

## ARTICLE OPEN



# *SLCO1B1*\*5 is protective against non-senile cataracts in cohort prescribed statins: analysis in a British-South Asian cohort

Emma F. Magavern<sup>1</sup>, David A. van Heel<sup>2</sup>, Genes & Health Research Team, Damian Smedley<sup>1</sup> and Mark J. Caulfield<sup>1</sup>✉

© The Author(s) 2023

**BACKGROUND:** Reported association between statin use and cataract risk is controversial. The *SLCO1B1* gene encodes a transport protein responsible for statin clearance. The aim of this study was to investigate a possible association between the *SLCO1B1*\*5 reduced function variant and cataract risk in statin users of South Asian ethnicity.

**METHODS:** The Genes & Health cohort consists of British-Bangladeshi and British-Pakistani participants from East London, Manchester and Bradford, UK. *SLCO1B1*\*5 genotype was assessed with the Illumina GSAMD-24v3-0-EA chip. Medication data from primary care health record linkage was used to compare those who had regularly used statins compared to those who had not. Multivariable logistic regression was used to test for association between statin use and cataracts, adjusting for population characteristics and potential confounders in 36,513 participants. Multivariable logistic regression was used to test association between *SLCO1B1*\*5 heterozygotes or homozygotes and cataracts, in subgroups having been regularly prescribed statins versus not.

**RESULTS:** Statins were prescribed to 35% (12,704) of participants (average age 41 years old, 45% male). Non-senile cataract was diagnosed in 5% (1686) of participants. An apparent association between statins and non-senile cataract (12% in statin users and 0.8% in non-statin users) was negated by inclusion of confounders. In those prescribed a statin, presence of the *SLCO1B1*\*5 genotype was independently associated with a decreased risk of non-senile cataract (OR 0.7 (CI 0.5–0.9, p 0.007)).

**CONCLUSIONS:** Our findings suggest that there is no independent association between statin use and non-senile cataract risk after adjusting for confounders. Among statin users, the *SLCO1B1*\*5 genotype is associated with a 30% risk reduction of non-senile cataracts. Stratification of on-drug cohorts by validated pharmacogenomic variants is a useful tool to support or repudiate adverse drug events in observational cohorts.

*The Pharmacogenomics Journal*; <https://doi.org/10.1038/s41397-023-00307-w>

## CLINICAL PERSPECTIVE

### What is new?

- Statin use is not independently associated with an increased risk of non-senile cataracts.
- Presence of the *SLCO1B1*\*5 genetic variant leading to increased exposure to statins is associated with a decreased risk of non-senile cataracts.
- Cross-sectional analysis of a cohort linking medication use and validated pharmacogenes with purported adverse drug reaction can support pharmacovigilance.

## WHAT ARE THE CLINICAL IMPLICATIONS?

- This study supports cardiovascular prescribers in counseling patients that statins are not independently associated with increased risk of non-senile cataracts and that in those with a *SLCO1B1*\*5 allele, use of statins is associated with decreased risk of a non-senile cataract.

- Low-income settings without ready access to cataract surgery and with high prevalence of cardiometabolic disease may prevent some non-senile cataracts by statin use in those with *SLCO1B1*\*5 allele.
- Addressing concerns about statin links with cataracts may enhance medication compliance in use of statins for primary and secondary cardiovascular prevention.

## INTRODUCTION

Statins are Hydroxymethylglutaryl-Coenzyme A (HMG-CoA) reductase inhibitors indicated in the treatment of primary and secondary prevention for cardiovascular disease as well as dyslipidemia [1]. They are among the most prescribed medications, second only to proton pump inhibitors in a study of English prescribing patterns [2]. A large USA based study showed that in 2013, 27.8% of adults over the age of 40 were prescribed a statin [3]. As a result of this widespread use, adverse drug reactions associated with statins have attracted significant attention [4].

<sup>1</sup>William Harvey Research Institute, Queen Mary University of London, London EC1M 6BQ, UK. <sup>2</sup>Blizard Institute, Queen Mary University of London, London E1 2AB, UK. A list of members and their affiliations appears in the Supplementary Information. ✉email: [m.j.caulfield@qmul.ac.uk](mailto:m.j.caulfield@qmul.ac.uk)

Received: 1 March 2023 Revised: 19 April 2023 Accepted: 21 April 2023

Published online: 23 May 2023

Cataracts are a leading cause of blindness world-wide, particularly problematic in low-income and middle-income countries with less access to surgical interventions [5]. The reported association between statin use and cataract risk is controversial and bi-directional [4]. While some large observational studies, randomized control trials, and meta-analyses have found statins to have a protective effect on cataracts, others have found an association with increased risk of cataracts, and many studies have found no significant association in either direction [4, 6–10]. Though a systematic review and meta-analysis of observational studies suggested a small increase in cataracts associated with statin use (OR: 1.11 (95% CI: 1.02–1.21);  $P = 0.017$ ), results were heterogeneous and likely impacted by residual confounding [11].

Observational studies can be confounded by the presence of cardio-metabolic risk factors for cataracts which are also indications for statins, and randomized controlled trials include selective populations and don't control for population level genetic differences. The sole study to use genetics as a tool to assess the relationship between statins and cataracts mimicked the LDL lowering effect of statins in isolation. As statins are known to have diverse mechanisms of action, including decreased inflammatory properties independent of LDL impact, this approach will model only one aspect of statin association with cataracts [12, 13].

The solute carrier organic anion transporter family member 1B1 (*SLCO1B1*) gene encodes the transporter protein OATP1B1 [14]. OATP1B1 is responsible for the active intrahepatic transport, and subsequent clearance, of statins [14]. *SLCO1B1\*5* is a polymorphism associated with increased exposure to statins as shown in pharmacokinetic studies, and increased risk of statin related adverse drug events, such as myopathy and myalgia [14–16]. *SLCO1B1\*5* prevalence varies substantially between different ethnic groups. The literature reports *SLCO1B1\*5* as present in 1% of African populations, 4% of South Asians, 12% of East Asians, 13% of Americans, and 16% of Europeans [17, 18]. South Asian ancestry populations suffer from a particularly high prevalence of cardiometabolic disease, therefore exploring statin related adverse drug reactions in this population is important [19].

Association between *SLCO1B1\*5* and cataract risk in statin users has not been characterized, and stratification by *SLCO1B1\*5* genotype in statin users and non-users offers a valuable approach to clarifying the relationship between statin use and cataracts.

The aim of our study was to use a genetic proxy for increased statin exposure by presence of the *SLCO1B1\*5* allele in a large cohort of more than 36 thousand participants to elucidate the relationship between statins and cataracts.

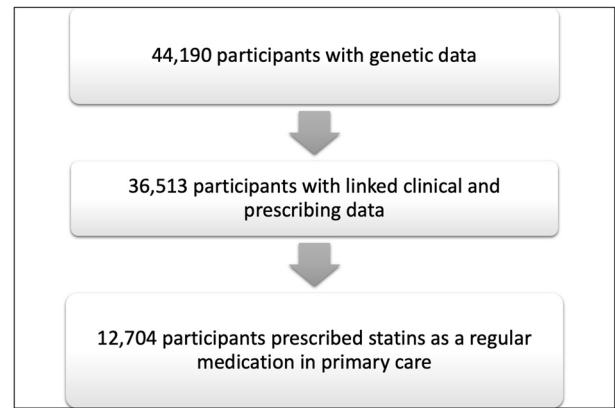
## METHODS

### The Genes & Health cohort

The Genes & Health cohort consists of British-Bangladeshi and British-Pakistani participants from East London, Manchester and Bradford [20]. The cohort has been broadly characterized in a prior publication [20]. The Genes & Health (G&H) cohort data was accessed with approval of the study executive committee. The G&H study has obtained ethical approval, 14/LO/1240, from London South East NRES Committee of the Health Research Authority, dated 16 September 2014. Volunteers provided DNA via a saliva sample and consented to link study data with electronic health records [20]. As described in prior publications, participants were genotyped using the Illumina GSAM-D-24v3-0-EA chip [21]. Human build 38 of the genome research consortium was used for this work.

### Characterization of *SLCO1B1* genotype in the G&H cohort

The *SLCO1B1\*5* genotype was extracted from the data set using PLINK 2.0 [22, 23]. The \*5 allele was defined as c.521T>C, rs4149056 (chr12:21178615 (GRCh38)). The population was in Hardy Weinberg equilibrium (HWE) for this single nucleotide polymorphisms (SNP) and there was no substantial missingness (Supplementary Table 1). The minor allele frequency (MAF) of the allele was 0.04 (Supplementary Table 1). Subsequent analysis was done in Rstudio [24].



**Fig. 1** Study cohort overview.

### Medication data from primary care

Medication use was assessed from linkage with primary care via participating clinical commissioning groups (CCGs) including Barking, Havering and Redbridge (BHR), Tower Hamlets (TH), City and Hackney, Waltham Forest (WF) and Newham (N). Our study population was constituted by  $N = 36,513$  individuals who had genetic and clinical data, including medications (Fig. 1). Participants were assigned to the ever-used statins group if they had any record of a statin on the ordinary prescription list (versus short term prescriptions) from primary care. Participants who did not meet this criterion were assigned to the never-used statin subgroup. Type of statins used included all those available in the UK: atorvastatin, simvastatin, rosuvastatin, pravastatin, fluvastatin (Fig. 2A). Pharmacokinetic data shows that there is agent specific variation in the increase in area under the curve associated with the *SLCO1B1\*5* allele and statin exposure [16, 25]. The effect is most pronounced for simvastatin and atorvastatin [26]. Pooling statins as a class exposure was chosen to optimize power but may also bias against or under-estimate agent specific signal detection.

### G&H curated phenotypes

G&H curated phenotypes were used in this analysis. These were generated using ICD10 codes, SNOMED codes, and Office of Population Censuses and Surveys (OPCS) codes from linkage with electronic health records, including Barts Health, NHS digital, Bradford teaching hospitals and primary care clinical commissioning groups (CCGs). The code used to generate the curated phenotypes is openly available in GitHub and signposted on the G&H website [27]. In summary, UK Biobank methodology for identifying first occurrence of diagnoses was followed with the addition of primary and secondary care SNOMED code data [27, 28]. Non-senile cataracts were defined by ICD10 code H26. Senile cataracts were defined by ICD10 code H25. Diabetes (DM) included E10; type 1 diabetes mellitus, E11; type 2 diabetes mellitus, E13; other specified diabetes mellitus, E14; unspecified diabetes mellitus. Dyslipidemia was defined by ICD 10 code E78. Obesity was defined by ICD10 code E66. Chronic Kidney Disease was defined by ICD10 code N18. Hypertension was defined by ICD 10 code I10. Ischemic heart disease (IHD) was defined by ICD 10 codes I21; acute myocardial infarction, I24; other acute ischemic heart diseases, and I25; chronic ischemic heart disease. Peripheral vascular disease (PVD) was defined by ICD10 code I73.

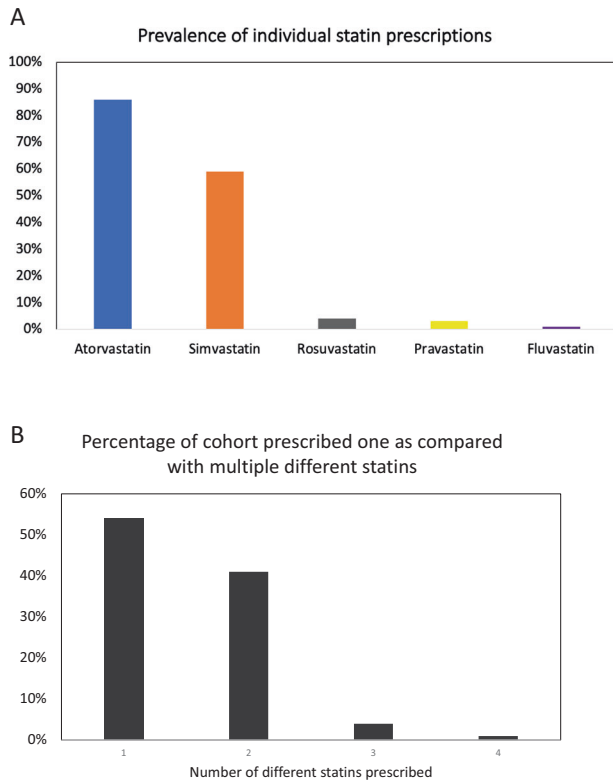
### G&H curated principal components

G&H has curated principal components as referenced in a prior publication [21]. The first two of these were used to control for population stratification in our analysis.

### Statistical methods

Multivariable logistic regression was used to test for association between statin use and cataracts, adjusting for population characteristics and potential confounders by inclusion of the listed cardio-metabolic conditions and characteristics as variables. Multivariable logistic regression was used to test association between *SLCO1B1\*5* containing diplotypes and cataracts, adjusting for age at recruitment, gender, cardiometabolic risk

factors, and two principal components in sub-groups having ever or never been prescribed statins. Fisher's exact test was used to compare cohort characteristics (Table 1 and Supplementary Table 2).



**Fig. 2** Statin prescribing prevalence and prevalence of exposure to multiple different statins. **A** Prevalance of exposure to specific statin agents within the cohort prescribed statins, in descending order of prevalence. **B** Prevalence of multiple different statin agent exposures in primary care. In the cohort prescribed statins (N 12704) it was common to have been prescribed 2 different agents within the class of statins. The odds of having been prescribed more than 1 different statin did not differ in presence or absence of *SLCO1B1\*5* allele using fisher exact test (OR 0.9 CI 0.8–1.0,  $p$  0.14).

## RESULTS

The average age at enrollment was 41 years old ( $\pm 14$  years) and 45% of participants were males. The cohort was characterized by a high prevalence of cardio-metabolic conditions including obesity (17%), diabetes (16%), hypertension (19%), and dyslipidemia (21%) (Table 1). 35% of G&H participants with linked medication data (12,704/36,513) had been prescribed a statin as an ordinary medication in primary care (Table 1, Fig. 1). Figure 2A illustrates the prevalence of individual medications within the statin class. Atorvastatin and simvastatin were the most commonly prescribed agents (Fig. 2A). Fluvastatin was not commonly prescribed. Some participants had been prescribed multiple different statins (Fig. 2B). 54% had only been prescribed 1 agents, while 41% had been prescribed 2 different statins and 4% had been prescribed 3 different types of statin. It was rare to have been prescribed 4 different types of statins (Fig. 2B).

1686 Participants (5%) had a non-senile cataract. 995 Participants (3%) had a senile cataract. 668 participants were diagnosed with both senile and non-senile cataracts. When stratified by statin use, 12% of participants who had been prescribed statins had a diagnosis of non-senile cataract, compared with 0.8% of those not prescribed a statin (Supplementary Table 2). The association between statin use and non-senile cataracts was not independent after controlling for confounding conditions associated with both CV and cataract risk, and population stratification (Table 2).

8% of the whole studied population had a *SLCO1B1\*5* allele (Fig. 3). Only 0.2% of the cohort were homozygous for the \*5 allele. There was no significant difference in *SLCO1B1\*5* genotype between those prescribed only 1 statin as compared with those prescribed more than 1 different type of statin (OR 0.9 CI 0.8–1.0,  $p$  0.14).

In the cohort who had been prescribed statins, 8% of those with a *SLCO1B1\*5* allele had a diagnosis of non-senile cataract as compared with 12% of those without a *SLCO1B1\*5* allele (Supplementary Table 2). The presence of the *SLCO1B1\*5* genotype was significantly associated with a lower risk of non-senile cataract, controlling for age at enrollment, gender, principal components, and co-morbidities (OR 0.7 CI 0.5–0.9,  $p$  0.007) (Table 3). The significant association between *SLCO1B1\*5* genotype and non-senile cataract diagnosis was not present in the subgroup not prescribed statins ( $p$  0.5). The association between *SLCO1B1\*5* genotype and senile cataract diagnosis was not significant ( $p$  0.2). Cohort characteristics stratified by statin use and by genotype are outlined in Supplementary Table 2.

**Table 1.** Cohort demographics, genotype, and disease prevalence.

Medication use	All participants (N = 36,513)	Prescribed Statin (N = 12,704)	Not Prescribed Statin (N = 23,809)	P value
Average age at enrollment (SD)	41 years old ( $\pm 14$ years)	53 years old ( $\pm 12$ years)	34 years old ( $\pm 10$ years)	<2.2 e–16
Male % (n)	45 (16,465)	58 (7395)	38 (9070)	<2.2 e–16
Obesity % (n)	17 (6140)	23 (2928)	13 (3212)	<2.2 e–16
Diabetes % (n)	16 (6024)	42 (5377)	3 (647)	<2.2 e–16
Hypertension % (n)	19 (6801)	46 (5861)	4 (940)	<2.2 e–16
Dyslipidemia % (n)	21 (7576)	54 (6831)	3 (745)	<2.2 e–16
CKD % (n)	5.9 (2155)	15 (1935)	0.9 (220)	<2.2 e–16
PVD % (n)	1.3 (477)	2 (274)	0.9 (203)	<2.2 e–16
IHD % (n)	7.5 (2741)	21 (2629)	0.5 (112)	<2.2 e–16
Cataracts (all) % (n)	5.4 (1973)	14 (1764)	0.9 (209)	<2.2 e–16
Cataracts, non-senile % (n)	4.6 (1686)	12 (1507)	0.8 (179)	<2.2 e–16
<i>SLCO1B1*5</i> homozygote or heterozygote % (n)	8.6 (3122)	8.8 (1115)	8.4 (2007)	0.3

CKD chronic kidney disease, IHD ischemic heart disease, PVD peripheral vascular disease.

## DISCUSSION

Our study found an association between statin use and a high prevalence of non-senile cataracts, but this appeared to be entirely explained by the burden of cardiometabolic risk factors (which was expectedly higher in individuals who were prescribed statins). In a large cohort of British South Asian ancestry participants who were prescribed statins, individuals carrying a *SLCO1B1\*5* allele had a 30% lower risk of developing non-senile cataract in comparison with individuals who did not carry this polymorphism known to lead to higher systemic exposure to statins (Fig. 3). As the *SLCO1B1\*5* allele has been linked with muscle related ADRs, we have undertaken this two-step analysis to show that the *SLCO1B1\*5*-statin association with decreased cataract prevalence is not likely to be due to decreased exposure to statin through lower doses or non-compliance in those with *SLCO1B1\*5*. If the protective SNP-drug effect seen were due to less statin exposure in those with a *SLCO1B1\*5* allele we would expect to see an association of statin use with cataracts independent of confounding factors in the first step of the analysis. We show that such an association is not present.

Our findings add a significant piece of the puzzle in the controversy regarding association between statins and cataracts,

**Table 2.** Association of statin use with non-senile cataracts adjusted for confounders: dyslipidemia, obesity, hypertension, CKD, PVD, diabetes, IHD, age at recruitment, sex, and two principal components.

	OR	CI	P value
<b>Statin use</b>	<b>1.0</b>	<b>0.8–1.2</b>	<b>0.97</b>
Dyslipidaemia	1.7	1.4–1.9	2.6e–10
Obesity	1.2	1.1–1.4	0.005
Hypertension	2.0	1.7–2.4	2.8e–16
CKD	1.3	1.1–1.5	0.0004
PVD	1.3	0.9–1.8	0.12
Diabetes	2.0	1.7–2.3	2e–16
IHD	0.8	0.7–0.9	0.003
Age at recruitment (Years)	1.1	1.1–1.1	<2e–16
Female sex	1.1	1.0–1.3	0.06

CKD chronic kidney disease, IHD ischemic heart disease, PVD peripheral vascular disease.

The exposure of interest for our analysis is shown in bold.

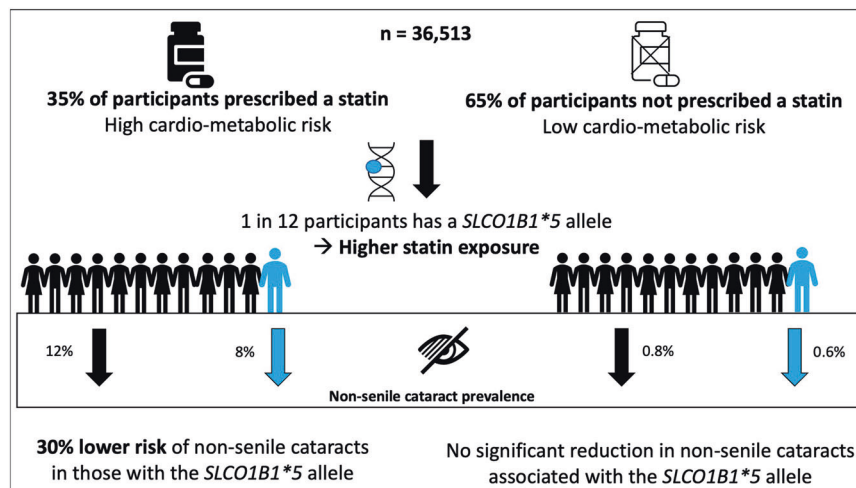
suggesting that pharmacogene association with statins are responsible for decreased non-senile cataract risk by higher exposure to statins. This is proof of concept that stratification by pharmacogene in observational on-drug cohorts can be helpful in clarifying drug association with putative adverse drug events. The G&H South Asian ancestry cohort is uniquely suited for this study due to high rates of cardiometabolic disease, and therefore a high-risk profile for non-senile cataracts.

These novel results represent the first exploration of pharmacogene-statin interaction in association with cataracts and are reassuring given the prevalence of statin prescription in the G&H community and broader population. They also account for conflicting results in the literature of bi-directional statin association with cataracts. Many prior studies have not reported on ethnic composition of the study cohort, so it is difficult to assess potential implications of our findings in interpretation of prior work. On a population level the protective effect of statins associated with *SLCO1B1\*5* would be amplified in European ancestry populations and minimal in African ancestry populations, due to diverse prevalence of *SLCO1B1\*5* in these populations, if prescription rates and co-morbidities are constant. Therefore, it seems unlikely to be accidental that the sole RCT reporting a protective effect of statins on cataracts included a 99.7% Caucasian cohort [9, 29]. It also seems quite likely that pooling studies from diverse populations without controlling for ancestry may yield conflicting results, particularly if both disease prevalence and allele prevalence vary across populations. None of the prior studies have included pharmacogenomic data.

Though the *SLCO1B1\*5* genotype was protective in association with non-senile cataracts there was no significant association with senile cataracts. This may be simply because the numbers of participants with senile cataracts were smaller and therefore this study was underpowered to find an association, if present, or may be because the pathophysiology of senile versus non-senile cataracts is different.

### Clinical implications

This study suggests that individuals with a *SLCO1B1\*5* allele who are prescribed statins are at lower risk of developing a non-senile cataract. Such individuals at high risk of non-senile cataract from cardiometabolic conditions may reduce this risk by a third if they take a statin. It thus highlights potential therapeutic opportunities in cataract prevention. It also underlines potential to use observational cohort data in conjunction with pharmacogene



**Fig. 3 Study results overview.** The *SLCO1B1\*5* allele was present in 8% of participants. 35% of the study population were prescribed statins. In the cohort prescribed statins, there was a 30% lower odds of non-senile cataracts diagnosis in those with a *SLCO1B1\*5* allele as compared with those who did not have a *SLCO1B1\*5* allele. There was no significant reduction in cataract risk in those with a *SLCO1B1\*5* not exposed to statins.

**Table 3.** Multivariable logistic regression assessing association between *SLCO1B1*\*5 genotype presence and non-senile cataract diagnosis in the on-statin cohort.

	OR	CI	P value
<b><i>SLCO1B1</i>*5</b>	<b>0.7</b>	<b>0.5–0.9</b>	<b>0.007</b>
Dyslipidaemia	1.7	1.4–2.0	2.3e–10
Obesity	1.2	1.0–1.4	0.02
Hypertension	1.9	1.6–2.2	7.1e–12
CKD	1.3	1.1–1.5	0.0006
PVD	1.2	0.8–1.6	0.3
Diabetes	2.0	1.7–2.3	<2e–16
IHD	0.8	0.7–1.0	0.01
Age at recruitment (per year)	1.1	1.1–1.1	<2e–16
Female sex	1.1	1.0–1.3	0.04

Adjusted for dyslipidaemia, obesity, hypertension, CKD, PVD, diabetes, IHD, as well as age at enrolment, gender, and two principal components. CKD chronic kidney disease, IHD ischemic heart disease, PVD peripheral vascular disease.

The exposure of interest for our analysis is shown in bold.

information to elucidate purported adverse drug reactions, an approach which had not been applied prior to this question.

Compliance to medications may be variable and dependant on numerous factors, including strength of counselling and depth of information available to patients. Certainly, knowledge of pharmacogenetic background and of dramatic reduction in risk of developing a potentially disabling condition may have a significant impact on patients' attitude toward statins, and thereby compliance. There is international consensus that pharmacogenomic testing is entering mainstream cardiovascular medicine, and therefore patients may well know if they have a *SLCO1B1*\*5 allele in the near future [26].

### Limitations

This study was not equipped to differentiate the association between individual statins and non-senile cataracts, as opposed to class effect. This was due to limited number of non-senile cataract events and unequal prescription of individual statins to individuals in the cohort. Likewise, due to unequal distribution of different statin agents and dosages as well as lack of timeline data this study was not equipped to assess effects of different dosages.

The results presented here have pooled participants who are homozygous for the *SLCO1B1*\*5 allele and those who are heterozygous. This was because of limited number of participants homozygous for the *SLCO1B1*\*5 allele in this population (only 26 participants with linked clinical and medication data were homozygous for *SLCO1B1*\*5 and had been prescribed a statin).

The dates of events were not available. Thus, we are unable to link time of statin use and time of cataract. We did not quantify time on statin prior to cataract for the same reason (due to lack of timeline data). However, the presence of the statin medication as a regular rather than short term medication assumes chronic use.

Despite these limitations, the relationship between the *SLCO1B1*\*5 genotype and reduced non-senile cataract risk only existed in the cohort who had been prescribed statins and was not apparent in the larger cohort of those not prescribed statins. This argues against a relationship between the genotype and the outcome which is not drug mediated.

### CONCLUSIONS

Our study shows an association between statin use and increased risk of non-senile cataracts is due to confounders linked with both cardiovascular/metabolic and cataract pathophysiology, in

keeping with previous research. We hereby demonstrate on a large cohort that the *SLCO1B1*\*5 genotype, known to lead to increased statin exposure, is significantly associated with decreased risk of non-senile cataracts in those taking statins. Although our novel results will need to be validated in other cohorts, they emphasize a new approach to a controversial question, utilizing a well characterized pharmacogene, and can provide two important clinical points. The first is re-assurance to patients and cardiometabolic clinicians who take and prescribe statins regularly, that this study agrees with several prior studies in concluding that statin use is not associated independently with increased risk of cataracts. The second is support for a protective association between statins and cataracts for those at high risk of non-senile cataracts due to comorbidities and exposed to higher concentration of drug. Furthermore, stratification of on-drug cohorts by validated pharmacogenomic variants is a useful tool to support or repudiate adverse drug events in observational cohorts. The population level protective effect of *SLCO1B1*\*5 in statin users, would be more pronounced in ethnic cohorts with higher prevalence of the \*5 allele, such as European ancestry populations, assuming equal prescribing prevalence and morbidity burden.

### DATA AVAILABILITY

All Genes & Health data can be accessed by application to the study access team <https://www.genesandhealth.org/research/scientists-using-genes-health-scientific-research>.

### REFERENCES

- British National Formulary. BNF: drugs: atorvastatin. National Institute for Health and Care Excellence; 2022. <https://bnf.nice.org.uk/drugs/atorvastatin/#indications-and-dose> accessed October 3, 2022.
- Audi S, Burrage DR, Lonsdale DO, Pontefract S, Coleman JJ, Hitchings AW, et al. The 'top 100' drugs and classes in England: an updated 'starter formulary' for trainee prescribers. *Br J Clin Pharmacol*. 2018;84:2562–71.
- Salami JA, Warraich H, Valero-Elizondo J, Spatz ES, Desai NR, Rana JS, et al. National trends in statin use and expenditures in the US adult population from 2002 to 2013. *JAMA Cardiol*. 2017;2:56.
- Mach F, Ray KK, Wiklund O, Corsini A, Catapano AL, Bruckert E, et al. Adverse effects of statin therapy: perception vs. the evidence – focus on glucose homeostasis, cognitive, renal and hepatic function, haemorrhagic stroke and cataract. *Eur Heart J*. 2018;39:2526–39.
- Liu Y-C, Wilkins M, Kim T, Malyugin B, Mehta JS. Cataracts. *Lancet*. 2017;390:600–12.
- Leuschen J, Mortensen EM, Frei CR, Mansi EA, Panday V, Mansi I. Association of statin use with cataracts. *JAMA Ophthalmol*. 2013;131:1427.
- Wise SJ, Nathoo NA, Etmann M, Mikelberg FS, Mancini GBJ. Statin use and risk for cataract: a nested case-control study of 2 populations in Canada and the United States. *Can J Cardiol*. 2014;30:1613–9.
- Kostis JB, Dobrzynski JM. Prevention of cataracts by statins. *J Cardiovasc Pharmacol Ther*. 2014;19:191–200.
- Bang CN, Greve AM, la Cour M, Boman K, Gohlke-Bärwolf C, Ray S, et al. Effect of randomized lipid lowering with simvastatin and ezetimibe on cataract development (from the Simvastatin and Ezetimibe in Aortic Stenosis Study). *Am J Cardiol*. 2015;116:1840–4.
- Yu S, Chu Y, Li G, Ren L, Zhang Q, Wu L. Statin use and the risk of cataracts: a systematic review and meta-analysis. *J Am Heart Assoc*. 2017;6. <https://doi.org/10.1161/JAHA.116.004180>.
- Alves C, Mendes D, Batel Marques F. Statins and risk of cataracts: a systematic review and meta-analysis of observational studies. *Cardiovasc Ther*. 2018;36:e12480.
- Ghouse J, Ahlberg G, Skov AG, Bundgaard H, Olesen MS. Association of common and rare genetic variation in the 3-hydroxy-3-methylglutaryl coenzyme A reductase gene and cataract risk. *J Am Heart Assoc*. 2022;11. <https://doi.org/10.1161/JAHA.122.025361>.
- Ridker PM, Cannon CP, Morrow D, Rifai N, Rose LM, McCabe CH, et al. C-Reactive protein levels and outcomes after statin therapy. *N Engl J Med*. 2005;352:20–28.
- Oshiro C, Mangravite L, Klein T, Altman R. PharmGKB very important pharmacogene: *SLCO1B1*. *Pharmacogenet Genomics*. 2010;20:211–6.

15. SEARCH Collaborative Group; Link E, Parish S, Armitage J, Bowman L, Heath S, et al. SLCO1B1 variants and statin-induced myopathy — a genome-wide study. *N Engl J Med*. 2008;359:789–99.
16. Pasanen MK, Fredrikson H, Neuvonen PJ, Niemi M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics of atorvastatin and rosuvastatin. *Clin Pharmacol Ther*. 2007;82:726–33.
17. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. Ensembl 2016. *Nucleic Acids Res*. 2016;44:D710–D716.
18. Ensembl. Variant: rs4149056; population genetics. ensembl.org. 2022. [https://www.ensembl.org/Homo\\_sapiens/Variation/Population?db=core;r=12:21178115-21179115;r=rs4149056;vdb=variation;vf=730080021](https://www.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=12:21178115-21179115;r=rs4149056;vdb=variation;vf=730080021). Accessed 3 Oct 2022.
19. George J, Mathur R, Shah AD, Pujades-Rodriguez M, Denaxas S, Smeeth L, et al. Ethnicity and the first diagnosis of a wide range of cardiovascular diseases: associations in a linked electronic health record cohort of 1 million patients. *PLoS ONE*. 2017;12:e0178945.
20. Finer S, Martin HC, Khan A, Hunt KA, MacLaughlin B, Ahmed Z, et al. Cohort profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int J Epidemiol*. 2020;49:20–21.
21. Chen M-H, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell*. 2020;182:1198–1213.e14.
22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–75.
23. Purcell S, Chang C. PLINK 2.0. [www.cog-genomics.org/plink/2.0/](http://www.cog-genomics.org/plink/2.0/).
24. RStudio Team. RStudio: integrated development environment for R. Boston, MA: PBC; 2022. <https://www.r-project.org/>.
25. Pasanen MK, Neuvonen M, Neuvonen PJ, Niemi M. SLCO1B1 polymorphism markedly affects the pharmacokinetics of simvastatin acid. *Pharmacogenet Genomics*. 2006;16:873–9.
26. Magavern EF, Kaski JC, Turner RM, Drexel H, Janmohamed A, Scourfield A, et al. The role of pharmacogenomics in contemporary cardiovascular therapy: a position statement from the European Society of Cardiology Working Group on Cardiovascular Pharmacotherapy. *Eur Heart J Cardiovasc Pharmacother*. 2022;8:85–99.
27. Genes & Health. Genes & Health: GeneAndHealth\_PHENOTYPES. 2022. [https://docs.google.com/spreadsheets/d/1ipwdf2j\\_owfr\\_QbkDYk1rk0TW3KtdfQYVQn-Vf-o38s/edit#gid=1517436704](https://docs.google.com/spreadsheets/d/1ipwdf2j_owfr_QbkDYk1rk0TW3KtdfQYVQn-Vf-o38s/edit#gid=1517436704). Accessed 29 Sep 2022.
28. Biobank UK. UK Biobank first occurrence of health outcomes defined by 3-character ICD10 code. 2019. [https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/first\\_occurrences\\_outcomes.pdf](https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/first_occurrences_outcomes.pdf).
29. Rossebø AB, Pedersen TR, Allen C, Boman K, Chambers J, Egstrup K, et al. Design and baseline characteristics of the Simvastatin and Ezetimibe in Aortic stenosis (SEAS) Study. *Am J Cardiol*. 2007;99:970–3.

## ACKNOWLEDGEMENTS

We acknowledge with thanks funding from Barts Biomedical Research Centre and funding from Barts Charity to support EFM. Genes & Health is/has recently been core-funded by Wellcome (WT102627, WT210561), the Medical Research Council (UK) (M009017, MR/X009777/1, MR/X009920/1), Higher Education Funding Council for England Catalyst, Barts Charity (845/1796), Health Data Research UK (for London substantive site), and research delivery support from the NHS National Institute for Health Research Clinical Research Network (North Thames). Genes & Health is/has recently been funded by Alnylam Pharmaceuticals, Genomics PLC; and a Life Sciences Industry Consortium of Astra Zeneca PLC, Bristol-Myers Squibb Company, GlaxoSmithKline Research and Development Limited, Maze Therapeutics Inc, Merck

Sharp & Dohme LLC, Novo Nordisk A/S, Pfizer Inc, Takeda Development Centre Americas Inc. We thank Social Action for Health, Centre of The Cell, members of our Community Advisory Group, and staff who have recruited and collected data from volunteers. We thank the NIHR National Biosample Centre (UK Biocentre), the Social Genetic & Developmental Psychiatry Centre (King's College London), Wellcome Sanger Institute, and Broad Institute for sample processing, genotyping, sequencing and variant annotation. We thank: Barts Health NHS Trust, NHS Clinical Commissioning Groups (City and Hackney, Waltham Forest, Tower Hamlets, Newham, Redbridge, Haverling, Barking and Dagenham), East London NHS Foundation Trust, Bradford Teaching Hospitals NHS Foundation Trust, Public Health England (especially David Wyllie), Discovery Data Service/Endeavour Health Charitable Trust (especially David Stables), Voror Health Technologies Ltd (especially Sophie Don), NHS England (for what was NHS Digital) - for GDPR-compliant data sharing backed by individual written informed consent. Most of all we thank all of the volunteers participating in Genes & Health.

## AUTHOR CONTRIBUTIONS

EFM conceived of and performed the analysis and drafted the manuscript. DVH provided expertise on the cohort and feedback on the drafted manuscript. DS and MC supervised the analysis and provided feedback on the drafted manuscript. MC provided expert conceptual supervision and secured the funding for the access application.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41397-023-00307-w>.

**Correspondence** and requests for materials should be addressed to Mark J. Caulfield.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023