GENETIC AND INTER-GENERATIONAL EPIDEMIOLOGY

Complement factor H genetic variant and age-related macular degeneration: effect size, modifiers and relationship to disease subtype

Reecha Sofat,1 Juan P Casas,2,3 Andrew R Webster,4,5 Alan C Bird,4,5 Samantha S Mann,4,5 John RW Yates4,5,6 Anthony T Moore,4,5 Tiina Sepp,4 Valentina Cipriani,4,5 Catery Bunce,4,5 Jane C Khan,6 Humma Shahid,6 Anand Swaroop,7,8,9 Gonçalo Abecasis,10 Kari E H Brantham,7 Sepideh Zareparsi,7 Arthur A Bergen,11,12,13 Caroline CW Klaver,13 Dominique C Baas,11 Kang Zhang,14,15 Yuhong Chen,14,15 Daniel Gibbs,15 Bernhard H F Weber,16 Claudia N Keilhauer,17 Lars G Fritzsche,16 Andrew Lotery,18 Angela J Cree,18 Helen L Griffiths,18 Shomi S Bhattacharya,4,5 Li L Chen,4,5 Sharon A Jenkins,4,5 Tunde Peto,4,5 Mark Lathrop,19 Thierry Leveillard,20 Michael B Gorin,21 Daniel E Weeks,22 Maria Carolina Ortube,21 Robert E Ferrell,22 Johanna Jakobsdottir,23 Yvette P Conley,23,24 Mati Rahu,25 Johan H Seland,26 Gisele Soubrane,27 Fotis Topouzis,28 Jesus Vioque,29 Laura Tomazzoli,30 Ian Young,31 John Whittaker,3,32 Usha Chakravarthy,33 Paulus T V M de Jong,11 Liam Smeeth,3 Astrid Fletcher3 and Aroon D Hingorani1,2,*

1Centre for Clinical Pharmacology, Department of Medicine, University College London, London, UK, 2Genetic Epidemiology Group, Department of Epidemiology and Public Health, 1-19 Torrington Place, University College London, London, UK, 3Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK, 4Institute of Ophthalmology, University College London, London, UK, 5Moorfields Eye Hospital, London, UK, 6Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK, 7Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, 8Neurobiology-Neurodegeneration & Repair Laboratory (N-NRL), National Eye Institute, National Institutes of Health, Bethesda, MD, USA, 9Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA, 10Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA, 11Department of Ophthalmogenetics, Meibergdreef 47, The Netherlands, 12The Netherlands Institute for Neuroscience (NIN-KNAW), 13Department of Clinical Genetics, Meibergdreef 47, The Netherlands, 14Ophthalmology, Academic Medical Centre Amsterdam (AMC/UVA), The Netherlands, 15Institute for Genomic Medicine and Department of Ophthalmology, University of California San Diego 9500 Gilman Drive, La Jolla, CA, USA, 16Department of Ophthalmology and Visual Sciences, Moran Eye Center, University of Utah School of Medicine, Salt Lake City, Utah, 17Institute of Human Genetics, University of Regensburg, Franz-Josef-Strauss-Allee 11, Germany, 18Eye Hospital, University of Wuerzburg Josef-Schneider-Str. 11, Germany, 19Division of Clinical Neurosciences, Southampton General Hospital, University of Southampton, 20Centre National de Genotypage, 21Centre National de Génétique, 22Department of Health Promotion & Development, School of Nursing, PA, USA, 23Department of Epidemiology and Biostatistics, National Institute for Health Development, Estonia, 24Stavanger University Hospital, Stavanger, University of Bergen, Norway, 25Department of Ophthalmology, Aristotle University of Thessaloniki, AHEPA Hospital, Thessaloniki, Greece, 26University Miguel Hernandez and CIBER en Epidemiología y Salud Pública (CIBERESP), Spain, Dpto Salud Publica. Campus San Juan. Ctra. Nacional 332 s/n. 03550-San Juan de Alicante, Spain, 27Section of Ophthalmology of the Dipartimento di Scienze Neurologiche e della Visione dell’Università degli Studi di Verona, Ospedale di Borgo Trento – Padiglione Geriatrico – P.Stefani 1 - 37124 Verona, 28Centre for Public Health, 1st Floor ICS B Block, Royal Victoria Hospital, Grosvenor Road, Belfast BT12 6BJ, 29Statistical Platforms and Technologies, GlaxoSmithKline, Medicines Research Centre, Maitstop 15101, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY and 30Centre for Vision and Vascular Science, Institute of Clinical Science, The Queen’s University of Belfast

*Corresponding author. Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London WC1E 6BT, UK. E-mail: a.hingorani@ucl.ac.uk

Accepted 16 November 2011

Background Variation in the complement factor H gene (CFH) is associated with risk of late age-related macular degeneration (AMD). Previous studies have been case-control studies in populations of European
ancestry with little differentiation in AMD subtype, and insufficient power to confirm or refute effect modification by smoking.

Methods

To precisely quantify the association of the single nucleotide polymorphism (SNP rs1061170, ‘Y402H’) with risk of AMD among studies with differing study designs, participant ancestry and AMD grade and to investigate effect modification by smoking, we report two unpublished genetic association studies ($n = 2759$) combined with data from 24 published studies (26 studies, 26,494 individuals, including 14,174 cases of AMD) of European ancestry, 10 of which provided individual-level data used to test gene–smoking interaction; and 16 published studies from non-European ancestry.

Results

In individuals of European ancestry, there was a significant association between Y402H and late-AMD with a per-allele odds ratio (OR) of 2.27 [95% confidence interval (CI) 2.10–2.45; $P = 1.1 \times 10^{-16}$]. There was no evidence of effect modification by smoking ($P = 0.75$). The frequency of Y402H varied by ancestral origin and the association with AMD in non-Europeans was less clear, limited by paucity of studies.

Conclusion

The Y402H variant confers a 2-fold higher risk of late-AMD per copy in individuals of European descent. This was stable to stratification by study design and AMD classification and not modified by smoking. The lack of association in non-Europeans requires further verification. These findings are of direct relevance for disease prediction. New research is needed to ascertain if differences in circulating levels, expression or activity of factor H protein explain the genetic association.

Keywords

Age-related macular degeneration (AMD), Complement factor H gene, meta-analysis

Introduction

Age-related macular degeneration (AMD) is an important cause of blindness in adults worldwide. Early identification of individuals at greater risk of disease and, in particular, late-stage AMD, which is most sight threatening, might help target preventative interventions to those at highest risk. Moreover, better understanding of the causes of AMD could lead to the development of improved treatments. AMD was the first disease in which a susceptibility locus was identified by a genome-wide association scan (GWAS). Carriers of variants in the complement factor H gene (CFH), a key regulator of the alternate complement pathway, were found to be at higher risk of AMD. The most replicated association has been with a single nucleotide polymorphism (SNP) rs1061170 encoding a tyrosine to histidine change at position 402 (Y402H).

As this is a common allele (minor allele frequency, 0.3) of large effect, it has been proposed that genotyping of this variant could be used to help predict risk of AMD. However, a number of existing uncertainties require resolution before a genetic test of this type could be recommended for use in clinical practice. First, the precise magnitude of the effect size requires better delineation because large effect sizes in initial ‘discovery’ studies can become attenuated as the literature matures. Secondly, it is unclear if the risk conferred by carriage of this variant is the same for all grades of AMD. Thirdly, although it has been suggested that smoking (a recognized risk factor for late AMD) could modify the effect of the Y402H polymorphism, individual studies have not been large enough to address this question reliably. Fourthly, it is uncertain whether Y402H is itself a causal variant or simply marks another causal site [because of linkage disequilibrium (LD)], either within or outside CFH. Since LD patterns vary among populations of differing ancestry, the magnitude of association and in turn the predictive performance of Y402H could be very different in populations of differing ancestry. Moreover, if Y402H marks a causal site in an adjacent gene, of which several exist in the region of complement activation (RCA) cluster on chromosome 1, targeting CFH or its downstream effects may not provide a useful means of treating or preventing AMD.

To address some of these uncertainties we tested the association of Y402H with AMD in two new studies,
Methods

New studies

MRC AMD case–control study
This was a prospective case–control study in 1469 unrelated individuals of European ancestry (1234 clinic-based cases and 194 population-based controls) undertaken at Moorfields Eye Hospital, a large ophthalmic hospital in London, United Kingdom. Each participant was interviewed specifically for the study, and a family history, smoking history and other medical history were taken. The ophthalmic examination included Snellen acuity, slit-lamp examination and bio-microscopic fundoscopy. Colour stereoscopic fundus photography of the macular region with grading of the photographs according to the International Classification System for age-related maculopathy (ICARMS) classification by two trained readers independently with any discrepancies being resolved by an ophthalmologist (Alan C Bird). Auto-fluorescence images were taken of the maculae and fluorescein angiography was performed when choroidal neovascularisation was suspected. For patients presenting with visual dysfunction in the second eye, retrospective data were gathered from hospital records concerning previous acuities. Moreover, any colour images or fluorescein images relating to previous visual loss were located from the hospital archive. All images were digitized. Cases were excluded if they had retino-choroidal inflammatory disease, diabetic retinopathy, branch retinal vein or artery occlusion or any other cause of visual loss other than amblyopia. Controls were all examined in a similar fashion and grading of the photographs according to the ICARMS classification at a single reading centre. Drusen were categorized according to their size, homogeneity of surface features and outlines, pigmentary irregularities were classified into either hypopigmentation or hyperpigmentation. Early AMD was defined as the presence of soft indistinct drusen (≥125 μm) or reticular drusen only or soft distinct drusen (≥63 μm) with pigimentary abnormalities or soft indistinct drusen (≥125 μm) or reticular drusen with pigimentary abnormalities. Late AMD subtypes were geographical atrophy (GA) or choroidal neovascularisation (CNV). GA was defined as any sharply demarcated round or oval area of apparent absence of the RPE, >175 μm, with visible choroidal vessels and no CNV. CNV was defined as the presence of a serous or haemorrhagic detachment of the RPE and/or a sub-retinal neovascular membrane and/or sub-retinal haemorrhage, and/or peri-retinal fibrous scarring, even with patches of GA. Blood was drawn at the time of interview into EDTA tubes and DNA was extracted and stored at −80°C. Genotyping was carried out at KBiosciences (http://www.kbioscience.co.uk) using KASPar chemistry, a competitive allele-specific PCR SNP genotyping system using FRET quencher cassette oligos (http://www.kbioscience.co.uk/genotyping/chemistry.htm).

Meta-analysis

Search strategy
MEDLINE (using Pubmed up to August 2008) and EMBASE (1981–2008) were searched, for studies evaluating the association between Y402H polymorphism and AMD. Free text and MeSH terms used were ‘age related macular degeneration’, ‘macular degeneration’, ‘complement factor H’ and ‘Y402H’, ‘rs1061170’. Searches were limited to ‘human’. Additional studies were identified through reference lists of publications and searching through ‘related articles’ link on Pubmed.

Inclusion criteria
To be included, studies had to be cohort or case–control (nested/prospective/retrospective), or nested in randomized controlled trials (RCT), studying unrelated individuals. Where studies contained duplicated results, the study with the smaller population was excluded. If there were clearly two populations (e.g. discovery and replication populations), both were included and notation used for these was as used in the primary study. Where studies contained data sets
from related and unrelated cases, information extracted was limited to the unrelated subset of the data.

**Data collection and management**

Data were extracted by one of the authors, uncertainties resolved by discussion with two others and differences resolved by consensus. Data were collected on AMD outcome from studies which reported any subtype of AMD, these were categorized either as early AMD or late AMD (including both GA and CNV, and mixed GA and CNV). Study quality was assessed according to HuGENET guidelines for genetic association studies (http://www.hugenet.org.uk). The following information was abstracted from published reports: grading scales used for AMD diagnosis, clinical categories of AMD as they were reported (early AMD, GA, CNV, late AMD, or mixed GA and CNV), if genotyping was carried out blinded to outcome, genotyping platform used, deviation from Hardy Weinberg Equilibrium (HWE) and ancestry of population under study. Corresponding authors from the studies that met inclusion criteria and which included a total of \( \geq 500 \) individuals were contacted on at least three occasions to request further limited tabular or individual level data. We requested information on age, gender, smoking status, Y402H genotype, AMD status and AMD subtype if affected and grading scale used for diagnosis (see Supplementary methods available at *IJE* online for details on harmonization of data). Individual level data were reconstructed from limited tabular data where necessary (see Supplementary methods available at *IJE* online).

**Study quality**

The Venice criteria were applied to the studies included in the meta-analysis to assess the strength of the cumulative evidence on genetic associations. These criteria are a semi-quantitative index which assigns three levels to the association under study. The levels assigned include first the amount of evidence, second the extent of replication and third the degree of protection from bias. Within each category, three scores from A to C were assigned, with A signifying the highest level indicating the strength of data in each category and C the lowest in each category, indicating paucity of data or poor data quality in each category. A composite score of assessment of association is generated which can then be graded as strong, moderate or weak for the association.

**Statistical analysis**

**Newly genotyped studies**

To evaluate differences between the groups, unpaired Student’s *t* - or \( \chi^2 \)-tests were used as appropriate. Endpoints analysed included all AMD and late AMD, and AMD subtypes (early, GA and CNV). Departure from HWE was evaluated using a \( \chi^2 \) analysis in controls. The principal *a priori* hypothesis was that the association between Y402H and late AMD follows an additive model according to the number of \( C \) alleles. However, recessive, dominant and genotypic models were also evaluated. For the additive model the OR was compared using logistic regression between cases and controls by assigning scores (0, 1 and 2) for different genotype groups and calculating the ORs and 95% CIs. For the EUREYE data set, an additional term was added in modelling the effect size, standard errors, \( P \)-values and corresponding 95% CIs, to take account of clustering of data by country. Concomitant subsidiary analyses assessed the association Y402H in the different subtypes of AMD. The joint contributions of the Y402H with smoking were assessed to evaluate any gene–environment interaction. Smoking was coded either as ever and never, or as current, ex-smoker and never smoked. Interaction on the multiplicative scale by age and gender was also explored. All data analysis was carried out using Stata (Version 11, Stata Corp LP College Station, Texas, USA).

**Meta-analysis of summary level data**

Summary data were used to quantify the effect of the Y402H variant on AMD risk in European and non-European ancestry individuals. Using data from European ancestry subjects we also evaluated the effect of AMD grading scale, study design and genotyping platform on effect estimates using the Der Simonian and Laird Q test and \( I^2 \), a measure to describe the percentage of variability in point estimates, and thus assess impact of heterogeneity on meta-analysis.

**Individual level meta-analysis**

In a subset of studies of European descent individuals providing individual level or detailed tabular data, we evaluated the effect of the Y402H variant on early and late forms of AMD and AMD subtype, and investigated the potential modifying effects of age, gender and smoking habit. For these analyses we used mixed effects logistic models treating ‘study’ as the random factor. Interactions of the Y402H–AMD association by age, gender and smoking were assessed on the multiplicative scale. We also used the case-only approach to further investigate the potential for effect modification by smoking and further tested the effect of intensity of smoking, as measured by pack-years of smoking on AMD risk, stratified by Y402H genotype, in studies where this information was available. Given the exploratory nature of these analyses, 99% CIs were used.

**Heterogeneity and effect modification**

Heterogeneity can arise because of errors and biases as well as true biological variation. We therefore evaluated the effect of study design, genotyping platform and grading scale for the diagnosis of late AMD (as sources of error and bias), and the effect of smoking and disease subtype, age and gender, as described.
above (as potential explanations for true biological variation). Using aggregate level meta-analysis, summary measures were calculated using a random effects model with inverse variance weights. Heterogeneity between estimates was assessed using the Der Simonian and Laird Q test and \( I^2 \) was used as a measure to describe the percentage of variability in point estimates. For smoking age, gender and AMD subtype, estimates were calculated as described above using mixed logistic regression models, and heterogeneity across groups was tested using a \( \chi^2 \) test.

Results

Newly genotyped studies

Clinical demographic data and distribution of Y402H genotype from the MRC case–control study and EUREYE are shown in Table 1. Allele frequencies in cases and controls were comparable to other published studies. There was no deviation from HWE in controls. An additive age-adjusted model (per C allele) demonstrated a significant association between Y402H and late AMD risk in both the MRC case–control study, OR 2.04 (95% CI 1.63–2.56; \( P = 1.77 \times 10^{-7} \)) and EUREYE, OR 2.43 (95% CI 2.04–2.90; \( P = 4.54 \times 10^{-9} \)) (Table 2). The effect was of the same order of magnitude whether or not clinical subtypes were coded as late AMD overall, or separately as GA and CNV, or when individuals who ever smoked were compared with never smokers (Table 2). There was no evidence for a gene–smoking interaction (\( P = 0.42 \) MRC case–control and \( P = 0.78 \) EUREYE for interaction).

Meta-analysis

The search identified 40 studies from which relevant information could be abstracted, 24 of which were in European populations, \(^2\)–5,13–32 six were in Japanese populations, \(^33\)–38 three in Chinese \(^39\)–41 (including mainland China and Hong Kong), three in Taiwanese, \(^42\)–44 one in Korean, \(^45\) one in Indian/South Asian, \(^46\) one in Latin American \(^47\) and one in Black South African subjects \(^48\) (Supplementary Figure S1, Supplementary Table S1). Ten studies, comprising a total 5804 cases of AMD including the newly genotyped studies, together provided more detailed limited tabular data or individual level data.\(^{14,19,23,28,29}\)

Main effect

Of a total of 26 studies (24 published and 2 newly genotyped), 18 reported studies which could be combined as late AMD (6231 cases and 10382 controls). The per-C allele OR for late AMD, from these studies, was OR 2.27 (99% CI 2.10–2.45; \( P = 1.1 \times 10^{-161} \)) (Figure 1) and for any AMD (early and late combined) the OR was 1.86 (99% CI 1.77–1.97; \( P = 1.58 \times 10^{-138} \); 26494 individuals; 14174 cases). Genotyping platform (\( \chi^2 P = 0.06 \) on 6 degrees of freedom, \( I^2 = 49.6\% \)), clinical grading scale (\( \chi^2 P = 0.33 \) on 5 degrees of freedom, \( I^2 = 13.3\% \)), and study design (\( \chi^2 P = 0.80 \) on 2 degrees of freedom, \( I^2 = 0.0\% \)) contributed to the heterogeneity in the effect size (Supplementary Figure S2).

Grade of AMD

The OR per-C allele for the association with early AMD was 1.47 (99% CI 1.37–1.58; 13 studies, 5224 cases); for GA was 2.50 (99% CI 2.17–2.99; 10 studies,
Effect modification by smoking

Using individual level data from five studies (2403 cases of late AMD) the per-allele OR for late AMD, when smoking was coded as current, ex-smokers and never smoked, was 2.43 (99% CI 1.94–3.05; \( P = 4.32 \times 10^{-7} \)) for never smokers, 2.25 (99% CI 1.83–2.76; \( P = 1.58 \times 10^{-4} \)) for ex-smokers vs never smokers and 2.50 (99% CI 1.58–3.96; \( P = 2.65 \times 10^{-7} \)) for current vs never smokers (Figure 2). When smoking was coded
as ever or never, the per-allele OR for late AMD in those who had never smoked was 2.30 (99% CI 2.03–3.71; \( P = 1.50 \times 10^{-30} \)) and in ever vs never smokers was 2.43 (99% CI 1.94–3.04; \( P = 2.27 \times 10^{-24} \)). There was no evidence for interaction using either categorization (P-value for interaction \( P = 0.75 \) and 0.63, respectively).

The findings were similar using any AMD as an outcome (nine studies, 5804 cases) and are summarized in Figure 2. Using a case only approach, limited to late AMD cases (\( n = 2403 \)) from five studies providing individual level data, the per-allele OR for association between smoking and genotype in late AMD cases only was 0.81 (99% CI 0.60–1.10), indicating no evidence for a gene–environment interaction. To test the assumption of independence between the genotype and smoking in the general population, the control group was also assessed, and the OR of association between smoking and genotype in controls only was 1.04 (99% CI 0.79–1.35). Further assessment of the effect of intensity of smoking, as measured by pack-years of smoking on AMD risk by genotype also did not show any evidence of interaction (Supplementary Figure S3).

**Effect modification by age and gender**

There was no evidence of effect modification by gender (P-value for interaction = 0.24), or age (P-value for interaction = 0.38), where age was fitted as a continuous variable. Both analyses used information from 10 studies (9603 individuals).

**Effect of ethnicity**

The frequency of the \( C \) allele for Y402H is lower in Japanese (MAF = 0.06) and Chinese (MAF = 0.07), compared with European subjects (MAF = 0.3) (Figure 3 and Supplementary Figure S4). Sixteen eligible genetic association studies were identified among non-European populations, with data reported in a form amenable to meta-analysis from 14 studies, with genotype counts being calculated from allele frequency data in one study.\(^{34}\) Data from one study were not extractable in a format to include per-allele estimates.\(^{46}\) The per-allele OR for AMD based on an analysis of data from six studies of Japanese participants (672 cases) was 1.10 (95% CI 0.72–1.66), and for individuals of Chinese ancestry (three studies, Hong Kong and mainland China, 447 cases), the per-allele odds of AMD was 1.43 (95% CI 0.93–2.22). The estimate from Taiwanese was higher at 3.46 (95% CI 2.38–5.03) (Supplementary Figure S5).

**Assessment of study quality based on the Venice Criteria**

Venice Criteria were applied to studies from populations of all ancestries, with the largest amount of
Table 3  Venice Criteria for study quality by ancestral origin of population studied

<table>
<thead>
<tr>
<th>Population</th>
<th>Amount of evidence</th>
<th>Replication</th>
<th>Protection from bias</th>
<th>Venice score</th>
</tr>
</thead>
<tbody>
<tr>
<td>White European</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>AAA</td>
</tr>
<tr>
<td>Japanese</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>BCB</td>
</tr>
<tr>
<td>Chinese</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>BCC</td>
</tr>
<tr>
<td>Indian (South Asian)</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>CCC</td>
</tr>
<tr>
<td>Latin American</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>CCC</td>
</tr>
<tr>
<td>Black African</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>CCC</td>
</tr>
</tbody>
</table>

A score of A indicates strong epidemiological credibility, B indicates moderate and C indicates weak, therefore a score of AAA in all categories indicates a large amount of evidence with extensive replication and protection from bias. Studies amongst European descent populations present strong epidemiological credibility although the relative paucity of data amongst other ancestries makes this association less so, although this may reflect inadequate power to detect an effect given the MAF particularly in Japanese and Chinese populations.
information coming from genetic association studies in those of White European ancestry. These findings are summarized in Table 3.

Discussion

Summary of main findings

By conducting a meta-analysis using a combination of new studies, published and unpublished individual level or limited tabular data, we have been able to quantify more precisely and reliably the risk of late AMD conferred by carriage of the Y402H variant in CFH. Our analysis indicates that this association has remained robust, with the best estimate of the effect size [a per-allele OR of 2.27 (99% CI 2.10–2.45)] remaining largely undiminished despite the continuing accrual of information on the association since the original GWAS. This estimate is more precise than the previously published meta-analysis which included 4856 cases from eight studies (OR for CC vs TT 6.32 95% CI (4.25–9.28); TC vs TT 2.50 95% CI (1.96–3.30)). Of the >500 GWAS of common diseases conducted to date, mainly in subjects of European ancestry, the association of CFH with AMD is amongst the largest for a common allele conferring susceptibility for a complex disease.

Exploiting the large size of the assembled data set, and in particular the availability of individual level or limited tabular data, we found no evidence for modification of the genetic association by smoking habit among Europeans, which has been reported previously in other, smaller (and therefore lower power) studies. We did identify evidence of a modest difference of genetic effect in comparisons of early vs late forms of AMD, but point estimates of effect size for disease subtype were broadly consistent. Notably, the association of Y402H was not replicated in individuals of non-European (mainly Chinese and Japanese) ancestry, though published information from all non-European populations is limited, which, coupled with the lower frequency of the risk allele, indicates that larger studies will be required to confirm or refute an effect of this variant on AMD risk in Chinese and Japanese subjects.

These findings from a meta-analysis of association studies that followed the initial discovery GWAS are of specific relevance to AMD, but the approach presented here in collating information from subsequent replication studies and probing interactions, is likely to be informative for other complex disorders where GWAS have provided new leads on causative genes that are subsequently assessed by additional studies in independent data sets, and has been recently demonstrated for coronary disease and variants in genetic markers on chromosome 9p21, which were first identified through GWAS.

Stability of the main effect estimate and implication for predictive genetic testing

Irrespective of whether Y402H is directly causal or not, the large size of the risk estimate from initial studies of Y402H and AMD risk in Europeans, and the high frequency of this allele in European populations have led to an interest in its use as a screening or predictive tool. This variant has been included, e.g. in a genetic risk profile offered by deCODE genetics (www.decode.com) and 23 and me (www.23andme.com), two commercial providers of genetic screening tests for complex diseases. However, initial effect estimates in genetic association studies, whether candidate based or GWAS, can be prone to the winner’s curse—inflation of the reported effect estimate over its true value. This can be shown to be a particular problem where the initial discovery study is small (and low in power) and declaration of association is based on its crossing a pre-specified P-value threshold. The initial GWAS in AMD was based on a collection of only 96 cases and 50 controls, which would be considered quite small for a GWAS of common disease by current standards (e.g. WTCCC1 included 2000 cases and controls). Thus, although the reported association of Y402H has proved robust, the initial effect estimate of an OR of 7.4 (95% CI 2.9–19.0, for homozygous individuals) is likely to be higher than the true effect. In our analysis of 6231 late AMD cases including the initial discovery study, we noted some attenuation of the effect size, as data on this association have expanded (Supplementary Figure S6), and we believe that the summary estimate identified by our meta-analysis provides the best current estimate of the true effect. We also noted similar effect estimates in the prospective longitudinal studies when compared with hospital-based case-control studies or cross-sectional studies (Supplementary Figure S2). Case-control studies are the preferred design for gene discovery and tend to be based on highly selected samples chosen to maximize the differences between cases and controls by choosing patients with younger onset or more advanced disease, and excluding controls whose fundoscopic examinations are not entirely normal, and may therefore overestimate risk in the general population. Prospective studies, more representative of general populations, collecting incident disease of AMD may therefore be better placed to inform on the clinical utility of this SNP in risk prediction and importantly the cost effectiveness of genetic testing of a common allele at a population level.

Genetic association studies that follow a GWAS tend to focus on a subset of SNPs, and often only the SNP(s) with the most extreme P-value are replicated, although these may not be the causal variants. The literature may also become prone to publication bias as positive replications may be more likely to be published than negative ones. However, this particular association does not seem to be substantially
influenced by publication bias as shown by the symmetry of the funnel plot (Supplementary Figure S7).

**Effect modification and heterogeneity**

One important attribute of a meta-analysis of genetic association is that it allows exploration of the existence and reasons for variation of effect size across data sets. In the analysis of the main effect of CFH genotype on AMD risk we found evidence for heterogeneity and explored reasons for this. We also tested potential modifiers of effect size with particular focus on investigating the effect of modification by smoking, which has aroused considerable interest. Study design, genotyping and grading scale all contributed to the observed heterogeneity but the influence was not large suggesting that these sources of variation do little to alter the summary effect estimate.

Neither gender nor smoking habit exhibited any interaction with genotype. The absence of a smoking interaction in this meta-analysis is at odds with the findings and interpretation from some individual studies. Our analysis of effect modification by smoking was conducted in a subset of the database in which information on effect size by smoking habit was available through direct contact with the study authors. Despite being limited to part of the dataset, the smoking analysis still incorporated data on up to 5804 AMD cases of which 3103 individuals were classified as current or prior smokers from nine studies. Quantitative data on smoking exposure (in terms of pack-years) were also available from 3279 individuals. In none of these analyses did we find positive evidence for a gene–smoking interaction. Further case-only analyses indicated an absence of gene–environment interaction, highlighting a consistent lack of evidence of interaction with smoking given the number of methods used to explore its presence. Advantages of the case-only approach are that it excludes potential biases arising from control selection and overcomes problems arising from combining different study designs, as well as having enhanced power compared with using cases and controls to test for interaction. Even with the large size of the data set, we recognize we may be underpowered to detect a modest interaction, but the analysis we have conducted suggests that the detection of smoking interaction in individual studies in this field is just as likely to be due to chance as a real finding.

The pathophysiological relationship between the different subtypes of AMD is poorly understood and the mechanisms may be different. For this reason we undertook a separate analysis of the association of Y402H with the two major late AMD phenotypes, an analysis that no individual study to date has been sufficiently powered to do. We found some evidence for a larger effect of this variant on risk of GA compared with neovascular AMD. However, this finding should be interpreted with caution because it is based on a subgroup analysis (though we tried to reduce the potential for type 1 error by pre-specifying the analysis and using 99% rather than 95% CIs), and the confidence limits for this estimate overlap estimates associated with other AMD subtypes. Furthermore, the apparently more modest effect of Y402H on early AMD may simply reflect the fact that prospective and cross-sectional observational studies with a surveillance approach to assessment of AMD contributed the majority of cases of this end-point. Although the difference is probably insufficiently large for this information to be useful for disease prediction, it may shed light on mechanistic differences in the disease subtypes that could be enhanced by meta-analyses of the effect of other variants in CFH and variants in other risk genes on disease subtype.

**Rationale for the specific interest in Y402H**

Multiple SNPs in and around CFH have demonstrated association with AMD, including a copy number variant and SNPs in downstream genes, other complement related genes (Factor B and C3), and an independent locus (ARMS2 on chromosome 10). However, it is this SNP that has been most widely studied based on the significance level of the initial association and the fact that it encodes a Y402H substitution in the encoded protein, which may alter function. However, evidence on the functionality of the variant is by no means conclusive and this SNP may simply mark another causal site in the gene or region, the RCA cluster, which has also been previously identified through linkage scans. The attenuated association of Y402H with AMD risk in patients of non-European ancestry, particularly Japanese individuals, in our meta-analysis provides some evidence of this, as differences in the LD structure of CFH (Supplementary Figure S4) could mean that Y402H may mark the putative causal variant in Europeans, but less well in non-European populations. One caveat to this may be that given the MAF of the risk allele in non-Europeans, we would also require a much larger sample in order to detect an effect estimate comparable with that in Europeans. Nevertheless, this finding suggests that other variants in this region are worthy of further investigation as identification of these would have important implications for use of genetic information in risk prediction of AMD, and importantly the development of therapeutic interventions that target the factor H protein and related pathways. Larger GWAS analyses in AMD, involving more dense next generation whole genome SNP arrays together with imputation of un-typed SNPs using standard methods are likely to be fruitful in refining the association signal in this region.

**Conclusions**

The Y402H variant of CFH is associated with the risk of AMD. As the evidence base on this association has matured in Europeans, the association remains
robust and in keeping with findings from the initial discovery study. There was no evidence for a true interaction of this variant with smoking though both exposures independently increase risk of AMD. The attenuated association of this variant with AMD in non-European subjects provides some evidence that Y402H could be a marker rather than a causal variant. Since the signals of association in the region span more than one gene both from GWAS and linkage studies, further genetic analyses, as well as studies that measure the CFH protein itself are required to assess whether circulating CFH itself or the product of another complement related gene in the near vicinity mediates the association with AMD risk. It is also important to study other loci identified which confer risk for AMD in a similar manner to what has been presented here, as it is for any other gene–disease associations. Together this information will be essential to convert the exciting genetic findings into potential new therapies and possible predictive tools for AMD. Further prospective studies in general populations with records of incident late AMD will be required to precisely define the absolute risk of AMD associated with the carriage of this and other SNPs which have been linked to AMD before a genotype-based predictive test could be adopted in clinical practice.

Supplementary Data
Supplementary Data are available at IJE online.

Funding
This work was supported by a Medical Research Council Biomarkers Award G0601354. R.S. was supported by a British Heart Foundation (Schillingford) Clinical Training Fellowship (FS/07/011). A.D.H. was supported by British Heart Foundation Senior Fellowship (FS 05/125). V.C. is funded by a grant from the Guide Dogs for the Blind Association. A.M., A.W. and J.Y. receive funding from the UK Department of Health’s NIHR Biomedical Research Centre at Moorfields Eye Hospital and UCL Institute of Ophthalmology. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health. The Moorfields Study in addition was funded by the Medical Research Council (Award number G0000682), the Mercer Fund (Fight for Sight UK). L.S. holds a Wellcome Trust Senior Research Fellowship. EUREYE was funded by EUQLK6-CT-1999-02094. B.H.F.W. is supported by grants from the German Research Foundation (WE1259/18-1, WE1259/19-1), the Ruth and Milton Steinbach Foundation, New York and the Alcon Research Institute, Fort Worth. Cambridge AMD Study was funded by Grant G000067 from the Medical Research Council, UK. AS and GA received funding from the NIH EY-016862. AS received funding from Macula Vision Research Foundation and the Harold Falls Professorship. A.A.B. is supported by the ANVVB, the Netherlands Macula Fund and the LSBS. The Pittsburgh Study was funded by the National Institutes of Health grant NIH R01 EY9859, Research for Prevent Blindness, New York, and the American Health Assistance Foundation and the Harold and Pauline Price Endowed Professorship (to Professor Gorin) at University of California, Los Angeles. Andrew Lotery is funded by the Macular Vision Research Foundation, The British Council for Prevention of Blindness and the Macular Disease Society. Mati Rahu is funded by the Estonian Ministry of Education and Science (target funding 0192112s02 and SF0940026s07). A.D.H. has provided non-remunerated advice to GlaxoSmithKline and London Genetics and has received honoraria for speaking at educational meetings on cardiovascular risk which have been donated in whole or part to charity. J.W. is 90% employed at GlaxoSmithKline whilst retaining a 10% appointment at the London School of Hygiene and Tropical Medicine. P.T.V.M. de.J. and B.H.F.W. have unrestricted research awards from Alcon.

Acknowledgments
We would like to acknowledge Melanie Hingorani FRCOphth for comments on study design and critical reading throughout.

Conflict of interest: None declared.
References

31. Xing C, Sivakumaran TA, Wang JJ et al. Complement factor H polymorphisms, renal phenotypes and


51 Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* 2007;447:661–78.


