4G/5G polymorphism influences both tissue and plasma PAI-1 levels, whilst the latter is also unduly influenced by environmental variables [41], this could explain the paradox. Furthermore, postulating that PAI-1 4G/5G exerts different phenotypic effects on cerebral versus other vascular beds might explain why its influences on ischemic stroke and ischemic heart disease are in opposite directions.

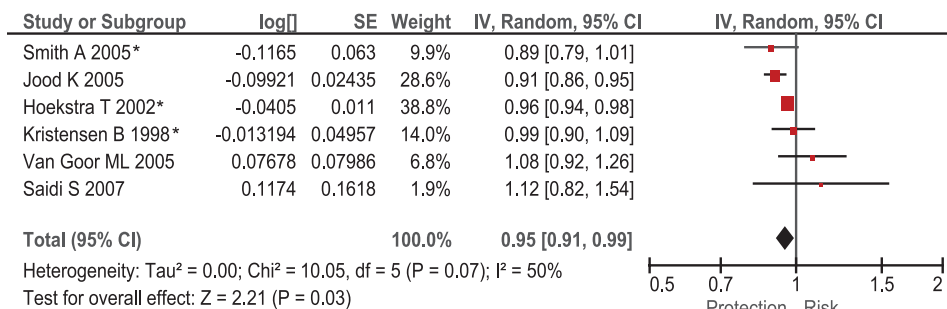
Conclusion

The current study provides a comprehensive meta-analysis of common genetic polymorphisms associated with ischemic stroke that were identified through the candidate gene approach. The results serve as an important comparator to emerging genome-wide association studies [9,31,32], which would be expected to converge

PAI-1 Gene – PAI-1 Levels* 5G5G vs 4G4G



PAI-1 Levels* - Stroke per 25.52% decrease



* Asterisked studies are based upon PAI-1 activity rather than levels

Figure 14. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: PAI-1 5G/4G and PAI-1 levels. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.

doi:10.1371/journal.pone.0009136.g014

upon similar genetic associations to those shown here, as well as to identify genes not previously implicated with cardiovascular disease pathogenesis. Since a major purpose of elucidating genetic influences on stroke (as for any complex disease) is to gain insights into its pathophysiology, we show here how pooled gene association data can be meaningfully compared with separate data relating the same genetic effects with both a pathophysiologically-related disease (here, ischemic cardiac disease) and biochemical intermediaries.

Methods

1: Genetic Polymorphism – Ischemic Stroke Association Meta-Analysis

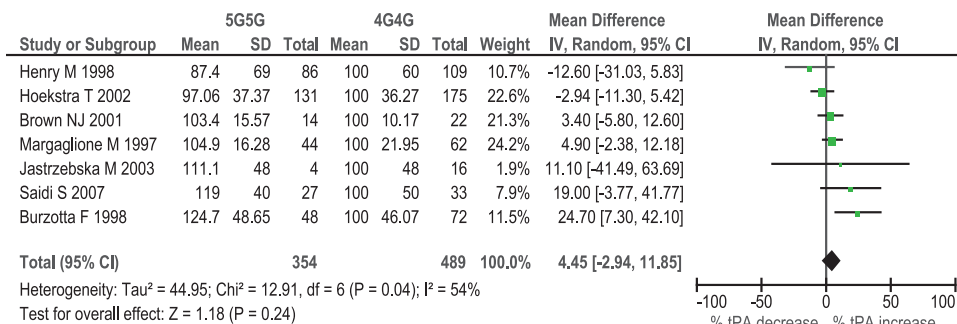
Electronic databases (Medline; EMBASE; Google Scholar) were searched upto January 1, 2009 for all case-control studies evaluating any candidate genetic polymorphism in ischemic stroke. Search words used were: *cerebrovascular disease, brain infarction, stroke and cerebral ischemia* in combination with *polymorphism, genetic,*

mutation, genotype and genes. All languages were searched and translated when necessary. Additional studies were sought from references, citations and from the PubMed option ‘Related Articles’, for each identified study.

Inclusion criteria were studies that: 1) employed case-control methods where ischemic stroke was analyzed as a dichotomous trait; both retrospective and prospective cohort designs were included; 2) confirmed the diagnosis of ischemic stroke with neuroimaging, and 3) were based in Caucasian populations, so as to minimise inter-racial heterogeneity [28]. Studies were excluded if: 1) patients were aged under 18 years; 2) only quantitative traits or intermediate phenotypes were being investigated, or 3) genotype frequency was not reported. For duplicate publications, the smaller data set(s) were discarded. For studies with more than one control group, the most appropriate control group was used.

Data were analyzed using software for preparing Cochrane reviews (Review Manager, v5; Comprehensive Meta Analysis v2.2.023). For each genetic polymorphism for which data were

PAI-1 Gene – tPA Levels 5G5G vs 4G4G



tPA Levels - Stroke 5G5G vs 4G4G

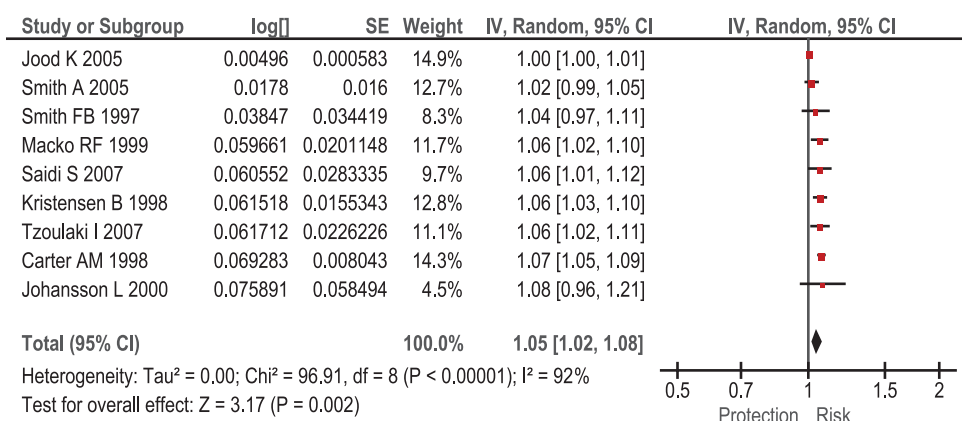


Figure 15. Forest plots showing quantitative relationship between genetic polymorphisms and associated biochemical variables for: PAI-1 5G/4G and tPA levels. Additional forest plots are shown in Figures 10, and 13– 15 that relate set changes in biochemical variables (determined from the first set of meta-analyses within each figure) with risk of stroke. For MTHFR and ACE this relationship is determined from a single study each.

doi:10.1371/journal.pone.0009136.g015

available from at least two studies, a meta-analysis was carried out. For each gene variant, a pooled odds ratio (OR) and 95% confidence intervals were calculated using random-effects [42] models, to estimate association strength. Tests for heterogeneity [43], and publication bias [12] were performed with significance set at $p < 0.05$. Random-effects results are reported throughout. In order to improve reliability, we only present here results for polymorphisms tested in >1000 pooled cases [5]. The proportion of cases in the population that could be attributed to a particular genetic variant (population attributable risk or PAR) was estimated as follows: $PAR = 100 \times [Prevalence(OR - 1)/(Prevalence(OR - 1) + 1)]$. The prevalence of exposure was estimated as the genotype frequency among pooled controls.

We also sought to determine temporal trends in the success of the candidate-gene approach by determining the pooled number of cases recruited into studies that tested for polymorphisms that were either associated, or not associated, with stroke according to our meta-analysis, and then plotting this against study publication year. Moreover, for every publication year, we calculated the probability that cases were tested for a polymorphism that demonstrated either a significant, or no, association by our meta-analysis. This analysis is restricted to polymorphisms for which there were >1000 pooled cases.

2: Ischemic Stroke – Ischemic Heart Disease Genetic Comparison

For each genetic polymorphism tested in ischemic stroke (with >1000 cases), we searched the literature for the most recent

meta-analysis testing for the same polymorphism in myocardial infarction and/or ischemic heart disease, based predominantly in Caucasian populations. For all polymorphisms showing a positive association with ischemic stroke there was an equivalent published meta-analysis in ischemic heart disease/myocardial infarction. For each polymorphism for which a cardiac meta-analysis existed we calculated the pooled random-effects OR of ischemic stroke for the equivalent genotype (or per-allele) comparison using our own meta-analysis.

3: Gene – Intermediate Phenotype Comparison

For each positive genetic association identified from the meta-analysis in part 1, we performed a separate analysis that produced an estimate of expected risk based upon genotype – biochemical, and biochemical – stroke, association studies, using the principle of mendelian randomisation [11]. Firstly, we searched the medical literature for two further types of study: 1) those relating each genetic polymorphism with a quantitative measure of the most strongly-associated biochemical or hematological marker - i.e. an intermediate phenotype, IP, for the gene in question - in populations free from cardiovascular disease, and 2) those relating an incremental change of the same IP with risk of ischemic stroke. For example, for the ACE polymorphism we searched for: 1) (ACE OR *angiotensin converting enzyme*) AND (*gene* OR *genetic* OR *genotype* OR *polymorphism* OR *mutation*), in combination with (ACE OR *angiotensin converting enzyme*) AND (*activity* OR *level*); and 2) (ACE OR *angiotensin converting enzyme*) AND (*activity* OR *level*) in conjunction

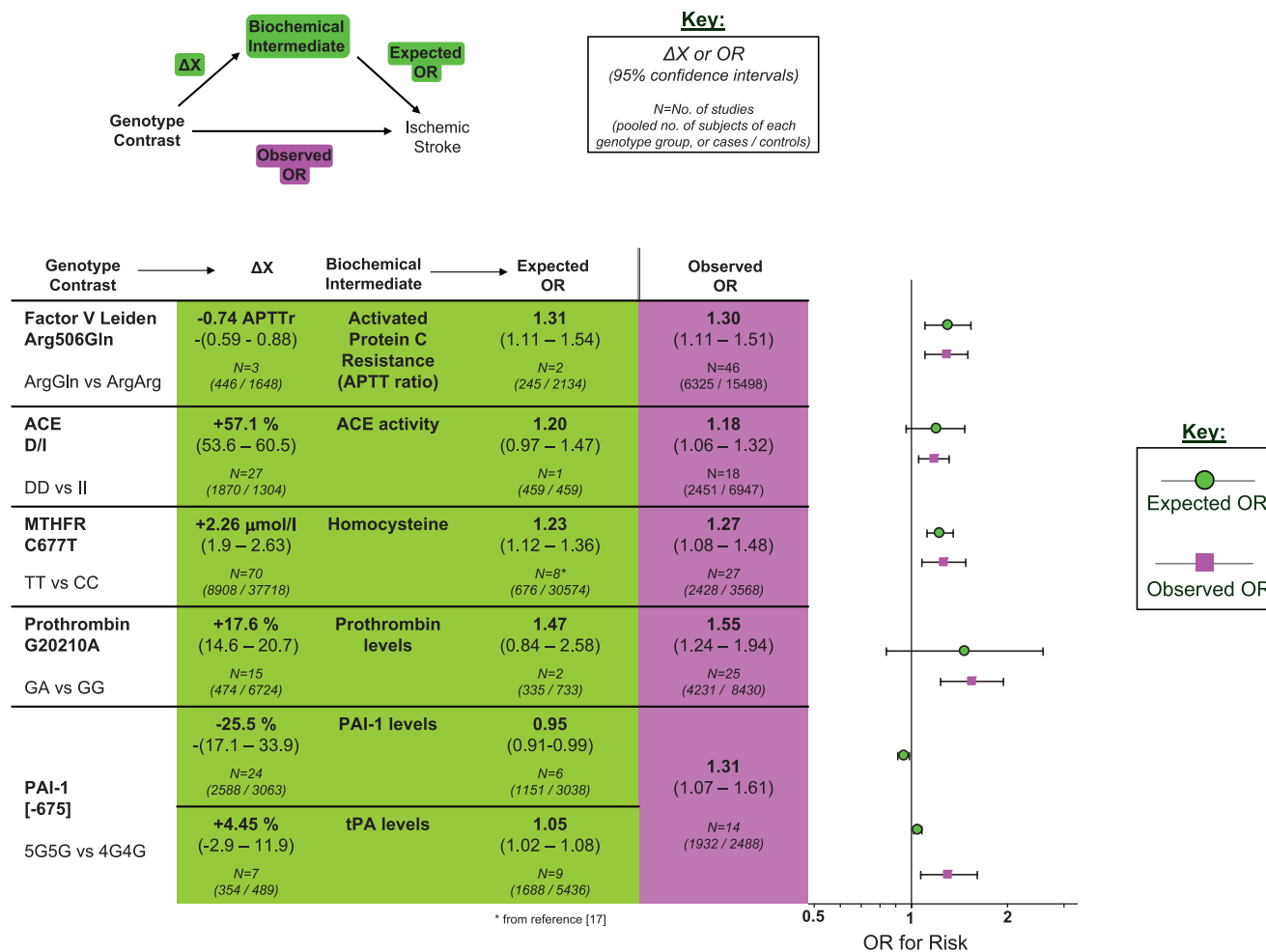


Figure 16. Comparison of estimated risk with observed risk for ischemic stroke-associated genetic polymorphisms. Estimated ORs (green) are calculated from 1) meta-analyses relating risk-genotype with biochemical variation (ΔX), and 2) single studies or meta-analyses relating biochemical variation with stroke risk (scaled log-linearly). Observed ORs are derived from meta-analyses of the equivalent genotype contrast using the current datasets.
 doi:10.1371/journal.pone.0009136.g016

with (*cerebrovascular disease OR brain infarction OR stroke OR cerebral ischemia*). As for part 1, we restricted studies to those based in adults from predominantly Caucasian populations.

From these two sets of studies we performed two types of analyses: 1) From the first set of studies we extracted from each the difference in levels (or biological activity) of the IP between the two homozygous variants, or between the wild-type and heterozygous variants in cases of dominant polymorphisms. These values were entered into a meta-analysis to obtain a weighted mean average using a random-effects model. For those IPs where different measurement units had been employed between studies (viz. ACE activity; prothrombin levels, plasminogen-activator inhibitor-1 levels and tissue plasminogen activator levels) we calculated the percentage change for the rarer, relative to the commoner, genotype. 2) From the second set of studies we obtained an OR of ischemic stroke for a given change in the level of that IP used in the first analysis. Where studies reported different ORs for different ranges of IP level, we chose the OR reported for the range closest to that applying to the control group (i.e. healthy population). The ORs between studies were pooled using a generic inverse variance procedure in which the logarithms of the ORs are weighted according to variance in a random-effects model. Since

the ORs reported between these studies usually refer to different amounts of IP change, we first scaled the OR values for each study in proportion to the pooled change in IP level for the genotype comparison of interest (from the first analysis) assuming a log-linear relationship [11]. For the relationship between homocysteine levels and ischemic stroke we used a summary OR from a previous meta-analysis [24].

Supporting Information

Figures S1 Supplementary Figures 1–10: Forest plots of genetic polymorphisms tested in more than 1000 pooled cases but showing no association with ischemic stroke.

Found at: doi:10.1371/journal.pone.0009136.s001 (0.08 MB PDF)

References S1 Supplementary References

Found at: doi:10.1371/journal.pone.0009136.s002 (0.29 MB PDF)

Acknowledgments

We are grateful for helpful suggestions from Aron Hingorani & Juan Pablo Casas during the course of this work.

Author Contributions

Conceived and designed the experiments: PB LS PS. Performed the experiments: PB GP PS. Analyzed the data: PB GP JW PS. Wrote the paper: PB GP LS JW PS.

References

- Johnston SC, Mendis S, Mathers CD (2009) Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *Lancet Neurol* 8: 345–354.
- Chiuvè SE, Rexrode KM, Spiegelman D, Logroscino G, Manson JE, et al. (2008) Primary prevention of stroke by healthy lifestyle. *Circulation* 118: 947–954.
- Arason GJ, Kramer J, Blaskó B, Kolka R, Thorbjörnsdóttir P, et al. (2007) Smoking and a complement gene polymorphism interact in promoting cardiovascular disease morbidity and mortality. *Clin Exp Immunol* 149: 132–138.
- Arnett DK, Baird AE, Barkley RA, Basson CT, Boerwinkle E, et al. (2007) American Heart Association Council on Epidemiology and Prevention; American Heart Association Stroke Council; Functional Genomics and Translational Biology Interdisciplinary Working Group; Relevance of genetics and genomics for prevention and treatment of cardiovascular disease. *Circulation* 115: 2878–901.
- Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, et al. (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 37: 120–122.
- Fisher M, Folland E (2008) Acute ischemic coronary artery disease and ischemic stroke: similarities and differences. *Am J Ther* 15: 137–149.
- Kingwell B, Boutouyrie P (2007) Genetic influences on the arterial wall. *Clin Exp Pharmacol Physiol* 34: 652–657.
- Jood K, Ladenvall P, Tjårlund-Wolf A, Ladenvall C, Andersson M, et al. (2005) Fibrinolytic gene polymorphism and ischemic stroke. *Stroke* 36: 2077–2081.
- Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, et al. (2009) Genome-wide association studies of stroke. *N Engl J Med* 360: 1718–1728.
- Rao R, Tah V, Casas JP, Hingorani A, Whittaker J, et al. (2009) Ischaemic stroke subtypes and their genetic basis: a comprehensive meta-analysis of small and large vessel stroke. *Eur Neurol* 61: 76–86.
- Casas JP, Bautista LE, Smeeth L, Sharma P, Hingorani AD (2005) Homocysteine and stroke: evidence on a causal link from mendelian randomisation. *Lancet* 365: 224–232.
- Egger M, Davey SG, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315: 629–634.
- Ye Z, Liu EH, Higgins JP, Keavney BD, Lowe GD, et al. (2006) Seven haemostatic gene polymorphisms in coronary disease: meta-analysis of 66,155 cases and 91,307 controls. *Lancet* 367(9511): 651–8.
- Agerholm-Larsen B, Nordestgaard BG, Tybjaerg-Hansen A (2000) ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. *Arterioscler Thromb Vasc Biol* 20(2): 484–92.
- Klerk M, Verhoef P, Clarke R, Blom HJ, Kok FJ, et al. (2002) MTHFR Studies Collaboration Group. MTHFR 677C->T polymorphism and risk of coronary heart disease: a meta-analysis. *JAMA* 288(16): 2023–31.
- Xu MQ, Ye Z, Hu FB, He L (2007) Quantitative assessment of the effect of angiotensinogen gene polymorphisms on the risk of coronary heart disease. *Circulation* 116(12): 1356–66.
- Song Y, Stampfer MJ, Liu S (2004) Meta-analysis: apolipoprotein E genotypes and risk for coronary heart disease. *Ann Intern Med* 141(2): 137–47.
- Vokó Z, Bereczky Z, Katona E, Adány R, Muszbek L (2007) Factor XIII Val34Leu variant protects against coronary artery disease. A meta-analysis. *Thromb Haemostasis* 97(3): 458–63.
- Ntzani EE, Rizos EC, Ioannidis JP (2007) Genetic effects versus bias for candidate polymorphisms in myocardial infarction: case study and overview of large-scale evidence. *Am J Epidemiol* 165(9): 973–84.
- Casas JP, Bautista LE, Humphries SE, Hingorani AD (2004) Endothelial nitric oxide synthase genotype and ischemic heart disease: meta-analysis of 26 studies involving 23028 subjects. *Circulation* 109(11): 1359–65.
- Sie MP, Sayed-Tabatabaei FA, Oei HH, Uitterlinden AG, Pols HA, et al. (2006) Interleukin 6 -174 g/c promoter polymorphism and risk of coronary heart disease: results from the Rotterdam study and a meta-analysis. *Arterioscler Thromb Vasc Biol* 26(1): 212–7.
- Ellervik C, Birgens H, Tybjaerg-Hansen A, Nordestgaard BG (2007) Hemochromatosis genotypes and risk of 31 disease endpoints: meta-analyses including 66,000 cases and 226,000 controls. *Hepatology* 46(4): 1071–80.
- Brenner D, Labreuche J, Poirier O, Cambien F, Amarenco P GENIC Investigators (2005) Renin-angiotensin-aldosterone system in brain infarction and vascular death. *Ann Neurol* 58(1): 131–8.
- Wald DS, Law M, Morris JK (2002) Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ* 325: 1202.
- Casas JP, Hingorani AD, Bautista LE, Sharma P (2004) Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol* 61: 1652–1661.
- Zintzaras E, Lau J (2008) Synthesis of genetic association studies for pertinent gene-disease associations requires appropriate methodological and statistical approaches. *J Clin Epidemiol* 61: 634–645.
- Tsantes AE, Nikolopoulos GK, Bagos PG, Tsiara CG, Kapsimali V, et al. (2007) Plasminogen activator inhibitor-1 4G/5G polymorphism and risk of ischemic stroke: a meta-analysis. *Blood Coagul Fibrinolysis* 18: 497–504.
- Ariyaratnam R, Casas JP, Whittaker J, Smeeth L, Hingorani AD, et al. (2007) Genetics of ischaemic stroke among persons of non-European descent: a meta-analysis of eight genes involving approximately 32,500 individuals. *PLoS Med* 4: e131.
- Peck G, Smeeth L, Whittaker J, Casas JP, Hingorani A, et al. (2008) The genetics of primary haemorrhagic stroke, subarachnoid haemorrhage and ruptured intracranial aneurysms in adults. *PLoS ONE* 3: e3691.
- Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD (2005) How many genes underlie the occurrence of common complex diseases in the population? *Int J Epidemiol* 34: 1129–1137.
- Matarin M, Brown WM, Scholz S, Simón-Sánchez J, Fung HC, et al. (2007) A genome-wide genotyping study in patients with ischemic stroke: initial analysis and data release. *Lancet Neurol* 6: 414–420.
- Karvanen J, Silander K, Kee F, Tirt L, Salomaa V, et al. (2009) MORGAM Project. The impact of newly identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. *Genet Epidemiol* 33: 237–246.
- Kannel WB, Wolf PA, Verter J, McNamara PM (1970) Epidemiologic assessment of the role of blood pressure in stroke. The Framingham study. *JAMA* 214: 301–310.
- Law MR, Morris JK, Wald NJ (2009) Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ* 338: b1665.
- Wolff T, Miller T, Ko S (2009) Aspirin for the primary prevention of cardiovascular events: an update of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 150: 405–410.
- Davey Smith G, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32: 1–22.
- Hamsten A, de Faire U, Walldius G, Dahlén G, Szamosi A, et al. (1987) Plasminogen activator inhibitor in plasma: risk factor for recurrent myocardial infarction. *Lancet* 2: 3–9.
- Thögersen AM, Jansson JH, Boman K, Nilsson TK, Weinhall L, et al. (1998) High plasminogen activator inhibitor and tissue plasminogen activator levels in plasma precede a first acute myocardial infarction in both men and women: evidence for the fibrinolytic system as an independent primary risk factor. *Circulation* 98: 2241–2247.
- Attia J, Thakkinian A, Wang Y, Lincz L, Parsons M, et al. (2007) The PAI-1 4G/5G gene polymorphism and ischemic stroke: an association study and meta-analysis. *J Stroke Cerebrovasc Dis* 16: 173–179.
- Hoekstra T, Geleijnse JM, Kluit C, Giltay EJ, Kok FJ, et al. (2003) 4G/4G genotype of PAI-1 gene is associated with reduced risk of stroke in elderly. *Stroke* 34: 2822–2828.
- Henry M, Tregouët DA, Alessi MC, Aillaud MF, Visvikis S, et al. (1998) Metabolic determinants are much more important than genetic polymorphisms in determining the PAI-1 activity and antigen plasma concentrations: a family study with part of the Stanislas Cohort. *Arterioscler Thromb Vasc Biol* 18: 84–91.
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7: 177–188.
- Deeks JJ, Altman DG, Bradburn MJ (2001) 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*. Annapolis Junction, Md: BMJ Publishing Group; 2001.