

Open

SLCO1B1 Genetic Variant Associated With Statin-Induced Myopathy: A Proof-of-Concept Study Using the Clinical Practice Research Datalink

DF Carr¹, H O'Meara¹, AL Jorgensen², J Campbell³, M Hobbs³, G McCann³, T van Staa³⁻⁵ and M Pirmohamed¹

This study aimed to determine whether patients with statin-induced myopathy could be identified using the United Kingdom Clinical Practice Research Datalink, whether DNA could be obtained, and whether previously reported associations of statin myopathy with the *SLCO1B1* c.521T>C and *COQ2* rs4693075 polymorphisms could be replicated. Seventy-seven statin-induced myopathy patients (serum creatine phosphokinase (CPK) > 4× upper limit of normal (ULN)) and 372 statin-tolerant controls were identified and recruited. Multiple logistic regression analysis showed the *SLCO1B1* c.521T>C single-nucleotide polymorphism to be a significant risk factor ($P = 0.009$), with an odds ratio (OR) per variant allele of 2.06 (1.32–3.15) for all myopathy and 4.09 (2.06–8.16) for severe myopathy (CPK > 10× ULN, and/or rhabdomyolysis; $n = 23$). *COQ2* rs4693075 was not associated with myopathy. Meta-analysis showed an association between c.521C>T and simvastatin-induced myopathy, although power for other statins was limited. Our data replicate the association of *SLCO1B1* variants with statin-induced myopathy. Furthermore, we demonstrate how electronic medical records provide a time- and cost-efficient means of recruiting patients with severe adverse drug reactions for pharmacogenetic studies.

The UK Clinical Practice Research Datalink (CPRD), formerly the General Practice Research Database, is a computerized database of anonymized longitudinal medical records from primary care. In March 2011 there were more than 12 million patient records contributing more than 64 million years of prospectively collected data; the number of records is to be increased to 52 million with the transition to CPRD.¹ The information collected includes patient demographics, medical diagnoses, prescription information, referrals, and health outcomes. Although the database has been widely used in observational studies, including reports on clinical epidemiology, disease patterns, drug utilization, and outcomes research, resulting in more than 800 publications, it has never been used to obtain patient samples for biomarker analysis.

To determine whether the CPRD could be used for biomarker analysis, we focused on the pharmacogenetics of statin-induced myopathy. This was chosen as the paradigm for several reasons: first, statins (inhibitors of

5-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase) are widely used, being the cornerstone of therapy for hyperlipidemia, with proven efficacy for both primary and secondary prevention of atherosclerotic arterial disease.² Although generally well tolerated, a few patients develop muscle-related adverse effects, ranging from muscle pains without any elevation of plasma creatine kinase (CK)—a biomarker for muscle injury—to rhabdomyolysis, in which CK is elevated to > 10 times the upper limit of normal (ULN), which may be associated with renal impairment.³ A systematic review of 21 clinical trials⁴ suggested that mild muscle pain, myopathy, and rhabdomyolysis attributable to statin therapy occurred at an incidence of 190, 5, and 1.6 per 100,000 patient years, respectively.

Second, functional variation of the hepatic uptake transporter *SLCO1B1* has been implicated in statin-induced myopathy. A genome-wide association study of 85 patients with incipient (CK level >3× ULN and >5× baseline) or definite myopathy

The first two authors contributed equally to this work.

¹Department of Molecular and Clinical Pharmacology, Wolfson Centre for Personalised Medicine, Liverpool, UK; ²Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Liverpool, UK; ³Clinical Practice Research Datalink, Medicines and Healthcare Products Regulatory Agency, London, UK; ⁴Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands; ⁵London School of Hygiene & Tropical Medicine, London, UK. Correspondence: M Pirmohamed (munirp@liv.ac.uk)

Received 14 May 2013; accepted 2 August 2013; advance online publication 2 October 2013. doi:10.1038/clpt.2013.161

(muscle symptoms with CK > 10× ULN) and 90 controls who were receiving 80 mg/day simvastatin showed a strong association with a noncoding single-nucleotide polymorphism (SNP; rs4363657).⁵ This was subsequently found to be in nearly complete linkage disequilibrium with a nonsynonymous c. 521T>C SNP (rs4149056) that encodes a valine to arginine amino acid substitution at residue 147 (p.V147L) and defines the *SLCO1B1**5 allele. This variant has subsequently been associated with statin-induced myopathy in a number of other studies.^{6–8} The incidence of statin-induced myopathy has been reported to be 19% in individuals without any *5 alleles, 27% in heterozygous individuals, and 50% in *5/*5 homozygous individuals.⁸

Recent studies have also suggested that variation in the coenzyme Q2 (COQ2) homologue gene may also predispose individuals to statin-induced myopathy. Puccetti *et al.* demonstrated an association between both rosuvastatin- and atorvastatin-induced myopathy and the rs4693075 polymorphism in the *COQ2* gene.⁹ An association of another *COQ2* variant (rs4693570) and statin-induced myalgia has also been described.¹⁰ Variants of the *COQ2* loci are directly involved in CoQ deficiency,¹⁰ a postulated mechanism of statin-induced myopathy.^{11,12}

Third, in randomized controlled trials, the incidence of statin-induced myopathy is very low. For example, of 6,031 patients receiving 80 mg simvastatin, the SEARCH (Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine) study⁵ identified just 49 (0.8%) patients who had developed myopathy (defined muscle symptoms with CPK > 10× ULN). Thus, it is important to explore other methods for recruiting patients from particular electronic records, in which the use of individual drugs is usually much higher than that in trials, and it represents real-world clinical practice, in which the incidence of severe adverse reactions is higher. This article thus describes the process by which the CPRD was used to identify and recruit a cohort of statin-receiving patients with and without an increase in CPK levels in the presence or absence of muscle symptoms. We have then undertaken genotyping for *SLCO1B1* and *COQ2* variants.

RESULTS

Statin-induced myopathy case recruitment

A total of 76 cases were recruited between June 2010 and November 2011, and a total of 372 controls were recruited from the General Practice Research Database between June 2010 and April 2012. Clinical data are summarized in **Table 1**. Within the first phase of recruitment (June 2010 onward), a total of 520 potential cases of statin-induced myopathy were identified on the patient list of recruited general practice clinics. Of these, 223 (42%) were deemed suitable by the physician for inclusion. As of November 2011, 76 (34%) patients had provided adequate biological samples (blood or saliva) to the receiving laboratory. Full recruitment statistics for the 36-month study period will be subsequently reported in a future publication.

Demography

At the time of the reported event, 59 of 76 (78%) myopathy patients were receiving simvastatin; 11 (14%) were on

atorvastatin, and 6 (8%) were using other statins (cerivastatin, pravastatin, rosuvastatin, or fluvastatin). In the control cohort, 222 of 372 (60%) were receiving simvastatin at the time of recruitment, 30% were on atorvastatin, and 10% received other statins (**Table 1**). Univariate binary logistic regression analyses (**Table 1**) showed borderline statistically significant differences between cases and controls in terms of the statin type ($P = 0.075$) and previous history of type 2 diabetes ($P = 0.046$), asthma ($P = 0.080$), and hypertension ($P = 0.087$). These four variables were all adjusted for in the SNP-association analyses. There was no difference in the use of CYP3A4 inhibitors between cases and controls.

SNP analysis

Both SNPs conformed to Hardy–Weinberg equilibrium ($P > 0.0001$). The two SNPs were successfully genotyped in 99.7% (rs4693075) and 100% (rs4149056) of individuals. For logistic regression analysis of *COQ2* rs4693075, 371 controls were included. On comparing the SNP model including the *SLCO1B1* c.521T>C SNP (rs4149056) with the baseline model, the likelihood ratio test gave a significant P value (**Table 2**) both when incorporating all statin-induced myopathy cases (76 cases, 372 controls; $P = 0.005$) and when limiting the analysis to just patients with severe myopathy (23 cases, 372 controls; $P = 0.0003$). Limiting analysis to only those individuals receiving atorvastatin ($n = 121$) demonstrated no significant association between *SLCO1B1* c.521T>C (rs4149056) and risk of either myopathy ($P = 0.613$) or severe myopathy ($P = 0.507$). However, in patients receiving simvastatin ($n = 281$), statistically significant associations between c.521T>C (rs4149056) and risk of both myopathy ($P = 0.014$) and severe myopathy ($P = 0.0004$) were observed. Addition of the *COQ2* rs4693075 to the baseline model did not give a statistically significant P value for either all myopathy ($P = 0.358$) or severe myopathy ($P = 0.937$).

Binary logistic regression (**Table 2**) demonstrated a significant risk per *SLCO1B1* c.521 C allele for all myopathy cases regardless of prescribed statin ($n = 76$; odds ratio (OR) = 2.08 (1.35–3.23), $P = 0.005$). This translates to an OR of 4.32 (1.82–10.43) for risk of all myopathy for CC carriers as compared with TT carriers. For cases with severe myopathy ($n = 18$), an even higher risk per C allele was observed (OR = 4.47 (1.84–10.84)), translating to an OR of 19.98 (3.38–117.50) in CC individuals vs. that in TT individuals.

Limiting this analysis to individuals receiving simvastatin only demonstrated a similar risk to that observed for all statins, with a per-C-allele OR for all myopathy ($n = 59$) of 2.13 (1.29–3.54; $P = 0.014$). For simvastatin-induced severe myopathy ($n = 18$), the OR was 4.97 (2.16–11.43). Stratification of simvastatin patients (all myopathy) into those receiving <40 mg/day ($n = 24$) or ≥40 mg/day ($n = 35$) showed an increased risk for c.521C-allele carriers in the ≥40-mg/day group (OR = 3.23 (1.74–5.99), $P = 0.0002$), whereas no significant risk was observed in the <40-mg group (OR = 1.03 (0.45–2.36), $P > 0.05$). For severe myopathy in patients receiving ≥40 mg/day simvastatin ($n = 13$), the OR per-C-allele was 6.28 (2.38–16.60; $P = 0.0004$). In patients

Table 1 Case-control comparison of nongenetic clinical variables

Variable	Controls (n = 372)	Cases (n = 76)	P value
Statin at index			
Simvastatin	222 (60%)	59 (75%)	0.075
Atorvastatin	110 (30%)	11 (14%)	
Rosuvastatin	21 (6%)	2 (3%)	
Fluvastatin	6 (2%)	1 (1%)	
Pravastatin	12 (3%)	3 (4%)	
Ceruvastatin	1 (<1%)	0 (0%)	
Mean daily dose, mg/day (SD)*	30.6 (±15.7)	33.2 (±15.7)	0.219
Mean age, years (SD)*	71.2 (±8.7)	69.9 (±10.4)	0.222
Gender	64% M/36% F	71% M/29% F	0.238
Mean BMI*	28.5 (±4.9)	29.3 (±5.4)	0.215
Smoking status ^a			
Nonsmoker	142 (41%)	28 (39%)	0.654
Ex-smoker	150 (43%)	34 (48%)	
Smoker	57 (16%)	9 (13%)	
Comedications in 6 months before index			
Antihypertensives	304 (82%)	60 (79%)	0.628
CYP3A4 inhibitors ^b	47 (12%)	12 (16%)	0.459
Known statin interactor (non-CYP3A4 substrate) ^c	24 (6%)	7 (9%)	0.454
Oral corticosteroids	15 (4%)	3 (4%)	1.000
Occurrence in previous 6 months or 2 weeks after index			
Cramps	1 (<1%)	1 (1%)	0.311
Myocardial infarction	2 (<1%)	1 (1%)	0.428
Renal failure	8 (2%)	4 (5%)	0.129
Trauma	1 (<1%)	1 (1%)	0.311
Previous history (any time before index)			
Type 2 diabetes	93 (25%)	28 (37%)	0.046
Alcohol dependence	21 (6%)	2 (3%)	0.396
Asthma	40 (11%)	14 (18%)	0.080
Atrial fibrillation	30 (8%)	10 (13%)	0.183
Chronic obstructive pulmonary disease	27 (7%)	5 (7%)	1.000
Hypertension	246 (66%)	42 (55%)	0.087
Hyperthyroidism	6 (2%)	2 (3%)	0.628
Hypothyroidism	26 (7%)	8 (11%)	0.339

All comparison analyses were undertaken using a χ^2 test, except those variables marked *, for which an independent-samples *t*-test was applied. Values in bold indicate $P < 0.1$ where variables were carried forward for inclusion in the binary logistic regression base model.

BMI, body mass index; F, female; M, male.

^aIndicates missing data (23 tolerant, 5 myopathy). ^bCYP3A4-interacting comedications were amiodarone, cyclosporine, azole antifungals, macrolide antibiotics, protease inhibitors, and calcium channel blockers. A definitive list of drugs is given in **Supplementary Table S1** online. ^cNon-CYP3A4-interacting comedications recorded were fenofibrate, gemfibrozil, digoxin, warfarin, and nicotinic acid.

receiving <40 mg/day ($n = 5$), no significant association was observed with severe myopathy (OR = 1.84 (0.34–9.86)).

Meta-analysis

A total of seven studies, including our own, were included in the initial meta-analysis of myopathy risk for *SLCO1B1* c.521C carriage for any statin (**Figure 1**). The overall OR for myopathy risk was 2.18 (1.39–3.43). Limiting the analysis to those studies ($n = 4$) reporting genotype frequency in patients receiving simvastatin, the combined OR was marginally higher at 3.25 (1.72–6.12). Three studies reported frequencies of *SLCO1B1* in atorvastatin-receiving patients. The combined OR for myopathy was not significant at 1.54 (0.80–2.97).

DISCUSSION

The recruitment of patients with severe adverse drug reactions to pharmacogenomic studies is complicated by the facts that these reactions are rare and there is no systematic process for identifying patients. The use of electronic health records therefore represents an opportunity to undertake such studies, but, to date, electronic health records have not been used to identify patients with severe and rare phenotypes. Part of the problem here is that the phenotypes in the databases may be inadequate, leading to capture of heterogeneous patient groups and thus the identification of no or weak associations. It is well known that phenotype standardization is crucial in order to disentangle the signals from noise.¹³

To evaluate whether electronic health care record databases can be used to recruit patients with severe adverse drug reactions, we first chose CPRD as the database to undertake this feasibility study because of the quality of data contained within, which has resulted in a large number of important drug safety findings (<http://www.cprd.com>). We then chose statin-induced myopathy as the paradigm adverse drug reaction. Although statin-induced myopathy can present with many different clinical manifestations,³ and indeed previous pharmacogenetic studies have used different end points (**Figure 1**), our inclusion criteria were simple, based on an increase in CPK levels. A previous study in Scotland using electronic records used a composite definition of intolerance based on increases greater than 50% from baseline in alanine transaminase and/or 1–3× ULN in CPK, with an accompanying prescription change.⁷ This perhaps represents a milder intolerance phenotype as compared with our definition of CPK > 4× ULN. The utility of our approach is shown by the fact that over a period of 16 months, after administrative startup, we were able to recruit 76 patients with statin-induced myopathy, of whom 23 were of a more severe phenotype, denoted by CPK > 10× ULN or rhabdomyolysis. The CPRD (as of October 2009) recorded 127,268 individuals receiving a statin with a concurrent CK measurement recorded. Of those, 953 (0.75%) had CK > 4× ULN concurrent with statin prescription (T.v.S., unpublished data), an incidence comparable with that reported previously.⁴

Our results show that the rs4149056 SNP in *SLCO1B1* is associated with statin-induced myopathy. This is in accordance with previous findings,^{5–8} confirming the utility of our approach.

Table 2 Multiple logistic regression analysis of statin-induced myopathy risk and *SLCO1B1* p.V174A and *COQ2* rs4693075 genetic variants

		<i>SLCO1B1</i> p.V174A					Per C-allele OR (95% CI)	<i>COQ2</i> rs4693075				
		Genotype frequency				<i>P</i>		Genotype frequency				<i>P</i>
<i>n</i>	T/T	T/C	C/C	<i>P</i>	G/G		G/C	C/C	<i>P</i>	Per C-allele OR (95% CI)		
All statins (<i>n</i> = 448)	Tolerant ^a	372	0.70	0.27	0.03	—	—	0.40	0.45	0.15	—	—
	All myopathy	76	0.53	0.39	0.08	0.005	2.08 (1.35–3.23)	0.34	0.45	0.21	0.358	1.27 (0.90–1.81)
	Severe myopathy	23	0.35	0.44	0.21	0.0003	4.47 (1.84–10.84)	0.44	0.39	0.17	0.937	0.99 (0.54–1.82)
Simvastatin only (<i>n</i> = 281)	Tolerant	222	0.66	0.32	0.02	—	—	0.43	0.41	0.16	—	—
	All myopathy	59	0.49	0.42	0.09	0.014	2.13 (1.29–3.54)	0.37	0.42	0.21	0.643	1.20 (0.81–1.78)
	<40 mg/day	24	0.63	0.37	0.00	0.997	1.03 (0.45–2.36)	0.42	0.42	0.16	0.956	1.09 (0.61–1.96)
	≥40 mg/day	35	0.40	0.46	0.14	0.0002	3.23 (1.74–5.99)	0.34	0.43	0.23	0.543	1.32 (0.81–2.14)
	Severe myopathy	18	0.28	0.50	0.22	0.0004	4.97 (2.16–11.43)	0.42	0.41	0.17	0.975	1.08 (0.56–2.09)
	<40 mg/day	5	0.40	0.60	0.00	0.778	1.84 (0.34–9.86)	0.80	0.20	0.00	0.215	0.22 (0.03–1.74)
	≥40 mg/day	13	0.23	0.46	0.31	0.0004	6.28 (2.38–16.60)	0.23	0.54	0.23	0.516	1.57 (0.73–3.37)
Atorvastatin only (<i>n</i> = 121)	Tolerant	110	0.78	0.2	0.02	—	—	0.36	0.53	0.11	—	—
	All myopathy	11	0.64	0.36	0.00	0.613	1.91 (0.56–6.54)	0.38	0.45	0.17	0.595	1.61 (0.60–4.33)
	Severe myopathy	3	1.00	0.00	0.00	0.507	N/A	0.67	0.00	0.33	0.956	0.86 (0.13–5.70)

Statistically significant associations ($P < 0.05$) are shown in bold. Allele frequencies for tolerant and myopathy phenotypes are also shown.

CI, confidence interval; N/A, not available; OR, odds ratio.

^aDenotes one missing genotype for tolerant group for *COQ2* rs4693075 analysis.

We have shown that possession of at least one copy of the C-allele (CT/CC) is a significant risk factor for statin-induced myopathy (CK > 4× ULN), with an observed OR per C allele of 2.09 (1.27–3.45). The risk per C allele of severe myopathy (CK > 10× ULN/rhabdomyolysis; $n = 23$) was greater still, with an OR of 4.47 (1.84–10.84).

Our data replicate those of Link *et al.*,⁵ who recruited cases and controls from a randomized trial setting, showing that our cases recruited through CPRD, an observational database, are comparable. However, our cases differ from those recruited by Link *et al.* in two important aspects: (i) the observations first made by Link *et al.* were in patients receiving 80 mg/day simvastatin, whereas the mean daily dose in this study was lower (33.4 ± 19.7 mg); and (ii) only 78% of our cases with myopathy were on simvastatin, with 22% receiving other statins, including atorvastatin (in 14% of cases). Limiting the analysis to those receiving simvastatin only demonstrated an association between *SLCO1B1* c.521T>C and both all myopathy cases (OR = 1.92 (1.08–3.42)) and those with severe myopathy (OR = 4.99 (1.72–14.50)). However, the association was observed only in those patients receiving ≥40 mg/day simvastatin (Table 2), indicating the importance of dose–genotype interaction. Despite the differences, the per-C-allele OR of 4.5 for high-dose (80 mg/day) simvastatin-induced myopathy (defined as CK > 3× ULN) by Link *et al.*⁵ was highly comparable with that observed in our study for the equivalent phenotype (CK > 4× ULN) with ≥40 mg/day simvastatin (4.97; 95% confidence interval: 2.16–11.43).

Atorvastatin was the second most common drug implicated in our case group, reflecting its usage in comparison with

simvastatin. However, unlike in simvastatin-treated patients, there was no significant association between the *SLCO1B1* c.521T>C variant and either myopathy or severe myopathy in atorvastatin-treated patients. This is consistent with a previous study that showed that the association was stronger for simvastatin than for atorvastatin.⁶ Our meta-analysis of studies in Caucasians, including our data (Figure 1), also shows that there was a higher risk with simvastatin (OR = 3.25 (1.72–6.12)) than with atorvastatin (OR = 1.54 (0.80–2.97)), regardless of daily dose, in carriers of the *SLCO1B1* polymorphism. Pathophysiologically, this would be consistent with the fact that this polymorphism has the greatest effect on simvastatin (area under the curve is 221% higher in patients with the c.521CC genotype than in patients with the c.521TT genotype) but also has a smaller effect on atorvastatin (mean increase in area under the curve of 173%), and a very small, if any, effect on the other statins.¹⁴ We did not have enough patients treated with the other statins to undertake any meaningful drug-specific analyses.

Recent studies^{9,15} have shown that variation in the *COQ2* gene also predisposes an individual to statin-induced myopathy. However, we could not replicate the association with the *COQ2* rs4693075 polymorphism in our patient group. Previous studies included patients mainly receiving atorvastatin and rosuvastatin.⁹ In our study, just 13 (17%) of the statin-intolerant patients and 4 (16%) of the severe myopathy cases were receiving either atorvastatin or rosuvastatin. As such, we did not have sufficient statistical power to test this particular hypothesis. On the basis of the minor allele frequency of 0.35 observed in our atorvastatin-tolerant patients, we would require 135 cases and controls in

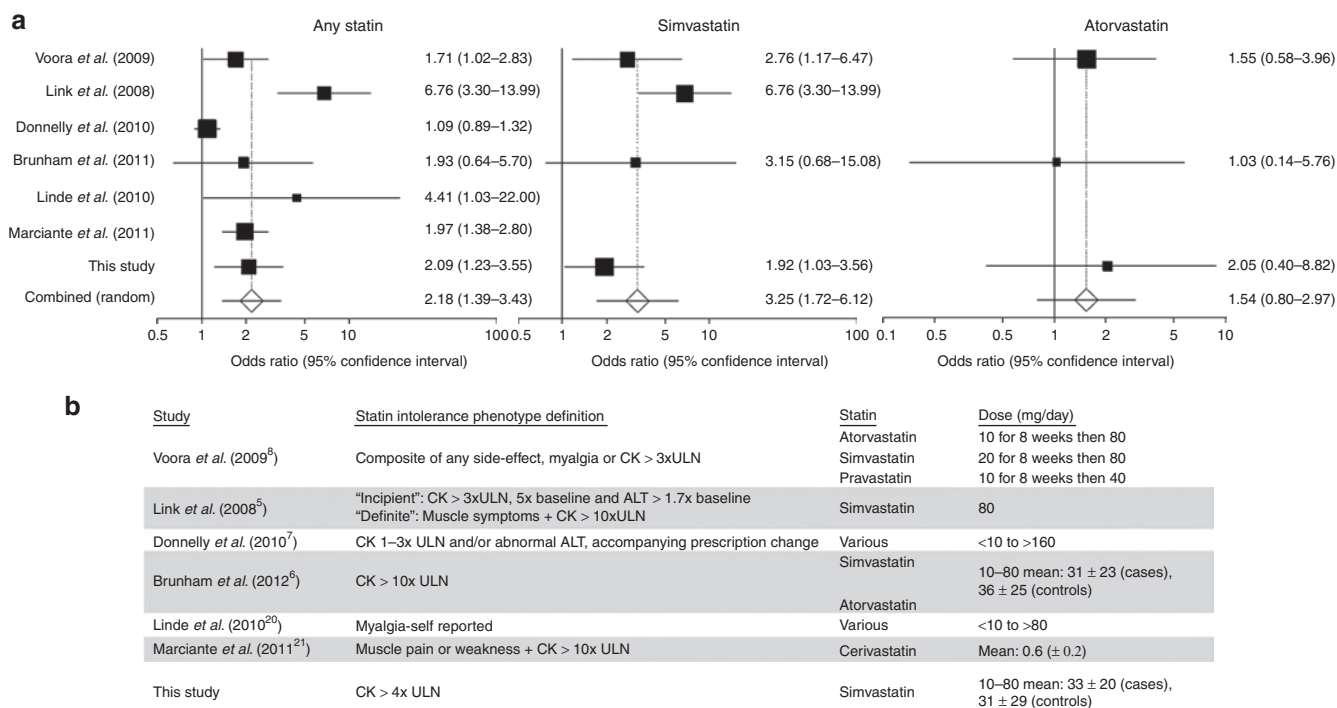


Figure 1 (a) Forest plot showing meta-analysis of all previously published studies of *SLCO1B1* c.521T>C association with statin-induced myopathy, including data from this study. Analysis was restricted to studies on Caucasian populations. (b) Details of the studies included in the analysis. ALT, alanine transaminase; ULN, upper limit of normal.

order to have a study with 80% power to detect an OR of 2 and a significance value of 0.05.

The percentage of suitable statin-induced myopathy patients, identified within general practices, from whom biological samples were ultimately received (34%) was actually better than we had expected (20–25%). A previous study using spontaneous reports under the UK yellow card scheme to obtain biological samples from patients with terodiline-induced cardiotoxicity demonstrated a success rate of 25%.¹⁶ Of course, we need to strive for higher recruitment rates for future studies, but interest in taking part in research studies by medical professionals is always tempered by the lack of time available. However, it should also be noted that a huge amount of time was saved through the more rapid identification of cases using the database, which would not have been possible through manual case-note searching.

In conclusion, there are clear time and cost benefits in using electronic patient records, such as the CPRD, for recruiting patients for genetic studies, particularly for rare phenotypes, such as statin-induced myopathy. There are also clinical benefits because the recruited patients will be from a real-world setting, and hence the effects of clinical factors such as concomitant medications can be evaluated. The electronic Medical Records and Genomics (eMERGE) network has already demonstrated the applicability of electronic medical records to identifying genomic loci associated with a population trait, white blood cell counts.¹⁷ Others have applied a similar methodology to the identification of patients for pharmacogenetic studies of drugs such as warfarin.¹⁸ In terms of the clinical utility of the genetic association between the *SLCO1B1* polymorphism and statin-induced

myopathy, there is now convincing evidence for simvastatin, but not for other statins, for which more studies are needed. A recent Clinical Pharmacogenetics Implementation Consortium guideline has made some recommendations regarding dosing and choice of statin in patients with the variant *SLCO1B1* genotype.¹⁹

METHODS

Study design

Patient identification and recruitment. From a cohort of ~600,000 patients receiving statins identified in the CPRD (<http://www.cprd.com>), a case-control design was used to identify suitable patients for the study. Participation was restricted to Caucasians ≥18 years of age and with the first-ever statin prescription at least 1 year after the start of CPRD data collection. Potential cases were selected from the database if they discontinued their implicated statin therapy and demonstrated an increase in CPK > 4x ULN.

Potential controls were selected if they had been receiving statins for at least 3 months with no previous above-normal serum CPK measurements. General practitioners were contacted with a list of potential cases and/or controls identified from their practices. After being given the opportunity to decline involvement, they were first asked to review the list and remove any patients they considered unsuitable. They were then asked to contact suitable patients by letter requesting participation. Consenting case patients were randomized and invited to provide either a saliva sample (by post) or a blood sample (by visit to the practice). Controls provided only blood samples. All samples were then forwarded to the University of Liverpool for processing. To preserve anonymity, patient and practice identifier codes were used throughout the recruitment process, and all patient contact was through the general practitioner only.

Study approval. Ethical approval was obtained from the National Research Ethics Committee North West 2—Liverpool Central, and approval to use the CPRD data was obtained from the Independent Scientific Advisory

Committee at the Medicines and Healthcare Products Regulatory Agency. In addition, site-specific approval to contact the GP practices was obtained for each of the 138 primary-care trusts across the United Kingdom. Local informed consent was obtained from all study subjects or their guardians in accordance with the Declaration of Helsinki.

DNA extraction and genotyping

Genomic DNA was extracted from 5 ml of whole blood or 2 ml of saliva (collected using the Oragene DNA Sampling kit, DNA Genotek, Ontario, Canada) using the Chemagic Magnetic Module 1 system per the manufacturer's protocol (Chemagen Biopolymer-Technologie, Baesweiler, Germany).

A total of 448 individuals were genotyped for the rs4149056 SNP in *SLCO1B1* and rs4693075 in *COQ2* using commercially available TaqMan real-time PCR SNP genotyping assays with 1× Genotyping Master Mix (both from Applied Biosystems, Carlsbad, CA). Subsequently, 20 ng of genomic DNA per reaction was genotyped according to the manufacturer's protocol using an ABI 7900HT real-time PCR system (Applied Biosystems, Carlsbad, CA); Ten percent of the samples were run in duplicate to ensure concordance of genotype.

Statistical analysis

A univariate analysis of association between all nongenetic variables considered to be of *a priori* interest and case-control status was first undertaken. The χ^2 test was used for categorical variables and Student's *t*-test for continuous variables. Any variable demonstrating a statistically significant association ($P < 0.10$) was carried forward and adjusted for in the SNP association analyses.

To test for association with each SNP in turn, two multiple logistic regression models were fitted. The first (the baseline model) included all univariately significant ($P < 0.10$) nongenetic variables. The second (the SNP model) was the same but also included a covariate to represent the SNP (either rs4149056 or rs4693075). An additive effect of the variant allele was assumed. Homozygote wild type was coded as "0," heterozygote as "1," and homozygote variant allele as "2."

To test for association with the SNP, the likelihood ratio test was used to compare the SNP model with the baseline model. A *P* value < 0.025 (0.05 corrected for two tests of associations using the Bonferroni approach) was assumed to represent statistical significance of the SNP.

Sensitivity analyses were undertaken by separately limiting cases to those classified as having either plasma CK $> 10 \times$ ULN or rhabdomyolysis ($n = 23$; termed "severe myopathy"). All statistical analyses were undertaken using SPSS version 17.0.

Meta-analysis

A search of PubMed (<http://www.ncbi.nlm.nih.gov/pubmed> accessed January 2012) using the search terms "SLCO1B1" and "statin" yielded 108 publications, of which 96 were original research articles. Inspection of titles and abstracts identified six research articles that defined the frequency of the SLCO1B1 rs4149056 polymorphism in an entirely, or predominantly, Caucasian population of statin-induced myopathy. Studies were included regardless of the suspect statin investigated, dose, and myopathy phenotype observed (as described in [Figure 1b](#)). Due to the high degree of heterogeneity among the included studies ($I^2 = 84.1\%$), a DerSimonian-Laird random effects model was applied to the meta-analysis in StatsDirect version 2.6.8 (StatsDirect, Altrincham, UK).

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at <http://www.nature.com/cpt>

ACKNOWLEDGMENTS

This work was funded by a grant from the e-Health Initiative funded jointly by the Medical Research Council (reference: MC_qA137929), the Wellcome Trust, the Engineering and Physical Sciences Research Council, and The Economic and Social Research Council. M.P. is a National Institute for Health Research senior investigator. The authors thank the contributors to the CPRD, the recruiting general practice physicians, and patients who participated in the study.

AUTHOR CONTRIBUTIONS

D.F.C. and H.O. wrote the manuscript. T.v.S., G.M., and M.P. designed the research. D.F.C., J.C., M.H., and H.O. performed the research. D.F.C., H.O., A.L.J., G.M., T.v.S., and M.P. analyzed the data.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

- ✓ Genetic factors may predispose certain individuals to statin-induced myopathy. Many such adverse drug reactions are rare, and large, costly prospective clinical studies are required to recruit sufficient cohorts for pharmacogenetic analysis purposes.

WHAT QUESTION DOES THIS STUDY ADDRESS?

- ✓ Using statin-induced myopathy as a paradigm, this study assessed how electronic patient medical records held in the CPRD could be used to identify, recruit, and obtain DNA samples from adverse drug reaction cases and controls. Replication of the *SLCO1B1* c.521T>C polymorphism association with statin-induced myopathy was used to validate this recruitment protocol for pharmacogenetic studies.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

- ✓ The use of the CPRD is a more cost- and time-effective method for the recruitment of patients for pharmacogenetic studies in comparison with traditional prospective recruitment methods.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

- ✓ Use of electronic medical records such as those contained in the CPRD could prove valuable in identifying and recruiting patients with a number of rare adverse drug reactions for pharmacogenetics studies and facilitate future identification of predisposing genetic biomarkers.

© 2013 American Society for Clinical Pharmacology and Therapeutics

1. Williams, E.K., van Staa, T., Puri, S. & Eaton, S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther. Adv. Drug Saf.* **3**, 89–99 (2012).
2. Sadowitz, B., Maier, K.G. & Gahtan, V. Basic science review: Statin therapy—Part I: The pleiotropic effects of statins in cardiovascular disease. *Vasc. Endovascular Surg.* **44**, 241–251 (2010).
3. Abd, T.T. & Jacobson, T.A. Statin-induced myopathy: a review and update. *Expert Opin. Drug Saf.* **10**, 373–387 (2011).
4. Law, M. & Rudnicka, A.R. Statin safety: a systematic review. *Am. J. Cardiol.* **97**, 52C–60C (2006).
5. Link, E. *et al.* SLCO1B1 variants and statin-induced myopathy—a genome-wide study. *N. Engl. J. Med.* **359**, 789–799 (2008).
6. Brunham, L.R. *et al.* Differential effect of the rs4149056 variant in SLCO1B1 on myopathy associated with simvastatin and atorvastatin. *Pharmacogenomics J.* **12**, 233–237 (2012).
7. Donnelly, L.A. *et al.* Common nonsynonymous substitutions in SLCO1B1 predispose to statin intolerance in routinely treated individuals with type 2 diabetes: a go-DARTS study. *Clin. Pharmacol. Ther.* **89**, 210–216 (2011).

8. Voora, D. *et al.* The SLCO1B1*5 genetic variant is associated with statin-induced side effects. *J. Am. Coll. Cardiol.* **54**, 1609–1616 (2009).
9. Puccetti, L., Ciani, F. & Auteri, A. Genetic involvement in statins induced myopathy. Preliminary data from an observational case-control study. *Atherosclerosis* **211**, 28–29 (2010).
10. Rúaño, G. *et al.* Mechanisms of statin-induced myalgia assessed by physiogenomic associations. *Atherosclerosis* **218**, 451–456 (2011).
11. Golomb, B.A. & Evans, M.A. Statin adverse effects: a review of the literature and evidence for a mitochondrial mechanism. *Am. J. Cardiovasc. Drugs* **8**, 373–418 (2008).
12. Marcoff, L. & Thompson, P.D. The role of coenzyme Q10 in statin-associated myopathy: a systematic review. *J. Am. Coll. Cardiol.* **49**, 2231–2237 (2007).
13. Pirmohamed, M., Aithal, G.P., Behr, E., Daly, A. & Roden, D. The phenotype standardization project: improving pharmacogenetic studies of serious adverse drug reactions. *Clin. Pharmacol. Ther.* **89**, 784–785 (2011).
14. Niemi, M. Transporter pharmacogenetics and statin toxicity. *Clin. Pharmacol. Ther.* **87**, 130–133 (2010).
15. Puccetti, L., Scarpini, F., Cappellone, R. & Auteri, A. Genetic influence in statin intolerance. *Clin. Pharmacol. Ther.* **90**, 365 (2011).
16. Ford, G.A., Wood, S.M. & Daly, A.K. CYP2D6 and CYP2C19 genotypes of patients with terodiline cardiotoxicity identified through the yellow card system. *Br. J. Clin. Pharmacol.* **50**, 77–80 (2000).
17. Crosslin, D.R. *et al.*; Electronic Medical Records and Genomics (eMERGE) Network. Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. *Hum. Genet.* **131**, 639–652 (2012).
18. Xu, H. *et al.* Facilitating pharmacogenetic studies using electronic health records and natural-language processing: a case study of warfarin. *J. Am. Med. Inform. Assoc.* **18**, 387–391 (2011).
19. Wilke, R.A. *et al.*; Clinical Pharmacogenomics Implementation Consortium (CPIC). The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy. *Clin. Pharmacol. Ther.* **92**, 112–117 (2012).
20. Linde, R., Peng, L., Desai, M. & Feldman, D. The role of vitamin D and SLCO1B1*5 gene polymorphism in statin-associated myalgias. *Dermatoendocrinol.* **2**, 77–84 (2010).
21. Marcianti, K.D. *et al.*; Cerivastatin, genetic variants, and the risk of rhabdomyolysis. *Pharmacogenet. Genomics.* **21**, 280–288 (2011).



This work is licensed under a Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>