THE UNIVERSITY
of EDINBURGH

Genetic associations with sporadic cerebral small vessel disease

Kristiina Rannikmäe

Doctor of Philosophy

The University of Edinburgh

2016

Contents

Table of tables

Table of figures

Declaration

I confirm that I composed this thesis by myself, and it is my own original work. The thesis has not been submitted in part or in whole for any other degree or professional qualification.

Kristiina Rannikmäe
30th March 2016

Acknowledgements

First and foremost, I would like to thank my main supervisor Professor Cathie Sudlow for inspiring me to come back to Edinburgh and for giving me the opportunity to undertake this research. I am genuinely grateful for her teaching, and extensive support and mentorship throughout the past few years.

I am also very grateful to my second supervisor Professor Rustam Al-Shahi Salman for his support, his contribution to the systematic reviews of CAA genetics and ICH classification systems, and for his comments on the earlier drafts of this thesis.

I have greatly appreciated the opportunity to undertake this research in the Centre for Clinical Brain Sciences, which provided a stimulating and friendly environment. I thank the Edinburgh Stroke Research Group for their comments that helped improve my work. I am incredibly lucky to have had research fellows Neshika Samarasekera, Rebecca Woodfield, Josephine Tsai, Pete Foley and Steven Makin as colleagues and friends. I thank Aidan Hutchinson, Christopher Matthews and Kirsten Shuler for their kind assistance with my various (and multiple) software and data manipulation queries. I would also like to express my gratitude to my colleagues from the University of Tartu in Estonia: Associate Professor Sulev Haldre for supporting and facilitating my decision to undertake a PhD abroad and Dr Mark Braschinsky for his research supervision before I moved to Edinburgh.

The work presented in this thesis would not have been possible without the collaboration of many individual research groups and consortia across the world:

The systematic review of ICH classification systems' reliability involved collaboration with Doctors Andreas Charidimou, Pipat Chiewvit, Jiann-Shing Jeng, Atte Meretoja, Frederic Palm, Jukka Putaala, Daniel Strbian, Marieke Wermer and Shin-Joe Yeh, and Professors Craig Anderson, Steven Greenberg, Gabriel Rinkel, Jonathan Rosand, Natalia Rost, Turgut Tatlisumak and David Werring, all of whom shared unpublished data and/or gave comments to improve the project. Dr Rebecca Woodfield's previous work on the reliability of ischaemic stroke classification systems provided me with a helpful framework significantly facilitating my project. Dr Josephine Tsai helped me with proportion meta-analyses..

Systematic review and meta-analyses of CAA genetics: Professors Rajesh Kalaria, Steven Greenberg, Helena Chui and Frederick Schmitt shared their study data. Dr Neshika Samarasekera and Professor Al-Shahi Salman helped select relevant studies and extract data, in addition to providing insightful comments on various aspects of the research. A part of this systematic review was initiated by Professor Cathie Sudlow's former PhD student Nahara Martînez-Gonzâlez and a medical student Robert Annan, and while I redesigned the study starting from the search strategy, it was helpful to have their previous work as an example.

The meta-analyses of collagen genes and small vessel disease were possible because of the ISGC community and extremely helpful collaboration of the Metastroke consortium, the CHARGE consortium, the ICH GWAS Study and the WMH in Ischemic Stroke GWAS Study members, who shared their data and advice. I am thankful to Professor Jonathan Rosand and his team – including Doctors Guido Falcone, Farid Radmanesh, William Devan and Christopher Anderson – for hosting me in their lab in the Center for Human Genetic Research, MGH; for facilitating data access; and for their useful comments on this project. I am truly grateful to Doctors Matthew Traylor, Steve Bevan and Rainer Malik for helping me – often at short notice and when I was in despair – with various aspects of genetic analysis. I thank Dr Tom van Agtmael from the University of Glasgow for his ideas earlier in the project. I am very grateful to Doctors Gail Davies and Pippa Thomson for their teaching of genetic analyses which was essential for this project's success.

I acknowledge funding received from the ENS, EFNS and Edinburgh Lothian Health Fund.

My final thanks goes to my family for their love and encouragement. Thank you Miia and Jack for being my role models and biggest supporters. Helena and Hanna: thank you for distracting me when I needed it, yet always listening and understanding at the important moments. And Stuart, for being there on every step of this journey with me. Finally, thank you Jack and Stuart for proofreading my thesis.

Publications and awards relating to this thesis

Papers in peer reviewed journals relating to this thesis

1. **Rannikmäe K**, Woodfield R, Anderson CS, Charidimou A, Chiewvit P, Greenberg SM, Jeng J-S, Meretoja A, Palm F, Putaala J, Rinkel GJE, Rosand J, Rost NS, Strbian D, Tatlisumak T, Tsai C-F, Wermer MJH, Werring D, Yeh S-J, Al-Shahi Salman R, Sudlow CLM. Reliability of intracerebral hemorrhage classification systems: a systematic review. International Journal of Stroke 2016. *In press.*
2. **Rannikmäe K**, Davies G, Thomson PA, Bevan S, Devan WJ, Falcone GJ, Traylor M, Anderson CD, Battey TW, Radmanesh F, Deka R, Woo JG, Martin LJ, Jimenez-Conde J, Selim M, Brown DL, Silliman SL, Kidwell CS, Montaner J, Langefeld CD, Slowik A, Hansen BM, Lindgren AG, Meschia JF, Fornage M, Bis JC, Debette S, Ikram MA, Longstreth WT, Schmidt R, Zhang CR, Yang Q, Sharma P, Kittner SJ, Mitchell BD, Holliday EG, Levi CR, Attia J, Rothwell PM, Poole DL, Boncoraglio GB, Psaty BM, Malik R, Rost N, Worrall BB, Dichgans M, Van Agtmael T, Woo D, Markus HS, Seshadri S, Rosand J, Sudlow CLM, METASTROKE Consortium, CHARGE WMH Group, ISGC ICH GWAS Study Collaboration, WMH in Ischemic Stroke GWAS Study Collaboration, International Stroke Genetics Consortium. Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease. Neurology 2015;84(9)918-926.
3. **Rannikmäe K**, Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Samarasekera N, Al-Shahi Salman R, Sudlow CLM. APOE associations with severe CAA-associated vasculopathic changes – collaborative meta-analysis. Journal of Neurology, Neurosurgery and Psychiatry 2014;85(3):300-305.
4. **Rannikmäe K**, Samarasekera N, Salman R, Martînez-Gonzâlez N, Sudlow CLM. Genetic associations of cerebral amyloid angiopathy – a systematic review and meta-analysis. Journal of Neurology, Neurosurgery and Psychiatry 2013;84(8):901-908.

Papers which I have contributed to during my PhD studies

1. NINDS Stroke Genetics Network (SiGN) and International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurology 2015; doi:10.1016/S1474-4422(15)00338-5.
2. Malik R, Freilinger T, Winsvold BS, Anttila V, Vander Heiden J, Traylor M, de Vries B, Holliday EG, Terwindt GM, Sturm J, Bis JC, Hopewell JC, Ferrari MD, **Rannikmäe K**, Wessman M, Kallela M, Kubisch C, Fornage M, Meschia JF, Lehtimäki T, Sudlow C, Clarke R, Chasman DI, Mitchell BD, Maguire J, Kaprio J, Farrall M, Raitakari OT, Kurth T, Ikram MA, Reiner AP, Longstreth WT Jr, Rothwell PM, Strachan DP, Sharma P, Seshadri S, Quaye L, Cherkas L, Schürks M, Rosand J, Ligthart L, Boncoraglio GB, Davey Smith G, van Duijn CM, Stefansson K, Worrall BB, Nyholt DR, Markus HS, van den Maagdenberg AM, Cotsapas C, Zwart JA, Palotie A, International Headache Genetics Consortium, Dichgans M, METASTROKE Collaboration of the International Stroke Genetics Consortium. Shared genetic basis for migraine and ischemic stroke: A genome-wide analysis of common variants. Neurology 2015;84(21):2132-2145.
3. Ay H, Arsava EM, Andsberg G, Benner T, Brown RD Jr, Chapman SN, Cole JW, Delavaran H, Dichgans M, Engström G, Giralt-Steinhauer E, Grewal RP, Gwinn K, Jern C, Jimenez-Conde J, Jood K, Katsnelson M, Kissela B, Kittner SJ, Kleindorfer DO, Labovitz DL, Lanfranconi S, Lee JM, Lehm M, Lemmens R, Levi C, Li L, Lindgren A, Markus HS, McArdle PF, Melander O, Norrving B, Peddareddygari LR, Pedersén A, Pera J, **Rannikmäe K**, Rexrode KM, Rhodes D, Rich SS, Roquer J, Rosand J, Rothwell PM, Rundek T, Sacco RL, Schmidt R, Schürks M, Seiler S, Sharma P, Slowik A, Sudlow C, Thijs V, Woodfield R, Worrall BB, Meschia JF. Pathogenic ischemic stroke phenotypes in the NINDS-stroke genetics network. Stroke 2014;45(12):3589-3596.

Other relevant publications

1. **Rannikmäe K**, Sudlow C. Book chapter "Genetics of Sporadic Cerebral Amyloid Angiopathy" in "Intracerebral Haemorrhage" by InTech. 2014; Chapter 5:51-72.

Conference presentations

1. **Rannikmäe K**, Woodfield R, Al-Shahi Salman R, Sudlow CLM on behalf of the ICH Classification Collaboration. Reliability of spontaneous intracerebral haemorrhage classification systems: a systematic review. Poster at the European Stroke Organisation Conference in Glasgow in 2015. (Published in *International Journal of Stroke*, 2015:10 Supplement 2:295.)
2. **Rannikmäe K**, Davies G, Thomson P, Devan W, Bevan S, Traylor M, Falcone GJ, Anderson C, Rost N, Woo D, Markus H, Seshadri S, Rosand J, Van Agtmael T, Sudlow CLM. The contribution of variants in collagen IV genes to common cerebrovascular phenotypes. Platform presentation at the European Stroke Conference in Nice in 2014. (Published in *Cerebrovascular Diseases*, 2014:37 Supplement 1:1–2.)
3. **Rannikmäe K,** Davies G, Devan W, Bevan S, Traylor M, Falcone G, Anderson C, Rost N, Markus H, Rosand J, Van Agtmael T, Sudlow CLM on behalf of the METASTROKE Consortium, CHARGE Consortium, and ISGC. The contribution of variants in collagen IV genes to common cerebrovascular phenotypes. Platform presentation at the International Stroke Genetics Consortium meeting in Paris in 2014.
4. **Rannikmäe K,** Davies G, Devan W, Bevan S, Traylor M, Falcone G, Anderson C, Rost N, Markus H, Rosand J, Van Agtmael T, Sudlow CLM on behalf of the METASTROKE Consortium, CHARGE Consortium, and ISGC. The contribution of variants in collagen IV genes to common cerebrovascular phenotypes. Poster at the UK Stroke Forum in Harrogate in 2013. (Published in *International Journal of Stroke*, 2013:8 Supplement 3:56.)
5. **Rannikmäe K,** Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Samarasekera N, Al-Shahi Salman R, Sudlow CLM. APOE allele-specific associations with severe CAA-associated vasculopathic changes – collaborative meta-analysis.

Poster at the UK Stroke Forum in Harrogate in 2013. (Published in *International Journal of Stroke*, 2013:8 Supplement 3:56.)

6. Woodfield R, **Rannikmäe K**, Grant I, Sudlow C. Reliability and feasibility of ischemic stroke classification systems for large epidemiological studies. Poster at the European Stroke Conference in London in 2013. (Published in *Cerebrovascular Diseases*, 2013: 35:159.)
7. **Rannikmäe K**, Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Sudlow CLM. Association between APOE genotype and cerebral amyloid angiopathy (CAA)-related vasculopathic changes: Collaborative metaanalysis. Platform presentation at the European Neurological Society meeting in Barcelona in 2013. (Published in *Journal of Neurology*, 2013:260:S40-S41.)
8. **Rannikmäe K**. Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Sudlow CLM. Association between APOE genotype and CAA related vasculopathic changes – collaborative meta-analysis. Poster at the European Stroke Conference in London in 2013. (Published in *Cerebrovascular Diseases*, 2013:35:435.)
9. **Rannikmäe K**. Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Sudlow CLM. Association between APOE genotype and CAA related vasculopathic changes – collaborative meta-analysis. Platform presentation at the International Stroke Genetics Consortium meeting in Charlottesville in 2013.
10. **Rannikmäe K**, Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Sudlow CLM. Association between APOE genotype and CAA related vasculopathic changes – collaborative meta-analysis. Platform presentation at the 3rd International Cerebral Amyloid Angiopathy Conference in Leiden in 2012.
11. Arsava EM, Barrett K, Biffi A, Brenner D, Brown R, Gamble D, Grazer A, Gschwendtner Andreas, Helenius J, Huq M, Katschnig P, Katsnelson M, Kittner S, Li L, Labovitz D, Malik S, Meschia J, Ossi R, Peddareddygari L, Pucek M, **Rannikmäe K**, Rhodes D, Saraf N, Sheikh H, Stoegerer Eva, Southerland A, Weissman D, Woo D, Worrall B, Ay H. The Utility of Causative Classification System: Correlation between Causative and Phenotypic Stroke Subtypes. Abstract at the International Stroke Conference in New Orleans in 2012. (Published in *Stroke*, 2012:43:2.)

12. **Rannikmäe K,** Samarasekera N, Al-Shahi Salman R, Martinez-Gonzalez N, Annan R, Sudlow CLM. Genetic associations of cerebral amyloid angiopathy – a systematic review and meta-analysis. Platform presentation at the European Stroke Conference in Hamburg in 2011. (Published in *Cerebrovascular Diseases*, 2011:31 Supplement 2:3-4).
13. **Rannikmäe K**, Samarasekera N, Al-Shahi Salman R, Martinez-Gonzalez N, Annan R, Sudlow CLM. Genetic associations of cerebral amyloid angiopathy – a systematic review and meta-analysis. Platform presentation at the UK Stroke Forum in Glasgow in 2011. (Published in *International Journal of Stroke*, 2011:6 Supplement 2:7.)
14. **Rannikmäe K,** Samarasekera N, Al-Shahi Salman R, Martinez-Gonzalez N, Annan R, Sudlow CLM. Genetic associations of cerebral amyloid angiopathy: a systematic review and meta-analysis. Platform presentation at the European Neurological Society meeting in Lisbon in 2011. (Published in *Journal of Neurology*, 2011:258 Supplement 1:S10.)
15. **Rannikmäe K,** Samarasekera N, Al-Shahi Salman R, Martinez-Gonzalez N, Annan R, Sudlow CLM. Genetic associations of cerebral amyloid angiopathy: a systematic review and meta-analysis. Platform presentation at the International Stroke Genetics Consortium meeting in Miami in 2011.

Awards

1. **Research funding**: European Federation of Neurological Sciences scientific fellowship; Newby Trust Award; European Neurological Societies scientific fellowship; Edinburgh Lothian Health Fund "Neurology" stipend; Edinburgh Lothian Health Fund "Stroke Endowments" stipend.
2. **Travel Grants** awarded by: The Guarantors of Brain; Chest, Heart & Stroke Scotland; European Social Fund's Doctoral Studies and Internationalisation Programme DoRa; Centre for Clinical Brain Sciences, University of Edinburgh.

Abbreviations

95% CI	95% confidence interval
AD	Alzheimer’s disease
APOE	Apolipoprotein E
APP	Amyloid precursor protein
CAA	Cerebral amyloid angiopathy
CE	Cardio-embolic
CT	Computed tomography
FAD	Familial Alzheimer's disease
GRRAS	Guidelines for Reporting Reliability and Agreement Studies
GWAS	genome-wide association study
HCAA	Hereditary cerebral amyloid angiopathy
HWE	Hardy-Weinberg equilibrium
I^2	Heterogeneity metric
ICH	Intracerebral haemorrhage
IS	Ischaemic stroke
IVH	Intraventricular haemorrhage
LD	Linkage disequilibrium
LVD	Large vessel disease
MRI	Magnetic resonance imaging
OR	Odds ratio
PSEN	Presenilin
SMD	Standardized mean difference
SNP	Single nucleotide polymorphism
SVD	Small vessel disease
TOAST	Trial of Org 10172 in Acute Stroke Treatment
WMH	White matter hyperintensity

Abstract

Background: Cerebral small vessel disease (SVD) causes substantial cognitive, psychiatric and physical disabilities. Despite its common nature, SVD pathogenesis and molecular mechanisms remain poorly understood, and prevention and treatment are probably suboptimal. Identifying the genetic determinants of SVD will improve understanding and may help identify novel treatment targets. The aim of this thesis is to better understand genetic associations with SVD through investigating its pathological, radiological and clinical phenotypes.

Methods: To unravel the genetic associations with SVD, I used three complementary approaches. First, I performed a systematic review looking at existing intracerebral haemorrhage (ICH) classification systems and their reliability, to help inform future studies of ICH genetics. Second, I performed a series of systematic reviews and meta-analyses, investigating associations between genetic polymorphisms and histopathologically confirmed cerebral amyloid angiopathy (CAA). Third, I performed meta-analyses of existing genome-wide datasets to determine associations of >1000 common single nucleotide polymorphisms (SNP) in the *COL4A1/COL4A2* genomic region with clinico-radiological SVD phenotypes: ICH and its subtypes, ischaemic stroke and its subtypes, and white matter hyperintensities.

Results: The reliability of existing ICH classification systems appeared excellent in eight studies conducted in specialist centres with experienced raters, although these existing systems have several limitations. In my systematic evaluation of CAA genetics, meta-analyses of 24 studies including 3520 participants showed robust evidence for a dose-dependent association between *APOE* ε4 and histopathological CAA. There was, however, no convincing association between *APOE* ε2 and presence of CAA in a meta-analysis of 11 studies including 1640 participants. Meta-analyses of five studies including 497 participants showed, contrary to an existing popular hypothesis, that while *APOE* ε4 may increase the risk of developing severe CAA vasculopathy, there is no clear evidence to support a role of ε2. There were few data about the role of *APOE* in hereditary CAA, but in the three studies that had

looked at this, there was no evidence for an association between *APOE* ε4 and CAA severity. There were too few studies and participants to draw firm conclusions about the effect of non-*APOE* ε2/ε3/ε4 genetic polymorphisms on CAA, but there were positive associations with *TGF-β1, TOMM40* and *CR1* genes in four studies. Finally, in my meta-analyses of the *COL4A1/COL4A2* genomic region, three intronic SNPs in *COL4A2* were associated with SVD phenotypes: significantly with deep ICH, and suggestively with lacunar ischaemic stroke and WMH.

Conclusions: I have shown that while existing ICH classification systems appear to have very good reliability, further research is needed to determine their performance in different settings. For large population-based prospective studies of ICH genetics, anatomical systems are likely to be more feasible, scalable and appropriate, although they have limitations and will need to be further developed. Using systematic reviews and meta-analyses, I have confirmed a dose-related association between *APOE* ε4 and histopathological CAA, but also demonstrated that, despite popular acceptance, there is insufficient data to draw firm conclusions about the association with *APOE* ε2. I found some positive associations with CAA in other genes, which merit replication in further larger studies, and showed that there is currently insufficient data about the role of *APOE* in hereditary CAA. Finally, I identified a novel association between a locus in a known hereditary SVD gene – *COL4A2* – and sporadic SVD. This highlights a new and successful approach for selecting candidate genes and can be expanded in future studies to include other known hereditary SVD genes.

Chapter 1: Introduction

Chapter contents:

1. Introduction

- Cerebral small vessel diseases (SVDs) are common and cause substantial cognitive and physical disability
- Cerebral amyloid angiopathy affects mainly superficial cortical and leptomeningeal small vessels, while arteriolosclerosis-associated SVD preferentially affects the deep perforating vessels
- SVDs are influenced by many environmental and genetic factors and interactions between these, but their pathogenesis remains poorly understood
- Studying genetic associations with SVDs will improve this understanding and may help identify novel treatment targets
- Systematic review and meta-analysis methods can be used to study genetic association data in a meaningful way, though neither is a replacement for large robust genetic association studies
- The aim of my work has been to better understand genetic associations with SVDs through investigating their pathological, radiological and clinical phenotypes through systematic review and meta-analysis methods

1.1 Cerebral small vessel diseases

1.1.1 Importance

In the past 15 years cerebral small vessel diseases (SVDs) have been recognised as a serious problem with a high cost to society. They are very common and cause certain subtypes of stroke, as well as contributing to a substantial proportion of dementias. Since their pathogenesis and molecular mechanisms remain poorly understood, prevention and treatment are suboptimal (Wardlaw et al. 2013a) and it is an important area for further study.

1.1.2 Definition

The term SVDs refers to a group of pathological processes with various aetiologies that affect the small vessels of the brain (Pantoni 2010). The definition of a cerebral small vessel has not been uniform amongst neuropathologists (Pantoni et al. 2006). A definition recently proposed, includes all small arteries and veins, arterioles, venules, and capillaries located in the brain parenchyma or in the subarachnoid space (Pantoni 2010). Current standard neuroimaging techniques used in practice do not allow the direct visualization of small vessels. Therefore, brain parenchyma lesions thought to be caused by small vessel alterations have been adopted as the marker of SVDs, and SVDs have become a synonym of certain brain parenchyma lesions (Pantoni 2010).

1.1.3 Classification

Pantoni (2010) has proposed a simplified aetiopathogenic classification of cerebral SVDs, including six types:

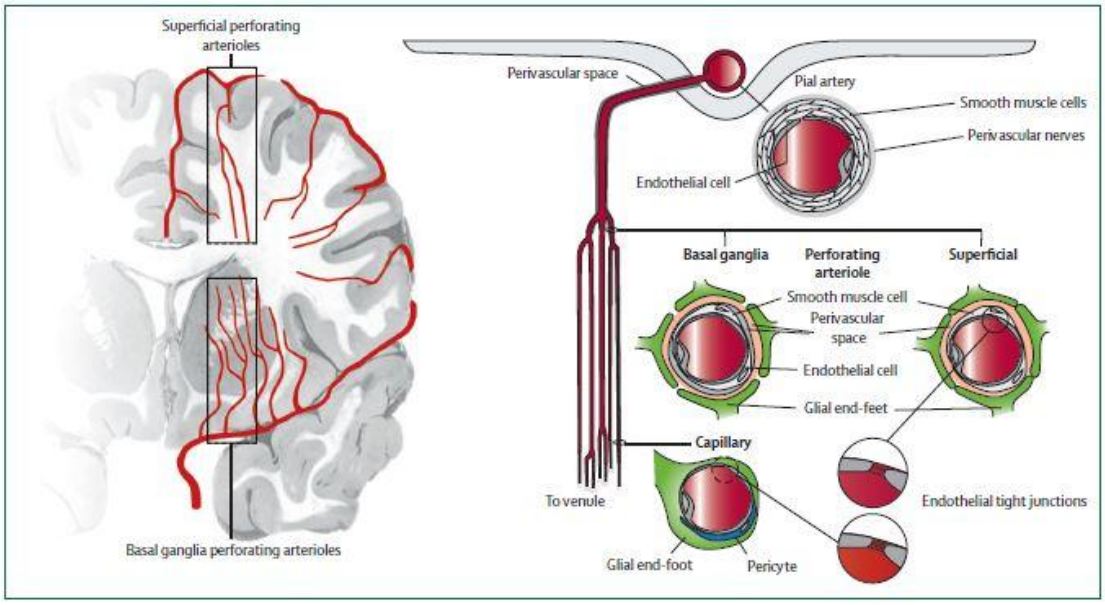
- Cerebral amyloid angiopathies (sporadic and hereditary);
- Arteriolosclerosis-associated SVD;
- Hereditary SVDs distinct from cerebral amyloid angiopathies;
- Inflammatory and immunologically mediated SVDs (vasculitis);
- Venous collagenosis;
- Other SVDs (e.g., post-radiation angiopathy).

The frequency of each of these SVD types varies considerably, with SVD due to sporadic cerebral amyloid angiopathy (CAA) and arteriolosclerosis-associated SVD being the two most common forms (Pantoni 2010). These two predominant SVD types are the focus of this thesis. Arteriolosclerosis-associated SVD is also known as non-amyloidogenic SVD and hypertensive SVD, but throughout the rest of this thesis, only the term arteriolosclerosis-associated SVD will be employed.

Cerebral arterial small vessels have two origins. Superficially, they stem from the subarachnoid circulation as the terminal vessels of medium-sized arteries, which originate from larger arteries. Deeper, at the base of the brain, they stem directly

from the large vessels as arterial perforators. These two systems converge towards each other and merge in the deepest areas of the subcortical white matter where there is a watershed area. Different underlying processes can preferentially affect either the superficial vessels or the deep perforating vessels (Figure 1.1) – while SVD due to CAA affects mainly superficial cortical and leptomeningeal small vessels, arteriolosclerosis-associated SVD preferentially affects the deep perforating vessels (please see 1.1.7.3 for more discussion on this).

Figure 1.1 Features of the arteriolar and capillary wall



Reprinted from The Lancet Neurology "Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging" 2013:12:483-497, Wardlaw JM, Smith C, Dichgans M, with permission from Elsevier.

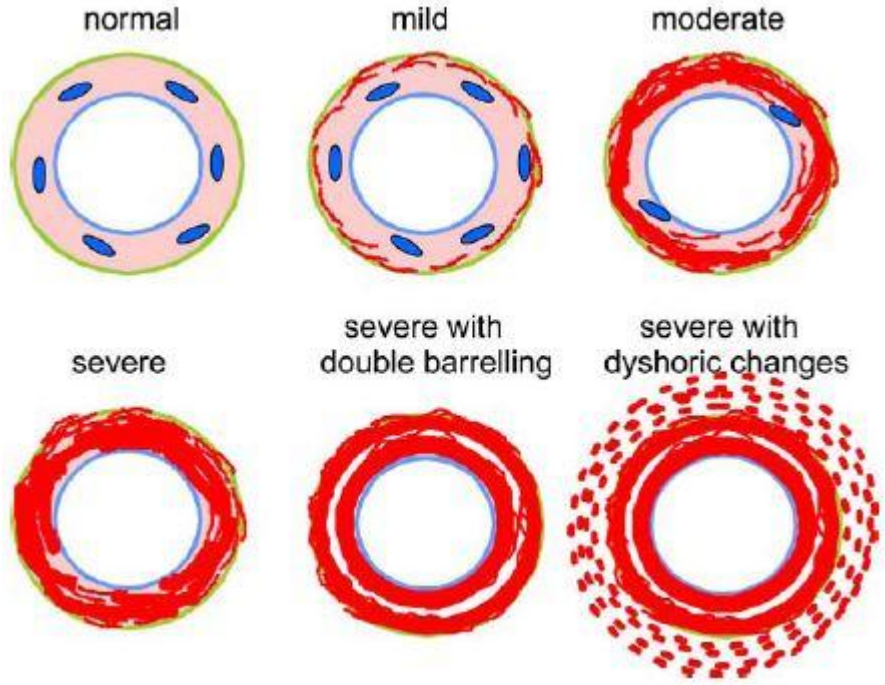
1.1.4 Pathological features

1.1.4.1 Pathological features of sporadic CAA

The term amyloid describes highly insoluble fibrils composed of protein polymers consisting of proteins rich in a β-pleated sheet secondary structure. In CAA, amyloid fibrils deposit in the walls of arteries and arterioles – and less frequently in veins and capillaries – of cortical and leptomeningeal vessels. Amyloid-β peptide is by far the most common amyloid subunit implicated in sporadic CAA (Revesz et al. 2009). CAA deposition is progressive, initially occurring in the basement membranes in the tunica media of the blood vessel wall. In the early stages, the vessel wall structure

remains intact, but as the severity of CAA increases, there is a loss of smooth muscle cells as amyloid-β infiltrates all layers of the wall (Keable et al. 2015). Finally, severe vasculopathic changes are seen: splitting or "double barrelling" of the vessel wall; microaneurysm formation; fibrinoid necrosis; and perivascular leakage (Attems 2005).

Figure 1.2 Progression of CAA



Mild: amyloid-β deposition in abluminal portions of the blood vessel wall; Moderate: abundant amyloid-β deposition in all layers of the blood vessel wall with loss of smooth muscle cells; Severe: blood vessel wall replaced by amyloid-β deposition, additional double barrelling and/or dyshoric changes (i.e. spread of amyloid-β into the surrounding synaptically dense brain areas composed mostly of unmyelinated axons, dendrites and glial cell processes). Reproduced from "Sporadic cerebral amyloid angiopathy: pathology, clinical implications, and possible pathomechanisms" in *Acta Neuropathol* 2005:110:345-359, Attems J, with permission of Springer.

Severe CAA is visible as acellular thickening of blood vessel walls on haematoxylin- and eosin-stained tissue sections, although this appearance is non-specific, occurring also with arteriolosclerosis-associated small vessel disease (Attems 2005). Stains for amyloid-β include Thioflavin S or T and Congo Red. In recent years, immunohistochemical stains for amyloid-β have become more widely used because they offer greater standardisation and increase the specificity of diagnosis (Dickson 2005).

Thal et al. (2002) have proposed that there are two different types of CAA: type-1 involves amyloid-β deposits in every type of leptomeningeal and intracortical blood vessel (including capillaries), and type-2, which is similar, but the capillaries are not involved (Thal et al. 2002).

1.1.4.2 Pathological features of arteriolosclerosis-associated SVD

Arteriolosclerosis-associated SVD is characterised by a loss of smooth muscle cells from the tunica media, fibrinoid necrosis, deposits of fibro-hyaline material, narrowing of the lumen, and thickening of the vessel wall (Fisher 1965, Ogata et al. 2014, Pantoni 2010).

Fibrinoid refers to material consisting of exudated plasma protein and necrotic smooth muscle cells. It deposits segmentally and commonly occupies a portion of the vessel. Consequently mural or occlusive thrombus formation, aneurysmal dilatation and leakage of blood components through the disintegrated vessel wall can appear.

Fibrinoid necrosis may precede hyalinosis, which involves degenerated collagen, smooth muscle cells and unspecified amorphous structures on electron microscopy. Fibrohyalinosis is used to describe arteries with thickened hyalinised walls that contain areas of fibrosis (Ogata et al. 2014).

Other possible pathological features of this form of microangiopathy are microatheroma and microaneurysms (Fisher 1965, Ogata et al. 2014, Pantoni 2010). Microatheroma encompasses atherosclerotic changes and deposition of cholesterol crystals (Ogata et al. 2014). Microaneurysms were originally described by Charcôt and Bouchard (Charcot and Bouchard 1868) and later by Miller Fisher in 1971. They are preferentially seen in the thalamus and basal ganglia in individuals with a deep intracerebral haemorrhage (Fisher 1971). The wall of the aneurysm lacks muscle or elastic tissue and sometimes shows fibrinoid necrosis.

1.1.5 Neuroimaging correlates

The consequences of SVDs on the brain parenchyma are heterogeneous and depend on whether superficial cortical or deep arterial perforator vessels are affected (Pantoni 2010). They include ischaemic and haemorrhagic manifestations.

1.1.5.1 Neuroimaging correlates of sporadic CAA

The neuroimaging correlates of CAA include lobar intracerebral haemorrhages (ICHs), brain microbleeds and white matter hyperintensities (WMH) (Charidimou et al. 2012a). More recently recognised potential imaging correlates of CAA are cortical superficial siderosis, convexity subarachnoid haemorrhage (Charidimou et al. 2012a, Linn et al. 2008), silent cortical ischaemic lesions (Charidimou et al. 2012a) and enlarged perivascular spaces (Charidimou et al. 2014).

Brain microbleeds are small hypointense foci on T2* weighted haem-sensitive MRI sequences, histopathologically corresponding to focal accumulations of hemosiderin laden macrophages (Charidimou et al. 2012a, Fazekas et al. 1999). WMH are seen on MRI as confluent areas in the hemispheric white matter that appear hyperintense on T2-weighted and fluid-attenuated inversion recovery sequences. On CT, WMH appear as hypodense periventricular or subcortical areas (Pantoni 2010). Superficial siderosis refers to cortical linear deposits of hemosiderin on T2*-weighted MRI images. Convexity subarachnoid haemorrhage describes localised bleeding in up to several adjacent sulci, without other subarachnoid bleeding at the base of the brain in the pattern typically associated with saccular aneurysm rupture.

1.1.5.2 Neuroimaging correlates of arteriolosclerosis-associated SVD

The imaging correlates of arteriolosclerosis-associated SVD include: acute lacunar infarcts; deep ICHs; lacunes (fluid-filled cavities thought to show old infarcts, many clinically silent) (Wardlaw et al. 2013a); WMH (Rost et al. 2010); enlarged perivascular spaces (Doubal et al. 2010); brain microbleeds (Cordonnier et al. 2007); and brain atrophy (Aribisala et al. 2013).

Terminology for imaging features of SVD and protocols for image acquisition and analysis vary widely. This variation inhibits cross-study comparisons and is a barrier to research on risk factors, pathophysiology, pathological correlations and clinical consequences of these lesions. To address this obstacle, an international working group from the Centres of Excellence in Neurodegeneration has developed definitions and imaging standards for markers and consequences of SVD under the acronym STRIVE (Standards for ReportIng Vascular changes on nEuroimaging) (Wardlaw et al. 2013b). They propose the following terms for neuroimaging features of SVD: (1) recent small subcortical infarcts; (2) lacune of presumed vascular origin; (3) white matter hyperintensity of presumed vascular origin; (4) perivascular space; (5) cerebral microbleed; (6) brain atrophy (Wardlaw et al. 2013b).

It has been suggested that a total SVD burden score may better capture the overall effects of SVD on the brain in comparison with only considering one or two individual features separately, but further testing of this score is required (Klarenbeek et al. 2013, Staals et al. 2014).

1.1.5.3 Shared manifestations

WMH are considered a radiological marker of arteriolosclerosis-associated SVD (Grinberg and Thal 2010, Rost et al. 2010, van Swieten et al. 1991). WMH are probably also a radiological marker of CAA (Charidimou et al. 2012a, Smith et al. 2010), however studies that have examined the relationship between CAA and white matter ischaemia at the pathological level in humans are limited (Esiri et al. 2015). There is some evidence to suggest a more posterior WMH distribution in association with CAA, compared to WMH in arteriolosclerosis-associated SVD. This is possibly explained by the CAA vascular pathology being more abundant in posterior regions of the brain, but needs confirmation in larger prospective studies (Zhu et al. 2012, Thanprasertsuk et al. 2014).

Enlarged perivascular spaces and brain microbleeds are other shared markers of different SVDs. Some evidence suggests that the location of enlarged perivascular spaces predominantly in the white matter and centrum semiovale region may be associated with CAA, while basal ganglia region enlarged perivascular spaces may

be a marker of arteriolosclerosis-associated SVD (Charidimou et al. 2014, Charidimou et al. 2013, Doubal et al. 2010, Martinez-Ramirez et al. 2013, Roher et al. 2003). Similarly, it has been suggested that the location of brain microbleeds may differ between CAA and arteriolosclerosis-associated SVD, with strictly lobar microbleeds being more suggestive of CAA (Charidimou et al. 2012a, Greenberg et al. 1999, Greenberg et al. 2009, Knudsen et al. 2001, Vernooij et al. 2008).

Another emerging feature of SVDs visible at higher MRI field strengths includes microinfarcts (Smith et al. 2012, Wardlaw et al. 2013a). Microinfarcts are mainly a pathological finding defined as sharply delimited microscopic regions of cellular death or tissue necrosis, sometimes with cavitation. They are distinct from small infarcts and lacunes and are undetectable on gross pathological examination or conventional structural MRI (Smith et al. 2012). Cortical microinfarcts have been associated with CAA (Haglund et al. 2006), while subcortical microinfarcts have been suggested as a potential marker of arteriolosclerosis-associated SVD (Smith et al. 2012).

1.1.6 Clinical manifestations

Clinical manifestations of SVD range from none to discrete focal neurological symptoms (e.g., stroke), to insidious global neurological dysfunction and dementia.

1.1.6.1 Ischaemic and haemorrhagic symptomatic stroke

SVDs are associated with ischaemic strokes and ICH, reflecting the distribution of the underlying microangiopathy. SVD due to CAA is associated mainly with (often multiple) lobar ICHs preferentially affecting the occipital and temporal lobes (Charidimou et al. 2012a, Samarasekera et al. 2012, Wardlaw et al. 2013a), while arteriolosclerosis-associated SVD manifests clinically as lacunar ischaemic strokes and deep subcortical and infratentorial ICHs.

1.1.6.2 Classification of stroke subtypes

Risk factors (including genetic risk factors), causal pathways, investigations, management and prognosis probably differ between stroke subtypes. Hence studies need to use classification systems that distinguish stroke subtypes accurately.

Existing ischaemic stroke classification systems are 'anatomical' or 'mechanistic'. The Oxfordshire Community Stroke Project classification (OCSP) is the most widely used anatomical system and predicts infarct site and size (Bamford et al. 1991). Mechanistic classification systems can be either 'single cause', assigning the single most likely stroke mechanism (e.g., TOAST [the Trial of Org 10172 in Acute Stroke Treatment] (Adams et al. 1993), CCS [Causative Classification] (Ay et al. 2005) or ASCO1 [Atherosclerosis-Small artery occlusion-Cardiac Embolism-Other-1] (Amarenco et al. 2009)), or 'descriptive' (e.g., the A-S-C-O or 'phenotypic CCS' systems (Arsava et al. 2010)), describing the contribution of each mechanism to the stroke.

Existing intracerebral haemorrhage (ICH) classification systems are also 'anatomical' or 'mechanistic'. Anatomical systems classify haemorrhages according to their anatomical origin or location as lobar, deep, infratentorial, intraventricular, and various combinations or modifications of these. Mechanistic systems integrate this anatomical information with clinical symptoms, signs and investigations, to assign a subtype based on presumed mechanism. Commonly used categories include hypertension, CAA, anticoagulation and structural causes. The only clearly structured, published, mechanistic system for ICH is SMASH-U (Structural lesion, Medication, Amyloid angiopathy, Systemic/other disease, Hypertension, Undetermined) (Meretoja et al. 2012).

An ideal classification system for large-scale research use, including for studies of stroke genetics, would assign the maximum number of cases to determined, valid subtypes, without sacrificing reliability or accuracy, and would be applicable in a range of different clinical settings. A recent systematic review explored existing ischaemic stroke classification systems, assessed their reliability, and made recommendations for future studies for classifying ischaemic stroke (Woodfield et al.

2013). However, to my knowledge, no such study assessing existing ICH classification systems exists.

1.1.6.3 Other clinical manifestations of SVD

Arteriolosclerosis-associated SVD causes substantial cognitive, psychiatric, and physical disabilities and contributes to up to 45% of dementias (Wardlaw et al. 2013a). There is increasing evidence that CAA is also an important contributor to cognitive impairment. In addition, SVD due to CAA is associated with transient neurological episodes called "amyloid spells", the prevalence of which in a recent multicentre CAA cohort (diagnosed using the Boston criteria) was 14% (Charidimou et al. 2012b). Finally, CAA is associated with rapidly progressive cognitive and neurological decline in the context of CAA-related inflammation, which typically affects older adults who present with acute to subacute cognitive decline, headache, behavioural change, seizures and focal neurological deficits (Charidimou et al. 2012a).

1.1.7 Pathogenesis

The pathogenic mechanisms linking the pathology of small vessels with the consequent parenchymal changes are largely unknown and likely very complex. Pathology often shows end-stage disease which makes identification of the earliest stages difficult (Wardlaw et al. 2013a). It is known that there are striking changes in small arteries and arterioles, summarised by the loss of smooth muscle cells and the replacement of wall constituents by other materials (e.g. collagen, amyloid, or hyaline material). The vessel lumen is restricted or dilated and the general architecture of the vascular tree disrupted with increased vessel tortuosity, elongation and formation of aneurysmal sacs. The probable subsequent mechanisms leading to the brain parenchymal damage are:

- reduced or dysregulated blood flow, which if focal and severe may lead to small infarcts;
- some areas suffer from chronic ischaemia causing selective damage to certain cell populations, rather than tissue pan-necrosis;

- vessel walls rupture and cause parenchymal haemorrhage;
- changes in small vessels lead to damage of the blood-brain barrier, which leads to extravasation of fluids and different molecules into the brain parenchyma.

The precise mechanisms leading from the vessel wall damage to supposedly ischaemic parenchymal insults remain unknown at present, as local thrombosis has not been documented and other possible mechanisms such as local vasospasm or lumen occlusion without thrombosis are difficult to prove. It is possible that thrombosis is a late secondary phenomenon, triggered by the disrupted arteriolar endothelium. Thickened stiffened arterioles may have impaired autoregulatory ability, and contribute further to brain damage through reduced ability to vasodilate when required. Both of these late-stage consequences could be superimposed on the background initiating process of endothelial failure. Besides ischaemia and blood-brain barrier dysfunction, other mechanisms to link the pathologic changes of small vessels to parenchymal damage have been proposed, such as venous dysfunction, increased arterial pulsatility and apoptosis. It is possible that these mechanisms are not mutually exclusive and act contemporaneously (Pantoni 2010, Wardlaw and Pantoni 2014, Wardlaw et al. 2013a).

Grinberg and Thal have put forward a theory suggesting arteriolosclerosis-associated SVD may exacerbate CAA by impairing perivascular clearance of amyloid-β. Proposed mechanisms include: (a) fibrosis of the vessel wall decreases the capacity for drainage of amyloid along the basement membranes, (b) increasing stiffness of the affected vessels leads to cessation of pulsation and a less effective outward transport of the perivascular fluid, (c) arteriolosclerosis-associated SVD induced leakage of plasma proteins into the vessel wall and the perivascular space competes for perivascular drainage of amyloid (Grinberg and Thal 2010).

1.1.7.1 Pathogenesis of sporadic CAA

Factors that initiate or promote amyloid-β deposition in sporadic CAA are not fully understood. However, research based on mouse models has provided valuable insights into its pathogenesis. It is thought that most vascular amyloid-β is derived from neurons (Herzig et al. 2006). CAA is likely to result from reduced amyloid-β clearance (rather than overproduction) through perivascular lymphatic drainage pathways in the brain (Weller et al. 2008). Theoretical models suggest that vessel pulsations supply the force for perivascular drainage of interstitial fluid and solutes. As arteries stiffen with age, the amplitude of pulsations is reduced and insoluble amyloid-β is deposited in the drainage pathways as CAA, thus further impeding the drainage of soluble amyloid-β (Charidimou et al. 2012a, Weller et al. 2008).

1.1.7.2 Pathogenesis of arteriolosclerosis-associated SVD

Though arteriolosclerosis-associated SVD is typically thought of as an age-related hypertensive vasculopathy, the association between vascular risk factors and SVD is still not completely understood (Wardlaw et al. 2013a). While exposure to vascular risk factors, particularly hypertension, will probably accelerate the vessel damage (Wardlaw and Pantoni 2014), epidemiological studies indicate that patients with lacunar stroke are equally likely to have a history of hypertension, diabetes, hypercholesterolaemia and smoking as are patients with non-lacunar ischaemic stroke (Jackson and Sudlow 2005, Jackson et al. 2010).

1.1.7.3 Differing structure of cortical and deep arteries

CAA affects mainly superficial cortical and leptomeningeal vessels (Ogata et al. 2014) which may be at least partly explained by the differing structures of the walls of cortical and deep arteries. Cortical arteries have only one leptomeningeal layer, and therefore lack the expandable perivascular space, which may make them more vulnerable to deposition of amyloid-β which cannot be drained (Figure 1.1). In contrast, deep arteries have a double layer of leptomeninges which can function as an expandable perivascular space (Pollock et al. 1997, Zhang et al. 1990).

Arteriolosclerosis-associated vasculopathy occurs mainly in the basal grey matter, pons and cerebellum (Ogata et al. 2014). The endothelial tight junctions are tightest in capillaries, in which the barrier function is most important, and looser in arteriolar and venular endothelium. The hypothetical consequence of this is that the effects of endothelial failure are seen earlier in larger proximal perforating arterioles than in the smaller capillaries. This difference could lead to differential development of proximal arteriolar disease and lacunar infarcts in the basal ganglia before more diffuse white matter hyperintensities in the centrum semiovale develop (in which tissue is served by smaller arterioles and capillaries).

1.1.8 Diagnosis

The reference standard for the diagnosis of CAA is pathological confirmation using biopsy or autopsy specimens (Charidimou et al. 2012a). Since patients with ICH may not undergo surgery or autopsy, the Boston criteria have been developed for an ante-mortem diagnosis of CAA. These criteria categorise ICH as definite, probable or possible CAA-related ICH by using pathological findings if available, radiographic features such as multiple ICHs or brain microbleeds, and clinical features such as the absence of coagulopathy (Knudsen et al. 2001). The Boston criteria were validated in a selected sample of 39 participants at a tertiary referral centre and their sensitivity for detection of CAA was 45% (95% confidence intervals [CI] 28-62%) and specificity 100% (95% CI 77-100%) (Knudsen et al. 2001). However, the number of neuropathological samples viewed per brain and their locations is unclear, which may influence the likelihood of finding other pathologies in the brain and determining the relative contribution of each to the ICH. More recently modified Boston criteria have been proposed, that include superficial siderosis in the radiological features, but this has not been shown to significantly increase their sensitivity (Linn et al. 2010). An important limitation for the use of the Boston criteria in population-based studies of ICH is that MRI is often unavailable or not tolerated.

To my knowledge, no specific criteria exist for the diagnosis of arteriolosclerosis-associated SVD, and the diagnosis is made based on typical radiological and clinical findings.

1.1.9 Hereditary forms

Monogenic or hereditary SVDs are caused by highly penetrant mutations in a single gene. In some of these disorders cerebral SVD is the main manifestation, in others SVD is part of a wider phenotypic spectrum (Opherk and Dichgans 2014).

1.1.9.1 Hereditary CAAs

Amyloid-β peptide is by far the most common amyloid subunit implicated in sporadic forms of CAA and Alzheimer's disease (AD). Amyloid-β CAA is also prominent in variants of familial AD due to mutations of the amyloid precursor protein (*APP*), presenilin-1 (*PSEN1*) or presenilin-2 (*PSEN2*) genes and in several forms of hereditary CAA (HCAA) due to mutations in the *APP* gene.

A number of other proteins may also be involved in rare familial diseases in which CAA is a characteristic morphological feature. These latter proteins include (Revesz et al. 2009):

- ABri and ADan subunits in *BRI2* gene-related familial British and familial Danish dementia;
- variant cystatin C in *cystatin C* gene-related Hereditary Cerebral Haemorrhage with Amyloidosis-Icelandic type;
- variant transthyretins in *TTR* gene-related meningo-vascular amyloidosis;
- disease-associated prion protein (PsP^{Sc}) in *PRPN* gene-related hereditary prion disease with premature stop codon mutations;
- mutated gelsolin (AGel) in *gelsolin* gene-related familial amyloidosis of Finnish type.

In the case of these rare familial diseases, in addition to the central nervous system, vascular deposits of the pathological protein can often be found in a variety of

peripheral tissues and systemic organs (Opherk and Dichgans 2014, Revesz et al. 2009).

1.1.9.2 Hereditary non-CAA SVDs

Currently mutations in at least seven genes (*NOTCH3, HTRA1, TREX1, CECR1, GLA, COL4A1* and *COL4A2*) are known to cause rare familial forms of non-CAA SVD:

- *NOTCH3* mutations cause cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) (Chabriat et al. 2009);
- *HTRA1* mutations are associated with cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL) (Hara et al. 2009);
- *TREX1* mutations cause retinal vasculopathy with cerebral leukodystrophy (RVCL) (Kavanagh et al. 2008);
- *CECR1* mutations cause deficiency of adenosine deaminase 2 (DADA2) (Zhou et al. 2014);
- *GLA* mutations cause Fabry's disease (Opherk and Dichgans 2014);
- *COL4A1* and *COL4A2* mutations cause hereditary angiopathy with nephropathy, aneurysms, and muscle cramps (HANAC) and other syndromes (Lanfranconi and Markus 2010, Opherk and Dichgans 2014).

One emerging view is that these familial SVDs are united by similar disease mechanisms involving perturbation of the cerebrovascular matrisome (i.e. the ensemble of proteins constituting the extracellular matrix as well as the associated proteins), which could be a convergent pathway that drives the functional and structural alterations of small brain vessels and disease manifestations, and that similar mechanisms could also play a role in sporadic disease (Joutel et al. 2015).

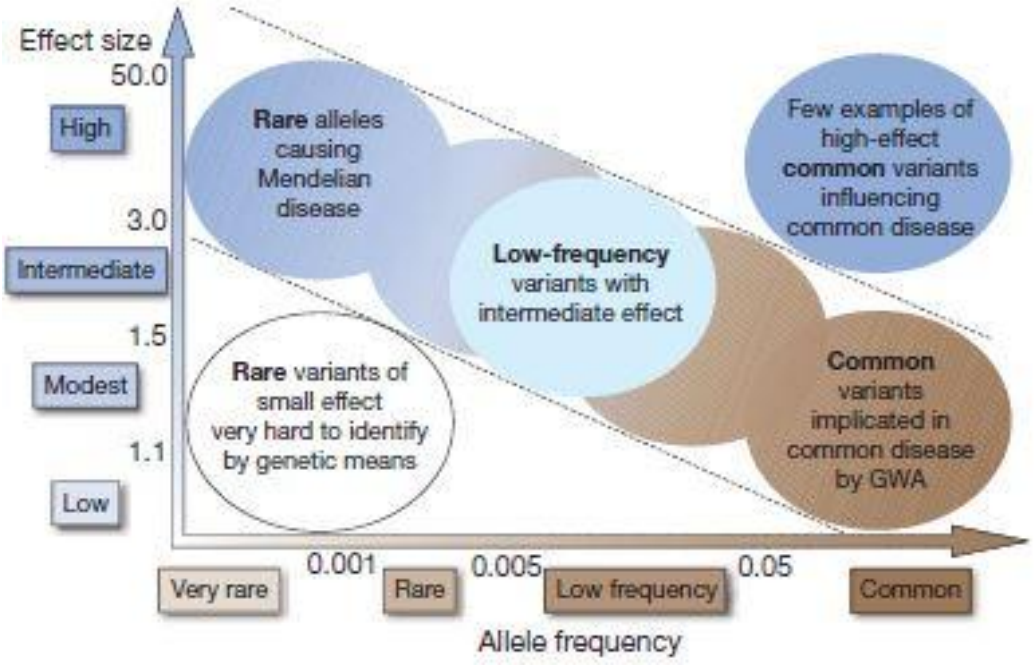
1.2 Studying genetic associations of complex disease

Common complex diseases, such as stroke and cerebral small vessel disease, are influenced by many environmental and genetic factors, and interactions between these. Studying genetic associations with complex diseases is important, because a substantial proportion of individual differences in disease susceptibility are known to be due to genetic factors. Understanding these may contribute to better prevention, diagnosis and treatment of the disease (Manolio et al. 2009).

Studies of rare monogenic disorders and common complex diseases have supported the correlation between variant frequency and effect size, whereby rare variants tend to associate with large and often deleterious effects, while common variants are more likely to have subtle effects on disease susceptibility (Figure 1.3). The relative impact of common and rare variants on common complex diseases remains an unanswered question (Gibson 2011, Manolio et al. 2009).

There are two principle approaches to the identification of disease genes: linkage analysis and association studies. Linkage analysis relies on the coinheritance of loci that lie near each other on the same chromosome. As a consequence, linkage studies require DNA and clinical information from more than one family member. This approach has been used successfully in monogenic disorders, but there are a number of challenges to its application in complex diseases (Dichgans and Markus 2005). These include low power of tests of linkage for genes of modest effect and assembling sufficient number of families (Risch and Merikangas 1996). While some genes of small effect could be uncovered by linkage analysis, it would be necessary to collect and analyse samples of 'unrealistic' size, to achieve adequate statistical power (Risch and Merikangas 1996).

Figure 1.3 Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio)



Most emphasis and interest lies in identifying associations with variants shown within diagonal dotted lines. Reprinted with permission from Macmillan Publishers Ltd: "Finding the missing heritability of complex diseases", Manolio et al, Nature 2009;461:747-753.

Association studies compare the frequency of specific DNA sequence variants in groups of individuals in a case-control design. An allele is said to be associated with disease if its frequency differs between cases and controls more than is expected by chance. For complex traits, association studies are generally considered to be more efficient than linkage studies because they have greater statistical power to detect several genes of small effect (Dichgans and Markus 2005, Risch and Merikangas 1996).

Association studies can use an *a priori* hypothesis-driven candidate-gene approach, or hypothesis-free genome-wide association study (GWAS), whole exome sequencing or whole genome sequencing approaches. Rather than relying on markers throughout the genome without regard to their function or context in a specific gene, candidate-gene studies focus on genes that are selected because of *a priori* hypotheses about their aetiological role in disease. Two important criticisms of this study design are the failure to replicate significant findings from many candidate-

gene studies and its dependence on the ability to predict functional candidate genes and polymorphisms (Tabor et al. 2002). GWAS studies allow examination of mostly common single nucleotide polymorphisms (SNPs) (minor allele frequency ≥1-5%) across the genome, but miss rare variation. In addition, variants identified by GWAS are not usually the causative variants themselves, but rather in linkage disequilibrium with the causative variant (Hirschhorn and Daly 2005). While GWAS studies primarily make use of markers that are intended to represent causal variation indirectly, whole exome sequencing and whole genome sequencing approaches can directly identify the causal (rare and common) variants. This is the central advantage of sequencing approaches over standardised GWAS panels, especially because complex diseases could be influenced by fairly infrequent variants in many different genes. However, despite the promise of next generation sequencing, the challenge is in appropriately analysing and interpreting the resulting sequencing data. For GWAS studies, a simple single statistical threshold to report significance between a given polymorphism and phenotype ($p < 5x10^{-8}$) has been derived based on the principle that it needs to account for the number of possible independent statistical tests with equal prior probabilities of true association with the phenotype (Goldstein et al. 2013). The same type of solution is not applicable to sequence data, as it reveals inherently different categories of variants that cannot be viewed as all having the same probability of influencing the disease, and treating them as such would constitute too great a cost in terms of power. Exome-sequencing data are based on the idea that mutations influencing human disease are more likely found in the coding sequence than elsewhere. Nonetheless, there is currently too little information about the full distribution of functional consequences of different variants in the human genome to allow simple quantitative assumptions to be applied to the variants in a universal fashion (Goldstein et al. 2013).

1.3 Systematic review and meta-analysis for studying genetic associations

A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses systematic

methods selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn (The Cochrane Collaboration 2011). Many systematic reviews contain meta-analyses. Meta-analysis uses quantitative methods for combining results of different studies on the same research question, thus increasing statistical strength and precision. It also measures and tries to explain the extent of inconsistency across the included studies. Both can be very useful tools in understanding the genetics of complex diseases, provided they are appropriately applied and interpreted.

Genetic association studies of candidate genes have shown that a substantial proportion of initial positive results cannot be replicated in subsequent studies, suggesting possible false-positive results. Equally, results may be false-negative with small genetic effects remaining undetectable (Ioannidis 2003). Meta-analysis combining data from several candidate-gene studies may help separate true effects from false. Meta-analysis has also become popular for synthesis of data from GWAS studies, given that most genetic effects are small and require the coalition of many teams to generate large-scale evidence. However, like any research design, meta-analysis has strengths but also weaknesses (Kavvoura and Ioannidis 2008).

Ideally, meta-analysis should be prospective, including all data that fulfil pre-set quality criteria. However if it is retrospective, pertinent data should be retrieved in an unbiased way, such as by use of systematic review methods. This should include searching for all relevant studies from appropriate databases and considering including unpublished data and non-English literature where possible (Kavvoura and Ioannidis 2008). The validity of including unpublished data that has not passed peer-review can be questioned, however the peer-review filter may be imperfect and there is some evidence to suggest that inclusion of unpublished data may lead associations to lose their statistical significance, suggesting that the published literature is shaped by selective reporting biases (Kyzas et al. 2005).

The reliability of meta-analysis is affected by the reliability of the included data. Hence the data should be examined for potential errors and biases, and these should be carefully recorded. Biases can be broadly divided into study-specific and field-wide biases. Study-specific biases include: selection bias (e.g. differential selection

of cases and controls creating unequally representative groups within study populations); information bias (e.g. poor quality control measures for genotyping and phenotype misclassification); and confounding (e.g. population stratification when the total population has been formed by admixture between subpopulations and admixture proportions vary between cases and controls). Field-wide biases include publication bias (preference to publish studies with "positive" findings i.e. statistically significant results) and reporting bias (when investigators publish only a subset of the analyses they have conducted with preference to most impressive results) (Kavvoura and Ioannidis 2008).

Some of the heterogeneity across studies may be explained by the errors and biases mentioned above, however there may also be genuine diversity in the genetic effects across studies. It is important to explore this and to use appropriate statistical heterogeneity metrics (i.e. Cochran's Q or χ^2statistic and I^2). The chi-squared (χ^2) test assesses whether observed differences in results are compatible with chance alone. A low p value of the χ^2 test or a large χ^2 statistic relative to its degree of freedom provides evidence of heterogeneity. It has low power when there are few studies and/or studies have a small sample size. Hence while a statistically significant result may indicate a problem with heterogeneity, a non-significant result must not be taken as evidence of no heterogeneity (The Cochrane Collaboration 2011). Some argue that since clinical and methodological diversity always occur in a meta-analysis, heterogeneity will always exist whether or not it can be detected using a statistical test (Higgins et al. 2003). Methods have been developed for quantifying inconsistency across studies that move the focus away from testing whether heterogeneity is present to assessing its impact on the meta-analysis. A useful statistic for quantifying inconsistency is I^2, which describes the percentage of the variability in effect estimates that is due to heterogeneity rather than chance (The Cochrane Collaboration 2011).

The fixed- and/or random-effects model should be chosen as appropriate (Kavvoura and Ioannidis 2008). It is generally recommended to use the fixed-effects method when the underlying assumption is that the true allele effect is the same in all studies included in the meta-analysis, and to use the random-effects method when the

underlying assumption is that the true allele effect is different in the included studies (Borenstein et al. 2010). When no heterogeneity exists, both models show similar effect sizes.

A number of diagnostics have been proposed to try and assess the robustness of summary results and explore heterogeneity. Sensitivity analyses (excluding specific studies) are a simple way of examining if the summary results and estimated heterogeneity depend on specific studies with possible errors, biases or special features. Meta-regression relates the effect size to one or more characteristics of the studies that may explain heterogeneity, but the covariates should be pre-specified and few (Thompson and Higgins 2002). There are several tests to explore potential publication bias (e.g. funnel plot asymmetry, "trim and fill"), however researchers need to be careful not to over-interpret and over-simplify their results. For example, significant results in asymmetry tests do not mean that publication bias is certain, because there are many other reasons why small studies may yield different results from larger ones. Equally, non-statistically significant results or a symmetrical funnel plot on visual inspection do not exclude publication bias. However, demonstration of small-study effects should lead to caution in interpreting summary results (Kavvoura and Ioannidis 2008). Finally, the potential effect of unpublished negative studies or studies not reporting the data required for meta-analysis (i.e. publication and reporting biases) can be assessed using a modified 'failsafe N' method. This method involves determining the size of a notional study with a null result required to bring any significant result to a just non-significant level (Rosenthal 1979).

Overall, while systematic reviews and meta-analyses are not suitable replacements for large robust genetic association studies, they provide useful additional tools to allow examination of data in a way that is both meaningful and less susceptible to bias than narrative reviews.

1.4 Known genetic associations with sporadic cerebral small vessel disease

Heritability is a measure indicating the proportion of variation in phenotype in a population explained by genetic factors. Traditionally, heritability has been assessed by studying related individuals. Family-based studies have demonstrated that WMH are highly heritable, with heritability measures ranging from 50% to 80% (Haffner et al. 2015, Opherk and Dichgans 2014). In clinically evident stroke, however, family-based studies are difficult to apply because of the relatively late age of onset and high case fatality. Therefore an alternative way of measuring heritability by using genome-wide complex trait analysis of GWAS data has been applied. This method allows calculation of the proportion of phenotypic variance explained by common genetic variants present on GWAS genotyping arrays. Using genome-wide complex trait analysis, the heritability of WMH volume in ischaemic stroke cases is estimated to be 21% (Adib-Samii et al. 2015). Lacunar ischaemic stroke heritability ranges from 16% to 25%, depending on the exact method used to measure heritability and the definition of lacunar stroke (TOAST classification versus MRI-defined) (Bevan et al. 2012, Traylor et al. 2015). Lobar ICH heritability is 73% (48% for non-*APOE* loci) and the respective deep ICH heritability is 34% (30% for non-*APOE* loci) (Devan et al. 2013). The discrepancy between high heritability estimates for WMH obtained in twin and family history studies and relatively low estimates derived from genome-wide complex trait analysis might indicate that much of the genetic contribution to SVD is because of rare genetic variants (Haffner et al. 2015). Also, this narrow-sense heritability does not capture gene-environment interactions or gene-gene interactions (denoted as broad-sense heritability).

Most genetic studies of SVD rely on either its radiological or clinical markers, or both, and none of these are specific for SVD. Also, the majority of studies have focused on individual markers of SVD rather than considering multiple markers in aggregate. In practice, large sample sizes are often used to "compensate" for the limited specificity of individual disease markers (Haffner et al. 2015).

So far, data about SVD genetic associations come from candidate-gene and GWAS studies, as no large whole exome or whole genome sequencing studies have been conducted.

White matter hyperintensities

GWAS of WMH burden in stroke-free individuals from the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium (CHARGE) revealed a risk locus on chromosome 17q25, where two lead SNPs belong to a cluster containing several genes (*TRIM65, TRIM 47, WBP2, MRPL38, FBF1, ACOX1) (Fornage et al. 2011).* A more recent multi-ethnic GWAS in stroke-free individuals replicated this locus and identified four novel loci associated with WMH burden: chromosome 10q24 (*PDCD11, NEURL, SH3PXD2A, TAF5, CALHM1)*; chromosome 2p21 (*HAAO, THADA)*; chromosome 1q22 (*PMF1)*; and chromosome 2p16 (*EFEMP1)* (Verhaaren et al. 2015). The 17q25 locus was also associated with WMH burden in ischaemic stroke cases but not with lacunar stroke, indicating that it acts through mechanisms more specific for the pathogenesis of WMH rather than SVD in general (Adib-Samii et al. 2013).

In a more recent GWAS of WMH burden in ischaemic stroke cases, there were no associations at genome-wide significance (Traylor et al. 2016). However in a subsequent meta-analysis including both stroke-free individuals and ischaemic stroke cases, six independent loci were significantly associated with WMH burden, four of these novel associations: rs72934505 (*NBEAL1*), rs941898 (*EVL*), rs962888 (*C1QL1*), rs9515201 (*COL4A2*) (Traylor et al. 2016). This suggests that some genetic associations with WMH burden are shared in stroke-free individuals and patients with stroke.

Many candidate-gene studies have attempted to find associations with WMH. These have been summarised in a systematic review and series of meta-analyses, which concluded that while numerous genes have been evaluated, there was no convincing evidence for an association with WMH. There did appear to be an association with an insertion/deletion polymorphism in the *ACE* gene, but this may have been due to small-study and other biases (Paternoster et al. 2009). A more recent systematic

review and meta-analysis specifically looking at the *APOE* gene demonstrated that both *APOE* ε4 and ε2 are associated with WMH burden (Schilling et al. 2013). Finally, variants in the *NOTCH3* gene were reported to be associated with WMH burden in a community-dwelling elderly cohort (Schmidt et al. 2011), but a more recent study in ischaemic stroke cases has not replicated this signal (Rutten-Jacobs et al. 2015).

Lacunar stroke

Although GWAS studies in ischaemic stroke patients have identified several loci for large artery and cardioembolic stroke, until recently, no GWAS significant loci for lacunar stroke have been demonstrated. This may be at least partly explained by insufficient sample size, sample heterogeneity and lack of robust standards for classifying patients as having a lacunar stroke. In the more recently published ischaemic stroke GWAS performed by the International Stroke Genetics Consortium and NINDS Stroke Genetics Network (SiGN), the 12q24.12 locus near *ALDH2*, which has been associated with all ischaemic stroke previously, exceeded genome-wide significance in the meta-analysis of lacunar stroke (NINDS Stroke Genetics Network (SiGN) and International Stroke Genetics Consortium (ISGC) 2015).

Candidate-gene studies have provided some evidence to suggest that a SNP in *PRKCH*, which is monomorphic in the European population, may be a risk locus for lacunar stroke exclusively in the Asian population (Kubo et al. 2007). The *PRKCH* locus has also been associated with subcortical silent brain infarcts and cerebral haemorrhage in the Chinese (Wu et al. 2009) (Haffner et al. 2015, Serizawa et al. 2008).

Brain microbleeds

To my knowledge, no GWAS of brain microbleeds has been published. A systematic review looking at genetic associations with brain microbleeds found that only the *APOE* ε2/ε3/ε4 polymorphism had been studied in >100 people. Meta-analyses showed a significant association between the ε4 allele and presence of brain microbleeds, especially in a strictly lobar location (Maxwell et al. 2011). This has

been confirmed by a more recent systematic review and meta-analysis (Schilling et al. 2013).

Intracerebral haemorrhage

ICH GWAS has identified a susceptibility locus for non-lobar ICH on chromosome 1q22 (lead SNP rs2984613), that overlaps with the *SLC25A44* and *PMF1* genes (Woo et al. 2014). Although not GWAS significant, the same 1q22 locus showed a trend towards association with WMH in stroke-free individuals (Fornage et al. 2011).

Candidate-gene studies have shown that *APOE* ε2 and ε4 alleles are associated with lobar ICH (Biffi et al. 2010). In addition, rare mutations in *COL4A1* and *COL4A2* have recently been shown to be associated with sporadic ICH (Jeanne et al. 2012, Schmidt et al. 2011, Weng et al. 2012). Finally, several variants within a larger set of oxidative phosphorylation genes were collectively associated with increased risk of both lacunar stroke and deep ICH (Anderson et al. 2013).

1.5 Aims of the thesis

Cerebral small vessel disease is very common, causing substantial cognitive, psychiatric and physical disabilities, and its cost to society is huge. Although substantial progress has been made in elucidating the molecular, cellular, and physiologic mechanisms underlying SVD, its pathogenesis still remains largely unknown. Understanding genetic associations with SVD may contribute to better understanding of its pathogenesis, and ultimately to prevention, diagnosis and treatment (Haffner et al. 2015, Wardlaw et al. 2013a). The aim of my work has been to better understand genetic associations with SVD phenotypes. In pursuit of this aim, the objectives of the work described in this thesis are:

- To perform a systematic review looking at the inter- and intra-rater reliability of existing ICH classification systems, to inform ICH subtyping in future large-scale prospective studies of ICH, including of ICH genetics (Chapter 2).

- To perform a systematic review and series of meta-analyses to assess the evidence for an association between the *APOE* ε2/ε3/ε4 polymorphism and histopathologically confirmed CAA (Chapter 3).
- To perform a systematic review and series of meta-analyses to assess the evidence for an association between the *APOE* ε2/ε3/ε4 polymorphism and progression to severe CAA vasculopathy (Chapter 4).
- To perform a systematic review to assess the evidence for an association between the *APOE* ε2/ε3/ε4 polymorphism and CAA severity in brain pathology samples of cases with genetically proven hereditary CAA (Chapter 5).
- To perform a systematic review to assess the evidence for an association between any (non-*APOE* ε2/ε3/ε4) genetic polymorphism and histopathologically confirmed CAA (Chapter 6).
- To conduct meta-analyses of existing GWAS data to determine if there is an association between common SNPs in the *COL4A1/COL4A2* genomic region and clinical / radiological manifestations of SVD (Chapter 7).

Chapter 2: Reliability of intracerebral haemorrhage classification systems

Chapter contents:

2. Reliability of intracerebral haemorrhage classification systems

- Accurately distinguishing ICH subtypes is important since they may have different risk factors, causal pathways, management and prognosis
- Existing ICH classification systems are 'anatomical' or 'mechanistic'
- An ideal classification system for sub-classifying ICH in large, population-based, prospective studies (including studies of ICH genetics) would assign the maximum number of cases to determined, valid subtypes without sacrificing reliability or accuracy, and be applicable in different clinical settings
- In this chapter I present my systematic review looking at existing ICH classification systems and their inter- and intra-rater reliability
- I used a comprehensive search strategy, critically appraised study quality, and assessed potential bias and factors affecting reliability
- I conclude that the reliability of existing classification systems appears excellent, but is unknown outside specialist centres with experienced raters

2.1 Introduction

Non-traumatic intracerebral haemorrhage (referred to in this chapter as ICH) accounts for 10-20% of strokes worldwide. Although age-standardized mortality rates for haemorrhagic stroke have decreased in the past two decades, the absolute number of those affected, of related deaths and of disability-adjusted life-years lost, is increasing (Krishnamurthi et al. 2013). Risk factors, causal pathways, investigations, management and prognosis may differ between ICH subtypes, and stratified approaches to treatment may be appropriate (Al-Shahi Salman et al. 2009). For example, some genetic associations differ between anatomical ICH subtypes, with *APOE* ε2/ε3/ε4 polymorphism being associated with lobar ICH (Biffi et al. 2010, Falcone et al. 2014) and a locus on chromosome 1q22 being associated with non-lobar ICH (Woo et al. 2014). Hence studies of ICH need to use classification

systems that distinguish subtypes accurately. Such studies also need to be very large for adequate statistical power and misclassification of subtypes would further reduce statistical power to detect differences between them (Burton et al. 2009).

Existing ICH classification systems are 'anatomical' or 'mechanistic'. Anatomical systems classify haemorrhages according to their anatomical origin or location on brain imaging (CT and/or MRI), or rarely, based on autopsy results as lobar, deep, infratentorial, intraventricular, and various combinations, or modifications of these. Mechanistic systems integrate this anatomical information with clinical symptoms, signs and investigations to assign a subtype based on presumed mechanism. Commonly used categories include hypertension, cerebral amyloid angiopathy (CAA), anticoagulation and structural causes.

The work described in this chapter was part of an initiative to develop scalable methods for sub-classifying ICH in large, population-based, prospective studies, including studies of ICH genetics. An ideal classification system for large-scale use would assign the maximum number of cases to determined, valid subtypes without sacrificing reliability or accuracy, and would be applicable in a range of different clinical settings.

Reliability and agreement (terms often used interchangeably) are ways of demonstrating reproducibility: the degree to which repeated measurements in stable study objects provide similar results. Reliability measures how well objects can be distinguished from each other despite measurement errors, while agreement assesses exactly how close the scores for repeated measurements are (de Vet et al. 2006). There are various measures of reliability depending on the data format and nature: the kappa statistic (measure of "true agreement" i.e. proportion of agreement beyond that expected by chance) used for nominal and ordinal data; ranked intra-class correlation used for ordinal data; and intraclass correlation coefficient used for continuous data (Kottner et al. 2011). Results of reliability and agreement studies provide information about the amount of error inherent in a classification which in turn may affect its validity (Kottner et al. 2011).

In this chapter, I aim to identify existing ICH classification systems, and assess their inter- and intra-rater reliability by performing a systematic review, incorporating a comprehensive search strategy and critical appraisal of reporting quality and potential bias. I also assess the proportion of cases assigned to each subtype to ensure that the included studies offer an appropriately representative sample.

2.2 Methods

2.2.1 Search strategy

I searched Ovid Medline (1946 - October 2014) and Embase (1980 - October 2014) for studies that assessed the inter- and/or intra-rater reliability of ICH classification systems in human adults, using a combination of search terms for haemorrhagic stroke, classification systems, and reliability (Appendix 1). I included conference abstracts and foreign language articles (obtaining translations where necessary), checked the bibliographies of all relevant studies and reviews identified, searched *Google Scholar* for studies citing relevant studies, and contacted researchers in the field for information about ongoing and unpublished studies. I assessed eligibility by reviewing all titles, abstracts and, where necessary, the full texts of potentially relevant articles. I resolved uncertainties through discussion and mutual consensus with a second, more senior researcher (Professor Cathie Sudlow).

2.2.2 Inclusion/exclusion criteria and contact with authors

I included all studies identified by October 2014, which reported inter- and/or intra-rater reliability of any anatomical or mechanistic ICH classification system. I also included studies identified by July 2013, which had not reported reliability, but had used an anatomical or mechanistic ICH classification system and reported the proportion of cases classified into mutually exclusive categories.

I excluded studies that used a classification system based on features other than anatomical or mechanistic (e.g. based on ICH severity or prognosis). I also excluded

studies conducted in highly selected patient populations (e.g. ICH limited to one anatomical territory), among cases with selected clinical features (e.g. study including only ICH cases with an epileptic seizure at onset) and studies that had sub-classified one mechanistic subgroup of ICH cases only (i.e. sub-classification of cases with ICH due to vascular malformations).

I contacted the authors of all included studies to obtain additional unpublished information about study characteristics that may affect reliability. To minimize publication and reporting bias (whereby positive results are more likely to be published or reported than negative ones), I also invited authors of studies which had used an anatomical or mechanistic classification system and reported the proportion of cases classified into mutually exclusive categories, but not reported reliability, to provide any unpublished reliability data. I created a structured data extraction form for all studies and pre-filled it as far as possible by entering data from the original publication (Appendix 2). I then asked the authors to complete the missing fields and where possible to provide additional unpublished reliability data or alternatively information that would enable me to calculate the reliability of the classification system myself.

2.2.3 Data extraction

I divided the included studies into those that had used either an anatomical or a mechanistic classification system.

From each study, I extracted data on:

- study characteristics: first author, publication year, study inclusion/exclusion criteria, country where study population recruited from, whether it was a population- or a hospital-based study;
- study population characteristics: number, mean age, ethnicity, gender, which investigations they had performed to enable classification;
- classification system used: categories, their definitions and classification rules;
- raters classifying the cases: number and qualifications;

- number of cases classified into each category;

For studies reporting on system reliability I also extracted data on:

- methods of reliability assessment: timing of classification in relation to symptom onset; raters' institutions; blinding of raters to being assessed; blinding of raters to other raters' and their own previous decisions; blinding of raters to patient's clinical history (for anatomical systems only) and outcome; data and data format presented to the raters;
- results of reliability assessment: kappa statistic with 95% confidence intervals (CI) and/or any other measure available.

2.2.4 Inter- and intra-rater reliability

I extracted data about any reliability measure reported for an anatomical or a mechanistic classification system. I displayed the most commonly reported inter- and intra-rater reliability measure (kappa [κ] statistics and 95% CI) of the more commonly reported classifications on a forest plot for visual inspection:

- lobar versus other location ICH;
- lobar versus deep versus cerebellar versus brainstem ± multiple location ICH;
- SMASH-U (Structural lesion, Medication, Amyloid angiopathy, Systemic/other disease, Hypertension, Undetermined) (Meretoja et al. 2012) and its modifications.

I considered kappa to indicate slight- (κ= 0.01–0.20), fair- (κ =0.21-0.40), moderate- (κ=0.41–0.60), substantial- (κ=0.61–0.80) or almost perfect reliability (κ=0.81–0.99) (Viera and Garrett 2005).

I assessed the quality of reliability reporting of included studies, using criteria based on the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) (Kottner et al. 2011). I assessed the potential risk of bias, using criteria developed specifically for this study, relating to study design features that I considered might influence the reliability results, such as:

- study investigators' independence of classification system development;
- raters conducting the rating independently or together;
- rater blinding to each other's classification / their own previous classification;
- raters' institutions same or different;
- single rater decisions or consensus opinions compared;
- rater blinding to being assessed;
- rater blinding to clinical information for anatomical systems and to patient outcome for mechanistic systems.

I explored factors that may affect the kappa statistic, that were available from all studies, such as the number of categories in the classification system and the number of raters.

2.2.5 Meta-analyses of ICH subtype proportions

I performed random-effects proportion meta-analyses for the more commonly used categories for anatomical classifications of ICH:

- supratentorial versus infratentorial location;
- lobar versus any other location;
- intraventricular haemorrhage (IVH) present versus absent;

and mechanistic classifications – ICH attributed versus not to:

- hypertension;
- CAA;
- anticoagulant use;
- vascular structural cause;
- undetermined cause.

I chose to use the DerSimonian and Laird random-effects method because of the substantial variability between the characteristics and settings of the individual studies contributing data, leading to the assumption that there is significant

heterogeneity and hence the true proportions may be different in the included studies (Borenstein et al. 2010). I assessed heterogeneity with I^2 statistics, interpreting the I^2 as follows (The Cochrane Collaboration 2011):

- I^2 = 0% to 40%: might not be important;
- I^2 = 30% to 60%: may represent moderate heterogeneity;
- I^2 = 50% to 90%: may represent substantial heterogeneity;
- I^2 = 75% to 100%: considerable heterogeneity;

I performed an additional sensitivity analysis of the proportions of ICH subtypes, including only those studies that also provided data for reliability, to ensure that results were consistent for this group of studies with the more inclusive analysis.

I also performed subgroup analyses by study mean age (≥70 versus <70 years, cut-off chosen based on spread of ages in included studies), country (Europe/USA/Australia versus Asia) and hospital- versus population-based study design. I chose to perform these subgroup analyses because data were available for these study characteristics from most included studies. I used the Q-test for heterogeneity, with a p-value of 0.1 indicating significant between-group difference, since heterogeneity tests are typically applied conservatively for meta-analysis (Petitti 2001). I performed all analyses with StatsDirect (http://www.statsdirect.com/).

2.3 Results

2.3.1 Included studies

From 2152 publications screened, I identified 20 eligible studies, 16 through electronic database searches (Anderson et al. 1994, Barton and Hemphill 2007, Chiewvit et al. 2009, Díaz-Guzmán et al. 2009, Labovitz et al. 2001, Lee et al. 2004, Lovelock et al. 2009, Mayda-Domaç et al. 2010, Meretoja et al. 2012, Nilsson et al.

2002, Rost et al. 2008, Zahuranec et al. 2012, Ziai et al. 2011, Takahashi et al. 2006, Wijman et al. 2010, Yeh et al. 2014) and four through other methods (hand search, bibliography screening, personal contact with researchers) (Bhattathiri et al. 2003, Charidimou et al. 2015b, Palm et al. 2013, Wermer et al. 2002) (Figure 2.1). One study was available in the form of an abstract only (Ziai et al. 2011) and one study was published in Spanish (Díaz-Guzmán et al. 2009).

2.3.2 Reliability of classification systems

Six studies provided data about the reliability of an anatomical classification system (Bhattathiri et al. 2003, Charidimou et al. 2015b, Chiewvit et al. 2009, Palm et al. 2013, Ziai et al. 2011, Wermer et al. 2002) and three about the reliability of a mechanistic classification system (Meretoja et al. 2012, Palm et al. 2013, Yeh et al. 2014) (Table 2.1, Figure 2.2). Authors from six studies provided additional unpublished information about classification system reliability and/or study characteristics (Charidimou et al. 2015b, Chiewvit et al. 2009, Meretoja et al. 2012, Palm et al. 2013, Wermer et al. 2002, Yeh et al. 2014).

2.3.2.1 Study characteristics

Most studies were hospital- rather than population-based (5/6 anatomical and 2/3 mechanistic system studies). The median number of cases was 76 for anatomical and 142 for mechanistic studies. Mean age of cases ranged from 57-75 years for anatomical and 61-71 years for mechanistic studies. The proportion of males ranged from 48-66%. Studies were conducted in Europe (4/6 anatomical, 2/3 mechanistic), Asia (1/6 anatomical, 1/3 mechanistic) and the USA (1/6 anatomical). Classifications were performed retrospectively in most studies (4/6 anatomical and 3/3 mechanistic). Where intra-rater reliability was assessed, the time interval between the two ratings ranged from 2-6 months in anatomical and 15 months to two years in mechanistic studies. There were 2-6 raters for anatomical and 2-3 for mechanistic studies. Raters in the anatomical studies included neuroradiologists, neurologists (some with special experience in stroke) and neurosurgeons; in the mechanistic studies they were all stroke neurologists. In 5/6 anatomical studies, CT scans were available to each rater, while this was not reported for one study. Access to medical records and imaging

reports was available to the raters in all mechanistic studies. Completeness of investigation varied across mechanistic studies, but in all studies most cases had had a CT brain scan, 15-20% of cases an MRI brain scan and 25-30% intracranial blood vessel imaging (CT angiography, MR angiography or digital subtraction angiography) (Table 2.2).

2.3.2.2 Quality of reporting and measures used to reduce bias

No study satisfied all the GRRAS criteria (Kottner et al. 2011) (Table 2.3). More than half of the studies lacked details in the original publication about the rater population of interest, rationale for the study, sampling method, measurement / rating process, statistical analysis and software used and about raters' and subjects' sample characteristics.

No study had used all possible measures to reduce potential bias when assessing reliability (Table 2.4). However reassuringly all studies but one, where the information was missing as it was only available as an abstract, reported that ratings were conducted independently, were single rater decisions and that the raters were blinded to their own previous decisions and to the decisions of other raters. Also, in the majority of the studies, the raters were blinded to clinical information that may have introduced bias. Common potential sources of bias were that the raters came from the same institution and that they were not blinded to the fact that they were being compared to other raters. The latter is important in order to remove the possibility of a Hawthorne effect i.e. ensuring that the rater's behaviour is not altered because of an awareness being observed (Kottner et al. 2011).

2.3.2.3 Inter-rater and intra-rater reliability

Anatomical classification systems

Six studies provided data about inter-rater reliability (Bhattathiri et al. 2003, Charidimou et al. 2015b, Chiewvit et al. 2009, Palm et al. 2013, Ziai et al. 2011, Wermer et al. 2002), three of these also had data about intra-rater reliability (Bhattathiri et al. 2003, Charidimou et al. 2015b, Wermer et al. 2002) (Table 2.1, Figure 2.2). This included unpublished reliability data from two studies (Charidimou

et al. 2015b, Palm et al. 2013). All studies had used the unweighted kappa statistic to measure reliability. Inter-rater reliability was substantial to almost perfect for classifying ICH as lobar versus any other location in four studies, (κ = 0.78 to 0.97) (Bhattathiri et al. 2003, Palm et al. 2013, Ziai et al. 2011, Wermer et al. 2002), and for classifying ICH into 4-5 categories (lobar, deep, cerebellar, brainstem ± multiple location categories) in two studies (κ = 0.81 to 0.87) (Charidimou et al. 2015b, Chiewvit et al. 2009). Intra-rater reliability for lobar versus any other ICH was almost perfect in two studies (κ = 0.85 to 1) (Bhattathiri et al. 2003, Wermer et al. 2002), and substantial in one study of ICH classified into four categories (lobar, deep, cerebellar, brainstem) (κ = 0.8) (Charidimou et al. 2015b).

Mechanistic classification systems

Three studies provided data about inter-rater reliability (Meretoja et al. 2012, Palm et al. 2013, Yeh et al. 2014) and two of these (Meretoja et al. 2012, Yeh et al. 2014) about intra-rater reliability of a mechanistic classification system (Table 2.1, Figure 2.2). The systems assessed were SMASH-U (Meretoja et al. 2012) and modifications of this. All studies provided additional unpublished data about study characteristics and reliability. All studies had used the unweighted kappa statistic to measure reliability. One study also reported the absolute agreement (Meretoja et al. 2012). Inter-rater reliability was almost perfect (κ = 0.89 to 0.93; absolute agreement 92/100), as was intra-rater reliability (κ = 0.92 to 0.93).

2.3.2.4 Factors affecting reliability

Although reliability generally appeared somewhat lower with a larger number of raters or categories, there were insufficient data to draw reliable conclusions about these or other factors potentially affecting reliability (Figure 2.2).

2.3.3 Proportions of ICH subtypes

These meta-analyses included 14 studies that used an anatomical classification system (Anderson et al. 1994, Charidimou et al. 2015b, Chiewvit et al. 2009, Díaz-Guzmán et al. 2009, Labovitz et al. 2001, Lee et al. 2004, Lovelock et al. 2009,

Mayda-Domaç et al. 2010, Nilsson et al. 2002, Palm et al. 2013, Rost et al. 2008, Zahuranec et al. 2012, Takahashi et al. 2006, Wijman et al. 2010) and seven that used a mechanistic system (Barton and Hemphill 2007, Chiewvit et al. 2009, Díaz-Guzmán et al. 2009, Meretoja et al. 2012, Palm et al. 2013, Wijman et al. 2010, Yeh et al. 2014). The number of studies included in different parts of meta-analyses varied, depending on availability of required data.

2.3.3.1 Study characteristics

Study characteristics were incompletely reported and inclusion criteria varied (Table 2.5, Table 2.6). 7/14 of the anatomical and 5/7 of the mechanistic classification studies were hospital-based, the remainder being population-based. There was a median of 115 cases per study for anatomical and 142 for mechanistic classifications. The mean age in each study ranged from 57-76 years for anatomical and 57-71 years for mechanistic classifications. 48-63% of cases were male. Studies recruited cases from Asia (4/14 anatomical, 2/7 mechanistic), Europe (5/14 anatomical, 3/7 mechanistic), the USA (4/14 anatomical, 2/7 mechanistic) and Australia (1/14 anatomical). The number of raters per study was 1-3 for anatomical and 1-2 for mechanistic studies. Rater expertise ranged from trainee to specialist neurologist or radiologist. To classify ICHs, 11/14 anatomical classification studies used CT brain scans, 3/14 used MRI brain scans and one also included autopsy data. Of the mechanistic classification studies, 5/7 provided raters with access to medical records and results of CT, MRI, and/or vascular imaging, while information about the investigations informing classification was unavailable for 2/7.

2.3.3.2 Classification system categories

Terminology varied across anatomical classification studies (e.g., the term “deep ICH” could include or exclude infratentorial ICH, or be replaced by “central ICH”; lobar ICH could include or exclude cerebellar cortical areas) (Table 2.7). Studies reported variable levels of detail about the rules used to define ICH location. 10/14 studies reported classifying based on “location of ICH”, 2/14 reported classifying based on “origin of ICH” and criteria were not clear for 2/14. Only two studies reported a category for multiple ICHs (Chiewvit et al. 2009, Wijman et al. 2010) and

only one (Anderson et al. 1994) for uncertain location. Clear definitions of classification categories were not available for all studies. Furthermore, except for studies using the SMASH-U system or its modifications, it was not clear how studies dealt with cases with more than one possible mechanistic cause (Table 2.7).

2.3.3.3 Results of proportion meta-analysis

Anatomical classification systems

The pooled proportion of ICH cases classified as supratentorial among ten studies was 0.88 (95% CI 0.86 to 0.89), with no detectable heterogeneity between studies (I^2=0%). The pooled proportion classified as lobar among 13 studies was 0.35 (95% CI 0.29 to 0.41), with substantial heterogeneity between studies (I^2=88%). Finally, the pooled proportion with IVH present among three studies was 0.47 (95% CI 0.41 to 0.52), with moderate heterogeneity between studies (I^2=50%) (Figure 2.3).

The pooled proportion estimates were very similar when only those studies contributing reliability data were analysed separately. The pooled proportion of ICH cases classified as supratentorial among two studies was 0.90 (95% CI 0.84 to 0.94), pooled proportion classified as lobar among three studies was 0.32 (95% CI 0.24 to 0.41) and the proportion with IVH present in one study was 0.42 (95% CI 0.34 to 0.50). There was substantial heterogeneity between studies in the lobar ICH category (I^2=60%). Heterogeneity could not be assessed in other categories, because there were too few studies.

Subgroup analyses showed that the proportion of cases classified as lobar was significantly higher in studies based in Europe, USA and Australia (0.41, 95% CI 0.36 to 0.45), compared to studies from Asia (0.24, 95% CI 0.21 to 0.27) ($p<0.0001$) and in population-based (0.41, 95% CI 0.33 to 0.49), compared to hospital-based studies (0.31, 95% CI 0.23 to 0.39) ($p=0.09$) (Table 2.8, Figure 2.4). There were no other significant subgroup differences (Table 2.8, Figure 2.5, Figure 2.6). Subgroup differences remained significant for lobar ICH and non-significant for supratentorial ICH when the Mayda-Domaç study conducted in Turkey was excluded from the Asian subgroup.

Mechanistic classification systems

Among seven studies, the pooled proportion of ICH cases classified as hypertensive was 0.51 (95% CI 0.40 to 0.61) with considerable heterogeneity between the studies (I^2=97%). The pooled proportion of ICH due to a vascular structural cause across seven studies was 0.06 (95% CI 0.05 to 0.08) with substantial heterogeneity (I^2=65%). Among five studies, the pooled proportion classified as CAA-related ICH was 0.16 (95% CI 0.10 to 0.23), as ICH due to anticoagulant use was 0.08 (95% CI 0.02 to 0.16) and as ICH of undetermined cause was 0.13 (95% CI 0.07 to 0.21) with considerable heterogeneity (I^2=95%, 97% and 96% respectively) (Figure 2.7).

The pooled proportion estimates were very similar when only those studies contributing reliability data were analysed separately. The pooled proportion of ICH cases classified as hypertensive among three studies was 0.47 (95% CI 0.32 to 0.62), as CAA-related was 0.20 (95% CI 0.12 to 0.29), as undetermined was 0.11 (95% CI 0.04 to 0.20), as due anticoagulant use was 0.09 (95% CI 0.02 to 0.20) and as due to vascular structural cause was 0.06 (95% CI 0.03 to 0.08). There was considerable heterogeneity between the studies in all categories (I^2= 87%-98.7%).

Subgroup analyses showed that studies with a higher mean age of cases (≥70 years), compared to a lower mean age of cases (<70 years), attributed a higher proportion to CAA (0.31, 95% CI 0.24 to 0.39 versus 0.13, 95% CI 0.08 to 0.19; p=0.0007), and a lower proportion to an undetermined cause (0.04, 95% CI 0.01 to 0.08 versus 0.16, 95% CI 0.09 to 0.25; p=0.06) (Table 2.8, Figure 2.8, Figure 2.9). Studies recruiting cases from Asia, compared to studies recruiting from Europe/USA/Australia, attributed a higher proportion to a vascular structural cause (0.08, 95% CI 0.07 to 0.09 versus 0.06, 95% CI 0.04 to 0.07; p=0.05) and a lower proportion to anticoagulant use (0.03, 95% CI 0.02 to 0.05 versus 0.14, 95% CI 0.12 to 0.16; p<0.0001) (Table 2.8, Figure 2.10, Figure 2.11). Population-based (versus hospital-based) studies attributed a higher proportion of cases to CAA (0.31, 95% CI 0.24 to 0.39 versus 0.13, 95% CI 0.08 to 0.19; p=0.0007) and a lower proportion to undetermined cause (0.04, 95% CI 0.01 to 0.08 versus 0.16, 95% CI 0.09 to 0.25; p=0.006) (Table 2.8, Figure 2.8, Figure 2.9). There were no other significant subgroup differences (Table 2.8, Figure 2.12).

2.4 Discussion

2.4.1 Main findings

My review found that both the inter- and intra-rater reliabilities of existing anatomical and mechanistic classification systems appear to be substantial to almost perfect. Although there was a suggestion that smaller numbers of raters or categories improved reliability, as would be expected (Sim and Wright 2005), there were insufficient data to formally assess factors that could influence reliability. The quality of reliability reporting was variable with no study completely following the GRRAS guidelines (Kottner et al. 2011), probably because measuring reliability was not the primary aim for all studies and also because these guidelines were published only recently. Furthermore, since no study had used all possible measures to reduce potential bias, reliability may have been over-estimated. Finally, all raters in these studies were experts in their field, limiting the generalisability of the results to less expert raters who might usefully contribute to large-scale research studies.

The proportions of ICH subtypes in anatomical categories were largely consistent with previous reports (Hanley 2009, Warlow et al. 2008), suggesting that the included studies are appropriately representative. Subgroup analyses suggested that the considerable heterogeneity between proportion of lobar ICH cases across studies was partly explained by study design and country. Substantial to considerable heterogeneity between proportions of mechanistic subtypes in included studies was partly explained by variation in mean age of cases, country and study design. I could not formally assess other potential contributing factors (including differences between diagnostic criteria and completeness of investigations). Contrary to previous findings (Shen et al. 2007), my meta-analyses found that ICH attributed to anticoagulant use was less frequent in Asian studies.

Existing classification systems have some limitations. For anatomical systems, these include:

- classification based on presumed site of ICH origin in some studies and on ICH location in others;

- lack of clarity and/or variation between studies in category definitions;
- surprisingly few classification systems with a separate category for bleeds in either multiple or uncertain locations.

For mechanistic systems, limitations include:

- assumptions about causal pathways (e.g., hypertension is commonly considered to be causally associated with deep ICH location and CAA with lobar ICH, despite doubts about the nature and/or strength of these associations) (Jackson and Sudlow 2006, Samarasekera et al. 2012);
- variable clarity of category definitions;
- the dependence on investigations undertaken to identify the potential cause (which vary considerably among specialties and countries, and with age, ICH location and blood pressure) (Cordonnier et al. 2010);
- varying definitions of primary and spontaneous ICH, with some studies excluding specific mechanistic categories;
- inability to assign a proportion of cases to a determined subtype;
- most cases do not have a single cause, but several interacting contributory factors (Warlow et al. 2008);
- classification systems do not always explain clearly how competing causes are dealt with.

2.4.2 Strengths and limitations

To my knowledge, no prior published systematic reviews of the reliability of ICH classification systems exist and thus this study is the first to summarise the total body of evidence. Other strengths of this study include a thorough search strategy and rigorous assessment of study characteristics, quality and bias indicators to inform the interpretation of the results. In addition, although I found relatively few relevant published studies, through contacting authors and establishing a collaborative group, I was able to include additional unpublished reliability results from four of the eight studies assessing reliability. Finally, this study has highlighted the limitations of

existing classification systems, which should help ensure their further refinement where needed and their appropriate use in diverse clinical and research settings.

This study has some limitations. I may have missed some publications where reliability was assessed but buried in a few words within the body of the text. I attempted to address this by manually searching through relevant review papers and reading full texts of all identified studies using an anatomical or mechanistic classification system. It is also important to emphasize that the pooled analyses of ICH subtype proportions was not the primary aim of this study, but a secondary analysis of eligible studies identified from a search which targeted studies of reliability. These analyses do not therefore include all published studies of the distribution of ICH subtypes. Furthermore, limited available data mean that conclusions about potential factors affecting the reliability and proportions of ICH subtypes are also inherently limited. In addition to bias, number of categories and raters, kappa can also be affected by disease prevalence, pattern of disagreement, time interval between ratings and other factors (Sim and Wright 2005). Finally, although reliability is an important feature of a classification system, it does not necessarily correlate with diagnostic accuracy or validity, which would require reference to a 'gold' standard.

2.4.3 Outstanding issues

While both anatomical and mechanistic systems appear to have excellent reliability, for large population-based, prospective epidemiological studies, anatomical classification systems are likely to be:

- **more feasible**, since less information from investigations is required for all cases to be classified;
- **more scalable**, since automated or semi-automated classification may be possible with a combination of International Classification of Diseases-coded hospital episodes and automated processing of brain imaging reports;
- **more appropriate** since they are free of assumptions about causal pathways.

The development of such methods for use at scale will require clear definitions, classification protocols, and the availability of categories for multiple or uncertain locations.

Mechanistic systems such as the SMASH-U and its modifications have the advantage of already having a very clear set of rules and excellent reliability (consistent with findings from ischemic stroke sub-classification studies showing that reliability increases with clear rules or protocols) (Woodfield et al. 2013). However, mechanistic systems could be further improved by integrating categories for cases with an uncertain and multiple overlapping mechanisms. Such systems are likely to be appropriate for:

- stratifying patients for clinical trials;
- some case-control studies, where stroke specialists collect the more detailed information required at the time of presentation with ICH;
- clinical practice to encourage a more systematic mechanistic work-up.

The feasibility of collecting this additional, more detailed information in large, prospective, population-based studies also needs further assessment, since it would complement the more basic information required for anatomical sub-classification and – potentially – allow nested case-control studies, based on not only anatomical, but also mechanistic information.

Reliability is unknown outside highly specialised centres with experienced readers. Future comparisons of reliability studies will be facilitated by studies following published GRRAS reporting guidelines. This should provide the detailed information necessary to understand how the study was designed and conducted, and how the results were obtained. Reliability is not a fixed property of the measurement tool but rather the product of interactions between the measurement tool, the objects, and the context of the assessment, affected by various sources of variability. Therefore study results are better interpretable when the study conditions are sufficiently described (Kottner et al. 2011).

2.4.4 Conclusions

To conclude, existing classification systems appear to have very good reliability in the settings in which they have been tested, but their reliability is unknown outside highly specialised centres with experienced readers. Future comparisons of reliability studies will be facilitated by studies following published GRRAS reporting guidelines (Kottner et al. 2011).

Tables

Table 2.1 Reliability of ICH classification systems

First author	Reported reliability measure	No. categories[1]	Inter-rater reliability			Intra-rater reliability		
			Kappa	Lower 95%	Upper 95%	Kappa	Lower 95%	Upper 95%
ANATOMICAL CLASSIFICATION SYSTEMS								
Chiewvit	Unweighted kappa	5	0.866	-	-	-	-	-
Wermer	Unweighted kappa	2	Mean across 3 ratings: 0.92 Rater I vs II: 0.88 Rater I vs III: 0.96 Rater II vs III: 0.92	- 0.75 0.88 0.81	- 1.00 1.00 1.00	1.00	-	-
Bhattathiri[2]	Unweighted subtype kappa	2 (thalamus or basal ganglia / other)	0.85	-	-	Mean kappa across six raters: 0.85 Range: 0.699 – 1[3]	-	-
	Unweighted subtype kappa	2 (lobar / other)	0.78			Mean kappa across six raters: 0.85 Range: 0.720 – 1[4]		
	Unweighted kappa	4 (frontal / temporal / parietal / occipital)	Range 0.61-0.8			-		
Palm	Unweighted Cohen´s kappa	2	0.94	0.88	1	-	-	-
Charidimou	Unweighted kappa	4	Combined across 3 ratings: 0.808 Rater I vs II: 0.791 Rater II vs III: 0.845 Rater I vs III: 0.791	0.672 0.753 0.787 0.779	0.892 0.853 0.874 0.808	0.804	0.76	0.971
Ziai	Kappa[5]	6 (lobar / caudate / globus pallidus / putamen / thalamus / IVH)	0.76	0.7	0.82	-	-	-
	Subtype kappa[5]	2 (subtype kappa for each location vs all other locations)	Caudate:0.56; Globus pallidus:0.03; Putamen:0.72; Thalamus:0.85; Lobar:0.97; Primary IVH:0.74	-	-	-	-	-
MECHANISTIC CLASSIFICATION SYSTEMS								
Meretoja	Unweighted Cohen´s kappa	6	0.89	0.82	0.96	0.93	0.89	0.97
Palm	Unweighted kappa	5	0.93	0.89	0.99	-	-	-
Yeh	Unweighted Cohen´s kappa	6	0.89	0.84	0.94	0.92	0.86	0.99

vs: versus; [1]If categories not defined in brackets then categories used as described in Table 2.7 (please see Table 2.7 for further information about category definitions); [2]Bhattathiri: it is not entirely clear from the article how the kappa has been calculated. The information in the table reflects my best guess; [3]Rater I: 0.950; Rater II: 0.898; Rater III: 0.779; Rater IV: 1; Rater V: 0.779; Rater VI: 0.699; [4] Rater I: 0.9; Rater II: 0.755; Rater III: 0.801; Rater IV: 1; Rater V: 0.720; Rater VI: 0.905; [5]Unclear if weighted or unweighted, assumption is that unweighted more likely. Meretoja: absolute agreement 92/100

Table 2.2 Characteristics of studies assessing reliability of ICH classification systems

First author Country Year	Classification system		Population							Raters						Data presented			
							Timing[4]			Number			Blinding[6]			Blinding[8]			
	Categories[1]	Clear rules Developed locally	No.[2] Ethnicity % male	Description	Age (mean ±SD)	Method[3]	Inter	Intra	Investigations performed	Inter	Intra	Expertise[5]	1	2	Institutions[7]	Hx	Outcome	Same[9]	Information available[10]
ANATOMICAL CLASSIFICATION SYSTEMS																			
Chiewvit Thailand 2009	·Lobar ·Thalamic-ganglionic ·Cerebellum ·Brainstem ·Multiple location	n/a n/a	84 Thai 63	Hospital case review	57±17	R	?	n/a	n/a	2	n/a	Experts: NRs	Y	N	Single	Y	Y	Y	Access to CT scans
Wermer Netherlands 2002	·Lobar ·Deep	n/a n/a	50/25 Caucasian 54	Selected (based on ICH volume) hospital case review	67	R	1-10 years	2 months	n/a	3	1	Experts: NRs	Y	N	Multiple	Y	Y	Y	Access to CT scans
Bhattathiri UK 2003	·Lobar ·BG/ thalamus ·Internal capsule	n/a n/a	43 ? ?	Selected (opportunity sample) trial case review	?	R	?	Min. 2 months	n/a	6	6	Experts: 4 NSs 2 NRs	Y	?	?	Y	Y	Y	Assume access to CT scans
Ziai USA 2011 *Abstract only*	·Lobar ·Caudate ·Globus pallidus ·Putamen ·Thalamus ·Primary IVH	n/a n/a	145 ? ?	Presumed hospital-based	?	?	?	n/a	n/a	3	n/a	Experts: NE NR CIL	?	?	?	?	?	?	?

First author Country Year	Classification system			Population							Raters						Data presented			
								Timing[4]			Number			Blinding[6]			Blinding[8]			
	Categories[1]	Clear rules	Developed locally	No.[2] Ethnicity % male	Description	Age (mean ±SD)	Method[3]	Inter	Intra	Investigations performed	Inter	Intra	Expertise[5]	1	2	Institutions[7]	Hx	Outcome	Same[9]	Information available[10]
Palm Germany 2013	·Lobar ·Deep	n/a	n/a	127 Caucasian 50.7	Population based stroke registry	71.3± 13.5	P	Most within 24 hours	n/a	n/a	2	n/a	Experts: SNs	Y	N	Single	Y	Y	Y	Access to CT scans
Charidimou UK 2015	·Lobar ·Deep ·Cerebellar ·Brainstem	n/a	n/a	69 92% Caucasian 48	Hospital-based, selected subset of cases	74.9± 12.3	R	>1 month	~6 months	n/a	3	1	Experts: RF NR SN	Y	N	Multiple	Y	Y	Y	Access to CT scans
MECHANISTIC CLASSIFICATION SYSTEMS																				
Meretoja Finland 2012	·Structural lesion ·Systemic/ other disease ·Medication ·CAA ·Hypertension ·Undetermined	Y	Y	100 Caucasian 58	Random selection of hospital cases	66±13	R	Several years	2 years	100% CT and/or MRI, 25% CTA,MRA or DSA	3	3	Experts: SNs	Y	N	Single	n/a	Y	Y	Full medical records, neurology opinion, imaging, all tests
Palm Germany 2013	·Structural vascular pathology ·OAK -related ·CAA ·Hypertension ·Undetermined	Y	Y[11]	142 Caucasian 51	Population based stroke registry	71.3± 13.5	R	6 months – 5 years	n/a	100% CT, 20% MRI, 26% CTA, MRA and/or DSA	2	n/a	Experts: SNs	Y	N	Single	n/a	?	Y	Anatomical classification, neuroimaging, age, medications on admission, INR, cardiovascular risk factors

First author Country Year	Classification system			Population							Raters						Data presented			
							Timing[4]				Number			Blinding[6]			Blinding[8]			
	Categories[1]	Clear rules	Developed locally	No.[2] Ethnicity % male	Description	Age (mean ±SD)	Method[3]	Inter	Intra	Investigations performed	Inter	Intra	Expertise[5]	1	2	Institutions[7]	Hx	Outcome	Same[9]	Information available[10]
Yeh Taiwan 2014	·Structural lesion ·Systemic/ other disease ·Medication ·CAA ·Hypertension ·Undetermined	Y	Y[11]	268 Chinese 66	Cases from the National Taiwan University Hospital Stroke Registry database	60.9 ± 16.0	R	7-19 months	15 months	100% CT, 15% MRI, 30% CTA, MRA and/or DSA	2	1	Experts: SNs	Y	N	Single	n/a	Y	Y	Past medical hx and medication hx, imaging results, clinical information about current admission, blood tests

?: unknown; Y: yes; N: no; NR: neuroradiologist; NS: neurosurgeon; NE: neurologist; SN: neurologist specialised in stroke; RF: research fellow; CIL: central imaging laboratory for a multicentre trial; DSA: digital subtraction angiography; [1]Definitions of classification system categories can be found in Table 2.7.

[2]Number of ICH cases classified for reliability assessment; [3]Method: R (retrospective): cases were classified after initial presentation, usually through retrieving and reviewing medical records; P (prospective): patients were classified at the time of or shortly after presenting and being recruited to the study; [4]Timing: time from symptom onset to classification for inter-rater reliability; time between two ratings for intra-rater reliability; [5]Expertise: Predefined categories were: Expert = qualified doctors at specialist level (trainee, research fellow or consultant); Less expert = physicians 1-4 years post registration, general practitioners, nurses or medical students (NB: no studies had less expert raters); [6]Blinding: 1: rater unaware of other raters' decisions and/or their own previous decision; 2: rater unaware of being compared to other raters; [7]Institutions: raters from single or multiple institutions; [8]Blinding: raters blind to each patient's clinical history (Hx) and/or outcome; [9]Same: same information available to each rater; [10]Information available: information available to the rater for classification purposes; [11]Classification system modified from SMASH-U locally;

Comments about specific studies: Chiewvit: mean age applies to a larger sample of 131 cases (including cases with subarachnoid haemorrhage, subdural haemorrhage and intraventricular haemorrhage [IVH]); Wermer: included only supratentorial cases; 50 cases rated by 3 raters to assess inter-rater agreement, 25/50 cases rated twice by one rater to assess intra-rater agreement, not specified how the 25 cases were chosen from amongst the 50 for intra-rater agreement; Bhattathiri: included only supratentorial ICH cases; classification assumed retrospective, though not specifically mentioned in the publication; Ziai: not stated

explicitly in the abstract but assumed to include only supratentorial ICH cases; not stated explicitly in the abstract but "CIL" assumed to use expert rater(s). Palm: % male, mean age and investigations performed applies to a larger sample of 152 non-traumatic ICH cases; Meretoja: each case classified by 2 of 3 raters, with 2 raters classifying 50 cases each and one rater classifying 100 cases; Investigations performed apply to larger sample of 1013 ICH cases; 100 cases for reliability selected at random from the whole sample of 1013 patients, but weighted to ensure appropriate representation of all SMASH-U classifications (the raters did not know of the weighting or the relative frequency of the SMASH-U classes at time of evaluation).

Table 2.3 Reporting quality of ICH classification system reliability studies based on the GRRAS guidelines

Guidelines for reporting reliability and agreement studies (GRRAS)		Chiewvit	Wermer[1]	Bhattathiri[2]	Meretoja[3]	Yeh
TITLE/ABSTRACT						
1. Identify in title or abstract that inter-rater/intra-rater reliability or agreement was investigated		●	●	●	●	○
INTRODUCTION						
2. Name and describe the diagnostic or measurement device of interest explicitly		○	●	●	●	●
3. Specify the subject population of interest		●	●	●	●	●
4. Specify the rater population of interest		○	●	●	○	⚬
5. Describe what is already known about reliability and agreement and provide a rationale for the study		○	⚬	●	●	○
METHODS						
6. Explain how the sample size was chosen. State the determined number of raters, subjects/objects, and replicate observations		●	●	⚬	⚬	●
7. Describe the sampling method		●	⚬	⚬	○	●
8. Describe the measurement / rating process	Data presented to adjudicators	●	●	●	●	●
	Time since index ICH	○	○	○	○	○
	Blinding between observers	○	○	●	○	○
	Blinding to study of reliability	○	○	○	○	○
	Time interval between successive ratings	○	⚬	⚬	○	○
	Single or consensus ratings	●	●	●	●	●

9. State whether the measurements/ratings were conducted independently (at the same time or not)		○	●	●	●	○
10. Describe the statistical analysis and software used		⚬	⚬	⚬	⚬	●
RESULTS						
11. State the actual number of raters and subjects which were included and the number of replicate observations which were conducted		○	●	●	●	●
12. Describe raters' and subjects' sample characteristics	Subject population, general characteristics including age	⚬	⚬	○	⚬	⚬
	% Investigations performed in subject population (mechanistic systems)				⚬	⚬
	Rater expertise[4]	●	●	●	●	○
	Rater experience with the classification system (mechanistic systems)				○	○
13. Report estimates of reliability and agreement including measures of statistical uncertainty		○	●	⚬	●	●
DISCUSSION						
14. Discuss the practical relevance of the results		⚬	●	⚬	●	○
SUPPLEMENTARY MATERIAL						
15. Provide detailed results if possible (eg., online)		○	○	○	○	○

Assessed only based on information available in the original publication, for some studies further information became available upon direct contact with authors; ●Criteria fulfilled in the original publication ○Criteria not fulfilled in the original publication; ⚬Criteria partly fulfilled in the original publication; [1]Wermer: do not explain how the 25 scans that were assessed for intra-rater agreement were chosen from amongst the 50 scans originally assessed; [2]Bhattathiri: in "results" section state that measurements conducted independently; [3]Meretoja: rater expertise described in "methods" section; [4]"Rater expertise" includes rater training and experience with the classification system.

Table 2.4 Risk of bias assessment for ICH classification system reliability studies

Measures to reduce bias	Charidimou	Wermer	Palm – anatomical	Palm – mechanistic	Meretoja	Chiewvit	Ziai	Bhattathiri	Yeh
1. Were the study investigators independent of classification system development?	•	•	•	⚬	○	•	?	•	⚬
2. Was rating conducted independently?	•	•	•	•	•	•	?	•	•
3. Were raters blind to each other's classification/their own previous classifications?	•	•	•	•	•	•	?	•	•
4. Did raters come from multiple institutions (versus a single institution)?	•	•	○	○	○	○	?	?	○
5. Were answers compared single rater decisions (versus consensus opinions compared)?	•	•	•	•	•	•	•	•	•
6. Were the raters blinded to being assessed?	○	○	○	○	○	○	?	?	○
7. Were the raters blinded to clinical information (for mechanistic systems blinded to patient outcome)?	•	•	•	?	•	•	?	•	•

Ziai: published as an abstract only; ● Criteria fulfilled in the study; ○ Criteria not fulfilled in the study; ⭘ Criteria partly fulfilled in the study; ?: unknown.

Table 2.5 Characteristics of studies using an anatomical ICH classification system

Study first author Publication year Country[1]	Study population characteristics[2]										Imaging used for rating[3]		Rater(s)[4]
	P or H	No. of cases	Age (mean ± SD)	% male	Ethnicity	INCLUDED CASES					CT	MRI	
						HT of IS	Tu	An	Tr	Additional exclusions			
Anderson 1994 Australia	P	47	68	52	Majority Anglo-European-Australian	-	?	-	-		80%	-	Neurology trainee
Labovitz 2001 USA	P	150	63.7	?	Multi-ethnic	?	-	-	-	• AVM related ICH	?	?	Committee of study neurologists
Nilsson 2002 Sweden	P	341	Median 74	56	?	-	-	-	-	• AVM related ICH	✓	-	Study author or R
Lee 2004 South Korea	H	83	66±11	60	?	?	-	?	-	• Multiple ICHs • AVM/cavernous angioma-related ICH • Cases with known bleeding disorders • Cases with prior neurosurgery	-	✓	?
Takahashi 2006 Japan	H	347	72±12	49	?	?	?	-	-		✓	-	2 raters

Study first author Publication year Country[1]	Study population characteristics[2]										Imaging used for rating[3]		Rater(s)[4]
	P or H	No. of cases	Age (mean ± SD)	% male	Ethnicity	INCLUDED CASES					CT	MRI	
						HT of IS	Tu	An	Tr	Additional exclusions			
Rost 2008 USA	H	629	72±12	54	≥80% Caucasian	-	-	?	-	• Vascular malformation-related ICH • Preceding trauma or ischemic stroke • Vasculitis • Blood dyscrasia • INR >3 due to anticoagulation • No outcome data	✓	?	?
Chiewvit 2009 Thailand	H	82	57±17	63	Thai	✓	?	✓	-		✓	-	2 NRs
Díaz-Guzmán 2009 Spain	P	15	?	53	?	?	?	?	?		?	?	N and a study monitor
Lovelock 2009 UK	P	74	76	56	?	-	-	-	-		✓	?	?

Study first author Publication year Country[1]	Study population characteristics[2]										Imaging used for rating[3]		Rater(s)[4]
	P or H	No. of cases	Age (mean ± SD)	% male	Ethnicity	INCLUDED CASES					CT	MRI	
						HT of IS	Tu	An	Tr	Additional exclusions			
Mayda-Domaç 2010 Turkey	H	202	62.2	59	?	?	?	-	-	• AVM/vascular abnormality-related ICH • Anti-aggregant/anti-coagulated patients • Cases with chronic medical problems • Cases with missing outcome data or admitted >24 h of onset	Some	Some	?
Wijman 2010 USA	H	58	62	60	27% Asian 66% White 7% other	✓	✓	?	-	• Prior known cause for ICH (i.e. vascular malformation) •No MRI or performed >30 days post ICH	✓	?	One or all of: SN, N, study investigator
Zahuranec 2012 USA	P	171	72±13	48	57% Mexican American, 43% non-Hispanic white	-	-	?	-		✓	-	N
Palm 2013 Germany	P	148	71.3± 13.5	51	Caucasian	-	-	?	-	• Previous IS or ICH	✓	Some	SN or location from clinical notes
Charidimou 2015 UK	H	69	74.9± 12.3	48	92% white British	-	-	-	-	• Known underlying structural cause	✓	-	NR

[1]Country: country where study cases recruited from; [2]Study population characteristics: P: population-based study, H: hospital-based study. No. of cases: number of ICH cases classified and included in the proportion meta-analyses; [3]Imaging used for rating: results reflect what was felt most likely based on publications, but it was often difficult to be certain; [4]Rater(s): number and qualification(s) of person (people) classifying cases;?: unknown; HT of IS = haemorrhagic transformation of ischemic stroke; Tu = intracranial tumour-related ICH; An = aneurysmal ICH; Tr = traumatic ICH; NR = neuroradiologist; R = radiologist; SN = neurologist specialised in stroke; SS = stroke specialist; Med. Records = medical records; Path. Data = pathology data; Comments about specific studies: Anderson: mean age and % male applies to a larger sample of 60 case presentations (including the 13 cases of large haemorrhage where impossible to distinguish between deep and lobar regions as site of origin); Chiewvit: mean age applies to a larger sample of 131 cases and % male applies to a larger sample of 84 cases (including cases with subarachnoid haemorrhage, subdural haemorrhage and intraventricular haemorrhage [IVH]); Mayda-Domaç: mean age and % male applies to a larger sample of 208 cases (including 6 cases with IVH); Wijman: mean age, % male and ethnicity applies to a larger sample of 70 cases (including 7 cases with multiple ICHs and 5 cases with IVH); Palm: mean age and % male applies for a larger sample of 152 non-traumatic ICH cases; Lovelock: mean age and % male applies to a larger sample of 78 cases (including the 4 cases of haemorrhagic transformation of an infarct); Rost: ethnicity is based on overall Massachusetts General Hospital ICH cohort ethnicity estimates;

Table 2.6 Characteristics of studies using a mechanistic ICH classification system

Study first author Publication year Country[1]	Study population characteristics[2]										Definition of classification categories[5]				Investigations performed				Rater(s)[4]
						INCLUDED CASES													
	P or H	No. cases	Age (mean ± SD)	% male	Ethnicity	HT of IS	Tu	An	Tr	Additional exclusions	HT	VSC	C	CAA	CT	MRI	CTAMRADSA	Other	
Barton 2007 USA	H	237	65±15	54	44% Asian, 34% White, 22% other	?	✓	✓	?		-	-	n/a	-	?	?	?	?	?
Chiewvit 2009 Thailand	H	84	57±17	63	Thai	✓	?	✓	-		-	-	-	n/a	✓	?	?	Med. records	?
Díaz-Guzmán 2009 Spain	P	15	?	53	?	?	?	?	?		-	-	-	n/a	?	?	?	?	?
Wijman 2010 USA	H	70	62	60	27% Asian, 66% White, 7% other	✓	✓	?	-	• Prior known cause for ICH i.e. vascular malformation • MRI not performed or performed >30 days post ICH	✓	-	n/a	✓	✓	✓	Some	Med. records & path. data	?

Study first author Publication year Country[1]	Study population characteristics[2]										Definition of classification categories[5]				Investigations performed				Rater(s)[4]
						INCLUDED CASES													
	P or H	No. cases	Age (mean ± SD)	% male	Ethnicity	HT of IS	Tu	An	Tr	Additional exclusions	HT	VSC	C	CAA	CT	MRI	CTAMRADSA	Other	
Meretoja 2012 Finland	H	1013	67	57	96% Caucasian 4% other	-	-	-	-		✓	✓	✓	✓	All had Imaging, 15% had MRI		25%	Med. records	≥1 SN
Palm 2013 Germany	P	142	71±14	51	Caucasian	-	-	?	-	• Previous IS or ICH	✓	✓	✓	✓	✓	20%	26%	Med. records	2 SNs
Yeh 2014 Taiwan	H	3785	59±17	63	Majority Chinese	-	-	-	-	• Previous stroke	✓	✓	✓	✓	All had Imaging (CT and/ or MRI)		25%	Med. records	2 SSs

[1]Country: country where study cases recruited from; [2]Study population characteristics: P: population-based study, H: hospital-based study. No. of cases: number of ICH cases classified and included in the proportion meta-analyses; [3]Imaging used for rating: results reflect what was felt most likely based on publications, but it was often difficult to be certain; [4]Rater(s): number and qualification(s) of person (people) classifying cases; [5]Definition of classification categories: HT: hypertension-related ICH; VSC: vascular structural cause for ICH; C: anticoagulant use-related ICH; CAA: cerebral amyloid angiopathy-related ICH; ?: unknown; HT of IS = haemorrhagic transformation of ischemic stroke; Tu = intracranial tumour-related ICH; An = aneurysmal ICH; Tr = traumatic ICH; NR = neuroradiologist; R = radiologist; SN = neurologist specialised in stroke; SS = stroke specialist; Med. Records = medical records; Path. Data = pathology data; Comments about specific studies: Chiewvit: mean age applies to a larger sample of 131 cases and % male applies to a larger sample of 84 cases (including cases with subarachnoid haemorrhage, subdural haemorrhage and intraventricular haemorrhage [IVH]); Wijman: mean age, % male and ethnicity applies to a larger sample of 70 cases (including 7 cases with multiple ICHs and 5 cases with IVH); Palm: mean age and % male applies for a larger sample of 152 non-traumatic ICH cases;

Table 2.7 Anatomical and mechanistic classification system categories and their definitions

ANATOMICAL CLASSIFICATION SYSTEM CATEGORIES AND THEIR DEFINITIONS	
(I) Anatomical category assigned based on origin of ICH	
Wermer	Included only supratentorial ICH cases: 1) lobar (origin was in one of the cerebral lobes) 2) deep (origin was in the basal ganglia)
Anderson	1) lobar (frontal, parietal, occipital, temporal, corona radiata) 2) deep (caudate, putamen, globus pallidus, thalamus, internal capsule, deep periventricular white matter) 3) cerebellar 4) brainstem 5) massive cortical or subcortical (cases of large haemorrhage where impossible to distinguish between deep and lobar regions as site of origin or localize to a lobe)
Rost	1) lobar (origin of haemorrhage appeared to be in cerebral hemispheres of the frontal, parietal, temporal, or occipital lobes superficial to the deep grey matter structures) 2) deep (haemorrhages originating in the thalamus, basal ganglia) 3) infratentorial
(II) Anatomical category assigned based on location of ICH	
Bhattathiri[1]	Included only supratentorial ICH cases: 1) lobar (frontal, temporal, parietal or occipital lobes involved) 2) basal ganglia, thalamus or its components 3) internal capsule
Chiewvit[2]	1) lobar (frontal, parietal, temporal, occipital) 2) thalamic-ganglionic (caudate, putaminal and thalamic) 3) cerebellum 4) brainstem 5) multiple location
Ziai (abstract)	Assume included only supratentorial ICH cases: 1) lobar 2) globus pallidus 3) putamen 4) caudate 5) thalamus 6) primary IVH
Palm	1) lobar (cortical or subcortical) 2) deep (basal ganglia, infratentorial) Recorded presence versus absence of IVH
Labovitz	1) lobar (defined as ICH involving the frontal, parietal, temporal, or occipital lobes or insular cortex) 2) deep

Nilsson	1) lobar (cortical or subcortical) 2) central (basal ganglia, thalamus, internal capsule, deep periventricular white matter or purely intraventricular) 3) cerebellar 4) brainstem
Lee	1) corticosubcortical area 2) deep grey matter (basal ganglia and thalamus) 3) infratentorial area (brainstem and cerebellum)
Takahashi	1) lobar 2) central (basal ganglia and thalamus) 3) cerebellum 4) brainstem Recorded presence versus absence of IVH
Lovelock	1) lobar 2) deep/posterior (involved basal ganglia, thalamus, cerebellum or brainstem)
Mayda-Domaç	1) lobar 2) putaminal 3) nucleus caudatus 4) thalamic 5) cerebellar 6) pontine 7) intraventricular
Wijman	1) lobar 2) basal ganglia 3) thalamus 4) cerebellum 5) brainstem 6) isolated IVH 7) multiple ICHs
Zahuranec	1) supratentorial 2) infratentorial Recorded presence versus absence of IVH
(III) Not clear whether anatomical category assigned based on origin or location of ICH	
Charidimou	1) lobar (cortical-subcortical not extending into deep structures) 2) deep (basal ganglia, thalamus) 3) cerebellar 4) brainstem
Díaz-Guzmán	1) lobar 2) deep (basal ganglia) 3) cerebellar

MECHANISTIC CLASSIFICATION SYSTEMS CATEGORIES AND THEIR DEFINITIONS	
(I) Potentially competing causes clearly dealt with (The SMASH-U system and its modifications)	
Meretoja	SMASH-U system based on a flowchart principle: 1) structural lesion: imaging or pathology confirmed structural vascular malformation diagnosed at ICH site 2) systemic /other disease: systemic or other determined cause for ICH except for anticoagulation, hypertension or CAA (liver cirrhosis when known liver disease combined with spontaneously elevated INR or LFTs >3x upper limit; thrombocytopenia if PLT<50) 3) Medication: warfarin with INR >2, use of novel oral anticoagulants within 3 days, full dose heparin or non-ischemic stroke systemic thrombolysis 4) CAA: Boston criteria, lobar, cortical or subcortical haemorrhage and age ≥55 5) hypertension: deep or infratentorial ICH with pre-ICH hypertension (most recent pre-ICH BP ≥160/100 mmHg or if BP n/k, mention of diagnosis by patient or records with LVH, or on therapy. 6) undetermined
Palm	Modified SMASH-U system, also based on a flowchart principle: 1) structural lesion: AVM or cavernoma 2) medication: on warfarin with INR >2 3) CAA: older than 54 years with lobar haematoma 4) hypertension: deep ICH with pre-ICH hypertension diagnosis or on medication or BP >140/90 on 2 or more measurements > 3 days after stroke 5) undetermined
Yeh	Modified SMASH-U system based on a flowchart principle, with renal failure (defined as chronic kidney disease stage 5 or requiring regular dialysis) added under the "systemic/other disease" category
(II) Unclear how potentially competing causes dealt with	
Barton	Cause of haemorrhage divided into following categories (not specified based on which criteria the categories were determined): 1) CAA 2) coagulopathy 3) hypertension 4) illicit drugs 5) vascular (AVM, aneurysm) 6) other lesion (tumour, abscess) 7) other

Chiewvit	Based on risk factors classified into one of the following categories: 1) hypertension 2) aneurysm 3) AVM 4) anticoagulants/thrombolytic therapy 5) alcohol abuse 6) unknown
Díaz-Guzmán	1) hypertension 2) vascular malformation 3) anticoagulants 4) CAA 5) other No further definitions of categories provided.
Wijman	Aetiology of spontaneous ICH classified as follows: 1) hypertension (patient considered to have hypertension [HT] if 1 of the 3 criteria met: well documented history of HT; on treatment for HT as an outpatient; LVH on ECG or ECHO; study neurologists were also instructed to look for BMBs in deep grey nuclei or brainstem) 2) CAA (neurologists also instructed to look for BMBs in subcortical white matter in an elderly patient with lobar ICH) 3) vascular malformation 4) neoplasm 5) haemorrhagic transformation of ischemic stroke (neurologists instructed to consider if areas of restricted diffusion substantially exceed commonly seen perihaemorrhagic diffusion artefacts, or additional areas of diffusion restriction in someone with considerable risk factors for ischemic stroke) 6) illicit drug use 7) coagulopathy (PLT <50; INR >3; APTT >80s) 8) other identified cause 9) unknown cause

[1]Bhattathiri: separately recorded "site of primary involvement" and "site of secondary involvement";

[2]Chiewvit: location of ICH based on the location of the largest blood clot.

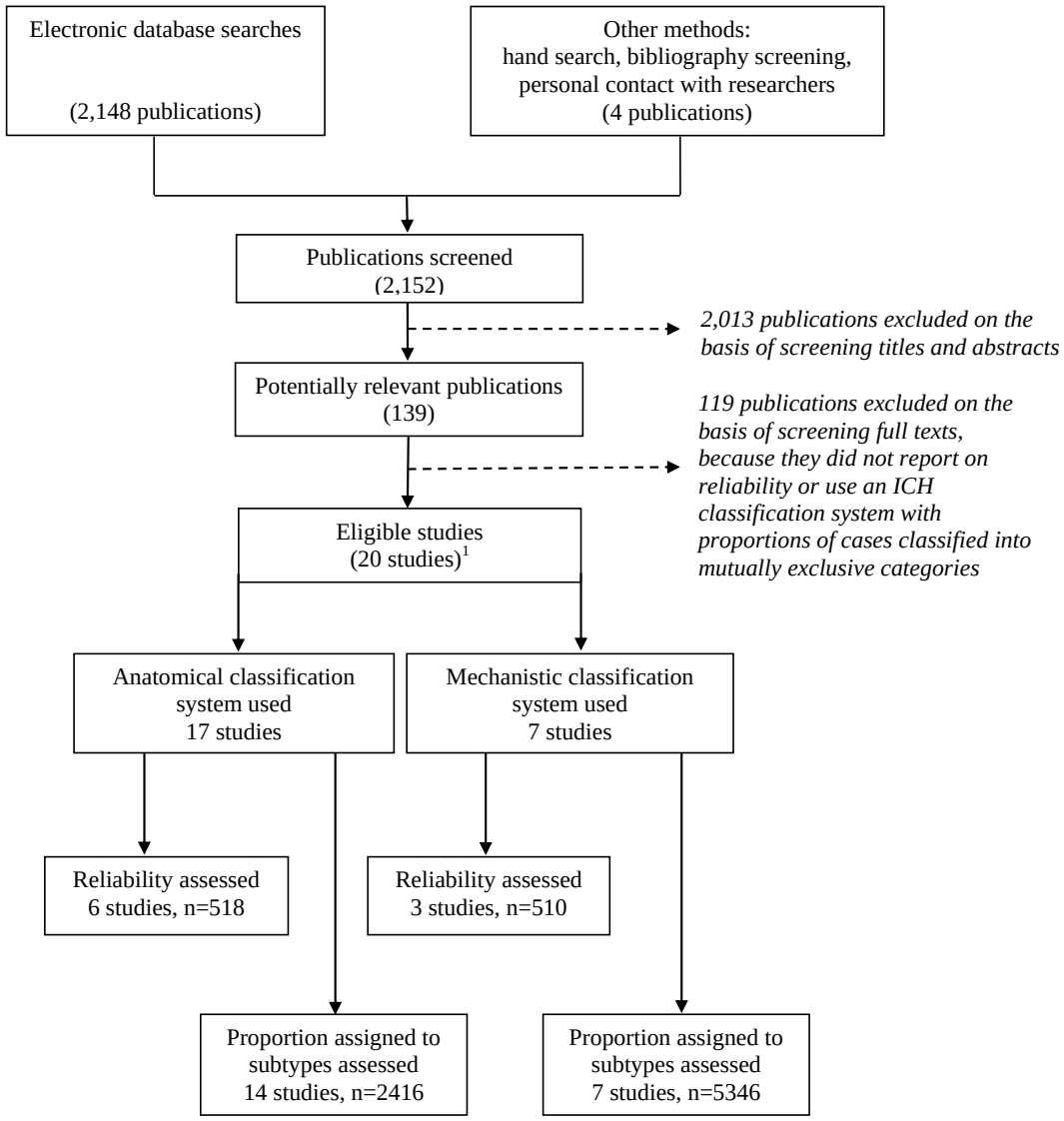
Table 2.8 Subgroup analyses of ICH subtype proportion meta-analysis

SUBGROUPS COMPARED		PROPORTION: ANATOMICAL CATEGORIES			PROPORTION: MECHANISTIC CATEGORIES				
		Supratentorial ICH	Lobar ICH	IVH present	Hypertension	CAA	Undetermined	Anticoagulant use	Vascular structural cause
Mean age of cases	≥70	0.88 (0.86-0.90)	0.35 (0.26-0.45)	n/a	0.51 (0.43-0.60)	**0.31 (0.24-0.39)**	**0.04 (0.01-0.08)**	0.11 (0.07-0.18)	0.03 (0.01-0.07)
	<70	0.86 (0.82-0.89)	0.31 (0.25-0.36)	n/a	0.51 (0.39-0.64)	**0.13 (0.08-0.19)**	**0.16 (0.09-0.25)**	0.07 (0.01-0.17)	0.07 (0.05-0.09)
	p-value	0.25	0.38	n/a	0.99	**<0.01**	**0.06**	0.52	0.14
Country cases recruited from	Eur/USA/ Australia	0.87 (0.84-0.90)	**0.41 (0.36-0.45)**	0.47 (0.37-0.58)	0.47 (0.30-0.65)	0.17 (0.08-0.28)	0.11 (0.02-0.26)	**0.14 (0.12-0.16)**	**0.06 (0.04-0.07)**
	Asia	0.88 (0.85-0.90)	**0.24 (0.21-0.27)**	0.46 (0.40-0.51)	0.55 (0.53-0.57)	0.12 (0.11-0.13)	0.16 (0.05-0.33)	**0.03 (0.02-0.05)**	**0.08 (0.07-0.09)**
	p-value	0.78	**<0.01**	0.77	0.45	0.29	0.64	**<0.01**	**0.05**
Study design	Hospital-based	0.88 (0.86-0.90)	**0.31 (0.23-0.39)**	0.46 (0.40-0.51)	0.52 (0.39-0.64)	**0.13 (0.08-0.19)**	**0.16 (0.09-0.25)**	0.07 (0.01-0.17)	0.07 (0.05-0.09)
	Population-based	0.84 (0.77-0.90)	**0.41 (0.33-0.49)**	0.47 (0.37-0.58)	0.50 (0.43-0.58)	**0.31 (0.24-0.39)**	**0.04 (0.01-0.08)**	0.11 (0.07-0.17)	0.04 (0.01-0.07)
	p-value	0.26	**0.09**	0.77	0.87	**<0.01**	**<0.01**	0.52	0.18

Numbers in the table represent pooled proportion (and 95% confidence interval) of cases in a category across all studies in a particular subgroup; p-value shows whether there is a significant difference between the pooled results of the two subgroups under comparison. Significant results ($p<0.1$) in bold; n/a: not applicable.

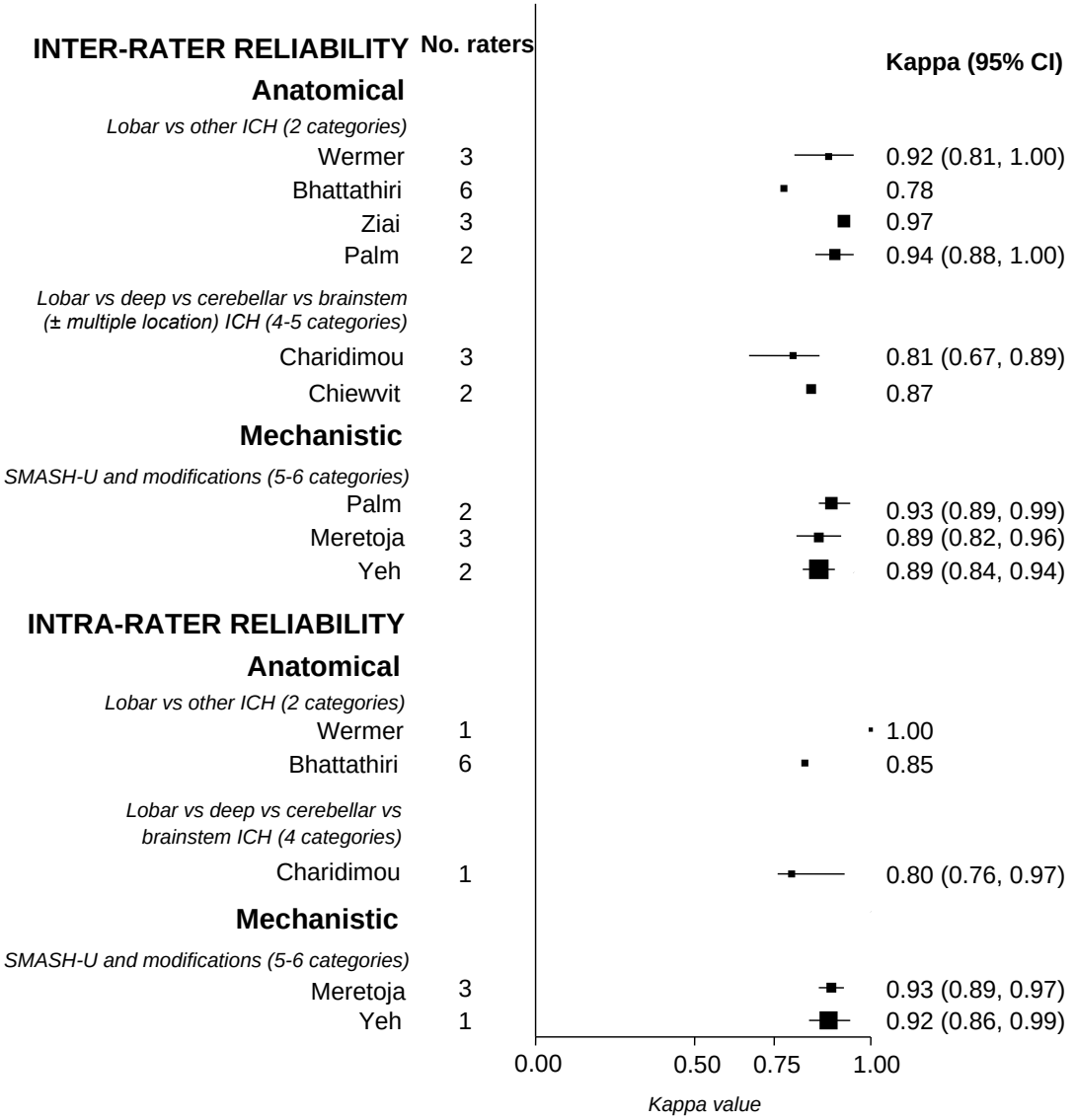
Figures

Figure 2.1 Selection of studies included in the systematic review of ICH classification system reliability



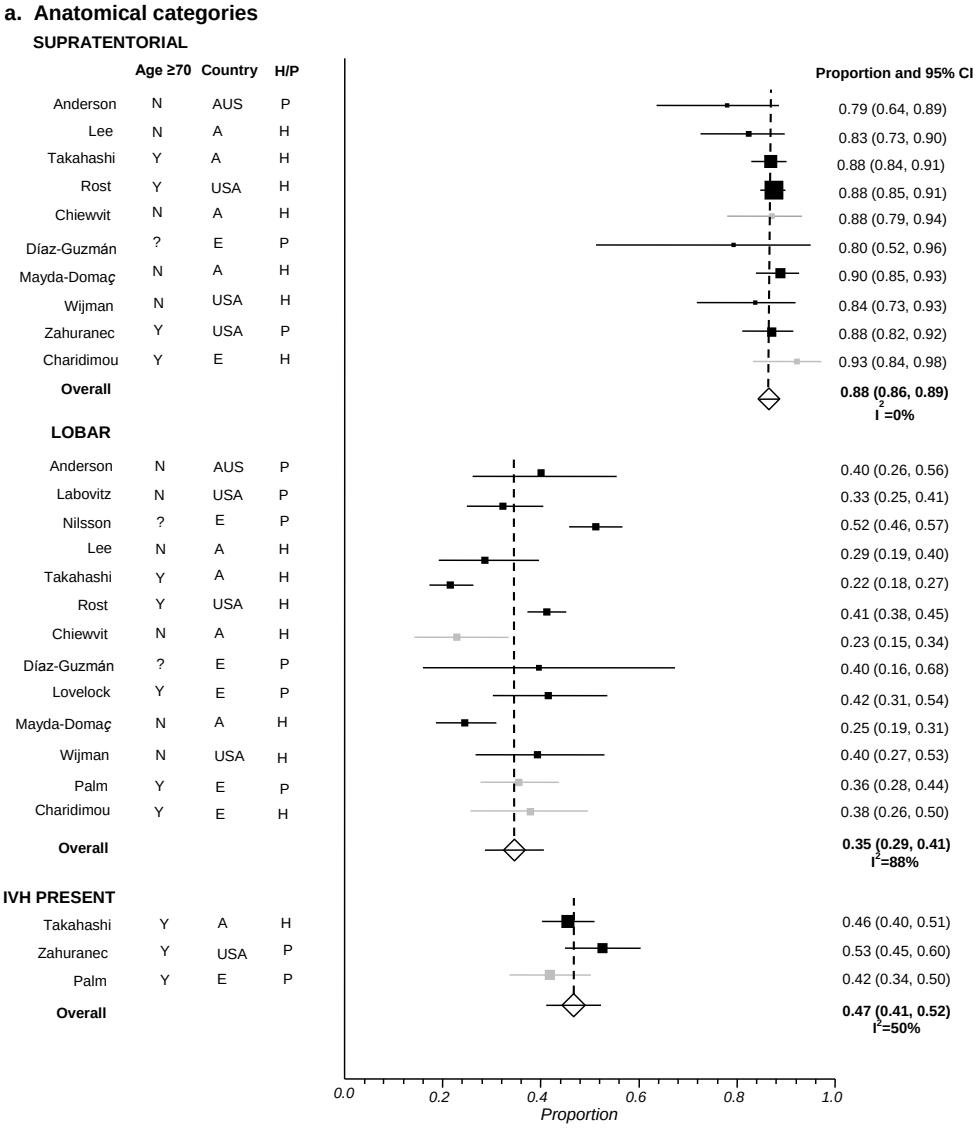
[1]Four studies had used both anatomical and mechanistic classification systems (Chiewvit et al. 2009, Díaz-Guzmán et al. 2009, Palm et al. 2013, Wijman et al. 2010); One study was in the form of an abstract only (Ziai et al. 2011); "n": overall number of ICH cases included in the studies.

Figure 2.2 Inter- and intra-rater reliability of ICH classification systems



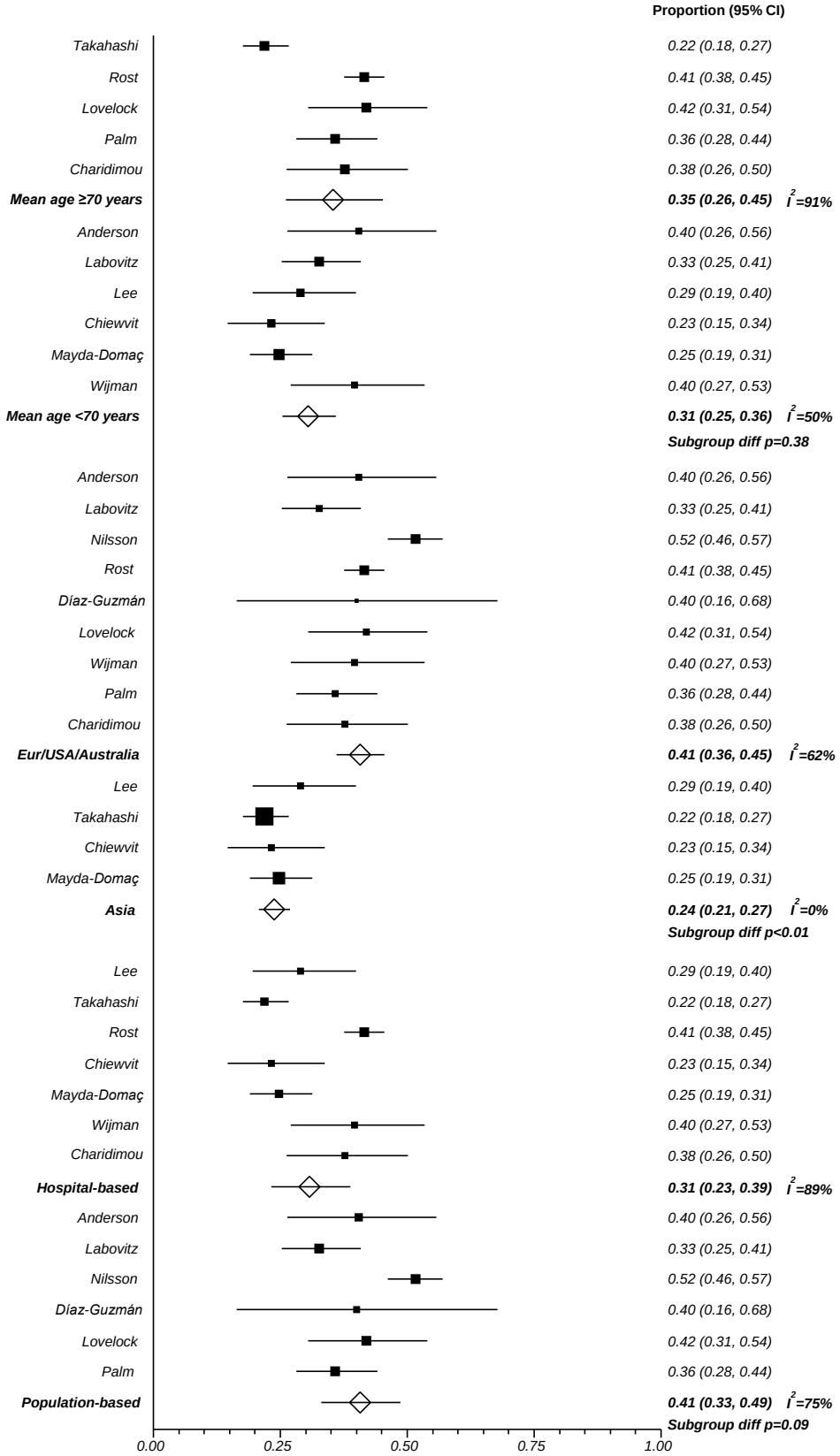
Squares represent individual study reliability estimates, and associated horizontal lines represent 95% confidence intervals; All kappas are unweighted; Wermer, Bhattathiri, Ziai: included only supratentorial ICH cases; Palm: included infratentorial as well as supratentorial ICH cases; Wermer: mean kappa based on 3 sets of ratings between 2 raters (different combination of raters from amongst 3 raters); Bhattathiri and Ziai: subtype kappa for lobar location versus any other location; Bhattathiri: Intra-rater reliability based on mean kappa for 6 raters all classifying each case twice; Meretoja: each case was classified by 2 raters, but there were 3 raters in total with 2 raters classifying 50 cases each, and one rater classifying 100 cases.

Figure 2.3 Proportions of ICH subtypes in anatomical categories



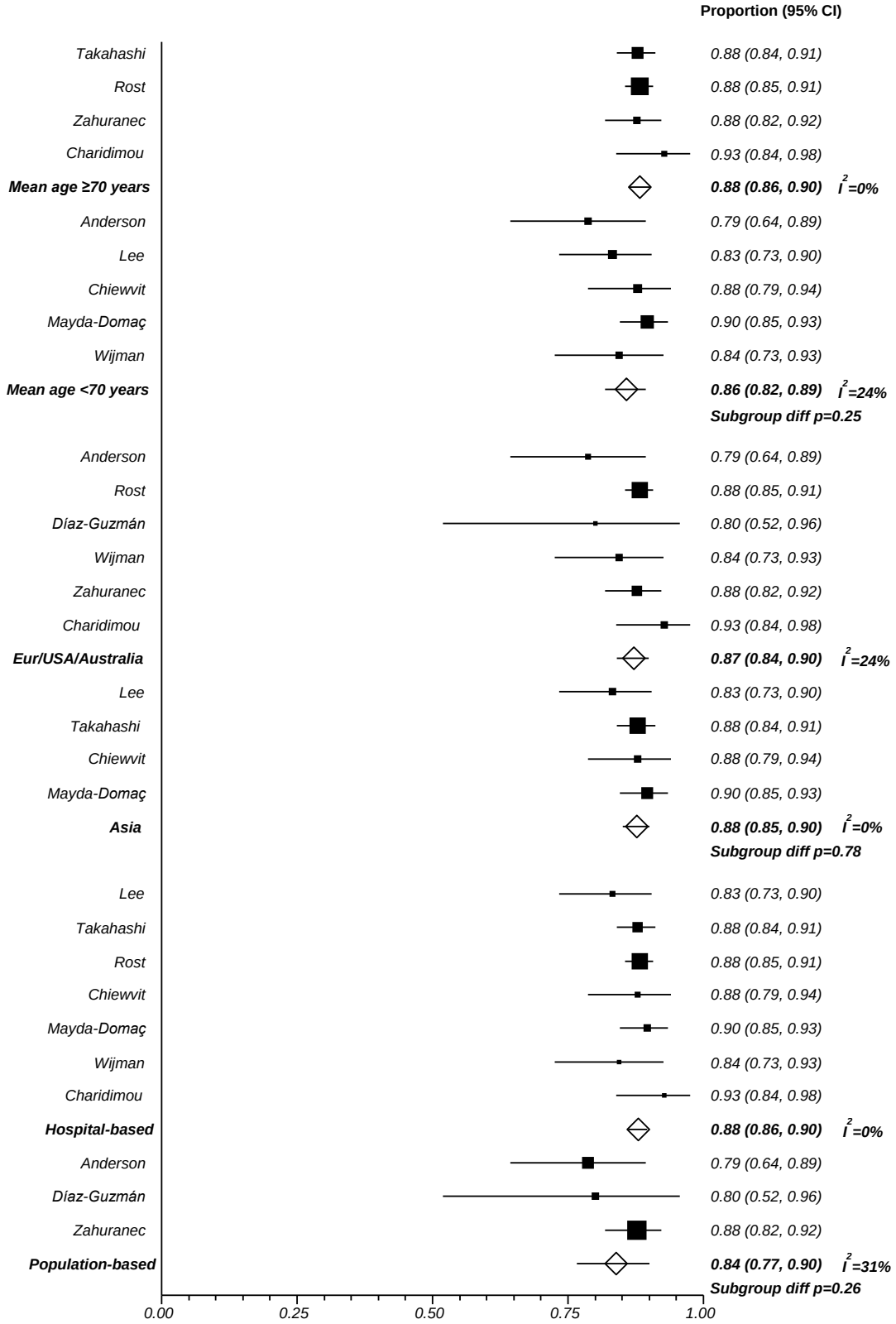
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent overall pooled proportion estimates; Studies contributing reliability data in grey; P: population-based; H: hospital-based; Y: yes; N: no; Country: country where study population recruited from; E: Europe (Spain, Finland, Germany, UK, Sweden); A: Asia (Thailand, Taiwan, South Korea, Turkey, Japan); AUS: Australia; Age: mean age of relevant study population; IVH: intraventricular haemorrhage. Where possible I excluded cases with multiple ICHs, cases in the uncertain ICH category, and cases with isolated IVH: *Anderson*: excluded 13 cases of large haemorrhage where impossible to distinguish between deep and lobar regions as site of origin; *Chiewvit*: excluded 2 cases with multiple ICHs and 4 cases with IVH; *Wijman*: excluded 7 cases with multiple ICHs and 5 cases with IVH; *Mayda-Domaç*: excluded 6 cases with IVH; *Charidimou:* excluded one case with IVH. Three studies that had used an anatomical classification system did not include sufficient data to be included in the meta-analysis (Bhattathiri et al. 2003, Ziai et al. 2011, Wermer et al. 2002). I^2 = within subgroup heterogeneity;

Figure 2.4 Lobar ICH proportion meta-analysis subgroups



Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

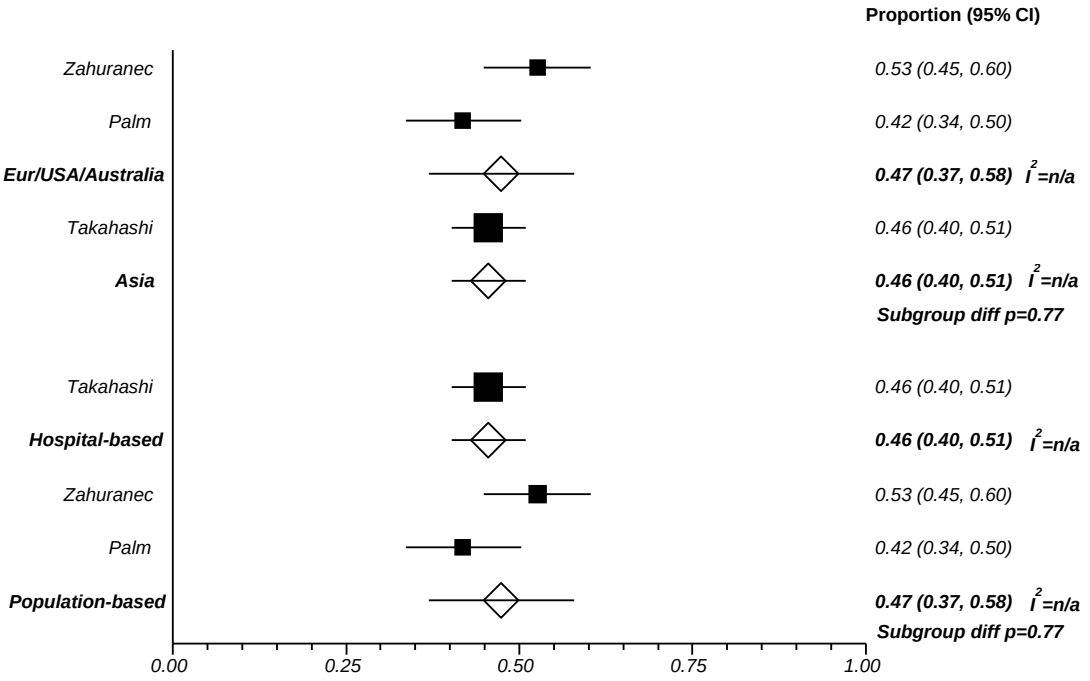
Figure 2.5 Supratentorial ICH proportion meta-analysis subgroups



Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion

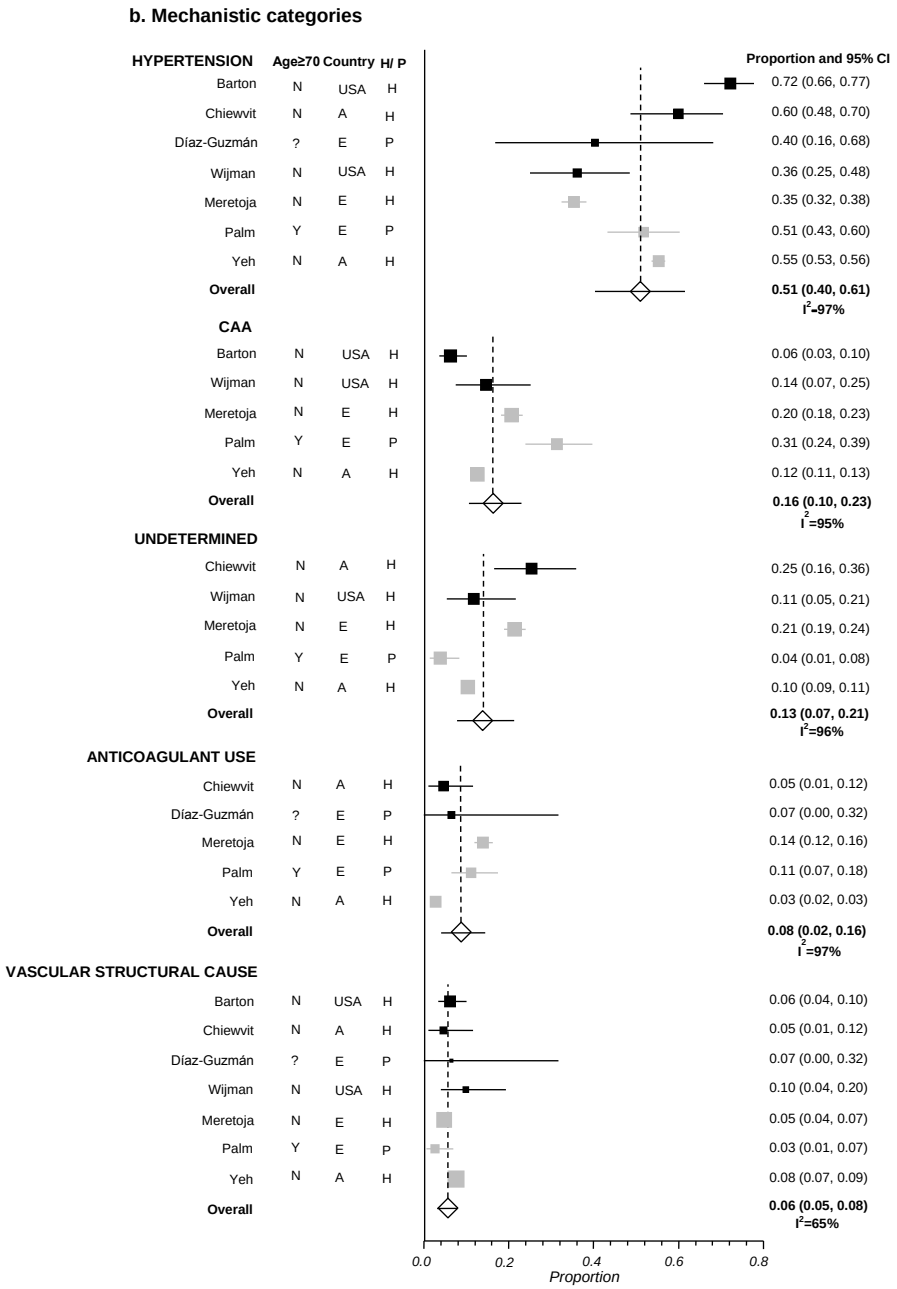
estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.6 ICH with IVH present proportion meta-analysis subgroups



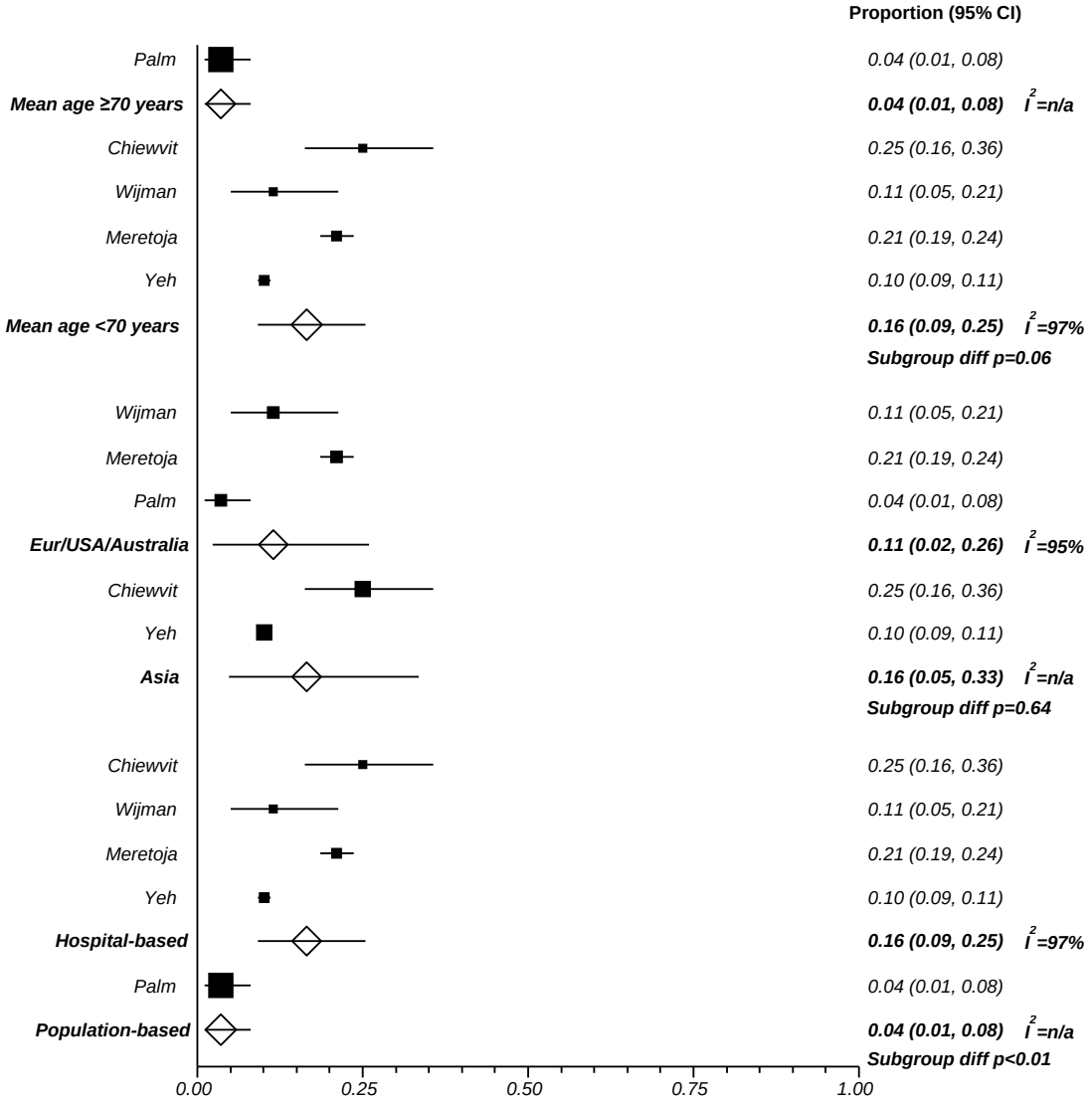
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.7 Proportions of ICH subtypes in mechanistic categories



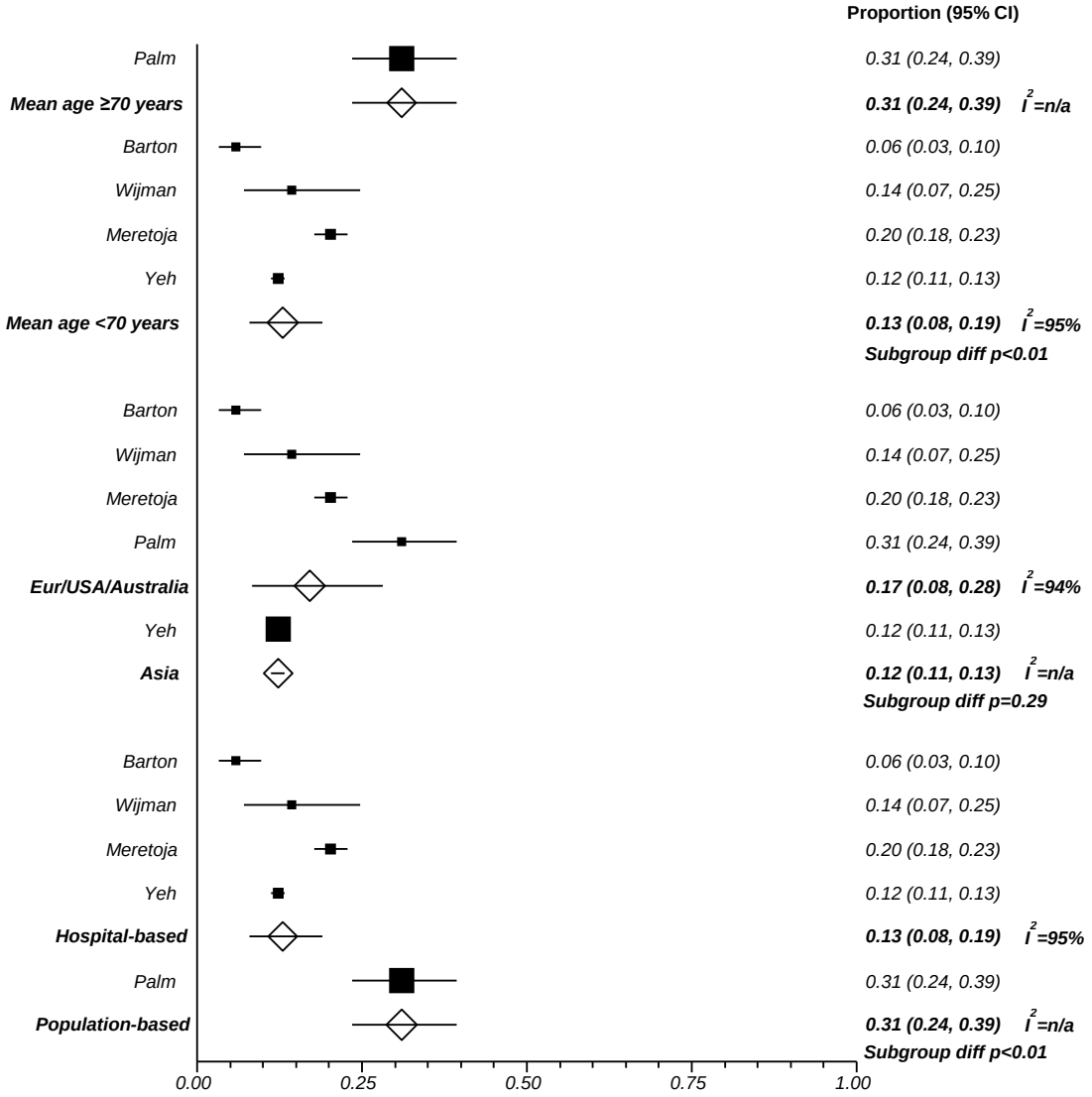
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent overall pooled proportion estimates; Studies contributing reliability data in grey; P: population-based; H: hospital-based; Y: yes; N: no; Country: country where study population recruited from; E: Europe (Spain, Finland, Germany, UK, Sweden); A: Asia (Thailand, Taiwan, South Korea, Turkey, Japan); AUS: Australia; Age: mean age of relevant study population; I^2 = within subgroup heterogeneity;

Figure 2.8 Undetermined ICH proportion meta-analysis subgroups



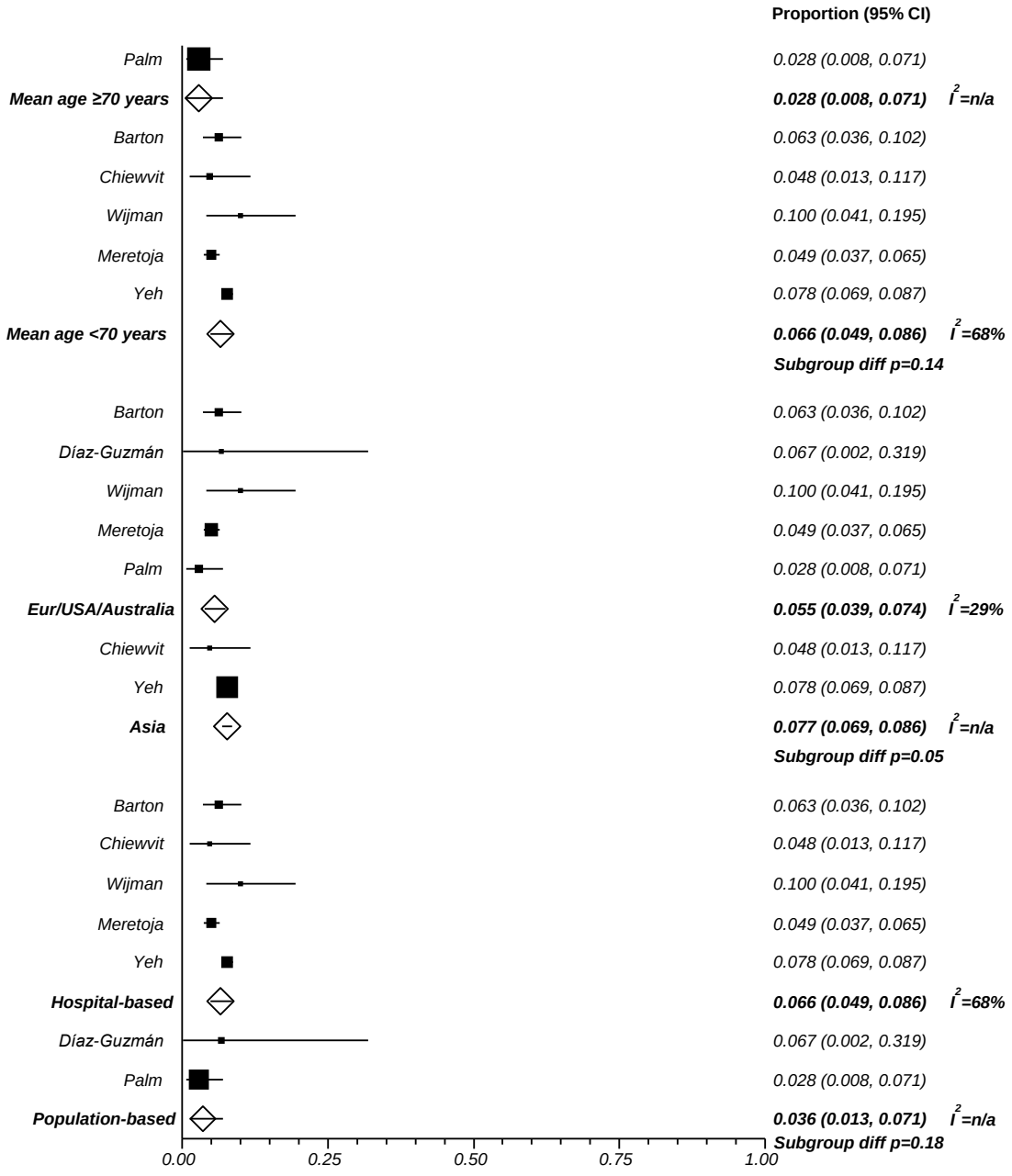
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.9 CAA-related ICH proportion meta-analysis subgroups



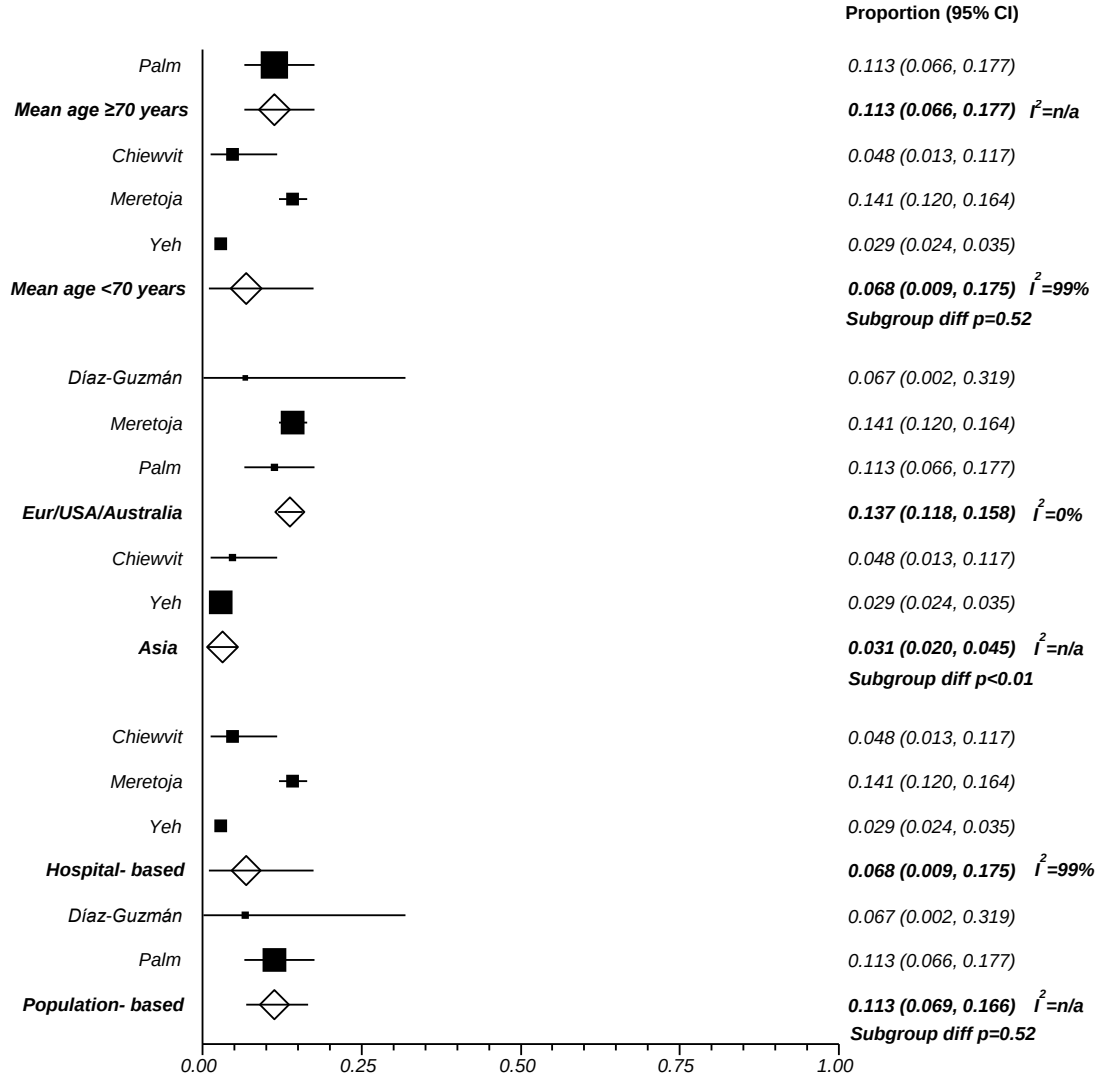
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.10 Vascular structural cause ICH proportion meta-analysis subgroups



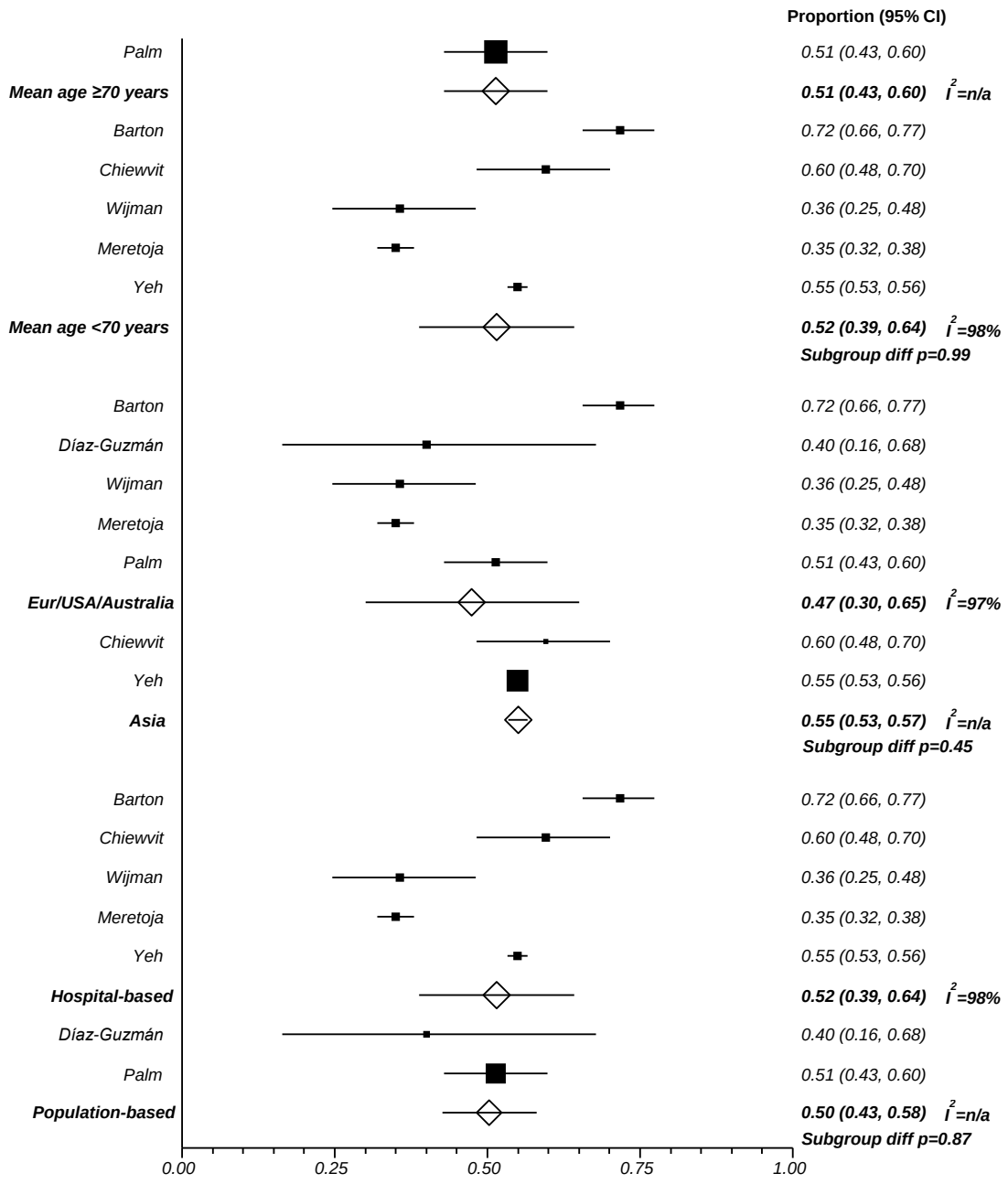
Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.11 Anticoagulant use-related ICH proportion meta-analysis subgroups



Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Figure 2.12 Hypertension-related ICH proportion meta-analysis subgroups



Random-effects method; Squares represent individual study proportion estimates, and associated horizontal lines represent 95% confidence intervals; Unfilled diamonds represent pooled proportion estimates for each subgroup; I^2: within subgroup heterogeneity; n/a: not possible to assess; Subgroup diff p: p-value showing whether there is a significant difference ($p<0.1$) between the pooled results of the two subgroups under comparison.

Chapter 3: *APOE* ε2/ε3/ε4 polymorphism associations with sporadic CAA

Chapter contents:

3. *APOE* ε2/ε3/ε4 polymorphism associations with sporadic CAA

- CAA, common in the ageing brain, is associated with dementia and lobar ICH
- *APOE* ε2/ε3/ε4 polymorphism is associated with many conditions in which CAA is involved, including ICH attributed to CAA based on clinical criteria
- An association between *APOE* and histopathologically confirmed CAA would explain these associations and increase understanding of underlying mechanisms
- In this chapter, I present my systematic review and series of meta-analyses looking at the association between *APOE* ε2/ε3/ε4 polymorphism and histopathologically confirmed presence of CAA
- I used a comprehensive search strategy, critically appraised study quality and assessed small-study biases, publication-, and reporting-biases
- I conclude that there is robust evidence for a dose-dependent association between *APOE* ε4 and histopathologically confirmed CAA, but there is no convincing association with *APOE* ε2

3.1 Introduction

Sporadic cerebral amyloid angiopathy (CAA) is characterised by deposition of amyloid-β protein in leptomeningeal and cortical blood vessels, with a prevalence in population-based autopsy studies of 20-40% in non-demented and 50-60% in demented elderly people (Charidimou et al. 2012a).

CAA is associated with increasing age, dementia, lobar intracerebral haemorrhage (ICH), lobar brain microbleeds, white matter hyperintensities, small cortical infarcts and superficial siderosis (Linn et al. 2010, Masuda et al. 1988, Neuropathology Group MRC CFAS 2001, Samarasekera et al. 2012). ICH is usually attributed to CAA when pathological examination reveals extensive CAA deposition and related vasculopathic changes (Vonsattel et al. 1991). The most commonly used criteria for

CAA diagnosis are the Boston criteria, which enables diagnosis of probable and possible CAA based on clinical and radiological criteria only, but require a full post-mortem examination for a definite diagnosis (Knudsen et al. 2001). Knudsen et al reported the Boston criteria to have a specificity of 100% and a sensitivity of 45%, however the number of participants included was only 39 (Knudsen et al. 2001). More recently, the modified Boston criteria have been shown to increase sensitivity, though not significantly (Linn et al. 2010). No specific treatment for CAA currently exists (Charidimou et al. 2012a).

Identifying genetic polymorphisms associated with histopathologically confirmed CAA should increase understanding of the mechanisms leading to CAA and associated diseases. The apolipoprotein E gene (*APOE*) on chromosome 19 codes for a protein of the same name, which is expressed in several organs including the brain. Single base changes in the gene determine the three common alleles ε2, ε3 and ε4, lead to differences in the amino acid sequence at two sites and changes in the function of the resulting protein isoforms. ApoE-3 is the most frequent (wild-type) isoform (Takeda et al. 2010). Polymorphisms in the *APOE* gene are associated with many conditions in which CAA may be involved, including subarachnoid haemorrhage, ICH, lobar brain microbleeds, and Alzheimer´s disease (AD) (Bertram et al. 2007, Biffi et al. 2011, Biffi et al. 2010, Linn et al. 2010, Maxwell et al. 2011, Paternoster et al. 2009, Sudlow et al. 2006). The best known function of the apolipoprotein E protein is the regulation of lipid metabolism, however in addition it may be involved in many other physiological and pathological processes. In vitro studies have shown that apolipoprotein E influences amyloid-β protein conformation, fibril formation and toxicity (Lambert et al. 1998, Wood et al. 1996), while in vivo mouse studies have confirmed a critical role for apolipoprotein E in amyloid-β protein deposition, toxicity and possibly clearance (Holtzman et al. 2000, Holtzman et al. 1999). The currently favoured view is that *APOE* ε4 enhances deposition of amyloid-β in cerebral blood vessel walls, while ε2 promotes haemorrhage from amyloid-laden blood vessels by increasing specific CAA-related vasculopathic changes (Greenberg et al. 1998, McCarron and Nicoll 2000, McCarron et al. 1999).

A clear association between *APOE* and histopathologically confirmed CAA would explain many of the observed associations between *APOE* and clinical outcomes. There is robust, large-scale evidence for an association of *APOE* with ICH attributed to CAA on the basis of clinical and radiological criteria. However, studies suggesting an association of *APOE* with histopathologically confirmed CAA have been limited by various methodological shortcomings, including small size. Furthermore, the possibly contrasting effects of the different *APOE* alleles on CAA have remained unclear.

In this chapter, I aim to assess the evidence for associations between the *APOE* ε2/ε3/ε4 genetic polymorphism and histopathologically confirmed CAA by carrying out a systematic review, incorporating a comprehensive search strategy, a thorough assessment of study quality, a series of meta-analyses, and an evaluation of the robustness of any positive findings to small-study and other methodological biases.

3.2 Methods

3.2.1 Search strategy

I sought all studies of adult humans published in any language, in which participants had been genotyped for any genetic polymorphism and had CAA assessed pathologically (using autopsy or biopsy), regardless of whether any association between the polymorphism and CAA was reported on.

I designed a search strategy aimed at picking up a broader range of studies than those included in this chapter, searching OVID Medline (1950 to March 2012) and Embase (1980 to March 2012), using a combination of search terms for APOE, genes and CAA (Appendix 3). I also checked the bibliographies of all relevant studies and reviews identified, and searched *Google Scholar* for studies citing relevant studies.

3.2.2 Inclusion/exclusion criteria

For the analyses reported in this chapter, I included studies looking at the *APOE* ε2/ε3/ε4 genetic polymorphism and:

- presence versus absence of sporadic CAA pathology;
- more versus less severe sporadic CAA pathology;
- or average CAA score on pathology.

I excluded studies that assessed genetic associations with CAA-related ICH versus CAA-free controls, because these would not be able to distinguish a genetic association with CAA from an association with ICH. However, I included studies that had recruited participants with CAA-related ICH as an unselected part of the spectrum of CAA severity. A clinical research fellow colleague (Dr Neshika Samarasekera) and I independently selected eligible studies, resolving disagreements by discussion and mutual consensus.

When two or more studies included overlapping sets of participants, I included only the study providing most data about the association among the largest number of participants. To avoid undue effort on relatively small studies, I further excluded studies with <35 participants, thus excluding <3% of the total participants in eligible studies.

3.2.3 Data extraction

For each study included, I extracted information on:

- first author;
- publication year;
- country in which the study was conducted;
- participant source and description;
- study size;
- participants' age, ethnicity and gender.

To assess the quality of reporting of genotyping based on the STREGA (Strengthening the Reporting of Genetic Association Studies) and MOOSE (Meta-analysis of Observational Studies in Epidemiology) recommendations (Little et al. 2009, Stroup et al. 2000), I extracted information on:

- source and storage of DNA;
- where genotyping was undertaken;
- whether genotypes were assigned using all data simultaneously or in batches;
- number of participants in whom genotyping was attempted and in whom it was successful;
- genotyping method;
- whether genotypes were in Hardy-Weinberg equilibrium (HWE);
- blinding of genotyping staff to pathology data.

To assess the quality and characteristics of CAA pathology assessment based on criteria I developed specifically for this study, I extracted data on:

- method and consistency of CAA assessment;
- qualification of CAA rater(s);
- brain locations examined;
- rating system used and its intra- and inter-rater reliability;
- blinding of CAA raters to genotyping results and relevant clinical information.

Where possible, I extracted data on the numbers of participants with ε4 or ε2 allele-containing (ε4+ or ε2+) genotypes and their CAA status, in a dichotomous (present or absent) or a continuous (mean severity score and its standard deviation [SD]) format. In studies without such data available, I recorded any qualitative statement about the association between *APOE* ε2/ε3/ε4 polymorphism and CAA. For each study, another researcher (one of Dr Neshika Samarasekera, Professor Rustam Al-Shahi Salman or Professor Cathie Sudlow) and I extracted these data, resolving disagreements by discussion and mutual consensus.

3.2.4 Assessment of study quality

I devised a system to assign the studies included in the meta-analyses an overall quality score (ranging from 0-9 points) based on the following:

- study size 0-2 points, considering larger studies to be superior (cut-offs were chosen to divide the total number of participants in all included studies into three roughly equal groups);
- blinding 0-2 points, considering studies using blinding to be superior;
- participant recruitment method 0-1 points, considering prospective studies superior to retrospective;
- format in which genotype-phenotype association data were provided 0-1 points, considering continuous data superior due to potential bias in selecting cut-offs for dichotomous data;
- quality of pathology assessment 0-2 points, based on the study reporting the method for assessing CAA, the rater qualification and number of brain locations examined for CAA;
- quality of genotyping reporting 0-1 points, based on whether the source of DNA and genotyping method were reported.

3.2.5 Meta-analyses

I used Cochrane Review Manager (version 5) for undertaking the meta-analyses. For studies presenting data in a dichotomous format, I calculated an unadjusted odds ratio (OR) with 95% confidence intervals (CIs) for presence versus absence of CAA among those with ε4+ versus ε4-, and ε2+ versus ε2- genotypes. I used the Mantel-Haenszel method which is the default method for fixed-effects meta-analysis in Review Manager and is considered to have better statistical properties than other methods when there are few events (The Cochrane Collaboration 2011). For studies presenting continuous data, I calculated standardized mean differences (SMD) with 95% CIs in CAA scores and I used the inverse variance fixed-effects method. I used SMDs because while the studies all assessed the same outcome (CAA burden), they measured it using a variety of scales and SMDs express the size of the intervention

effect in each study relative to the variability observed in that study, thus standardizing the results of the studies to a uniform scale before they are combined (Borenstein et al. 2010).

$$\text{SMD} = \frac{\text{difference in mean outcome between 2 groups in the study}}{\text{study standard deviation}}$$

Then, to analyse the data from dichotomous and continuous studies together, I used the generic inverse variance fixed-effects method, that calculates pooled ORs for all studies, after inputting the natural logarithm of the OR (logOR) and its standard error [SE(logOR)].

I used Microsoft Excel to calculate the logOR and SE (logOR) for dichotomous studies, using the "LN" function and the equation below:

	Number of participants with an *APOE* ε4 or ε2 allele present	Number of participants with an *APOE* ε4 or ε2 allele absent
Number of participants with CAA present	a	c
Number of participants with CAA absent	b	d

$$\text{SE(logOR)} = \sqrt{\frac{1}{\text{a}} + \frac{1}{\text{b}} + \frac{1}{\text{c}} + \frac{1}{\text{d}}}$$

I used Microsoft Excel and the method described by Chinn to convert SMDs and 95% CIs to logORs and SE(logOR) (Chinn 2000) as per equations:

$$\text{SE(SMD)} = \frac{\text{upper 95\% CI} - \text{lower 95\% CI}}{3.92}$$

$$\log \text{OR} = \text{SMD} * \left(\frac{\pi}{\sqrt{3}}\right)$$

$$\text{SE(logOR)} = \text{SE(SMD)} * \left(\frac{\pi}{\sqrt{3}}\right)$$

I repeated the analyses, comparing the ε4+ and ε2+ genotypes with the wild-type ε3/ε3 genotype only, to confirm that any significant associations were robust and not driven by the ε2/ε3 and ε2/ε2 genotypes in the ε4- group, or by the ε4/ε3 and ε4/ε4 genotypes in the ε2- group. I also compared participants' CAA status by ε4 homozygous and heterozygous genotypes. I assessed heterogeneity with I^2 and χ^2 statistics, interpreting the I^2 as follows (The Cochrane Collaboration 2011):

- I^2 = 0% to 40%: might not be important;
- I^2 = 30% to 60%: may represent moderate heterogeneity;
- I^2 = 50% to 90%: may represent substantial heterogeneity;
- I^2 = 75% to 100%: considerable heterogeneity.

I considered $p<0.05$ to imply statistical significance. I repeated all analyses using a random-effects method (DerSimonian and Laird method for studies presenting data in a dichotomous format, inverse-variance random-effects method for studies presenting data in a continuous format and for combined analyses of dichotomous and continuous data) (The Cochrane Collaboration 2011).

3.2.6 Subgroup analyses

For presence versus absence of CAA among those with ε4+ versus ε4- genotypes, I conducted pre-specified subgroup analyses based on:

- participants' clinical dementia status: studies with clinically demented participants versus those with clinically non-demented participants versus studies with both demented and non-demented participants (or where dementia status was not specified);
- neuropathologically confirmed AD status: studies of neuropathologically confirmed AD participants versus those with clinically non-demented participants versus studies with both demented and non-demented participants (or where dementia status was not specified);
- ethnicity: studies with Asian versus studies with Caucasian participants;
- study quality score: studies with up to 2 points versus those with 3-4 points versus studies with 5-9 points;

- each quality criterion included in the overall study quality score: study size, blinding, design, data format, pathology assessment and genotyping reporting quality.

I assessed the subgroup differences based on analyses results from Review Manager that undertakes a standard test for heterogeneity across subgroup results, considering $p<0.1$ to imply statistical significance (Petitti 2001).

3.2.7 Assessment of potential publication and reporting bias

For any statistically significant result for the association between *APOE* genotype and CAA, I assessed the potential effect of unpublished negative studies or studies not reporting the data required for meta-analysis (i.e. publication and reporting biases), using a modified 'failsafe N' method: I determined the size of a notional study with a null result (OR=1) required to bring any significant result ($p<0.05$) to a just non-significant level ($p=0.05$), assuming the overall prevalence of CAA and distribution of genotypes to be the average of these for the studies included (Rosenthal 1979). I did this by calculating the logOR and SE(logOR) for a hypothetical study with n participants in Microsoft Excel, and then including the data in the relevant Review Manager meta-analysis to investigate the effect of the hypothetical study on the overall OR and p-value. I increased the size (n) of the hypothetical study gradually until the overall p-value increased to 0.05. I used the size of this notional study as a guide to whether there might plausibly be enough participants in unpublished, unreported or otherwise unretrieved null studies to make an apparently significant result non-significant, if data from such studies were available for inclusion in my meta-analyses.

I also visually inspected funnel plots for the ε4+ versus ε4-, and ε2+ versus ε2- analysis.

3.3 Results

3.3.1 Included studies

Of 1754 publications identified for screening, 107 were relevant to my analyses of *APOE* genotype and sporadic CAA (Figure 3.1). From these 107, I excluded 61 (49 because they included participants overlapping with other included studies, and 12 because of very small size [< 35 participants]), leaving 46 studies, including 6645 participants in total (Figure 3.1) (Alafuzoff et al. 2009, Alafuzoff et al. 1999, Attems et al. 2005, Attems et al. 2008, Berg et al. 1998, Caselli et al. 2010, Chalmers et al. 2003, Christoforidis et al. 2005, Chui et al. 2006, Cruz-Sánchez et al. 2000, Davidson et al. 2006, Durany et al. 2000, Etiene et al. 1998, Greenberg et al. 1995, Greenberg et al. 1998, Honig et al. 2005, Jicha et al. 2008, Lashley et al. 2008, Leclercq et al. 2005, Lewis et al. 2006, Love et al. 2003, Mortimer et al. 2009, Nelson et al. 2010, Nicoll et al. 2011, Olichney et al. 2000, Oyama et al. 1995, Pahnke et al. 2003, Petrovitch et al. 2008, Peuralinna et al. 2011, Pfeifer et al. 2002, Premkumar et al. 1996, Roher et al. 2003, Schneider et al. 2005, Sonnen et al. 2010, Stopa et al. 2008, Zarow et al. 1999, Zipser et al. 2007, Zubenko et al. 1994, Tanskanen et al. 2005, Thal et al. 2002, Thomas et al. 2000, Tiraboschi et al. 2004, Walker et al. 2000, Yamada 2002, Yamaguchi et al. 2001, Yip et al. 2005).

3.3.2 Study characteristics

All studies used autopsy brains from a brain bank, clinical autopsies, or brain pathological material (mainly autopsy, some biopsy) arising during the course of a population-based prospective study. The median number of participants per study who were both genotyped and assessed pathologically was 100. Mean age was 70 to 85 years in most studies. About half of all participants were male. 19 studies (2660 participants) were conducted in white populations in Europe, 21 (3225 participants) in whites in the USA, five (714 participants) in Asian populations (three in Japan, two in Japanese-Americans in the USA) and for one study this information was unavailable. About 30% of participants had clinical dementia (mainly AD), about

10% were known not to be demented and dementia status was not specified for about 60% (Table 3.1).

Genotyping reporting quality was generally limited when assessed against current reporting standards (Little et al. 2009, Stroup et al. 2000) (Table 3.2). 10 studies (1723 participants) did not report the source of DNA, 25 (2993 participants) used brain tissue, seven (1040 participants) used blood samples, three (632 participants) used both, and one used brain tissue or buccal cells. No study described how DNA was stored or how data were used to assign genotypes. Only two studies (350 participants) reported where genotyping had been done and two (294 participants) reported on the number of participants in whom genotyping had been attempted and in whom it had been successful. Six studies (1380 participants) did not report on the method of genotyping, while the rest used PCR-based techniques. In 22 studies (2884 participants) genotypes were either stated or calculated to be in HWE, for three (326 participants) they were not in HWE, 19 (3328 participants) provided no information on HWE and for two studies (117 participants) it was not applicable because included participants were chosen to provide a mixture of *APOE* genotypes or CAA severity. Only two studies (333 participants) reported that genotypers were blind to pathology data.

Methods for pathological assessment were very variable (Table 3.3). With regard to CAA assessment, 34 studies (4570 participants) used immunohistochemistry (IHC) in all or some samples, 10 (1291 participants) used an alternative (Congo Red ± polarized light, Hematoxylin & Eosin, or Thioflavin S) and two (794 participants) did not report on staining method. Pathological assessment methods were used consistently in 44 studies, meaning that the same method was used throughout the study and did not vary with time, recruiting centre etc. 14 studies (2231 participants) specified that a neuropathologist rated CAA, but most did not report on this. A variety of CAA rating scales were used: 10 studies used scales devised by Olichney (Olichney et al. 1995), Vonsattel (Vonsattel et al. 1991), and/or Attems (Attems 2005); 28 studies used an alternative severity rating measure or scale; five did not report the scale or severity measure used; and three did not rate CAA severity. Only two studies reported the inter- and intra-rater reliability of the scale used (or

referenced a study of these). 31 studies (4600 participants) assessed CAA in occipital and/or parietal cortex, frequently in addition to other areas. Only eight studies (890 participants) reported that pathology raters were blind to genotyping results and six (1096 participants) that they were blind to relevant clinical information.

3.3.3 Assessment of study quality

For the studies in the meta-analysis, I calculated quality scores (ranging from 0 to 9) based on my pre-set criteria. There was substantial variation in overall study quality. No study fulfilled all criteria, but larger studies tended to have higher quality scores (Table 3.4).

3.3.4 Meta-analyses results

3.3.4.1 Association between *APOE* ε4 and CAA

24 studies out of 46 (3520 of 6645 participants) provided sufficient quantitative data for meta-analyses (14 in a dichotomous and 10 in a continuous format) (Figure 3.2, Figure 3.3) (Alafuzoff et al. 1999, Attems et al. 2005, Caselli et al. 2010, Chalmers et al. 2003, Christoforidis et al. 2005, Chui et al. 2006, Cruz-Sánchez et al. 2000, Davidson et al. 2006, Greenberg et al. 1995, Leclercq et al. 2005, Mortimer et al. 2009, Nicoll et al. 2011, Olichney et al. 2000, Pfeifer et al. 2002, Premkumar et al. 1996, Schneider et al. 2005, Zarow et al. 1999, Zubenko et al. 1994, Tanskanen et al. 2005, Thal et al. 2002, Walker et al. 2000, Yamada 2002, Yamaguchi et al. 2001, Yip et al. 2005). 22 studies could not be included in my meta-analysis: six (443 participants) of these made a qualitative statement (Berg et al. 1998, Jicha et al. 2008, Oyama et al. 1995, Roher et al. 2003, Stopa et al. 2008, Thomas et al. 2000), while 16 (2682 participants) provided no data about the association (Alafuzoff et al. 2009, Attems et al. 2008, Durany et al. 2000, Etiene et al. 1998, Greenberg et al. 1998, Honig et al. 2005, Lashley et al. 2008, Lewis et al. 2006, Love et al. 2003, Nelson et al. 2010, Pahnke et al. 2003, Petrovitch et al. 2008, Peuralinna et al. 2011, Sonnen et al. 2010, Tiraboschi et al. 2004, Zipser et al. 2007).

Figure 3.2 shows study specific and pooled ORs for ε4+ versus ε4- genotypes based on the fixed-effects method. The pooled OR showed a significantly increased odds of having CAA for ε4+ genotypes (OR 2.89, 95% CI 2.51 to 3.33; p<0.00001), with significant heterogeneity between the studies' results (I^2=71%; χ^2_{23df}=80.68; p<0.00001). When I meta-analysed studies providing continuous and dichotomous data separately, pooled results were in the same direction as the overall pooled OR and were separately significant (continuous studies: SMD 0.50, 95% CI 0.40 to 0.60; p<0.00001; dichotomous studies: OR 4.25, 95% CI 3.33 to 5.43; p<0.00001). The association was of similar magnitude and remained significant when I used the random-effects method (OR 3.08, 95% CI 2.32 to 4.09; p<0.00001) (Figure 3.3) and when I compared the ε4+ versus ε3/ε3 genotypes (OR 2.68, 95% CI 2.15 to 3.32; p<0.00001) (Figure 3.4, Table 3.5).

Subgroup analysis based on clinical dementia status showed no significant subgroup differences (I^2=0%; χ^2_{2df}=1.91; p=0.38) (Figure 3.5), with similar results when I limited the dementia subgroup to neuropathologically confirmed AD cases only (I^2=27.3%; χ^2_{2df}=2.75; p=0.25) (Figure 3.6, Figure 3.7, Table 3.6). I detected no significant differences in subgroup analyses based on ethnicity (I^2=57.1%; χ^2_{1df}=2.33; p=0.13) (Figure 3.6, Figure 3.8). The association did not vary significantly by study quality score (I^2=41.9%; χ^2_{2df}=3.44; p=0.18) (Figure 3.6, Figure 3.9). Nor did I detect significant subgroup differences based on the individual quality criteria of study size (I^2=0%; χ^2_{2df}=1.59; p=0.45) (Figure 3.6, Figure 3.10), blinding (I^2=56.4%; χ^2_{1df}=2.29; p=0.13) (Figure 3.6, Figure 3.11) or quality of genotyping reporting (I^2=56.5%; χ^2_{1df}=2.30; p=0.13) (Figure 3.6, Figure 3.12). The association was slightly larger in studies with prospective recruitment of participants compared to retrospective studies (OR 3.66, 95% CI 2.87 to 4.68 versus OR 2.56, 95% CI 2.15 to 3.05; test for subgroup differences I^2=81.8%; χ^2_{1df}=5.48; p=0.02) (Figure 3.6, Figure 3.13**)**, in studies presenting data in dichotomous versus continuous format (OR 3.64, 95% CI 2.91 to 4.55 versus OR 2.48, 95% CI 2.06 to 2.98; test for subgroup differences I^2=85.2%; χ^2_{1df}=6.77; p=0.009) (Figure 3.6, Figure 3.14) and in studies with higher quality of pathology assessment (studies with 2 points: OR 4.20, 95% CI 3.37 to 5.24 versus studies with 1 point: OR 2.22, 95% CI

1.84 to 2.67; test for subgroup differences I^2=94.7%; χ^2_{1df}=18.91; p<0.0001) (Figure 3.6, Figure 3.15). The subgroup differences were similar when I used the random-effects method, although no longer statistically significant by study design (p=0.18) and data format (p=0.11) (Table 3.7).

Studies providing only qualitative data reported either no significant association or a trend towards association with *APOE* ε4 (Berg et al. 1998, Jicha et al. 2008, Oyama et al. 1995, Roher et al. 2003, Stopa et al. 2008, Thomas et al. 2000) (Table 3.8).

12 studies provided quantitative data for meta-analysis of the association of *APOE* ε4 allele dose with CAA, showing a significant increase in the odds of having CAA with increasing dose of the ε4 allele (Attems et al. 2005, Chalmers et al. 2003, Cruz-Sánchez et al. 2000, Davidson et al. 2006, Leclercq et al. 2005, Nicoll et al. 2011, Olichney et al. 2000, Premkumar et al. 1996, Zarow et al. 1999, Zubenko et al. 1994, Thal et al. 2002, Walker et al. 2000). The OR for participants with one ε4 allele versus no ε4 allele was 2.09 (95% CI 1.69 to 2.58; p<0.00001), for ε4 homozygotes versus ε4 heterozygotes was 3.38 (95% CI 2.34 to 4.88; p=0.04) and for the ε4 homozygotes versus participants with no ε4 allele was 7.04 (95% CI 4.81 to 10.29; p=0.0003) (Figure 3.16). Two studies (117 participants) providing a qualitative statement about the association supported this result (Greenberg et al. 1995, Roher et al. 2003) (Table 3.8).

3.3.4.2 Association between *APOE* ε2 and CAA

11 studies (1640 participants) provided quantitative data for meta-analysis of *APOE* ε2+ versus ε2- genotypes with CAA (Attems et al. 2005, Christoforidis et al. 2005, Cruz-Sánchez et al. 2000, Davidson et al. 2006, Leclercq et al. 2005, Nicoll et al. 2011, Olichney et al. 2000, Premkumar et al. 1996, Tanskanen et al. 2005, Thal et al. 2002, Walker et al. 2000). Figure 3.17 shows study specific and pooled ORs for ε2+ versus ε2- genotypes based on the fixed-effects method. The pooled OR showed significantly decreased odds of having CAA for ε2+ genotypes (OR 0.64, 95% CI 0.47 to 0.86, p=0.003). When I meta-analysed studies providing continuous and dichotomous data separately, pooled results were in the same direction as the overall pooled OR (continuous studies: SMD -0.27, 95% CI -0.5 to -0.03; p=0.03;

dichotomous studies: OR 0.63, 95% CI 0.43 to 0.94; p=0.03). There was moderate heterogeneity between the studies' results (I^2=46%; χ^2_{10df}=18.57; p=0.05). However the association was no longer significant when I used the random-effects method (OR 0.68, 95% CI 0.42 to 1.08; p=0.1) (Figure 3.18) or when I compared the ε2+ genotypes with ε3/ε3 genotypes (OR 0.90, 95% CI 0.65 to 1.26; p=0.55) (Table 3.5, Figure 3.19).

Two studies (213 participants) provided a qualitative statement; neither reported a significant association (Chalmers et al. 2003, Greenberg et al. 1995) (Table 3.8).

3.3.4.3 Bias assessment

Failsafe N calculations showed that a null study of >177,000 participants would be required to bring the association of ε4+ genotypes with CAA to a just statistically non-significant level (i.e. p=0.05). It would require a null study of >7600 participants to bring the stronger association with CAA of ε4 homozygous versus heterozygous genotypes to a just non-significant level.

The funnel plots for the ε4+ versus ε4- analysis and for the ε2+ versus ε2- analysis were only slightly asymmetrical (Figure 3.20, Figure 3.21).

3.4 Discussion

3.4.1 Main findings

I found a highly significant, dose-dependent association between *APOE* ε4 and pathologically proven CAA. The association remained robust when using both fixed-effects and random-effects methods, and when using both the ε4- genotypes and the ε3/ε3 genotypes as the comparison group. The association did not vary significantly with dementia status, ethnicity, or study quality. I found no overall robust association between *APOE* ε2 and presence of pathological CAA (Rannikmäe et al. 2013) (Appendix 4). While analysis using the fixed-effects method suggested a significant protective effect of ε2, this was not confirmed by analysis using the random-effects method, and by analyses using the ε3/ε3 genotypes (rather than the ε2- genotypes) as

the comparison group. The latter would suggest that the apparent protective effect of ε2 could be driven by the ε4/ε3 and the ε4/ε4 genotypes in the ε2- group.

To my knowledge, this is the first systematic review and meta-analysis of associations between *APOE* ε2/ε3/ε4 genetic polymorphisms and pathologically proven CAA, although there have been meta-analyses of *APOE* genetic associations with deep and lobar ICH, ICH severity and outcome, and brain microbleeds (Biffi et al. 2011, Biffi et al. 2010, Maxwell et al. 2011). The ε4 may influence the risk of CAA by increasing the amyloid-β 40:42 ratio, shifting amyloid deposition to the vessels instead of brain parenchyma (Fryer et al. 2005), and may reduce the efficiency of efflux of amyloid-β along perivascular channels (Thal et al. 2007) (Charidimou et al. 2012a).

The prevalence of CAA in AD is over 70% but the relationship between CAA and AD is still poorly understood. Although the diagnostic criteria for dementia and the participant inclusion criteria varied between studies (with some excluding cases with severe dementia), the demonstration of a similar association in those with and without clinical dementia suggests that the association of *APOE* ε4 with CAA is independent of its known association with dementia (mainly AD). This is further supported by similar results when I limited the dementia subgroup to neuropathologically verified AD cases only. Relative to other apoE isoforms, apoE4 is thought to increase aggregation or impair clearance of amyloid-β, or both. While this mechanistic pathway is probably common to both CAA and AD, specific mechanisms might also occur. For example, apoE isoform-specific neurotoxic effects may contribute to neurodegeneration in AD, independent of interactions with amyloid-β (Verghese et al. 2011).

I performed the analyses using the fixed-effects method and then repeated all analyses using the random-effects method to confirm robustness of any significant results. It is generally recommended to use the fixed-effects method when the underlying assumption is that the true allele effect is the same in all studies included in the meta-analysis, and to use the random-effects method when the underlying assumption is that the true allele effect is different in the included studies (Borenstein

et al. 2010). I chose to use both methods to confirm robustness of any significant results and also because there was moderate to substantial heterogeneity between the studies. This approach is supported by the Cochrane Collaboration, which recommends comparing the fixed- and random-effects estimates if one is concerned about the influence of small-study effects on the results of meta-analysis in which there is evidence of between-study heterogeneity (The Cochrane Collaboration 2011). If the estimates are similar, then any small-study effects have little effect on the effect estimate.

Since this systematic review, at least one further study looking at the association between *APOE* ε2/ε3/ε4 polymorphism and neuropathological CAA in 1062 subjects has been published (Yu et al. 2015). Yu et al study confirms a significant association between ε4 and CAA, though there is some overlap with the cases included in my systematic review. Interestingly, Yu et al also show a significant association between CAA severity and ε2, but only after adjusting for Alzheimer's pathology. There is no data available about proportion of cases with an intracerebral haemorrhage in this cohort. There was not sufficient data to perform a meaningful subgroup analysis for ε2 based on participants' dementia status in my systematic review. However when including data from Yu et al (2015) to the overall meta-analysis of *APOE* ε2 association with CAA, the association remains non-significant (fixed-effects analysis OR 1.25, 95% CI 0.99 to 1.57; p=0.06; random-effects analysis OR 1.11, 95% CI 0.71 to 1.74; p=0.63). Further studies are needed to investigate if there is an association of ε2 with neuropathological CAA when adjusted for AD pathology.

3.4.2 Strengths and limitations

My study benefited from thorough ascertainment and critical appraisal of pertinent studies including a large number of participants. Lack of variation in the effect of *APOE* ε4 by study size and the very large failsafe N strongly suggest that this association could not plausibly be explained by publication, reporting or any other small-study bias. This is despite the fact that the failsafe N method assumes that the mean effect size in the hidden studies is zero, when in fact it could be negative, which would require fewer studies to nullify the effect (Borenstein et al. 2009). In

addition, although study quality was generally limited when assessed against current reporting standards (Little et al. 2009, Stroup et al. 2000), there were – reassuringly – no significant subgroup differences by overall study quality score.

There were some limitations. Pathological assessment was very variable. CAA assessment location varied widely, possibly influencing the rate of CAA detection, since a greater burden of CAA is generally reported in the occipital or parietal lobes, albeit with a higher frequency of frontal lobe involvement reported in studies from China and Japan (Attems 2005, Chen et al. 2010). Indeed, at the time these studies were conducted, there was no widely accepted, standardized histopathological grading system for CAA (Chen et al. 2010), and no comparative studies to determine the most accurate method for assessing CAA. Recently, a new consensus protocol and scoring scheme for the assessment of CAA and associated vasculopathic abnormalities in post-mortem brain tissue has been developed (Love et al. 2014). Most included studies did not address the issue of potential confounding due to population stratification. However, such confounding is unlikely because these were not case-control studies but cross-sectional studies across a range of CAA severities. Finally, there were a large number of studies that I could not include in the meta-analysis because they did not report the necessary quantitative data about the association between the genetic polymorphism and CAA. However, my systematic review is strengthened by the identification and detailed characterisation of these studies. By contrast, many meta-analyses do not account for otherwise relevant studies without the necessary data within their publications and so risk undetected reporting bias, since many studies report only significant findings and fail to mention non-significant ones. I have shown that my findings are robust to these missing data, since the large failsafe N (>177,000) greatly exceeded the total number of participants excluded due to lack of available data (3,125).

3.4.3 Outstanding issues

Since *APOE* effects on ICH may vary with ethnicity, there may be ethnic variation in genetic associations with CAA, but these have not been widely studied in non-white populations (Tzourio et al. 2008). Genetic associations may also differ by CAA

location and subtype. For example, there is preliminary evidence that *APOE* ε4 may be associated with CAA type 1 (where CAA is found in cortical capillaries), and ε2 with CAA type 2 (where amyloid is deposited in leptomeningeal and cortical vessels with the exception of cortical capillaries) (Thal et al. 2002). Finally, the suggestion of different genetic influences on amyloid deposition in the vessel wall and progression to advanced vasculopathic changes leading to ICH requires further investigation, in particular the proposed differential effects of ε4 and ε2 alleles.

3.4.4 Conclusions

In conclusion, despite study quality issues, a large body of evidence supports an association of *APOE* ε4 with the presence and severity of histopathologically confirmed CAA, at least in white populations. There is no convincing robust association of histopathologically confirmed CAA with *APOE* ε2. Future research efforts require methodologically robust studies adhering to a standardised CAA assessment protocol (Love et al. 2014) and current reporting standards, facilitating comparisons between studies and collaborative data pooling efforts. These should focus on the differential effects of *APOE* ε2 and ε4, variation in genetic effects by ethnicity and CAA location, and on the potential influence of other polymorphisms on this clinically important but as yet incompletely understood phenotype.

Tables

Table 3.1 Characteristics of studies included in the systematic review of *APOE* genotype associations with CAA

First author, publication year	Country[1]	Brief description of participants[2]	Study size[3]	Mean age or age range (years)	Ethnicity	% male
Alafuzoff 1999	Finland	Brains from a brain bank	209	Dementia:82 No dementia:75	Caucasian[4]	33
Attems 2005	Austria	Brains from a large hospital pathology department	53	84[5]	Caucasian[4]	39[5]
Caselli 2010	USA	Cognitively healthy adults mainly from retirement communities enrolled in the Sun Health Research Institute Brain Donation program 1991-2007	179	83	Majority Caucasian	55
Chalmers 2003	UK	Clinically diagnosed AD and 34 elderly normal controls from a brain bank	120	78[5]	Caucasian[4]	47[5]
Christoforidis 2005	Germany	Routine autopsy brains	118	74	Caucasian	-
Chui 2006	USA	People from the Ischemic Vascular Dementia Programme (longitudinal study of subcortical ischemic vascular dementia, AD and cognitively normal elderly people)	73	83[5]	86% Caucasian	57[5]
Cruz-Sánchez 2000	-	Cases with clinically diagnosed dementia & controls without cognitive impairment	73	Dementia:76 No dementia:67	Caucasian[4]	60
Davidson 2006	UK	Neuropathologically confirmed AD cases	146	-	Caucasian	-
Greenberg 1995	USA	Brains selected from a brain tissue resource centre to provide cases each of CAA grade	93	79	Caucasian[4]	-
Leclercq 2005	UK	Traumatic head injury cases	88	30[6]	Caucasian[4]	74
Mortimer 2009	USA	Participants from the Nun Study (prospective	267	91	Caucasian[4]	0

First author, publication year	Country[1]	Brief description of participants[2]	Study size[3]	Mean age or age range (years)	Ethnicity	% male
		study which recruited members of the School Sisters of Notre Dame congregation 1991-1992)				
Nicoll 2011	UK	Participants from MRC CFAS neuropathology cohort of 456 participants (UK population-based study that started in early 1990s, and recruited people aged over 65 years from the community)	310	Dementia:89[6] No dementia:84[6]	Caucasian[4]	40
Olichney 2000	USA	AD cases from an autopsy series	247	80[5]	Caucasian[4]	53[5]
Pfeifer 2002	USA	Men from the Honolulu-Asia Aging Study (prospective study recruiting Japanese-American men born 1900-1919, living in Oahu 1965-1968)	201	85[5]	Japanese-American	100
Premkumar 1996	USA	People from an AD research programme	240	Normal controls:70 Neurological controls:73 AD cases:79	Caucasian[4]	45
Schneider 2005	USA	People from the Religious Orders Study (prospective study in older Catholic nuns, priests, and monks with no dementia at baseline)	208	86[5]	Majority Caucasian	45[5]
Tanskanen 2005	Finland	People aged ≥ 95 years from the Vantaa 85+ study (Finnish population-based study recruiting people from the community born before 1906, living in Vantaa on April 1, 1991)	71	97[5]	Caucasian[4]	18[5]
Thal 2002	-	Cases with CAA & controls without	56	77	-	52
Walker 2000	Germany	Routine autopsy brains	244	67	Caucasian[4]	55

First author, publication year	Country[1]	Brief description of participants[2]	Study size[3]	Mean age or age range (years)	Ethnicity	% male
Yamada 2002	Japan	Brains from autopsy series of a large geriatric hospital in Japan, excluding brains with neurodegenerative diseases except AD	201	86	Asian	40
Yamaguchi 2001	Japan	People who died from cancer	101	65	Asian[4]	-
Yip 2005	USA	AD cases	99	75	Majority Caucasian	97
Zarow 1999	USA	Cases with definite AD	42	77[5]	Caucasian[4]	50[5]
Zubenko 1994	USA	Neuropathologically confirmed AD cases	91	76	Caucasian[4]	39
Berg 1998	USA	Brains from a consecutive autopsy series of prospectively studied probable AD or incipient dementia participants and controls	176	81[5]	Caucasian[4]	46[5]
Jicha 2008	USA	AD patients from an AD research centre and a community-based study	81	88	Caucasian[4]	26
Oyama 1995	Japan	Consecutive autopsy brains from a geriatric hospital and psychiatric hospital	37	50-101	Asian[4]	-
Roher 2003	USA	Brains from a brain donation programme	24	82	Caucasian[4]	54
Stopa 2008	USA	Brains from a brain bank and AD research centre	75	78	Caucasian	38
Thomas 2000	-	Elderly Norwegian AD patients	50	86	Caucasian[4]	28
Alafuzoff 2009	Finland	Participants from a university hospital	492	74[5]	Caucasian[4]	46[5]
Attems 2008	Austria	Consecutive autopsy brains from 3 large hospitals	157	79[5]	Caucasian[4]	40[5]
Durany 2000	-	AD patients, CAA cases without other AD features and non-neurological controls from a	15	74[5]	Caucasian	56[5]

First author, publication year	Country[1]	Brief description of participants[2]	Study size[3]	Mean age or age range (years)	Ethnicity	% male
		general hospital				
Etiene 1998	USA	AD cases from an AD research centre brain bank	100	80	Caucasian[4]	36
Greenberg 1998	USA	Autopsy brains with moderate or severe CAA chosen from a brain tissue resource centre	47	79[5]	Majority Caucasian	-
Honig 2005	USA	Demented and control participants	538	80	Majority Caucasian	47
Lashley 2008	UK	PD cases and normal control participants randomly selected from a brain bank	57	78[5]	Caucasian[4]	53[5]
Lewis 2006	UK	AD cases, vascular dementia cases, and controls from a brain tissue resource centre	98	81	Caucasian[4]	44
Love 2003	UK	Brain donors without evidence of AD (CERAD criteria)	118	60-102[5]	Caucasian[4]	51[5]
Nelson 2010	USA	Participants from an AD centre autopsy cohort	334	84	Majority Caucasian	41
Pahnke 2003	-	Routine autopsy brains of non-AD participants	30	67[5]	Caucasian[4]	54[5]
Petrovitch 2008	USA	Men from the Honolulu-Asia Aging Study (prospective study recruiting Japanese-American men born 1900-1919, living in Oahu 1965-1968)	174	86[5]	Japanese-American	100
Peuralinna 2011	Finland	People from the Vantaa 85+ study (Finnish population-based study that recruited people from the community born before 1906 and living in Vantaa on April 1, 1991)	211	93[5]	Caucasian[4]	17[5]
Sonnen 2010	USA	Participants from the Adult Changes in Thought	256	93 participants≤75	Majority	42[5]

First author, publication year	Country[1]	Brief description of participants[2]	Study size[3]	Mean age or age range (years)	Ethnicity	% male
		Study (prospective study enrolling 3392 cognitively intact community-dwelling participants aged ≥65 in 1994-2003)		157 participants:76-85 43 participants:≥ 86[5]	Caucasian	
Tiraboschi 2004	USA	Brains with a neuropathological diagnosis of AD	48	80	Caucasian[4]	42
Zipser 2007	USA	AD participants	7	51-97[5]	Caucasian[4]	39[5]

Dark grey shading: studies providing data for the meta-analyses; light grey shading: studies providing a qualitative statement about the ε4 allele association with CAA; no shading: studies known to have *APOE* genotype and CAA data but providing no data about association between *APOE* genotype and CAA; "-": data not available; AD: Alzheimer's disease; PD: Parkinson's disease; CAA: cerebral amyloid angiopathy; [1] Country: country where study conducted (if not stated explicitly, then it was assumed from authors' affiliations); [2] All studies included autopsy brains; [3] Study size: number of non-overlapping participants genotyped for the *APOE* polymorphism and assessed for CAA; [4] Ethnicity not explicitly stated, but assumed from context; [5] Data provided for a larger or smaller group of participants, not just those genotyped for the *APOE* polymorphism and assessed for CAA; [6] Median age.

Table 3.2 Genotyping assessment and reporting characteristics of studies included in the systematic review of *APOE* genotype associations with CAA

First author, publication year	DNA source	DNA storage	Genotyping centre	Genotypes assigned simultaneously or in batches	No. genotyping attempted / no. successful[1]	Genotyping method	Genotypes in HWE	Blinding[2]
Alafuzoff 1999	Brain or blood	-	-	-	-	PCR-based	-	-
Attems 2005	Brain	-	-	-	-	PCR-based	Yes	-
Caselli 2010	-	-	-	-	-	PCR-based	Yes	-
Chalmers 2003	Brain	-	-	-	-	PCR-based	-	-
Christoforidis 2005	Brain	-	-	-	118/125	PCR-based	Yes	-
Chui 2006	Blood	-	-	-	-	-	-	-
Cruz-Sánchez 2000	Brain	-	-	-	-	PCR-based	Yes	-
Davidson 2006	Brain	-	-	-	-	PCR-based	Yes	-
Greenberg 1995	Brain	-	-	-	-	PCR-based	n/a	Yes
Leclercq 2005	Brain	-	-	-	-	PCR-based	Yes	-
Mortimer 2009	Brain or buccal cells	-	-	-	-	PCR-based	-	-
Nicoll 2011	Brain	-	-	-	-	PCR-based	Yes	-
Olichney 2000	Brain or blood	-	-	-	-	PCR-based	Yes	-
Pfeifer 2002	Blood	-	-	-	-	PCR-based	-	-
Premkumar 1996	Brain	-	-	-	-	PCR-based	Yes	Yes
Schneider 2005	Blood	-	-	-	-	PCR-based	Yes	-
Tanskanen 2005	Blood	-	-	-	-	PCR-based	Yes	-

First author, publication year	DNA source	DNA storage	Genotyping centre	Genotypes assigned simultaneously or in batches	No. genotyping attempted / no. successful[1]	Genotyping method	Genotypes in HWE	Blinding[2]
Thal 2002	Brain	-	-	-	-	PCR-based	Yes	-
Walker 2000	Brain	-	-	-	-	PCR-based	No	-
Yamada 2002	Brain	-	-	-	-	PCR-based	Yes	-
Yamaguchi 2001	-	-	-	-	-	-	-	-
Yip 2005	-	-	-	-	-	-	-	-
Zarow 1999	Brain	-	-	-	-	PCR-based	Yes	-
Zubenko 1994	Brain	-	-	-	-	PCR-based	Yes	-
Berg 1998	Brain or blood	-	Stated	-	176/199	PCR-based	Yes	-
Jicha 2008	Blood	-	-	-	-	PCR-based	-	-
Oyama 1995	-	-	-	-	-	PCR-based	Yes	-
Roher 2003	Brain	-	-	-	-	PCR-based	n/a	-
Stopa 2008	-	-	-	-	-	PCR-based	No	-
Thomas 2000	Brain	-	-	-	-	PCR-based	Yes	-
Alafuzoff 2009	Brain	-	-	-	-	PCR-based	-	-
Attems 2008	Brain	-	-	-	-	PCR-based	Yes	-
Durany 2000	Brain	-	-	-	-	PCR-based	-	-
Etiene 1998	Brain	-	-	-	-	PCR-based	Yes	-
Greenberg 1998	Brain	-	-	-	-	PCR-based	-	-
Honig 2005	-	-	-	-	-	-	-	-

First author, publication year	DNA source	DNA storage	Genotyping centre	Genotypes assigned simultaneously or in batches	No. genotyping attempted / no. successful[1]	Genotyping method	Genotypes in HWE	Blinding[2]
Lashley 2008	Brain	-	-	-	-	PCR-based	-	-
Lewis 2006	Brain	-	-	-	-	PCR-based	-	-
Love 2003	-	-	-	-	-	PCR-based	-	-
Nelson 2010	-	-	-	-	-	-	-	-
Pahnke 2003	Brain	-	-	-	-	PCR-based	Yes	-
Petrovitch 2008	Blood	-	Stated	-	-	PCR-based	-	-
Peuralinna 2011	Blood	-	-	-	-	PCR-based	Yes	-
Sonnen 2010	Blood	-	-	-	-	PCR-based	-	-
Tiraboschi 2004	Brain	-	-	-	-	PCR-based	-	-
Zipser 2007	-	-	-	-	-	PCR-based	No	-

Dark grey shading: studies providing data for the meta-analyses; light grey shading: studies providing a qualitative statement about the ε4 allele association with CAA; no shading: studies known to have *APOE* genotype and CAA data but providing no data about association between *APOE* genotype and CAA; "-": data not available; n/a: not applicable; HWE: Hardy-Weinberg equilibrium; PCR: polymerase chain reaction; [1] Number of participants in whom genotyping was attempted and number in whom it was successful; [2] Blinding: blinding of genotypers to pathology data.

Table 3.3 Pathology assessment characteristics of studies included in the systematic review of *APOE* genotype associations with CAA

First author, publication year	Method for assessing CAA				Consistency[1]	Rater qualification[2]	Locations rated[3]	Grading system[4]	System reliability[5]	Blinding to genotype[6]	Blinding to clinical information[7]
	IHC	CR	HE	ThS							
Alafuzoff 1999	✓				Yes	-	P; LM;	Other	-	-	-
Attems 2005	✓				Yes	-	F; Ot; O; LM;	Olichney Attems	-	-	-
Caselli 2010			some	some	Yes	-	F; T; P; O;	Other	-	-	-
Chalmers 2003	✓				Yes	Neuro-pathologist	F; T; P; LM;	Olichney	-	-	-
Christoforidis 2005	✓				Yes	-	T; O; LM;	Other	-	-	-
Chui 2006	some	✓	✓		Yes	Neuro-pathologist	F; T; P; O; Ot;	Vonsattel	-	Yes	Yes
Cruz-Sánchez 2000	✓	✓			Yes	-	Ot; O; T;	n/a	n/a	-	-
Davidson 2006	✓				Yes	-	F	Other	"Good"[8]	-	-
Greenberg 1995		✓			Yes	Neuro-pathologist	F; P; T; Ot; O; C;	Vonsattel	-	Yes	-
Leclercq 2005	✓	✓			Yes	-	F; T; LM;	Other	-	Yes[9]	Yes[9]
Mortimer 2009	✓				Yes	Neuro-pathologist	F; P; T; O; Ot; LM;	Other	-	-	Yes
Nicoll 2011		✓			Yes	Neuro-pathologist	F; T; P; O; Ot; LM;	Other	-	-	Yes
Olichney 2000			✓	✓	Yes	-	Ot; F; P; T; LM;	Olichney	-	-	-
Pfeifer 2002	✓				Yes	Neuro-pathologist	F; T; P; O; LM;	After Vinters	-	-	Yes
Premkumar 1996	✓	some		some	Yes	Neuro-pathologist[10]	F; T; P; O;	Other	-	Yes	-

First author, publication year	Method for assessing CAA				Consistency[1]	Rater qualification[2]	Locations rated[3]	Grading system[4]	System reliability[5]	Blinding to genotype[6]	Blinding to clinical information[7]
	IHC	CR	HE	ThS							
Schneider 2005	✓				Yes	Neuro-pathologist	F; O; LM;	After Attems	-	-	-
Tanskanen 2005	some	✓			Yes	-	F; T; P; Ot; C; LM;	Other	-	-	-
Thal 2002	✓				Yes	-	T; O; LM	Other	-	-	-
Walker 2000	✓				Yes	-	Ot; T; LM;	Other	-	Yes	-
Yamada 2002	some	✓			Yes	-	O; LM;	Other	-	-	-
Yamaguchi 2001	✓				Yes	-	F	-	n/a	-	-
Yip 2005	✓				Yes	Neuro-pathologist	O	Other	-	-	-
Zarow 1999				✓	Yes	Neuro-pathologist	F; P; O; T; Ot; LM;	Other	intra-rater: κ=0.76; inter-rater: κ=0.77	-	-
Zubenko 1994		✓			Yes	Neuro-pathologist	-	n/a	n/a	-	-
Berg 1998				✓	Yes	-	Unknown for CAA but F; T; P; LM;	Other	-	-	-
Jicha 2008			✓		Yes	-	Unknown for CAA but F; T; P; O; C; Ot; LM;	Other	-	-	-
Oyama 1995	✓[11]				Yes	-	F; T; O;	Other	-	-	-
Roher 2003				✓	Yes	-	LM; P;	Other	-	-	-
Stopa 2008	✓				Yes	-	F	Other	-	Yes	-
Thomas 2000	✓				Yes	Image analysis	T	Other	-	-	-

First author, publication year	Method for assessing CAA				Consistency[1]	Rater qualification[2]	Locations rated[3]	Grading system[4]	System reliability[5]	Blinding to genotype[6]	Blinding to clinical information[7]
	IHC	CR	HE	ThS							
Alafuzoff 2009	✓				Yes	-	F; T; P;	n/a	n/a	-	-
Attems 2008	✓				Yes	-	F; Ot; O; LM;	Olichney	-	-	Yes
Durany 2000	✓	✓	✓		Yes	-	O; T; Ot; LM;	Other	-	-	-
Etiene 1998	✓				Yes	-	T; LM;	Other	-	-	-
Greenberg 1998	some	✓	✓		Yes	-	Cerebral cortex and cerebellum	Vonsattel	-	Yes	-
Honig 2005	-				-	-	-	-	-	-	-
Lashley 2008	✓				Yes	Neuro-pathologist	F; T; P; O; LM;	Olichney	-	-	-
Lewis 2006	✓				Yes	Image analysis	T; F;	Other	-	-	-
Love 2003	✓				Yes	-	F; T;	Olichney	-	-	-
Nelson 2010	✓		✓		Yes	-	Not known for CAA	-	n/a	-	-
Pahnke 2003	✓				Yes	-	T; Ot; LM;	Other	-	Yes	-
Petrovitch 2008	✓				Yes	Neuro-pathologist	F; T; P; O; Ot; LM;	Other	-	-	-
Peuralinna 2011	✓	✓			Yes	-	F; P; T; O; Ot; C	Other	-	-	-
Sonnen 2010	-				-	Neuro-pathologist	-	-	-	-	-
Tiraboschi 2004				✓	Yes	-	F; P; T; Ot; LM;	Other	-	-	-
Zipser 2007	✓				Yes	-	F	-	-	-	-

Dark grey shading: studies providing data for the meta-analyses; light grey shading: studies providing a qualitative statement about the ε4 allele association with CAA; no shading: studies known to have *APOE* genotype and CAA data but providing no data about association between *APOE* genotype and CAA; "-": data not available;

n/a: not applicable; IHC: immunohistochemistry; CR: congo red; HE: Hematoxylin and Eosin; ThS: thioflavin S; [1] Consistency: were the same pathological assessment methods used consistently? (i.e. was the same method used throughout the study or did it vary with time, recruiting centre etc.); [2] Rater qualification: qualification of the person assessing CAA; [3] Locations rated: F=frontal cortex; T=temporal cortex; P=parietal cortex; O=occipital cortex; Ot=other; C=cerebellar cortex; LM=leptomeninges (assumed leptomeninges were assessed when the scale used in the study included assessment of leptomeningeal vessels); [4] Grading system: scale or severity grading system or amyloid quantification system used in the study; [5] System reliability: inter- and intra-rater reliability of the grading system used; [6] Blinding to genotype: blinding of CAA raters to genotyping results; [7] Blinding to clinical information: blinding of CAA raters to relevant clinical information such as participants' dementia and Alzheimer's disease status, history of intracerebral haemorrhage, age; [8] Inter-rater reliability was looked at in a previous study and described as "good" but using methods that were unclear; [9] Leclercq 2005 – only IHC results assessed blind to genotype and clinical information; [10] Premkumar 1996 – some rated by neuropathologist, not known for some; [11] Oyama 1995 – assumed IHC was used based on a reference to a previous study on same participants; If an included study referenced another study regarding how pathology assessment was done, then I generally assumed the method and scale used were the same, however, other details of pathology assessment had to be mentioned in the original study.

Table 3.4 Quality scores of studies included in the meta-analyses of *APOE* genotype associations with CAA

Study first author, publication year	Study size			Blinding[1]			Study design[2]		Data format[3]		Pathology assessment[4]			Genotyping reporting[5]		N[6]	Total score
	<140	140-239	>239	none	any	complete	R	P	D	C							
	0	1	2	0	1	2	0	1	0	1	0	1	2	0	1		(0-9)
Alafuzoff 1999		✓		✓			✓			✓		✓			✓	209	4
Attems 2005	✓			✓			✓		✓			✓			✓	53	2
Caselli 2010		✓		✓				✓		✓		✓		✓		179	4
Chalmers 2003	✓			✓			✓			✓			✓		✓	120	4
Christoforidis 2005	✓			✓			✓		✓			✓			✓	118	2
Chui 2006	✓				✓			✓	✓				✓	✓		73	4
Cruz-Sanchez 2000	✓			✓			✓		✓			✓			✓	73	2
Davidson 2006		✓		✓			✓			✓		✓			✓	146	4
Greenberg 1995	✓				✓		✓		✓				✓		✓	93	4
Leclercq 2005	✓				✓		✓		✓			✓			✓	88	3
Mortimer 2009			✓		✓			✓		✓			✓		✓	267	8
Nicoll 2011			✓		✓			✓	✓				✓		✓	310	7
Olichney 2000			✓	✓			✓			✓		✓			✓	247	5
Pfeifer 2002		✓			✓			✓	✓				✓		✓	201	6
Premkumar 1996			✓		✓		✓		✓				✓		✓	240	6
Schneider 2005		✓		✓				✓	✓				✓		✓	208	5
Tanskanen 2005	✓			✓				✓	✓			✓			✓	71	3
Thal 2002	✓			✓			✓		✓			✓			✓	56	2

Study first author, publication year	Study size			Blinding[1]			Study design[2]		Data format[3]		Pathology assessment[4]			Genotyping reporting[5]		N[6]	Total score
	<140	140-239	>239	none	any	complete	R	P	D	C							
	0	1	2	0	1	2	0	1	0	1	0	1	2	0	1		(0-9)
Walker 2000			✓		✓		✓			✓		✓			✓	244	6
Yamada 2002		✓		✓			✓			✓		✓			✓	201	4
Yamaguchi 2001	✓			✓			✓		✓			✓		✓		101	1
Yip 2005	✓			✓			✓			✓		✓		✓		99	2
Zarow 1999	✓			✓			✓			✓			✓		✓	42	4
Zubenko 1994	✓			✓			✓		✓			✓			✓	91	2

[1] Blinding: genotypers to pathology data, CAA raters to genotyping data, CAA raters to clinical information (0: no blinding, 1: blinding at least one way, 2: blinding all three ways); [2] Study design: prospective [P] or retrospective [R]; [3] Data format: dichotomous [D] or continuous [C]; [4] Pathology assessment: points assigned based on whether method for assessing CAA reported, neuropathologist-rated CAA, >1 brain location examined (0: no criteria fulfilled; 1: one to two criteria fulfilled; 2: all three criteria fulfilled); [5] Genotyping reporting: points assigned based on whether source of DNA and genotyping method reported (0: none or one criteria reported; 1: both criteria reported); [6] N: number of participants.

Table 3.5 Summary of meta-analyses of association between *APOE* genotype and presence versus absence of CAA

Comparison		CAA present versus CAA absent	Number of dichotomous studies: n/N (ε4+ /ε2+) versus n/N (ε4-/ε2-/ε3ε3)
ε4+ versus ε4- genotypes	FE	OR 2.9 (95% CI 2.5 to 3.3); p<0.00001	14 studies: 382/594 versus 388/1100
	RE	OR 3.1 (95% CI 2.3 to 4.1); p<0.00001	
ε4+ versus ε3/ε3 genotypes	FE	OR 2.7 (95% CI 2.2 to 3.3); p<0.00001	9 studies: 262/405 versus 209/672
	RE	OR 3.3 (95% CI 2.0 to 5.5); p<0.00001	
ε2+ versus ε2- genotypes	FE	OR 0.6 (95% CI 0.5 to 0.9); p=0.003	8 studies: 40/134 versus 372/871
	RE	OR 0.7 (95% CI 0.4 to 1.1); p=0.1	
ε2+ versus ε3/ε3 genotypes	FE	OR 0.9 (95% CI 0.7 to 1.3); p=0.55	8 studies: 40/134 versus 147/512
	RE	OR 1.0 (95% CI 0.6 to 1.6); p=0.95	

n: number of participants with CAA present; N: total number of participants; Dark grey shading: statistically significant results; FE: results using the fixed-effects method; RE: results using the random-effects method.

Table 3.6 Subgroups for meta-analyses of *APOE* genotype associations with CAA based on participants' clinical dementia and Alzheimer's disease status

First author and publication year	Participants' description by subgroup
Alafuzoff 1999	**D**: 106 participants with clinical and pathological (CERAD) diagnosis of AD **ND**: 103 clinically not demented participants
Attems 2005	**M**: 20/53 participants had a clinical and pathological diagnosis of AD, 33/53 not diagnosed with AD
Caselli 2010	**ND**: 158 clinically not demented participants and 21 participants with mild cognitive impairment
Chalmers 2003	**D**: 86 participants clinical and pathological (CERAD) diagnosis of AD **ND**: 34 participants with no clinical or pathological diagnosis of AD
Christoforidis 2005	**M**: 8/125 participants had a clinical and pathological diagnosis of AD. Data provided for only 116 of the 125 participants
Chui 2006	**M**: Severe dementia participants were excluded. 13/79 were not demented, 13/79 had cognitive impairment, 53/79 were clinically demented. Data provided for 73 participants
Cruz-Sánchez 2000	**D**: 35 participants: 28/35 participants with clinical and pathological (Braak) diagnosis of AD; 7/35 participants with clinical and pathological diagnosis of small vessel disease dementia **ND**: 38 clinically not demented participants
Davidson 2006	**D**: 146 participants with clinical and pathological (CERAD) diagnosis of AD
Greenberg 1995	**M**: 88 clinically demented and non-demented participants
Leclercq 2005	**M**: 88 participants whose dementia status not known
Mortimer 2009	**M**: 135/267 participants were clinically demented at death, 166/267 met the study´s neuropathologic criteria for AD
Nicoll 2011	**M**: 151/308 participants had a clinical diagnosis of dementia; 135/308 did not have clinical dementia; dementia status for 22/308 not known
Olichney 2000	**D**: 246 participants with clinical and pathological (Khachaturian and CERAD) diagnosis of AD

Pfeifer 2002	**M**: 138/211 participants were clinically not demented, 22/211 participants had a clinical and pathological diagnosis of AD, 27/211 participants diagnosed with vascular dementia, 24/211 mixed or other dementia. Data provided for 201/211 participants
Premkumar 1996	**M**: 34 participants with neurological conditions other than AD (i.e. ALS, CJD), questionable dementia **D**: 190 participants with clinical and pathological (Khachaturian and CERAD) diagnosis of AD **ND**: 16 clinically not demented participants
Schneider 2005	**M**: 102/214 participants clinically diagnosed with AD. Data provided for 208 participants.
Tanskanen 2005	**D**: 47 participants clinically demented **ND**: 24 clinically not demented participants
Thal 2002	**D**: 15 participants with clinical and pathological (Braak, NIA-Reagan Institute criteria) diagnosis of AD **ND**: 41 clinically not demented participants
Walker 2000	**M**: 244 participants, cases with clinical evidence of overt dementia were excluded
Yamada 2002	**M**: 82/201 participants had a pathologically confirmed diagnosis of AD
Yamaguchi 2001	**M**: 101 participants who had not been formally assessed for dementia, but showed no signs of dementia in daily life
Yip 2005	**D**: 99 participants with clinical and pathological (CERAD, NIA-Reagan Institute criteria) diagnosis of AD
Zarow 1999	**D**: 42 participants with clinical and pathological (Khachaturian, CERAD) diagnosis of AD
Zubenko 1994	**D**: 91 participants with clinical and pathological (Khachaturian) diagnosis of AD

D: clinically demented participants; ND: clinically not demented participants (though may have some AD-related changes on neuropathology); M: mixed demented and not demented participants (studies where data could not be extracted for demented and not demented participants separately);

Table 3.7 Summary of subgroup analyses results for meta-analyses of association of *APOE* ε4+ versus ε4- genotypes with CAA

Subgroup		Fixed-effects method		Random-effects method	
		OR (95% CI)	Subgroup difference	OR (95% CI)	Subgroup difference
Dementia status	Clinically demented participants	2.37 (1.85-3.05)	$p_{diff} = 0.38$	2.74 (1.53-4.93)	$p_{diff} = 0.89$
	Clinically not demented participants	3.32 (2.16-5.11)		3.25 (1.88-5.63)	
	Mixed participants	2.74 (2.25-3.33)		3.23 (2.20-4.74)	
Alzheimer's disease (AD) status	Participants with pathologically confirmed AD	2.25 (1.75-2.90)	$p_{diff} = 0.25$	2.33 (1.29-4.24)	$p_{diff} = 0.64$
	Clinically not demented participants	3.32 (2.16-5.11)		3.25 (1.88-5.63)	
	Mixed participants	2.74 (2.25-3.33)		3.23 (2.20-4.74)	
Ethnicity	Asian participants	2.08 (1.33-3.25)	$p_{diff} = 0.13$	2.25 (1.18-4.28)	$p_{diff} = 0.33$
	Caucasian participants	3.00 (2.58-3.48)		3.21 (2.36-4.38)	
Study quality score	0-2 points	2.48 (1.69-3.63)	$p_{diff} = 0.18$	2.48 (1.69-3.63)	$p_{diff} = 0.43$
	3-4 points	2.61 (2.10-3.24)		2.74 (1.83-4.11)	
	5-9 points	3.37 (2.71-4.18)		4.00 (2.15-7.45)	
Study size	< 140	2.82 (2.20-3.62)	$p_{diff} = 0.45$	2.89 (2.12-3.95)	$p_{diff} = 0.8$
	140-239	2.58 (1.98-3.35)		2.74 (1.63-4.60)	
	>239	3.21 (2.56-4.04)		3.72 (1.73-8.02)	
Blinding	Any blinding	3.28 (2.64-4.06)	$p_{diff} = 0.13$	3.85 (2.13-6.96)	$p_{diff} = 0.3$
	No blinding	2.62 (2.17-3.17)		2.71 (2.02-3.65)	

Study design	Prospective study design	3.66 (2.87-4.68)	p_{diff} = 0.02	3.97 (2.63-5.99)	p_{diff} = 0.18
	Retrospective study design	2.56 (2.15-3.05)		2.72 (1.89-3.92)	
Data format	Continuous data format	2.48 (2.06-2.98)	p_{diff} = 0.009	2.43 (1.59-3.72)	p_{diff} = 0.11
	Dichotomous data format	3.64 (2.91-4.55)		3.84 (2.68-5.50)	
Pathology assessment quality	1 point	2.22 (1.84-2.67)	p_{diff} <0.0001	2.40 (1.79-3.22)	p_{diff} = 0.05
	2 points	4.20 (3.37-5.24)		4.16 (2.62-6.62)	
Genotyping reporting quality	0 points	2.29 (1.64-3.19)	p_{diff} = 0.13	2.29 (1.64-3.19)	p_{diff} = 0.12
	1 point	3.04 (2.60-3.56)		3.33 (2.37-4.67)	

OR =odds ratio; CI = confidence interval; p_{diff} = subgroup difference p-value.

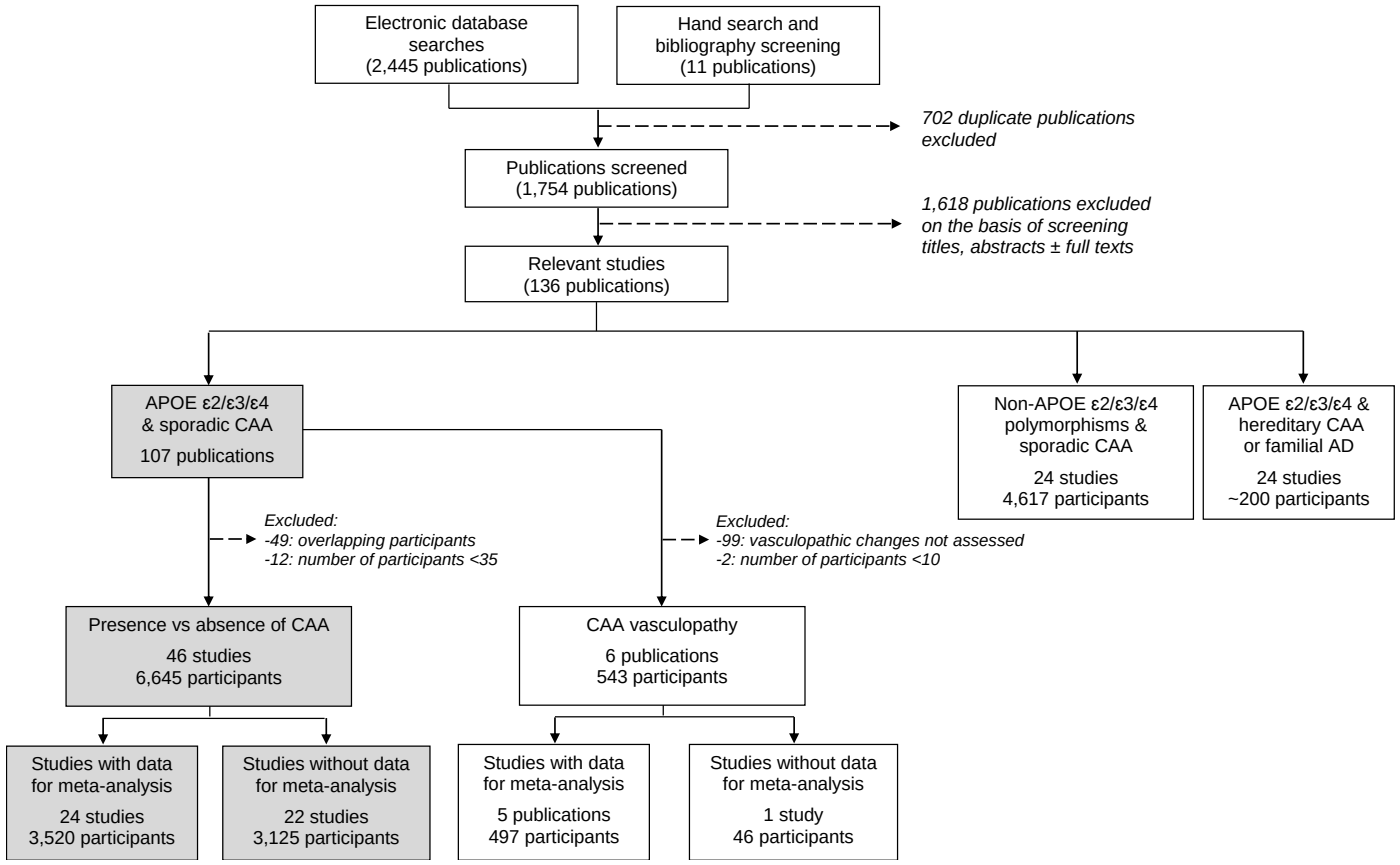
Table 3.8 Summary of studies providing qualitative data about the association between *APOE* genotype and CAA

First author, publication year	Participants´ dementia status	Number of participants[1]	Reported data about the association between CAA and *APOE* genotype
Berg 1998	M	176	*APOE* ε4 allele possession associated with increased severity of CAA in the neocortex but not in hippocampal or entorhinal regions
Jicha 2008	D	81	No overall association with *APOE* ε4 allele possession
Oyama 1995	-	37	*APOE* ε4 allele possession associated with a trend towards increased severity of CAA
Roher 2003	D	24	Homozygosity for the *APOE* ε4 allele associated with increased severity of CAA
Stopa 2008	M	75	No overall association with *APOE* genotype
Thomas 2000	D	50	No overall association with *APOE* ε4 allele possession
Greenberg 1995[2]	M	93	Homozygosity for the *APOE* ε4 allele associated with increased severity of CAA *APOE* ε2 allele possession associated with a trend towards less severe CAA
Chalmers 2003[2]	M	120	No association with homozygosity for the *APOE* ε2 allele

"-": data not available; M: mixed clinically demented and non-demented participants; D: clinically demented participants; [1] Number of participants in the study for whom association between *APOE* genotype and CAA assessed; [2] Study also provided data for meta-analysis of the association between *APOE* ε4 allele and CAA.

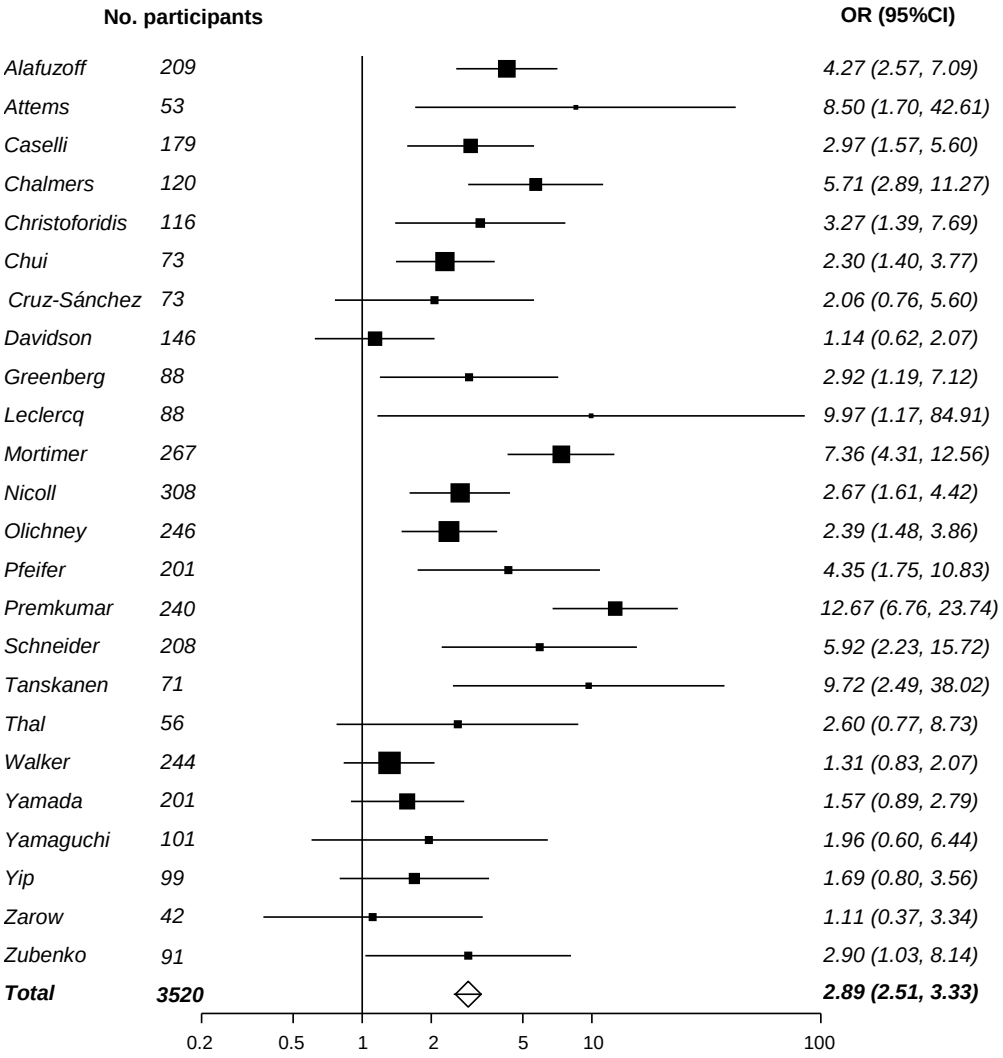
Figures

Figure 3.1 Selection of studies included in the systematic review of *APOE* genotype associations with CAA



Grey boxes represent the selection of studies relevant for Chapter 3.

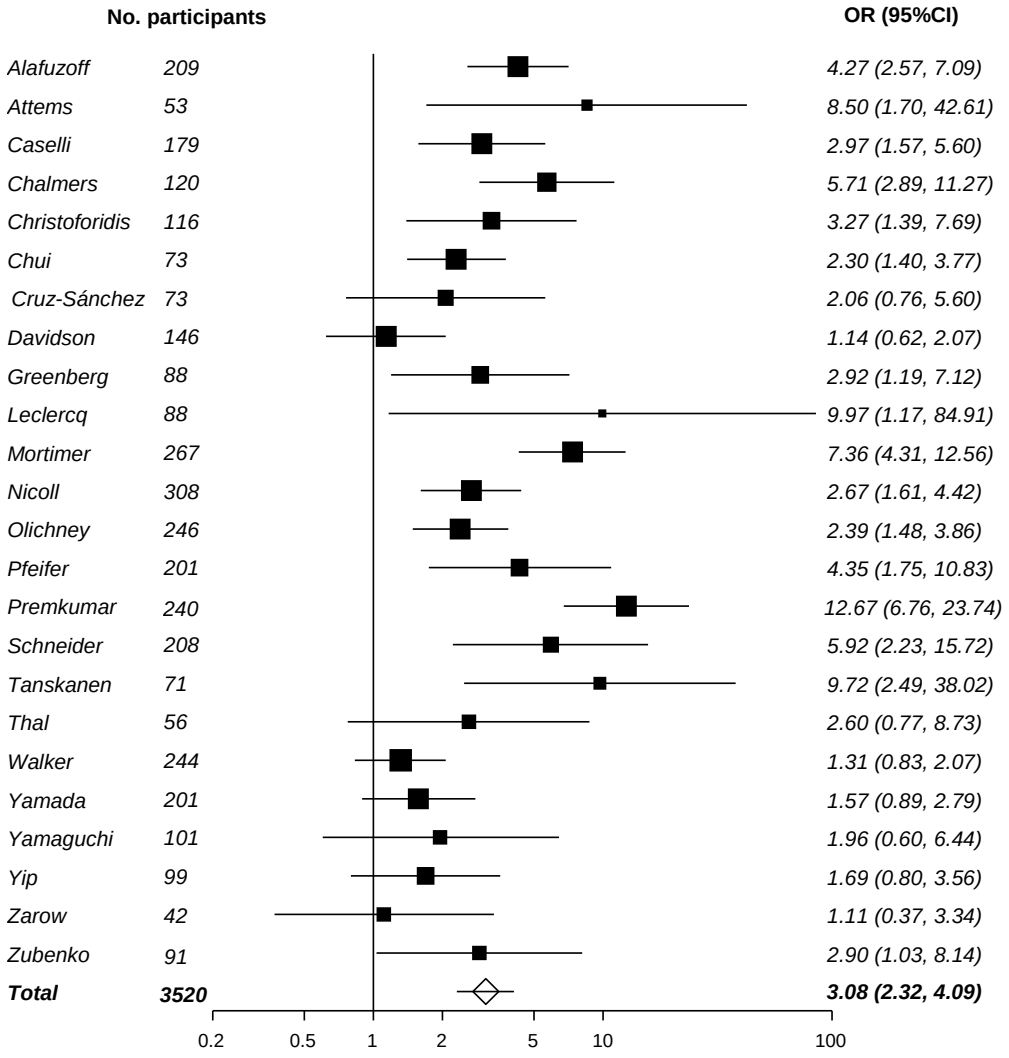
Figure 3.2 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA (fixed-effects method)



Generic inverse variance fixed-effects method; $p_{(\text{overall effect})}<0.00001$; Heterogeneity: $I^2=71\%$; $\chi^2_{23df}=80.68$; $p<0.00001$;

OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR.

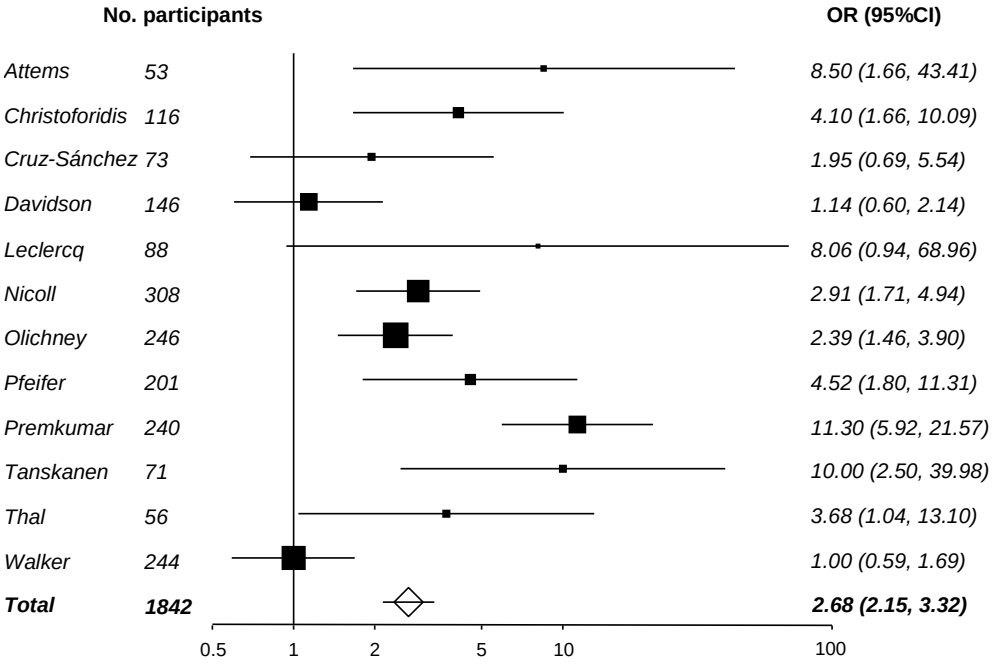
Figure 3.3 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA (random-effects method)



Generic inverse variance random-effects method; $p_{(\text{overall effect})}<0.00001$; Heterogeneity: $I^2=71\%$; $\chi^2_{23df}=80.68$; $p<0.00001$;

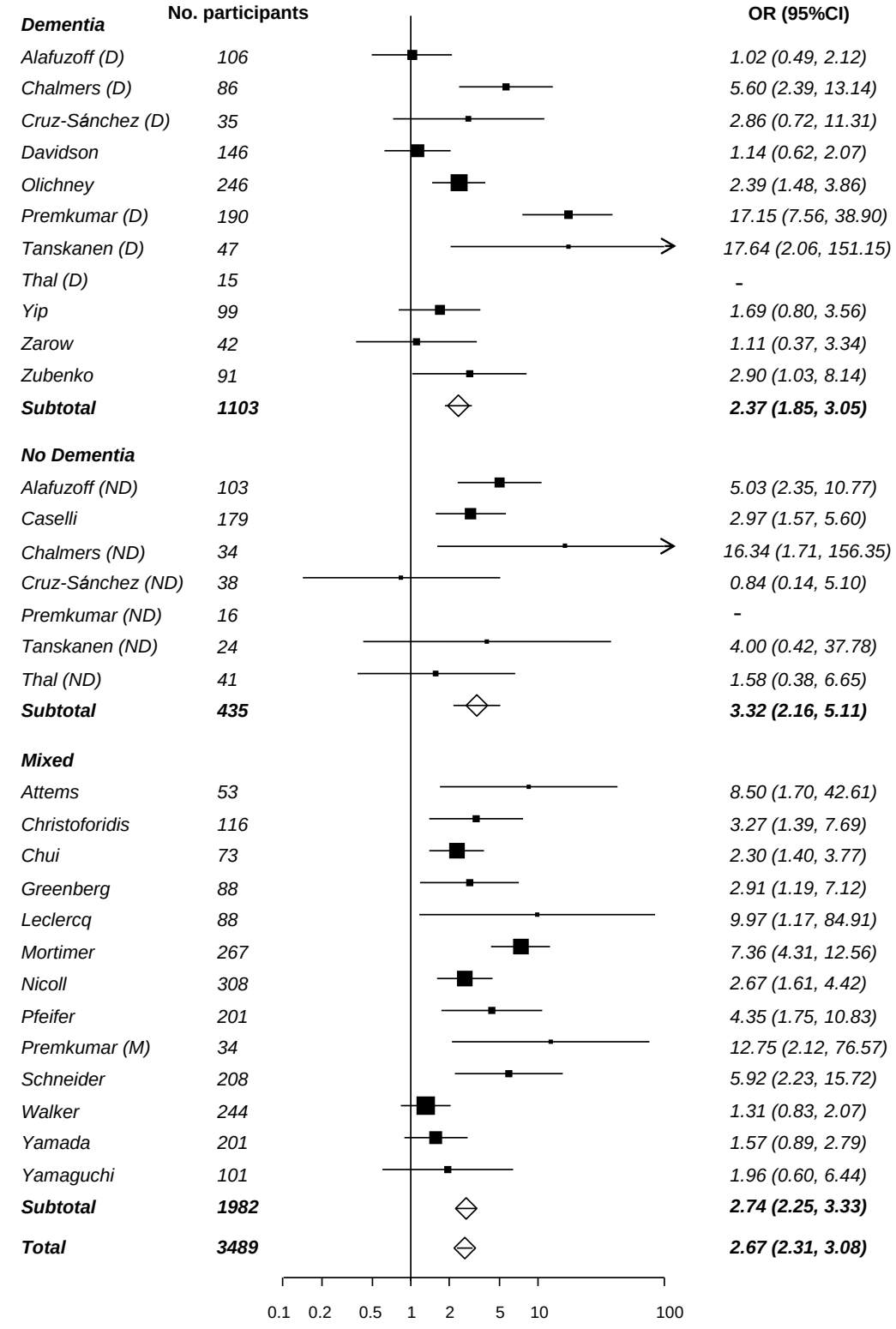
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the random-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR.

Figure 3.4 Meta-analysis of association of *APOE* ε4+ versus ε3/ε3 genotypes with CAA (fixed-effects method)



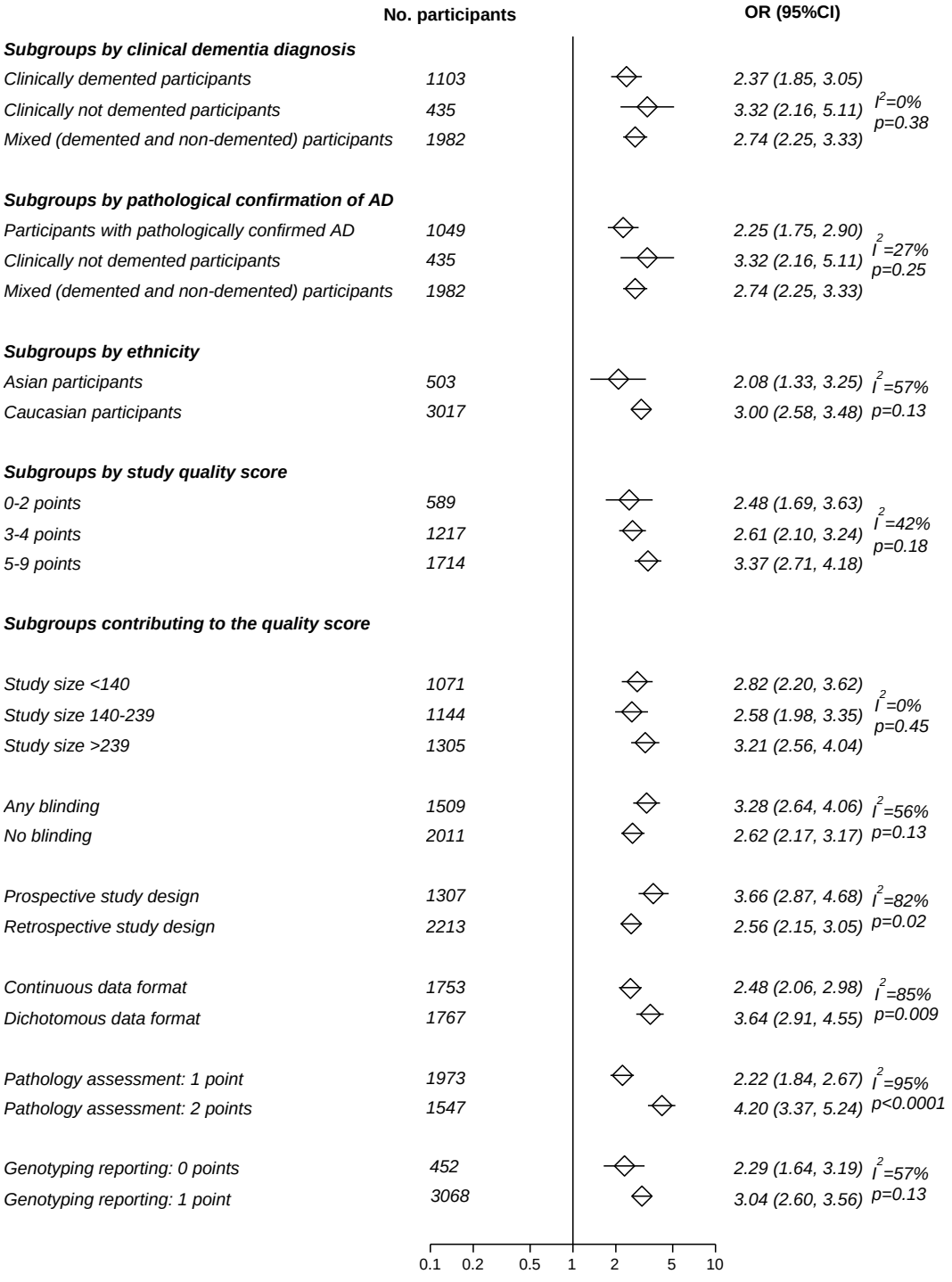
Generic inverse variance fixed-effects method; $p_{(\text{overall effect})} < 0.00001$; Heterogeneity: $I^2=78\%$; $\chi^2_{11df}=48.94$; $p<0.00001$; No. participants = total number of participants in the study (5 participants excluded because of missing data: Christoforidis 2 participants, Nicoll 2 participants, Olichney 1 participant); OR = odds ratio; CI = confidence interval; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects model; Horizontal lines represent 95% CIs. Diamond represents the pooled OR.

Figure 3.5 Meta-analysis of association of APOE ε4+ versus ε4- genotypes with CAA by clinical dementia status



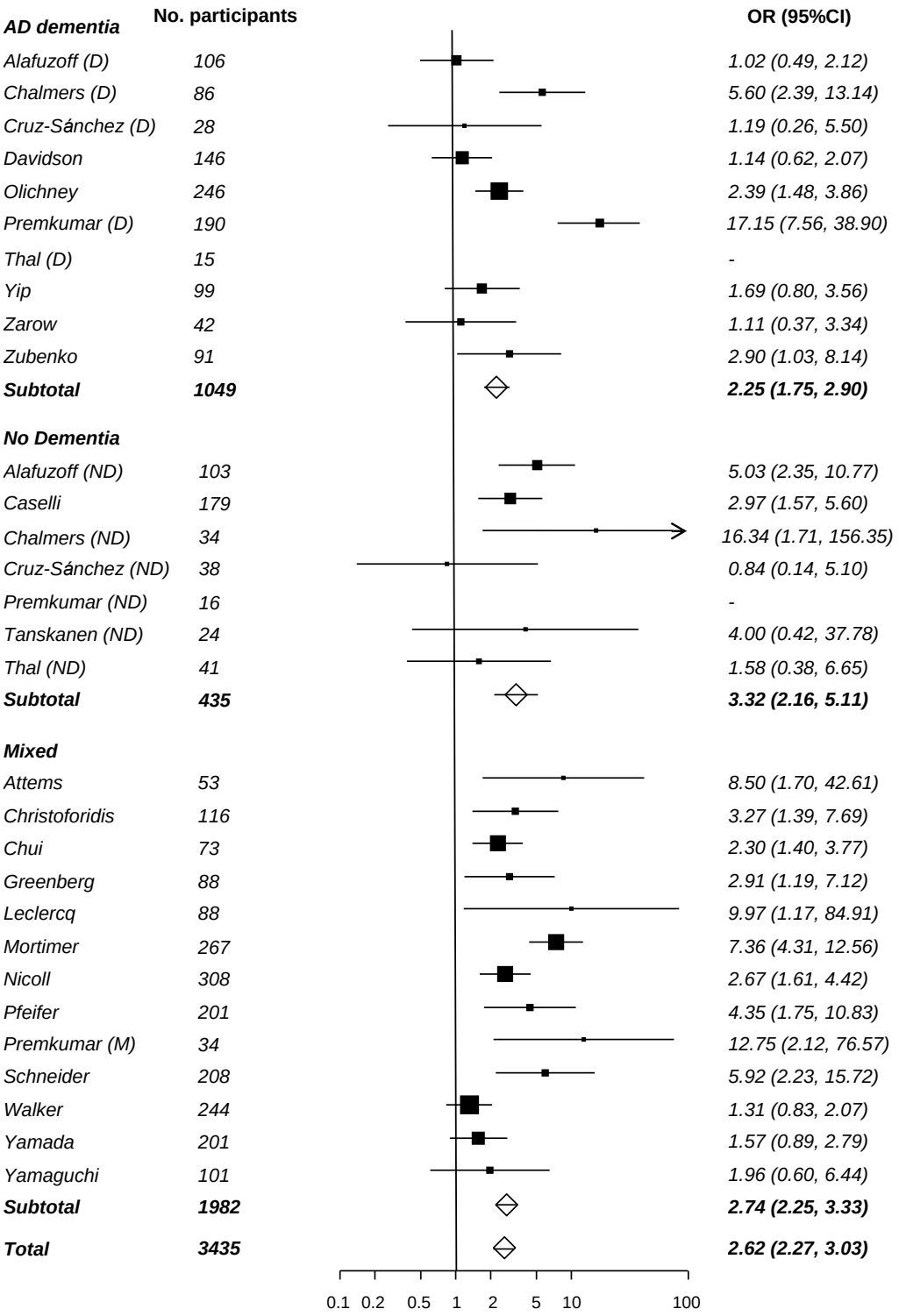
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Dementia: I^2=79%; χ^2_{9df}=43.42; p<0.00001; No dementia: I^2=22%; χ^2_{5df}=6.45; p=0.27; Mixed: I^2=68%; χ^2_{12df}=37.29; p=0.0002; Test for subgroup differences: I^2=0%; χ^2_{2df}=1.91; p_{diff}=0.38;
OR = odds ratio; CI = confidence interval; Dementia (D) = participants with clinical dementia; No dementia (ND) = participants with no clinical dementia; Mixed (M) = participants with and without clinical dementia; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); OR not estimable for 15 participants in Thal dementia group and for 16 participants in Premkumar no dementia group, because either all participants in both genotype groups had CAA, or no participant in either genotype group had CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.6 Summary of meta-analyses of association of *APOE* ε4+ versus ε4- genotypes with CAA by subgroups



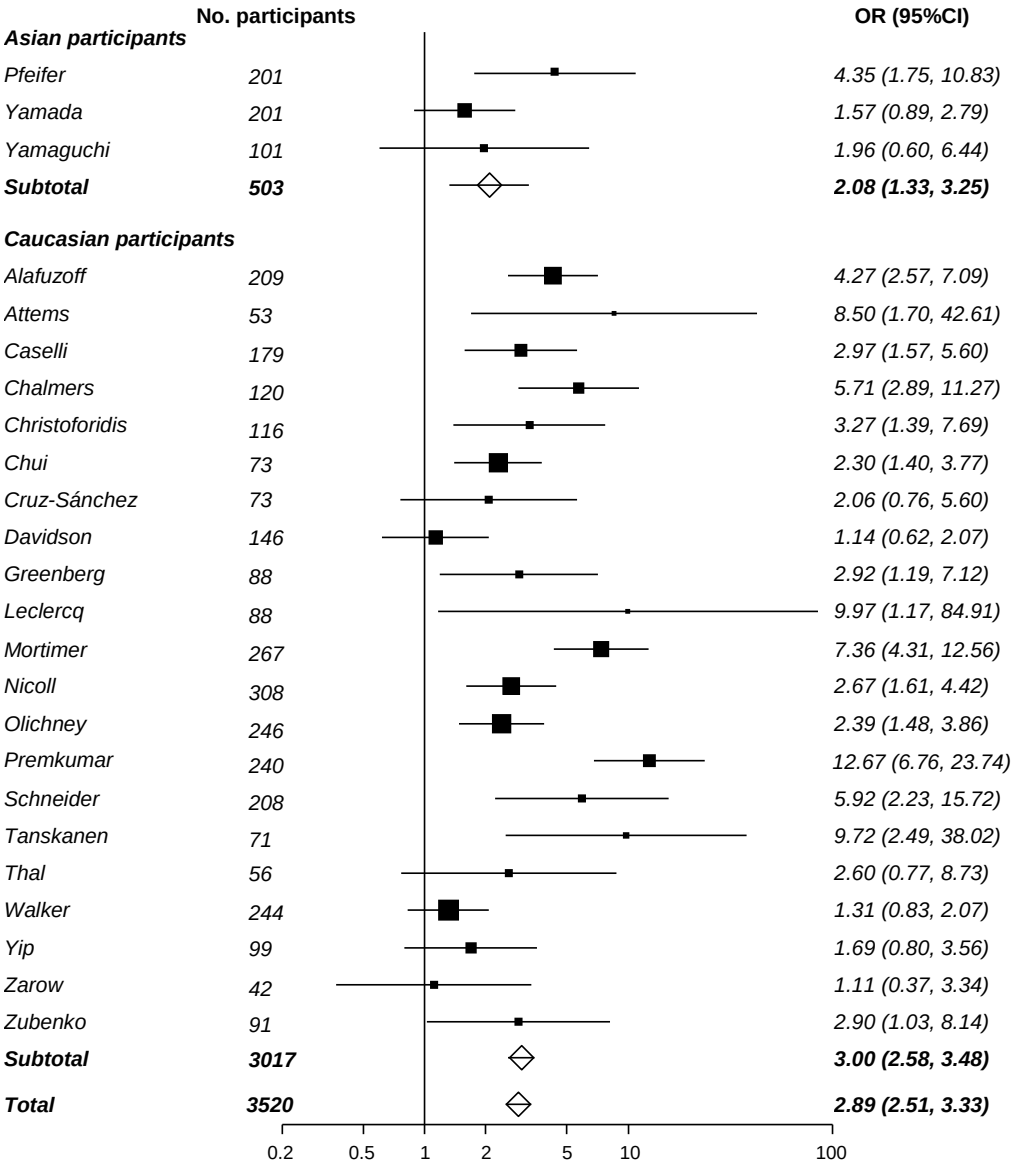
AD = Alzheimer's disease; OR = odds ratio; CI = confidence interval; Diamonds represent the pooled ORs; Horizontal lines represent 95% CIs; I^2 = heterogeneity between subgroups; p = statistical significance of between subgroup difference.

Figure 3.7 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by pathologically confirmed AD diagnosis



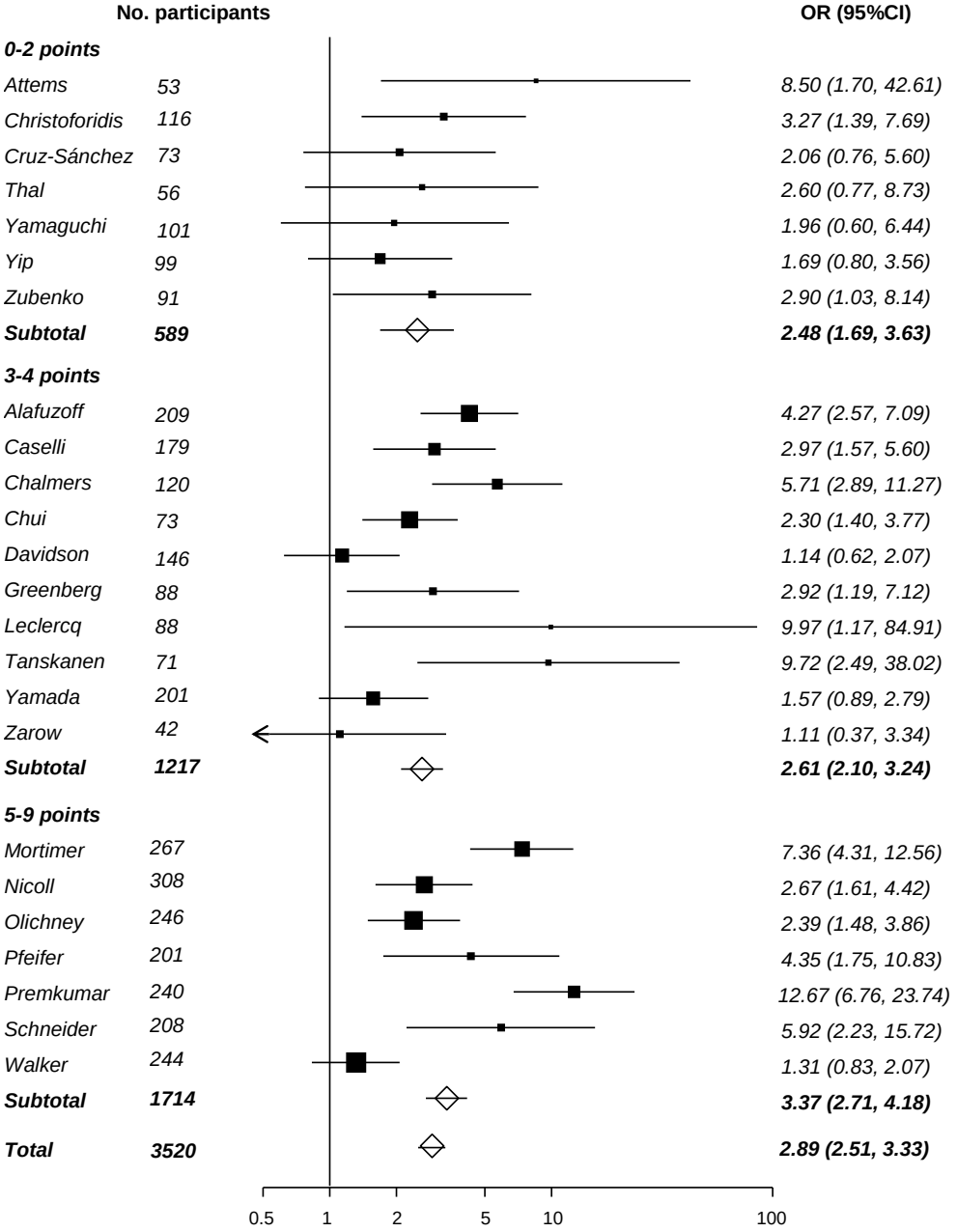
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: AD dementia I^2=80%; χ^2_{8df}=40.61; p<0.00001; No dementia: I^2=22%; χ^2_{5df}=6.45; p=0.27; Mixed: I^2=68%; χ^2_{12df}=37.29; p=0.0002; Test for subgroup differences: I^2=27.3%; χ^2_{2df}=2.75; p_{diff}=0.25;
OR = odds ratio; CI = confidence interval; AD = Alzheimer's disease; D = neuropathologically confirmed AD participants; ND = clinically not demented participants; M = demented and not demented participants; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); OR not estimable for 15 participants in Thal dementia group and for 16 participants in Premkumar no dementia group because either all participants in both genotype groups had CAA, or no participant in either genotype group had CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs. Diamonds represent the pooled ORs.

Figure 3.8 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by ethnicity



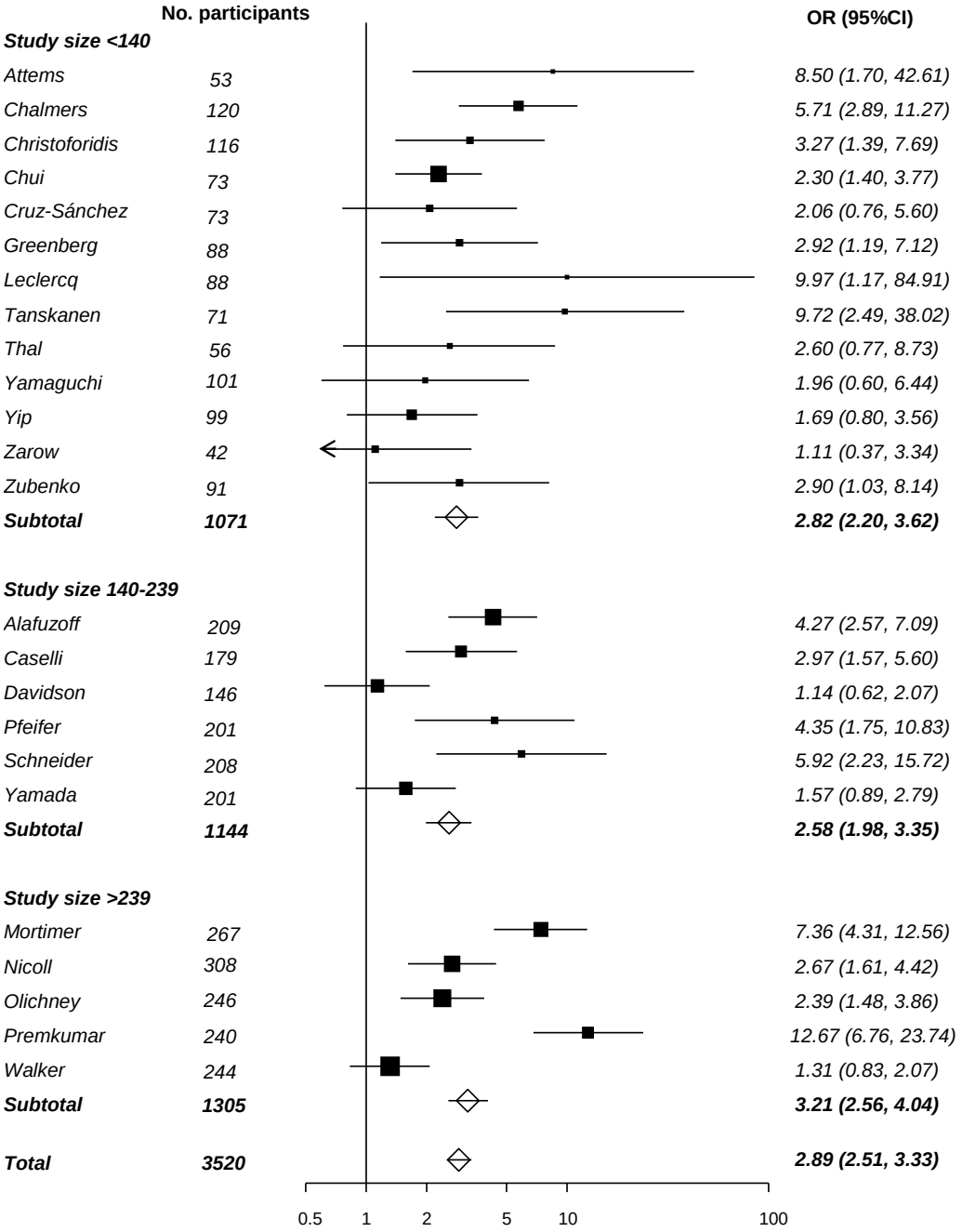
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Asian participants: I^2=42%; χ^2_{2df}=3.43; p=0.18; Caucasian participants: I^2=73%; χ^2_{20df}=74.92; p<0.00001; Test for subgroup differences: I^2=57.1%; χ^2_{1df}=2.33; p_{diff}=0.13;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.9 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by study quality score



Generic inverse variance fixed-effects method; Within subgroup heterogeneity: 0-2 points: I^2=0%; χ^2_{6df}=4.04; p=0.67; 3-4 points: I^2=67%; χ^2_{9df}=26.97; p=0.001; 5-9 points: I^2=87%; χ^2_{6df}=46.23; p<0.00001; Test for subgroup differences: I^2=41.9%; χ^2_{2df}=3.44; p_{diff}=0.18;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

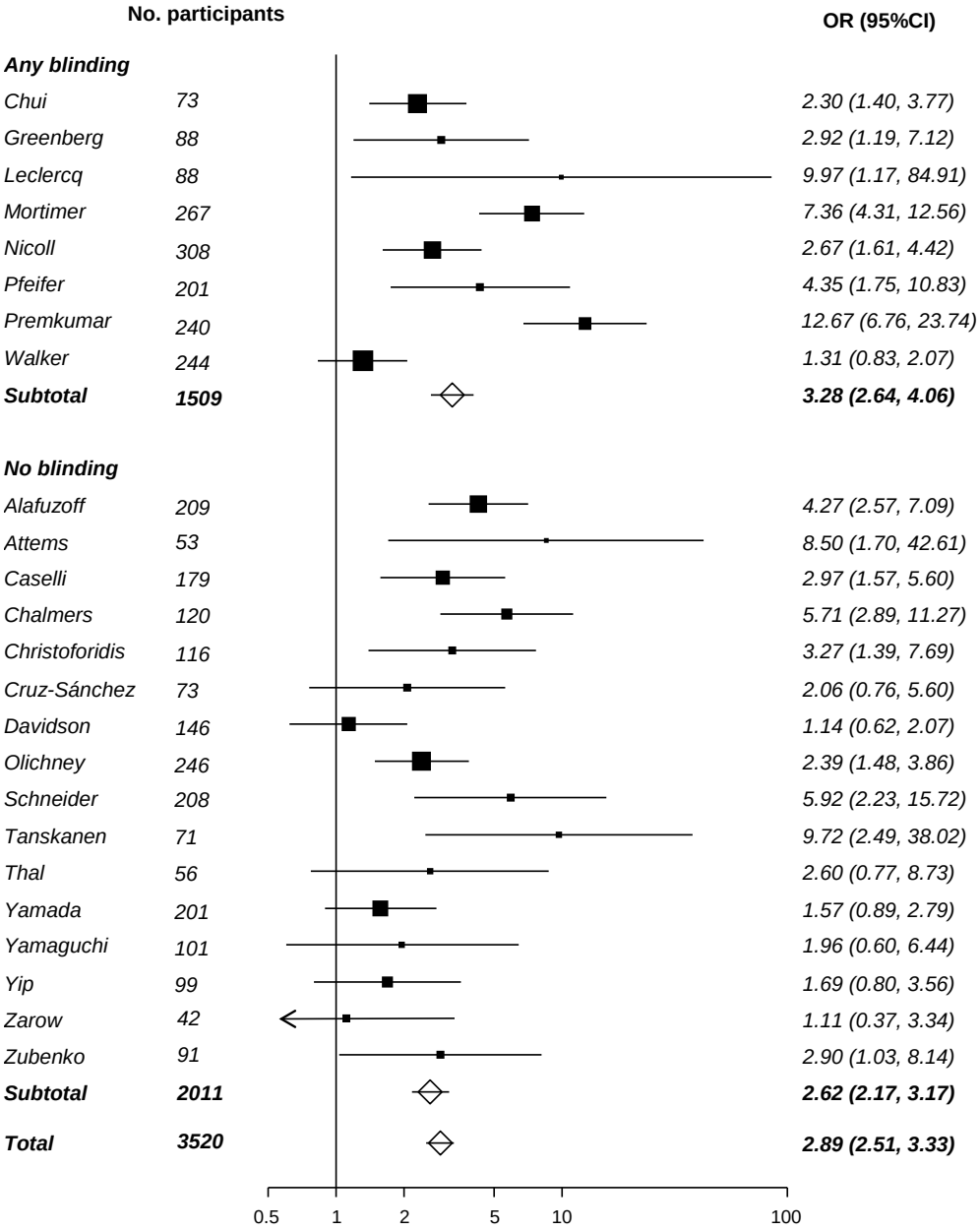
Figure 3.10 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by study size



Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Study size <140: I^2=27%; χ^2_{12df}=16.51; p=0.17; Study size 140-239: I^2=72%; χ^2_{5df}=18.10; p=0.003; Study size >239: I^2=91%; χ^2_{4df}=44.48; p<0.00001; Test for subgroup differences: I^2=0%; χ^2_{2df}=1.59; p_{diff}=0.45; OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-

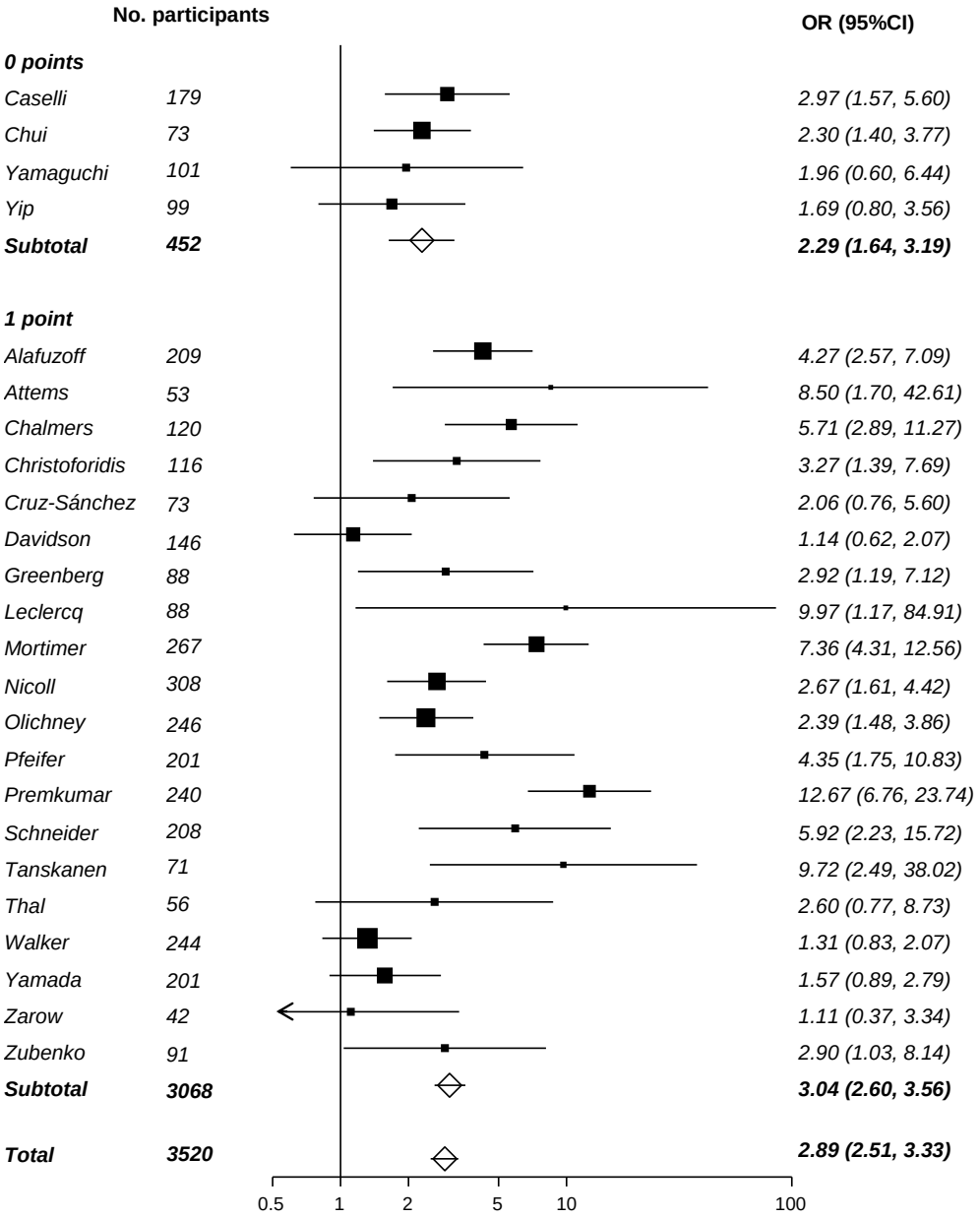
specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.11 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by blinding



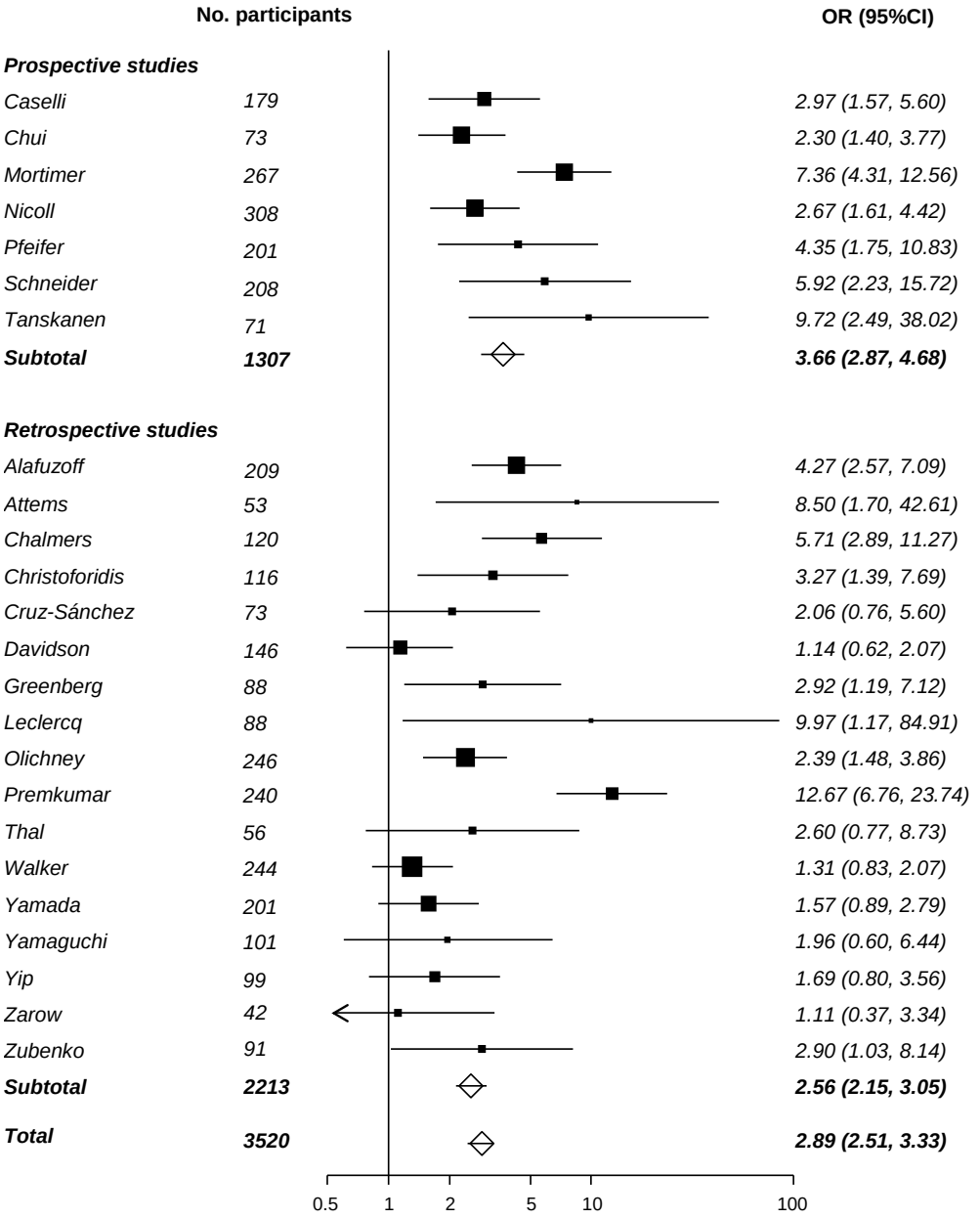
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Any blinding: I^2=85%; χ^2_{7df}=46.28; p<0.00001; No blinding: I^2=53%; χ^2_{15df}=32.10; p=0.006; Test for subgroup differences: I^2=56.4%; χ^2_{1df}=2.29; p_{diff}=0.13;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.12 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by genotyping reporting quality



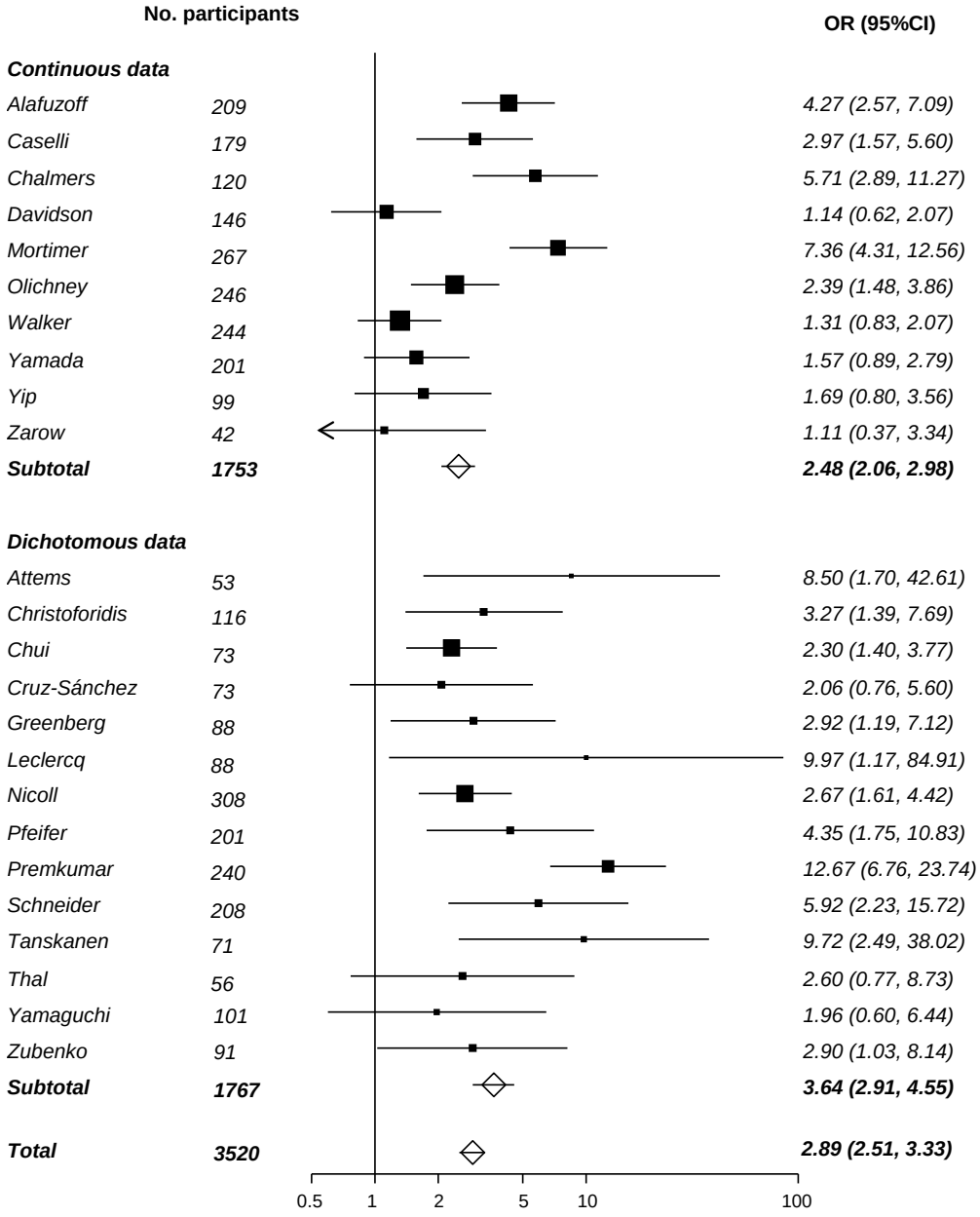
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Studies with 0 points: I^2=0%; χ^2_{3df}=1.34; p=0.72; Studies with 1 point: I^2=75%; χ^2_{19df}=77.03; p<0.00001; Test for subgroup differences: I^2=56.5%; χ^2_{1df}=2.30; p_{diff}=0.13; OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.13 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by study design



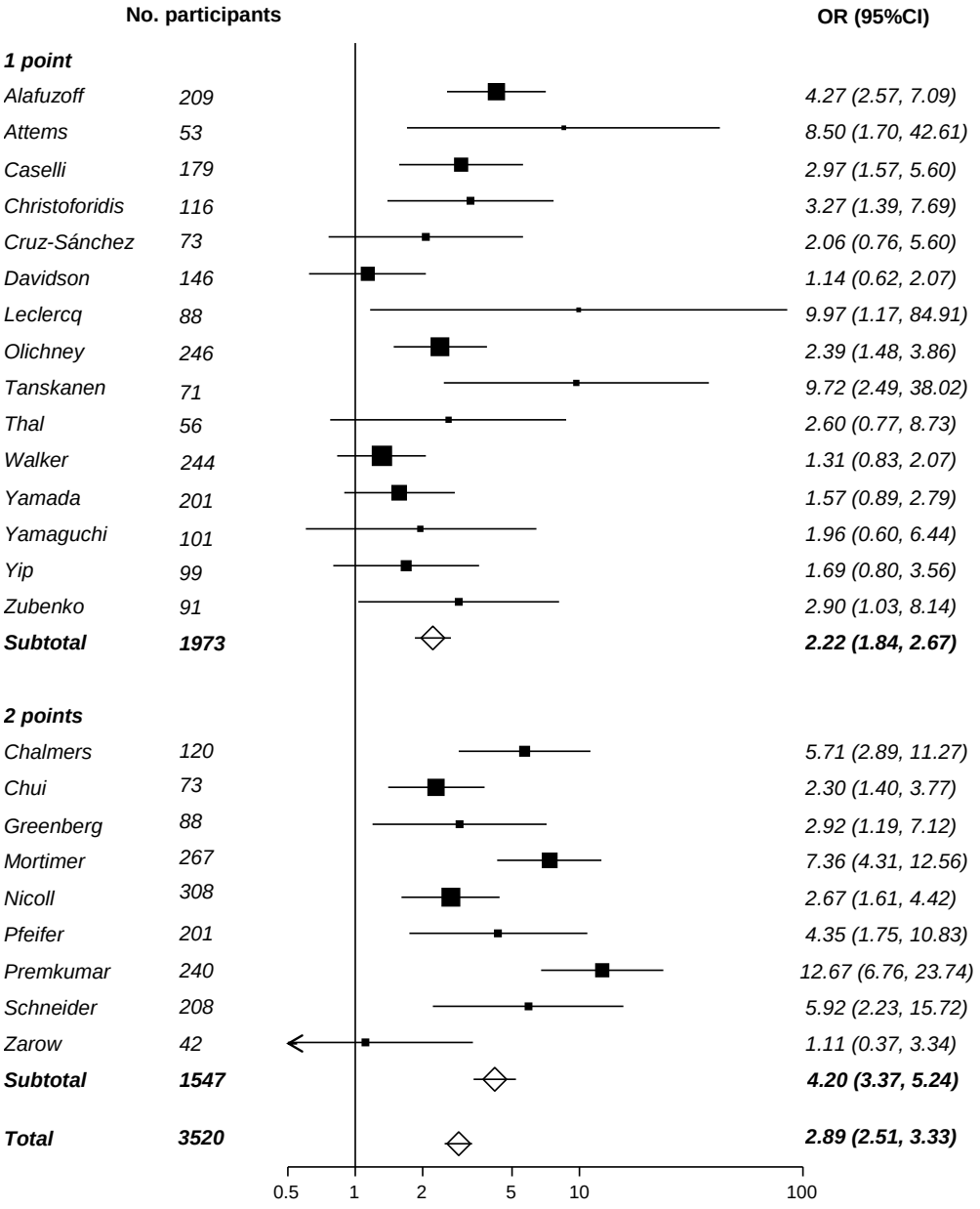
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Prospective studies: I^2=60%; χ^2_{6df}=14.88; p=0.02; Retrospective studies: I^2=73%; χ^2_{16df}=60.32; p<0.00001; Test for subgroup differences: I^2=81.8%; χ^2_{1df}=5.48; p_{diff}=0.02;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.14 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by study data format



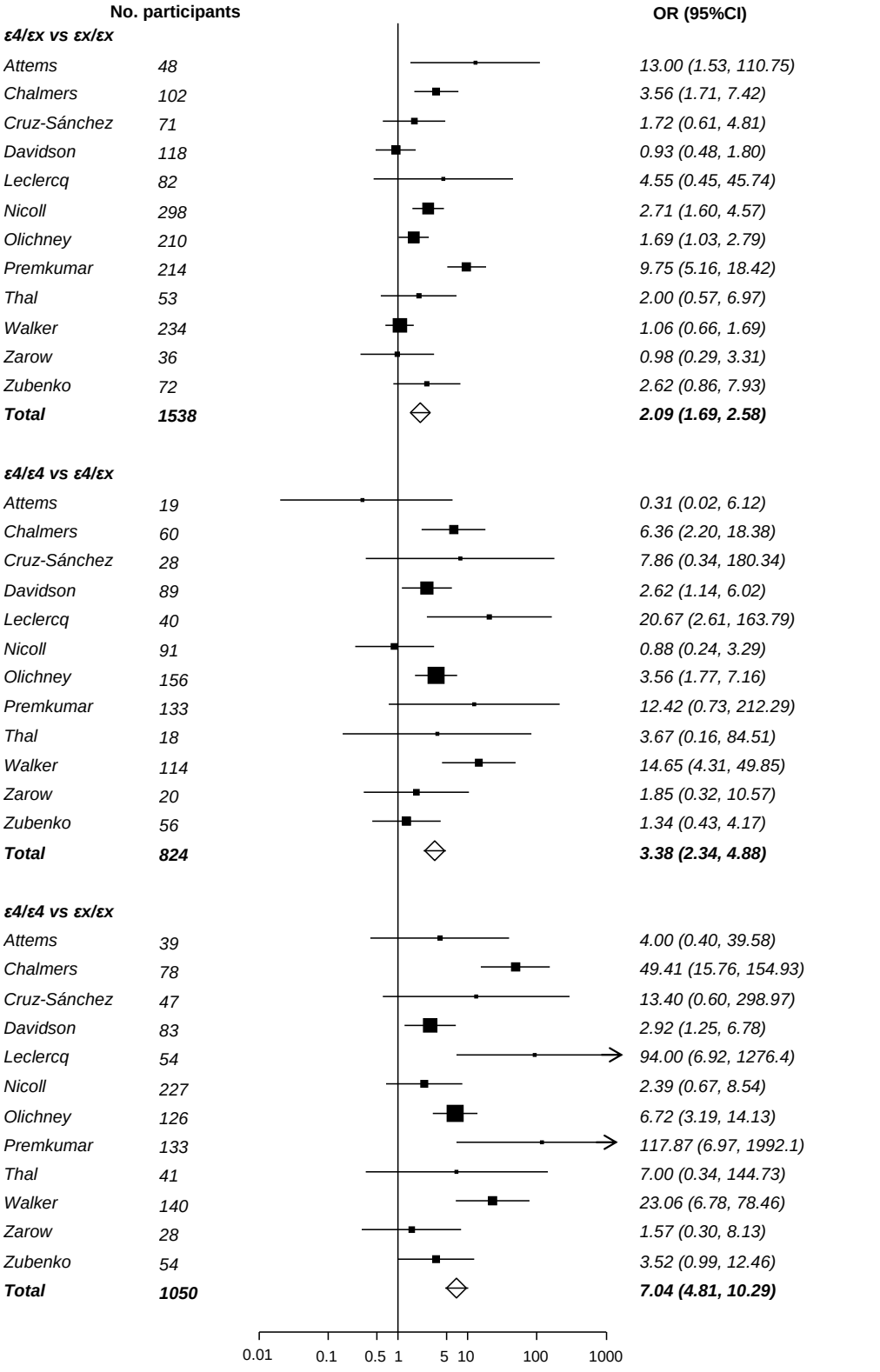
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Continuous data: I^2=80%; χ^2_{9df}=45.93; p<0.00001; Dichotomous data: I^2=54%; χ^2_{13df}=27.98; p=0.009; Test for subgroup differences: I^2=85.2%; χ^2_{1df}=6.77; p_{diff}=0.009;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.15 Meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA by pathology assessment quality



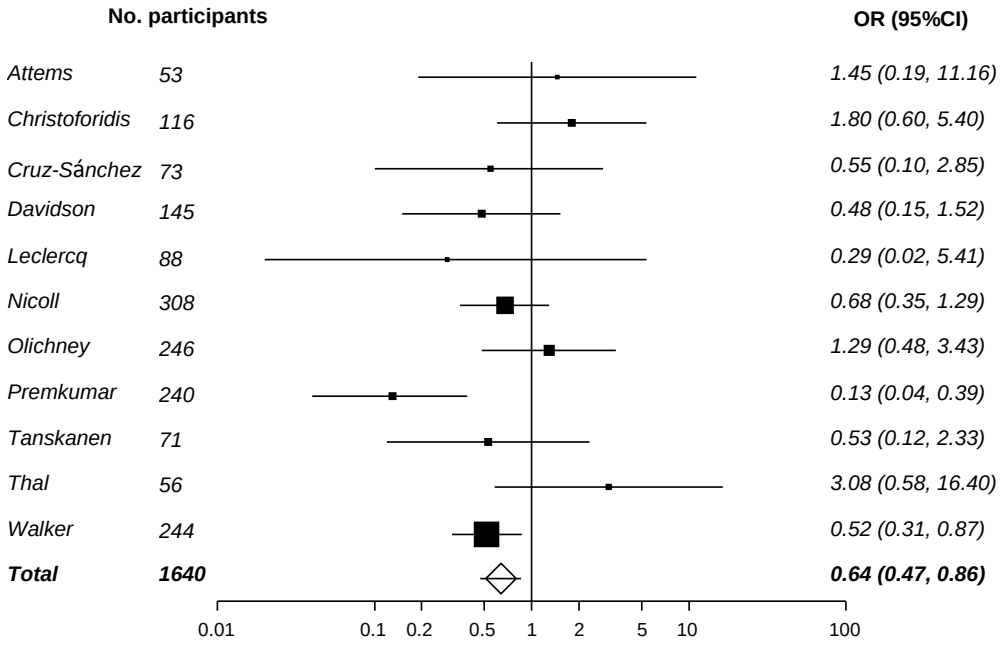
Generic inverse variance fixed-effects method; Within subgroup heterogeneity: Studies with 1 point: I^2=52%; χ^2_{14df}=29.38; p=0.009; Studies with 2 points: I^2=75%; χ^2_{8df}=32.38; p<0.0001; Test for subgroup differences: I^2=94.7%; χ^2_{1df}=18.91; p_{diff}<0.0001;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (10 participants excluded because of missing data: Christoforidis 2 participants, Greenberg 5 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

Figure 3.16 Meta-analysis of effects of *APOE* ε4 dose on presence versus absence of CAA



Generic inverse variance fixed-effects method; Within group heterogeneity: ε4/εx versus εx/εx I^2=76%; χ^2_{11df}=44.99; p<0.00001; ε4/ε4 versus ε4/εx : I^2=47%; χ^2_{11df}=20.77; p=0.04; ε4/ε4 versus εx/εx: I^2=68%; χ^2_{11df}=34.11; p=0.0003;
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (excluded due to missing data: Nicoll 2011: 2 participants; Davidson 2006 and Olichney 2000: 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamonds represent the pooled ORs.

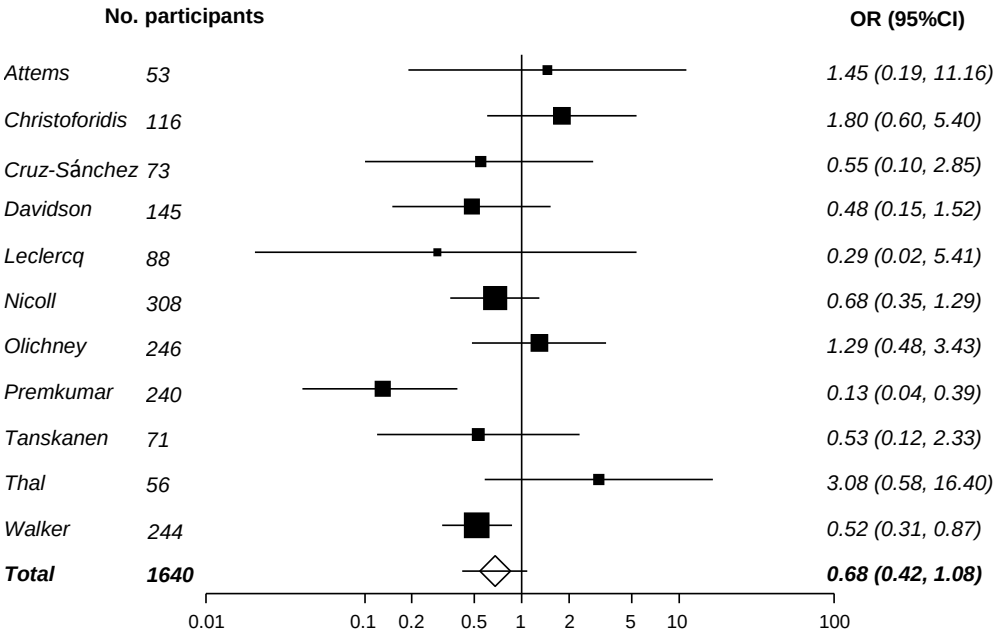
Figure 3.17 Meta-analysis of association of *APOE* ε2+ versus ε2- genotypes with CAA (fixed-effects method)



Generic inverse variance fixed-effects method; $p_{(\text{overall effect})} = 0.003$; Heterogeneity: I^2=46%; χ^2_{10df}=18.57; p=0.05;

OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (5 participants excluded because of missing data: Christoforidis 2 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CI; Diamond represents the pooled OR.

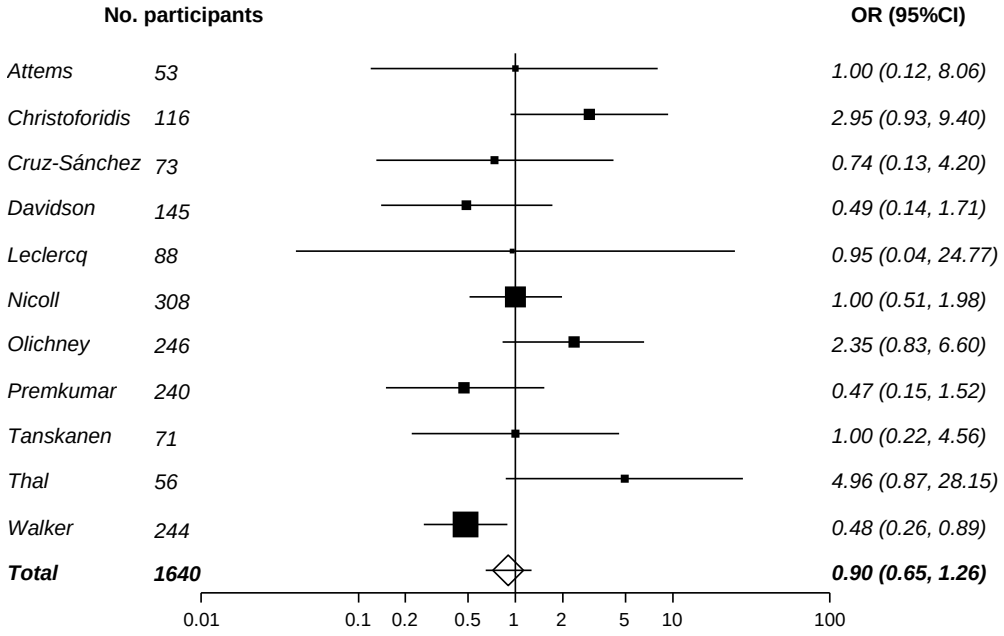
Figure 3.18 Meta-analysis of association of *APOE* ε2+ versus ε2- genotypes with CAA (random-effects method)



Generic inverse variance random-effects method; $p_{(\text{overall effect})} = 0.1$; Heterogeneity: I^2=46%; χ^2_{10df}=18.57; p=0.05;

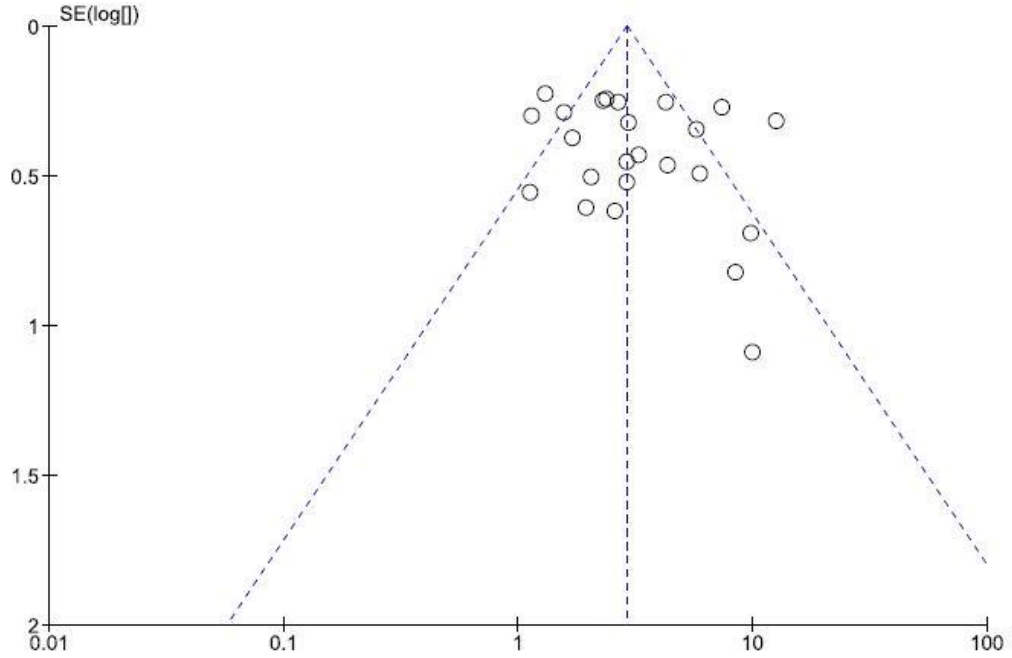
OR = odds ratio; CI = confidence interval; No. participants = total number of participants contributing to the analysis (5 participants excluded because of missing data: Christoforidis 2 participants, Nicoll 2 participants, Olichney 1 participant); Squares represent study-specific ORs, with their size proportional to the study weight based on the random-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR.

Figure 3.19 Meta-analysis of association of *APOE* ε2+ versus ε3/ε3 genotypes with CAA (fixed-effects method)



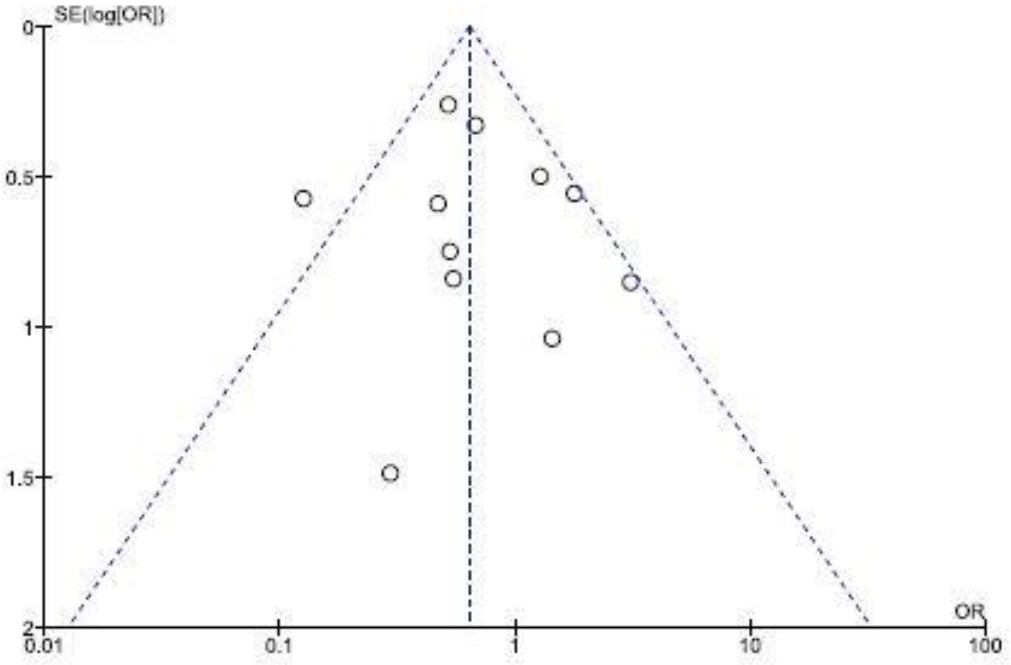
Generic inverse variance fixed-effects method; $p_{(\text{overall effect})}$ =0.55; Heterogeneity: I^2=42%; χ^2_{10df}=17.30; p=0.07; No. participants = total number of participants in the study (6 participants excluded because of missing data: Christoforidis 2 participants, Nicoll 2 participants, Olichney 1 participant, Davidson 1 participant); OR = odds ratio; CI = confidence interval; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects model; Horizontal lines represent 95% CIs. Diamond represents the pooled OR.

Figure 3.20 Funnel plot for the meta-analysis of association of *APOE* ε4+ versus ε4- genotypes with CAA



Dots represent individual studies contributing data for the meta-analysis, the position of the study on the x-axis shows the study's logOR and the position on the y-axis the standard error (SE[logOR]).

Figure 3.21 Funnel plot for the meta-analysis of association of *APOE* ε2+ versus ε2- genotypes with CAA



Dots represent individual studies contributing data for the meta-analysis, the position of the study on the x-axis shows the study's logOR and the position on the y-axis the standard error (SE[logOR]).

Chapter 4: *APOE* ε2/ε3/ε4 polymorphism associations with severe CAA vasculopathy

Chapter contents:

4. *APOE* ε2/ε3/ε4 polymorphism associations with severe CAA vasculopathy

- Neuropathological studies suggest that CAA causes lobar ICH
- While only the *APOE* ε4 allele is associated with presence of histopathological CAA (shown in Chapter 3), both *APOE* ε4 and ε2 are associated with lobar ICH
- A popular explanation is that ε4 promotes amyloid deposition, while ε2 promotes progression to severe CAA vasculopathy that causes vessel rupture and ICH
- In this chapter, I present my systematic review and series of meta-analyses, looking at the association between the *APOE* ε2/ε3/ε4 polymorphism and progression to severe CAA vasculopathy
- I used a comprehensive search strategy, critically appraised study quality and obtained unpublished data by contacting authors
- I conclude that while ε4 may increase the risk of developing severe CAA vasculopathy, current evidence does not support the theory of ε2 role in this
- Before reliable conclusions can be drawn about *APOE* ε2 specific effects, larger numbers of individuals will need to be included in CAA histopathology studies

4.1 Introduction

Sporadic cerebral amyloid angiopathy (CAA) is characterised by deposition of amyloid-β protein in leptomeningeal and cortical blood vessels, with a prevalence in population-based autopsy studies of 20-40% in non-demented and 50-60% in demented elderly people (Charidimou et al. 2012a). Neuropathological case-control and cross-sectional studies, as well as the increased incidence of intracerebral haemorrhage (ICH) in patients with Alzheimer's disease, suggest that CAA causes lobar ICH (Chi et al. 2013, Samarasekera et al. 2012).

CAA is thought to be responsible for up to a third of all ICH in elderly people (Pezzini and Padovani 2008). The prevalence of ICH in cases with CAA in a recent systematic review was about 11% (Samarasekera et al. 2012). It is unknown why only a few people with CAA pathology develop an ICH, but it seems likely to involve biological pathways additional to and distinct from those involved in vascular amyloid deposition. Cases of CAA with ICH not only have a greater proportion of amyloid-laden blood vessels (McCarron et al. 1999), but also more often demonstrate severe CAA with associated vasculopathy (Figure 4.1) (Mandybur 1986, McCarron et al. 1999, Okazaki et al. 1979, Vonsattel et al. 1991).

Apolipoprotein E genotype (*APOE*) is associated with both histopathologically-confirmed CAA and CAA-related clinical phenotypes, including lobar ICH. My systematic review and meta-analyses in Chapter 3 demonstrated a dose-dependent association between histopathologically confirmed CAA and ε4 allele but there was no convincing association with ε2-containing genotypes (Rannikmäe et al. 2013). However, a recent large-scale genetic association study found that both ε4- and ε2-containing genotypes were associated with lobar ICH, particularly when attributed to CAA (Biffi et al. 2010). Furthermore, *APOE* ε2 but not ε4 predicted initial hematoma volume, haematoma expansion, increased mortality and poor functional outcome of lobar ICH (Biffi et al. 2011, Brouwers et al. 2012). The generally accepted explanation for these findings is that while *APOE* ε4 promotes vascular amyloid deposition, ε2 promotes progression to severe CAA with associated vasculopathy leading to vessel rupture and ICH (Figure 4.1) (Greenberg et al. 1998, McCarron et al. 1999).

In this chapter, I aim to study the evidence behind these proposed allele-specific effects, by performing a systematic review, incorporating a comprehensive search strategy, a thorough assessment of study quality, and a series of meta-analyses looking at the association between the *APOE* ε2/ε3/ε4 polymorphism and progression to severe CAA vasculopathy.

4.2 Methods

4.2.1 Search strategy

I sought all studies of adult humans published in any language, which had conducted both *APOE* genotyping and histopathological assessment for CAA, including assessment for severe CAA with associated vasculopathic changes. From now on when using the term 'severe CAA', I am referring to severe CAA with associated vasculopathic changes.

I used the same search strategy as in Chapter 3, searching OVID *Medline* (1950 to March 2012) and *Embase* (1980 to March 2012), using a combination of search terms for APOE, genes and CAA (Appendix 3). I also checked the bibliographies of all relevant studies and reviews identified, and searched *Google Scholar* for other studies citing relevant studies.

4.2.2 Inclusion/exclusion criteria

For the analyses reported in this chapter, I included all studies of adult humans published in any language, which had conducted both *APOE* genotyping and histopathological assessment for CAA, including assessment for severe CAA with associated vasculopathic changes (blood vessel dilatation; microaneurysm formation; fibrinoid degeneration; cracking and double-barrelling of the vessel wall; and paravascular leakage of blood). The assessment for severe CAA could have occurred either as part of the Vonsattel grading scale, which includes such changes in its 'severe' category (Appendix 5), or through specifically reporting on some or all of the relevant histopathological characteristics.

To avoid the effects of possible reporting bias (whereby positive results are more likely to be included within publications than negative ones) and to increase the size of the relevant dataset, I considered studies to be eligible for inclusion whether or not they actually reported on any association of *APOE* with severe CAA.

My focus was on assessing the potential influence of *APOE* genotypes (in particular *APOE* ε2) on severe CAA preceding rupture. Because *APOE* ε2 is already known to be associated with CAA-related ICH (perhaps – but not definitely – causally), and severe CAA is more commonly found in such cases, I reasoned that including brains with ICH in my analyses would inevitably yield an association of *APOE* ε2 with severe CAA, without necessarily meaning that *APOE* ε2 influences risk through promoting the most severe stages of CAA pathology that precede rupture. Hence, to avoid selection bias, I excluded studies that had selected participants on the basis of having had a CAA-related ICH. For the same reason, I excluded participants selected on the basis of having had a CAA-related ICH from the included studies. However I included studies that had recruited participants with CAA-related ICH as an unselected part of the spectrum of CAA severity.

I also excluded studies of hereditary CAA cases and those with less than ten eligible participants (thereby excluding <3% of overall eligible participants).

One of my three colleagues (clinical research fellow Dr Neshika Samarasekera, Professor Rustam Al-Shahi Salman or Professor Cathie Sudlow) and I independently selected eligible studies, resolving disagreements by discussion and mutual consensus.

4.2.3 Data Extraction

For each study included, I extracted information on:

- first author;
- publication year;
- country in which the study was conducted;
- source, description, mean age, ethnicity and gender of participants;
- study size.

I assessed selected genotyping characteristics based on the STREGA (Strengthening the Reporting of Genetic Association Studies) and MOOSE (Meta-analysis of

Observational Studies in Epidemiology) recommendations (Little et al. 2009, Stroup et al. 2000). I extracted information on:

- DNA source;
- number of participants in whom genotyping was attempted and in whom it was successful;
- genotyping method;
- blinding of genotyping staff to pathology data.

I assessed the quality and characteristics of CAA pathology assessment using my own previously developed criteria (Chapter 3). I extracted data on:

- the method of CAA assessment;
- qualification of CAA rater(s);
- brain locations examined;
- blinding of CAA raters to genotyping results and relevant clinical information.

For the planned meta-analyses, I required summary data on the numbers of participants with each *APOE* genotype (ε3ε3, ε2ε3, ε3ε4, ε4ε4, ε2ε2, ε2ε4) and whether or not CAA was present or absent. For those with CAA, I required, for each genotype, their CAA severity on the Vonsattel scale and/or data on the presence or absence of specific vasculopathic changes associated with severe CAA. These data were not all provided in the publications I identified.

To facilitate the sharing and analysis of the unpublished data required, I formed a collaborative group including the principal investigators of the relevant studies. I created a structured data extraction form and completed it as far as possible by entering data from the relevant publication(s) (Appendix 6). Principal investigators and/or their colleagues then checked the information entered, made any necessary corrections, and entered additional unpublished data if required and available.

4.2.4 Meta-analyses

For the meta-analyses I used Cochrane Review Manager (version 5) software. I included only individuals with CAA present on histopathological assessment, calculating study-specific, fixed-effects and random-effects pooled, unadjusted odds ratios (ORs) with 95% confidence intervals (CIs) for:

- severe versus mild or moderate CAA among ε4 carriers (ε4+) versus those with other genotypes (ε4-), and
- severe versus mild or moderate CAA among ε2 carriers (ε2+) versus those with other genotypes (ε2-).

I used the Mantel-Haenszel method which is the default method for fixed-effects meta-analysis in Review Manager and is considered to have better statistical properties than other methods when there are few events (The Cochrane Collaboration) and the DerSimonian and Laird method for the random-effects analyses.

I then repeated the analyses comparing:

- severe versus moderate CAA (excluding mild CAA and so focussing on the severe end of the CAA spectrum);
- severe versus mild/moderate CAA among ε4+ and among ε2+ versus ε3ε3
- severe versus moderate CAA among ε4+ and among ε2+ versus ε3ε3

I used the wild-type ε3ε3 as a comparison group to avoid potential bias from mixed effects of ε2 and ε4 in the comparison group.

I also repeated some of the analyses performed in Chapter 3 for the subset of studies included here: I included individuals both with and without CAA, comparing the presence versus absence of CAA among ε4+ versus ε4-, and versus ε3ε3, and among ε2+ versus ε2-, and versus ε3ε3 genotypes.

I considered $p<0.05$ to imply statistical significance. I assessed heterogeneity with I^2 and χ^2 statistics.

4.3 Results

4.3.1 Included studies

From 1754 publications identified for screening, eight studies were relevant (Figure 4.2). I excluded two with <10 eligible participants (Eng et al. 2004, van Horssen et al. 2005), leaving six eligible studies. These six studies included 645 unique participants, of whom 102 had been selected on the basis of having had an ICH, leaving 543 that were eligible for my study (Chui et al. 2006, Davis et al. 1999, Greenberg et al. 1995, Greenberg et al. 1998, Kalaria and Premkumar 1995, McCarron et al. 1999). Only one of the six studies had previously reported on the association between the *APOE* genotype and severe CAA (assessed using Vonsattel scale) (Appendix 5), finding a significantly greater frequency of *APOE* ε2 in severe versus moderate CAA cases (Greenberg et al. 1998). This study and four others that had rated CAA on the Vonsattel scale, between them including 497 eligible participants (92% of all 543 potentially eligible participants) (Chui et al. 2006, Davis et al. 1999, Greenberg et al. 1995, Kalaria and Premkumar 1995), were able to share their unpublished data in collaborative meta-analyses. Data were unavailable from one additional study (46 eligible participants) that had assessed various different CAA-associated vasculopathic changes (double barrelling of the vessel wall, fibrinoid necrosis and microaneurysm formation), including those that are not part of the Vonsattel scale (McCarron et al. 1999).

4.3.2 Study characteristics

All of the five studies included in the meta-analyses used autopsy brains from brain tissue banks or a population-based prospective study with an autopsy component. There were between 57 and 227 eligible participants per study. Mean age at death was between 77 and 84 years and about half of all participants were male. Three

studies (357 participants) were conducted in predominantly white populations in the USA while information on ethnicity was unavailable for two studies (140 participants). About 50% of participants had clinical dementia (mainly neuropathologically confirmed AD), about 20% were known not to be demented, and in the remaining 30% dementia status was unknown (Table 4.1).

DNA source was brain tissue in three studies, blood in one study, and brain tissue or blood in one study. Number in whom genotyping was attempted and in whom it was successful was available in four studies. Genotyping method was PCR and genotypers were blinded to pathology data in all studies (Table 4.2).

Methods for pathological assessment were variable, reflecting a lack of agreed standards for CAA pathology assessment at the time these studies were conducted. Raters were blinded to genotyping data in all studies, and to relevant clinical information in four studies. In one study pathologists had access to minimal clinical information. Raters in all studies were qualified neuropathologists. In all studies locations rated included the frontal-, temporal-, parietal- and occipital cortex. In one study the leptomeningeal vessels, and in two studies the cerebellar cortical vessels were also rated (Table 4.2).

4.3.3 Meta-analyses results

4.3.3.1 Association between *APOE* ε4 and CAA severity on the Vonsattel scale

Fixed-effects meta-analysis of data from 353 participants with CAA included in five studies showed, for ε4+ versus ε4- genotypes, a significant association with severe versus mild/moderate CAA (OR 2.55, 95% CI 1.44 to 4.52; p=0.001) but no significant association with severe versus moderate CAA (OR 1.64, 95% CI 0.88 to 3.07; p=0.12) (Figure 4.3, Figure 4.4) There was no significant heterogeneity between individual studies' results (I^2=0%; χ^2_{3df}=1.22; p=0.75 and I^2=0%; χ^2_{3df}=1.19; p=0.76 respectively).

Results were similar and conclusions unchanged when I used ε3ε3 genotypes as the comparison group (Figure 4.5, Figure 4.6). Comparing severe CAA with mild/moderate CAA showed an OR 2.88 (95% CI 1.55 to 5.35; p=0.0008) and comparing severe CAA with moderate CAA an OR of 1.95 (95% CI 0.98 to 3.87; p=0.06).

Results were also similar and conclusions unchanged when I performed the analyses using the random-effects method (Table 4.3).

Associations with the presence versus absence of CAA were consistent with results from my previous analysis in Chapter 3 showing a clearly significant association with ε4+ (ε4+ versus ε4-: OR 4.91, 95% CI 3.12 to 7.73; p<0.00001) (Figure 4.7).

4.3.3.2 Association between *APOE* ε2 and CAA severity on the Vonsattel scale

For ε2+ versus ε2- genotypes, the associations were non-significant. ε2+ genotypes compared to ε2- genotypes for association with severe versus mild/moderate CAA with the fixed-effects method showed an OR of 1.74 (95% CI 0.70 to 4.32; p=0.23) and for association with severe versus moderate CAA an OR of 2.78 (95% CI 0.86 to 8.98; p=0.09) (Figure 4.8, Figure 4.9). The confidence intervals were wide due to small numbers of participants, particularly in the ε2+ group, which included 22 and 12 individuals respectively, only seven of whom had severe CAA. There was moderate heterogeneity between individual studies' results for severe CAA versus mild/moderate CAA (I^2=52%; χ^2_{3df}=6.24; p=0.1) and minimal heterogeneity for severe versus moderate CAA (I^2=11%; χ^2_{3df}=3.38; p=0.34).

Results were similar and conclusions unchanged when I used ε3ε3 genotypes as the comparison group (rather than ε2-) or when I performed the analyses using the random-effects method (Table 4.3, Figure 4.10, Figure 4.11, Figure 4.12). Comparing ε2+ versus ε3/ε3 genotypes for association with severe CAA versus mild/moderate CAA using the fixed-effects method showed a marginally significant association (OR 3.18, 95% CI 1.14 to 8.89; p=0.03) but the association was no

longer significant when analysed with the random-effects method (OR 3.40, 95% CI 0.89 to 12.96; p=0.07). Comparing severe CAA with moderate CAA using the fixed-effects method showed a non-significant OR of 3.45 (95% CI 0.89 to 13.35; p=0.07) and the result was similar when using the random-effects method (Table 4.3).

Associations with the presence versus absence of CAA were consistent with results from my previous analysis in Chapter 3 showing no robust association with ε2 (Figure 4.13, Figure 4.14, Figure 4.15).

4.4 Discussion

4.4.1 Main findings

After collecting and analysing data from five out of six existing relevant studies identified through a systematic search (including >90% of all eligible participants), meta-analyses suggested a possible association of *APOE* ε4 with progression to severe CAA as well as the expected association of ε4 (and not of ε2) with the presence of CAA (Rannikmäe et al. 2014, Rannikmäe et al. 2013). However, I was unable to confirm a statistically significant association between *APOE* ε2+ genotypes and severe CAA. It should be emphasised that, although I included all available data from relevant publications, the relative rarity of ε2+ genotypes (only 22 of 353 individuals with CAA in the collaborative dataset) made the ε2-based analyses very imprecise.

Since my systematic review, at least two further relevant studies have been published, suggesting that the theory of the *APOE* ε2 allele driving CAA vasculopathy and in that way promoting ICH is not yet confirmed. One study in 65 cases found no overall association between the ε2 or ε4 alleles and CAA vasculopathy (Love et al. 2014). Another study in 105 pathologically confirmed CAA cases found that vasculopathy was not more common in cases with ICH compared to cases without ICH. In addition they found a trend towards the ε2 allele being associated with ICH, which suggests that ε2 is associated with ICH through

mechanisms other than promoting vasculopathy (Charidimou et al. 2015a). Unfortunately these studies do not provide raw data to be included in the meta-analyses performed in this chapter.

I performed the analyses using the fixed-effects method and then repeated all analyses using the random-effects method to confirm the robustness of any significant results. As discussed in Chapter 3, it is generally recommended to use the fixed-effects method when the underlying assumption is that the true allele effect is the same in all studies included in the meta-analysis, and to use the random-effects method when the underlying assumption is that the true allele effect is different in the included studies (Borenstein et al. 2010). I chose to use both methods to confirm robustness of any significant results and also because there was heterogeneity between some studies. In addition, even when no apparent heterogeneity can be detected with the statistical tests, this does not exclude the presence of heterogeneity because the test for heterogeneity often has poor power, and therefore can be non-significant even if the true level of heterogeneity is substantial (Borenstein et al. 2010). This approach is supported by the Cochrane Collaboration, which recommends comparing the fixed- and random-effects estimates if one is concerned about the influence of small-study effects on the results of meta-analysis in which there is evidence of between-study heterogeneity (The Cochrane Collaboration 2011). If the estimates are similar, then any small-study effects have little impact on the effect estimate.

4.4.2 Strengths and limitations

The strengths of this study are the thorough search methods along with critical appraisal of the quality of included studies (which was generally good), inclusion of unpublished data through establishing a collaborative group to share data, and inclusion of these data in meta-analyses. In addition, I was careful to avoid selection bias by excluding cases selected on the basis of having had an ICH, since this phenotype has known associations both with severe CAA and with *APOE* ε2 and ε4.

This study has a number of limitations. First, although the analysis includes data from >90% of eligible cases identified, I was unable to include cases from one small study (McCarron et al. 1999). In that study, *APOE* ε2 allele-carriers had a statistically significant excess of fibrinoid necrosis compared with non-ε2 carriers, but CAA-related ICH cases were included in these analyses. Second, despite including data from almost all relevant cases from the published literature, total numbers were relatively small and confidence intervals wide, especially for analyses of the effects of *APOE* ε2. Third, methods for histopathological assessment varied between studies, potentially introducing heterogeneity and reducing the likelihood of detecting a consistent effect across studies. Fourth, *APOE* allele-specific effects on severe CAA may differ according to the presence or absence of Alzheimer's disease, particularly for *APOE* ε2, which has been associated with a decreased risk of Alzheimer's dementia (Verghese et al. 2011). I could not perform informative subgroup analysis, however, because of the small overall numbers of participants and because dementia status was unknown for a large number of participants. Fifth, while the studies included here assessed those severe CAA-associated vasculopathic changes that are specifically alluded to in the Vonsattel scale, other vasculopathic changes may also be relevant. Sixth, both *APOE* allele-specific and other genetic associations may differ by CAA subtype. For example, there is preliminary evidence that *APOE* ε4 may be associated with CAA type 1 (where CAA is found in cortical capillaries), and ε2 with CAA type 2 (where amyloid is deposited in leptomeningeal and cortical vessels with the exception of cortical capillaries) (Thal et al. 2002). If CAA types 1 and 2 represent different pathological entities, the mechanisms and genetic risk factors for severe CAA and ICH could also differ. I did not have the necessary data to explore this in my study. Finally, the validity of including unpublished data that have not passed peer-review can be questioned, however the peer-review filter may be imperfect and there is some evidence from other fields to suggest that inclusion of unpublished data may lead associations to lose their statistical significance, suggesting that the published literature is shaped by selective reporting biases (Kyzas et al. 2005).

This study provides an example of the importance of systematic review and meta-analysis methods in studying genetic associations, and how narrative reviews not informed by systematic reviews can be affected by citation bias (Greenberg 2009). Although studies in the area of CAA have continued to reference particular studies selectively, it is to be hoped that increasing recognition of the importance of systematic and unbiased assessments of the body of evidence on a particular topic will increasingly inform and direct future and ongoing research efforts.

4.4.3 Outstanding issues

Further large methodologically robust studies adhering to current reporting standards are needed to definitively answer the question about the effects of the *APOE* genotype on progression to severe CAA vasculopathy. Also, the allele-specific effects on severe CAA may differ by CAA subtype and according to the presence or absence of Alzheimer's disease. Finally, there may be other genetic influences that interact with the *APOE* genotype to increase risk of or protect against severe CAA and ICH.

4.4.4 Conclusions

This study confirms existing strong evidence that *APOE* ε4 promotes cerebral amyloid angiopathy, and further suggests that ε4 may increase the risk of developing severe CAA among those with CAA. However, while these findings do not exclude a biologically meaningful association between the *APOE* ε2 allele and progression to severe CAA, I did not find convincing evidence to support this. Much larger numbers of individuals will need to be included in CAA histopathology studies before reliable conclusions can be drawn about the specific effects of *APOE* ε2 on CAA. In particular, the hypothesis that *APOE* ε2 influences risk of ICH through promoting progression of CAA to its severe form is not supported by the existing, relevant, unbiased data. Future research efforts in this area will also be helped substantially by the use of standardised histopathological grading system for CAA (including assessment of CAA types 1 and 2) (Love et al. 2014), and by the consistent reporting

of dementia – and specifically Alzheimer's disease – status among individuals included in histopathology studies (Dubois et al. 2013).

Tables

Table 4.1 Characteristics of studies included in the meta-analyses of *APOE* genotype associations with severe CAA vasculopathy

First author, publication year and country	Brief description of participants	Study size[1]	Clinically demented participants	Mean age (SD)	Ethnicity	% male
Davis 1999 USA	Prospective study of > 60 year old cognitively normal volunteers from community	57	0/0	83.9 (7.4)	Caucasian	46%
Greenberg 1995 and 1998 USA	Brains selected from a brain tissue resource centre	140	-	78.9 (8.9)[3]	-	46%
Chui 2006 USA	People from a longitudinal study of ischaemic vascular dementia, AD and cognitively normal elderly people	73	42/34[4]	82.8	86% Caucasian	57%
Kalaria 1995 USA	Brains from research based brain tissue banks	227[5]	211/177[6]	77-79	Caucasian	45%

[1]Number of eligible participants genotyped & assessed for CAA; [3]Mean age for 138 participants; [4]42 out of a larger sample of 79 participants were clinically demented and 34/79 had clinically diagnosed and neuropathologically confirmed Alzheimer's disease; [5]Authors provided unpublished data for additional 34 controls not included in the original publication; [6]211 participants clinically demented, 177 participants with clinically diagnosed and neuropathologically confirmed Alzheimer's disease; "-": data not available.

Table 4.2 Genotyping and pathology characteristics and quality of studies included in the meta-analyses of *APOE* genotype associations with severe CAA vasculopathy

First author	DNA source	No. genotyping attempted/ no. successful[1]	Genotyping method	Method for assessing CAA[2]	Blinding[3]			Rater qualification[4]	Locations rated[5]
					G to P	P to G	P to C		
Davis	Blood or brain tissue	57/57	PCR	IHC	✓	✓	✓	NP	F,T,P, O,LM
Greenberg[6]	Brain tissue	-	PCR	CR (all) H&E (some) IHC (some)	✓	✓	✓	NP	F,T,P, O,C
Chui	Blood	73/73	PCR	H&E (all) IHC (some) CR (some)	✓	✓	✓	NP	F,T,P,O
Kalaria	Brain tissue	227/227	PCR	CR (PL) ThS	✓	✓	-[7]	NP	F,T,P,O

[1]Number of participants in whom genotyping attempted and in whom it was successful; [2]Method for assessing CAA from the pathology samples: IHC = immunohistochemistry; CR = Congo Red; H&E = Hematoxylin and Eosin; PL = polarised light; ThS = Thioflavin S; [3]Blinding: G to P = genotypers to pathology data; P to G = pathologists to genotyping data; P to C = pathologists to relevant clinical information such as participants' dementia and Alzheimer's disease status, history of intracerebral haemorrhage, age; "✓" done; "-" not done; [4]Qualification of the person rating CAA; NP = neuropathologist; [5]Locations examined for CAA vasculopathic changes: F= frontal cortex; T= temporal cortex; P= parietal cortex; O= occipital cortex; C= cerebellar cortex; LM= leptomeninges; [6]Data applies to both Greenberg 1995 and 1998 publications; [7]Minimal clinical information available to pathologists; "-": data not available.

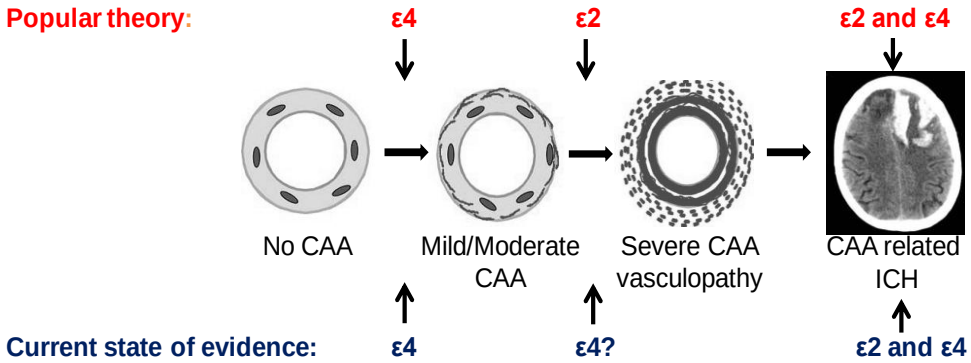
Table 4.3 Summary of meta-analyses of *APOE* genotype associations with severe CAA vasculopathy

Comparison		Severe CAA versus mild/moderate CAA	n/N	Severe CAA versus moderate CAA	n/N
ε4+ versus ε4- genotypes	FE	OR 2.6 (95% CI 1.4 to 4.5); p=0.001	74/224 versus 19/129	OR 1.6 (95% CI 0.9 to 3.1); p=0.12	74/169 versus 19/69
	RE	OR 2.5 (95% CI 1.4 to 4.5); p=0.002		OR 1.7 (95% CI 0.9 to 3.1); p=0.11	
ε4+ versus ε3/ε3 genotypes	FE	OR 2.9 (95% CI 1.6 to 5.4); p=0.0008	74/224 versus 15/114	OR 2.0 (95% CI 0.98 to 3.9); p=0.06	74/169 versus 15/61
	RE	OR 2.9 (95% CI 1.5 to 5.3); p=0.0009		OR 1.96 (95% CI 0.99 to 3.9); p=0.05	
ε2+ versus ε2- genotypes	FE	OR 1.7 (95% CI 0.7 to 4.3); p=0.23	7/22 versus 86/331	OR 2.8 (95% CI 0.9 to 9.0); p=0.09	7/12 versus 86/226
	RE	OR 2.3 (95% CI 0.5 to 11.3); p=0.3		OR 2.7 (95% CI 0.6 to 11.4); p=0.19	
ε2+ versus ε3/ε3 genotypes	FE	OR 3.2 (95% CI 1.1 to 8.9); p=0.03	7/22 versus 15/114	OR 3.5 (95% CI 0.9 to 13.4); p=0.07	7/12 versus 15/61
	RE	OR 3.4 (95% CI 0.9 to 13); p=0.07		OR 3.2 (95% CI 0.5 to 20.5); p=0.22	

n: number of participants with severe CAA; N: total number of participants; Dark grey shading: statistically significant results; FE: results using the fixed-effects method; RE: results using the random-effects method.

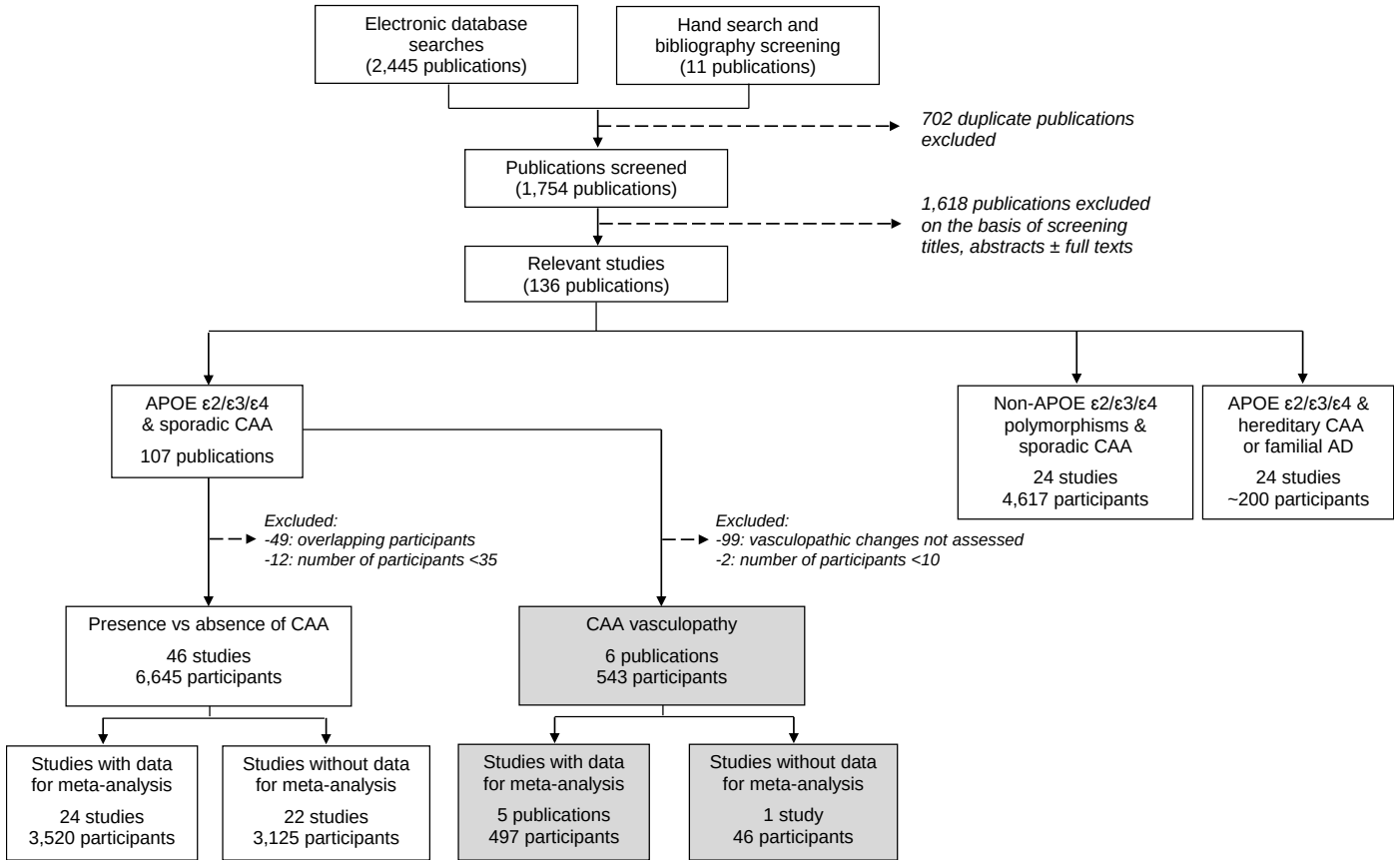
Figures

Figure 4.1 Proposed theory and current state of evidence about associations between *APOE* and CAA phenotype



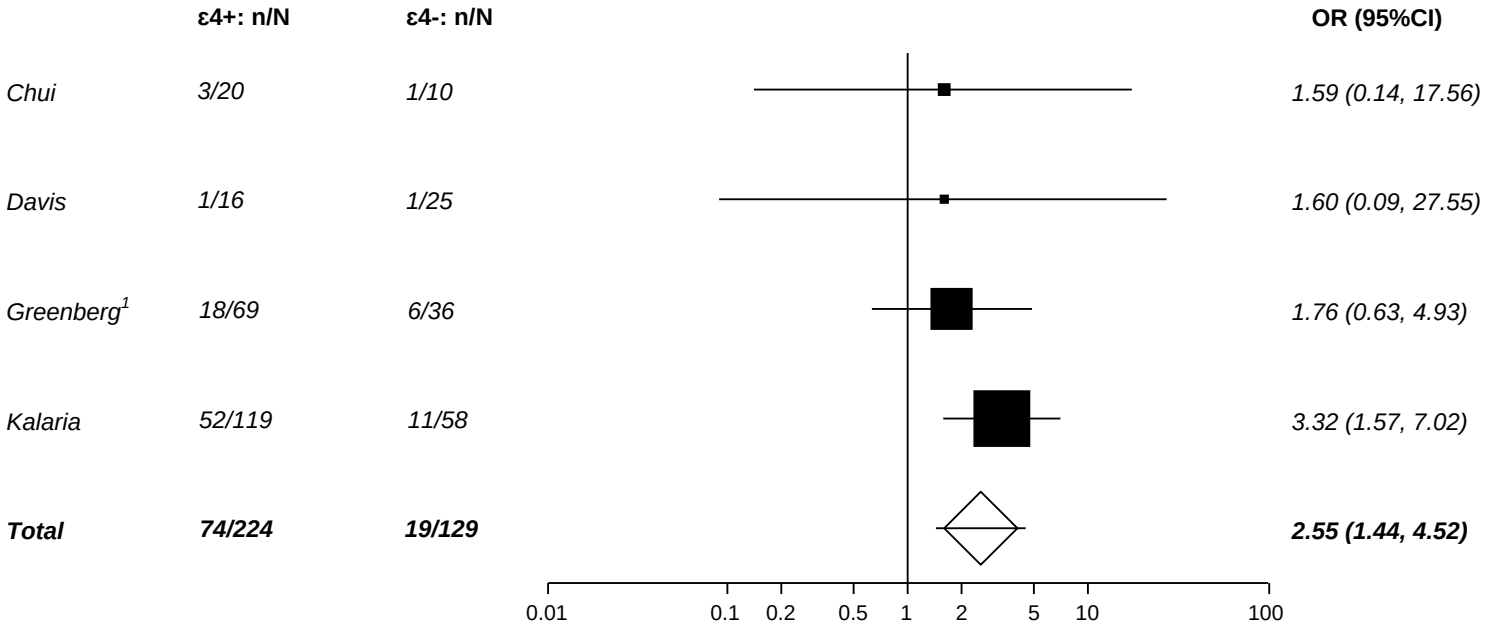
Adapted from Figure 1 in Acta Neuropathologica 2005;110: 345–359 "Sporadic cerebral amyloid angiopathy: pathology, clinical implications, and possible pathomechanisms", Johannes Attems, with kind permission from Springer Science and Business Media and Professor Attems.

Figure 4.2 Selection of studies included in the systematic review of *APOE* genotype associations with severe CAA vasculopathy



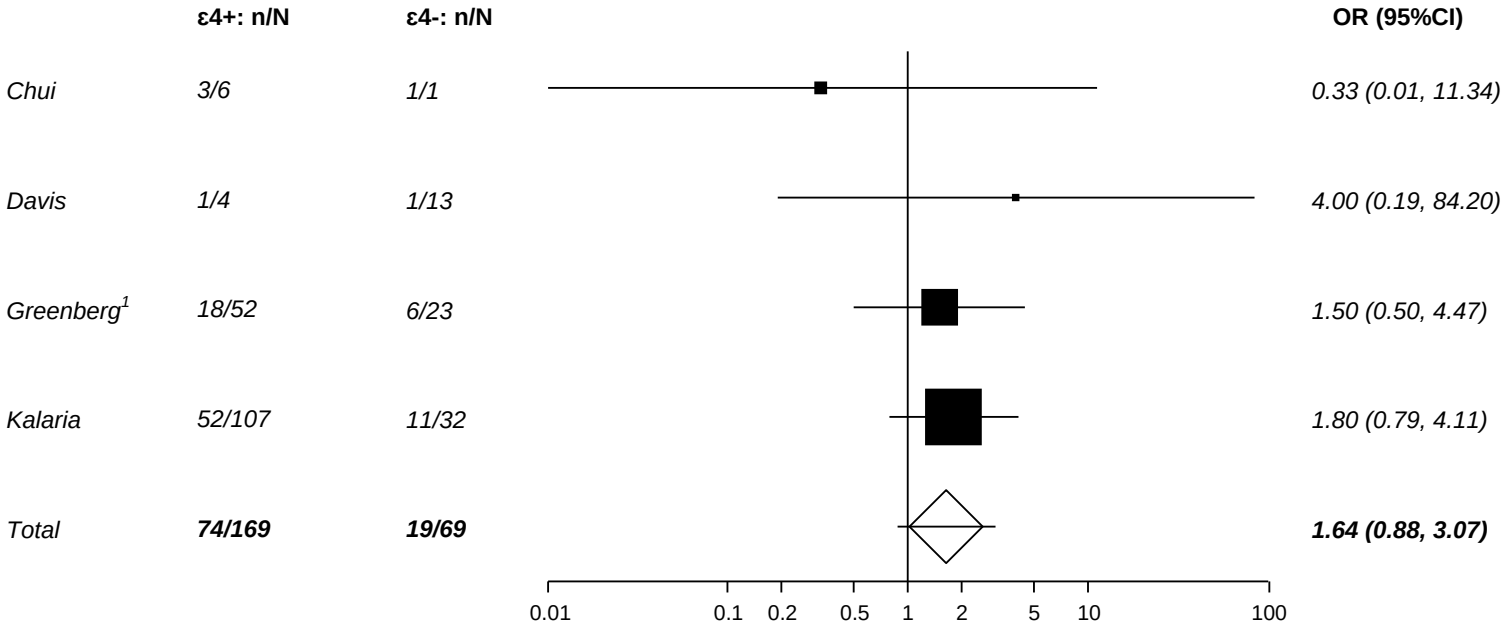
Grey boxes represent the selection of studies relevant for Chapter 4.

Figure 4.3 Meta-analysis of association of *APOE* ε4+ versus *APOE* ε4- genotypes with severe CAA versus mild/moderate CAA



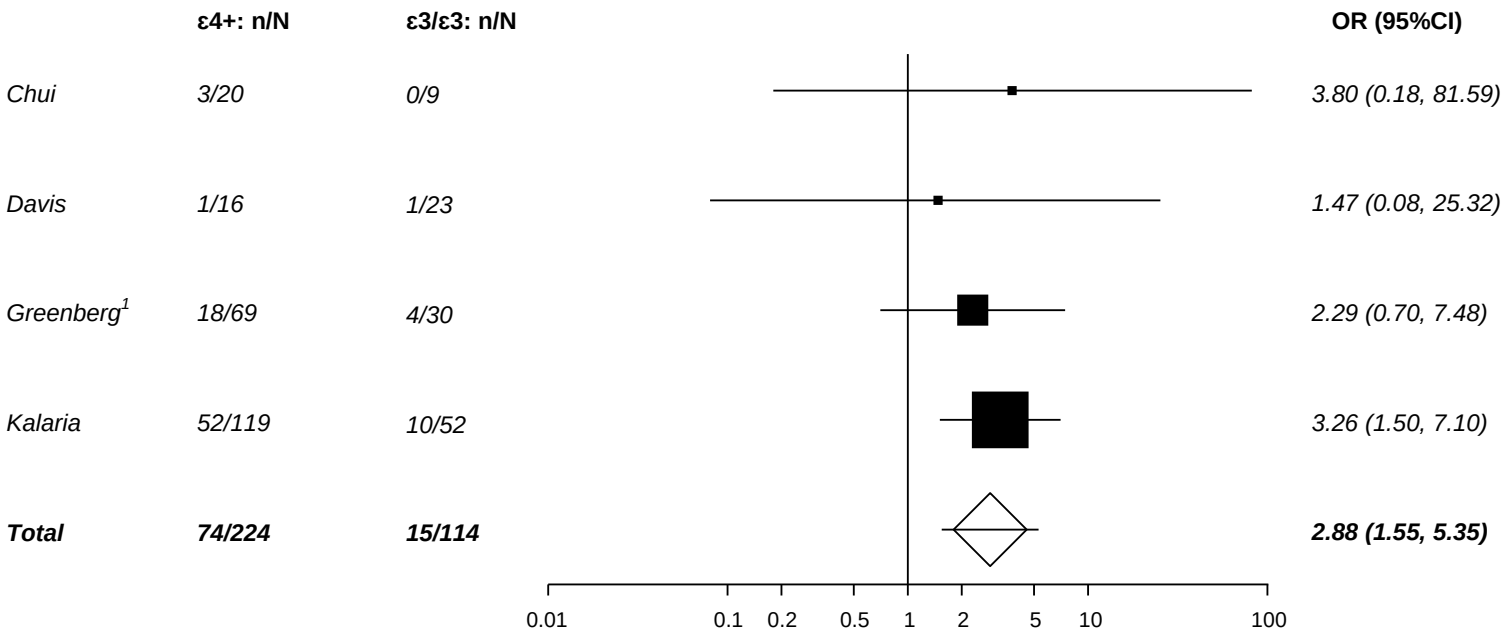
Fixed-effects method; $p_{(overall\ effect)}$=0.001; Heterogeneity: I^2=0%; χ^2_{3df}=1.22; p=0.75; OR = odds ratio; CI = confidence interval; ε4+ = participants with ε4 containing genotypes; ε4- = participants without ε4 containing genotypes; n = number of participants with severe CAA; N = total number of participants with CAA of any severity (mild, moderate or severe); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.4 Meta-analysis of association of *APOE* ε4+ versus *APOE* ε4- genotypes with severe CAA versus moderate CAA



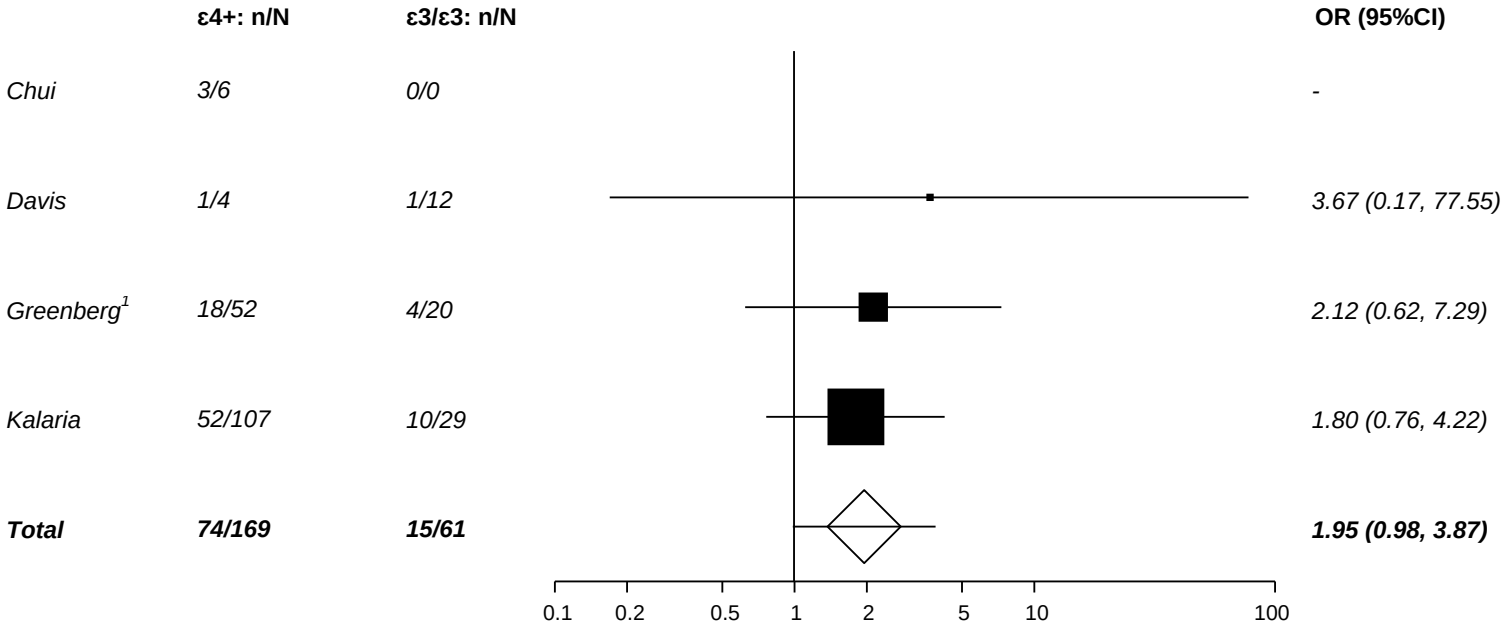
Fixed-effects method; $p_{(overall\ effect)}$=0.12; Heterogeneity: I^2=0%; χ^2_{3df}=1.19; p=0.76; OR = odds ratio; CI = confidence interval; ε4+ = participants with ε4 containing genotypes; ε4- = participants without ε4 containing genotypes; n = number of participants with severe CAA; N = total number of participants with moderate and severe CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.5 Meta-analysis of association of *APOE* ε4+ versus *APOE* ε3/ε3 genotypes with severe CAA versus mild/moderate CAA



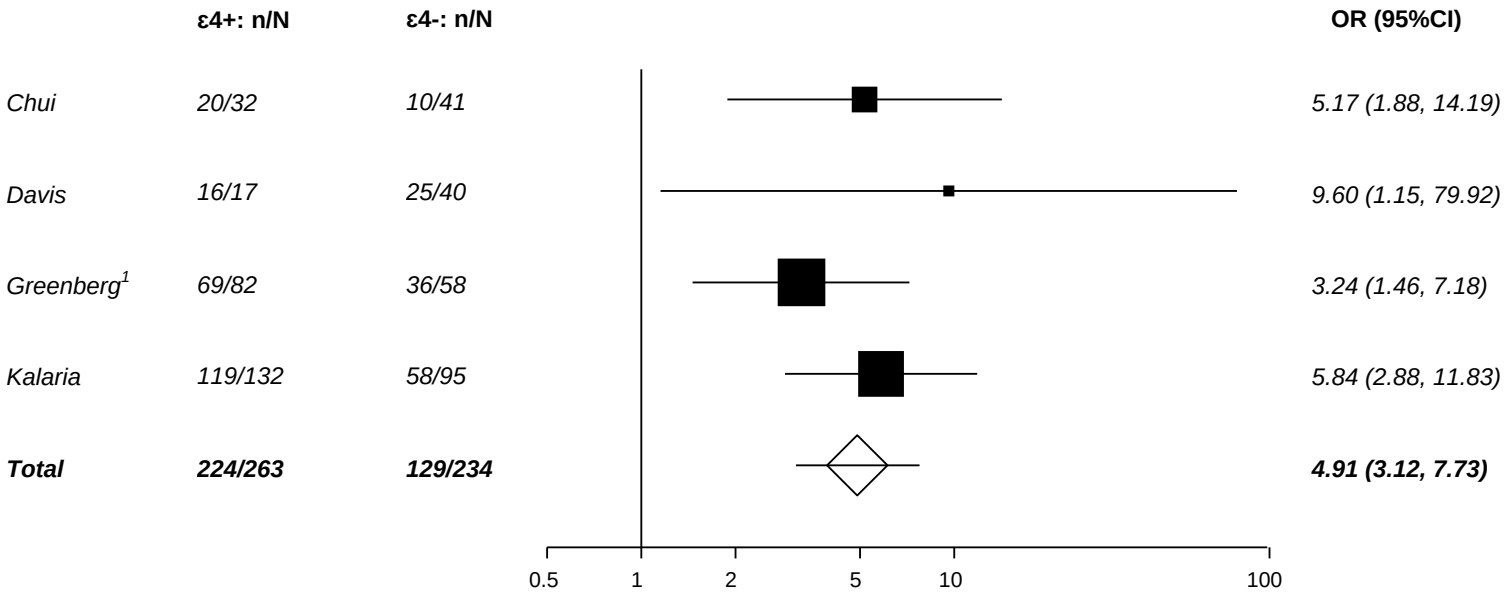
Fixed-effects method; $p_{(overall\ effect)}$=0.0008; Heterogeneity: I^2=0%; χ^2_{3df}=0.49; p=0.92; OR = odds ratio; CI = confidence interval; ε4+ = participants with ε4 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with severe CAA; N = total number of participants with CAA of any severity (mild, moderate or severe); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.6 Meta-analysis of association of *APOE* ε4+ versus *APOE* ε3/ε3 genotypes with severe CAA versus moderate CAA



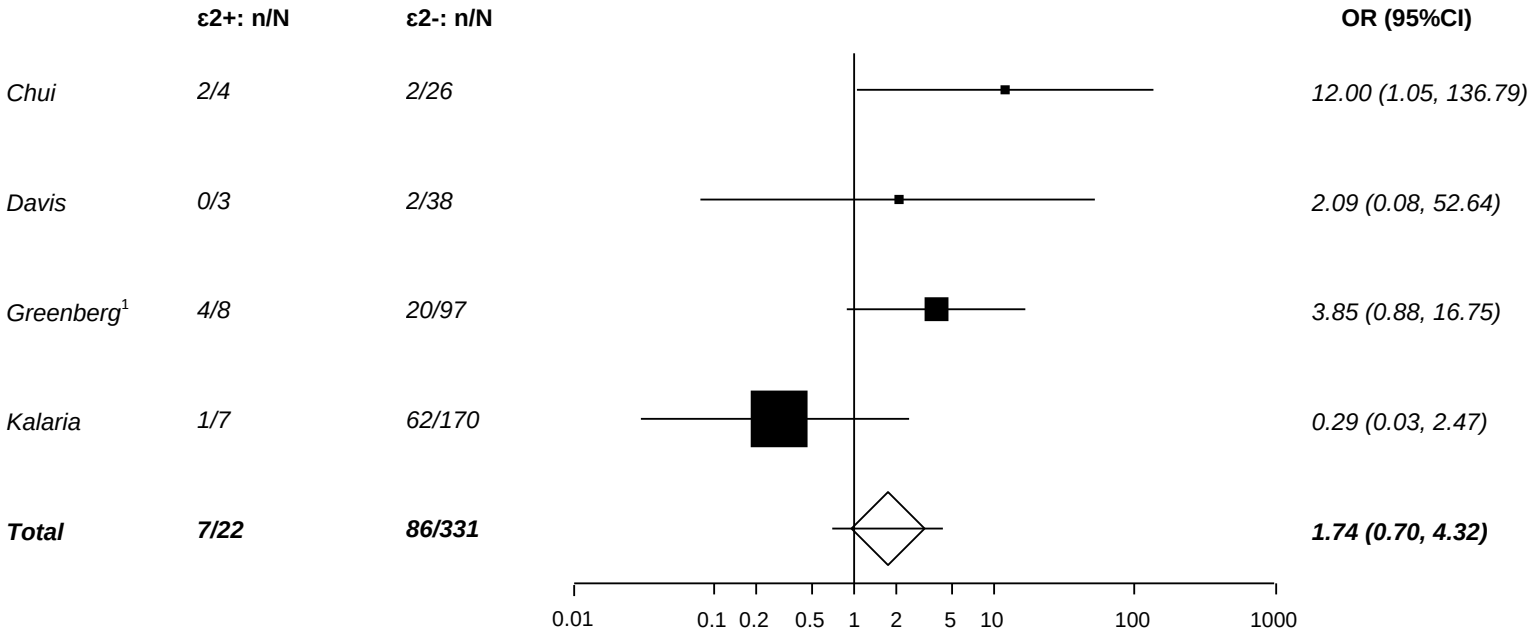
Fixed-effects method; $p_{(overall\ effect)}$=0.06; Heterogeneity: I^2=0%; χ^2_{2df}=0.22; p=0.90;
OR = odds ratio; CI = confidence interval; ε4+ = participants with ε4 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with severe CAA; N = total number of participants with moderate and severe CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.7 Meta-analysis of association of *APOE* ε4+ versus *APOE* ε4- genotypes with presence versus absence of CAA



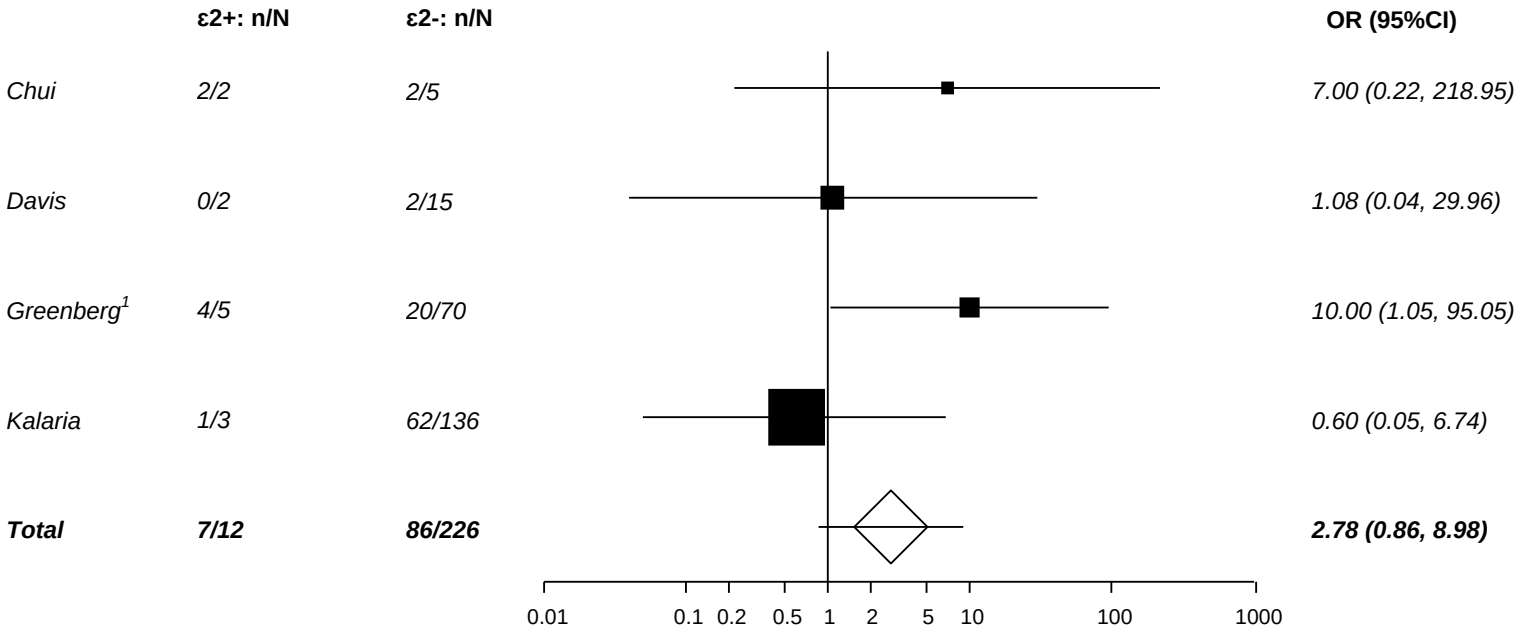
Fixed-effects method; $p_{(\text{overall effect})}<0.00001$; Heterogeneity: $I^2=0\%$; $\chi^2_{3df}=1.67$; p=0.64;
OR = odds ratio; CI = confidence interval; ε4+ = participants with ε4 containing genotypes; ε4- = participants without ε4 containing genotypes; n = number of participants with CAA present; N = total number of study participants; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.8 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε2- genotypes with severe CAA versus mild/moderate CAA



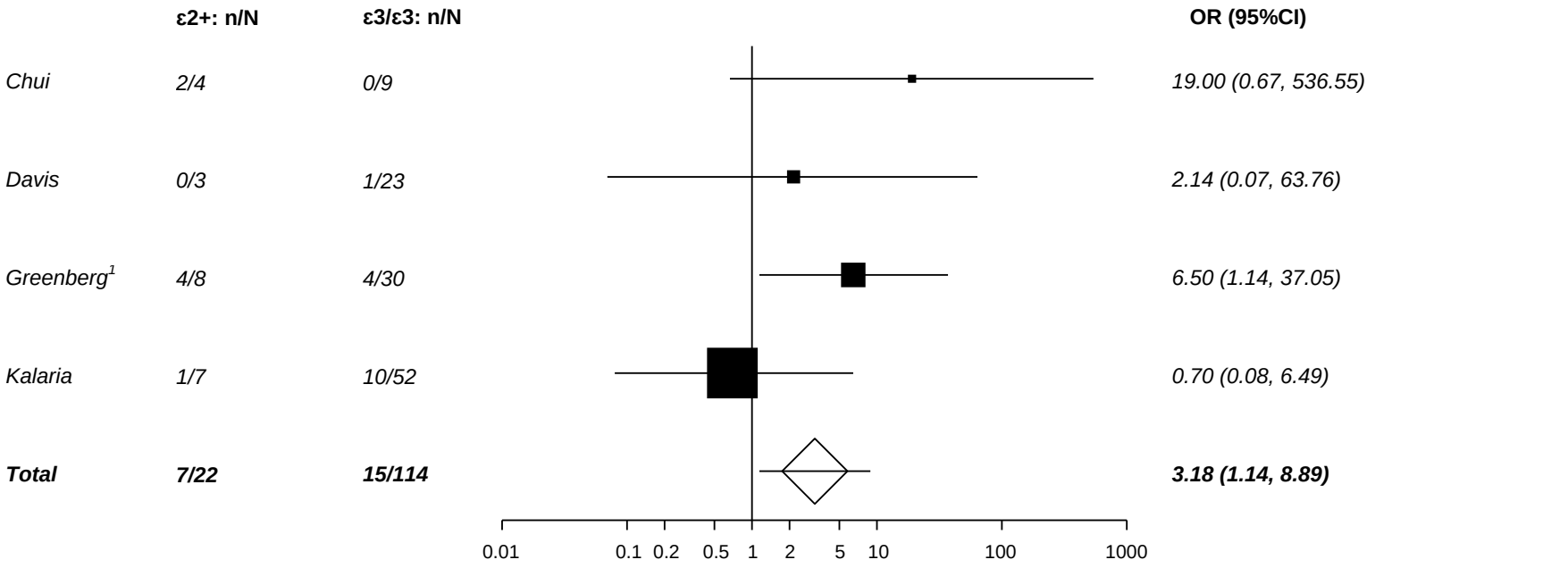
Fixed-effects method; $p_{(\text{overall effect})}$=0.23; Heterogeneity: I^2=52%; χ^2_{3df}=6.24; p=0.10;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε2- = participants without ε2 containing genotypes; n = number of participants with severe CAA; N = total number of participants with CAA of any severity (mild, moderate or severe); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.9 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε2- genotypes with severe CAA versus moderate CAA



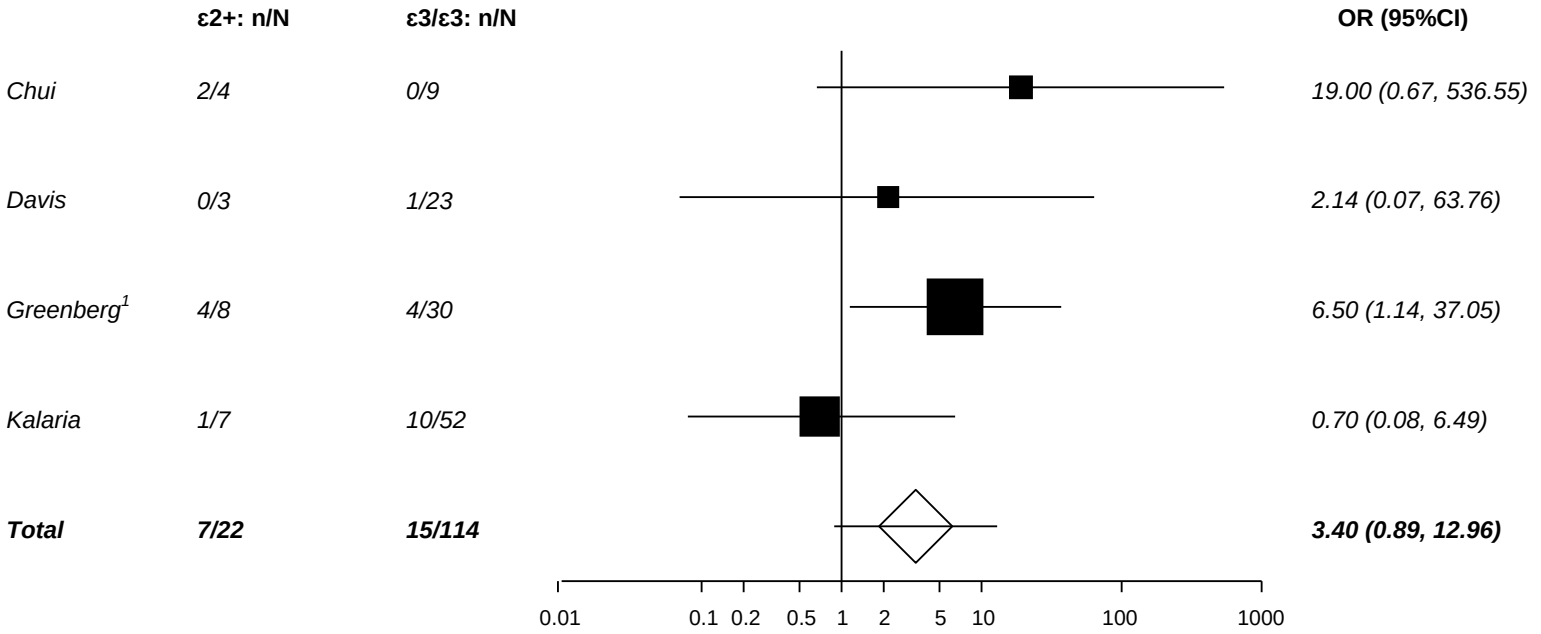
Fixed-effects method; $p_{(overall\ effect)}$=0.09; Heterogeneity: I^2=11%; χ^2_{3df}=3.38; p=0.34;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε2- = participants without ε2 containing genotypes; n = number of participants with severe CAA; N = total number of participants with moderate and severe CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.10 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε3/ε3 genotypes with severe CAA versus mild/moderate CAA (fixed-effects method)



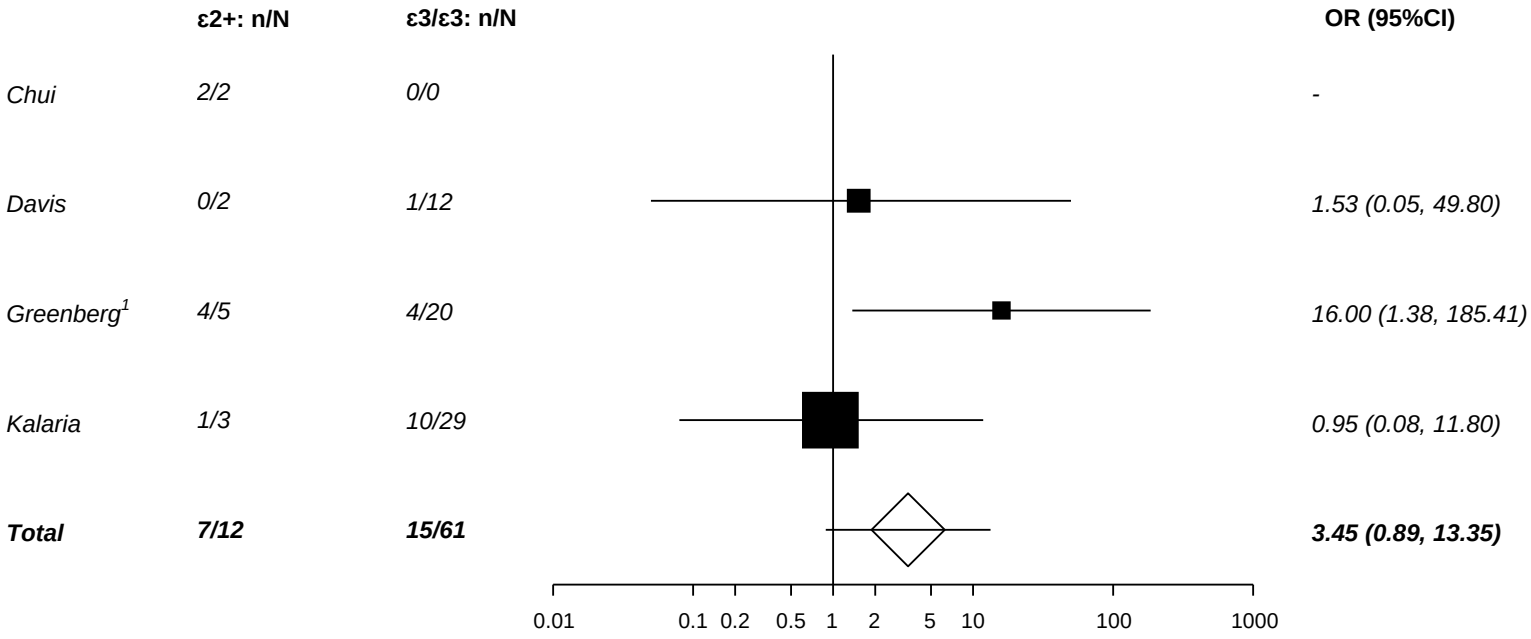
Fixed-effects method; $p_{(overall\ effect)}$=0.03; Heterogeneity: I^2=16%; χ^2_{3df}=3.58; p=0.31;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with severe CAA; N = total number of participants with CAA of any severity (mild, moderate or severe); Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.11 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε3/ε3 genotypes with severe CAA versus mild/moderate CAA (random-effects method)



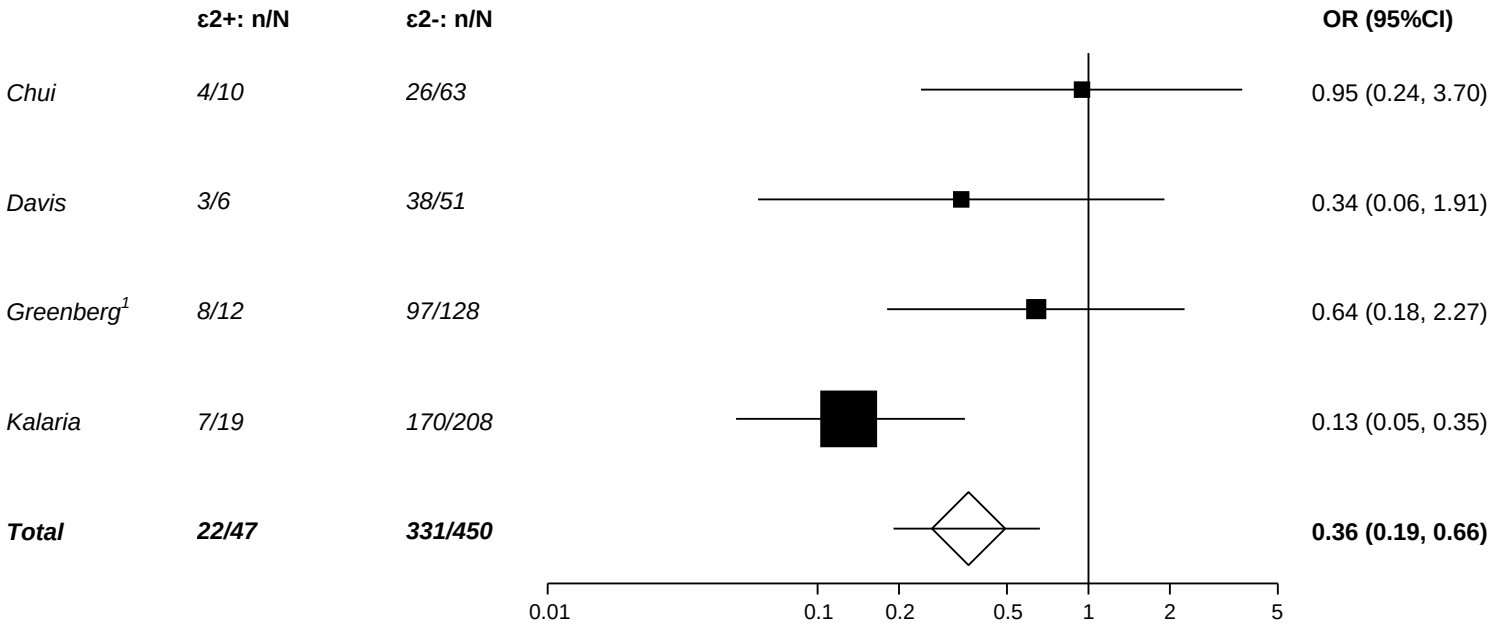
Random-effects method; $p_{(overall\ effect)}$=0.07; Heterogeneity: I^2=16%; χ^2_{3df}=3.58; p=0.31;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with severe CAA; N = total number of participants with CAA of any severity (mild, moderate or severe); Squares represent study-specific ORs, with their size proportional to the study weight based on the random-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.12 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε3/ε3 genotypes with severe CAA versus moderate CAA



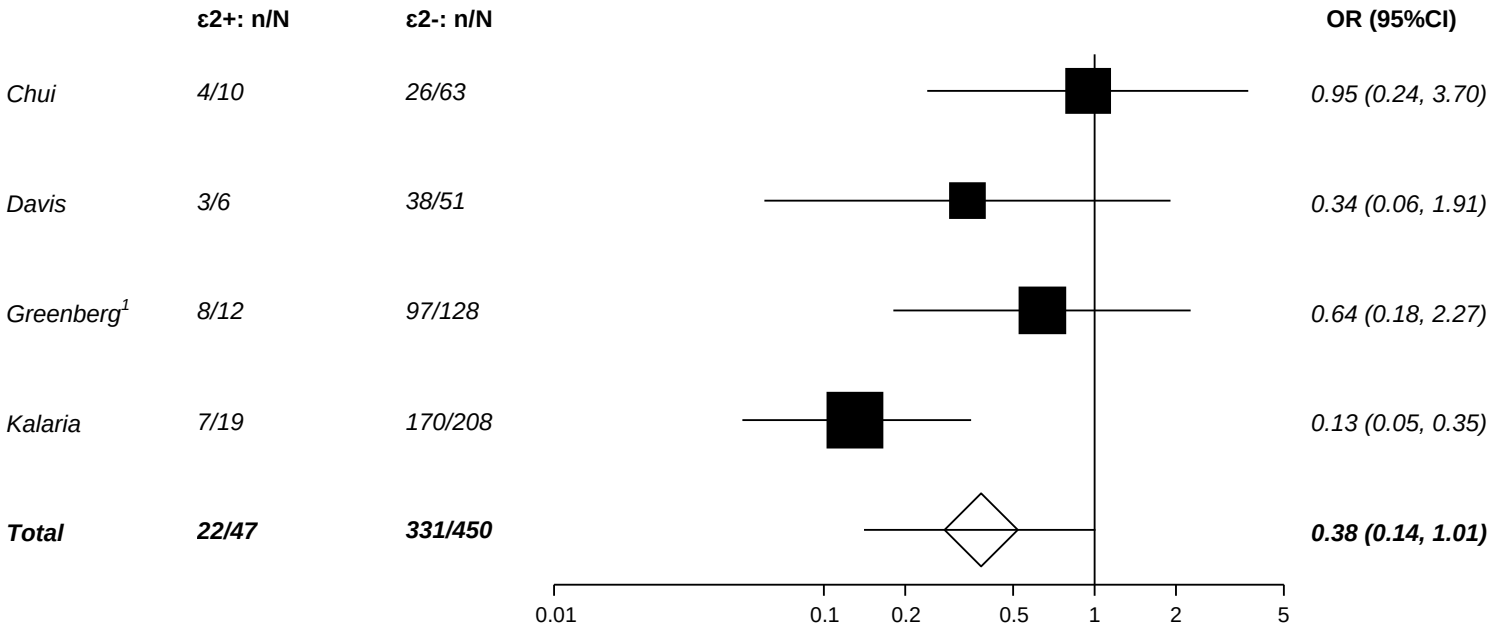
Fixed-effects method; $p_{(overall\ effect)}$=0.07; Heterogeneity: I^2=27%; χ^2_{2df}=2.72; p=0.26;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with severe CAA; N = total number of participants with moderate and severe CAA; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.13 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε2- genotypes with presence versus absence of CAA (fixed-effects method)



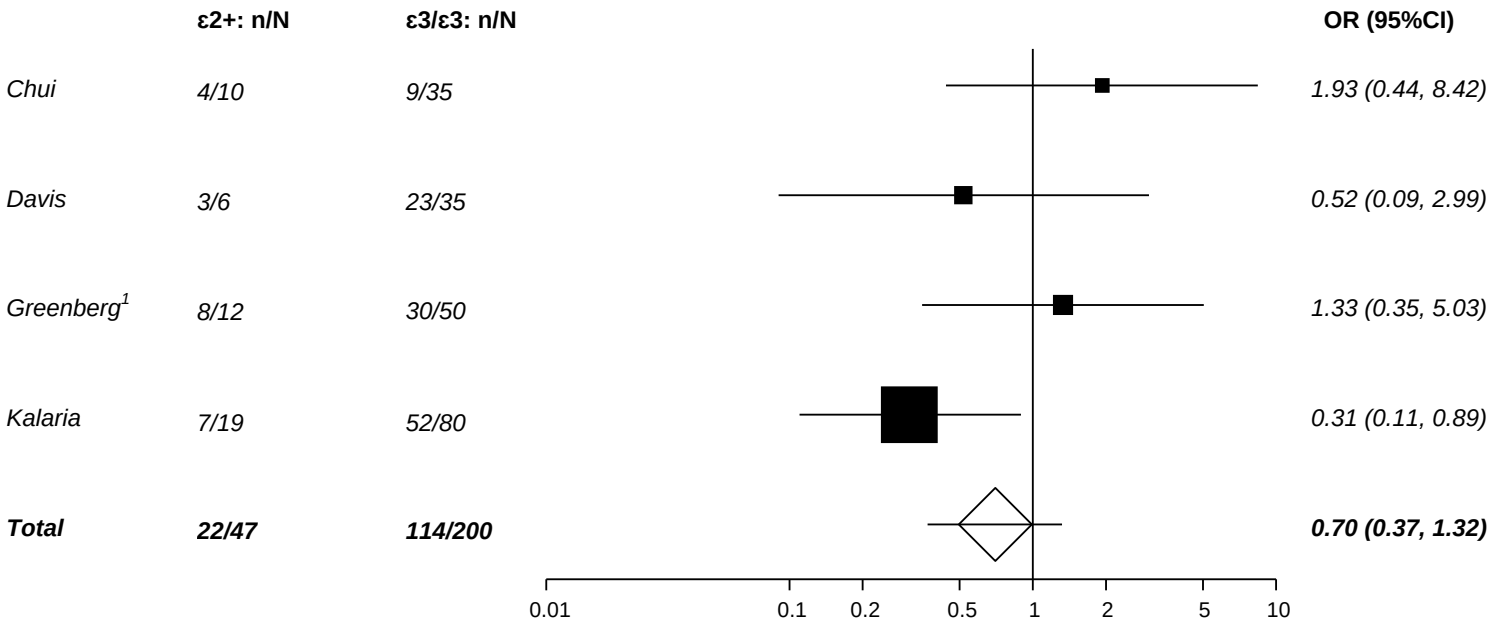
Fixed-effects method; $p_{(overall\ effect)}$=0.001; Heterogeneity: I^2=55%; χ^2_{3df}=6.72; p=0.08;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε2- = participants without ε2 containing genotype; n = number of participants with CAA present; N = total number of study participants; Squares represent study-specific ORs, with their size proportional to the study weight based on the fixed-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.14 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε2- genotypes with presence versus absence of CAA (random-effects method)



Random-effects method; $p_{(overall\ effect)}$=0.05; Heterogeneity: I^2=55%; χ^2_{3df}=6.72; p=0.08;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε2- = participants without ε2 containing genotypes; n = number of participants with CAA present; N = total number of study participants; Squares represent study-specific ORs, with their size proportional to the study weight based on the random-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Figure 4.15 Meta-analysis of association of *APOE* ε2+ versus *APOE* ε3/ε3 genotypes with presence versus absence of CAA



Fixed-effects method; $p_{(\text{overall effect})}$=0.27; Heterogeneity: I^2=41%; χ^2_{3df}=5.11; p=0.16;
OR = odds ratio; CI = confidence interval; ε2+ = participants with ε2 containing genotypes; ε3/ε3 = participants with an ε3/ε3 genotype; n = number of participants with CAA present; N = total number of study participants; Squares represent study-specific ORs, with their size proportional to the study weight based on the random-effects method; Horizontal lines represent 95% CIs; Diamond represents the pooled OR; [1]Results of two studies conducted in one centre were combined for the analyses (Greenberg et al. 1995, Greenberg et al. 1998).

Chapter 5: *APOE* ε2/ε3/ε4 polymorphism associations with severity of hereditary CAA

Chapter contents:

5. *APOE* ε2/ε3/ε4 polymorphism associations with severity of hereditary CAA

- Many types of hereditary CAA exist, distinguished by their genetic cause
- Phenotypic variability between cases suggests modulation by additional factors
- Understanding hereditary CAA mechanisms may contribute towards understanding those of sporadic CAA
- *APOE* ε4 allele is associated with presence, and possibly severity of sporadic CAA (shown in Chapters 3 and 4)
- In this chapter, I aimed to study whether *APOE* also affects histopathological severity of hereditary CAA
- I performed a systematic review of published studies assessing the association of *APOE* ε2/ε3/ε4 polymorphism with histopathological CAA severity in hereditary CAA cases
- I conclude that there is no evidence for an association of *APOE* ε4 with CAA severity in hereditary CAA cases. However there were too few studies and participants, and limited data about study quality to draw firm conclusions
- There are no published data about the association between *APOE* ε2 and CAA severity in hereditary CAA

5.1 Introduction

Amyloid-β peptide is by far the most common amyloid subunit implicated in sporadic forms of CAA and Alzheimer's disease (AD). Amyloid-β CAA is also prominent in variants of familial AD (FAD) due to mutations of the amyloid precursor protein (*APP*), presenilin-1 (*PSEN1*) or presenilin-2 (*PSEN2*) genes, and in several forms of hereditary CAA (HCAA) due to mutations in the *APP* gene. Missense mutations of the *APP* can be within or just outside the coding region of the amyloid-β peptide. Mutations localised close to the β-secretase or γ-secretase

cleavage sites with amino acid substitutions flanking the amyloid-β sequence result in clinico-pathological phenotypes of early-onset FAD, while those resulting in amino acid substitution within the amyloid-β sequence are more often associated with a neuropathological phenotype that includes prominent CAA (HCAAs) (Revesz et al. 2009). Eight types of HCAAs (some also known as hereditary cerebral haemorrhage with amyloidosis or HCHWA) caused by mutations in six sites on the *APP* gene have been described (Kamp et al. 2014, Zhang-Nunes et al. 2006):

- HCAA due to mutation in codon 678;
- HCAA Flemish-type (mutation in codon 692);
- HCAA Dutch-type (mutation in codon 693);
- HCAA Italian-type (mutation in codon 693);
- HCAA Arctic-type (mutation in codon 693);
- HCAA Iowa-type (mutation in codon 694);
- HCAA Piedmont-type or Italian-type II (mutation in codon 705);
- HCAA Italian-type III (mutation in codon 713).

HCAAs of Italian-, Piedmont- and Dutch-types are associated primarily with lobar cerebral haemorrhages, white matter lesions, and cognitive impairment, and only variable extents of the plaque and neurofibrillary pathologies characteristic of AD (Zhang-Nunes et al. 2006). HCAAs of Italian III-, Iowa-, Arctic-, Flemish- types and HCAA associated with codon 678 mutation are associated both with CAA (and sometimes ICH) and with dementia, often of the AD type (Kamp et al. 2014). Compared to sporadic CAA, the HCAAs generally have a more severe clinical course, with an earlier age of onset and earlier age of death (Zhang-Nunes et al. 2006). Overall the similarities between HCAAs and sporadic forms of CAA are more striking than the differences.

Over-expression of the wild-type APP without amino acid substitution in the protein sequence also results in severe parenchymal and vascular amyloid-β deposition. This can be seen in FAD caused by duplication of the APP gene, and in Down's syndrome (Revesz et al. 2009).

A number of other proteins are also involved in rare familial diseases in which CAA is a characteristic morphological feature. These proteins include: the ABri and ADan subunits in *BRI2* gene-related familial British and familial Danish dementia; variant cystatin C in *cystatin C* gene-related HCHWA-Icelandic type; variant transthyretins in *TTR* gene-related meningo-vascular amyloidosis; disease-associated prion protein (PsP^{Sc}) in *PRPN* gene-related hereditary prion disease with premature stop codon mutations; and mutated gelsolin (AGel) in *gelsolin* gene-related familial amyloidosis of Finnish type (Revesz et al. 2009). In the case of these rare familial diseases, mainly vascular deposits of the pathological protein can be found not only in the central nervous system, but also in a variety of peripheral tissues and systemic organs (Revesz et al. 2009).

There is phenotypic variability between HCAAs caused by different mutations, which is at least partly explained by altered *APP* processing and the varying properties of the respective mutant peptides compared to the wild-type amyloid-β (Zhang-Nunes et al. 2006). However, the phenotypic expression is also very variable between cases affected by the same mutation, suggesting modulation by additional factors. For example, in cases with HCAA-Dutch type, the first and best described HCAA, affecting three large families originating from the coastal villages of Katwijk and Scheveningen in the Netherlands, clinical features such as age at onset, age at death, occurrence of dementia and the number of strokes vary considerably between the cases (Bornebroek et al. 1997b, Kamp et al. 2014, Zhang-Nunes et al. 2006). In addition, extent and severity of radiological and pathological features of HCAA can vary (Bornebroek et al. 1997a, Natté et al. 2001). Another example of phenotypic heterogeneity between carriers of the same mutation is HCAA-Iowa type. Affected members of the Iowa family carrying the mutation reported no episodes of clinically manifest ICH, while a second family from Spain carrying the same mutation demonstrated ICH in three of the four affected members (Zhang-Nunes et al. 2006) and a recently reported Polish family also presented with ICH (Iwanowski et al. 2015)

Various factors, hypothesised to modulate the disease phenotype in cases with HCAA Dutch-type, have been studied (Zhang-Nunes et al. 2006). The mortality rate

is shown to be higher in females, and when the disease has been paternally transmitted (Bornebroek et al. 1997b). Genetic polymorphisms in the *APOE* and *PS1* genes have also been investigated, and shown not to influence the clinical phenotype of the disease (Bornebroek et al. 1997a, Haan et al. 1994, Zhang-Nunes et al. 2006).

Understanding the mechanisms of hereditary forms of a disease may contribute towards better understanding the mechanisms of sporadic forms of the same disease. My systematic review and meta-analyses in Chapter 3 demonstrated a dose-dependent association between histopathologically confirmed sporadic CAA and *APOE* ε4 containing genotypes. I therefore aimed to study how the *APOE* genotype affects the variation in phenotypic presentation (in this case the severity of cerebral amyloid angiopathy) of hereditary diseases where CAA is a prominent feature.

In this chapter, I aim to assess the evidence for an association between the *APOE* ε2/ε3/ε4 polymorphism and severity of CAA in brain pathology samples of cases with genetically proven HCAA and FAD by performing a systematic review of relevant published data.

5.2 Methods

5.2.1 Search strategy

I sought all studies of adult humans published in any language, in which cases with a confirmed genetic mutation causing HCAA and/or FAD had been genotyped for the APOE ε2/ε3/ε4 polymorphism and had CAA assessed pathologically (using autopsy or biopsy).

I used the same search strategy as in Chapter 3, searching OVID Medline (1950 to March 2012) and Embase (1980 to March 2012), using a combination of search terms for APOE, genes and CAA (Appendix 3). I also checked the bibliographies of all relevant studies and reviews identified, and searched *Google Scholar* for studies citing relevant studies.

5.2.2 Inclusion/exclusion criteria

For the analyses reported in this chapter, I included all studies in which >1 case with a genetic mutation causing HCAA and/or FAD had been genotyped for the *APOE* ε2/ε3/ε4 polymorphism and had CAA assessed pathologically (using autopsy or biopsy), regardless of whether any association between the polymorphism and CAA was reported on. A clinical research fellow colleague (Dr Neshika Samarasekera) and I independently selected eligible studies, resolving disagreements by discussion.

5.2.3 Data extraction

For each study included, I extracted information on:

- first author;
- publication year;
- participant source and description;
- genetic mutation causing HCAA and/or FAD;
- number of cases in whom the *APOE* ε2/ε3/ε4 polymorphism was genotyped and CAA was assessed pathologically;
- if available, any data about studying the association between the *APOE* ε2/ε3/ε4 polymorphism and severity of CAA on pathological assessment, and information about the scale used to assess CAA severity.

5.3 Results

5.3.1 Included studies

Of 1754 publications identified for screening I included 24 relevant studies (Figure 5.1) (Bornebroek et al. 1997a, Bornebroek et al. 1997b, Bugiani et al. 2010, Cabrejo et al. 2006, Dorfman et al. 2010, Fukutani et al. 1997, Gustafson et al. 1998, Halliday et al. 1997, Kumar-Singh et al. 2002, Lemere et al. 1996, Mann et al. 1996a, Mann et al. 1996b, Mann et al. 1996c, Mann et al. 2001, Natté et al. 2001, Nochlin et al. 1998,

Roks et al. 2000, Rosenberg et al. 2000, Rovelet-Lecrux et al. 2006, Singleton et al. 2000, Tomidokoro et al. 2010, Vinters et al. 1998, Yasuda et al. 2000, Yokota et al. 2002).

5.3.2 Study characteristics

Most studies were either clinico-pathologic descriptions of affected families, or pathologic studies of archival autopsy specimens of individuals carrying a mutation causing HCAA/FAD.

The total number of eligible non-overlapping cases (genetically confirmed mutation causing HCAA or FAD, genotyped for the *APOE* ε2/ε3/ε4 polymorphism and CAA assessed pathologically) across studies was between 92 and 197 (Table 5.1). There was uncertainty about the exact number of cases because the total number of cases in two studies was unclear, and the number of non-overlapping eligible cases across studies could not be confidently calculated due to limited information available from the publications. The number of eligible cases per study ranged from two to 52. The median number of eligible cases per study was four (Table 5.1).

Eight studies included cases of HCAA due to a mutation in codons 692-694 of the *APP* gene on chromosome 21: four studies included cases of HCAA-Dutch type; two studies cases of HCAA-Flemish type; and two studies cases of HCAA-Italian and HCAA-Iowa type. 16 studies included cases of FAD due to a mutation in the *APP* (codons 670, 671, 717 or duplications), *PS1* or *PS2* genes on chromosomes 21, 14 or 1 respectively: six studies included cases of FAD due to a mutation in *APP*; nine studies due to a mutation in *PS1*; and one study due to a mutation in *PS2*.

5.3.3 Association between *APOE* ε2/ε3/ε4 polymorphism and severity of hereditary CAA

Only three of the 24 eligible studies reported that they had studied the association between *APOE* ε2/ε3/ε4 polymorphism and severity of CAA on pathology (Bornebroek et al. 1997a, Natté et al. 2001, Mann et al. 2001).

The first study included eight eligible cases of HCAA-Dutch type (Bornebroek et al. 1997a), the second study included 14 to 16 eligible cases of HCAA-Dutch type (Natté et al. 2001), and the third study included 52 cases with FAD due to a mutation in the *PS1* gene (Mann et al. 2001). Two of these three studies included partly overlapping cases of HCAA-Dutch type (Bornebroek et al. 1997a, Natté et al. 2001). The total number of non-overlapping cases in these three studies was between 66 and 75, but the exact number of non-overlapping cases could not be identified based on information available from the publications.

In the Bornebroek et al. (1997a) study, CAA severity was rated semi-quantitatively on a previously published scale (Ellis et al. 1996) as follows:

- 0 = no CAA;
- 1 = sparse or mild CAA;
- 2 = moderate CAA;
- 3 = frequent or severe CAA.

In the Natté et al. (2001) study, CAA amount was measured semi-quantitatively with computerised morphometry following immunohistochemistry by calculating:

- the percent of cortical area occupied by amyloid-β positive vessel walls;
- the lumen/vessel diameter ratio of amyloid-β positive cortical vessels.

In the Mann et al. (2001) study, CAA severity was measured on a scale as follows:

- 1 = few leptomeningeal vessels weakly or patchily stained;
- 2 = few leptomeningeal vessels strongly or evenly stained, with mild intracortical vascular involvement;
- 3 = many leptomeningeal and intracortical vessels patchily or strongly stained;
- 4 = many leptomeningeal and intracortical vessels strongly or evenly stained, with dysphoric angiopathy.

The Bornebroek et al. study found that all included cases had severe CAA in all cortical areas and concluded that modulation by *APOE* genotype was impossible to detect, but unlikely (Bornebroek et al. 1997a). The Natté et al. and Mann et al. studies made a general statement about the association between the *APOE* genotype and CAA severity, but did not show raw data or describe exact statistical methods used to test the association, probably because this was not the primary aim of the studies (Natté et al. 2001, Mann et al. 2001). Natté et al. (2001) reported that cases with an *APOE* ε4 allele did not have a higher amount of CAA compared to cases without an *APOE* ε4 allele. Mann et al. (2001) reported that there was no correlation between possession of the ε4 allele and CAA severity across all cases. It was unclear from the publications if the analyses looking at the association between CAA severity and *APOE* genotype had taken into account the possible effect of age of the cases, the duration of illness, and in the case of Mann et al. study, the location of the mutation along the *PS1* gene. All three studies may have included cases with CAA-related ICH as an unselected part of the spectrum of CAA severity, but the exact numbers of CAA-related ICH cases were not available.

No study commented about the association between the *APOE* ε2 allele and CAA severity.

5.4 Discussion

5.4.1 Main findings

With my thorough systematic search I identified 24 studies in >90 individuals with HCAA or FAD, with data on both *APOE* genotype and CAA histopathology assessment. Only three studies including >65 cases had investigated and reported on the association between *APOE* ε2/ε3/ε4 polymorphism and CAA severity in cases with HCAA and FAD. These three studies did not find an association between possession of the *APOE* ε4 allele and CAA severity in cases with HCAA Dutch-type and *PS1*-related FAD. There were too few studies and cases, and insufficient data

about the methods used to test for an association, to draw firm conclusions. There was no data about the association between *APOE* ε2 and CAA severity.

To my knowledge, this is the first systematic review looking at the association between *APOE* ε2/ε3/ε4 genetic polymorphism and pathological CAA severity in cases with HCAA and FAD. Previous studies in cases with HCAA Dutch-type have shown that age of onset, age at death, occurrence of dementia, number of strokes, WMH lesion volume, focal lesions on MRI scan, and number of amyloid-β plaques in brain tissue are not influenced by the *APOE* ε2/ε3/ε4 genotype (Bornebroek et al. 1997a, Haan et al. 1995, Haan et al. 1994). Furthermore, no association between the *APOE* ε4 allele and amyloid-β plasma levels was found in 22 HCAA Dutch-type patients (Bornebroek et al. 2003). Also for HCAA Flemish-type, there is no evidence that *APOE* genotype influences the age at onset of dementia or ICH, though the statistical power of analyses in these studies has been low (Haan et al. 1994, Roks et al. 2000). In contrast with these findings however, is a study showing that cultures of human brain pericytes with an ε4/ε4 genotype exhibited more Dutch-type amyloid β induced cell death than cultures with other *APOE* genotypes (Verbeek et al. 2000).

Since my systematic review, I am aware of one further published study where association between CAA severity and presence of an *APOE* ε4 allele was investigated in ten FAD cases carrying a *PS1* mutation. It reported a significantly greater proportion of cortical vessels (but not leptomeningeal vessels) affected by CAA in ε4 carriers. However the study included only three *APOE* ε4 carriers and correction for multiple testing was not undertaken, so these results should be interpreted with caution (Ryan et al. 2015).

One possible explanation for the lack of apparent influence of the *APOE* genotype on CAA severity is that apoE may have a lower binding affinity for the mutant amyloid-β and therefore won't influence amyloid-β deposition (Bornebroek et al. 1997a). Alternatively, the studies that investigated this were not large enough to detect the association.

Interestingly, a recent study on genetic modifiers in CADASIL, a non-amyloidogenic hereditary SVD, found the *APOE* ε2 allele (but not the ε4 allele) to be an independent risk factor for higher WMH volumes in *NOTCH3* mutation carriers. This further supports the need to investigate the associations of HCAA and FAD severity with not only ε4, but also ε2 allele, and include larger numbers of affected cases (Gesierich et al. 2015).

5.4.2 Strengths and limitations

The strengths of my study include a thorough search strategy and strict case inclusion criteria requiring confirmation of the mutation on genetic testing. My systematic review is also strengthened by the identification and characterisation of the 21 studies that did not report investigating the association between the *APOE* genotype and CAA severity, but had the necessary data to do so. By contrast, many meta-analyses do not account for otherwise relevant studies without the necessary data within their publications and so risk undetected reporting bias, since many studies report only significant findings and fail to mention non-significant ones. It was not possible to calculate the exact proportion of eligible cases for whom there was data about the association, out of all eligible cases in the 24 studies, but it was in the range of 30% to 80%. Identifying these studies also opens up the possibility to contact the authors of the studies directly, to enquire about any unpublished relevant data. In the case of my study, I decided contacting the authors of the 21 studies was extremely unlikely to change my conclusions, considering the number of included cases in these 21 studies was very small, ranging from two to eight (median 3.5).

My study has some limitations. First, despite including data from all relevant studies from the published literature, because HCAA and FAD are relatively rare conditions, the number of included cases was small. Second, there were no data about the association between *APOE* and CAA severity in cases with HCAA and FAD due to mutations other than *PS1* and *APP* Dutch-type. Third, the three studies that had looked at the association probably included cases with CAA-related ICH as a spectrum of CAA severity. The exact proportion of these cases was not available from the publications, but it is likely to be high (especially in the case of HCAA

Dutch-type) considering that CAA-related ICH is one of the reasons leading to death and autopsy in these cases. Depending on the proportion of ICH cases, these studies may not be able to distinguish a genetic association with CAA severity from an association with CAA-related ICH. Fourth, the scales of assessing CAA severity in pathology samples were variable, and as demonstrated by the Bornebroek et al study, cases of HCAA often have severe CAA anyway (Bornebroek et al. 1997a) so modulation by *APOE* genotype may be difficult to detect. It may therefore be more appropriate to use a scale that also takes into account vasculopathic changes associated with severe CAA as a way of quantifying severe CAA, as suggested by Vinters et al (1998). Fifth, it was unclear from the publications if the analyses looking at the association between CAA severity and *APOE* genotype had taken into account the possible effect of age of the cases and the duration of illness. Sixth, the study that looked at the association in cases with a *PS1* mutation included 52 cases with mutations at different locations along the *PS1* gene, which may mean slightly different phenotypes were being compared. Mann et al (2001) even report an association between CAA severity and mutation location in the *PS1* gene (mutations between codons one to 200 were associated with less severe CAA, compared to mutations after codon 200). It was unclear from the publication if this was taken into account when the association of CAA severity with *APOE* was studied.

5.4.3 Outstanding issues

Since cases of HCAA and FAD often have severe CAA anyway, it may be more appropriate for future studies to use a scale that quantifies severe CAA further by taking into account vasculopathic changes. This may help detecting modulation by *APOE* genotype and other factors.

Studies did not report looking at the association between *APOE* ε2 and CAA severity. This would be an interesting question to look at, considering the generally accepted but not proven theory addressed in Chapter 4, suggesting that while *APOE* ε4 promotes vascular amyloid deposition, ε2 promotes progression to severe CAA with associated vasculopathy leading to vessel rupture and ICH. One intriguing

observation is the different phenotype of the same HCAA Iowa-type mutation in families in Iowa and Spain – three out of four affected Spanish family members presented with haemorrhagic stroke while none of the 12 Iowa family members had a haemorrhagic stroke. It has been speculated that this is possibly explained by the presence of *APOE* ε2 in the Spanish proband with haemorrhagic stroke and *APOE* ε4 in the Spanish brother and the Iowa family members without ICH. However, in the absence of *APOE* data from other affected members in both pedigrees, the involvement of *APOE* in determining the phenotype of the Iowa mutation remains speculative (Greenberg et al. 2003). More recently, a Polish family with the Iowa mutation presenting with ICH has been reported, but the *APOE* genotype is not available from the publication (Iwanowski et al. 2015).

Future studies could also investigate the association between *APOE* and CAA severity in cases with HCAA and FAD due to mutations other than *PS1* and *APP* Dutch-type.

5.4.4 Conclusions

Unlike in the case of sporadic CAA, there is no evidence for an association between *APOE* ε4 and CAA severity in cases with HCAA and FAD. However, the data are sparse, largely due to the relative rarity of these conditions, and the studies that have investigated this have provided limited data in their publications, which limits interpreting the robustness of the results. There are no data about the association between *APOE* ε2 and CAA severity.

Table 5.1 Characteristics of studies included in the systematic review of *APOE* genotype associations with hereditary CAA severity

Phenotype	Mutation	First author, publication year	Brief description of study participants	No. confirmed cases: *APOE* & histopathology available	Reported association with *APOE*
HCAA-Dutch		Bornebroek 1997[1]	Case series of 420 individuals from six families in the Netherlands, including 187 affected individuals	? (max 37)[3]	-
	Chr.21 *APP* codon 693 Glu22Gln	**Bornebroek 1997[2]**	**31 cases in whom diagnosis was based on confirmation of a genetic mutation**	**8[3]**	**All eight brains had severe CAA in all cortical areas. This made modulation by *APOE* genotype impossible to detect but unlikely.**
		Vinters 1998	Archival autopsy specimens of 29 cases[4]	9[3]	-
		Natté 2001	**19 cases for whom brain tissue was available and clinical information allowed dementia to be either diagnosed or excluded**	**14-16[3]**	**Cases with *APOE* ε4 allele did not have a higher amount of CAA than cases without *APOE* ε4 allele**
HCAA-Flemish	Chr. 21 *APP* codon 692 Ala21Gly	Roks 2000	Nine symptomatic cases from one family and 21 relatives (five of them presymptomatic carriers of the mutation)	3[5]	-
		Kumar-Singh 2002	Three cases[4] from one family with biopsy and/or autopsy brain specimens available	3[5]	-
HCAA-Italian	Chr. 21 *APP* codon 693 Glu22Lys	Bugiani 2010	Case series of 20 symptomatic individuals and their relatives from four Italian families carrying the mutation	2	-

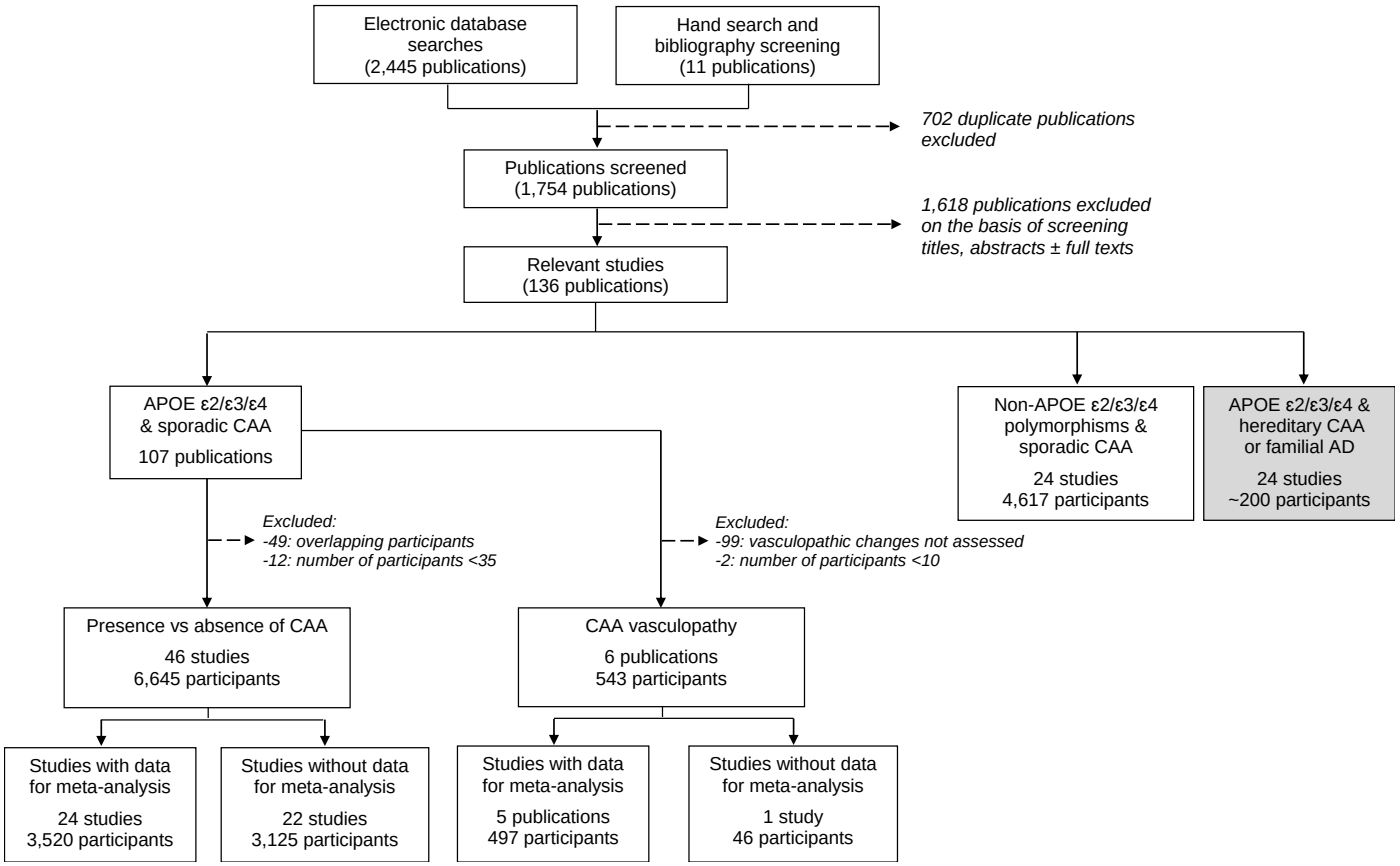
Phenotype	Mutation	First author, publication year	Brief description of study participants	No. confirmed cases: *APOE* & histopathology available	Reported association with *APOE*
HCAA-Iowa	Chr. 21 *APP* codon 694 Asp23Asn	Tomidokoro 2010	Frozen brain tissue obtained at autopsy from two cases with the mutation, and from three neuropathologically confirmed AD cases was studied and compared	2	-
FAD	Chr. 21 *APP* codon 717 and/or 670/671	Mann 1996[6]	Brain tissue from five autopsied cases of AD due to APP codon 717 mutation (4 Val=>Ile, 1 Val=>Gly), three cases of AD due to APP codon 670/671 mutation	8[7]	-
		Fukutani 1997	Compared cerebellar tissue from 10 sporadic AD cases, 10 FAD cases (including two AD cases due to APP codon 717 mutation (Val=>Ile)) and 10 controls	2[7]	-
		Halliday 1997	Studied brain tissue from two affected siblings from a family with the APP codon 717 (Val=>Ile) mutation	2	-
		Rosenberg 2000	Four genetically confirmed FAD cases from a family with APP codon 717 (Val=>Ile) mutation	4[7]	-
FAD	Chr. 14 *PS1*	Mann 1996[8]	156 consecutive typical AD cases and five cases with FAD (due to confirmed *PS1* mutation) examined histologically and compared	5[10]	-
		Mann 1996[9]	Brains of eight members from five families with chromosome 14 linked FAD	8[10]	-

Phenotype	Mutation	First author, publication year	Brief description of study participants	No. confirmed cases: *APOE* & histopathology available	Reported association with *APOE*
		Lemere 1996	Four autopsy brains of cases from a Colombian FAD family and 12 sporadic AD cases studied neuropathologically	3[10]	-
		Gustafson 1998	Describe a Swedish family with six cases of FAD with a novel mutation in *PS1*	3[10]	-
		Singleton 2000	Four cases of FAD due to a mutation in PS1 studied	4[10]	-
		Yasuda 2000	Report a case with a novel *PS1* mutation, and compare it to a case from a known FAD family carrying a different *PS1* mutation	2[10]	-
		Mann 2001	**54 cases from 28 families with *PS1* AD (25 separate mutations covering the length of *PS1)* studied and compared to 25 cases of sporadic AD**	**52[10]**	**Although some cases with severe CAA were bearers of the *APOE* ε4 allele, there was no correlation between possession of this and the extent of CAA across all cases**
		Yokota 2002	Clinicopathologic study of two AD cases from a pedigree with a *PS1* mutation	2[10]	-
		Dorfman 2010	Immunohistochemistry study of eight sporadic AD cases, eight FAD cases with confirmed *PS1* mutations and 11 controls	8[10]	-

Phenotype	Mutation	First author, publication year	Brief description of study participants	No. confirmed cases: *APOE* & histopathology available	Reported association with *APOE*
FAD	Chr. 1 *PS2*	Nochlin 1998	Report neuropathological features of a Volga German family with FAD due to a *PS2* mutation	4	-
FAD	Chr. 21 *APP* duplication	Cabrejo 2006	Describe the phenotype of 21 cases of FAD due to *APP* duplication from five affected French families	5[11]	-
		Rovelet-Lecrux 2006	Report five French families with FAD due to *APP* duplication	5[11]	-

HCAA: hereditary cerebral amyloid angiopathy; FAD: familial Alzheimer´s disease; No: number; APP: amyloid β precursor protein; PS: presenilin; [1] (Bornebroek et al. 1997b); [2] (Bornebroek et al. 1997a); [3,5,7,10,11] Include overlapping cases – cannot identify exact number of overlaps based on information available from the publications; [4] based on context and references, assume genetically confirmed cases, though not explicitly stated in the publication; [6] (Mann et al. 1996b); [8] (Mann et al. 1996c); [9] (Mann et al. 1996a).

Figure 5.1 Selection of studies included in the systematic review of *APOE* genotype associations with hereditary CAA severity



Grey box represents the selection of studies relevant for Chapter 5.

Chapter 6: Other genetic polymorphisms (non-*APOE* ε2/ε3/ε4) associated with sporadic CAA

Chapter contents:

6. Other genetic polymorphisms (non-*APOE* ε2/ε3/ε4) associated with sporadic CAA

- Identifying genetic polymorphisms associated with histopathologically confirmed CAA should increase understanding of the underlying mechanisms
- Several candidate genes have been investigated, often selected based on their known involvement with AD and the amyloid processing pathway
- I performed a systematic review to study the associations between any (non-*APOE* ε2/ε3/ε4) genetic polymorphism and histopathological CAA
- I conclude that there are too few studies and participants to draw firm conclusions about the effect of non-*APOE* ε2/ε3/ε4 genetic polymorphisms on CAA, but some positive associations (with *TGF-β1, TOMM40* and *CR1* genes*)* merit replication in further larger studies

6.1 Introduction

Sporadic cerebral amyloid angiopathy (CAA) is characterised by deposition of amyloid-β protein in leptomeningeal and cortical blood vessels, with a prevalence in population-based autopsy studies of 20-40% in non-demented and 50-60% in demented elderly people (Charidimou et al. 2012a). CAA is associated with increasing age, dementia, lobar intracerebral haemorrhage (ICH), lobar brain microbleeds, white matter hyperintensities, small cortical infarcts and superficial siderosis (Linn et al. 2010, Masuda et al. 1988, Neuropathology Group MRC CFAS 2001, Samarasekera et al. 2012). Identifying genetic polymorphisms associated with histopathologically confirmed CAA should increase understanding of the mechanisms leading to CAA and associated diseases.

While *APOE* is the most widely studied candidate gene, several other candidate genes have also been investigated for an association with pathologically confirmed CAA. These candidate genes have mainly been selected based on their known

association with Alzheimer's disease, or in some cases, based on their involvement in the amyloid processing pathway.

In this chapter, I aim to assess the evidence for associations between any non-*APOE* ε2/ε3/ε4 genetic polymorphisms and histopathologically confirmed CAA by carrying out a systematic review.

6.2 Methods

6.2.1 Search strategy

I sought all studies of adult humans published in any language, in which participants had been genotyped for any genetic polymorphism and had CAA pathology assessed (using autopsy or biopsy), regardless of whether any association between the polymorphism and CAA was reported on.

I used the same search strategy as in Chapter 3, searching OVID *Medline* (1950 to March 2012) and *Embase* (1980 to March 2012), using a combination of search terms for APOE, genes and CAA (Appendix 3). I also checked the bibliographies of all relevant studies and reviews identified, and searched *Google Scholar* for studies citing relevant studies.

6.2.2 Inclusion/exclusion criteria

For the analyses reported in this chapter, I included studies where participants had been genotyped for any non-*APOE* ε2/ε3/ε4 genetic polymorphism and had CAA assessed pathologically. I excluded studies that assessed genetic associations with CAA-related ICH versus CAA-free controls, because these would not be able to distinguish a genetic association with CAA from an association with ICH. However, I included studies that had recruited participants with CAA-related ICH as an unselected part of the spectrum of CAA severity. A clinical research fellow colleague (Dr Neshika Samarasekera) and I independently selected eligible studies, resolving disagreements by discussion and mutual consensus.

6.2.3 Data extraction

For each study included I extracted information on:

- first author;
- publication year;
- participant source and description;
- study size;
- gene(s) and polymorphism(s) studied;
- association between the polymorphism studied and pathological CAA.

6.3 Results

6.3.1 Included studies

Of 1754 publications identified for screening, 26 studies were relevant (Figure 6.1). I excluded two studies because they compared genetic associations with CAA-related ICH versus CAA-related ICH free controls (McCarron et al. 2003, Zunarelli et al. 1997). I included the remaining 24 studies in 4617 participants (Figure 6.1) (Biffi et al. 2012, Chalmers et al. 2004, Chapuis et al. 2009, Chapuis et al. 2006, Christoforidis et al. 2005, Cuenco et al. 2008, Davidson et al. 2006, Durany et al. 2000, Hamaguchi et al. 2005, Kölsch et al. 2007, Lendon et al. 2002, Pahnke et al. 2003, Peila et al. 2007, Peuralinna et al. 2011, Shi et al. 2006, Zubenko et al. 1999, Thal et al. 2010, Valant et al. 2012, Yamada et al. 2002, Yamada et al. 1997, Yamada et al. 1998a, Yamada et al. 1998b, Yamada et al. 1999, Yamada et al. 2003).

6.3.2 Study characteristics

All studies used autopsy brains from a brain bank, clinical autopsies or brain pathological material (mainly autopsy, some biopsy) arising during the course of a population-based prospective study. The median number of eligible participants per study who were both genotyped and assessed pathologically was 150. Across the 24

studies polymorphisms in 25 different genes were studied. A polymorphism in the *LRP1* gene was investigated in three studies including 597 participants in total. Polymorphisms in five genes (*TGF-β1*, *ACT*, *CYP46*, *ACE1*, *APOE* promoter) were investigated in two studies each, including 187 to 524 participants per gene in total. All other polymorphisms were investigated in only one study and the number of participants per study ranged from 50 to 723 per study (Table 6.1).

6.3.3 Association between non-*APOE* ε2/ε3/ε4 polymorphisms and CAA

There were insufficient data for formal meta-analyses, but there were some positive associations with CAA. There was a consistent trend towards an association with the T allele of rs1800470 in the transforming growth factor β1 gene (*TGF-β1*) in two studies including 449 participants. One study in 723 participants demonstrated an association between 14 SNPs in the translocase of outer mitochondrial membrane 40 gene (*TOMM40*) and CAA severity. One study found that the A allele of rs6656401 in the complement component receptor 1 gene (*CR1*) was associated with more severe CAA. Other studies found no overall significant associations although some reported associations in particular subgroups (Table 6.1, Table 6.2).

6.4 Discussion

6.4.1 Main findings

There were too few studies and participants to draw firm conclusions about the effects of non-*APOE* ε2/ε3/ε4 genetic polymorphisms on histopathological CAA. There were some positive associations – with *TGF-β1*, *TOMM40* and *CR1* genes – which merit replication in further larger studies or pooled datasets (Rannikmäe et al. 2013) (Appendix 7).

The transforming growth factor β1 is multifunctional cytokine, which is expressed in the periphery, and by glia and neurons in the brain. It is a potent regulator of injury

and inflammatory responses in the central nervous system. In experimental studies, TGF-β1 mRNA levels and protein production increase after a cerebral ischaemic event. It appears to protect the area surrounding the ischaemic region by maintaining neuronal viability and function through an interaction with microglia and astrocyte metabolism. In contrast to its neuroprotective properties after ischaemic events, *TGF-β1* on chromosome 19 has been implicated in the neurotoxic events associated with the Alzheimer's disease amyloid cascade. However, its exact role in the process is still under investigation. The C allele of rs1800470 causes an amino acid substitution at codon 10, and has been associated with a higher level of TGF-β1 mRNA and protein compared to the wild-type T allele. The C allele has been associated with a reduced risk of myocardial infarction, rheumatoid arthritis and osteoporosis. The possible reported association between the T allele of the *TGF-β1* gene and CAA may be mediated by the lack of protective action on the endothelium, by influence on Aβ clearance and deposition through activation of astrocytes and microglia, or by some other as yet unknown mechanism (Hamaguchi et al. 2005, Peila et al. 2007).

The *TOMM40* gene on chromosome 19 encodes a subunit of the translocase of the outer membrane complex Tom40. It lies in the mitochondria, and is involved in regulating protein traffic across the outer mitochondrial membrane. Some studies have suggested that genetic variation in *TOMM40* is associated with AD risk. Valant et al. showed that genetic variants in *TOMM40* were also associated with CAA severity, but not with CAA-related ICH. This is interesting, and supports the theory that the biological processes involved in vessel rupture differ from those implicated in the initial vascular amyloid deposition. The effect of *TOMM40* on CAA may be mediated by its effects on mitochondrial metabolism (Valant et al. 2012).

The complement component receptor 1 gene *(CR1)* is a member of the receptors of complement activation (RCA) family and is located in the 'cluster RCA' region of chromosome 1. It is associated with AD in GWAS studies. The association of a *CR1* gene polymorphism with CAA may be via clearance of Aβ peptide (Chibnik et al. 2011).

Since this systematic review, a genome-wide association meta-analysis of neuropathologic features of AD, including of cerebral amyloid angiopathy, has been undertaken, including just over 2800 individuals (Beecham et al. 2014). It does not confirm the association of *CR1* with neuropathological CAA. This may be due to a relatively modest sample size for a GWAS study, or suggest a spurious association in the previous study (which likely includes overlapping subjects) (Biffi et al. 2012). A genome-wide significant association with CAA is identified for rs6857 in the *APOE* region. This SNP is also significantly associated with AD and other neuropathological features of AD, so the association with CAA may be explained by ascertainment bias. No other significant genome-wide associations are observed for CAA.

6.4.2 Strengths and limitations

To my knowledge, this is the first systematic review of associations between non-*APOE* ε2/ε3/ε4 genetic polymorphisms and pathologically proven CAA. My study benefits from thorough ascertainment of pertinent studies, which overcomes potential citation bias (the citation or non-citation of research findings, depending on the nature and direction of the results) that is very likely to affect the conclusions in non-systematic reviews (Greenberg 2009).

The main limitation was that there were insufficient data for formal meta-analyses and for drawing firm conclusions about associations. There were also not enough data to draw conclusions about potential biases associated with these candidate-gene studies, but reporting biases (when the dissemination of research findings is influenced by the nature and direction of results) are likely to occur (The Cochrane Collaboration 2011). More specifically, the results of this systematic review may be affected by:

- publication bias: the publication or non-publication of findings depending on the nature and direction of the results

and

- outcome reporting bias: selective reporting of some outcomes but not others, depending on the nature and direction of the results.

Finally, the included studies were all candidate-gene studies, where small relative risks may not be detected as confounding and misclassification are more likely to obscure small to moderate relative risks than larger relative risks. Another possible limitation of these studies relates to the selection of candidate genes, which is often done on the basis of imperfect understanding of biological pathways (Manolio et al. 2009, Tabor et al. 2002) and is prone to be affected by citation bias.

6.4.3 Conclusions

There were too few studies and participants to draw firm conclusions about the effect of non-*APOE* ε2/ε3/ε4 genetic polymorphisms on CAA, but there were some positive associations which merit replication in further larger studies or pooled datasets.

Table 6.1 Characteristics of studies included in the systematic review of non-*APOE* ε2/ε3/ε4 genetic polymorphisms and CAA

First author, publication year	Brief description of participants	Study size[1]	Gene(s) studied	Polymorphism(s) studied	Reported association with CAA[2]
Biffi 2012	Participants from two longitudinal studies (Religious Orders Study, Rush Memory and Aging Project)	544	*CR1*	rs6656401	A allele associated with more severe CAA
Chalmers 2004	Brain tissue from the South-West Brain Bank (86 AD cases, 58 controls)	144	*APOE* promoter	rs449647 rs405509	No association with CAA severity
			ACE1	287 bp ins/del in intron 16	
			CYP46	rs754203	
Chapuis 2006	Autopsy brains from AD cases in UK Greater Manchester region	76	*VEGF*	rs699947	No association with CAA severity
Chapuis 2009	Autopsy brains from AD cases	114	*IL-33*	rs1157505 rs11792633 rs7044343	No overall association with CAA severity (rs11792633 and rs7044343 associated with CAA severity only in non-*APOE* ε4 carriers)

First author, publication year	Brief description of participants	Study size[1]	Gene(s) studied	Polymorphism(s) studied	Reported association with CAA[2]
Christoforidis 2005	125 consecutive routine autopsy brains from University of Leipzig, (8 neuropathologically confirmed AD cases)	125	*LRP1*	rs1799986	No overall association with CAA severity (C/C genotype associated with CAA severity only in leptomeningeal vessels)
				5' region tetranucleotide repeat TTTC (83,87,91,95 bp)	No overall association with CAA severity (87bp polymorphism associated with CAA presence versus absence in cortical vessels)
Cuenco 2008	Pathologically confirmed AD cases from Edith Norse Rogers Veterans Affairs Center (USA)	69	*SORL1*	6 three SNP combinations[3]	No associations with CAA
Davidson 2006	Neuropathologically confirmed AD cases from the UK	146	*CTSD*	rs17571	No association with CAA severity
Durany 2000	Neuropathologically confirmed AD cases from a general hospital	32	*ACT*	probably rs4943[4]	No association with presence of CAA
Hamaguchi 2005	Consecutive autopsy brains from a Japanese geriatric hospital (73 neuropathologically confirmed AD and 94 non-AD cases)	167	*TGF-β1*	rs1800470	Dose-dependent association between T allele and CAA severity across all cases
Kölsch 2007	Consecutive autopsy brains from participants ≥60 years	79	*GSTO1*	rs4925	No association with CAA severity
			IL-1A	rs1800587	No data about association reported
			IL-1B	rs1143634	No data about association reported
			IL-1B	rs3087258	No data about association reported

First author, publication year	Brief description of participants	Study size[1]	Gene(s) studied	Polymorphism(s) studied	Reported association with CAA[2]
Lendon 2002	Autopsy brains of AD cases	95	*ACE1*	287 bp ins/del in intron 16	No association with CAA severity
Pahnke 2003	Routine autopsy brains from clinics in Germany (participants free of overt dementia)	274	*APOE* promoter	rs449647	No data about association reported
Peila 2007	Autopsy samples from the community-based longitudinal Honolulu-Asia Aging Study	282	*TGF-β1*	rs1800470	TC and CC genotypes associated with reduced risk for CAA
				rs1800469	No association with CAA
Peuralinna 2011	Population-based autopsy sample of Finns ≥80 years from the Vantaa 85+ study	282	*AβPP*	60 SNPs[5]	No associations with CAA
			AβPP promoter	Sequenced	
Shi 2006	Autopsy series of AD cases from the Greater Manchester region of UK during 1986-2003	87	*OLR1*	rs12316150	No association with CAA severity
		89	*OLR1*	rs1050283	No overall association with CAA severity (CC genotype associated with increased Aβ40 load as CAA in the frontal cortex in *APOE* ε4- participants)
		92	*LRP1*	rs1799986	No association with CAA severity

First author, publication year	Brief description of participants	Study size[1]	Gene(s) studied	Polymorphism(s) studied	Reported association with CAA[2]
Thal 2010	Non-selected autopsy cases aged ≥50 years (71 AD cases and 309 controls) who had died in a European Union or USA hospital	380	*LRP1*	rs1799986	No overall association with CAA (T allele frequency increased in AD cases with non-capillary CAA/no CAA compared to capillary CAA cases)[6]
			CYP46	rs754203 rs4900442 rs7157609	No significant associations with CAA
			*CH25H*1*	rs13500	No significant association with CAA
			*CH25H*2*	rs1131706	T allele associated with capillary CAA presence
Valant 2012	Pathology samples from the Religious Orders Study and the Rush Memory and Aging Project	723	*TOMM40*	14 SNPs[7]	Associated with CAA severity[6]
Yamada 1997	Autopsy series[8] (36 neuropathological AD and 101 non-AD cases)	137	*PS1*	rs165932	No overall association with CAA severity (2/2 genotype less frequent in moderate/ severe CAA compared to slight/no CAA)
Yamada 1998a	Autopsy series[8] (48 neuropathological AD cases)	155	*ACT*	probably rs4943[4]	No overall association with CAA severity (correlation between A allele frequency and CAA severity in AD cases)
Yamada 1998b	Autopsy series[8] (48 neuropathological AD cases)	155	*BCHE*	rs1803274	No association with CAA severity

First author, publication year	Brief description of participants	Study size[1]	Gene(s) studied	Polymorphism(s) studied	Reported association with CAA[2]
Yamada 1999	Autopsy series[8] (68 neuropathological AD cases)	178	*A2M*	Exon 18 del (5 nucleotides in splice acceptor)	No association with CAA severity
Yamada 2002	Consecutive autopsy cases from a large geriatric hospital in Japan (47 neuropathological AD cases)	154	*PON1*	rs662	No association with CAA severity
Yamada 2003	Consecutive autopsy cases from a large geriatric hospital in Japan (75 neuropathological AD cases)	164	*NEP*	GT repeat in enhancer/ promoter region	No overall association with CAA severity (but increased severity in participants with ≤40 repeats compared to those with >40 repeats total)
Zubenko 1999	Autopsy cases from the Presbyterian University Hospital and University of Pittsburgh	50	*DXS1047* locus	202 bp allele	No association with CAA

Ins/del: insertion and deletion polymorphism; AD: Alzheimer's disease

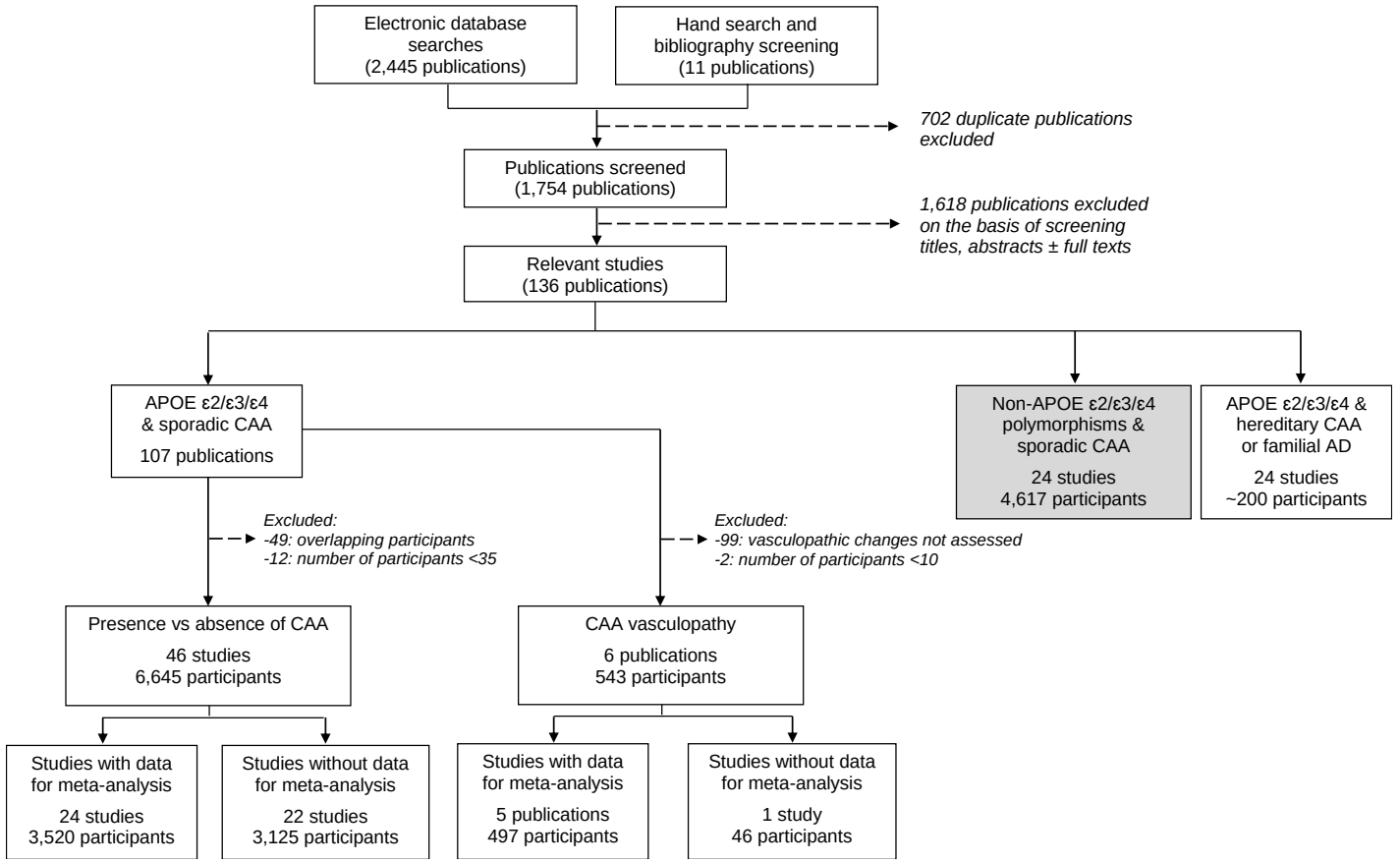
[1] Study size: number of eligible participants genotyped and assessed for pathologically-confirmed CAA; [2] If an association is reported, then the authors reported this as a statistically significant association; [3] rs4935774, rs578506, rs582446, rs668387, rs689021, rs641120, rs1699102, rs2276412, rs3824968, rs2282649, rs1010159, rs1784933; [4] Signal region of the gene → A/T alleles that determine the amino acid alanine or threonine; [5] rs1571660, rs2829929, rs1394839, rs2829946, rs2829949, rs7282634, rs2040273, rs8126493, rs4175, rs2829961, rs3787620, rs373521, rs1145, rs2829973, rs1783025, rs380417, rs1787438, rs1783016, rs214488, rs383700, rs2829984, rs2234983, rs216779, rs400603, rs440666, rs2014146, rs216762, rs2070657, rs2830000, rs7278838, rs2830008, rs2830012, rs768039, rs2830028, rs3991, rs2830033, rs2830036, rs2830038, rs7283136, rs2830040, rs2830044, rs2830051, rs2830052, rs11702267, rs2830054, rs3827166, rs2830058, rs2830068, rs466609, rs2830088, rs2186302, rs2830104, rs459543, rs45476095, rs466448, rs438031, rs463946, rs11911934, rs2156079, rs463164; [6]capillary CAA (or CAA-type 1): capillary Aβ-deposition present; non-capillary CAA (or CAA-type 2): no capillary Aβ-deposition present; [7] rs2075650, rs34404554, rs11556505, rs769449, rs12972156, rs12972970, rs157582, rs184017, rs157581, rs283815, rs157580, rs439401, rs34095326, rs10119; Unclear what the p-value for association with rs157580 was (contradictory information in the original publication); [8] from the Yokufukai Geriatric Hospital in Japan.

Table 6.2 Summary of reported associations between non-*APOE* ε2/ε3/ε4 genetic polymorphisms and CAA

Gene	Location / Polymorphism	No. of studies	No. of participants	Summary of results[1]
TGF-β1	rs1800470	2	449	Consistent trend for positive association between T allele and CAA
TOMM40	14 SNPs[2]	1	723	SNPs associated with CAA severity
CR1	rs6656401	1	544	A allele associated with more severe CAA
LRP1	rs1799986	3	597	No overall associations (inconsistent trends and in some cases associations in subgroups)
ACT	Probably rs4943[3]	2	187	
CYP46	rs754203	2	524	
ACE1	287 bp insertion/deletion in intron 16	2	239	
APOE promoter	rs449647	2	418	
Gene				
PS1 (presenilin-1)*; BCHE* (butyrylcholinesterase)*; DXS1047* locus*; APOE* promoter*; A2M* (α_2 macroglobulin)*; PON1* (paraoxonase)*; NEP* (neprilysin)*; OLR1* (oxidized low-density lipoprotein receptor 1)*; CYP46; CH25H*1* and *CH25H*2* (cholesterol 25-hydroxylase genes); *VEGF* (vascular endothelial growth factor); *IL-33* (interleukin genes); *GSTO1* (glutathione S-transferase omega-1); *SORL1* (sortilin related receptor); *CTSD* (cathepsin D); *AβPP* and *AβPP* promoter		18	50-380[4]	

TGF-β1: transforming growth factor β1; *TOMM40*: translocase of outer mitochondrial membrane 40; *CR1*: complement component receptor 1; *LRP1*: low-density lipoprotein receptor 1; *ACT*: α1 antichymotrypsin; *CYP46*: cholesterol 24-hydroxylase gene; *ACE1*: angiotensin 1 converting enzyme; *APOE*: apolipoprotein E;
[1]If an association is reported, then the authors have reported this as a statistically significant association
[2]rs2075650, rs34404554, rs11556505, rs769449, rs12972156, rs12972970, rs157582, rs184017, rs157581, rs283815, rs157580, rs439401, rs34095326, rs10119; Unclear what the p-value for association with rs157580 is (contradictory information in the original publication);
[3]Signal region of the gene → A/T alleles that determine the amino acid alanine or threonine
[4]Range of participant numbers in individual studies.

Figure 6.1 Selection of studies included in the systematic review of non-*APOE* ε2/ε3/ε4 genetic polymorphisms and CAA



Grey box represents the selection of studies relevant for Chapter 6.

Chapter 7: Contribution of variants in *COL4A1* and *COL4A2* genes to sporadic SVD

Chapter contents:

7. Contribution of variants in *COL4A1* and *COL4A2* genes to sporadic SVD

- Understanding genetic associations with SVD would provide insight into underlying mechanisms
- Genes causing rare Mendelian forms of a disease may also contain variants conferring risk for sporadic forms of a disease with a similar phenotype
- Mutations in *COL4A1* and *COL4A2* genes cause Mendelian subcortical SVD and hence I hypothesised that common variants in *COL4A1* and *COL4A2* are associated with sporadic subcortical SVD
- I conducted meta-analyses of GWAS data to determine associations of common SNPs in *COL4A1/COL4A2* with clinical and radiological manifestations of SVD
- I conclude that there is an association between common variation in *COL4A2* and sporadic cerebral SVD, particularly deep intracerebral haemorrhage
- These findings merit replication studies

7.1 Introduction

Genes causing rare familial forms of a disease may also contain variants conferring risk for sporadic forms of a disease with a similar phenotype. There are several examples of this. Polymorphisms in the *APP* gene associated with autosomal dominant familial Alzheimer's disease (AD) confer risk to sporadic forms of AD and age-related cognitive decline (Jonsson et al. 2012). Variants in the *TREX1* gene are associated with both familial chilblain lupus and sporadic forms of systemic lupus erythematosus (Lee-Kirsch et al. 2007). Common variants near *ABCG8* and *LCAT* genes reached genome-wide significance for an association with lipid levels, while rare variants in these genes are known to cause Mendelian forms of dyslipidaemia (Kathiresan et al. 2009). Whether genes involved in familial forms of cerebral SVD also confer risk to sporadic SVD remains uncertain. It has been suggested that rare variants in the *NOTCH3* gene, mutations in which cause cerebral autosomal

dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), are associated with white matter hyperintensities (WMH) in the general population (Schmidt et al. 2011). However, a more recent study does not support a role for common *NOTCH3* variation in the risk of sporadic SVD (Rutten-Jacobs et al. 2015).

Collagen IV is a major component of the basement membrane, a specialised extracellular matrix structure that compartmentalises tissues, provides structural support and influences cell behaviour. In blood vessel walls, the basement membrane separates smooth muscle cells from the endothelial cells lining the vessel lumen. Vertebrates express six collagen type IV chains (α1-α6(IV)), encoded by the genes *COL4A1-6.* In the endoplasmic reticulum, three distinct obligatory heterotrimeric molecules (protomers) are formed, each containing three alpha chains, α1.α1.α2(IV), α3.α4.α5(IV) and α5.α5.α6(IV), but only α1.α1.α2(IV) is present in the vasculature. The alpha chains contain a N-terminal 7S domain, a collagen domain consisting of G-X-Y repeats in which every third amino acid is a highly conserved glycine residue, and a C-terminal NC1 domain that initiates protomer formation. After secretion, protomers interact to form a collagen IV network within the basement membrane (Khoshnoodi et al. 2008).

The *COL4A1* and *COL4A2* genes are located in tandem on chromosome 13q34 and have a shared communal bi-directional promoter (Khoshnoodi et al. 2008). Dominant missense *COL4A1/COL4A2* mutations lead to a spectrum of Mendelian disorders, characterised by eye, kidney and cerebrovascular defects, including intracranial aneurysms and SVD. These disorders manifest clinically as porencephaly (cerebrospinal fluid filled cavities due to ICH), deep ICH, and lacunar ischaemic strokes, and radiologically as WMH and mainly subcortical brain microbleeds (Vahedi and Alamowitch 2011). The clinical phenotype appears to be influenced by the type and position of the mutation, as well as genetic and environmental modifiers (Van Agtmael and Bruckner-Tuderman 2010).

The mutations are associated with basement membrane defects, but can also lead to intracellular retention of mutant protein, swollen endoplasmic reticulum vesicles and

increased expression of chaperones and apoptosis. This suggests that endoplasmic reticulum stress, which is pathogenic when chronic, may also play a role in disease pathogenesis (Jeanne et al. 2012, Murray et al. 2014, Van Agtmael et al. 2010, Van Agtmael et al. 2005). However, the exact contributions of endoplasmic reticulum stress and basement membrane defects to disease development remain unknown. Recent data suggest that manipulation of chaperone levels, intracellular collagen accumulation, and endoplasmic reticulum stress (e.g. with 4-phenyl butyric acid), are potential therapeutic options for collagen IV diseases including haemorrhagic stroke (Murray et al. 2014).

There are some data to suggest that variants in *COL4A1* and *COL4A2* may also be associated with sporadic cardiovascular diseases. Common SNPs in *COL4A1* may be associated with the presence of intracranial aneurysms (Ruigrok et al. 2006), vascular stiffness, coronary artery calcification and coronary artery disease (Adi et al. 2013, O'Donnell et al. 2011, Schunkert et al. 2011, Tarasov et al. 2009, Yamada et al. 2008). Furthermore, rare mutations in *COL4A1* and *COL4A2* have recently been shown to be associated with sporadic ICH (Jeanne et al. 2012, Schmidt et al. 2011, Weng et al. 2012).

I therefore hypothesized that variants in the *COL4A1* and *COL4A2* genes may confer risk for sporadic cerebral SVD.

In this chapter, I aim to investigate whether there is an association between common SNPs in collagen genes and sporadic SVD. I do this by conducting meta-analyses of GWAS data among individuals of European ancestry, to determine associations of common SNPs in the *COL4A1/COL4A2* genomic region with clinical and radiological manifestations of SVD (deep ICH, lacunar ischaemic stroke and white matter hyperintensities).

7.2 Methods

7.2.1 Identifying participating studies

Through the International Stroke Genetics Consortium (ISGC) (http://www.strokegenetics.org/)[1], a non-systematic literature review, and personal contact with researchers in the field of stroke genetics, I identified the large consortia that hold GWAS data for the phenotypes that were of interest to me: intracerebral haemorrhage (ICH) and ischaemic stroke (IS) including their subtypes, and WMH. I approached these consortia with formal collaboration proposals and/or by emailing the principal investigator(s) directly, describing my proposed project and inviting them to collaborate by sharing data. I participated in the relevant teleconference calls to introduce my project and answer any questions.

The consortia and investigators approached were all very positive and agreed to participate in this project. As a result I established an international collaborative group that included four large consortia (partly united under the ISGC umbrella): the ISGC ICH GWAS consortium (Woo et al. 2014), the Metastroke consortium (Traylor et al. 2012), the WMH in Ischaemic Stroke GWAS Collaboration (Adib-Samii et al. 2013) and the CHARGE consortium WMH Group (Fornage et al. 2011) [http://web.chargeconsortium.com/]). These consortia included data from the majority of large GWAS studies of relevant cerebrovascular phenotypes in individuals of European ancestry, available at the time of starting this project in 2012: ICH and its subtypes (deep and lobar), ischaemic stroke and its TOAST subtypes (large vessel disease [LVD], cardio-embolic [CE] and small vessel disease/lacunar IS), WMH volume in ischaemic stroke cases and WMH volume in population-based studies (Table 7.1).

[1] Professor Cathie Sudlow is on the ISGC steering committee and I am an ISGC member.

7.2.2 Data collection

I collected study characteristics data and genotype-phenotype association data from relevant publications, by direct communication with study and/or consortium leads, and other study team members. I also undertook a visit to a collaborator's (Professor Jonathan Rosand) laboratory in the Center for Human Genetic Research at the Massachusetts General Hospital in Boston, USA, to access some of the data and perform some of the relevant analyses.

7.2.2.1 Study characteristics

I extracted information for each study contributing data to the analysis about:

- overall study characteristics: study design, inclusion and exclusion criteria, phenotype definition and diagnostic criteria, and, where relevant, methods of phenotype measurement;
- study participants' characteristics: ancestry, numbers of cases ± controls for each phenotype, mean age and standard deviation of cases ± controls;
- genotyping and association analysis methods: GWAS panel used, quality control steps, imputation (software, reference panel, quality parameters), software and statistical model used for association analyses, handling of population stratification, Hardy-Weinberg equilibrium, adjusting for cryptic or overt relatedness, strand and build of the human genome on which results provided.

7.2.2.2 Genotype data

I collected existing GWAS data from participating studies for the *COL4A1* and *COL4A2* genomic region of ~464 kbp, which includes >1000 single nucleotide polymorphisms (SNPs) covering the bi-directional communal promoter, introns, exons and a 50kbp upstream and downstream flanking region (Human Genome reference build 19 [hg19] coordinates 110.751.310 – 111.215.373). For each phenotype (deep ICH, lobar ICH, all ICH, LVD IS, CE IS, lacunar IS, all IS, WMH

volume in IS and WMH volume in population-based studies) I collected the following study-level summary genotype-phenotype association data for all directly genotyped and imputed SNPs in the region:

- SNP reference (rs) number and position;
- effect allele/non-effect allele and their frequencies;
- association effect size estimate (ß-coefficient);
- standard error of the ß-coefficient;
- p-value for the association;
- for imputed SNPs, the imputation quality measure used and its value.

7.2.3 Data analysis

7.2.3.1 Calculating the significance threshold

To allow for multiple testing, I used a modified version of the Nyholt method (Li and Ji 2005). This method takes into account the linkage disequilibrium (LD) between SNPs: based on the spectral decomposition of matrices of pairwise LD between SNPs, it calculates the 'effective number' (N) of SNPs tested. It uses this number to calculate the statistical significance threshold by applying the Bonferroni correction for N independent tests ($p<0.05/N$). I used the publicly available HapMap II CEU (release 23) dataset (Utah residents of northern and western European ancestry) genotype information to calculate N using the web interface http://gump.qimr.edu.au/general/daleN/SNPSpD/.

The modified Nyholt method controls accurately for the error rate in evaluations of real and simulated data. I decided it was more appropriate for my data than the standard Bonferroni correction (which would be overly conservative due to the assumption that all SNPs are independent), or the conventional GWAS significance p-value (which would be overly conservative because it assumes the study is hypothesis-free) (Li and Ji 2005, Nyholt 2004, Nyholt 2005, Salyakina et al. 2005).

7.2.3.2 Meta-analyses of *COL4A1/COL4A2* SNPs for each phenotype

I performed meta-analyses of the genotype summary data from each contributing study, assessing associations of *COL4A1* and *COL4A2* SNPs with each of the cerebrovascular phenotypes available – both those assumed to represent cerebral SVD specifically (deep ICH, lacunar IS, WMH volume in IS and in population-based studies), and others (lobar ICH, all ICH, CE IS, LVD IS, all IS). I hypothesised that associations would be specific to – or at least strongest with – cerebral SVD phenotypes. I used a fixed-effects inverse-variance-based model in METAL genetic meta-analysis software, which weights the ß-coefficients by their estimated standard errors (Willer et al. 2010). I used a case-control approach for ICH, ischaemic stroke and their subtypes, generating, for each SNP, odds ratios (ORs) per additional minor allele for being a case versus control. I used a quantitative trait approach for WMH analysis, assessing, for each SNP, the effect per additional minor allele on the natural log-transformed WMH volume.

Quality control measures had been applied to all individual studies prior to provision of data (Table 7.2, Table 7.3, Table 7.4, Table 7.5). Data from the included studies had been imputed to different reference datasets and their releases (HapMap II, HapMap III or 1000 Genomes) using a range of imputation software (IMPUTE, MACH and BimBam) (Marchini and Howie 2010) and were provided with reference to one of two different Human Genome reference builds (hg18 or hg19) (Abecasis et al. 2010, Altshuler et al. 2010, Frazer et al. 2007) (Table 7.2, Table 7.3, Table 7.4, Table 7.5). To permit consistent analyses and interpretation of the data across all included phenotypes and studies, I converted all the data to the same reference build (hg19) and limited the analyses to SNPs present in the HapMap II release 23 CEU population (Utah residents of northern and western European ancestry) only. This strategy provided the maximum set of 1070 SNPs represented in all studies and phenotypes.

I filtered these data based on imputation quality scores ($r^2 \geq 0.3$ [MACH, BimBam, IMPUTE]; info score ≥ 0.7 [PLINK]) and minor allele frequency (MAF) $\geq 1\%$. The imputation quality score cut-offs were chosen based on what is traditionally used and

recommended by the respective websites. The MACH and Impute r^2 scores are highly correlated with each other (Marchini and Howie 2010).

7.2.3.3 Further exploration of significant SNPs across all phenotypes

For any SNPs significantly associated with any of the nine phenotypes assessed, I examined associations with all other phenotypes. I focussed on comparisons of findings for cerebral SVD phenotypes (deep ICH, lacunar IS, WMH) versus non-SVD phenotypes (lobar ICH, all ICH, CE IS, LVD IS, all IS). To display the data across all phenotypes in a uniform way, I performed data manipulations to display ORs together with their 95% confidence intervals (CIs) and p-values using forest plots, to indicate the size and direction of associations across phenotypes.

To calculate ORs per additional minor allele for each SNP for being a case versus control for the ICH and ischaemic stroke phenotypes, I did the following: I calculated the ORs from the β-coefficients by using the EXP function in Excel (inverse of the natural logarithm). I calculated the 95% CIs of the ORs by first calculating the error factor (Excel formula: EF=EXP[1.96*SE]) and then dividing and multiplying the OR with the EF to calculate the lower and upper 95% CIs respectively (Table 7.6).

I calculated the OR per 1 unit standard deviation (SD) change in WMH volume in ischaemic stroke cohorts from the β-coefficient by using the EXP function in Excel, as the WMH values had a SD of 1. For WMH in population cohorts, I divided the ß-coefficients and SEs by their respective pooled SD across cohorts (0.684) first and then used the transformed β-coefficient and SE to calculate the respective ORs and their 95% CIs (Table 7.6).

I assessed the consistency of results for significantly associated SNPs through conducting formal I^2 and χ^2 tests for heterogeneity of results across individual studies (Higgins et al. 2003). METAL calculated the χ^2 and the p-value, from which I derived the I^2 value using the formula:

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

7.2.3.4 Functional annotation of SNPs

For SNPs with significant associations, I extracted a list of all other SNPs in moderate LD (defined as $r^2 \geq 0.3$ with a maximum allowed distance limit of 500 bp) from the web-based SNP Annotation and Proxy Search tool (SNAP) (http://www.broadinstitute.org/mpg/snap/). This was based on CEU population data from the 1000 Genomes pilot 1 (Johnson et al. 2008). I then obtained the functional annotation data for these SNPs from the Ensembl genome browser (http://www.ensembl.org/index.html), Seattle Seq (http://snp.gs.washington.edu/SeattleSeqAnnotation138/) and Haploreg v2 (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) databases (Flicek et al. 2014, Ng et al. 2009, Ward and Kellis 2012), using all three databases with partly overlapping sources of information to ensure consistency of results.

I further explored the functional significance of SNPs that were in high LD ($r^2 \geq 0.9$ with a maximum allowed distance limit of 500 bp) with the significant SNPs from my analyses using the online Genotype-Tissue Expression Portal (GTEx) eQTL (expression quantitative trait locus) browser (http://www.broadinstitute.org/gtex/) and the Regulome db database (http://regulome.stanford.edu/) (Boyle et al. 2012, GTEx Consortium 2013). An eQTL represents a locus in the genome in which variation between individuals is associated with a quantitative gene expression trait, often measured as mRNA abundance. The Regulome db database identifies whether the SNPs are located within known or predicted regulatory elements in the intergenic regions of the human genome. I focused particularly on data obtained from experiments using tissues relevant to cerebrovascular phenotypes, specifically central

nervous system (brain and spinal cord) and vasculature. The r^2 cut-offs were chosen based on discussions with statistical genetics colleagues about what would be a generally accepted reasonable cut-off to use, and based on the manageability of data.

I also explored whether SNPs with significant associations were in LD with the rare SNPs previously reported from sequencing studies to be associated with sporadic ICH (Jeanne et al. 2012, Weng et al. 2012). In addition, I looked up the LD between the significant SNPs from my analysis and the SNPs known to be associated with intracranial aneurysms, coronary artery calcification, myocardial infarction and vascular stiffness in cases of European origin (O'Donnell et al. 2011, Ruigrok et al. 2006, Tarasov et al. 2009). I used the web-based tool SNAP (Johnson et al. 2008). If the SNP was not included in SNAP, I used the Haploview programme to check whether the SNP (based on its position) was in the same haplotype block as the previously reported rare SNPs (Barrett et al. 2005).

7.3 Results

7.3.1 Included data

The entire dataset consisted of the following (Table 7.1):

- three cohorts of 1545 patients with ICH and 1485 controls, including information on the main ICH subtypes (874 deep and 671 lobar);
- 15 cohorts of 12 389 patients with ischaemic stroke, including information on the main determined TOAST subtypes (2167 LVD, 2405 CE and 1854 lacunar IS) (Adams et al. 1993), and 62 004 controls;
- nine cohorts of 2733 ischaemic stroke patients;
- seven cohorts of 9361 individuals from population-based studies with brain MRI-based measures of WMH volume.

7.3.2 Study characteristics

Table 7.2, Table 7.3, Table 7.4 and Table 7.5 include detailed descriptions of the design and characteristics of the participating studies (Adib-Samii et al. 2013, Falcone et al. 2013, Fornage et al. 2011, Rost et al. 2010, Traylor et al. 2012, Woo et al. 2002, Woo et al. 2014).

Two ICH cohorts were from prospective hospital-based case-control studies, and one was from a prospective population-based case-control study. All studies included primary spontaneous ICH cases of European descent only, excluding trauma-, tumour- and vascular malformation-related ICH cases. Lobar ICH was defined as involving predominantly the cortex and underlying white matter, and deep ICH as involving predominantly the basal ganglia, periventricular white matter, internal capsule or infratentorial area. The mean age range was 67 to 74 years for all ICH cases, 65 to 71 years for deep ICH cases and 71 to 76 years for lobar ICH cases. The mean age range for controls was 66 to 73 years (Table 7.2).

The ischaemic stroke sample consisted of 15 cohorts of patients of European ancestry. All studies used a case-control methodology. 11 studies were cross-sectional and four were in large prospective population-based cohorts. Stroke was defined as a typical clinical syndrome with radiological confirmation and the subtyping was done with the TOAST classification system (Adams et al. 1993). The mean age range for patients was 41 to 82 years, and for controls 40 to 86 years (Table 7.3).

The studies contributing data for WMH in ischaemic stroke phenotype were all hospital-based. They included cases of ischaemic stroke (European ancestry) of any subtype and excluded cases of CADASIL, vasculitis, demyelinating and mitochondrial diseases. WMH was assessed in the hemisphere contralateral to acute ischaemic stroke. All supratentorial white matter lesions were included in WMH volume measurement with the exception of WMH corresponding to lacunar infarcts. WMH volume was analysed using semi-automated software. In the majority of

studies measurements were done from 1.5 T MRI FLAIR sequences. Mean age range was 58 to 71 years (Table 7.4).

The seven studies providing data for the WMH in the population phenotype were all prospective population-based cohort studies. They excluded cases who had suffered from a stroke or TIA prior to the MRI. In five studies, WMH volume was estimated on a quantitative scale using custom-written computer programs based on automatic or semi-automatic analysis. In two studies WMH volume was estimated on a semi-quantitative 10-point scale by visual comparison with templates. Majority of the studies had used a 1.5T MRI scanner and T1, T2 and FLAIR sequences. Mean age range was 63 to 76 years (Table 7.5).

7.3.3 Significance threshold

Based on the genotype data from the HapMap II CEU release 23 dataset, the 'effective number' (N) of SNPs tested was 598, giving a final Nyholt significance threshold of p = 0.000084. This compares to the Bonferroni corrected p-value of p=0.00005 and a genome-wide significance p-value of $p=5x10^{-8}$.

7.3.4 Meta-analyses of *COL4A1/COL4A2* SNPs for each phenotype

Based on the significance threshold of p=0.000084, three common SNPs in *COL4A2* were significantly associated with the deep ICH phenotype (rs9521732: OR per additional A allele = 1.28, 95% CI 1.13 to 1.44, p=0.00007; rs9521733: OR per additional C allele = 1.29, 95% CI 1.14 to 1.46, p=0.00003; rs9515199: OR per additional C allele = 1.28, 95% CI 1.14 to 1.44, p=0.00006) (Figure 7.1, Table 7.7). These three SNPs were all intronic and in high LD with each other. One SNP (rs9521733) was directly genotyped in the majority of the cohorts, while two were imputed in all (rs951599) or in the majority (rs9521732) of the cohorts (Table 7.8). The associations across individual cohorts included in the deep ICH meta-analysis were highly consistent, with no significant heterogeneity (I^2=0%, p>0.9) (Figure

7.2). There were no statistically significant associations of common SNPs in *COL4A1/COL4A2* with any of the other phenotypes (Figure 7.3, Figure 7.4, Figure 7.5, Figure 7.6, Figure 7.7, Figure 7.8, Figure 7.9, Figure 7.10).

7.3.5 Associations with other cerebrovascular phenotypes of the three *COL4A2* SNPs associated with deep ICH

Figure 7.11 shows association results for the three SNPs significantly associated with deep ICH across all nine phenotypes assessed. Although these three SNPs were significantly associated only with deep ICH, there were suggestive associations with two other cerebral SVD phenotypes: lacunar ischaemic stroke (rs9521732: OR 1.09, 95% CI 1.02 to 1.17, p=0.01639; rs9521733: OR 1.10, 95% CI 1.03 to 1.18, p=0.00734; rs9515199: OR 1.09, 95% CI 1.02 to 1.17, p=0.0145); and WMH volume in symptomatic ischaemic stroke cases (rs9521732: OR 1.07, 95% CI 1.01 to 1.14, p=0.01442; rs9521733: OR 1.07, 95% CI 1.01 to 1.13, p=0.01642; rs9515199: OR 1.07, 95% CI 1.01 to 1.14, per 1 SD change in WMH volume; p=0.0145). There was no evidence for association of these SNPs with other non-SVD stroke subtypes (lobar ICH, CE and LVD IS), or with WMH volume in population-based studies (Figure 7.11) As expected, ORs for these SNPs with the all ICH and all ischaemic stroke phenotypes were intermediate between those for deep and lobar ICH, and for lacunar and non-SVD ischaemic stroke subtypes respectively, suggesting associations with these combined phenotypes were driven by results for deep ICH and lacunar ischaemic stroke (Figure 7.11).

There was no or minimal heterogeneity between the individual cohorts' results for lacunar ischaemic stroke (I^2=0%; p>0.8) and for WMH in ischaemic stroke (I^2=17%-22%; p>0.2) (Figure 7.12, Figure 7.13).

7.3.6 Functional annotation of SNPs

I identified 92 unique SNPs in moderate LD ($r^2 \geq 0.3$) with the three significant SNPs from my analyses. All of these SNPs (including the three significant SNPs) were intronic (Table 7.9).

From these, I then selected all five SNPs that were in high LD ($r^2 > 0.9$) with the three significant SNPs, giving a total of eight SNPs for analysis of functional annotation. The GTEx eQTL browser search revealed no significant eQTLs for any of these eight SNPs. However, the Regulome db database search revealed minimal binding evidence for four of the eight SNPs, suggesting that these SNPs are located in areas of the genome that may have some regulatory functions (Table 7.10). When I considered only results based on experiments carried out on tissues most relevant to cerebrovascular phenotypes (brain, spinal cord and blood vessels), there was still some evidence for these four SNPs being located in genomic areas of potential functional significance (Table 7.11).

The three rare exonic *COL4A2* SNPs reported in a previous sequencing study to be associated with sporadic ICH are rs117412802 (hg19 position: 111143601), rs62621875 (hg19 position: 111143681) and A1690T (hg19 position: 111164467) (Jeanne et al. 2012). Based on SNAP, rs117412802 was not in LD with the three SNPs significantly associated with deep ICH in my analyses ($r^2=0$). rs62621875 and A1690T were not represented in SNAP but, there is unlikely to be LD between the SNPs associated with deep ICH in my analysis and these as they are not in the same haplotype block and distance from the three ICH SNPs is >100000 bases (Figure 7.14). LD between the three significant SNPs and the three SNPs known to be associated with intracranial aneurysms, coronary artery calcification, myocardial infarction and vascular stiffness in cases with European origin was very low ($r^2=0$ to 0.01) (Table 7.12).

7.4 Discussion

7.4.1 Main findings

Here I present the first evidence for a significant association between an intronic locus in *COL4A2* and deep ICH, a symptomatic cerebral SVD phenotype. I also found suggestive associations in the same direction for this locus with other cerebral SVD phenotypes: lacunar ischaemic stroke and WMH in ischaemic stroke cases. These results indicate that the *COL4A2* gene is a locus conferring risk for sporadic forms of symptomatic cerebral SVD, particularly deep ICH (Rannikmäe et al. 2015)

I did not find evidence for an association between these three SNPs (or any others in the *COL4A1/COL4A2* region) and WMH assessed in largely asymptomatic individuals in population-based studies. The difference in results between WMH assessed among symptomatic people with ischaemic stroke and among those in population-based studies (both of which are often regarded as cerebral SVD phenotypes) could be due to genuine differences in the nature of the phenotype represented by WMH in these different groups and its genetic influences, differences in scanning and measurement techniques used in the contributing cohorts, or the play of chance. Interestingly, a recent GWAS study demonstrated shared genetic associations with white matter hyperintensity volumes in healthy individuals and patients with stroke, indicating common genetic susceptibility (Traylor et al. 2016).

The SNPs identified in my study, and all SNPs in moderate LD with them, are intronic, and may be located in regulatory areas of the gene. The three SNPs from my analysis are not in close LD with the variants reported from previous sequence analyses in sporadic ICH. This demonstrates that they do not represent the same signal as that observed from the previously described exonic rare variants, but instead identify a novel intronic region of *COL4A2* associated with deep ICH (and – possibly – with other cerebral SVD phenotypes). Interestingly, a recently published GWAS of WMH identified a *COL4A2* SNP (rs9515201) significantly associated with WMH. This SNP is in close LD (r^2>0.8) with the three SNPs identified in my study (Traylor et al. 2016).

7.4.2 Strengths and limitations

This study has several strengths. I investigated a specific, pre-specified hypothesis, clearly defining the clinical phenotypes and candidate genes of interest, based on pre-existing supporting data. I used a novel approach, investigating genetic associations across a range of cerebrovascular phenotypes, but focusing on a specific cerebrovascular sub-phenotype – cerebral SVD – thought to have distinct pathophysiological mechanisms and risk factors. Furthermore, through a network of collaborative groups, largely under the umbrella of the International Stroke Genetics Consortium, I was able to include in my *in silico* analyses the majority of the currently available data representing the phenotypes of interest in individuals of European ancestry.

This study has some limitations. I have shown that SNPs in *COL4A2* are associated with deep ICH by analysing data for the specific candidate *COL4A1/2* region and by correcting appropriately for multiple testing within that region using the generally accepted modified Nyholt method (Li and Ji 2005, Nyholt 2004, Nyholt 2005, Salyakina et al. 2005, Traylor et al. 2012). However this association did not reach a genome-wide level of significance (possible reasons include study size) (Woo et al. 2014), while associations with other cerebral SVD phenotypes (lacunar ischaemic stroke and WMH in ischaemic stroke cases) were suggestive, but not independently robust to multiple testing. In addition, although I corrected appropriately for the number of SNPs in the region, I did not adjust the statistical threshold for the number of phenotypes investigated. I considered that this would be overly conservative since these phenotypes are not completely independent of each other. As in any meta-analysis, this study is dependent on accurate allocation of cases and controls, and accurate phenotype measurements in the original studies.

7.4.3 Conclusions

These results indicate an association between common variation in the *COL4A2* gene and symptomatic small vessel disease, particularly deep intracerebral haemorrhage.

These findings merit replication studies, including in ethnic groups of non-European ancestry.

Tables

Table 7.1 Studies included in the meta-analyses of association between *COL4A1/COL4A2* SNPs and SVD

Phenotype	Collections	Sample size[1]	
		Cases	Controls
ICH	3 cohorts from the ISGC ICH GWAS consortium	1,545 all ICH • **874 deep** • 671 lobar	1,485
Ischaemic stroke (IS)	15 cohorts from the ISGC / Metastroke consortium of ischaemic stroke case-control studies	12,389 all IS • 2,405 CE • 2,167 LVD • **1,854 lacunar**	62,004
WMH in ischaemic stroke	9 ischaemic stroke cohorts, that included individuals with MRI data on WMH, from the WMH in Ischaemic Stroke GWAS Collaboration	**2,733 individuals**	
WMH in population	7 population-based cohorts, that included individuals with MRI data on WMH, from the CHARGE Consortium WMH group	**9,361 individuals**	

[1] Numbers in bold indicate small vessel disease phenotypes. ISGC: International Stroke Genetics Consortium; CHARGE: Cohorts for Heart and Aging Research in Genomic Epidemiology; WTCCC2: Wellcome Trust Case Control Consortium 2;

Table 7.2 Cohorts contributing data for the meta-analyses of association between *COL4A1/COL4A2* SNPs and ICH: characteristics and summary statistics

Cohort name	GOCHA	ISGC	GERFHS
	STUDY INFORMATION		
Study design	Prospective hospital-based case-control	3 separate prospective hospital-based case-control cohorts	Prospective population-based case-control
Study inclusion and exclusion criteria	Included: primary acute ICH cases presenting to participating centres, aged >55 years, confirmation of primary ICH on CT. Excluded: warfarin-, trauma-, brain tumour-, vascular malformation-related ICH, haemorrhagic transformation of ischaemic stroke, other secondary causes.		Included: spontaneous ICH in the Greater Cincinnati/ Northern Kentucky region; age ≥18 years. Excluded: trauma-, brain tumour-, vascular malformation-related ICH.
ICH diagnosis criteria	Non-traumatic abrupt onset of severe headache, altered level of consciousness, and/or focal neurological deficit that is associated with a focal collection of blood within the brain parenchyma as observed on CT or at autopsy and is not due to haemorrhagic transformation of an infarction.		
Deep ICH diagnosis criteria	Involving predominantly the basal ganglia, periventricular white matter, or internal capsule, and infratentorial ICH		
Lobar ICH diagnosis criteria	Involving predominantly the cortex and underlying white matter		
Key study references	Falcone et al. Stroke 2013	Woo et al. Am J Hum Genet 2014	Woo et al. Stroke 2002
	SUBJECTS' CHARACTERISTICS		
Ancestry of study sample	European descent		
Total number of ICH cases	298	450	797
Age of all cases with ICH (mean±SD)	74±10	72±12	67±15
Number of deep ICH cases	125	279	470
Age of cases with deep ICH (mean ± SD)	71±13	69±14	65±16
Number of lobar ICH cases	173	171	327
Age of cases with lobar ICH (mean±SD)	76±11	74±13	71±12

Cohort name	**GOCHA**	**ISGC**	**GERFHS**
Number of controls	457	489	539
Age of controls (mean±SD)	72±8	73±12	66±15
	GENOTYPING INFORMATION		
GWAS panel used	Illumina HumanHap 610 Quad		Affymetrix 6.0
QC steps done before imputation	QC steps before imputation done in every centre, including removal of population outliers, missing data and HWE departures		
Imputation software	IMPUTE2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html)		
Which reference imputation done to	1000-genomes (June 2011)		
Imputation quality metric reported	IMPUTE info score & PLINK info score (http://pngu.mgh.harvard.edu/~purcell/plink/)		
QC filtering applied to imputed variants	Removed SNPs with PLINK info score of <0.7 and minimum allele frequency (MAF) <1%		
Software and statistical model used for association analysed	PLINK ICH = bo + b1*SNP + b2*age + b3*sex + b4*PC1 ... PC4		
Handling of population stratification	Principal component (PC) analysis. Remotion of population outliers by visual inspection of PC1*PC2 plot, followed by inclusion of PC1-PC4 in regression models		
Hardy-Weinberg Equilibrium (HWE)	Excluded if HWE $p<1x10^{-6}$		
Adjusting for cryptic or overt relatedness	Excluded individuals with an inferred first- or second-degree relative in the sample identified on the basis of pairwise allele sharing estimates (estimated genome proportion shared identical by descent) $\pi>0.1875$		
Strand and build of the human genome on which results are provided	+ strand, build 37/hg19		

GOCHA: Genetics of Cerebral Hemorrhage with Anticoagulation; ISGC: International Stroke Genetics Consortium; GERFHS: Genetic and Environmental Risk Factors for Hemorrhagic Stroke

Table 7.3 Cohorts contributing data for the meta-analyses of association between *COL4A1/COL4A2* SNPs and ischaemic stroke: characteristics and summary statistics

Cohort name	**WTCCC2_UK**	**WTCCC2_D**	**ISGS & SWISS**	**Rotterdam**	**Milano**	**ARIC**	**ASGC**	**Brains**
	STUDY INFORMATION							
Study design[1]	CS	CS	CS	PB	CS	PB	CS	CS
Inclusion and exclusion criteria	Included: Patients with ischaemic stroke who were of European ancestry from Europe, North America, and Australia, together with controls of matched ancestry. All studies used a case-control methodology.							
IS definition	Stroke was defined as a typical clinical syndrome with radiological confirmation. Stroke subtyping was done with the TOAST classification system.							
Reference	Traylor et al, 2012							
	SUBJECTS' CHARACTERISTICS							
Ancestry of study sample	European ancestry (from Europe, North-America, Australia)							
Total no. cases/ controls	2374/ 5175	1174/ 797	1070/ 2329	367/ 5396	372/ 407	385/ 8803	1162/ 1244	361/ 444
No. CE cases	460	330	247	-	65	93	240	29
No. LVD cases	498	346	229	-	74	31	421	120
No. SVD cases	474	106	201	-	25	63	310	97
Age of cases/ controls (mean±SD)	72±13/52[1]	67±13/ 63±11	67±14/ 65±13	71±8/ 69±9	57±16/ 51±8	57±5/ 54±6	73±13/ 70±12	74±14/ ≥65
	GENOTYPING INFORMATION							
GWAS panel used	Illumina 660		Illumina 550/610/ 660	Illumina 550	Illumina 610/660	Affymetrix 6.0	Illumina 610	
QC filters applied before imputation	QC steps before imputation done in every centre, including removal of ancestry outliers defined by principal component analysis and poorly typed individuals.							
Imputation	MaCH		MaCH v1.0.16	MaCH		MaCH v1.0.16		

Cohort name	**WTCCC2_UK**	**WTCCC2_D**	**ISGS & SWISS**	**Rotterdam**	**Milano**	**ARIC**	**ASGC**	**Brains**
software								
Reference imputed to	HapMap2 CEU		1000G (Aug 2010)	HapMap#22	HapMap2 CEU	HapMap2 CEU	HapMapII #24 CEU	
Imputation quality metric reported	MaCH oevar*							
QC filtering applied to imputed variants	Included SNPs with imputation quality ≥0.3 (O/E variance) and MAF ≥1%							
Software and statistical model used for association analysed	Logistic regression for all cohorts with a cross-sectional study design to model the multiplicative SNP effects on risk of the dichotomous outcome of stroke against ancestry-matched controls. Cox proportional-hazards models for prospective studies to assess time to first stroke, fitting an additive model relating genotype dose to the stroke outcome.							
Handling of population stratification	Three cohorts used ancestry-informative principal components as covariates to correct for population stratification (ISGS/SWISS, ASGC, Brains). Age and sex were included as covariates in two centres (ISGS/SWISS and Brains). In all other centres no covariates were included.							
HWE	All centres screened SNPs for HWE errors prior to imputation or analysis.							
Adjusting for cryptic or overt relatedness	All centres screened individuals for relatedness prior to imputation or analysis.							

Continued

Cohort name	CHS	deCODE	FHS	GASROS	GEOS	HPS	HVH
	STUDY INFORMATION						
Study design[1]	PB	CS	PB	CS	CS	CS	CS
Inclusion and exclusion criteria	Included: Patients with ischaemic stroke who were of European ancestry from Europe, North America, and Australia, together with controls of matched ancestry. All studies used a case-control methodology.						
IS definition	Stroke was defined as a typical clinical syndrome with radiological confirmation. Stroke subtyping was done with the TOAST classification system.						
Reference	Traylor et al, 2012						
	SUBJECTS' CHARACTERISTICS						
Ancestry of study sample	European ancestry (from Europe, North-America, Australia)						
Total no. cases/controls	454/ 2817	2391/ 26970	171/ 4164	516/ 1202	448/ 498	578/ 468	566/ 1290
No. CE cases	147	399	48	169	90	-	88
No. LVD cases	-	255	-	95	37	-	61
No. SVD cases	73	240	-	38	54	-	173
Age of cases/ controls (mean±SD)	82±6/ 86±6	73±12/ 57±21	80±11/ 72±12	67±15/ 48±9	41±7/ 40±7	65±8/ 59±9	69±9/ 67±9
	GENOTYPING INFORMATION						
GWAS panel used	Illumina 370	Illumina 317/370	Affymetrix 550	Affymetrix 6.0	Illumina H Omni	Illumina 610	Illumina 370
QC filters applied before imputation	QC steps before imputation done in every centre, including removal of ancestry outliers defined by principal component analysis and poorly typed individuals.						
Imputation software	BIM-BAM	IMPUTE	MaCH v1.0.15	MaCH v1.0.16	Not imputed	MaCH	BIM-BAM
Reference imputed to	NCBI b35	HapMap2 CEU	HapMap#22	HapMap3 CEU+TSI training set	Not imputed	HapMap2 #22	

Cohort name	**CHS**	**deCODE**	**FHS**	**GASROS**	**GEOS**	**HPS**	**HVH**
Imputation quality metric reported	O/E ratio	IMPUTE info	MaCH oevar		n/a	MaCH oevar	O/E ratio
QC filtering applied to imputed variants	Included SNPs with imputation quality ≥0.3 (O/E variance or IMPUTE info score) and MAF ≥1%.						
Software and statistical model used for association analysed	Logistic regression for all cohorts with a cross-sectional study design to model the multiplicative SNP effects on risk of the dichotomous outcome of stroke against ancestry-matched controls. Cox proportional-hazards models for prospective studies to assess time to first stroke, fitting an additive model relating genotype dose to the stroke outcome.						
Handling of population stratification	One cohort used ancestry-informative principal components as covariates to correct for population stratification (GEOS). Sex was used as a covariate in one centre (GASROS) and one centre used recruitment phase (1 or 2) as a covariate (GEOS). In all other centres no covariates were included.						
HWE	All centres screened SNPs for HWE errors prior to imputation or analysis.						
Adjusting for cryptic or overt relatedness	All centres screened individuals for relatedness prior to imputation or analysis.						

CS: cross-sectional study; PB: prospective population-based study. [1] The approximate age of genotyping of the 2738 controls from the 1958 Birth Cohort. Age was not available for the remaining controls. MaCH oevar: the measure is the ratio of the empirically observed variance of the allele dosage to the expected binomial variance at Hardy-Weinberg equilibrium. WTCCC2: Wellcome Trust Case Control Consortium 2; UK: The United Kingdom of Great Britain and Northern Ireland; D: Germany; ISGS: Ischaemic stroke Genetics Study; SWISS: Sibling with Ischaemic Stroke Study; ARIC: Atherosclerosis Risk in Communities study; ASGC: Australian Stroke Genetics Collaborative; BRAINS: Bio-Repository of DNA in stroke; CHS: Cardiovascular Health Study; FHS: Framingham Heart Study; GASROS: Genes Affecting Stroke Risk and Outcome Study; GEOS: Genetics of Early-Onset Stroke: HPS: Heart Protection Study; HVH: Heart and Vascular Health Study

Table 7.4 Cohorts contributing data for the meta-analyses of association between *COL4A1/COL4A2* SNPs and WMH in ischaemic stroke: characteristics and summary statistics

Cohort name	**ASGC**	**GASROS**	**SWISS**	**ISGS**	**Edinburgh**	**Milano**	**SGUL**	**Oxford FLAIR**	**Oxford T2**	**Munich FLAIR**	**Munich T2**
	STUDY INFORMATION										
Study design	HB	HB	HB	HB	HB	HB	HB	PB	PB	HB	HB
Study inclusion and exclusion criteria	Included: age >18 years; diagnosis of ischaemic stroke of any subtype Excluded: cases of CADASIL, vasculitis, demyelinating and mitochondrial disorders										
WMH quantification method used	WMH assessed in hemisphere contralateral to acute ischaemic stroke. All supratentorial WM lesions were included in WMH volume (WMHV) with the exception of WMH corresponding to lacunar infarcts; sequential axial FLAIR images analysed using a validated, semi-automated protocol (MRIcro, Un. of Nottingham).				WMH assessed in the hemisphere contralateral to acute ischaemic stroke. All supratentorial white matter lesions were included in WMH volume (WMHV) with the exception of WMH corresponding to lacunar infarcts; Analysed using DISPunc semi-automated lesion drawing software.						
MRI scanner used	1.5T Siemens Magnetom Avanto	1.5T GE Medical Signa			1.5T GE Medical Signa, 1.5 T Siemens	1.5T Siemens, 0.5T Philips	1.5 T Philips, 1.5T GE Signa LX	1.5T GE Medical Signa, 1.5T Philips		1T or 1.5T Siemens Magnetom, 1.5T GE Medical Signa	1T or 1.5T Siemens Magnetom 3T and 1.5T GE Medical Signa
MRI sequence used for assessing WMH	Axial FLAIR				Axial FLAIR	Coronal FLAIR or Axial FLAIR	Axial FLAIR	Coronal FLAIR	Axial T2	Axial FLAIR	Axial T2
Key study references	Adib-Samii et al, 2013; Rost et al, 2010										

Cohort name	ASGC	GASROS	SWISS	ISGS	Edinburgh	Milano	SGUL	Oxford FLAIR	Oxford T2	Munich FLAIR	Munich T2
	SUBJECTS' CHARACTERISTICS										
Ancestry of study sample	European ancestry										
No. assessed	104	975	115	209	65	152	323	65	75	447	203
Age (mean±SD)	65±13	66±14	66±11	68±14	69±13	58±14	71±13	65±15	68±13	66±12	67±12
GENOTYPING INFORMATION											
GWAS panel used	Human610-Quad	Affymetrix 6.0, Illumina Human 610-Quad or Illumina OmniExp. beadchips	Illumina650K-Quad		Illumina Human660W-Quad	Illumina Human610-Quad or Human660W-Quad	Illumina Human660W-Quad				
QC filters applied to genotype data before imputation	Individuals removed if inferred sex discordant with recorded sex; if >5% missing genotype data.				Individuals removed if inferred sex discordant with recorded sex; if >5% missing genotype data; SNPs were excluded for minor allele frequency <1% or >5% missing data.						
Imputation software	IMPUTE2										
Which reference imputation done to	HapMap3 and 1000G (June	HapMap3 and 1000G (June 2010 or 2011)	HapMap3 and 1000 G (June 2010)		1000G Phase 1 integrated variant set (March 2012)						

Cohort name	ASGC	GASROS	SWISS	ISGS	Edinburgh	Milano	SGUL	Oxford FLAIR	Oxford T2	Munich FLAIR	Munich T2
	2010)										
Imputation quality metric reported	IMPUTE info score & PLINK info score										
QC filtering applied to imputed variants	Removed SNPs with IMPUTE info score <0.3 and SNPs with MAF <1%										
Software and statistical model used for association analysed	WMHV from the hemisphere contralateral to acute ischaemic stroke was doubled to obtain whole brain values and adjusted for normal inter-individual variation in head size. Values were natural log transformed to a normal distribution. Within each group, rank-transformed residuals were derived from a linear regression model predicting WMHV with age, sex, and the first 2 ancestry principle components as covariates in GenABEL. Thus, the phenotype was adjusted for age because WMHV is highly age dependent. Principal components, derived using EIGENSTRAT, were included to correct for potential population stratification. Association analysis was undertaken in PLINK using pseudodosages, a fractional count of 0 to 1 alleles for each genotype weighted by imputation probability, within a linear regression (additive) model.										
Handling of population stratification	Principal component analysis. Remotion of population outliers by visual inspection of PC1*PC2 plot, followed by inclusion of PC1-PC4 into regression models.				2 ancestry informative principal components covariates were included in the model used to derive the WMH phenotype.						
HWE	Excluded if HWE $p<1x10^{-6}$										
Adjusting for cryptic or overt relatedness	Excluded individuals with an inferred first- or second-degree relative in the sample identified on the basis of pairwise allele sharing estimates (genome proportion shared identical by descent) $\pi>0.2$.				To obtain a set of putatively unrelated individuals, a hidden Markov model (HMM) was used to infer identity by descent and then individuals were removed iteratively to obtain a set with pairwise identity by descent of <5%.						
Strand and build of the human genome	+ strand, build 37/hg19										

HB: hospital-based stroke study; PB: population-based stroke study; SGUL: St. George's University, London

Table 7.5 Cohorts contributing data for the meta-analyses of association between *COL4A1/COL4A2* SNPs and WMH in the population: characteristics and summary statistics

Cohort name	**ARIC**	**CHS**	**FHS**	**Rotterdam Study I**	**Rotterdam Study II**	**AGES**	**ASPS**
	STUDY INFORMATION						
Study design	Prospective population-based cohort studies						
Study inclusion and exclusion criteria	In all cohorts, participants were excluded if they lacked information on MRI, genotypes, or both, or if they suffered a stroke or transient ischaemic attack prior to MRI. In addition, CHS did not genotype participants with clinical cardiovascular disease at baseline. ASPS and RS did not perform MRI scans in participants with dementia, and FHS analyses excluded participants who had dementia at the time of MRI.						
WMH quantification method used	In AGES-Reykjavik, ASPS, FHS, and Rotterdam, WMH volume was estimated on a quantitative scale using custom-written computer programs based on an automatic segmentation algorithm or a semiautomatic segmentation analysis involving operator-guided removal of non-brain elements. In ARIC and CHS, WMH volume was estimated on a semi-quantitative 10-point scale by visual comparison with 8 templates that successively increased from barely detectable white matter lesions to extensive, confluent abnormalities. Study participants' brain images were compared with the reference standards after aligning them to approximately the same apparent size. Hence, visual grades are inherently normalized for brain size.						
MRI scanner used	General Electric or Picker 1.5 Tesla scanners	1.5T General Electric or Picker or 0.35 T Toshiba	1 or 1.5 T Siemens Magnetom scanner	1.5 T Siemens Vision scanner	1.5 T GE Healthcare scanner	1.5 T Signa TwinSpeed system	1.5 T Gyroscan S15 and ACS
MRI sequence used for assessing WMH	T1- and T2-weighted scans in the axial plane were obtained for all participants. These were complemented by either scans obtained with fluid attenuation inversion recovery or proton density sequences to allow better separation of white matter hyperintensities and cerebrospinal fluid.						
Key study references	Fornage M et al, 2011						
	SUBJECTS' CHARACTERISTICS						
Ancestry of study sample	European ancestry						
Number of cases	798	2184	2200	380	567	2467	765

Cohort name	**ARIC**	**CHS**	**FHS**	**Rotterdam Study I**	**Rotterdam Study II**	**AGES**	**ASPS**
Age of cases (mean±SD)	63±4	72±5	64± 11	73±8	67±5	76±5	65±8
	GENOTYPING INFORMATION						
GWAS panel used	Affymetrix GeneChip SNP Array 6.0	Illumina Human 370-Duo BeadChip	Affymetrix GeneChip Human mapping 500K Array Set and 50K Human gene Focused panel	Illumina HumanHap550 Duo BeadChip	Illumina HumanHap550 Duo BeadChip	Illumina Human 370-Duo BeadChip	Illumina 610-Quad BeadChip
QC filters applied to genotype data before imputation	Participant-specific quality control filters were applied based on missing call rate, cryptic relatedness, and number of Mendelian errors per individual. SNP-specific quality controls included filters for call rate, minor allele frequency, Hardy-Weinberg equilibrium, differential missingness by outcome or genotype.						
Imputation software	MaCH (v1.0.15 or 1.0.16)	BIM-BAM 15	MaCH (v1.0.15 or 1.0.16)				
Which reference imputation done to	HapMap2 CEU #22; +strand of NCBI build 36	+strand of NCBI build 35	HapMap2 CEU #22; +strand of NCBI build 36				
Imputation quality metric reported	O/E ratio and oevar						
QC filtering applied to imputed variants	Excluded SNPs with O/E ratio & oevar <0.3 and MAF<1%						
Software and statistical model used for association	Within each study, a linear regression model was used to evaluate the association of the natural log-transformed volume of WMH (log[WMH + 1]) with the number of minor alleles (0 to 2) at each SNP. Analyses were adjusted only for age, sex, and total intracranial volume (except in ASPS, ARIC, and CHS). ARIC and CHS also adjusted for study site, and FHS adjusted for familial						

Cohort name	ARIC	CHS	FHS	Rotterdam Study I	Rotterdam Study II	AGES	ASPS
analysed	structure.						
HWE	Excluded if Hardy-Weinberg $p<1x10^{-5}$ in CHS and $p<1x10^{-6}$ in AGES-Reykjavik, ARIC, ASPS, FHS, and Rotterdam						
Adjusting for cryptic or overt relatedness and population stratification	Studies were screened for latent population substructure, including cryptic relatedness, using suitable programs: EIGENSTRAT in ARIC, FHS and AGES-Reykjavik, an IBD matrix in ASPS and Rotterdam, and using principal component analysis in CHS.						
Strand and build of the human genome	+ strand; hg18						

AGES: Aging Gene-Environment Susceptibility-Reykjavik Study; ASPS: Austrian Stroke Prevention Study

Table 7.6 Original measures of association (β) from the meta-analyses of association between *COL4A1/COL4A2* SNPs and SVD, and the derived odds ratios (ORs)

	Phenotype	β (minor allele)	SE of β	OR (minor allele)	95% CI of OR
rs 9521732 (minor allele A)	**All ICH**	0.1548	0.0519	1.17	1.05 to 1.29
	Deep ICH	0.2451	0.0614	1.28	1.13 to 1.44
	Lobar ICH	0.0489	0.068	1.05	0.92 to 1.20
	All IS	0.0261	0.0161	1.03	0.99 to 1.06
	CE IS	0.022	0.0321	1.02	0.96 to 1.09
	LVD IS	-0.0465	0.0354	0.95	0.89 to 1.02
	Lacunar IS	0.0858	0.0357	1.09	1.02 to 1.17
	WMH in IS	0.0707	0.0289	1.07	1.01 to 1.14
	WMH in population	0.0088	0.008	1.01	0.99 to 1.04
rs 9521733 (minor allele C)	**All ICH**	0.1706	0.0521	1.19	1.07 to 1.31
	Deep ICH	0.2547	0.0614	1.29	1.14 to 1.46
	Lobar ICH	0.0674	0.068	1.07	0.94 to 1.22
	All IS	0.0294	0.0164	1.03	1.00 to 1.06
	CE IS	0.0151	0.0326	1.02	0.95 to 1.08
	LVD IS	-0.0466	0.0353	0.95	0.89 to 1.02
	Lacunar IS	0.0979	0.0365	1.10	1.03 to 1.18
	WMH in IS	0.0691	0.0288	1.07	1.01 to 1.13
rs 9515199 (minor allele C)	**All ICH**	0.1554	0.0519	1.17	1.06 to 1.29
	Deep ICH	0.2465	0.0614	1.28	1.14 to 1.44
	Lobar ICH	0.0487	0.068	1.05	0.92 to 1.20
	All IS	0.0308	0.0165	1.03	1.00 to 1.07
	CE IS	0.02	0.0332	1.02	0.96 to 1.09
	LVD IS	-0.0406	0.036	0.96	0.89 to 1.03
	Lacunar IS	0.0891	0.0365	1.09	1.02 to 1.17
	WMH in IS	0.0707	0.0289	1.07	1.01 to 1.14
	WMH in population	0.0088	0.008	1.01	0.99 to 1.04

Table 7.7 *COL4A1/COL4A2* SNPs significantly associated with the deep ICH phenotype

SNP	Gene	Alleles[1]	Functional annotation	P-value	MAF	r^2
rs9521732	*COL4A2*	**A**/C	Intronic	0.0000652	A=0.413	rs9521733 = 0.896 rs9515199 = 1
rs9521733	*COL4A2*	**C**/T	Intronic	0.0000339	C=0.404	rs9515199 = 0.896 rs9521732 = 0.896
rs9515199	*COL4A2*	**C**/T	Intronic	0.0000599	C=0.414	rs9521733 = 0.896 rs9521732 = 1

[1] Risk allele is marked in bold; MAF: minor allele frequency; r^2: measure of linkage disequilibrium; Intronic: located in a non-coding region of the gene

Table 7.8 Imputed and directly genotyped *COL4A2* SNPs

Phenotype	Study	Imputation quality measures		
		rs9521732	**rs9521733**	**rs9515199**
ICH	ISGC	1[a]	G	1[a]
	GOCHA	1[a]	G	0.99[a]
	GERFHS	G	G	1[a]
ISCHAEMIC STROKE	ARIC	G	n/a	1[b]
	ASGC	0.96[b]	G	0.97[b]
	Brains	0.99[b]	G	0.98[b]
	CHS	0.92[c]	G	0.92[c]
	deCODE	0.99[d]	G	0.99[d]
	GASROS	1[b]	0.99[b]	n/a
	GEOS	G	n/a	n/a
	HVH	0.92[c]	G	0.91[c]
	ISGS & SWISS	0.98[b]	G	0.98[b]
	Milano	0.97[b]	1[b]	0.97[b]
	WTCCC2_D	0.98[b]	G	0.97[b]
	WTCCC2_UK	0.99[b]	G	0.99[b]
	FHS	0.94[b]	0.94[b]	0.94[b]
	HPS	0.97[b]	G	0.97[b]
	Rotterdam	0.99[b]	1[b]	0.99[b]
WMH in symptomatic stroke	Edinburgh	0.97[d]	G	0.97[d]
	Munich_T2	0.90[d]	G	0.90[d]
	Munich_FLAIR	1[d]	G	1[d]
	Oxford_T2	0.94[d]	G	0.94[d]
	Oxford_FLAIR	0.77[d]	G	0.77[d]
	SGUL	0.98[d]	G	0.98[d]
	Milano	0.80[d]	0.84[d]	0.80[d]
	ASGC	1[d]	1[d]	1[d]
	GASROS_affy	G	G	1[d]
	GASROS_illumina	1[d]	G	1[d]

	GASROS_omni	G	1[d]	1[d]
	ISGC-SWISS	0.93[d]	G	0.93[d]
WMH in population	AGES	0.99[b]	n/a	0.98[b]
	ARIC	G	n/a	1[b]
	CHS	0.92[c]	n/a	0.92[c]
	FHS	0.94[b]	n/a	0.94[b]
	Rotterdam Study I	0.99[b]	n/a	0.99[b]
	Rotterdam Study II	0.99[b]	n/a	0.99[b]
	ASPS	0.98[b]	n/a	0.98[b]

G: directly genotyped; n/a: SNP not available for the study.
Imputation quality numbers provided reflect the minimum quality measure across sub-phenotypes, rounded to two decimal places.
[a] PLINK info [b] MaCH oevar [c] BIMBAM O/E ratio [d] IMPUTE info

Table 7.9 Functional annotation of SNPs in moderate LD (r2>0.3) with rs9521732, rs9521733 and rs9515199*

Gene	hg 19	Consequence Ensembl	Consequence SeattleSeq	Haploreg: dbSNP functional annotation
COL4A2	rs11126924	-	-	intronic
COL4A2	rs11594556	-	-	intronic
COL4A2	rs11619425	intron_variant	intron-variant	intronic
COL4A2	rs11619427	intron_variant	intron-variant	intronic
COL4A2	rs11619430	intron_variant	intron-variant	intronic
COL4A2	rs11838637	intron_variant	intron-variant	intronic
COL4A2	rs11838776	intron_variant	intron-variant	intronic
COL4A2	rs12853693	intron_variant	intron-variant	intronic
COL4A2	rs1888004	intron_variant	intron-variant	intronic
COL4A2	rs1888005	intron_variant	intron-variant	intronic
COL4A2	rs1927342	intron_variant	intron-variant	intronic
COL4A2	rs1927343	-	intron-variant	intronic
COL4A2	rs1927344	intron_variant	intron-variant	intronic
COL4A2	rs1927345	intron_variant	intron-variant	intronic
COL4A2	rs1927346	intron_variant	intron-variant	intronic
COL4A2	rs1927347	intron_variant	intron-variant	intronic
COL4A2	rs1927349	intron_variant	intron-variant	intronic
COL4A2	rs1927355	intronic downstream_ gene_variant[1]	-	intronic
COL4A2	rs1999013	intron_variant	intron-variant	intronic
COL4A2	rs2149067	intron_variant	intron-variant	intronic
COL4A2	rs2391825	intron_variant	intron-variant	intronic
COL4A2	rs34402154	intron_variant	intron-variant	intronic
COL4A2	rs34992019	intron_variant	intron-variant	intronic
COL4A2	rs3899318		intron-variant	intronic
COL4A2	rs4283091	intron_variant	intron-variant	intronic
COL4A2	rs4492912	intron_variant	intron-variant	intronic
COL4A2	rs4547215	intron_variant	intron-variant	intronic
COL4A2	rs4586292	intron_variant	intron-variant	intronic
COL4A2	rs4771674	intron_variant	intron-variant	intronic
COL4A2	rs4771675	intron_variant	intron-variant	intronic
COL4A2	rs4771676	intron_variant	intron-variant	intronic
COL4A2	rs4773157	intron_variant	intron-variant	intronic
COL4A2	rs4773169	intron_variant	intron-variant	intronic
COL4A2	rs4773170	intron_variant	intron-variant	intronic
COL4A2	rs4773171	-	intron-variant	intronic
COL4A2	rs4773173	intron_variant	intron-variant	intronic
COL4A2	rs4773174	intron_variant	intron-variant	intronic
COL4A2	rs4773177	intron_variant	intron-variant	intronic
COL4A2	rs55940034	intron_variant	intron-variant	intronic
COL4A2	rs61963197	intron_variant	intron-variant	intronic

Gene	hg 19	Consequence Ensembl	Consequence SeattleSeq	Haploreg: dbSNP functional annotation
COL4A2	rs7318424	-	intron-variant	intronic
COL4A2	rs7318742	intron_variant	intron-variant	intronic
COL4A2	rs7321362	intron_variant	intron-variant	intronic
COL4A2	rs7323228	intron_variant	intron-variant	intronic
COL4A2	rs7326145	-	intron-variant	intronic
COL4A2	rs7328731	intron_variant	intron-variant	intronic
COL4A2	rs7333596	intron_variant	intron-variant	intronic
COL4A2	rs7333748	-	intron-variant	intronic
COL4A2	rs7334022	intron_variant	intron-variant	intronic
COL4A2	rs750598	intron_variant	intron-variant	intronic
COL4A2	rs7982993	-	intron-variant	intronic
COL4A2	rs7983374	intron_variant	intron-variant	intronic
COL4A2	rs7990844	intron_variant	intron-variant	intronic
COL4A2	rs7991842	intron_variant	intron-variant	intronic
COL4A2	rs7999034	intron_variant	intron-variant	intronic
COL4A2	rs872587	-	intron-variant	intronic
COL4A2	rs872588	intron_variant	intron-variant	intronic
COL4A2	rs872589	intron_variant	intron-variant	intronic
COL4A2	rs913746	-	intron-variant	intronic
COL4A2	rs9284253	intron_variant	intron-variant	intronic
COL4A2	rs9301454	intron_variant	intron-variant	intronic
COL4A2	rs9515195	-	intron-variant	intronic
COL4A2	rs9515196	-	intron-variant	intronic
COL4A2	rs9515197	intron_variant	intron-variant	intronic
COL4A2	**rs9515199**	**intron_variant**	**intron-variant**	**intronic**
COL4A2	rs9515201	-	intron-variant	intronic
COL4A2	rs9515204	intron_variant	intron-variant	intronic
COL4A2	rs9521717	-	intron-variant	intronic
COL4A2	rs9521718	-	intron-variant	intronic
COL4A2	rs9521719	intron_variant	intron-variant	intronic
COL4A2	rs9521720	intron_variant	intron-variant	intronic
COL4A2	rs9521721	-	intron-variant	intronic
COL4A2	rs9521729	intron_variant	intron-variant	intronic
COL4A2	rs9521730	-	intron-variant	intronic
COL4A2	**rs9521732**	**intron_variant**	**intron-variant**	**intronic**
COL4A2	**rs9521733**	**intron_variant**	**intron-variant**	**intronic**
COL4A2	rs9521734	intron_variant	intron-variant	intronic
COL4A2	rs9521735	intron_variant	intron-variant	intronic
COL4A2	rs9521739	intron_variant	intron-variant	intronic
COL4A2	rs9521740	intron_variant	intron-variant	intronic
COL4A2	rs9521742	-	intron-variant	intronic
COL4A2	rs9521743	intron_variant	intron-variant	intronic
COL4A2	rs9521744	intron_variant	intron-variant	intronic

Gene	hg 19	Consequence Ensembl	Consequence SeattleSeq	Haploreg: dbSNP functional annotation
COL4A2	rs9521746	intron_variant	intron-variant	intronic
COL4A2	rs9521747	intron_variant	intron-variant	intronic
COL4A2	rs9521748	intron_variant	intron-variant	intronic
COL4A2	rs952359	intron_variant	intron-variant	intronic
COL4A2	rs9555692	-	intron-variant	intronic
COL4A2	rs9555694	-	intron-variant	intronic
COL4A2	rs9555695	intron_variant	intron-variant	intronic
COL4A2	rs9559780	-	intron-variant	intronic
COL4A2	rs9559781	-	intron-variant	intronic
COL4A2	rs9559788	-	intron-variant	intronic
COL4A2	rs9583489	intron_variant	intron-variant	intronic
COL4A2	rs9588151	intron_variant	intron-variant	Intronic

-:SNP not included in database
* the three SNPs significantly associated with deep ICH are shown in bold.
[1] A sequence variant located 3' of a gene

Table 7.10 Functional annotation of SNPs in high LD (r2>0.9) with rs9521732, rs9521733 and rs9515199: Regulome db database results

hg19	Regulome db score	Motifs[1]	Histone modification[2]	Protein binding[3]	Chromatin structure[4]
rs9521732	no data	n/a	n/a	n/a	n/a
rs9515199	no data	n/a	n/a	n/a	n/a
rs9521735	no data	n/a	n/a	n/a	n/a
rs4771674	no data	n/a	n/a	n/a	n/a
rs9521733	5	-	✓	-	✓
rs9521734	6	✓	✓	-	-
rs1999013	5	✓	✓	-	✓
rs9555695	4	-	✓	✓	✓

n/a: not applicable; - : no evidence; ✓: evidence
Grey boxes indicate that there is some evidence for this functional annotation based on experiments done on more relevant tissues: brain, spinal cord or blood vessel tissue.
Regulome db score: The scoring system represents with increasing confidence that a variant lies in a functional location and is likely to result in a functional consequence (lower scores indicate increasing evidence for a variant to be located in a functional region).
1: Likely to affect binding and linked to expression of a gene target
2: Likely to affect binding
3: Less likely to affect binding
4-6: Minimal binding evidence (lack evidence of the variant actually disrupting the site of binding)
[1]Motifs: the SNP is located in an area of short recurring patterns in DNA (motif) thought to have a regulatory function and hence may predict transcription factor binding sites.
[2]Histone modification: the SNP is located in an area of histone modification. Histones are proteins that associate with DNA in the nucleus and help condense it into chromatin. Histone modifications are a range of post-translational modifications to the N-terminal tails of the histones in chromatin, which include a series of methylations and acetylations at defined lysine and arginine residues. Histone modification profiles are associated with differences in gene transcription and hence can be used as a generic tool to identify functional elements in the genome.
[3]Protein binding: the SNP is located in an area binding a transcription factor.
[4]Chromatin structure: the SNP is located in an area of possible chromatin accessibility, suggesting the area has a regulatory function. Chromatin is a complex of DNA and proteins that forms chromosomes within the nucleus of eukaryotic cells.

Table 7.11 Functional annotation of SNPs: Regulome db database results sorted by tissue

Cell line Symbol[1]	Tissue[2]	rs9555695			rs9521733		rs9521734	rs1999013		
		histone modification	protein binding	chromatin structure	histone modification	chromatin structure	histone modification	histone modification	chromatin structure	motifs
Hmvecdblad	blood vessel: adult blood microvascular endothelial cells, dermal-derived T3	-	-	✓	-	-	-	-	-	-
Huvec	blood vessel: umbilical vein endothelial cells T2	✓	-	-	✓	✓	✓	✓	-	-
Nha	astrocytes from normal brain T3	✓	-	-	✓	-	✓	✓	-	-
Astrocytes	Astrocytes	-	-	-	-	✓	-	-	-	-
HASP	astrocytes normal spinal cord T3	-	-	-	-	✓	-	-	-	-
Sknshra	neuroblastoma cell line - cancer brain T3	✓	-	-	-	-	-	-	-	-
A549	epithelial cell line derived from a lung carcinoma tissue T2	✓	-	-	✓	-	✓	✓	-	-
Cd20ro01778	blood B cells T2	-	-	-	-	-	-	✓	-	-
Chorion	chorion cells fetal membrane T3	-	-	✓	-	-	-	-	✓	✓
Cll	chronic lymphocytic leukemia cell, T-cell lymphocyte T3	-	-	-	-	-	-	-	✓	-
Dnd41	leukemia T cell T3	✓	-	-	✓	-	✓	✓	-	-
Gm12878	B lymphocyte T1	✓	-	-	✓	-	✓	✓	✓	-
Gm12891	B-lymphocyte T3	-	-	-	-	-	-	-	✓	✓
Gm12892	B-lymphocyte T3	-	-	-	-	-	-	-	✓	✓
Gm18507	lymphoblastoid T3	-	-	-	-	-	-	-	✓	-

Cell line Symbol[1]	Tissue[2]	rs9555695			rs9521733		rs9521734	rs1999013		
		histone modification	protein binding	chromatin structure	histone modification	chromatin structure	histone modification	histone modification	chromatin structure	motifs
Gm19238	B lymphocyte T3	-	-	-	-	-	-	-	✓	-
Gm19239	B-lymphocyte T3	-	-	-	-	-	-	-	✓	-
H1hesc	embryonic stem cells T1	✓	-	-	✓	-	✓	✓	✓	-
H7es	undifferentiated embryonic stem cells	✓	-	-	-	-	-	-	✓	-
H9es	embryonic stem cell T3	-	-	-	-	-	-	-	✓	-
Helas3	cervical carcinoma cells T2	✓	-	-	✓	-	✓	✓	-	-
Hepg2	hepatocellular carcinoma T2	✓	-	-	✓	-	✓	✓	-	-
Hmec	normal breast tissue T3	✓	-	-	✓	-	✓	✓	✓	-
Hpde6e6e7	pancreatic duct T3	-	-	✓	-	-	-	-	✓	-
Hsmm	muscle T3	✓	-	-	✓	-	✓	✓	✓	-
Hsmmt	muscle T3	✓	-	-	✓	-	✓	✓	-	-
Ips	induced pluripotent stem cell derived from skin fibroblast T3	-	-	-	-	-	-	-	✓	-
K562	leukemia T1	✓	-	-	✓	-	✓	✓	-	-
K562b	Leukemia	-	-	-	-	-	-	✓	-	-
Monocd14ro 1746	monocytes T2	✓	-	-	✓	-	✓	✓	-	
NB4	blood: acute promyelocytic leukemia cell line T3	-	✓	-	-	-	-	-	-	-
Nhdfad	adult dermal fibroblasts T3	✓	-	-	✓	-	✓	✓	-	-

Cell line Symbol[1]	Tissue[2]	rs9555695			rs9521733		rs9521734	rs1999013		
		histone modification	protein binding	chromatin structure	histone modification	chromatin structure	histone modification	histone modification	chromatin structure	motifs
Nhek	skin epidermal keratinocytes T3	✓	-	-	✓	-	✓	✓	✓	✓
Nhlf	lung fibroblast T3	✓	-	-	✓	-	✓	✓	-	-
Osteo		✓	-	-	✓	-	✓	✓	-	-
Osteobl	bone osteoblasts T3	✓	-	-	✓	-	✓	✓	-	-
Panislets	normal pancreas T3	-	-	-	-	-	-	-	✓	-
Phte	primary tracheal epithelial cells T3	-	-	-	-	-	-	-	✓	-
Saec	small airway epithelial cells	-	-	-	-	-	-	✓	-	-
Stellate	liver T3	-	-	-	-	-	-	-	✓	-
Urotsa		-	-	-	-	-	-	-	✓	-
UrotsaUt189		-	-	-	-	-	-	-	✓	-

[1]cell line where experiment performed; [2]tissue where cell line from;
Dark grey cells: evidence from brain, spinal cord or blood vessel tissue.
"-" : no evidence; "✓": evidence

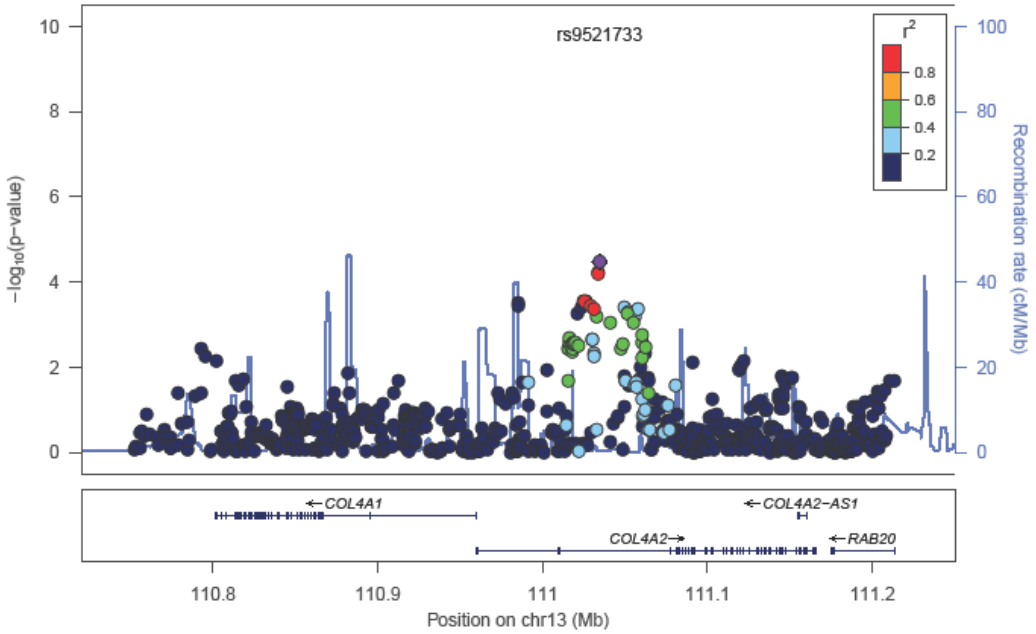
Table 7.12 LD between three significant *COL4A2* SNPs and SNPs associated with other cardiovascular phenotypes

First author Year	Ethnicity of cases	Phenotype	SNP	LD with three ICH SNPs
Tarasov et al, 2009	Sardinian & Old-Order Amish (distant founder population of European ancestry)	Vascular stiffness	rs3742207	r^2: 0.004 to 0.01
O'Donnel et al, 2011	European ancestry	Coronary artery calcification and myocardial infarction	rs3809346 & rs874203	r^2: 0 to 0.002
Schunkert et al, 2011	Majority European ancestry	Coronary artery disease	rs4773144	r^2: 0
Ruigrok et al, 2006	European ancestry	Intracranial aneurysm	rs3783107	r^2: 0.003

LD: linkage disequilibrium

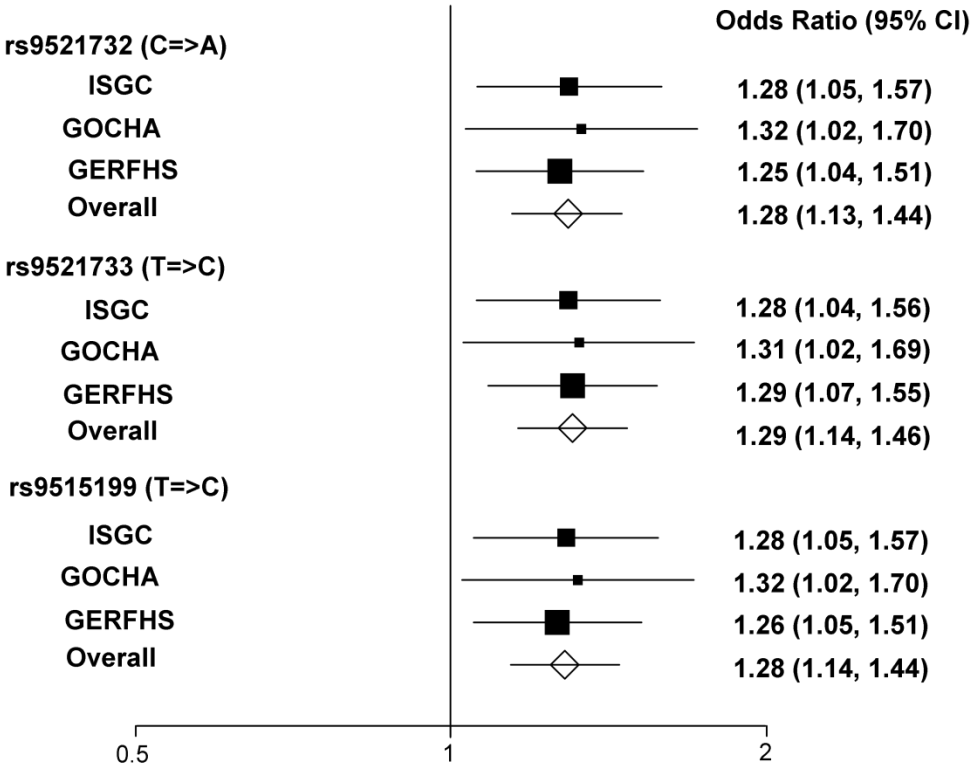
Figures

Figure 7.1 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and deep ICH



Dots mark individual SNPs; y-axis: p-value for association between each SNP and deep ICH; x-axis: position of SNPs. Most significantly associated SNP shown in purple (rs number on top) – colours for other SNPs depend on LD with this lead SNP (see r^2 colour coding on figure). Recombination: exchange of a segment of DNA between two homologous chromosomes during meiosis leading to a novel combination of genetic material in the offspring. Recombination rate: measured as frequency of exchange per unit physical distance (centiMorgan [cM] / Mega base pair [Mb]). cM: unit of linkage that refers to the distance between two gene loci determined by the frequency with which recombination occurs between them. Two loci are said to be one cM apart if recombination is observed between them in 1% of meioses. Mb: a unit of length of nucleic acids, equal to one million base pairs.

Figure 7.2 Associations of significant *COL4A2* SNPs across cohorts included in the meta-analysis of deep ICH



Heterogeneity:

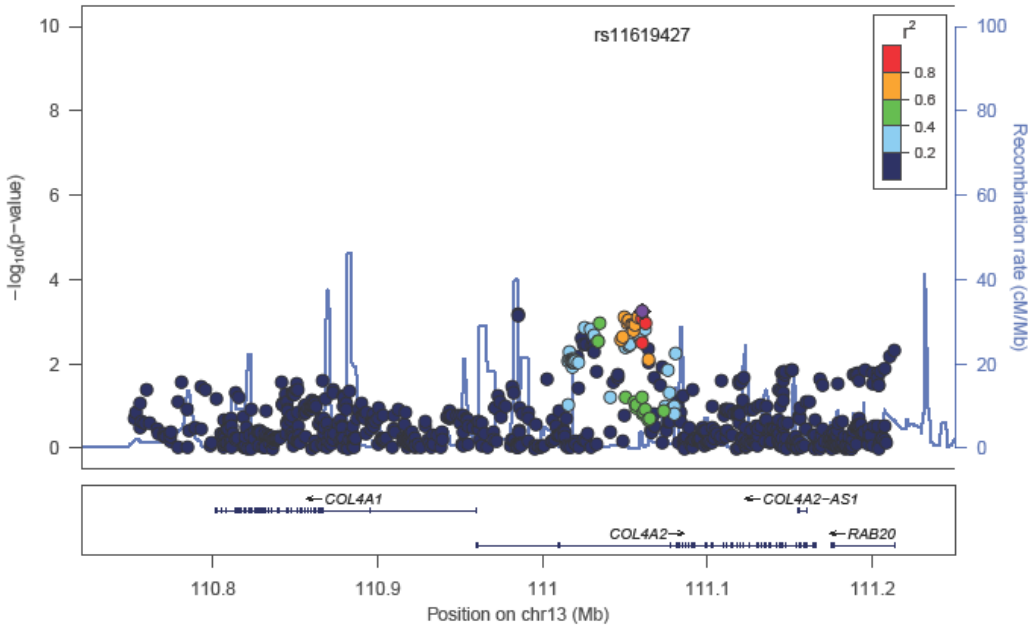
rs9521732 χ^2_{2df} = 0.09; p=0.9563 I^2= 0%

rs9521733 χ^2_{2df} = 0.02; p=0.9903 I^2= 0%

rs9515199 χ^2_{2df} = 0.08; p=0.9607 I^2= 0%

Reproduced from Rannikmäe et al, Neurology 2015; 84(9):918-26. Wolters Kluwer Health Lippincott Williams & Wilkins©.

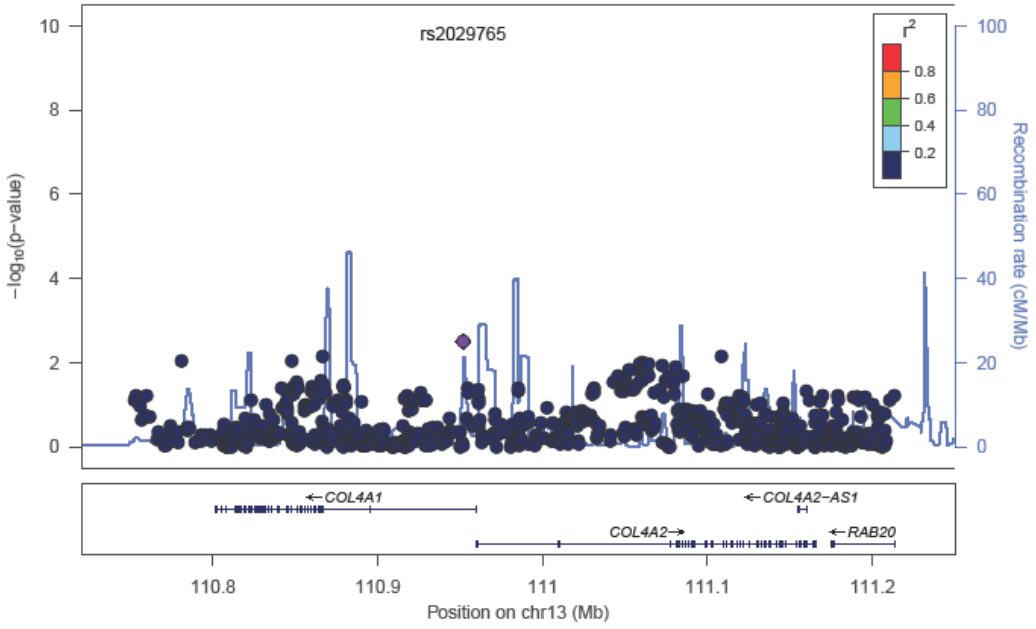
Figure 7.3 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and all ICH



Legend as for Figure 7.1

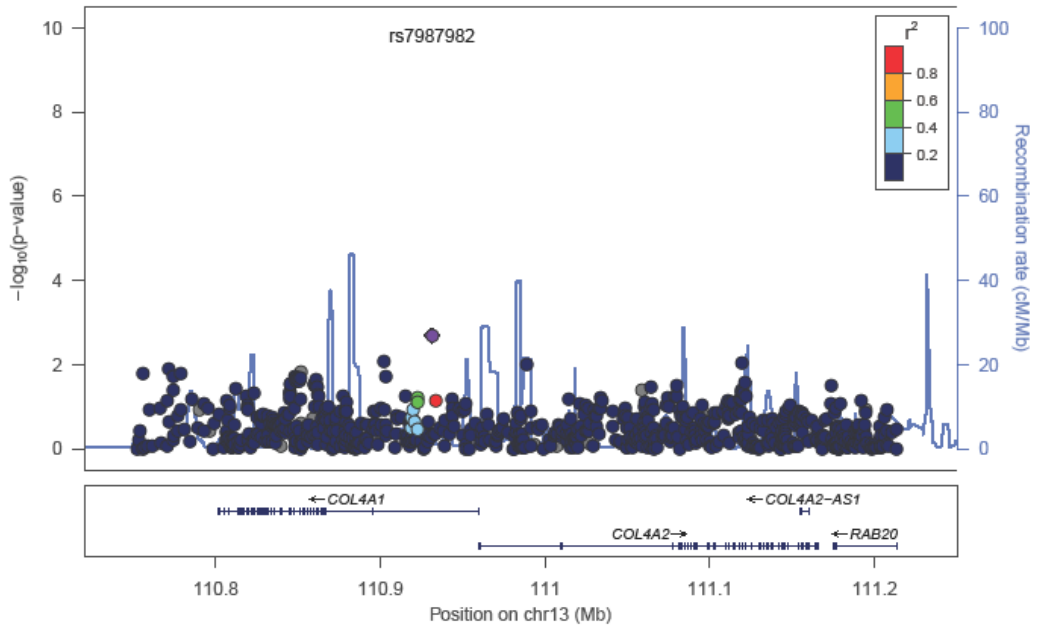
Dots mark individual SNPs; y-axis: p-value for association between each SNP and phenotype; x-axis: position of SNPs.

Figure 7.4 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and lobar ICH



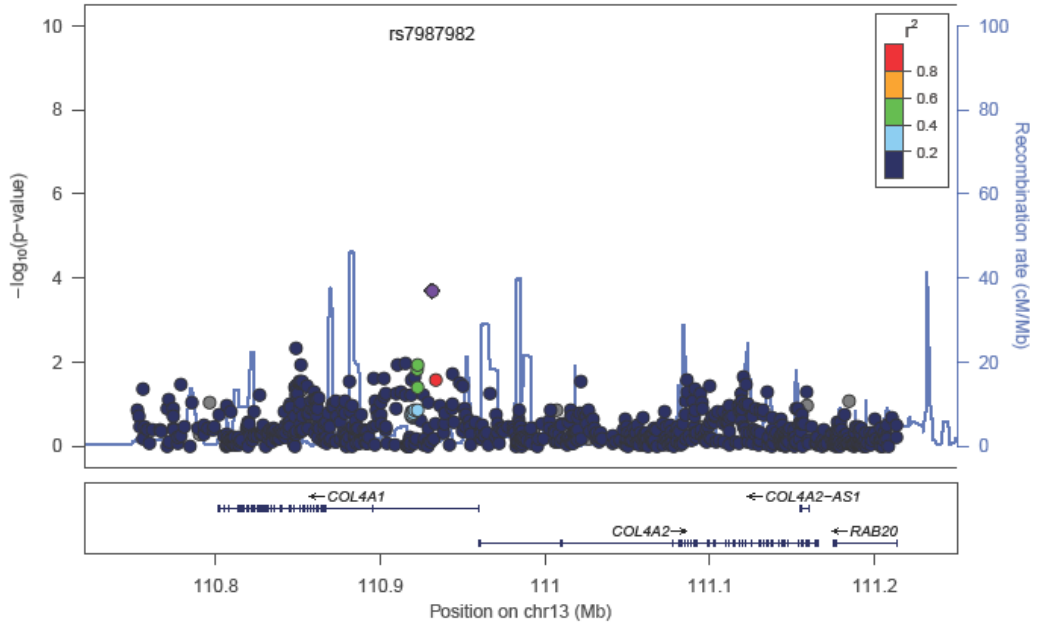
Legend as for Figure 7.1

Figure 7.5 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and all IS



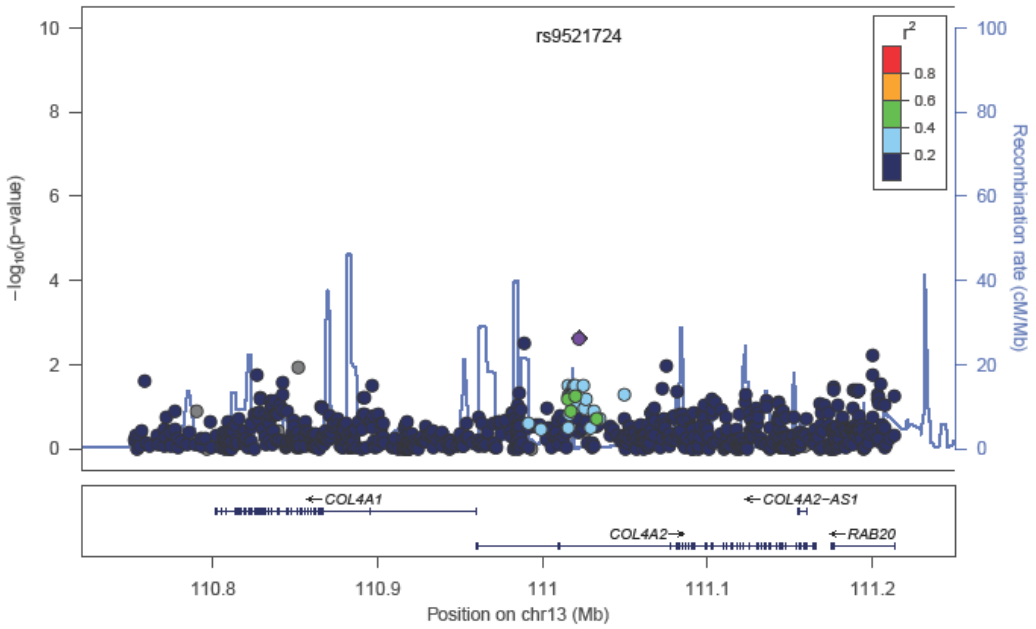
Legend as for Figure 7.1

Figure 7.6 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and CE IS



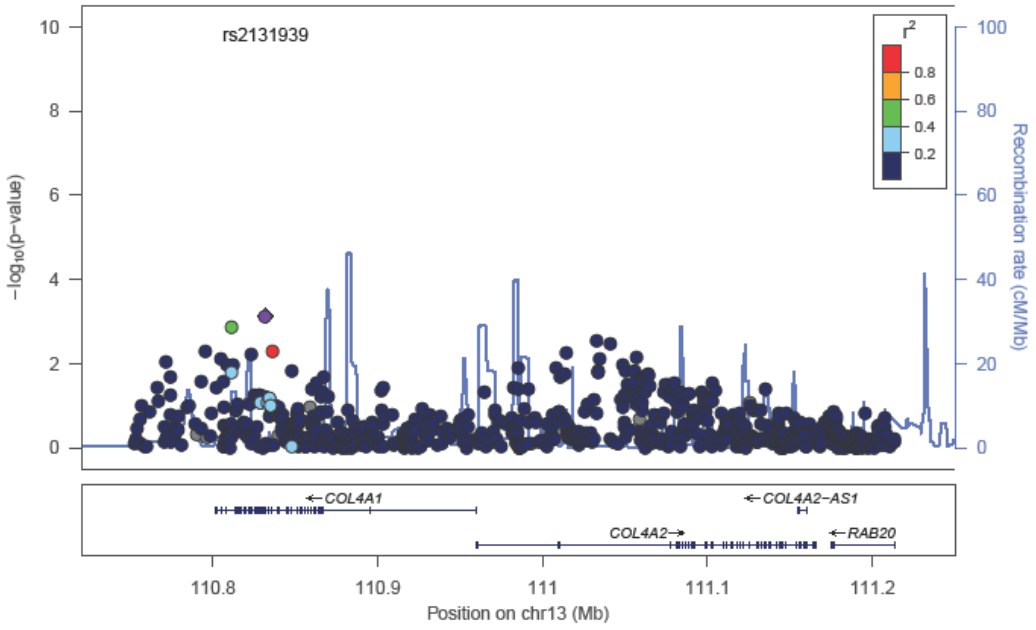
Legend as for Figure 7.1

Figure 7.7 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and LVD IS



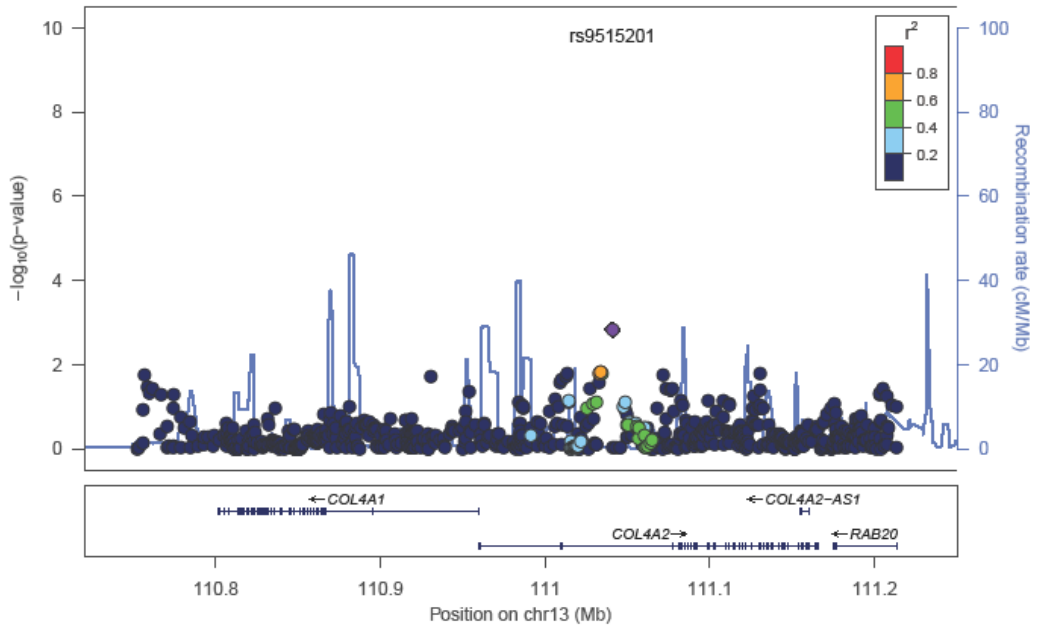
Legend as for Figure 7.1

Figure 7.8 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and lacunar IS



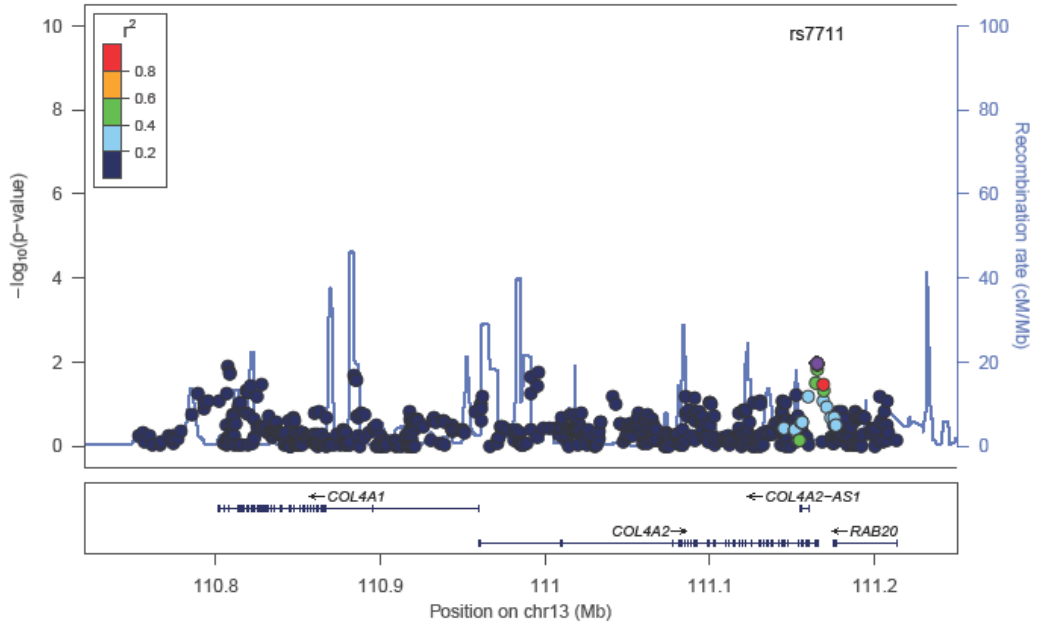
Legend as for Figure 7.1

Figure 7.9 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and WMH in ischaemic stroke



Legend as for Figure 7.1

Figure 7.10 Regional association plot for meta-analysis of association between *COL4A1/COL4A2* SNPs and WMH in the population



Legend as for Figure 7.1

Figure 7.11 Associations between the three *COL4A2* SNPs significantly associated with deep ICH and all included phenotypes

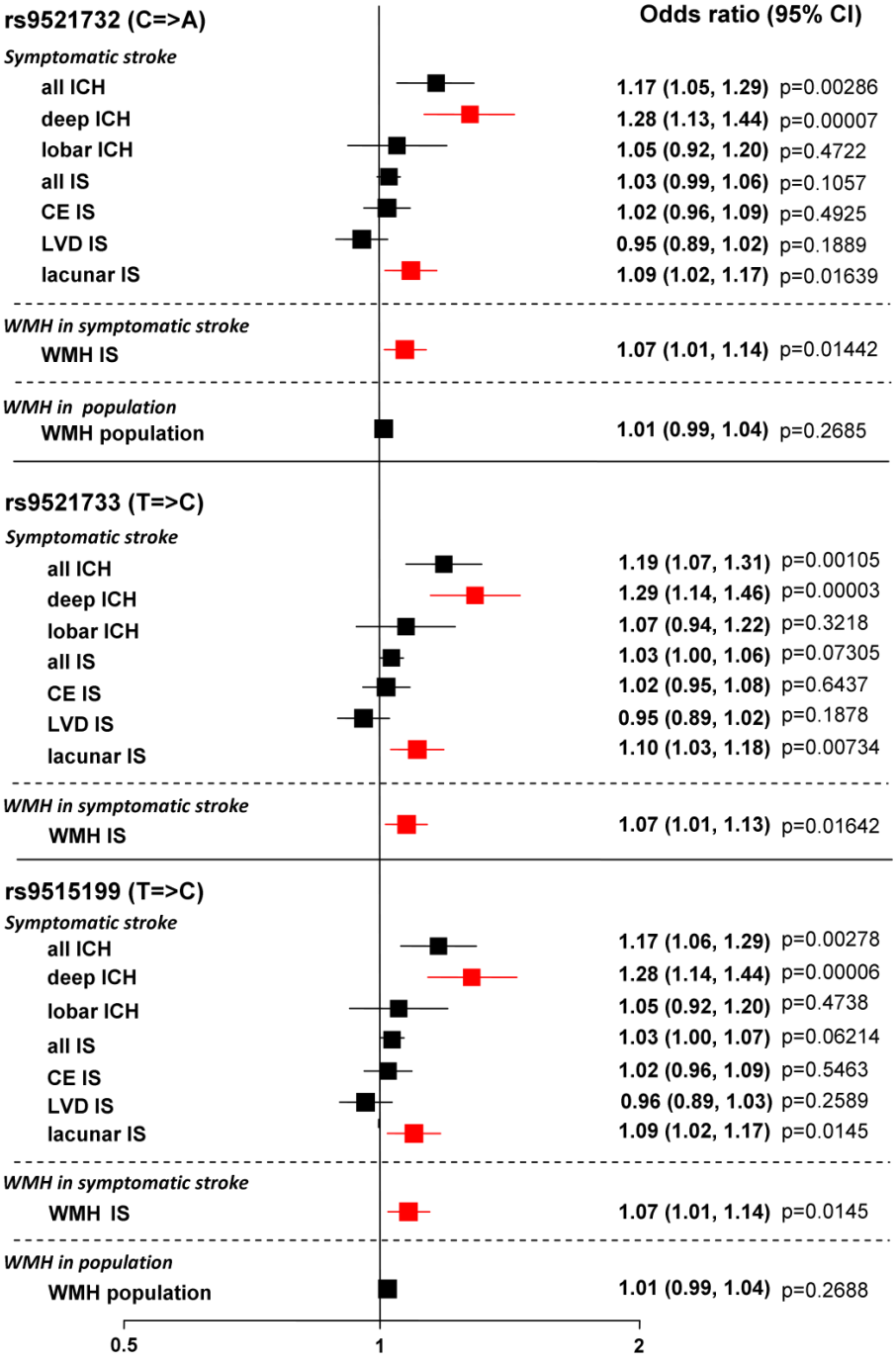
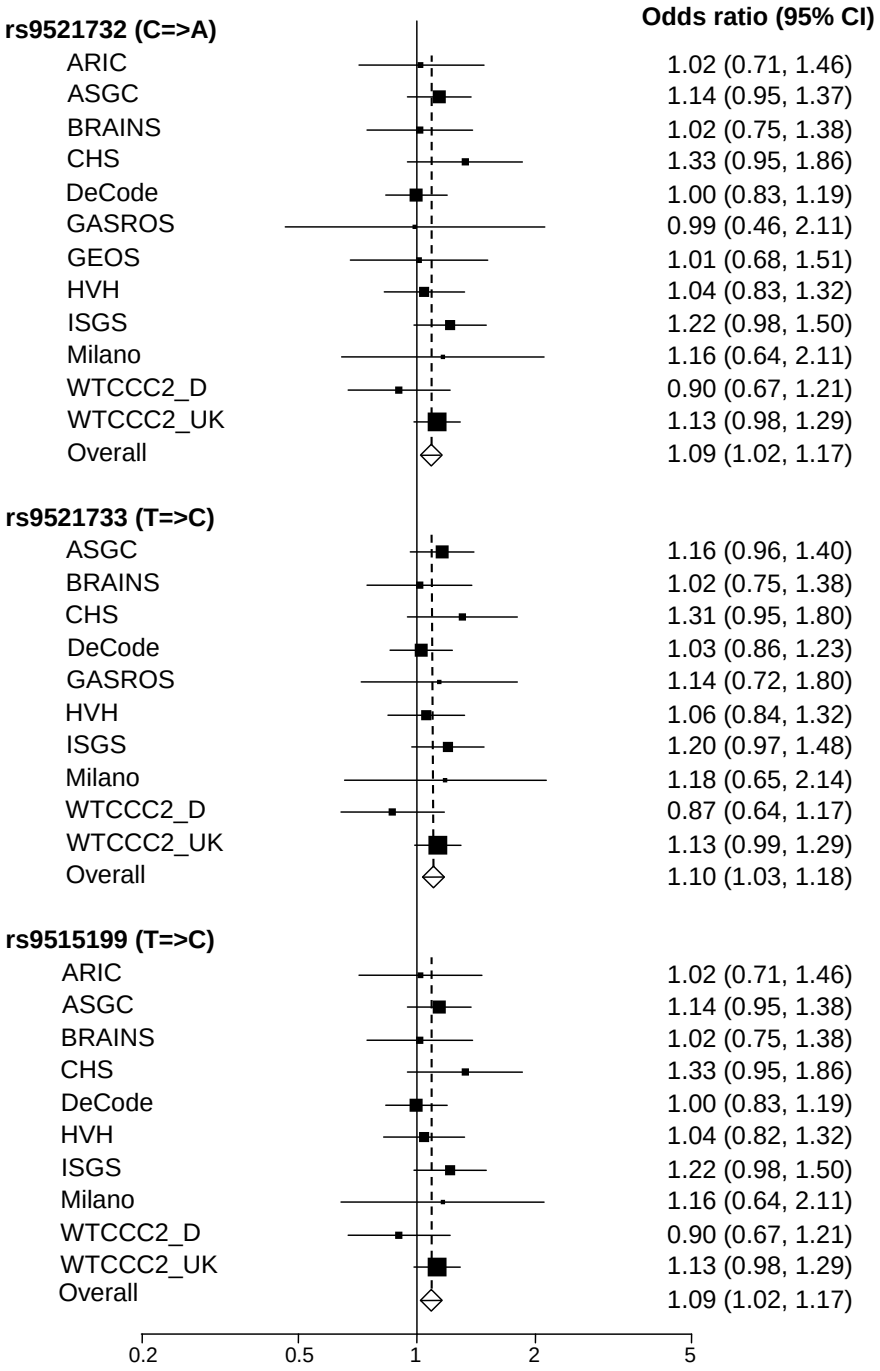


Figure represents pooled ORs across all cohorts. Significant or suggested associations in cerebral SVD phenotypes are shown in red; ICH=intracerebral haemorrhage; IS=ischaemic stroke; CE=cardio-embolic; LVD=large vessel disease; Data for rs9521733 was unavailable in the WMH in population cohorts. Reproduced from Rannikmäe *et al*, *Neurology* 2015; 84(9):918-26. Wolters Kluwer Health Lippincott Williams & Wilkins©.

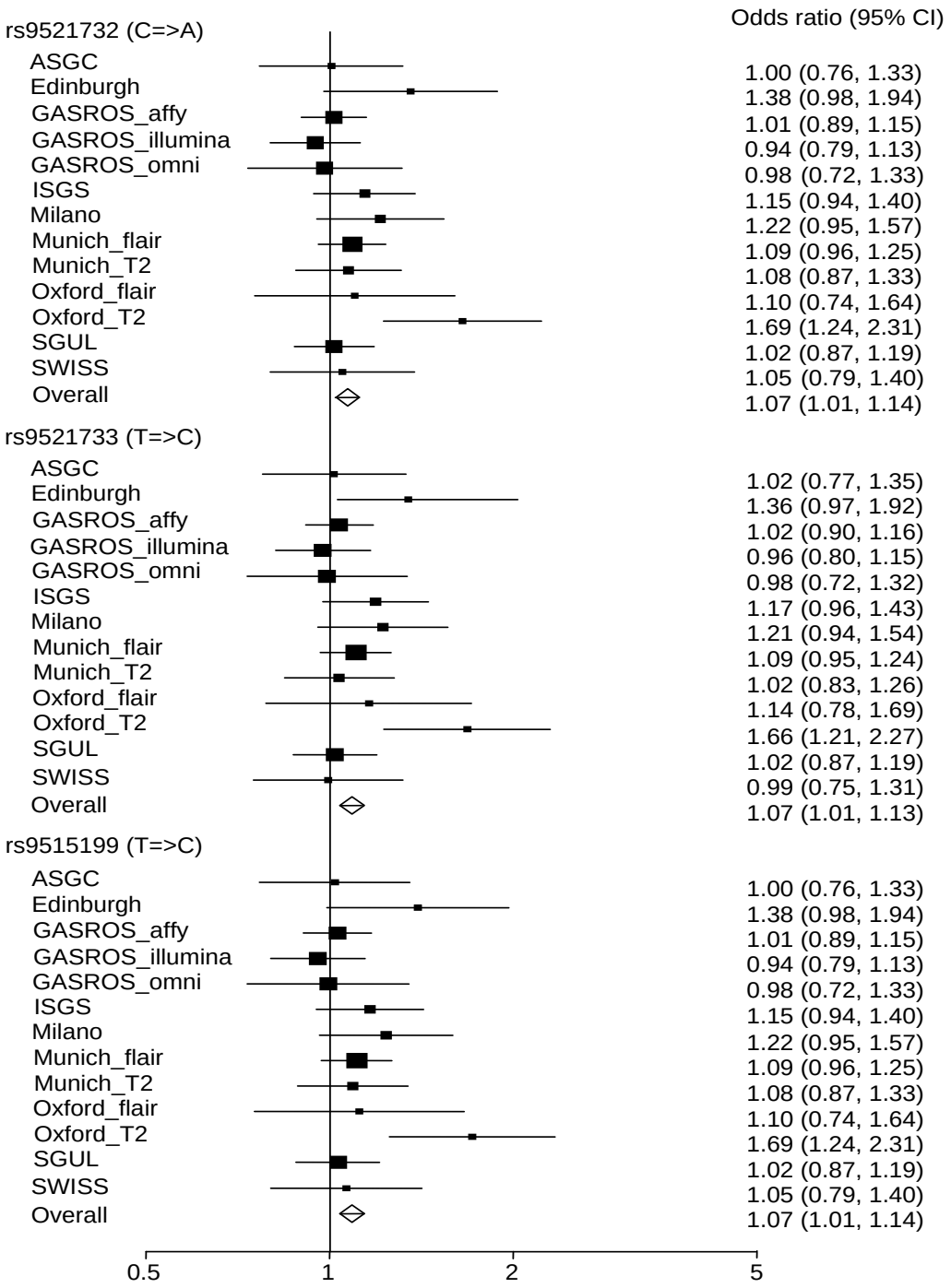
Figure 7.12 Associations between the three *COL4A2* SNPs significantly associated with deep ICH and all cohorts included in the meta-analysis of lacunar IS



Heterogeneity:

rs9521732 χ^2_{11df}= 5.939; p=0.8774 I^2=0%; rs9521733 χ^2_{9df} = 5.607; p=0.7786 I^2=0%; rs9515199 χ^2_{9df} = 5.769; p=0.7628 I^2=0%

Figure 7.13 Associations between the three *COL4A2* SNPs significantly associated with deep ICH and all cohorts included in the meta-analysis of WMH in IS

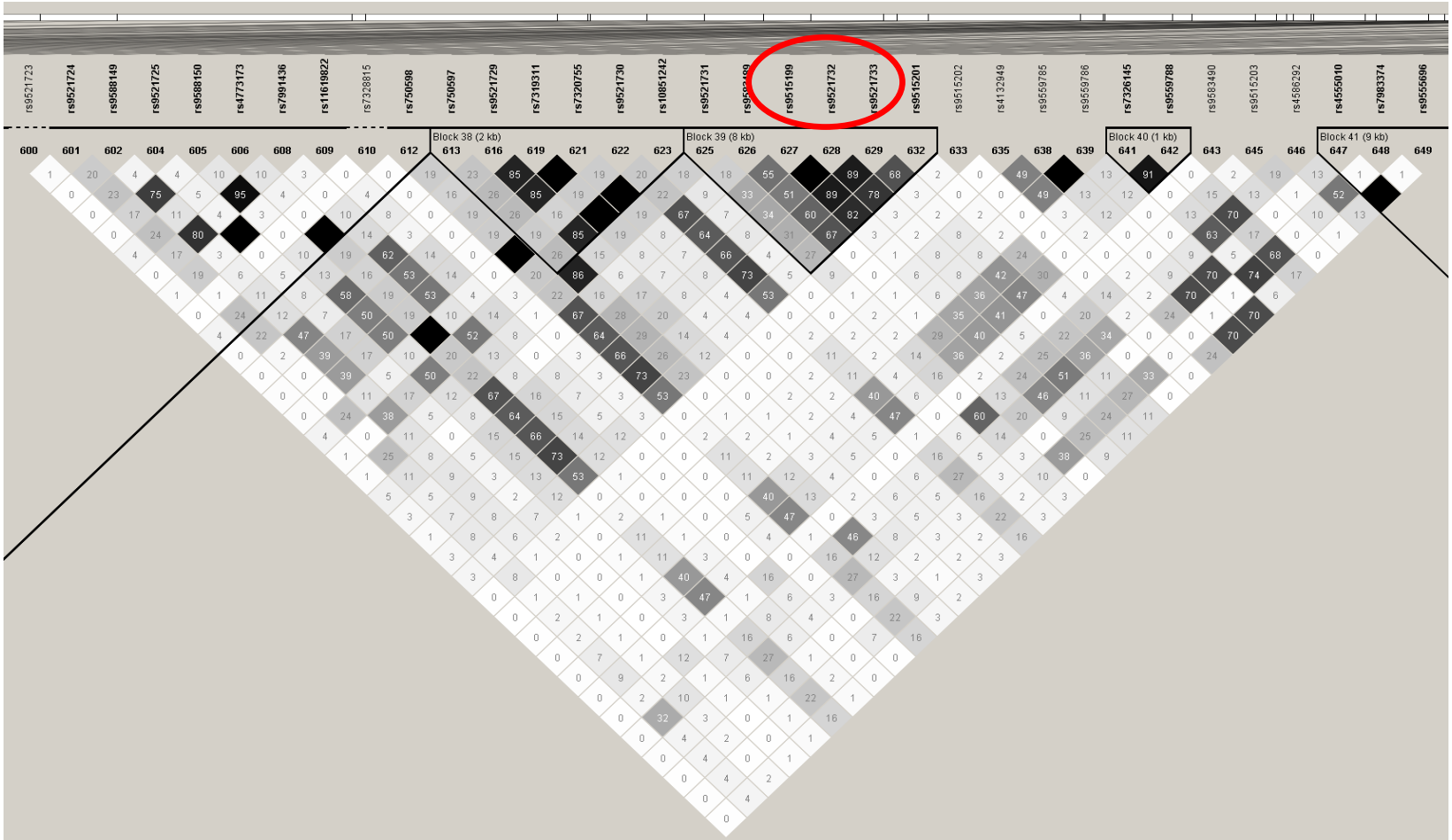


Heterogeneity:

rs9521732 χ^2_{12df} = 15.397; p=0.2204 I^2=22%; rs9521733 χ^2_{12df} = 14.421; p=0.2747 I^2=17%;

rs9515199 χ^2_{12df} = 15.371; p=0.2218 I^2=22%

Figure 7.14 *COL4A2* LD structure in the CEU population



The default algorithm: 95% confidence bounds on D prime are generated and each comparison is called "strong LD", "inconclusive" or "strong recombination". A block is created if 95% of informative (i.e. non-inconclusive) comparisons are "strong LD" (Gabriel et al. 2002). Numbers inside the “diamonds” represent the r^2/100. Colour scheme: white if r^2=0; grey if r^2 >0 and <1; black if r^2=1. The red circle marks the three SNPs significantly associated with a deep ICH phenotype in my analyses. CEU: Utah residents with ancestry from northern and western Europe**.**

Chapter 8: Summary and implications for future research

Chapter contents:

8. Summary and implications for future research

8.1 Main findings of thesis

8.1.1 Reliability of ICH classification systems

My objective in Chapter 2 was to perform a systematic review looking at the inter- and intra-rater reliability of existing ICH classification systems, to inform ICH subtyping in future large-scale prospective studies of ICH, including of ICH genetics.

Among eight relevant studies, the inter- and intra-rater reliabilities of existing anatomical and mechanistic classification systems appeared to be substantial to almost perfect. However, there were insufficient data to assess factors potentially influencing reliability, reliability reporting quality was variable, and because no study had used all possible measures to reduce potential bias, reliability may have been overestimated. Furthermore, since all raters in these studies were experts, uncertainty remains about the generalisability of the results to less expert raters, who might usefully contribute to large-scale research studies.

8.1.2 Genetic associations with CAA

My objective in Chapter 3 was to perform a systematic review and series of meta-analyses to assess the evidence for an association between the APOE ε2/ε3/ε4 polymorphism and histopathologically confirmed CAA.

Among 46 relevant studies in 6645 participants, overall study quality and pathological assessment methods varied widely, while genotyping reporting quality was generally limited. Meta-analyses showed a highly significant, dose-dependent association between *APOE* ε4 and pathologically proven CAA, robust to publication- and other small-study biases, and consistent across subgroups determined by

dementia status, ethnicity, and study quality. Meta-analyses showed no overall robust association between *APOE* ε2 and presence of pathological CAA. I concluded that despite study quality issues, a large body of evidence supports an association of *APOE* ε4 with the presence and severity of histopathologically confirmed CAA, at least in white populations, but there is no convincing robust association of histopathologically confirmed CAA with *APOE* ε2.

My objective in Chapter 4 was to perform a systematic review and series of meta-analyses to assess the evidence for an association between the APOE ε2/ε3/ε4 polymorphism and progression to severe CAA vasculopathy.

Among five included relevant studies in 497 participants study quality was very good. Meta-analyses suggested a possible association of *APOE* ε4, but not of *APOE* ε2, with progression to severe CAA. Despite inclusion of all available relevant data, small numbers of individuals with ε2+ genotypes made the ε2-based analyses very imprecise. While I could not exclude a biologically meaningful association between the *APOE* ε2 allele and progression to severe CAA, there was no convincing evidence to support the popular hypothesis that *APOE* ε2 influences risk of ICH through promoting progression of CAA to its severe form. Much larger numbers of individuals will need to be included in CAA histopathology studies before reliable conclusions can be drawn about *APOE* allele-specific effects on progression to severe CAA.

My objective in Chapter 5 was to perform a systematic review to assess the evidence for an association between the APOE ε2/ε3/ε4 polymorphism and CAA severity in brain pathology samples of cases with genetically proven hereditary CAA.

There were very few relevant studies, data were sparse and incomplete reporting limited interpretation of study quality. There was no reported association between possession of the *APOE* ε4 allele and CAA severity in cases with hereditary CAA Dutch-type and presenilin-1 gene-related familial AD in the three relevant studies. No study reported looking at the association between *APOE* ε2 and CAA severity.

My objective in Chapter 6 was to perform a systematic review to assess the evidence for an association between any (non-APOE ε2/ε3/ε4) genetic polymorphism and histopathologically confirmed CAA.

Among 24 relevant studies (4617 participants) where 25 different candidate genes were studied, there were some positive associations with CAA, but insufficient data for formal meta-analyses. Three reasonably convincing findings merit replication in further larger studies: there was a consistent trend towards an association with a SNP in the transforming growth factor β1 gene (*TGF-β1*) in two studies including 449 participants; one study in 723 participants demonstrated an association between 14 SNPs in the translocase of outer mitochondrial membrane 40 gene (*TOMM40*) and CAA severity; and one study in 544 participants found that a SNP in the complement component receptor 1 gene (*CR1*) was associated with more severe CAA. Otherwise there were no overall significant associations, although some studies reported associations in particular subgroups.

8.1.3 Contribution of *COL4A1* and *COL4A2* to sporadic SVD

My objective in Chapter 7 was to conduct meta-analyses of existing GWAS data to determine if there is an association between common SNPs in the COL4A1/COL4A2 genomic region and clinical / radiological manifestations of SVD.

I conducted a large collaborative individual cohort summary data meta-analysis involving: 1545 ICH cases and 1485 controls; 12 389 ischaemic stroke cases and 62 004 controls; and 12 094 individuals with WMH assessed on MRI scans. I found a significant association between three SNPs in an intronic locus in *COL4A2* and deep ICH, and suggestive associations in the same direction for these SNPs with other cerebral SVD phenotypes: lacunar ischaemic stroke and WMHs in ischaemic stroke cases. These results indicate that the *COL4A2* gene is a locus conferring risk for sporadic forms of symptomatic cerebral SVD, particularly deep ICH. These findings merit replication studies, including in ethnic groups of non-European ancestry.

8.2 Implications for future research

8.2.1 Intracerebral haemorrhage classification

As shown in Chapter 2, reliability of existing ICH classification systems appears very good, but future studies should investigate their performance outside highly specialised centres with less experienced readers. Future comparisons of reliability studies should also be facilitated by studies following published reporting guidelines.

In Chapter 2, I identified and highlighted several limitations of existing ICH classification systems, which, if addressed in future studies, should ensure their further refinement where needed and their appropriate use in diverse clinical and research settings.

Overall, for large population-based prospective studies, including studies of ICH genetics, anatomical classification systems are likely to be more feasible, scalable, and appropriate, since they are free of assumptions about causal pathways. The development of such methods for use at scale will require clear definitions, classification protocols, and the availability of categories for multiple or uncertain locations. Mechanistic systems such as the SMASH-U and its modifications are likely to be appropriate for stratifying patients for clinical trials, some case-control studies, and in clinical practice to encourage a more systematic mechanistic work-up. Mechanistic systems have the advantage of already having a very clear set of rules and excellent reliability, but they could be further improved by including categories for cases with an uncertain and multiple overlapping mechanisms. The feasibility of collecting additional, more detailed information in large, prospective, population-based studies also needs further assessment, since it would complement the more basic information required for anatomical sub-classification and – potentially – allow nested case-control studies based on not only anatomical but also mechanistic information.

8.2.2 Genetic associations with SVD

There may be ethnic variation in genetic associations with CAA, but these have not been widely studied in non-white populations (Tzourio et al. 2008). Genetic associations may also differ by CAA location and subtype. For example, there is preliminary evidence that *APOE* ε4 may be associated with CAA type 1 and ε2 with CAA type 2 (Thal et al. 2002). The suggestion of different genetic influences on amyloid deposition in the vessel wall and progression to advanced vasculopathic changes leading to ICH requires further investigation, in particular the proposed differential effects of ε4 and ε2 alleles. It is also unclear what additional modulating factors – genetic and non-genetic – affect the variable phenotypic expression between cases affected by the same mutation causing hereditary CAA. Future studies should try and replicate associations between *TGF-β1*, *TOMM40* and *CR1* and CAA. Findings of an association between the *COL4A1/COL4A2* locus and sporadic SVD also merit replication, including in ethnic groups of non-European ancestry.

The estimated heritability of common cerebral small vessel disease clinical and radiological phenotypes (lacunar ischaemic stroke, ICH and white matter hyperintensities) ranges from 15% to 80%, but we are only scratching the surface in terms of identifying their genetic architecture. Many explanations for the missing heritability of complex diseases, such as SVD, have been suggested (Manolio et al. 2009). These include:

- small size of existing GWASs meaning much larger number of common variants of smaller effect are yet to be found;
- role of rarer variants with possibly larger effects poorly detected by GWASs;
- role of structural variants poorly captured by existing arrays;
- low power to detect gene-gene and gene-environment interactions.

While larger hypothesis-free GWAS studies are ongoing, alongside preparing for and running the whole exome and genome sequencing studies in large numbers of cases and controls, these could be complemented by hypothesis-driven candidate-gene studies.

Candidate-gene studies have been criticized firstly, for their failure to replicate significant findings in follow-up studies, and secondly, for their dependence on the ability to predict functional candidate genes and polymorphisms.

Regarding the first criticism, there are many potential reasons for their lack of reproducibility, such as differences in the study design, population, and definition of the phenotype. There may also be real biological differences. In addition, sometimes the selected polymorphisms are not likely to be causal having been selected for various other reasons (e.g., ease of genotyping), and hence may not be in linkage disequilibrium with the causal variant across different populations. Future studies could address these criticisms by following proposed methodological criteria to try and reduce potential reasons for lack of reproducibility (Dichgans and Markus 2005).

The second criticism relates to the selection of candidate genes on the basis of imperfect understanding of biological pathways (Manolio et al. 2009, Tabor et al. 2002). I and others (Jeanne et al. 2012, Weng et al. 2012) have demonstrated a potential association between loci in the *COL4A1/COL4A2* region with a cerebral SVD phenotype, although it awaits confirmation in replication studies. This successful approach – selecting candidate genes for a common disease based on their association with a Mendelian form of the same disease – could be one approach used in future studies of SVD genetics.

Where appropriate, identified genetic associations should be followed up by functional studies to prove causation and better understand disease mechanisms.

Future studies should also further investigate the correlation between various pathological, radiological and clinical markers of small vessel diseases. The classification of SVDs into 'SVD associated with CAA' and 'arteriolosclerosis-associated SVD' may be too simplistic, as they may share some risk factors and pathological/radiological/clinical markers. Identifying genetic associations with SVD may be further improved if methods to assess the total SVD load on imaging are developed and used to avoid over-reliance on just one feature, as has occurred in most studies so far (Staals et al. 2014, Wardlaw et al. 2013a).

Finally, ongoing and future research efforts should be informed and directed by up-to-date systematic and unbiased assessments of the body of evidence in the form of systematic reviews and, where appropriate, meta-analyses.

References

Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA & 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010; 467:1061-1073.

Adams HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL & Marsh EE. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 1993; 24:35-41.

Adi D, Xie X, Ma YT, Fu ZY, Yang YN, Li XM, Xiang Y, Liu F & Chen BD. Association of COL4A1 genetic polymorphisms with coronary artery disease in Uygur population in Xinjiang, China. Lipids Health Dis 2013; 12:153.

Adib-Samii P, Devan W, Traylor M, Lanfranconi S, Zhang CR, Cloonan L, Falcone GJ, Radmanesh F, Fitzpatrick K, Kanakis A, Rothwell PM, Sudlow C, Boncoraglio GB, Meschia JF, Levi C, Dichgans M, Bevan S, Rosand J, Rost NS & Markus HS. Genetic architecture of white matter hyperintensities differs in hypertensive and nonhypertensive ischemic stroke. Stroke 2015; 46:348-353.

Adib-Samii P, Rost N, Traylor M, Devan W, Biffi A, Lanfranconi S, Fitzpatrick K, Bevan S, Kanakis A, Valant V, Gschwendtner A, Malik R, Richie A, Gamble D, Segal H, Parati EA, Ciusani E, Holliday EG, Maguire J, Wardlaw J, Worrall B, Bis J, Wiggins KL, Longstreth W, Kittner SJ, Cheng YC, Mosley T, Falcone GJ, Furie KL, Leiva-Salinas C, Lau BC, Saleem Khan M, Sharma P, Fornage M, Mitchell BD, Psaty BM, Sudlow C, Levi C, Boncoraglio GB, Rothwell PM, Meschia J, Dichgans M, Rosand J, Markus HS, Australian Stroke Genetics Collaborative, Wellcome Trust Case-Control Consortium-2 (WTCCC2), METASTROKE & International Stroke Genetics Consortium. 17q25 Locus is associated with white matter hyperintensity volume in ischemic stroke, but not with lacunar stroke status. Stroke 2013; 44:1609-1615.

Al-Shahi Salman R, Labovitz DL & Stapf C. Spontaneous intracerebral haemorrhage. BMJ 2009; 2009; 339:b2586.

Alafuzoff I, Aho L, Helisalmi S, Mannermaa A & Soininen H. Beta-amyloid deposition in brains of subjects with diabetes. Neuropathol Appl Neurobiol 2009; 35:60-68.

Alafuzoff I, Seppo H, Mannermaa A, Riekkinen PS & Soininen H. Beta-amyloid load is not influenced by the severity of cardiovascular disease in aged and demented patients. Stroke 1999; 30:613-618.

Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo

N, Gonzaga-Jauregui C, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE & International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. Nature 2010; 467:52-58.

Amarenco P, Bogousslavsky J, Caplan LR, Donnan GA & Hennerici MG. New approach to stroke subtyping: the A-S-C-O (phenotypic) classification of stroke. Cerebrovasc Dis 2009; 27:502-508.

Anderson CD, Biffi A, Nalls MA, Devan WJ, Schwab K, Ayres AM, Valant V, Ross OA, Rost NS, Saxena R, Viswanathan A, Worrall BB, Brott TG, Goldstein JN, Brown D, Broderick JP, Norrving B, Greenberg SM, Silliman SL, Hansen BM, Tirschwell DL, Lindgren A, Slowik A, Schmidt R, Selim M, Roquer J, Montaner J, Singleton AB, Kidwell CS, Woo D, Furie KL, Meschia JF, Rosand J & the International Stroke Genetics Consortium. Common variants within oxidative phosphorylation genes influence risk of ischemic stroke and intracerebral hemorrhage. Stroke 2013; 44:612-619.

Anderson CS, Chakera TMH, Stewart-Wynne EG & Jamrozik KD. Spectrum of primary intracerebral haemorrhage in Perth, Western Australia, 1989-90: incidence and outcome. J Neurol Neurosurg Psychiatry 1994; 57:936-940.

Aribisala BS, Valdés Hernández MC, Royle NA, Morris Z, Muñoz Maniega S, Bastin ME, Deary IJ & Wardlaw JM. Brain atrophy associations with white matter lesions in the ageing brain: the Lothian Birth Cohort 1936. Eur Radiol 2013; 23:1084-1092.

Arsava EM, Ballabio E, Benner T, Cole JW, Delgado-Martinez MP, Dichgans M, Fazekas F, Furie KL, Illoh K, Jood K, Kittner S, Lindgren AG, Majersik JJ, Macleod MJ, Meurer WJ, Montaner J, Olugbodi AA, Pasdar A, Redfors P, Schmidt R, Sharma P, Singhal AB, Sorensen AG, Sudlow C, Thijs V, Worrall BB, Rosand J, Ay H & the International Stroke Genetics Consortium. The Causative Classification of Stroke system: an international reliability and optimization study. Neurology 2010; 75:1277-1284.

Attems J. Sporadic cerebral amyloid angiopathy: Pathology, clinical implications, and possible pathomechanisms. Acta Neuropathol 2005; 110:345-359.

Attems J, Jellinger KA & Lintner F. Alzheimer's disease pathology influences severity and topographical distribution of cerebral amyloid angiopathy. Acta Neuropathol 2005; 110:222-231.

Attems J, Lauda F & Jellinger KA. Unexpectedly low prevalence of intracerebral hemorrhages in sporadic cerebral amyloid angiopathy: an autopsy study. J Neurol 2008; 255:70-76.

Ay H, Furie KL, Singhal A, Smith WS, Sorensen AG & Koroshetz WJ. An evidence-based causative classification system for acute ischemic stroke. Ann Neurol 2005; 58:688-697.

Bamford J, Sandercock P, Dennis M, Burn J & Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 1991; 337:1521-1526.

Barrett JC, Fry B, Maller J & Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21:263-265.

Barton CW & Hemphill JC, III. Cumulative Dose of Hypertension Predicts Outcome in Intracranial Hemorrhage Better than American Heart Association Guidelines. Acad Emerg Med 2007; 14:695-701.

Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, Corneveaux JJ, Hardy J, Vonsattel JP, Younkin SG, Bennett DA, De Jager PL, Larson EB, Crane PK, Kamboh MI, Kofler JK, Mash DC, Duque L, Gilbert JR, Gwirtsman H, Buxbaum JD, Kramer P, Dickson DW, Farrer LA, Frosch MP, Ghetti B, Haines JL, Hyman BT, Kukull WA, Mayeux RP, Pericak-Vance MA, Schneider JA, Trojanowski JQ, Reiman EM, Schellenberg GD, Montine TJ & Alzheimer's Disease Genetics Consortium (ADGC). Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet 2014; 10:e1004606.

Berg L, McKeel DWJ, Miller JP, Storandt M, Rubin EH, Morris JC, Baty J, Coats M, Norton J, Goate AM, Price JL, Gearing M, Mirra SS & Saunders AM. Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. Arch Neurol 1998; 55:326-335.

Bertram L, McQueen MB, Mullin K, Blacker D & Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 2007; 39:17-23.

Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dichgans M & Markus HS. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke 2012; 43:3161-3167.

Bhattathiri SP, Gregson B, Prasad KS, Mitchell P, Soh C, Mitra D, Gholkar A & Mendelow AD. Reliability assessment of computerized tomography scanning measurements in intracerebral hematoma. Neurosurg Focus 2003; 15:E6.

Biffi A, Anderson CD, Jagiella JM, Schmidt H, Kissela B, Hansen BM, Jimenez-Conde J, Pires CR, Ayres AM, Schwab K, Cortellini L, Pera J, Urbanik A, Romero JM, Rost NS, Goldstein JN, Viswanathan A, Pichler A, Enzinger C, Rabionet R, Norrving B, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Broderick JP, Greenberg SM, Roquer J, Lindgren A, Slowik A, Schmidt R, Woo D, Rosand J & the International Stroke Genetics Consortium.

APOE genotype and extent of bleeding and outcome in lobar intracerebral haemorrhage: a genetic association study. Lancet Neurol 2011; 10:702-709.

Biffi A, Shulman JM, Jagiella JM, Cortellini L, Ayres AM, Schwab K, Brown DL, Silliman SL, Selim M, Worrall BB, Meschia JF, Slowik A, De Jager PL, Greenberg SM, Schneider JA, Bennett DA & Rosand J. Genetic variation at CR1 increases risk of cerebral amyloid angiopathy. Neurology 2012; 78:334-341.

Biffi A, Sonni A, Anderson CD, Kissela B, Jagiella JM, Schmidt H, Jimenez-Conde J, Hansen BM, Fernandez-Cadenas I, Cortellini L, Ayres A, Schwab K, Juchniewicz K, Urbanik A, Rost NS, Viswanathan A, Seifert-Held T, Stoegerer EM, Tomás M, Rabionet R, Estivill X, Brown DL, Silliman SL, Selim M, Worrall BB, Meschia JF, Montaner J, Lindgren A, Roquer J, Schmidt R, Greenberg SM, Slowik A, Broderick JP, Woo D, Rosand J & the International Stroke Genetics Consortium. Variants of APOE influence risk of deep and lobar intracerebral hemorrhage. Ann Neurol 2010; 68:934-943.

Borenstein M, Hedges LV, Higgins J & Rothstein H. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Syn Meth 2010; 1:97-111.

Borenstein M, Hedges LV, Higgins JPT & Rothstein HR. 2009. "Publication bias" in *Introduction to Meta-Analysis*. Printed in the UK by TJ International, Padstow, Cornwall: John Wiley & Sons ltd, 277-291.

Bornebroek M, De Jonghe C, Haan J, Kumar-Singh S, Younkin S, Roos R & Van Broeckhoven C. Hereditary cerebral hemorrhage with amyloidosis Dutch type (AbetaPP 693): decreased plasma amyloid-beta 42 concentration. Neurobiol Dis 2003; 14:619-623.

Bornebroek M, Haan J, Van Duinen SG, Maat-Schieman MLC, Van Buchem MA, Bakker E, Van Broeckhoven C & Roos RAC. Dutch hereditary cerebral amyloid angiopathy: structural lesions and apolipoprotein E genotype. Ann Neurol 1997a; 41:695-698.

Bornebroek M, Westendorp RG, Haan J, Bakker E, Timmers WF, Van Broeckhoven C & Roos RAC. Mortality from hereditary cerebral haemorrhage with amyloidosis-Dutch type. The impact of sex, parental transmission and year of birth. Brain 1997b; 120:2243-2249.

Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM & Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012; 22:1790-1797.

Brouwers HB, Biffi A, Ayres AM, Schwab K, Cortellini L, Romero JM, Rost NS, Viswanathan A, Greenberg SM, Rosand J & Goldstein JN. Apolipoprotein E genotype predicts hematoma expansion in lobar intracerebral hemorrhage. Stroke 2012; 43:1490-1495.

Bugiani O, Giaccone G, Rossi G, Mangieri M, Capobianco R, Morbin M, Mazzoleni G, Cupidi C, Marcon G, Giovagnoli A, Bizzi A, Di Fede G, Puoti G, Carella F, Salmaggi A, Romorini A, Patruno GM, Magoni M, Padovani A & Tagliavini F.

Hereditary cerebral hemorrhage with amyloidosis associated with the E693K mutation of APP. Arch Neurol 2010; 67:987-995.

Burton PR, Hansell AL, Fortier I, Manolio TA, Khoury MJ, Little J & Elliott P 2009. Size matters: just how big is BIG?: Quantifying realistic sample size requirements for human genome epidemiology. Int J Epidemiol 2009; 38:263-273.

Cabrejo L, Guyant-Maréchal L, Laquerrière A, Vercelletto M, De La Fournière F, Thomas-Antérion C, Verny C, Letournel F, Pasquier F, Vital A, Checler F, Frebourg T, Campion D & Hannequin D. Phenotype associated with APP duplication in five families. Brain 2006; 129:2966-2976.

Caselli RJ, Walker D, Sue L, Sabbagh M & Beach T. Amyloid load in nondemented brains correlates with APOE e4. Neurosci Lett 2010; 473:168-171.

Chabriat H, Joutel A, Dichgans M, Tournier-Lasserve E & Bousser MG. Cadasil. Lancet Neurol 2009; 8:643-653.

Chalmers K, Wilcock GK & Love S. APOE epsilon 4 influences the pathological phenotype of Alzheimer's disease by favouring cerebrovascular over parenchymal accumulation of A beta protein. Neuropathol Appl Neurobiol 2003; 29:231-238.

Chalmers KA, Culpan D, Kehoe PG, Wilcock GK, Hughes A & Love S. APOE promoter, ACE1 and CYP46 polymorphisms and β-amyloid in Alzheimer's disease. NeuroReport 2004; 15:95-98.

Chapuis J, Hot D, Hansmannel F, Kerdraon O, Ferreira S, Hubans C, Maurage CA, Huot L, Bensemain F, Laumet G, Ayral AM, Fievet N, Hauw JJ, Dekosky ST, Lemoine Y, Iwatsubo T, Wavrant-Devrieze F, Dartigues JF, Tzourio C, Buee L, Pasquier F, Berr C, Mann D, Lendon C, Alperovitch A, Kamboh MI, Amouyel P & Lambert JC. Transcriptomic and genetic studies identify IL-33 as a candidate gene for Alzheimer's disease. Mol Psychiatry 2009; 14:1004-1016.

Chapuis J, Tian J, Shi J, Bensemain F, Cottel D, Lendon C, Amouyel P, Mann D & Lambert JC. Association study of the vascular endothelial growth factor gene with the risk of developing Alzheimer's disease. Neurobiol aging 2006; 27:1212-1215.

Charcot JM & Bouchard C. Nouvelles recherches sur la pathogenie de l'hemorrhagie cerebrale Arch Physiol Norm Pathol 1868; 1:110-127.

Charidimou A, Gang Q & Werring DJ. Sporadic cerebral amyloid angiopathy revisited: recent insights into pathophysiology and clinical spectrum. J Neurol Neurosurg Psychiatry 2012a; 83:124-137.

Charidimou A, Jaunmuktane Z, Baron JC, Burnell M, Varlet P, Peeters A, Xuereb J, Jäger R, Brandner S & Werring DJ. White matter perivascular spaces: an MRI marker in pathology-proven cerebral amyloid angiopathy? Neurology 2014; 82:57-62.

Charidimou A, Martinez-Ramirez S, Shoamanesh A, Oliveira-Filho J, Frosch M, Vashkevich A, Ayres A, Rosand J, Gurol ME, Greenberg SM & Viswanathan A.

Cerebral amyloid angiopathy with and without hemorrhage: evidence for different disease phenotypes. Neurology 2015a; 84:1206-1212.

Charidimou A, Meegahage R, Fox Z, Peeters A, Vandermeeren Y, Laloux P, Baron JC, Jäger HR & Werring DJ. Enlarged perivascular spaces as a marker of underlying arteriopathy in intracerebral haemorrhage: a multicentre MRI cohort study. J Neurol Neurosurg Psychiatry 2013; 84:624-629.

Charidimou A, Peeters A, Fox Z, Gregoire SM, Vandermeeren Y, Laloux P, Jäger HR, Baron JC & Werring DJ. Spectrum of transient focal neurological episodes in cerebral amyloid angiopathy: multicentre magnetic resonance imaging cohort study and meta-analysis. Stroke 2012b; 43:2324-2330.

Charidimou A, Wilson D, Shakeshaft C, Ambler G, White M, Cohen H, Yousry T, Al-Shahi Salman R, Lip G, Houlden H, Jäger HR, Brown MM & Werring DJ. The Clinical Relevance of Microbleeds in Stroke study (CROMIS-2): rationale, design, and methods. Int J Stroke 2015b; 10 Suppl A100:155-61.

Chen YW, Lee MJ & Smith EE. Cerebral amyloid angiopathy in East and West. Int J Stroke 2010; 5:403-411.

Chi NF, Chien LN, Ku HL, Hu CJ & Chiou HY. Alzheimer disease and risk of stroke: a population-based cohort study. Neurology 2013; 80:705-711.

Chibnik LB, Shulman JM, Leurgans SE, Schneider JA, Wilson RS, Tran D, Aubin C, Buchman AS, Heward CB, Myers AJ, Hardy JA, Huentelman MJ, Corneveaux JJ, Reiman EM, Evans DA, Bennett DA & De Jager PL. CR1 is associated with amyloid plaque burden and age-related cognitive decline. Ann Neurol 2011; 69:560-569.

Chiewvit P, Danchaivijitr N, Nilanont Y & Poungvarin N. Computed Tomographic Findings in Non-traumatic Hemorrhagic Stroke. J Med Assoc Thai 2009; 92:73-86.

Chinn S. A simple method for converting an odds ratio to effect size for use in meta analysis. Stat Med 2000; 19:3127-3131.

Christoforidis M, Schober R & Krohn K. Genetic-morphologic association study: association between the low density lipoprotein-receptor related protein (LRP) and cerebral amyloid angiopathy. Neuropathol Appl Neurobiol 2005; 31:11-19.

Chui HC, Zarow C, Mack WJ, Ellis WG, Zheng L, Jagust WJ, Mungas D, Reed BR, Kramer JH, DeCarli CC, Weiner MW & Vinters HV. Cognitive impact of subcortical vascular and Alzheimer's disease pathology. Ann Neurol 2006; 60:677-687.

Cordonnier C, Al-Shahi Salman R & Wardlaw J. Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting. Brain 2007; 130:1988-2003.

Cordonnier C, Klijn CJ, van Beijnum J & Al-Shahi Salman R. Radiological Investigation of Spontaneous Intracerebral Hemorrhage: Systematic Review and Trinational Survey. Stroke 2010; 41:685-690.

Cruz-Sánchez FF, Durany N, Thome J, Riederer P & Zambón D. Correlation between Apolipoprotein-E polymorphism and Alzheimer's disease pathology. J Alzheimers Dis 2000; 2:223-229.

Cuenco KT, Lunetta KL, Baldwin CT, McKee AC, Guo J, Cupples LA, Green RC, George-Hyslop PH, Chui H, DeCarli C & Farrer LA for the Mirage Study Group. Association of distinct variants in SORL1 with cerebrovascular and neurodegenerative changes related to Alzheimer disease. Arch Neurol 2008; 65:1640-1648.

Davidson Y, Gibbons L, Pritchard A, Hardicre J, Wren J, Tian J, Shi J, Stopford C, Julien C, Thompson J, Payton A, Thaker U, Hayes AJ, Iwatsubo T, Pickering-Brown SM, Pendleton N, Horan MA, Burns A, Purandare N, Lendon CL, Neary D, Snowden JS & Mann DM. Genetic associations between cathepsin D exon 2 C->T polymorphism and Alzheimer's disease, and pathological correlations with genotype. J Neurol Neurosurg Psychiatry 2006; 77:515-517.

Davis DG, Schmitt FA, Wekstein DR & Markesbery WR. Alzheimer neuropathologic alterations in aged cognitively normal subjects. J Neuropathol Exp Neurol 1999; 58:376-388.

de Vet HC, Terwee CB, Knol DL & Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006; 59:1033-1039.

Devan WJ, Falcone GJ, Anderson CD, Jagiella JM, Schmidt H, Hansen BM, Jimenez-Conde J, Giralt-Steinhauer E, Cuadrado-Godia E, Soriano C, Ayres AM, Schwab K, Kassis SB, Valant V, Pera J, Urbanik A, Viswanathan A, Rost NS, Goldstein JN, Freudenberger P, Stögerer EM, Norrving B, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Montaner J, Fernandez-Cadenas I, Delgado P, Greenberg SM, Roquer J, Lindgren A, Slowik A, Schmidt R, Woo D, Rosand J, Biffi A & the International Stroke Genetics Consortium. Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage. Stroke 2013; 44:1578-1583.

Díaz-Guzmán J, Egido-Herrero JA, Fuentes B, Fernández-Pérez C, Gabriel-Sánchez R, Barberà G, Abilleira S & Proyecto Ictus del Grupo de Estudio de Enfermedades Cerebrovasculares de la Sociedad Española de Neurología. Incidencia de ictus en España: estudio Iberictus. Datos del estudio piloto. [Incidence of strokes in Spain: the Iberictus study. Data from the pilot study]. Rev Neurol 2009; 48:61-65.

Dichgans M & Markus HS. Genetic association studies in stroke: methodological issues and proposed standard criteria. Stroke 2005; 36:2027-2031.

Dickson DW. Required techniques and useful molecular markers in the neuropathologic diagnosis of neurodegenerative diseases. Acta Neuropathol 2005; 109:14-24.

Dorfman VB, Pasquini L, Riudavets M, López-Costa JJ, Villegas A, Troncoso JC, Lopera F, Castaño EM & Morelli L. Differential cerebral deposition of IDE and NEP in sporadic and familial Alzheimer's disease. Neurobiol Aging 2010; 31:1743-1757.

Doubal FN, MacLullich AM, Ferguson KJ, Dennis MS & Wardlaw JM. Enlarged perivascular spaces on MRI are a feature of cerebral small vessel disease. Stroke 2010; 41:450-454.

Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ & Scheltens P. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurology 2013; 9:1118-1127.

Durany N, Ravid R, Riederer P & Cruz-Sánchez FF. Increased frequency of the α1-antichymotrypsin T allele in cerebral amyloid angiopathy. Neuropathology 2000; 20:184-189.

Ellis RJ, Olichney JM, Thal LJ, Mirra SS, Morris JC, Beekly D & Heyman A. Cerebral amyloid angiopathy in the brains of patients with Alzheimer's disease: the CERAD experience, Part XV. Neurology 1996; 46:1592-1596.

Eng JA, Frosch MP, Choi K, Rebeck GW & Greenberg SM. Clinical manifestations of cerebral amyloid angiopathy-related inflammation. Ann Neurol 2004; 55:250-256.

Esiri M, Chance S, Joachim C, Warden D, Smallwood A, Sloan C, Christie S, Wilcock G & Smith AD. Cerebral amyloid angiopathy, subcortical white matter disease and dementia: literature review and study in OPTIMA. Brain Pathol 2015; 25:51-62.

Etiene D, Kraft J, Ganju N, Gomez-Isla T, Gemelli B, Hyman BT, Hedley-Whyte ET, Wands JR & De La Monte SM. Cerebrovascular pathology contributes to the heterogeneity of Alzheimer's disease. J Alzheimers Dis 1998; 1:119-134.

Falcone GJ, Biffi A, Devan WJ, Brouwers HB, Anderson CD, Valant V, Ayres AM, Schwab K, Rost NS, Goldstein JN, Viswanathan A, Greenberg SM, Selim M, Meschia JF, Brown DL, Worrall BB, Silliman SL, Tirschwell DL, Rosand J & GOCHA Investigators. Burden of blood pressure-related alleles is associated with larger hematoma volume and worse outcome in intracerebral hemorrhage. Stroke 2013; 44:321-326.

Falcone GJ, Malik R, Dichgans M & Rosand J. Current concepts and clinical applications of stroke genetics. Lancet Neurol 2014; 13:405-418.

Fazekas F, Kleinert R, Roob G, Kleinert G, Kapeller P, Schmidt R & Hartung HP. Histopathologic analysis of foci of signal loss on gradient-echo T2*-weighted MR images in patients with spontaneous intracerebral hemorrhage: evidence of microangiopathy-related microbleeds. AJNR Am J Neuroradiol 1999; 20:637-642.

Fisher CM. Lacunes: small, deep cerebral infarcts. Neurology 1965; 15:774-784.

Fisher CM. Pathological observations in hypertensive cerebral hemorrhage. J Neuropathol Exp Neurol 1971; 30:536-550.

Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Girón CG, Gordon L, Hourlier T, Hunt S,

Johnson N, Juettemann T, Kähäri AK, Keenan S, Kulesha E, Martin FJ, Maurel T, McLaren WM, Murphy DN, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ruffier M, Sheppard D, Taylor K, Thormann A, Trevanion SJ, Vullo A, Wilder SP, Wilson M, Zadissa A, Aken BL, Birney E, Cunningham F, Harrow J, Herrero J, Hubbard TJ, Kinsella R, Muffato M, Parker A, Spudich G, Yates A, Zerbino DR & Searle SM. Ensembl 2014. Nucleic Acids Res 2014; 42(Database issue):D749-755.

Fornage M, Debette S, Bis JC, Schmidt H, Ikram MA, Dufouil C, Sigurdsson S, Lumley T, DeStefano AL, Fazekas F, Vrooman HA, Shibata DK, Maillard P, Zijdenbos A, Smith AV, Gudnason H, de Boer R, Cushman M, Mazoyer B, Heiss G, Vernooij MW, Enzinger C, Glazer NL, Beiser A, Knopman DS, Cavalieri M, Niessen WJ, Harris TB, Petrovic K, Lopez OL, Au R, Lambert JC, Hofman A, Gottesman RF, Garcia M, Heckbert SR, Atwood LD, Catellier DJ, Uitterlinden AG, Yang Q, Smith NL, Aspelund T, Romero JR, Rice K, Taylor KD, Nalls MA, Rotter JI, Sharrett R, van Duijn CM, Amouyel P, Wolf PA, Gudnason V, van der Lugt A, Boerwinkle E, Psaty BM, Seshadri S, Tzourio C, Breteler MM, Mosley TH, Schmidt R, Longstreth WT, DeCarli C & Launer LJ. Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. Ann Neurol 2011; 69:928-939.

Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones

MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J & International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449:851-861.

Fryer JD, Simmons K, Parsadanian M, Bales KR, Paul SM, Sullivan PM & Holtzman DM. Human apolipoprotein E4 alters the amyloid-beta 40:42 ratio and promotes the formation of cerebral amyloid angiopathy in an amyloid precursor protein transgenic model. J Neurosci 2005; 25:2803-2810.

Fukutani Y, Cairns NJ, Rossor MN & Lantos PL. Cerebellar pathology in sporadic and familial Alzheimer's disease including APP 717 (Val→Ile) mutation cases: A morphometric investigation. J Neurol Sci 1997; 149:177-184.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ & Altshuler D. The structure of haplotype blocks in the human genome. Science 2002; 296:2225-2229.

Gesierich B, Opherk C, Rosand J, Gonik M, Malik R, Jouvent E, Hervé D, Adib-Samii P, Bevan S, Pianese L, Silvestri S, Dotti MT, De Stefano N, van der Grond J, Boon EM, Pescini F, Rost N, Pantoni L, Lesnik Oberstein SA, Federico A, Ragno M, Markus HS, Tournier-Lasserve E, Chabriat H, Dichgans M, Duering M & Ewers M. APOE ε2 is associated with white matter hyperintensity volume in CADASIL. J Cereb Blood Flow Metab 2015; doi: 10.1038/jcbfm.2015.85.

Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet 2011; 13:135-145.

Goldstein DB, Allen A, Keebler J, Margulies EH, Petrou S, Petrovski S & Sunyaev S. Sequencing studies in human genetics: design and interpretation. Nat Rev Genet 2013; 14:460-470.

Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ 2009; 339:b2680.

Greenberg SM, O'Donnell HC, Schaefer PW & Kraft E. MRI detection of new hemorrhages: potential marker of progression in cerebral amyloid angiopathy. Neurology 1999; 53:1135-1138.

Greenberg SM, Rebeck GW, Vonsattel JP, Gomez-Isla T & Hyman BT. Apolipoprotein E epsilon 4 and cerebral hemorrhage associated with amyloid angiopathy. Ann Neurol 1995; 38:254-259.

Greenberg SM, Shin Y, Grabowski TJ, Cooper GE, Rebeck GW, Iglesias S, Chapon F, Tournier-Lasserve E & Baron JC. Hemorrhagic stroke associated with the Iowa amyloid precursor protein mutation. Neurology 2003; 60:1020-1022.

Greenberg SM, Vernooij MW, Cordonnier C, Viswanathan A, Al-Shahi Salman R, Warach S, Launer LJ, Van Buchem MA, Breteler MM & Microbleed Study Group.

Cerebral microbleeds: a guide to detection and interpretation. Lancet Neurol 2009; 8:165-174.

Greenberg SM, Vonsattel JP, Segal AZ, Chiu RI, Clatworthy AE, Liao A, Hyman BT & Rebeck GW. Association of apolipoprotein E epsilon2 and vasculopathy in cerebral amyloid angiopathy. Neurology 1998; 50:961-965.

Grinberg LT & Thal DR. Vascular pathology in the aged human brain. Acta Neuropathol 2010; 119:277-290.

GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013; 45:580-585.

Gustafson L, Brun A, Englund E, Hagnell O, Nilsson K, Stensmyr M, Öhlin AK & Abrahamson M. A 50-year perspective of a family with chromosome-14-linked Alzheimer's disease. Hum Genet 1998; 102:253-257.

Haan J, Roos RA & Bakker E. No protective effect of apolipoprotein E epsilon 2 allele in Dutch hereditary cerebral amyloid angiopathy. Ann Neurol 1995; 37:282.

Haan J, Van Broeckhoven C, van Duijn CM, Voorhoeve E, van Harskamp F, van Swieten JC, Maat-Schieman ML, Roos RA & Bakker E. The apolipoprotein E epsilon 4 allele does not influence the clinical expression of the amyloid precursor protein gene codon 693 or 692 mutations. Ann Neurol 1994; 36:434-437.

Haffner C, Malik R & Dichgans M. Genetic factors in cerebral small vessel disease and their impact on stroke and dementia. J Cereb Blood Flow Metab 2015; doi: 10.1038/jcbfm.2015.71.

Haglund M, Passant U, Sjöbeck M, Ghebremedhin E & Englund E. Cerebral amyloid angiopathy and cortical microinfarcts as putative substrates of vascular dementia. Int J Geriatr Psychiatry 2006; 21:681-687.

Halliday G, Brooks W, Arthur H, Creasey H & Broe GA. Further evidence for an association between a mutation in the APP gene and Lewy body formation. Neurosci Lett 1997; 227:49-52.

Hamaguchi T, Okino S, Sodeyama N, Itoh Y, Takahashi A, Otomo E, Matsushita M, Mizusawa H & Yamada M. Association of a polymorphism of the transforming growth factor-β 1 gene with cerebral amyloid angiopathy. J Neurol Neurosurg Psychiatry 2005; 76:696-699.

Hanley DF 2009. Intraventricular hemorrhage: severity factor and treatment target in spontaneous intracerebral hemorrhage. Stroke 2009; 40:1533-1538.

Hara K, Shiga A, Fukutake T, Nozaki H, Miyashita A, Yokoseki A, Kawata H, Koyama A, Arima K, Takahashi T, Ikeda M, Shiota H, Tamura M, Shimoe Y, Hirayama M, Arisato T, Yanagawa S, Tanaka A, Nakano I, Ikeda S, Yoshida Y, Yamamoto T, Ikeuchi T, Kuwano R, Nishizawa M, Tsuji S & Onodera O. Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease. N Engl J Med 2009; 360:1729-1739.

Herzig MC, Van Nostrand WE & Jucker M. Mechanism of cerebral beta-amyloid angiopathy: murine and cellular models. Brain Pathol 2006; 16:40-54.

Higgins JP, Thompson SG, Deeks JJ & Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327:557-560.

Hirschhorn JN & Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005; 6:95-108.

Holtzman DM, Bales KR, Tenkova T, Fagan AM, Parsadanian M, Sartorius LJ, Mackey B, Olney J, McKeel D, Wozniak D & Paul SM. Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease. Proc Natl Acad Sci USA 2000; 97:2892-2897.

Holtzman DM, Bales KR, Wu S, Bhat P, Parsadanian M, Fagan AM, Chang LK, Sun Y & Paul SM. Expression of human apolipoprotein E reduces amyloid-beta deposition in a mouse model of Alzheimer's disease. J Clin Invest 1999; 103:R15-R21.

Honig LS, Kukull W & Mayeux R. Atherosclerosis and AD: analysis of data from the US National Alzheimer's Coordinating Center. Neurology 2005; 64:494-500.

Ioannidis JP. Genetic associations: false or true? Trends Mol Med 2003; 9:135-138.

Iwanowski P, Kozubski W & Losy J. Iowa-type hereditary cerebral amyloid angiopathy in a Polish family. J Neurol Sci 2015; 356:202-204.

Jackson C & Sudlow C. Are lacunar strokes really different? A systematic review of differences in risk factor profiles between lacunar and nonlacunar infarcts. Stroke 2005; 36:891-901.

Jackson CA, Hutchison A, Dennis MS, Wardlaw JM, Lindgren A, Norrving B, Anderson CS, Hankey GJ, Jamrozik K, Appelros P & Sudlow CL. Differing risk factor profiles of ischemic stroke subtypes: evidence for a distinct lacunar arteriopathy? Stroke 2010; 41:624-629.

Jackson CA & Sudlow CL 2006. Is hypertension a more frequent risk factor for deep than lobar supratentorial intracerebral haemorrhage? J Neurol Neurosurg Psychiatry 2006; 77:1244-1252.

Jeanne M, Labelle-Dumais C, Jorgensen J, Kauffman WB, Mancini GM, Favor J, Valant V, Greenberg SM, Rosand J & Gould DB. COL4A2 mutations impair COL4A1 and COL4A2 secretion and cause hemorrhagic stroke. Am J Hum Genet 2012; 90:91-101.

Jicha GA, Parisi JE, Dickson DW, Cha RH, Johnson KA, Smith GE, Boeve BF, Petersen RC & Knopman DS. Age and apoE associations with complex pathologic features in Alzheimer's disease. J Neurol Sci 2008; 273:34-39.

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ & de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008; 24:2938-2939.

Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, Stefansson H, Sulem P, Gudbjartsson D, Maloney J, Hoyte K, Gustafson A, Liu Y, Lu Y, Bhangale T, Graham RR, Huttenlocher J, Bjornsdottir G, Andreassen OA, Jönsson EG, Palotie A, Behrens TW, Magnusson OT, Kong A, Thorsteinsdottir U, Watts RJ & Stefansson K. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature 2012; 488:96-99.

Joutel A, Haddad I, Ratelade J & Nelson MT. Perturbations of the cerebrovascular matrisome: a convergent mechanism in small vessel disease of the brain? J Cereb Blood Flow Metab 2015; doi: 10.1038/jcbfm.2015.62.

Kalaria RN & Premkumar DRD. Apolipoprotein E genotype and cerebral amyloid angiopathy. Lancet 1995; 346:1424.

Kamp JA, Moursel LG, Haan J, Terwindt GM, Lesnik Oberstein SA, van Duinen SG & van Roon-Mom WM. Amyloid β in hereditary cerebral hemorrhage with amyloidosis-Dutch type. Rev Neurosci 2014; 25:641-651.

Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burtt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS, Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL & Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet 2009; 41:56-65.

Kavanagh D, Spitzer D, Kothari PH, Shaikh A, Liszewski MK, Richards A & Atkinson JP. New roles for the major human 3'-5' exonuclease TREX1 in human disease. Cell Cycle 2008; 7:1718-1725.

Kavvoura FK & Ioannidis JP. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Hum Genet 2008; 123:1-14.

Keable A, Fenna K, Yuen HM, Johnston DA, Smyth NR, Smith C, Al-Shahi Salman R, Samarasekera N, Nicoll JA, Attems J, Kalaria RN, Weller RO & Carare RO. Deposition of amyloid β in the walls of human leptomeningeal arteries in relation to perivascular drainage pathways in cerebral amyloid angiopathy. Biochim Biophys Acta 2015; doi: 10.1016/j.bbadis.2015.08.024.

Khoshnoodi J, Pedchenko V & Hudson BG. Mammalian collagen IV. Microsc Res Tech 2008; 71:357-370.

Klarenbeek P, van Oostenbrugge RJ, Rouhl RP, Knottnerus IL & Staals J. Ambulatory blood pressure in patients with lacunar stroke: association with total MRI burden of cerebral small vessel disease. Stroke 2013; 44:2995-2999.

Knudsen KA, Rosand J, Karluk D & Greenberg SM. Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. Neurology 2001; 56:537-539.

Kottner J, Audigé, L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, Roberts C, Shoukri M & Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011; 64:96-106.

Krishnamurthi RV, Feigin VL, Forouzanfar MH, Mensah GA, Connor M, Bennett DA, Moran AE, Sacco RL, Anderson LM, Truelsen T, O'Donnell M, Venketasubramanian N, Barker-Collo S, Lawes CM, Wang W, Shinohara Y, Witt E, Ezzati M, Naghavi M, Murray C & Global Burden of Diseases, Injuries, Risk Factors Study 2010 & GBD Stroke Experts Group. Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet Glob Health 2013; 1:e259-281.

Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, Matsushita T, Yamazaki K, Ohnishi Y, Saito S, Kitazono T, Ibayashi S, Sueishi K, Iida M, Nakamura Y & Kiyohara Y. A nonsynonymous SNP in PRKCH (protein kinase C eta) increases the risk of cerebral infarction. Nat Genet 2007; 39:212-217.

Kumar-Singh S, Cras P, Wang R, Kros JM, van Swieten J, Lübke U, Ceuterick C, Serneels S, Vennekens K, Timmermans JP, Van Marck E, Martin JJ, van Duijn CM & Van Broeckhoven C. Dense-core senile plaques in the Flemish variant of Alzheimer's disease are vasocentric. Am J Pathol 2002; 161:507-520.

Kölsch H, Larionov S, Dedeck O, Orantes M, Birkenmeier G, Griffin WS & Thal DR. Association of the glutathione S-transferase omega-1 Ala140Asp polymorphism with cerebrovascular atherosclerosis and plaque-associated interleukin-1 alpha expression. Stroke 2007; 38:2847-2850.

Kyzas PA, Loizou KT & Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. J Natl Cancer Inst 2005; 97:1043-1055.

Labovitz DL, Hauser WA & Sacco RL. Prevalence and predictors of early seizure and status epilepticus after first stroke. Neurology 2001; 57:200-206.

Lambert MP, Barlow AK, Chromy BA, Edwards C, Freed R, Liosatos M, Morgan TE, Rozovsky I, Trommer B, Viola KM, Wals P, Zhang C, Finch CE, Krafft GA & Klein WL. Diffusible, nonfibrillar ligands derived from Abeta1-42 are potent central nervous system neurotoxins. Proc Natl Acad Sci USA 1998; 95:6448-6453.

Lanfranconi S & Markus HS. COL4A1 mutations as a monogenic cause of cerebral small vessel disease: a systematic review. Stroke 2010; 41:e513-518.

Lashley T, Holton JL, Gray E, Kirkham K, O'Sullivan SS, Hilbig A, Wood NW, Lees AJ & Revesz T. Cortical alpha-synuclein load is associated with amyloid-beta plaque burden in a subset of Parkinson's disease patients. Acta Neuropathol 2008; 115:417-425.

Leclercq PD, Murray LS, Smith C, Graham DI, Nicoll JA & Gentleman SM. Cerebral amyloid angiopathy in traumatic brain injury: association with apolipoprotein E genotype. J Neurol Neurosurg Psychiatry 2005; 76:229-233.

Lee-Kirsch MA, Gong M, Chowdhury D, Senenko L, Engel K, Lee YA, de Silva U, Bailey SL, Witte T, Vyse TJ, Kere J, Pfeiffer C, Harvey S, Wong A, Koskenmies S,

Hummel O, Rohde K, Schmidt RE, Dominiczak AF, Gahr M, Hollis T, Perrino FW, Lieberman J & Hübner N. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nat Genet 2007; 39:1065-1067.

Lee SH, Bae HJ, Kwon SJ, Kim H, Kim YH, Yoon BW & Roh JK. Cerebral microbleeds are regionally associated with intracerebral hemorrhage. Neurology 2004; 62:72-76.

Lemere CA, Lopera F, Kosik KS, Lendon CL, Ossa J, Saido TC, Yamaguchi H, Ruiz A, Martinez A, Madrigal L, Hincapie L, Arango L JC, Anthony DC, Koo EH, Goate AM, Selkoe DJ & Arango V JC. The E280A presenilin 1 Alzheimer mutation produces increased Ab42 deposition and severe cerebellar pathology. Nat Med 1996; 2:1146-1150.

Lendon CL, Thaker U, Harris JM, McDonagh AM, Lambert JC, Chartier-Harlin MC, Iwatsubo T, Pickering-Brown SM & Mann DM. The angiotensin 1-converting enzyme insertion (I)/deletion (D) polymorphism does not influence the extent of amyloid or tau pathology in patients with sporadic Alzheimer's disease. Neurosci Lett 2002; 328:314-318.

Lewis H, Beher D, Cookson N, Oakley A, Piggott M, Morris CM, Jaros E, Perry R, Ince P, Kenny RA, Ballard CG, Shearman MS & Kalaria RN. Quantification of Alzheimer pathology in ageing and dementia: age-related accumulation of amyloid-beta(42) peptide in vascular dementia. Neuropathol Appl Neurobiol 2006; 32:103-118.

Li J & Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb) 2005; 95:221-227.

Linn J, Halpin A, Demaerel P, Ruhland J, Giese AD, Dichgans M, van Buchem MA, Bruckmann H & Greenberg SM. Prevalence of superficial siderosis in patients with cerebral amyloid angiopathy. Neurology 2010; 74:1346-1350.

Linn J, Herms J, Dichgans M, Brückmann H, Fesl G, Freilinger T & Wiesmann M. Subarachnoid hemosiderosis and superficial cortical hemosiderosis in cerebral amyloid angiopathy. AJNR Am J Neuroradiol 2008; 29:184-186.

Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N & STrenghtening the REporting of Genetic Association Studies. STrenghtening the REporting of Genetic Association Studies (STREGA) - an extension of the STROBE Statement. PLoS med 2009; 6:e22.

Love S, Chalmers K, Ince P, Esiri M, Attems J, Jellinger K, Yamada M, McCarron M, Minett T, Matthews F, Greenberg S, Mann D & Kehoe PG. Development, appraisal, validation and implementation of a consensus protocol for the assessment of cerebral amyloid angiopathy in post-mortem brain tissue. Am J Neurodegener Dis 2014; 3:19-32.

Love S, Nicoll JA, Hughes A & Wilcock GK. APOE and cerebral amyloid angiopathy in the elderly. Neuroreport 2003; 14:1535-1536.

Lovelock CE, Anslow P, Molyneux AJ, Byrne JV, Kuker W, Pretorius PM, Coull A & Rothwell PM. Substantial Observer Variability in the Differentiation Between Primary Intracerebral Hemorrhage and Hemorrhagic Transformation of Infarction on CT Brain Imaging. Stroke 2009; 40:3763-3767.

Mandybur TI. Cerebral amyloid angiopathy: the vascular pathology and complications. Journal of Neuropathology and Experimental Neurology 1986; 45:79-90.

Mann DMA, Iwatsubo T, Cairns NJ, Lantos PL, Nochlin D, Sumi SM, Bird TD, Poorkaj P, Hardy J, Hutton M, Prihar G, Crook R, Rossor MN & Haltia M. Amyloid beeta protein deposition in chromosome 14-linked Alzheimer´s disease: predominance of abeeta42(43). Ann Neurol 1996a; 40:149-156.

Mann DMA, Iwatsubo T, Ihara Y, Cairns NJ, Lantos PL, Bogdanovic N, Lannfelt L, Winblad B, Maat-Schieman MLC & Rossor MN. Predominant deposition of amyloid-beta42(43) in plaques in cases of Alzheimer's disease and hereditary cerebral hemorrhage associated with mutations in the amyloid precursor protein gene. Am J Pathol 1996b; 148:1257-1266.

Mann DMA, Iwatsubo T & Snowden JS. Atypical amyloid (Abeta) deposition in the cerebellum in Alzheimer's disease: an immunohistochemical study using end-specific Abeta monoclonal antibodies. Acta Neuropathol 1996c; 91:647-653.

Mann DMA, Pickering-Brown SM, Takeuchi A, Iwatsubo T & Members of the Familial Alzhemier's Disease Pathology Study Group. Amyloid angiopathy and variability in amyloid beta deposition is determined by mutation position in presenilin-1-linked Alzheimer's disease. Am J Pathol 2001; 158:2165-2175.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA & Visscher PM. Finding the missing heritability of complex diseases. Nature 2009; 461:747-753.

Marchini J & Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010; 11:499-511.

Martinez-Ramirez S, Pontes-Neto OM, Dumas AP, Auriel E, Halpin A, Quimby M, Gurol ME, Greenberg SM & Viswanathan A. Topography of dilated perivascular spaces in subjects from a memory clinic cohort. Neurology 2013; 80:1551-1556.

Masuda J, Tanaka K, Ueda K & Omae T. Autopsy study of incidence and distribution of cerebral amyloid angiopathy in Hisayama, Japan. Stroke 1988; 19:205-210.

Maxwell SS, Jackson CA, Paternoster L, Cordonnier C, Thijs V, Al-Shahi Salman R & Sudlow CLM. Genetic associations with brain microbleeds: Systematic review and meta-analyses. Neurology 2011; 77:158-167.

Mayda-Domaç F, Misirli H & Yilmaz M. Prognostic role of mean platelet volume and platelet count in ischemic and hemorrhagic stroke. J Stroke Cerebrovasc Dis 2010; 19:66-72.

McCarron MO & Nicoll JA. Apolipoprotein E genotype and cerebral amyloid angiopathy-related hemorrhage. Ann N Y Acad Sci 2000; 903:176-179.

McCarron MO, Nicoll JA, Stewart J, Ironside JW, Mann DM, Love S, Graham DI & Dewar D. The apolipoprotein E epsilon2 allele and the pathological features in cerebral amyloid angiopathy-related hemorrhage. J Neuropathol Exp Neurol 1999; 58:711-718.

McCarron MO, Stewart J, McCarron P, Love S, Vinters HV, Ironside JW, Mann DM, Graham DI & Nicoll JA. Association between interleukin-1A polymorphism and cerebral amyloid angiopathy-related hemorrhage. Stroke 2003; 34:e193-195.

Meretoja A, Strbian D, Putaala J, Curtze S, Haapaniemi E, Mustanoja S, Sairanen T, Satopää J, Silvennoinen H, Niemelä M, Kaste M & Tatlisumak T. SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. Stroke 2012; 43:2592-2597.

Mortimer JA, Snowdon DA & Markesbery WR. The effect of APOE-epsilon4 on dementia is mediated by Alzheimer neuropathology. Alzheimer Dis Assoc Disord 2009; 23:152-157.

Murray LS, Lu Y, Taggart A, Van Regemorter N, Vilain C, Abramowicz M, Kadler KE & Van Agtmael T. Chemical chaperone treatment reduces intracellular accumulation of mutant collagen IV and ameliorates the cellular phenotype of a COL4A2 mutation that causes haemorrhagic stroke. Hum Mol Genet 2014; 23:283-292.

Natté R, Maat-Schieman ML, Haan J, Bornebroek M, Roos RAC & van Duinen SG. Dementia in hereditary cerebral hemorrhage with amyloidosis-Dutch type is associated with cerebral amyloid angiopathy but is independent of plaques and neurofibrillary tangles. Ann Neurol 2001; 50:765-772.

Nelson PT, Abner EL, Schmitt FA, Kryscio RJ, Jicha GA, Smith CD, Davis DG, Poduska JW, Patel E, Mendiondo MS & Markesbery WR. Modeling the association between 43 different clinical and pathological variables and the severity of cognitive impairment in a large autopsy cohort of elderly persons. Brain Pathol 2010; 20:66-79.

Neuropathology Group MRC CFAS. Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). Lancet 2001; 357:169-175.

Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA & Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009; 461:272-276.

Nicoll JA, Savva GM, Stewart J, Matthews FE, Brayne C, Ince P & Medical Research Council Cognitive Function and Aging Study. Association between APOE genotype, neuropathology and dementia in the older population of England and Wales. Neuropathol Appl Neurobiol 2011; 37:285-294.

Nilsson OG, Lindgren A, Brandt L & Säveland H. Prediction of death in patients with primary intracerebral hemorrhage: a prospective study of a defined population. J Neurosurg 2002; 97:531-536.

NINDS Stroke Genetics Network (SiGN) & International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol 2015; doi: 10.1016/S1474-4422(15)00338-5.

Nochlin D, Bird TD, Nemens EJ, Ball MJ & Sumi SM. Amyloid angiopathy in a Volga German family with Alzheimer's disease and a presenilin-2 mutation (N141I). Ann Neurol 1998; 43:131-135.

Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 2004; 74:765-769.

Nyholt DR. Evaluation of Nyholt's procedure for multiple testing correction - author's reply. Hum Hered 2005; 60:61-62.

O'Donnell CJ, Kavousi M, Smith AV, Kardia SL, Feitosa MF, Hwang SJ, Sun YV, Province MA, Aspelund T, Dehghan A, Hoffmann U, Bielak LF, Zhang Q, Eiriksdottir G, van Duijn CM, Fox CS, de Andrade M, Kraja AT, Sigurdsson S, Elias-Smale SE, Murabito JM, Launer LJ, van der Lugt A, Kathiresan S, Krestin GP, Herrington DM, Howard TD, Liu Y, Post W, Mitchell BD, O'Connell JR, Shen H, Shuldiner AR, Altshuler D, Elosua R, Salomaa V, Schwartz SM, Siscovick DS, Voight BF, Bis JC, Glazer NL, Psaty BM, Boerwinkle E, Heiss G, Blankenberg S, Zeller T, Wild PS, Schnabel RB, Schillert A, Ziegler A, Münzel TF, White CC, Rotter JI, Nalls M, Oudkerk M, Johnson AD, Newman AB, Uitterlinden AG, Massaro JM, Cunningham J, Harris TB, Hofman A, Peyser PA, Borecki IB, Cupples LA, Gudnason V, Witteman JC & CARDIoGRAM Consortium. Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. Circulation 2011; 124:2855-2864.

Ogata J, Yamanishi H & Ishibashi-Ueda H. 2014. "Pathology of cerebral small vessel disease" in *Cerebral small vessel disease*, Pantoni, L & Gorelick, PB (eds.) 1 ed. Cambridge, United Kingdom: Cambridge University Press, 4-15.

Okazaki H, Reagan TJ & Campbell RJ. Clinicopathological studies of primary cerebral amyloid angiopathy. Mayo Clinic Proceedings 1979; 54:22-31.

Olichney JM, Hansen LA, Hofstetter CR, Grundman M, Katzman R & Thal LJ. Cerebral infarction in Alzheimer's disease is associated with severe amyloid angiopathy and hypertension. Arch Neurol 1995; 52:702-708.

Olichney JM, Hansen LA, Hofstetter CR, Lee JH, Katzman R & Thal LJ. Association between severe cerebral amyloid angiopathy and cerebrovascular lesions

in Alzheimer disease is not a spurious one attributable to apolipoprotein E4. Arch Neurol 2000; 57:869-874.

Opherk C & Dichgans M. 2014. "Hereditary cerebral small vessel disease and the role of genetic factors" in *Cerebral small vessel disease*, Panoni, L & Gorelick, PB (eds.) 1 ed. Cambridge, UK: Cambridge University Press, 82-91.

Oyama F, Shimada H, Oyama R & Ihara Y. Apolipoprotein E genotype, Alzheimer's pathologies and related gene expression in the aged population. Brain Res Mol Brain Res 1995; 29:92-98.

Pahnke J, Walker LC, Schroeder E, Vogelgesang S, Stausske D, Walther R & Warzok RW. Cerebral β-amyloid deposition is augmented by the -491AA promoter polymorphism in non-demented elderly individuals bearing the apolipoprotein E ε4 allele. Acta Neuropathol 2003; 105:25-29.

Palm F, Henschke N, Wolf J, Zimmer K, Safer A, Schröder RJ, Inselmann G, Brenke C, Becher H & Grau AJ. Intracerebral haemorrhage in a population-based stroke registry (LuSSt): incidence, aetiology, functional outcome and mortality. J Neurol 2013; 260:2541-2550.

Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol 2010; 9:689-701.

Pantoni L, Sarti C, Alafuzoff I, Jellinger K, Munoz DG, Ogata J & Palumbo V. Postmortem examination of vascular lesions in cognitive impairment: a survey among neuropathological services. Stroke 2006; 37:1005-1009.

Paternoster L, Chen W & Sudlow CLM. Genetic determinants of white matter hyperintensities on brain scans: a systematic assessment of 19 candidate gene polymorphisms in 46 studies in 19,000 subjects. Stroke 2009; 40:2020-2026.

Peila R, Yucesoy B, White LR, Johnson V, Kashon ML, Wu K, Petrovitch H, Luster M & Launer LJ. A TGF-β1 polymorphism association with dementia and neuropathologies: the HAAS. Neurobiol Aging 2007; 28:1367-1373.

Pezzini A & Padovani A. Cerebral amyloid angiopathy-related hemorrhages. Neurological Sciences 2008; 29:Suppl 2:S260-263.

Petitti DB. Approaches to heterogeneity in meta-analysis. Stat Med 2001; 20:3625-3633.

Petrovitch H, Ross GW, He Q, Uyehara-Lock J, Markesbery W, Davis D, Nelson J, Masaki K, Launer L & White LR. Characterization of Japanese-American men with a single neocortical AD lesion type. Neurobiol Aging 2008; 29:1448-1455.

Peuralinna T, Tanskanen M, Mäkelä M, Polvikoski T, Paetau A, Kalimo H, Sulkava R, Hardy J, Lai SL, Arepalli S, Hernandez D, Traynor BJ, Singleton A, Tienari PJ & Myllykangas L. APOE and AβPP gene variation in cortical and cerebrovascular amyloid-β pathology and Alzheimer's Disease: a population-based analysis. J Alzheimers Dis 2011; 26:377-385.

Pfeifer LA, White LR, Ross GW, Petrovitch H & Launer LJ. Cerebral amyloid angiopathy and cognitive function: the HAAS autopsy study. Neurology 2002; 58:1629-1634.

Pollock H, Hutchings M, Weller RO & Zhang ET. Perivascular spaces in the basal ganglia of the human brain: their relationship to lacunes. J Anat 1997; 191(Pt 3):337-346.

Premkumar DR, Cohen DL, Hedera P, Friedland RP & Kalaria RN. Apolipoprotein E-epsilon4 alleles in cerebral amyloid angiopathy and cerebrovascular pathology associated with Alzheimer's disease. Am J Pathol 1996; 148:2083-2095.

Rannikmäe K, Davies G, Thomson PA, Bevan S, Devan WJ, Falcone GJ, Traylor M, Anderson CD, Battey TW, Radmanesh F, Deka R, Woo JG, Martin LJ, Jimenez-Conde J, Selim M, Brown DL, Silliman SL, Kidwell CS, Montaner J, Langefeld CD, Slowik A, Hansen BM, Lindgren AG, Meschia JF, Fornage M, Bis JC, Debette S, Ikram MA, Longstreth WT, Schmidt R, Zhang CR, Yang Q, Sharma P, Kittner SJ, Mitchell BD, Holliday EG, Levi CR, Attia J, Rothwell PM, Poole DL, Boncoraglio GB, Psaty BM, Malik R, Rost N, Worrall BB, Dichgans M, Van Agtmael T, Woo D, Markus HS, Seshadri S, Rosand J, Sudlow CL, METASTROKE Consortium, the CHARGE WMH Group, the ISGC ICH GWAS Study Collaboration, the WMH in Ischemic Stroke GWAS Study Collaboration & and the International Stroke Genetics Consortium. Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease. Neurology 2015; 84:918-926.

Rannikmäe K, Kalaria RN, Greenberg SM, Chui HC, Schmitt FA, Samarasekera N, Al-Shahi Salman R & Sudlow CLM. APOE associations with severe CAA-associated vasculpathic changes: collaborative meta-analysis. J Neurol Neurosurg Psychiatry 2014; 85:300-305.

Rannikmäe K, Samarasekera N, Martînez-Gonzâlez NA, Al-Shahi Salman R & Sudlow C. Genetics of cerebral amyloid angiopathy: systematic review and meta-analysis. J Neurol Neurosurg Psychiatry 2013; 84:901-908.

Revesz T, Holton JL, Lashley T, Plant G, Frangione B, Rostagno A & Ghiso J. Genetics and molecular pathogenesis of sporadic and hereditary cerebral amyloid angiopathies. Acta Neuropathol 2009; 118:115-130.

Risch N & Merikangas K. The future of genetic studies of complex human diseases. Science 1996; 273:1516-1517.

Roher AE, Kuo YM, Esh C, Knebel C, Weiss N, Kalback W, Luehrs DC, Childress JL, Beach TG, Weller RO & Kokjohn TA. Cortical and leptomeningeal cerebrovascular amyloid and white matter pathology in Alzheimer's disease. Mol Med 2003; 9:112-122.

Roks G, Van Harskamp F, De Koning I, Cruts M, De Jonghe C, Kumar-Singh S, Tibben A, Tanghe H, Niermeijer MF, Hofman A, Van Swieten JC, Van Broeckhoven C & van Duijn CM. Presentation of amyloidosis in carriers of the codon 692 mutation in the amyloid precursor protein gene (APP692). Brain 2000; 123:2130-2140.

Rosenberg CK, Pericak-Vance MA, Saunders AM, Gilbert JR, Gaskell PC & Hulette CM. Lewy body and Alzheimer pathology in a family with the amyloid-beta precursor protein APP717 gene mutation. Acta Neuropathol 2000; 100:145-152.

Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull 1979; 86:638-641.

Rost NS, Rahman RM, Biffi A, Smith EE, Kanakis A, Fitzpatrick K, Lima F, Worrall BB, Meschia JF, Brown RD, Brott TG, Sorensen AG, Greenberg SM, Furie KL & Rosand J. White matter hyperintensity volume is increased in small vessel stroke subtypes. Neurology 2010; 75:1670-1677.

Rost NS, Smith EE, Chang Y, Snider RW, Chanderraj R, Schwab K, FitzMaurice E, Wendell L, Goldstein JN, Greenberg SM & Rosand J. Prediction of Functional Outcome in Patients with Primary Intracerebral Hemorrhage: The FUNC Score. Stroke 2008; 39:2304-2309.

Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerrière A, Vital A, Dumanchin C, Feuillette S, Brice A, Vercelletto M, Dubas F, Frebourg T & Campion D. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet 2006; 38:24-26.

Ruigrok YM, Rinkel GJ, van't Slot R, Wolfs M, Tang S & Wijmenga C. Evidence in favor of the contribution of genes involved in the maintenance of the extracellular matrix of the arterial wall to the development of intracranial aneurysms. Hum Mol Genet 2006; 15:3361-3368.

Rutten-Jacobs LC, Traylor M, Adib-Samii P, Thijs V, Sudlow C, Rothwell PM, Boncoraglio G, Dichgans M, Bevan S, Meschia J, Levi C, Rost NS, Rosand J, Hassan A & Markus HS. Common NOTCH3 Variants and Cerebral Small-Vessel Disease. Stroke 2015; 46:1482-1487.

Ryan NS, Biessels GJ, Kim L, Nicholas JM, Barber PA, Walsh P, Gami P, Morris HR, Bastos-Leite AJ, Schott JM, Beck J, Mead S, Chavez-Gutierrez L, de Strooper B, Rossor MN, Revesz T, Lashley T & Fox NC. Genetic determinants of white matter hyperintensities and amyloid angiopathy in familial Alzheimer's disease. Neurobiol Aging 2015; 36:3140-3151.

Salyakina D, Seaman SR, Browning BL, Dudbridge F & Muller-Myhsok B. Evaluation of Nyholt's procedure for multiple testing correction. Hum Hered 2005; 60:19-25.

Samarasekera N, Smith C & Al-Shahi Salman R. The association between cerebral amyloid angiopathy and intracerebral haemorrhage: systematic review and meta-analysis. J Neurol Neurosurg Psychiatry 2012; 83:275-281.

Schilling S, DeStefano AL, Sachdev PS, Choi SH, Mather KA, DeCarli CD, Wen W, Høgh P, Raz N, Au R, Beiser A, Wolf PA, Romero JR, Zhu YC, Lunetta KL, Farrer L, Dufouil C, Kuller LH, Mazoyer B, Seshadri S, Tzourio C & Debette S. APOE genotype and MRI markers of cerebrovascular disease: systematic review and meta-analysis. Neurology 2013; 81:292-300.

Schmidt H, Zeginigg M, Wiltgen M, Freudenberger P, Petrovic K, Cavalieri M, Gider P, Enzinger C, Fornage M, Debette S, Rotter JI, Ikram MA, Launer LJ, Schmidt R & CHARGE consortium Neurology working group. Genetic variants of the NOTCH3 gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel disease. Brain 2011; 134:3384-3397.

Schneider JA, Bienias JL, Wilson RS, Berry-Kravis E, Evans DA & Bennett DA. The apolipoprotein E epsilon4 allele increases the odds of chronic cerebral infarction detected at autopsy in older persons. Stroke 2005; 36:954-959.

Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ & CARDIoGRAM Consortium. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011; 43:333-338.

Serizawa M, Nabika T, Ochiai Y, Takahashi K, Yamaguchi S, Makaya M, Kobayashi S & Kato N. Association between PRKCH gene polymorphisms and subcortical silent brain infarction. Atherosclerosis 2008; 199:340-345.

Shen AY, Yao JF, Brar SS, Jorgensen MB & Chen W. Racial/ethnic differences in the risk of intracranial hemorrhage among patients with atrial fibrillation. J Am Coll Cardiol 2007; 50:309-315.

Shi J, Tian J, Pritchard A, Lendon C, Lambert JC, Iwatsubo T & Mann DM. A 3'-UTR polymorphism in the oxidized LDL receptor 1 gene increases Aβ40 load as

cerebral amyloid angiopathy in Alzheimer's disease. Acta Neuropathol 2006; 111:15-20.

Sim J & Wright CC 2005. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. Phys Ther 2005; 85:257-268.

Singleton AB, Hall R, Ballard CG, Perry RH, Xuereb JH, Rubinsztein DC, Tysoe C, Matthews P, Cordell B, Kumar-Singh S, De Jonghe C, Cruts M, Van Broeckhoven C & Morris CM. Pathology of early-onset Alzheimer's disease cases bearing the Thr113-114ins presenilin-1 mutation. Brain 2000; 123:2467-2474.

Smith EE, Nandigam KR, Chen YW, Jeng J, Salat D, Halpin A, Frosch M, Wendell L, Fazen L, Rosand J, Viswanathan A & Greenberg SM. MRI markers of small vessel disease in lobar and deep hemispheric intracerebral hemorrhage. Stroke 2010; 41:1933-1938.

Smith EE, Schneider JA, Wardlaw JM & Greenberg SM. Cerebral microinfarcts: the invisible lesions. Lancet Neurol 2012; 11:272-282.

Sonnen JA, Larson EB, Walker RL, Haneuse S, Crane PK, Gray SL, Breitner JC & Montine TJ. Nonsteroidal anti-inflammatory drugs are associated with increased neuritic plaques. Neurology 2010; 75:1203-1210.

Staals J, Makin SD, Doubal FN, Dennis MS & Wardlaw JM. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. Neurology 2014; 83:1228-1234.

Stopa EG, Butala P, Salloway S, Johanson CE, Gonzalez L, Tavares R, Hovanesian V, Hulette CM, Vitek MP & Cohen RA. Cerebral cortical arteriolar angiopathy, vascular beta-amyloid, smooth muscle actin, Braak stage, and APOE genotype. Stroke 2008; 39:814-821.

Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA & Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; 283:2008-2012.

Sudlow C, Martínez González N, Kim J & Clark C. Does Apolipoprotein E genotype influence the risk of ischemic stroke, intracerebral hemorrhage, or subarachnoid hemorrhage? Systematic review and meta-analyses of 31 studies among 5961 cases and 17,965 controls. Stroke 2006; 37:364-370.

Tabor HK, Risch NJ & Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002; 3:391-397.

Takahashi O, Cook EF, Nakamura T, Saito J, Ikawa F & Fukui T. Risk stratification for in-hospital mortality in spontaneous intracerebral haemorrhage: a Classification and Regression Tree Analysis. QJM 2006; 99:743-750.

Takeda M, Martínez R, Kudo T, Tanaka T, Okochi M, Tagami S, Morihara T, Hashimoto R & Cacabelos R. Apolipoprotein E and central nervous system disorders: reviews of clinical findings. Psychiatry Clin Neurosci 2010; 64:592-607.

Tanskanen M, Lindsberg PJ, Tienari PJ, Polvikoski T, Sulkava R, Verkkoniemi A, Rastas S, Paetau A & Kiuru-Enari S. Cerebral amyloid angiopathy in a 95+ cohort: complement activation and apolipoprotein E (ApoE) genotype. Neuropathol Appl Neurobiol 2005; 31:589-599.

Tarasov KV, Sanna S, Scuteri A, Strait JB, Orrù M, Parsa A, Lin PI, Maschio A, Lai S, Piras MG, Masala M, Tanaka T, Post W, O'Connell JR, Schlessinger D, Cao A, Nagaraja R, Mitchell BD, Abecasis GR, Shuldiner AR, Uda M, Lakatta EG & Najjar SS. COL4A1 is associated with arterial stiffness by genome-wide association scan. Circ Cardiovasc Genet 2009; 2:151-158.

Thal DR, Ghebremedhin ER, U., Yamaguchi H, Del Tredici K & Braak H. Two types of sporadic cerebral amyloid angiopathy. J Neuropathol Exp Neurol 2002; 61:282-293.

Thal DR, Larionov S, Abramowski D, Wiederhold KH, Van Dooren T, Yamaguchi H, Haass C, Van Leuven F, Staufenbiel M & Capetillo-Zarate E. Occurrence and co-localization of amyloid beta-protein and apolipoprotein E in perivascular drainage channels of wild-type and APP-transgenic mice. Neurobiol Aging 2007; 28:1221-1230.

Thal DR, Papassotiropoulos A, Saido TC, Griffin WS, Mrak RE, Kölsch H, Tredici KD, Attems J & Ghebremedhin E. Capillary cerebral amyloid angiopathy identifies a distinct APOE ε4-associated subtype of sporadic Alzheimer's disease. Acta Neuropathol 2010; 120:169-183.

Thanprasertsuk S, Martinez-Ramirez S, Pontes-Neto OM, Ni J, Ayres A, Reed A, Swords K, Gurol ME, Greenberg SM & Viswanathan A. Posterior white matter disease distribution as a predictor of amyloid angiopathy. Neurology 2014; 83:794-800.

The Cochrane Collaboration 2011. Cochrane Handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. Available from *www.cochrane-handbook.org.*

Thomas AJ, Morris CM, Ferrier IN & Kalaria RN. Distribution of amyloid beta 42 in relation to the cerebral microvasculature in an elderly cohort with Alzheimer's disease. Ann N Y Acad Sci 2000; 903:83-88.

Thompson SG & Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002; 21:1559-1573.

Tiraboschi P, Sabbagh MN, Hansen LA, Salmon DP, Merdes A, Gamst A, Masliah E, Alford M, Thal LJ & Corey-Bloom J. Alzheimer disease without neocortical neurofibrillary tangles: "a second look". Neurology 2004; 62:1141-1147.

Tomidokoro Y, Rostagno A, Neubert TA, Lu Y, Rebeck GW, Frangione B, Greenberg SM & Ghiso J. Iowa variant of Familial Alzheimer´s Disease. Accumulation of posttranslationally modified AbetaD23N in parenchymal and cerebrovascular amyloid deposits. Am J Pathol 2010; 176:1841-1854.

Traylor M, Bevan S, Baron JC, Hassan A, Lewis CM & Markus HS. Genetic Architecture of Lacunar Stroke. Stroke 2015; 46:2407-2412.

Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, Fornage M, Ikram MA, Malik R, Bevan S, Thorsteinsdottir U, Nalls MA, Longstreth W, Wiggins KL, Yadav S, Parati EA, Destefano AL, Worrall BB, Kittner SJ, Khan MS, Reiner AP, Helgadottir A, Achterberg S, Fernandez-Cadenas I, Abboud S, Schmidt R, Walters M, Chen WM, Ringelstein EB, O'Donnell M, Ho WK, Pera J, Lemmens R, Norrving B, Higgins P, Benn M, Sale M, Kuhlenbäumer G, Doney AS, Vicente AM, Delavaran H, Algra A, Davies G, Oliveira SA, Palmer CN, Deary I, Schmidt H, Pandolfo M, Montaner J, Carty C, de Bakker PI, Kostulas K, Ferro JM, van Zuydam NR, Valdimarsson E, Nordestgaard BG, Lindgren A, Thijs V, Slowik A, Saleheen D, Paré G, Berger K, Thorleifsson G, Hofman A, Mosley TH, Mitchell BD, Furie K, Clarke R, Levi C, Seshadri S, Gschwendtner A, Boncoraglio GB, Sharma P, Bis JC, Gretarsdottir S, Psaty BM, Rothwell PM, Rosand J, Meschia JF, Stefansson K, Dichgans M, Markus HS, Australian Stroke Genetics Collaborative, Wellcome Trust Case Control Consortium 2 (WTCCC2) & International Stroke Genetics Consortium. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol 2012; 11:951-962.

Traylor M, Zhang CR, Adib-Samii P, Devan WJ, Parsons OE, Lanfranconi S, Gregory S, Cloonan L, Falcone GJ, Radmanesh F, Fitzpatrick K, Kanakis A, Barrick TR, Moynihan B, Lewis CM, Boncoraglio GB, Lemmens R, Thijs V, Sudlow C, Wardlaw J, Rothwell PM, Meschia JF, Worrall BB, Levi C, Bevan S, Furie KL, Dichgans M, Rosand J, Markus HS, Rost N & the International Stroke Genetics Consortium. Genome-wide meta-analysis of cerebral white matter hyperintensities in patients with stroke. Neurology 2016; 86:146-153.

Tzourio C, Arima H, Harrap S, Anderson C, Godin O, Woodward M, Neal B, Bousser MG, Chalmers J, Cambien F & MacMahon S. APOE genotype, ethnicity, and the risk of cerebral hemorrhage. Neurology 2008; 70:1322-1328.

Vahedi K & Alamowitch S. Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease. Curr Opin Neurol 2011; 24:63-68.

Valant V, Keenan BT, Anderson CD, Shulman JM, Devan WJ, Ayres AM, Schwab K, Goldstein JN, Viswanathan A, Greenberg SM, Bennett DA, De Jager PL, Rosand J, Biffi A & the Alzheimer's Disease Neuroimaging Initiative (ADNI). TOMM40 in cerebral amyloid angiopathy related intracerebral hemorrhage: comparative genetic analysis with Alzheimer's disease. Transl Stroke Res 2012; 3:S102-S112.

Walker LC, Pahnke J, Madauss M, Vogelgesang S, Pahnke A, Herbst EW, Stausske D, Walther R, Kessler C & Warzok RW. Apolipoprotein E4 promotes the early deposition of Abeta42 and then Abeta40 in the elderly. Acta Neuropathol 2000; 100:36-42.

Van Agtmael T, Bailey MA, Schlötzer-Schrehardt U, Craigie E, Jackson IJ, Brownstein DG, Megson IL & Mullins JJ. Col4a1 mutation in mice causes defects in

vascular function and low blood pressure associated with reduced red blood cell volume. Hum Mol Genet 2010; 19:1119-1128.

Van Agtmael T & Bruckner-Tuderman L. Basement membranes and human disease. Cell Tissue Res 2010; 339:167-188.

Van Agtmael T, Schlötzer-Schrehardt U, McKie L, Brownstein DG, Lee AW, Cross SH, Sado Y, Mullins JJ, Pöschl E & Jackson IJ. Dominant mutations of Col4a1 result in basement membrane defects which lead to anterior segment dysgenesis and glomerulopathy. Hum Mol Genet 2005; 14:3161-3168.

van Horssen J, de Jong D, de Waal RM, Maass C, Otte-Holler I, Kremer B, Verbeek MM & Wesseling P. Cerebral amyloid angiopathy with severe secondary vascular pathology: a histopathological study. Dement Geriatr Cogn Disord 2005; 20:321-330.

van Swieten JC, van den Hout JH, van Ketel BA, Hijdra A, Wokke JH & van Gijn J. Periventricular lesions in the white matter on magnetic resonance imaging in the elderly. A morphometric correlation with arteriolosclerosis and dilated perivascular spaces. Brain 1991; 114(Pt 2):761-774.

Ward LD & Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012; 40:D930-934.

Wardlaw JM & Pantoni L. 2014. "Sporadic cerebral small vessel disease: pathogenic aspects." in *Cerebral small vessel disease.* Pantoni, L & Gorelick, PB (eds.) 1 ed. Cambridge, United Kingdom: Cambridge University Press, 52-63.

Wardlaw JM, Smith C & Dichgans M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. Lancet Neurol 2013a; 12:483-497.

Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, Lindley RI, O'Brien JT, Barkhof F, Benavente OR, Black SE, Brayne C, Breteler M, Chabriat H, Decarli C, de Leeuw FE, Doubal F, Duering M, Fox NC, Greenberg S, Hachinski V, Kilimann I, Mok V, Oostenbrugge R, Pantoni L, Speck O, Stephan BC, Teipel S, Viswanathan A, Werring D, Chen C, Smith C, van Buchem M, Norrving B, Gorelick PB, Dichgans M & STandards for ReportIng Vascular changes on nEuroimaging (STRIVE v1). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol 2013b; 12:822-838.

Warlow C, van Gijn J, Dennis M, Wardlaw J, Bamford J, Hankey G, Sandercock P, Rinkel G, Langhorne P, Sudlow C & Rothwell P. 2008. "What caused this intracerebral haemorrhage?" in *Stroke: practical management.* 3rd ed. Blackwell Publishing. 411-456.

Weller RO, Subash M, Preston SD, Mazanti I & Carare RO. Perivascular drainage of amyloid-beta peptides from the brain and its failure in cerebral amyloid angiopathy and Alzheimer's disease. Brain Pathol 2008; 18:253-266.

Weng YC, Sonni A, Labelle-Dumais C, de Leau M, Kauffman WB, Jeanne M, Biffi A, Greenberg SM, Rosand J & Gould DB. COL4A1 mutations in patients with sporadic late-onset intracerebral hemorrhage. Ann Neurol 2012; 71:470-477.

Verbeek MM, Van Nostrand WE, Otte-Höller I, Wesseling P & De Waal RM. Amyloid-beta-induced degeneration of human brain pericytes is dependent on the apolipoprotein E genotype. Ann N Y Acad Sci 2000; 903:187-199.

Verghese PB, Castellano JM & Holtzman DM. Apolipoprotein E in Alzheimer's disease and other neurological disorders. Lancet Neurol 2011; 10:241-252.

Verhaaren BF, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, Beecham AH, Rajan KB, Lopez LM, Barral S, van Buchem MA, van der Grond J, Smith AV, Hegenscheid K, Aggarwal NT, de Andrade M, Atkinson EJ, Beekman M, Beiser AS, Blanton SH, Boerwinkle E, Brickman AM, Bryan RN, Chauhan G, Chen CP, Chouraki V, de Craen AJ, Crivello F, Deary IJ, Deelen J, De Jager PL, Dufouil C, Elkind MS, Evans DA, Freudenberger P, Gottesman RF, Guðnason V, Habes M, Heckbert SR, Heiss G, Hilal S, Hofer E, Hofman A, Ibrahim-Verbaas CA, Knopman DS, Lewis CE, Liao J, Liewald DC, Luciano M, van der Lugt A, Martinez OO, Mayeux R, Mazoyer B, Nalls M, Nauck M, Niessen WJ, Oostra BA, Psaty BM, Rice KM, Rotter JI, von Sarnowski B, Schmidt H, Schreiner PJ, Schuur M, Sidney SS, Sigurdsson S, Slagboom PE, Stott DJ, van Swieten JC, Teumer A, Töglhofer AM, Traylor M, Trompet S, Turner ST, Tzourio C, Uh HW, Uitterlinden AG, Vernooij MW, Wang JJ, Wong TY, Wardlaw JM, Windham BG, Wittfeld K, Wolf C, Wright CB, Yang Q, Zhao W, Zijdenbos A, Jukema JW, Sacco RL, Kardia SL, Amouyel P, Mosley TH, Longstreth WT, DeCarli CC, van Duijn CM, Schmidt R, Launer LJ, Grabe HJ, Seshadri SS, Ikram MA & Fornage M. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. Circ Cardiovasc Genet 2015; 8:398-409.

Wermer MJH, Rinkel GJE, van Rooij WJJ, Witkamp TD, Ziedses des Plantes BG & Algra A 2002. Interobserver agreement in the assessment of lobar versus deep location of intracerebral haematomas on CT. J Neuroradiol 2002; 29:271-274.

Vernooij MW, van der Lugt A, Ikram MA, Wielopolski PA, Niessen WJ, Hofman A, Krestin GP & Breteler MM. Prevalence and risk factors of cerebral microbleeds: the Rotterdam Scan Study. Neurology 2008; 70:1208-1214.

Viera AJ & Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005; 37:360-363.

Wijman CA, Venkatasubramanian C, Bruins S, Fischbein N & Schwartz N. Utility of Early MRI in the Diagnosis and Management of Acute Spontaneous Intracerebral Hemorrhage. Cerebrovasc Dis 2010; 30:456-463.

Willer CJ, Li Y & Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26:2190-2191.

Vinters HV, Natté, R, Maat-Schieman MLC, Van Duinen SG, Hegeman-Kleinn I, Welling-Graafland C, Haan J & Roos RAC. Secondary microvascular degeneration

in amyloid angiopathy of patients with hereditary cerebral hemorrhage with amyloidosis, Dutch type (HCHWA-D). Acta Neuropathol 1998; 95:235-244.

Vonsattel JP, Myers RH, Hedley-Whyte ET, Ropper AH, Bird ED & Richardson EPJ. Cerebral amyloid angiopathy without and with cerebral hemorrhages: a comparative histological study. Ann Neurol 1991; 30:637-649.

Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, Anderson CD, Brouwers HB, Valant V, Battey TW, Radmanesh F, Raffeld MR, Baedorf-Kassis S, Deka R, Woo JG, Martin LJ, Haverbusch M, Moomaw CJ, Sun G, Broderick JP, Flaherty ML, Martini SR, Kleindorfer DO, Kissela B, Comeau ME, Jagiella JM, Schmidt H, Freudenberger P, Pichler A, Enzinger C, Hansen BM, Norrving B, Jimenez-Conde J, Giralt-Steinhauer E, Elosua R, Cuadrado-Godia E, Soriano C, Roquer J, Kraft P, Ayres AM, Schwab K, McCauley JL, Pera J, Urbanik A, Rost NS, Goldstein JN, Viswanathan A, Stögerer EM, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Montaner J, Fernandez-Cadenas I, Delgado P, Malik R, Dichgans M, Greenberg SM, Rothwell PM, Lindgren A, Slowik A, Schmidt R, Langefeld CD, Rosand J & the International Stroke Genetics Consortium. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J Hum Genet 2014; 94:511-521.

Woo D, Sauerbeck LR, Kissela BM, Khoury JC, Szaflarski JP, Gebel J, Shukla R, Pancioli AM, Jauch EC, Menon AG, Deka R, Carrozzella JA, Moomaw CJ, Fontaine RN & Broderick JP. Genetic and environmental risk factors for intracerebral hemorrhage: preliminary results of a population-based study. Stroke 2002; 33:1190-1195.

Wood SJ, Chan W & Wetzel R. Seeding of A beta fibril formation is inhibited by all three isotypes of apolipoprotein E. Biochemistry 1996; 35:12623-12628.

Woodfield R, Rannikmäe K, Grant I & Sudlow C. Reliability and feasibility of ischemic stroke classification systems for large epidemiological studies. Cerebrovasc Dis 2013; 35:(Supplement 3):P159.

Wu L, Shen Y, Liu X, Ma X, Xi B, Mi J, Lindpaintner K, Tan X & Wang X. The 1425G/A SNP in PRKCH is associated with ischemic stroke and cerebral hemorrhage in a Chinese population. Stroke 2009; 40:2973-2976.

Yamada M. Risk factors for cerebral amyloid angiopathy in the elderly. Ann N Y Acad Sci 2002; 977:37-44.

Yamada M, Sodeyama N, Itoh Y, Otomo E, Matsushita M & Mizusawa H. No association of paraoxonase genotype or atherosclerosis with cerebral amyloid angiopathy. Stroke 2002; 33:896-900.

Yamada M, Sodeyama N, Itoh Y, Suematsu N, Otomo E, Matsushita M & Mizusawa H. Association of presenilin-1 polymorphism with cerebral amyloid angiopathy in the elderly. Stroke 1997; 28:2219-2221.

Yamada M, Sodeyama N, Itoh Y, Suematsu N, Otomo E, Matsushita M & Mizusawa H. Association of α1 antichymotrypsin polymorphism with cerebral amyloid angiopathy. Ann Neurol 1998a; 44:129-131.

Yamada M, Sodeyama N, Itoh Y, Suematsu N, Otomo E, Matsushita M & Mizusawa H. Butyrylcholinesterase K variant and cerebral amyloid angiopathy. Stroke 1998b; 29:2488-2490.

Yamada M, Sodeyama N, Itoh Y, Suematsu N, Otomo E, Matsushita M & Mizusawa H. A deletion polymorphism of alpha2 macroglobulin gene and cerebral amyloid angiopathy. Stroke 1999; 30:2277-2279.

Yamada M, Sodeyama N, Itoh Y, Takahashi A, Otomo E, Matsushita M & Mizusawa H. Association of neprilysin polymorphism with cerebral amyloid angiopathy. J Neurol Neurosurg Psychiatry 2003; 74:749-751.

Yamada Y, Kato K, Oguri M, Fujimaki T, Yokoi K, Matsuo H, Watanabe S, Metoki N, Yoshida H, Satoh K, Ichihara S, Aoyagi Y, Yasunaga A, Park H, Tanaka M & Nozawa Y. Genetic risk for myocardial infarction determined by polymorphisms of candidate genes in a Japanese population. J Med Genet 2008; 45:216-221.

Yamaguchi H, Sugihara S, Ogawa A, Oshima N & Ihara Y. Alzheimer beta amyloid deposition enhanced by ApoE epsilon4 gene precedes neurofibrillary pathology in the frontal association cortex of nondemented senior subjects. J Neuropathol Exp Neurol 2001; 60:731-739.

Yasuda M, Maeda S, Kawamata T, Tamaoka A, Yamamoto Y, Kuroda S, Maeda K & Tanaka C. Novel presenilin-1 mutation with widespread cortical amyloid deposition but limited cerebral amyloid angiopathy. Journal Neurol Neurosurg Psychiatry 2000; 68:220-223.

Yeh S-Y, Tang S-C, Tsai L-K & Jeng J-S. Pathogenetical Subtypes of Recurrent Intracerebral Hemorrhage. Stroke 2014; 45:2636-2642.

Yip AG, McKee AC, Green RC, Wells J, Young H, Cupples LA & Farrer LA. APOE, vascular pathology, and the AD brain. Neurology 2005; 65:259-265.

Yokota O, Terada S, Ishizu H, Ujike H, Ishihara T, Nakashima H, Yasuda M, Kitamura Y, Uéda K, Checler F & Kuroda S. NACP/alpha-synuclein, NAC, and beta-amyloid pathology of familial Alzheimer's disease with the E184D presenilin-1 mutation: A clinicopathological study of two autopsy cases. Acta Neuropathol 2002; 104:637-648.

Yu L, Boyle PA, Nag S, Leurgans S, Buchman AS, Wilson RS, Arvanitakis Z, Farfel JM, De Jager PL, Bennett DA & Schneider JA. APOE and cerebral amyloid angiopathy in community-dwelling older persons. Neurobiol Aging 2015; 36:2946-2953.

Zahuranec DB, Sánchez BN, Brown DL, Wing JJ, Smith MA, Garcia NM, Meurer WJ, Morgenstern LB & Lisabeth LD. Computed Tomography Findings for Intracerebral Hemorrhage Have Little Incremental Impact on Post-Stroke Mortality Prediction Model Performance. Cerebrovasc Dis 2012; 34:86-92.

Zarow C, Zaias B, Lyness SA & Chui H. Cerebral amyloid angiopathy in Alzheimer disease is associated with apolipoprotein E4 and cortical neuron loss. Alzheimer Dis Assoc Disord 1999; 13:1-8.

Zhang-Nunes SX, Maat-Schieman ML, van Duinen SG, Roos RA, Frosch MP & Greenberg SM. The cerebral beta-amyloid angiopathies: hereditary and sporadic. Brain Pathol 2006; 16:30-39.

Zhang ET, Inman CB & Weller RO. Interrelationships of the pia mater and the perivascular (Virchow-Robin) spaces in the human cerebrum. J Anat 1990; 170:111-123.

Zhou Q, Yang D, Ombrello AK, Zavialov AV, Toro C, Stone DL, Chae JJ, Rosenzweig SD, Bishop K, Barron KS, Kuehn HS, Hoffmann P, Negro A, Tsai WL, Cowen EW, Pei W, Milner JD, Silvin C, Heller T, Chin DT, Patronas NJ, Barber JS, Lee CC, Wood GM, Ling A, Kelly SJ, Kleiner DE, Mullikin JC, Ganson NJ, Kong HH, Hambleton S, Candotti F, Quezado MM, Calvo KR, Alao H, Barham BK, Jones A, Meschia JF, Worrall BB, Kasner SE, Rich SS, Goldbach-Mansky R, Abinun M, Chalom E, Gotte AC, Punaro M, Pascual V, Verbsky JW, Torgerson TR, Singer NG, Gershon TR, Ozen S, Karadag O, Fleisher TA, Remmers EF, Burgess SM, Moir SL, Gadina M, Sood R, Hershfield MS, Boehm M, Kastner DL & Aksentijevich I. Early-onset stroke and vasculopathy associated with mutations in ADA2. N Engl J Med 2014; 370:911-920.

Zhu YC, Chabriat H, Godin O, Dufouil C, Rosand J, Greenberg SM, Smith EE, Tzourio C & Viswanathan A. Distribution of white matter hyperintensity in cerebral hemorrhage and healthy aging. J Neurol 2012; 259:530-536.

Ziai W, Gill R, Ullman N, Gandhi D & Hanley D. Reliability of spontaneous intracerebral hemorrhage localization on admission and follow-up computed tomography. Neurocritical Care Conference: 9th Annual Meeting of the Neurocritical care Society.Montreal, QC, Canada 2011; 1 Suppl 1:pp S102.

Zipser BD, Johanson CE, Gonzalez L, Berzin TM, Tavares R, Hulette CM, Vitek MP, Hovanesian V & Stopa EG. Microvascular injury and blood-brain barrier leakage in Alzheimer's disease. Neurobiol Aging 2007; 28:977-986.

Zubenko GS, Hughes HB & Stiffler JS. Clinical and neurobiological correlates of DXS1047 genotype in Alzheimer's disease. Biol Psychiatry 1999; 46:173-181.

Zubenko GS, Stiffler S, Stabler S, Kopp U, Hughes HB, Cohen BM & Moossy J. Association of the apolipoprotein E epsilon 4 allele with clinical subtypes of autopsy-confirmed Alzheimer's disease. Am J Med Genet 1994; 54:199-205.

Zunarelli E, Nicoll JAR, Graham DI, Dewar D, Love S, Ironside SL, Stewart J & Vinters HV. Presenilin-1 polymorphism and hemorrhage due to cerebral amyloid angiopathy. Alzheimer's Res 1997; 3:205-209.

Appendices

Appendix 1

Ovid Embase search strategy (1980 to October 2014)

1. cerebrovascular disease/
2. basal ganglion hemorrhage/ or cerebrovascular accident/ or stroke/ or exp brain hematoma/ or exp brain hemorrhage/
3. ((brain or cerebr$ or cerebell$ or intracerebr$ or intracran$ or parenchyma$ or intraventricular or infratentorial or supratentorial or basal gang$ or ganglion$ or putaminal or putamen or posterior fossa or brain?stem or intra?axial or lobar or deep or thalam$ or cortical or superficial or vertebrobasil$ or front$ or tempor$ or pariet$ or occipit$ or subarachn$) adj5 (haemorrhage$ or hemorrhage$ or haematoma$ or hematoma$ or bleed$)).tw.
4. ((h?emorrhag$ or aneurysm$) adj6 (stroke$ or cerebrovasc$ or cerebr?vasc$ or cerebral vasc$ or brain vasc$ or cva$ or apoplex$ or attack$ or event$ or insult$ or intracranial)).tw.
5. 1 or 2 or 3 or 4
6. (type$ or subtype$ or classification$ or categor$ or SMASH-U).tw.
7. 5 and 6
8. interrater reliability/ or intrarater reliability/ or observer variation/ or reproducibility/ or predictive value/
9. ((observer or interobserver or inter-observer or intra-observer or intraobserver or interrater or rater or inter-rater or intra-rater or intrarater) adj5 (variation$ or variabilit$ or bias or reliability or agreement or comparison$ or error$ or concordance or feasib$)).tw.
10. 8 or 9
11. 7 and 10

Ovid Medline search strategy (1946 to October 2014)

1. Stroke/
2. Cerebrovascular Disorders/
3. exp intracranial hemorrhages/
4. ((brain$ or cerebr$ or cerebell$ or intracerebr$ or intracran$ or parenchyma$ or intraventricular or infratentorial or supratentorial or basal gang$ or ganglion$ or putaminal or putamen or posterior fossa or brain?stem or intra?axial or lobar or deep or thalam$ or cortical or superficial or vertebrobasil$ or front$ or tempor$ or pariet$ or occipit$ or subarachn$) adj5 (haemorrhage$ or hemorrhage$ or haematoma$ or hematoma$ or bleed$)).tw.
5. ((h?emorrhag$ or aneurysm$) adj6 (stroke$ or cerebrovasc$ or cerebr?vasc$ or cerebral vasc$ or brain vasc$ or cva$ or apoplex$ or attack$ or event$ or insult$ or intracranial)).tw.
6. 1 or 2 or 3 or 4 or 5
7. (type$ or subtype$ or classification$ or categor$ or SMASH-U).tw.
8. "reproducibility of results"/
9. Observer Variation/
10. "Predictive Value of Tests"/
11. ((observer or interobserver or inter-observer or intra-observer or intraobserver or interrater or rater or inter-rater or intra-rater or intrarater) adj5 (variation$ or variabilit$ or bias or reliability or agreement or comparison$ or error$ or concordance or feasib$)).tw.
12. 6 and 7
13. 8 or 9 or 10 or 11
14. 12 and 13

Appendix 2 Example letter to contact authors

Dear Dr Palm,

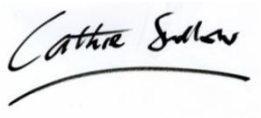
Reliability of spontaneous ICH classification systems – invitation to join a collaborative systematic review

"Intracerebral haemorrhage in a population-based stroke registry (LuSSt): incidence, aetiology, functional outcome and mortality." F. Palm, N. Henschke, J. Wolf, K. Zimmer et al.; J Neurol 2013; Oct;260(10):2541-50.

We would like to invite you to join a collaborative systematic review of the reliability of spontaneous intracerebral haemorrhage (sICH) classification systems. We have recently undertaken a thorough systematic literature search aiming to identify existing studies of the reliability of various sICH classification systems but were able to identify only four relevant publications. These published studies have found the inter-rater reliability to be good, but we want to avoid publication bias by including studies such as yours (above), if you have some additional unpublished data about inter- and/or intrarater reliability of the classification system you used.

We plan to include the relevant person/people from any contributing group as co-authors on any publication arising from this work. If you are interested in collaborating and could provide the relevant data, all you need to do is simply fill in the attached summary tables with data from your study above.

We are circulating this invitation by both email and post, so please contact us by whichever means is most convenient for you. Please respond by filling in and returning the attached forms **by 9th December 2013.** We will be delighted to discuss any aspect of this proposal with you by email or telephone.



Yours sincerely,

Professor Cathie Sudlow
Professor of Neurology and Clinical Epidemiology
Honorary Consultant Neurologist
Chief Scientist & Senior Epidemiologist, UK Biobank
Division of Clinical Neurosciences
University of Edinburgh, WGH, Crewe Road,
Edinburgh EH4 2XU, UK
Email: cathie.sudlow@ed.ac.uk

Dr Kristiina Rannikmäe
Clinical Research Training Fellow
Division of Clinical Neurosciences
University of Edinburgh
Western General Hospital
Crewe Road, Edinburgh EH4 2XU, UK
Tel: +44 (0)7 808 664 850
Email: kristiina.rannikmae@ed.ac.uk

Collaborative systematic review of the reliability of spontaneous ICH classification systems – data collection tables

Please complete as much of these summary tables A - D as you can (we do not expect you to be able to fill in all the boxes). Please fill in or tick the box with the appropriate answer in the grey-shaded cells. Please amend or supplement the information in the prefilled unshaded cells as appropriate.

Please return the completed tables to the following address:

Email: kristiina.rannikmae@ed.ac.uk

and / or

Address:

Dr Kristiina Rannikmäe

Clinical Research Training Fellow

Bramwell Dott Building

Division of Clinical Neurosciences

University of Edinburgh

Western General Hospital

Crewe Road, Edinburgh EH4 2XU, UK

and / or

Fax: (0)131 332 5150 (FAO: Kristiina Rannikmäe)

TABLE A: Study characteristics.

TABLE A	
1. First author and publication title	F. Palm: Intracerebral haemorrhage in a population-based stroke registry (LuSSt): incidence, aetiology, functional outcome and mortality.
2. Publication year	2013
3. Country where study conducted	Germany
4. Primary study aim	To assess incidence, mortality, functional outcome and mortality of ICH based on a prospective population based stroke registry (the Ludwigshafen Stroke Study).
5. Inclusion/exclusion criteria	Excluded recurrent strokes, primary subdural and epidural haematomas, traumatic ICHs, tumour-related ICH, haemorrhagic transformation of an ischaemic stroke. Included cases with first-ever primary ICH over 5 years (between 2006-2010).
6. Number of included ICH cases	152 ICH cases
7. Gender of ICH cases	50.7% male
8. Ethnicity of ICH cases	
9. Mean age ± standard deviation of ICH cases	71.3 ± 13.5
10. Completeness of investigations for ICH	CT for all; MRI for 20.4%; CTA/MRA/DSA for 26.3%
11. Classification of ICH used in study	All ICH cases divided into the following mutually exclusive categories based on <u>location of the ICH (anatomical classification)</u>: 1) Lobar 2) Deep And based on the <u>underlying aetiology of the ICH (aetiological classification)</u>: 1) Structural vascular pathology 2) OAK-related 3) CAA 4) Hypertension 5) Undetermined
12. Definitions of the classification categories	<u>Anatomical:</u> 1) lobar: cortical or subcortical bleeds 2) deep: basal ganglia, infratentorial. <u>Aetiological:</u> 1) Structural vascular pathology: AVM or cavernoma 2) OAK-related: on warfarin with INR >2 3) CAA: older than 54 years with lobar haematoma site 4) Hypertension: deep ICH with pre-ICH hypertension diagnosis (on medication, BP >140/90 on 2 or more measurements > 3 days after stroke) 5) Undetermined
13. Did more than 1 rater[1] classify the ICH cases into one of the above mentioned categories?	☐ Yes ☐ No
14. Did any one rater classify the same cases more than once at a different timepoint?	☐ Yes ☐ No
If the answer to question number 13 and/or 14 is "yes" then please proceed to question 15.	

15. Was the inter-rater reliability of the anatomical ICH classification system assessed?	☐ Yes ☐ No
16. Was the inter-rater reliability of the aetiological ICH classification system assessed?	☐ Yes ☐ No
17. Was the intra-rater reliability of the anatomical ICH classification system assessed?	☐ Yes ☐ No
18. Was the intra-rater reliability of the aetiological ICH classification system assessed?	☐ Yes ☐ No

Questions regarding assessing inter-rater reliability
(if you have not calculated reliability but can provide numeric data in Table C, please continue)

	Anatomical classification	**Aetiological classification**
19. What statistical measure was used to assess reliability?		
20. Please provide the relevant value(s) of the statistical measure used (e.g. the kappa value with 95% confidence intervals where available)		
21. How many raters classified the cases?		
22. What was the experience and qualification of the raters?		
23. What institution(s) were the raters from?		
24. What was the information available to the raters (clinical information and information regarding investigations)?		
25. Was the same information available to each rater?		
26. What (if any) were the decision rules provided to the raters for classifying the cases?		
27. What format was the information provided in? (i.e. clinical notes, extracted information from clinical notes, computerised format etc)		
28. How long from symptom onset was classification done?		
29. Was classification done prospectively or retrospectively?[3]		
30. Was there blinding between raters? (i.e.was each rater unaware of the other rater's decisions?)		
31. Was each rater aware of being compared to other rater(s)?		

Questions regarding assessing intra-rater reliability[2]
(if you have not calculated reliability but can provide numeric data in Table D, please continue)

	Anatomical classification	**Aetiological classification**
32. What statistical measure was used to assess reliability?		
33. Please provide the relevant value(s) of the statistical measure used (e.g. the kappa value with 95% confidence intervals where available)		
34. How many times did the rater classify cases?		
35. What was the experience and qualification of the rater?		
36. What institution was the rater from?		
37. What information was available to the rater (clinical information and information regarding		

investigations)?		
38. Was the same information available to the rater each time?		
39. What (if any) were the decision rules provided to the rater for classifying the cases?		
40. What format was the information provided in? (i.e. clinical notes, extracted information from clinical notes, computerised format etc)		
41. How long from symptom onset where the classifications done?		
42. What was the time interval between the ratings?		
43. Was classification done prospectively or retrospectively?[3]		
44. Was there blinding? (i.e., was the rater unaware of his/her own previous decision and other raters' decisions)		
45. Was the rater aware that intra-rater reliability was assessed (i.e., was he/she aware that his/her decisions where being compared to his/her previous decisions)?		

[1] For the purposes of this study a "rater" is defined as a person assigning an ICH case into one of the two anatomical or five aetiological categories based on the classification of intracerebral haemorrhage used here. [2] If there are data for intra-observer reliability for more than one rater, please fill in this table for each rater separately. [3]Prospective: i.e. patients were classified at the time or shortly after presenting and being recruited to the study. Retrospective: i.e. cases were classified some time after initial presentation, usually through retrieving and reviewing medical records

TABLE B: Information regarding the proportions of cases classified to each category.

Please insert the number of cases classified into each category (if more than one rater classified the cases please explain briefly in the space below the table how consensus was reached to produce these numbers).

TABLE B_1 (anatomical classification)			
	Total	Lobar	Deep
Number of cases in each category	152		

TABLE B_2 (aetiological classification)						
	Total	Hypertensive	CAA	OAK-related	Structural vascular pathology	Undetermined
Number of cases in each category	152					

TABLE C: Numeric data to enable calculation of inter-rater reliability. Please complete a similar table to "example table C" for the 152 ICH cases in your study (for example in an Excel spreadsheet or in Microsoft word). You are also very welcome to provide us with the relevant information in any other format you may have. We are interested in both the anatomical and aetiological classification system used.

Please insert to each cell the number of raters that assigned case X to category Y. Please see an example table below: this table has been populated based on the following assumptions: there were 3 raters assigning 2 ICH cases to one of 2 categories. For example rater number 1 thought that case number 1 was a lobar bleed, but raters 2 and 3 thought that case number 1 was deep. Raters 1 and 3 thought that case number 2 was deep but rater number 2 thought that it was lobar. Etc.

EXAMPLE TABLE C		
Categories →	**Lobar**	**Deep**
2 ICH cases ↓		
Case 1	1	2
Case 2	1	2

TABLE D: Numeric data to enable calculation of intra-rater reliability. Please complete a similar table to "example table D" for the 152 ICH cases in your study (for example in an Excel spreadsheet or in Microsoft word). You are also very welcome to provide us with the relevant information in any other format you may have. We are interested in both the anatomical and aetiological classification system used.

Please insert to each cell the number of times a rater assigned case X to category Y. Please see an example table below: this table has been populated based on the following assumptions: one rater assigned 2 ICH cases to one of 2 categories, the rater did that twice with a 1 month interval. For example on first assessment the rater thought that case number 1 was a lobar bleed, but on the second assessment a month later classified it as deep. The rater thought that case number 2 was a deep bleed on both occasions. Etc.

EXAMPLE TABLE D[4]		
Categories →	**Lobar**	**Deep**
2 ICH cases ↓		
Case 1	1	1
Case 2	0	2

[4] If there is data for intra-rater reliability for more than one rater, please complete Table D for each rater separately.

Appendix 3

Ovid Embase search strategy (1980 to March 2012)

1. exp vascular amyloidosis/
2. exp Congo Red/ or exp amyloid/ or exp amyloid beta protein/ or exp amyloid precursor protein/
3. exp cerebrovascular disease/
4. 2 and 3
5. (amyloid angiopath$ or congophil$ angiopath$ or cerebral amyloid$ or cerebral congo?red or cerebral A?beta or cerebral beta?amyloid).tw.
6. 1 or 4 or 5
7. exp genetics/ or exp genetic disorder/ or genetic epidemiology/ or exp genetic analysis/ or exp population genetic parameters/ or quantitative trait/ or exp molecular genetics/ or exp genetic parameters/ or exp gene mapping/ or exp APOLIPOPROTEIN E2/ or exp APOLIPOPROTEIN/ or exp APOLIPOPROTEIN E3/ or exp APOLIPOPROTEIN E/ or exp APOLIPOPROTEIN E4/
8. (polymorphi$ or genotype$ or gene or genes or genetic$ or allele$ or mutat$ or apolipoprotein$ or apoprotein$ or APO?E$).tw.
9. 7 or 8
10. 6 and 9
11. limit 10 to human

Ovid Medline search strategy (1950 to March 2012)

1. exp Cerebral Amyloid Angiopathy/
2. amyloidosis/ or amyloidosis, familial/ or exp Amyloid beta-Protein/ or exp Amyloid/ or exp Congo Red/
3. exp cerebrovascular disorders/
4. 2 and 3
5. (amyloid angiopath$ or congophil$ angiopath$ or cerebral amyloid$ or cerebral congo?red or cerebral A?beta or cerebral beta?amyloid).tw.
6. 1 or 4 or 5
7. exp genetics/ or exp genotype/ or exp inheritance patterns/ or exp linkage genetics/ or exp genes/ or exp genome/ or apolipoproteins/ or exp apolipoproteins e/
8. (polymorphi$ or genotype$ or gene or genes or genetic$ or allele$ or mutat$ or apolipoprotein$ or apoprotein$ or APO?E$).tw.
9. Cerebral Amyloid Angiopathy/ge [Genetics]
10. 7 or 8
11. 6 and 10
12. 9 or 11
13. limit 12 to humans

Appendix 4

I am currently undergoing discussion with the editors of JNNP with regards to the necessity to make corrections to the published article (Rannikmäe et al. 2013). The published data in the article includes an error in the calculations for the fixed-effects meta-analysis of associations of *APOE* ε2+ versus ε2- genotypes with presence versus absence of CAA. The published overall pooled odds ratio for this analysis was 0.73 (95% confidence intervals 0.53 to 1.00) and the p-value for this analysis was 0.05. The corrected overall pooled odds ratio for this analysis is 0.64 (95% confidence intervals 0.47 to 0.86) and the p-value is 0.003. The error resulted from a mistake in the Microsoft Excel formula (omission of number 1) when calculating the standard error of the logarithm of the odds ratio (SE[logOR]) for the Premkumar study. This error resulted in entering the incorrect SE(logOR) to the Review Manager meta-analysis programme (incorrect value 4.1557; correct value 0.5734) which then assigned an incorrect weight to the study when meta-analysing the data across all included studies (incorrect weight 0.1%; correct weight 7.2%). The meta-analyses included in the thesis for the effects of *APOE* ε4 dose on presence versus absence of CAA also demonstrate different overall odds ratios compared to the published results (Rannikmäe et al. 2013) because since publishing the data, further statistical manipulations have enabled additional data to be included. The overall conclusions remain unchanged.

Appendix 5

Vonsattel scale for grading CAA:

- Mild: amyloid is restricted to a congophilic rim around normal or atrophic smooth muscle fibers in the media of otherwise normal vessels
- Moderate: media is replaced by amyloid and is thicker than normal, with no evidence of remote or recent blood leakage.
- Severe: extensive amyloid deposition with focal wall fragmentation and at least one focus of paravascular leakage evidenced by the presence of erythrocytes or hemosiderin or both.

Appendix 6

Example letter to contact authors

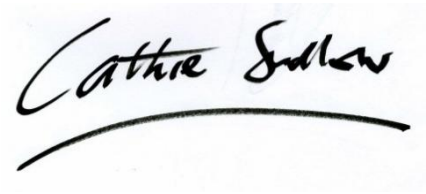
Dear Dr Chui,

Re: Collaborative meta-analysis of APOE ε2 and CAA related vasculopathic changes

We would like to invite you to join a collaborative meta-analysis of the association between APOE genotype and cerebral amyloid angiopathy related vasculopathic changes that lead to CAA-haemorrhage. We propose including as co-authors on any publication arising from this the relevant person/people from each contributing group. All that is required is for you to simply fill in the attached summary tables with data from your study described in: *Chui HC, Zarow C, Mack WJ et al. Cognitive impact of subcortical vascular and Alzheimer's disease pathology. Ann Neurol 2006;60:677-687.*

Our interest in this topic has arisen after our recently completed systematic review and series of meta-analyses looking at the genetic associations of CAA (please find a brief summary of the results and the scientific background for this work attached).

We are circulating this invitation by both email and post, so please contact us by whichever means is most convenient for you. Please respond by filling in and returning the attached forms by July 9th 2012. We will be delighted to discuss any aspect of this proposal with you by email or telephone.



Yours sincerely,

Dr Cathie Sudlow
Clinical Reader & Honorary Consultant Neurologist,
Chief Scientist & Senior Epidemiologist, UK Biobank
Division of Clinical Neurosciences
University of Edinburgh
Western General Hospital
Crewe Road, Edinburgh EH4 2XU, UK
Email: cathie.sudlow@ed.ac.uk

Dr Kristiina Rannikmäe
Clinical Research Training Fellow
Division of Clinical Neurosciences
University of Edinburgh
Western General Hospital
Crewe Road, Edinburgh EH4 2XU, UK
Tel: +44 (0)7 808 664 850
Email: kristiina.rannikmae@ed.ac.uk

Collaborative meta-analysis of APOE ε2 and CAA related vasculopathic changes - data collection tables

Please complete as much of these summary tables as you can (we do not expect you to be able to fill in all the boxes). Please fill in or mark the appropriate answer in the grey-shaded cells. If the information in the prefilled unshaded cells is incorrect, please amend as appropriate. We would also like to once more draw you attention to the fact that we are only collecting data about participants without CAA related haemorrhage.

Please return the completed tables to the following address:

Address:

Dr Kristiina Rannikmäe

Clinical Research Training Fellow

Bramwell Dott Building

Division of Clinical Neurosciences

University of Edinburgh
Western General Hospital
Crewe Road, Edinburgh EH4 2XU, UK
and / or
Email: kristiina.rannikmae@ed.ac.uk

Table 1 – study characteristics and methods

First author and year of publication	Chui 2006
Country	USA
Brief description of participants	People from the Ischaemic Vascular Dementia Programme (longitudinal study of subcortical ischaemic vascular dementia, AD and cognitively normal elderly people)
No. of participants clinically demented	
No. of participants not clinically demented	
No. of participants for whom dementia status not known	
No. of participants with clinical and neuropathological diagnosis of AD	
No. of participants genotyped & assessed for CAA	
Mean age of participants at death	
Ethnicity of participants	86 % Caucasian
% participants male	
Source of DNA	Blood
Were genotypes assigned using all data simultaneously or in batches?	
No. in whom genotyping attempted/in whom successful	
Genotyping method	- PCR - Other (please specify)
Were genotypers blinded to pathology data?	- Yes - No
Method for assessing CAA on Vonsattel scale	Congo Red and/or H&E for all participants, IHC for some
Method for assessing double-barrelling	
Method for assessing blood vessel dilatation and microaneurysms	
Method for assessing fibrinoid necrosis	
Qualification of the person who rated CAA and vasculopathic changes	Neuropathologist
Locations examined for CAA severity	Frontal-, temporal-, parietal-, occipital cortex
Locations examined for CAA related vasculopathic changes	
Were pathologists blinded to genotype?	Yes
Were pathologists blinded to relevant clinical information (i.e.dementia/AD status, age, etc of the participants) ?	Yes

Table 2 – study data

Please enter the number of participants with each genotype/phenotype feature in the grey cells.

Phenotype – pathological features	Genotype					
	ε2/ε2	ε2/ε3	ε2/ε4	ε4/ε4	ε4/ε3	ε3/ε3
Severity of CAA on the <u>Vonsattel scale</u>*						
Severe CAA						
Moderate CAA						
Mild CAA						
No CAA						
Total no. of participants						
CAA related changes						
Any severity of CAA						
Present						
Absent						
Total no. of participants						
Vasculopathic CAA: fibrinoid necrosis#						
Present						
Absent						
Total no. of participants						
Vasculopathic CAA: blood vessel dilatation and/or microaneurysm formation						
Present						
Absent						
Total no. of participants						
Vasculopathic CAA: double-barrelling of the vessel wall ("vessel within a vessel" appearance)						
Present						
Absent						
Total no. of participants						

*The Vonsattel scale is defined as follows: "Mild"- amyloid is restricted to a congophilic rim around normal or atrophic smooth muscle fibres in the media of otherwise normal vessels. "Moderate" - media is replaced by amyloid and is thicker than normal, with no evidence of remote or recent blood leakage. "Severe" - extensive amyloid deposition with focal wall fragmentation and at least one focus of paravascular leakage evidenced by the presence of erythrocytes or hemosiderin or both. Each participant should be graded according to the most advanced degree of CAA present among all vessels examined. If a different scoring method has been used, please explain. #Fibrinoid necrosis is defined as homogeneous discrete foci or segments of vascular wall with smudgy eosinophilic material that stain dark blue with PTAH (posphotungstic acid hematoxylin stain) or red with MSB (Martius Scarlett blue trichrome stain) and that obscure cytoarchitecture.

Appendix 7

The overall number of participants for the other polymorphisms part varies slightly between the published manuscript (4703 participants) (Rannikmäe et al, 2013) and the thesis data (4617 participants). This is because for the Chapuis et al 2006 paper, different number of cases who have been genotyped and had pathological CAA assessment is reported in text and in table 2 (114 and 76 respectively). After further discussions, I have decided to use the smaller number for the purposes of the thesis, to make sure I do not overestimate the effect of this study. For the Durany et al 2000 paper, similarly it is unclear which groups of participants have been included in various parts of the analyses and after further discussions and review of the paper, I have decided that it is more appropriate for the purposes of the thesis to use the smaller number of AD cases (32) compared to the overall number of 80 participants that include CAA-related ICH cases and controls. The overall conclusions remain unchanged.