DOCTOR OF PHILOSOPHY

Genetic factors in statin intolerance

Siddiqui, Moneeza Kalhan

Award date:
2016

Genetic factors in statin intolerance

Moneeza Kalhan Siddiqui M.Sc., M.P.H.

For the degree of Doctor of Philosophy

University of Dundee

November 2016

Table of Contents

List of Tables

List of figures

List of Abbreviations

Abbreviation	Full form
Arg	Arginine (amino acid)
Asp	Aspartic acid (amino acid)
CAD	Coronary Artery Disease
CK	Creatine phosphokinase (enzyme)
CKM	Creatine Kinase Muscle (gene)
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular Disease
eQTL	expression Quantitatvie Trait Loci
GAUSS3	Goal Achievement after Utilizing an anti-PCSK9 Antibody in Statin -Intolerant Subjects -3
Glu	Glutamic acid (amino acid)
Gly	Glycine (amino acid)
GoDARTS	Genetics of Diabetes Audit and Research Tayside, Scotland
GRO	General Register Office
HDL	High Density Lipoprotein
HIC	Health Informatics Centre
His	Histidine (amino acid)
HLA	Human Leukocyte Antigen
HMG-CoA	Hydroxy-3-methylglutaryl-CoA
HMGCR	Hydroxy-3-methylglutaryl-CoA Reductase
JUPITER	Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin
LD	Linkage disequilibrium
LDL	Low Density Lipoprotein
LILR	Leukocyte Immunoglobulin-Like Receptor (gene family)
LILRA	Leukocyte Immunoglobulin-Like Receptor - subfamily A (gene subfamily)
LILRB	Leukocyte Immunoglobulin-Like Receptor - subfamily B (gene subfamily)
PCSK9	Proprotein Convertase Subtilisin/Kexin type 3 (enzyme)
PREDICTION-ADR	Prediction-Adverse Drug Reactions
PSD	Practioner Services Division
RCT	Randomized Controlled Trial
SAMS	Statin-Associated Muscle Symptoms
SF	Statin Failure
SI	Statin intolerance
SIM	Statin-Induced Myositis
SNP	Single Nucleotide Polymorphism
T2D	Type 2 Diabetes
TLR	Toll-Like Receptor

ULN	Upper Limit of Normal
UoD	University of Dundee
UoL	University of Liverpool
WTCCC	Wellcome Trust Case Control Consortium
XULN	Times Upper Limit of Normal
Abbreviation	Full form
Arg	Arginine (amino acid)
Asp	Aspartic acid (amino acid)
CAD	Coronary Artery Disease
CK	Creatine phosphokinase (enzyme)
CKM	Creatine Kinase Muscle (gene)
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular Disease
eQTL	expression Quantitatvie Trait Loci
GAUSS3	Goal Achievement after Utilizing an anti-PCSK9 Antibody in Statin-Intolerant Subjects -3
Glu	Glutamic acid (amino acid)
Gly	Glycine (amino acid)
GoDARTS	Genetics of Diabetes Audit and Research Tayside, Scotland
GRO	General Register Office
HDL	High Density Lipoprotein
HIC	Health Informatics Centre
His	Histidine (amino acid)
HLA	Human Leukocyte Antigen
HMG-CoA	Hydroxy-3-methylglutaryl-CoA
HMGCR	Hydroxy-3-methylglutaryl-CoA Reductase
JUPITER	Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin
LD	Linkage disequilibrium
LDL	Low Density Lipoprotein
LILR	Leukocyte Immunoglobulin-Like Receptor (gene family)
LILRA	Leukocyte Immunoglobulin-Like Receptor - subfamily A (gene subfamily)
LILRB	Leukocyte Immunoglobulin-Like Receptor subfamily B (gene subfamily)
PCSK9	Proprotein Convertase Subtilisin/Kexin type 3 (enzyme)
PREDICTION-ADR	Prediction-Adverse Drug Reactions
PSD	Practioner Services Division
RCT	Randomized Controlled Trial
SAMS	Statin-Associated Muscle Symptoms
SF	Statin Failure
SI	Statin intolerance
SIM	Statin-Induced Myositis
SNP	Single Nucleotide Polymorphism
T2D	Type 2 Diabetes

TLR	Toll-Like Receptor
ULN	Upper Limit of Normal
UoD	University of Dundee
UoL	University of Liverpool
WTCCC	Wellcome Trust Case Control Consortium
XULN	Times Upper Limit of Normal

Declaration of candidate

I declare that this thesis is based on results obtained from investigations which I have carried out in the Division of Molecular and Clinical Medicine, University of Dundee, between October 2013 and July 2016 using funding provided by the PREDICTION-ADR consortium (European Commission, FP7). I declare that the entire thesis is my own composition. Any work other than my own is clearly stated in the test and acknowledged with reference to the relevant investigators or contributors. This thesis has never been presented previously, in whole or in part, for the award of any higher degree. I have consulted all the references cited within the text of this thesis.

Signed: _____

Moneeza Kalhan Siddiqui
Date: 22nd November 2016

Declaration of the supervisor

I certify that Moneeza Kalhan Siddiqui has completed the equivalent of 9 terms of experimental research and that she has fulfilled the conditions of the University of Dundee, so that she is qualified to submit this thesis in application for the degree of Doctor of Philosophy.

Signed: _____

Professor Colin N A Palmer
Date: 22nd November 2016

Dedication

For my grandmother – her omnipresent love, playfulness and humour stay with me everyday.
For my father – my safety net.
For my uncle – for pushing me to take flight. Twice.
For my grandfather – who always carefully answered “this?” and taught me how to think
For my mother – who truly deserves all the credit.

Acknowledgements

This PhD would have been unimaginably different without the mentorship of Prof. Colin Palmer. His enthusiasm for science and joy in knowledge is pure, infectious and inspiring. Thank you!
The PREDICTION-ADR consortium for this opportunity.
Dr. Dan Chasman for his prompt and engaging collaboration.
My Thesis Monitoring Committee: Professors Alexander Jovanovic, Mike Ashford and Ewan Pearson for their support. Dr. Alex Doney and Dr. Jacob George for their help and insight in this project.
Caroline Glen, co-conspirator par excellence. I am ever grateful for your help.
My wonderful colleagues, Abirami, Alison, Mike, Adem, Cyrielle, Myra, Maggie, Enrique, Laura, Fiona Carr, Roger and Karen for making the office so convivial.
Dr. Louise Donnelly for her indispensable guidance, Phil Appleby and Abirami Veluchamy for their work on the genetic data, making it so accessible.
I would especially like to thank the following people:
Adem for making me endless cups of coffee.
Cyrielle for her excellent company. And for being there when I fell.
Polly, my trusty co-chef for keeping me on track by always being more spaced out. And for feeding me the best cakes, cookies, pies and bread in the whole wide world. To her conservatory, thank you for the naps.
Abigail and Steven for their invaluable friendship and helping me find my feet in Dundee
Lauren Tebay, Olivia, Daniel and Sharon for all the food and irreplaceable warmth.
Elya, for being the louder voice in my head and the staunchest of friends.
Lauren, for ruining me for any future flat-mates.
Elya, Lauren, Ami, Serra and Mo – for the visits, holidays, meals, commiseration and good times. To my friends from home, Stef, Shreya, Esther, Shaheen… it's been such a long journey and I've had the best companions for it.

And finally, dear, dear Aijaz for being the first responder in Dundee.

Abstract

Background: There are approximately 12 million statin users in the United Kingdom. Reports of statin intolerance occurs between 7 and 29% of users, manifesting as muscle ache, fatigue or more seriously, muscle breakdown leading to myopathy. Creatine phosphokinase (CK) levels are used as a biomarker of statin-induced muscle damage. Non-adherence or discontinuation of therapy is a common result of intolerance and can result in negative cardiovascular disease-related outcomes.

Aim: This thesis attempts to identify trends in record-linked medical data in a Scottish Caucasian cohort (GoDARTS) that best represent statin intolerance in order to study associated genetic factors.

Methods: Prescribing trends such as switching or discontinuation of statin therapy were examined, and thresholds created to select true cases of intolerance. Information on CK levels was gathered from medical records and appropriate test results were utilized. Genotypic data was gathered for the variants and genetic regions of interest using a variety of methods including chip-based genotyping followed by imputation, TAQMAN genotyping, and exome sequencing. Subsequently hypothesis-based association analyses were conducted, including linear and logistic regressions, followed by meta-analyses, regional GWAS followed by a regional meta –analysis.

Results: The phenotypes of statin intolerance were validated both internally and externally. Previously reported missense variants in LILRB5 (Asp247Gly) and CKM (Glu83Gly) were replicated and shown to be associated with CK levels irrespective of statin usage in the GoDARTS cohort and the clinical trial setting (JUPITER). Further, the CKM variant was also associated with inducibility of CK at times of tissue injury. The Asp247 genotype in LILRB5 was associated with increased risk of statin intolerance, and was replicated in associations with non-compliance to statin therapy and the development of myalgia in the JUPITER trial. The association with myalgia showed a stratified effect based on therapy (statin or placebo), with those on placebo showing the genotype effect. Further, the variant was also associated with increased risk of statin-induced myositis, cases of which had been clinically adjudicated and exome sequenced for the PREDICTION-ADR consortium. Further exploration of the LILR gene region showed an association with variants in LILRB2 (His20Arg and Val235Met) which were in strong LD with each other but were not in linkage with the variant in LILRB5. Stratified analysis revealed that the risk for carriers of the LILRB2 variants was increased depending on the genotype carried at the LILRB5 variant.

Conclusions: This study characterizes novel genetic factors associated with statin intolerance impacting adherence. The findings point to the immunomodulatory effects of statins. The results suggest that true statin-induced myalgia and non-specific myalgia are distinct, with a possible role for the immune system in their development. The findings encourage further investigation into the immune-physiology of statin-induced muscle damage and identifies genetically susceptible groups who are more likely to be statin intolerant.

1 Introduction

1.1 Cholesterol in the etiology of cardiovascular diseases

Cardiovascular diseases (CVD) affect the heart and blood vessels, and are a huge healthcare burden. They are the number one cause of deaths globally, accounting for over 30% of all deaths, according to the World Health Organization (WHO) (1). In the United Kingdom, the National Health Service (NHS) estimates that CVD are the cause of over 160,000 deaths annually (2). CVD have many etiologies; chief amongst them is hypercholesterolemia (high amounts of circulating cholesterol). The transport of cholesterol from the liver (where it is synthesized) to the cells (where it is required as an essential structural and functional component) is carried out by Low Density Lipoprotein cholesterol (LDL-C). Excess circulating cholesterol is deposited along the walls of arteries leading to the formation of plaques or atheroma and the subsequent hardening and narrowing of the walls, this condition is known as atherosclerosis. A detailed schematic showing the molecular mechanism of plaque formation as demonstrated by Choudhury *et al.* is seen in Figure 1.1 (3). Depending on its location, plaque formation can restrict blood flow to surrounding organs and negatively impact their functioning. If a plaque ruptures it can cause a blood clot, if such a clot were to arrest the supply of blood to the heart it can lead to a myocardial infarction (MI), if blood supply to the brain is cut off it can lead to a stroke or a mini-stroke, also known as Transient Ischemic Attacks (TIA). In addition, coronary heart disease (CHD) and peripheral artery disease (PAD) are also caused by hypercholesterolemia.

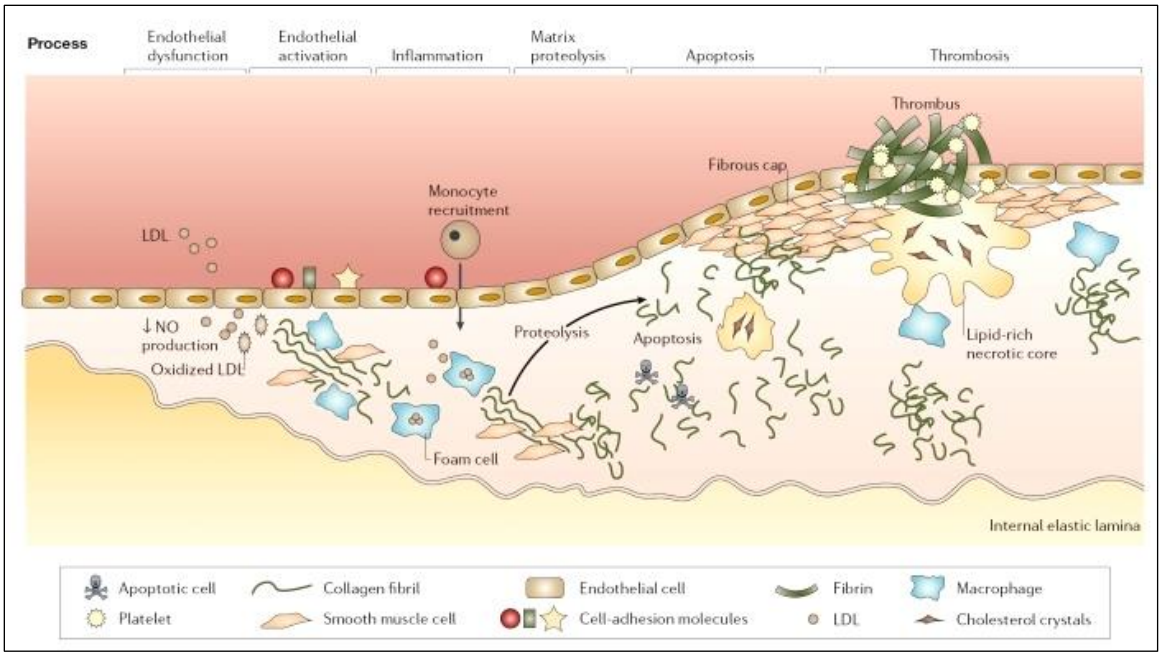


Figure 1.1 Figure from Choudhury *et al.* showing the molecular mechanism of plaque formation in the artery. © 2004 Nature Publishing Group Choudhury, R. P. *et al.* Molecular, cellular and functional imaging of atherothrombosis. Nature Reviews Drug Discovery 3, 914 (2004). All rights reserved.

1.2 Statins: mechanism of action

Cholesterol is produced by the mevalonate pathway. Hydroxy-3-methylglutaryl-CoA (HMG-CoA) is converted to mevalonate in the cytoplasm, in a reaction catalyzed by HMG-CoA-reductase (Figure 1.2). This is the rate-limiting step in the biosynthesis of cholesterol. Statins act by inhibiting HMG-CoA-reductase, thereby forestalling the production of cholesterol (4–7).

After the discovery of the role of LDL-C in increasing the risk of CHD, and the role of HMG-CoA reductase in its synthesis, efforts began to find a successful HMG-CoA reductase inhibitor. In 1971, Endo *et al.* produced a compound known as ML-236B or compactin which was to become the first successful inhibitor (8); it was later called Mevastatin (7,8). After successful demonstrations of the compound's ability to lower LDL-C in plant and animal models, in 1990 analogues Lovastatin, Simvastatin and Pravastatin were produced by pharmaceutical companies (4,7).

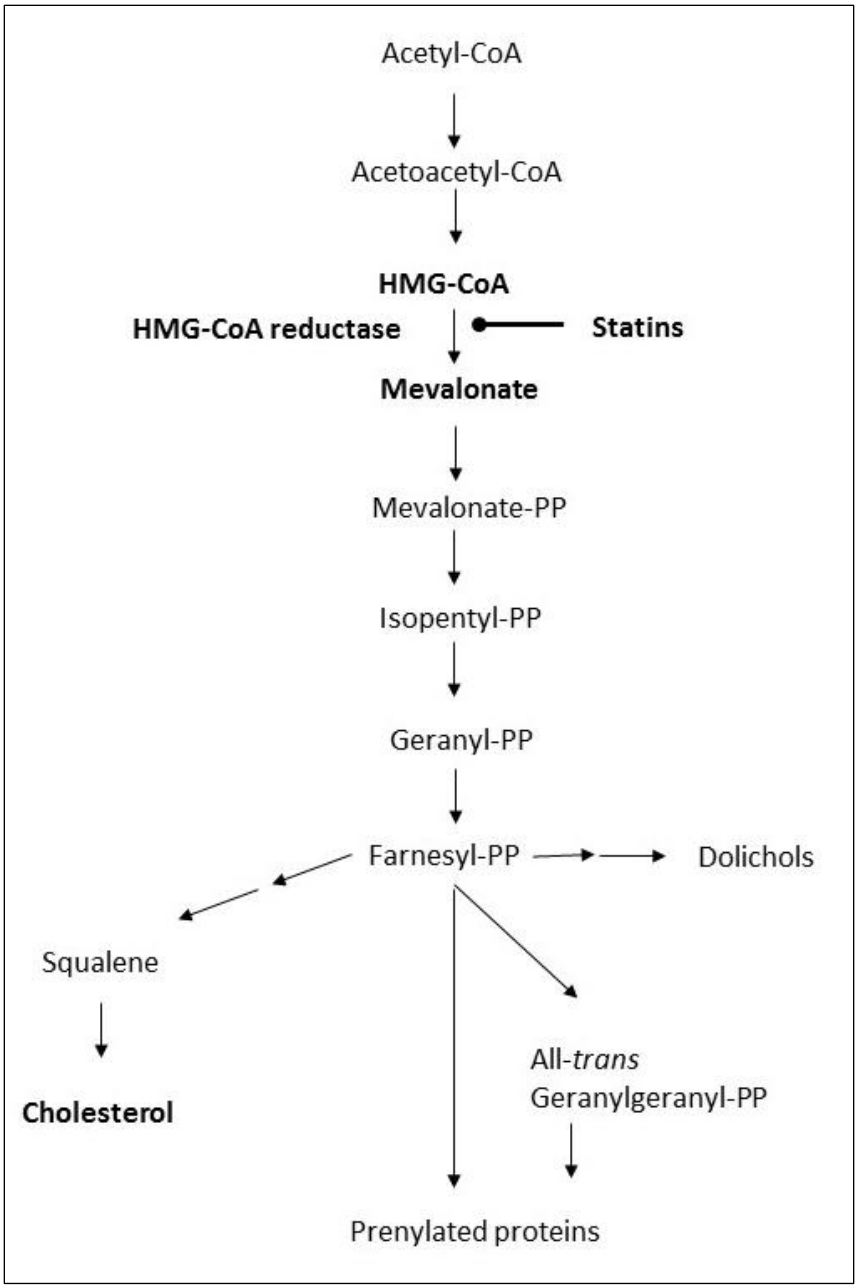


Figure 1.2 The mammalian mevalonate pathway. PP – pyrophosphate. Adapted from Corsini *et al.* by M. Schachter (5,9)

1.3 Pharmacoepidemiology of statins

Presently, the types of statins available in the United Kingdom are Simvastatin, Rosuvastatin, Atorvastatin, Pravastatin and Fluvastatin (10). Cerivastatin has been discontinued due to the high incidence of rhabdomyolysis and death associated with its use (11,12). Their relative efficacies in LDL reduction are shown in Table 1.1 (13,14)

Table 1.1 Statins and associated cholesterol and LDL reduction

Statin drug (mg)	Year of introduction	Approved daily dose (mg)	Change in total cholesterol (%)	Change in LDL cholesterol (%)
Simvastatin (10)	1991	10-40	-22	-27
Atorvastatin (10)	1996	10-80	-27	-34
Pravastatin (10)	1991	20-80	-11	-13.5
Fluvastatin (10)	1993	20-80	-5.5	-6.75
Rosuvastatin (10)	2003	10-40	-88	-10.8
Cerivastatin (0.2)	1997 *	0.2-0.3	-22	-27

Recalled in 2001
Source: adapted from Maron et. al.

The United Kingdom is already one of the largest users of statins worldwide (15), and with revised National Institute for Health and Care Excellence (NICE) guidelines an estimated 12 million people will now be prescribed statins by 2020 (16,17). The Organisation for Economic Co-operation and Development (OECD) estimates that statin use in the United Kingdom is over 40% higher than the OECD average, which includes countries such as France, Germany, Canada and the United States (18). NICE recommends statin use for people with pre-existing heart disease, people who have familial hypercholesterolemia and those who are presently healthy but are at a 20% or higher risk of developing heart disease. Risk of heart disease is determined using the QRISK2 tool (https://www.qrisk.org/) that makes the assessment based on known risk factors such as age, sex, body mass index (BMI), ethnicity, family history, chronic diseases such as diabetes, kidney disease, high blood pressure, atrial fibrillation or rheumatoid arthritis, an individual's blood cholesterol levels and blood pressure (19). At present statins are the first choice of medication in cholesterol control for the prevention and management of CVD (20).

1.4 Pharmacokinetics of statins

Transport of statins

For statins to effectively reduce hepatic production of cholesterol, they have to be transported from the portal blood into the liver (where the cytochrome (CYP) P450 enzyme systems can metabolize them via the hepatocellular membrane's phospholipid bilayer, as seen in Figure 1.3. This transport primarily occurs through the actions of the organic anion-transporting polypeptide (OATP1B1) influx transporter, which is expressed on the basolateral membrane of human hepatocytes (21,22). The gene *SLCO1B1* on chromosome 12 encodes OATP1B1. Disruption in the functioning of the transporter can result in decreased hepatic uptake, and greater systemic plasma concentrations of statins (23); this increased exposure to statins, could in turn increase the risk of development of adverse reactions.

Members of the adenosine triphosphate –ATP- binding cassette (ABC) protein family mediate the efflux of statins and their metabolites out of the liver (22,24,25). Reduced efflux of statins increases their plasma concentration, which could in turn increase the risk of development of adverse reactions.

A representation of the generic pharmacokinetics of statins is presented in Figure 1.3 (21). It shows the entry of the drug into the liver via the intestines, CYP450 enzymes act on the drug in both locations but the vast majority of metabolism is hepatic. Active and inactive metabolites are released in both locations. From the liver these are transported out by efflux transporters (ABC family members) via bile or via the bloodstream. From the bloodstream, they are cleared through the kidneys by members of the same protein family.

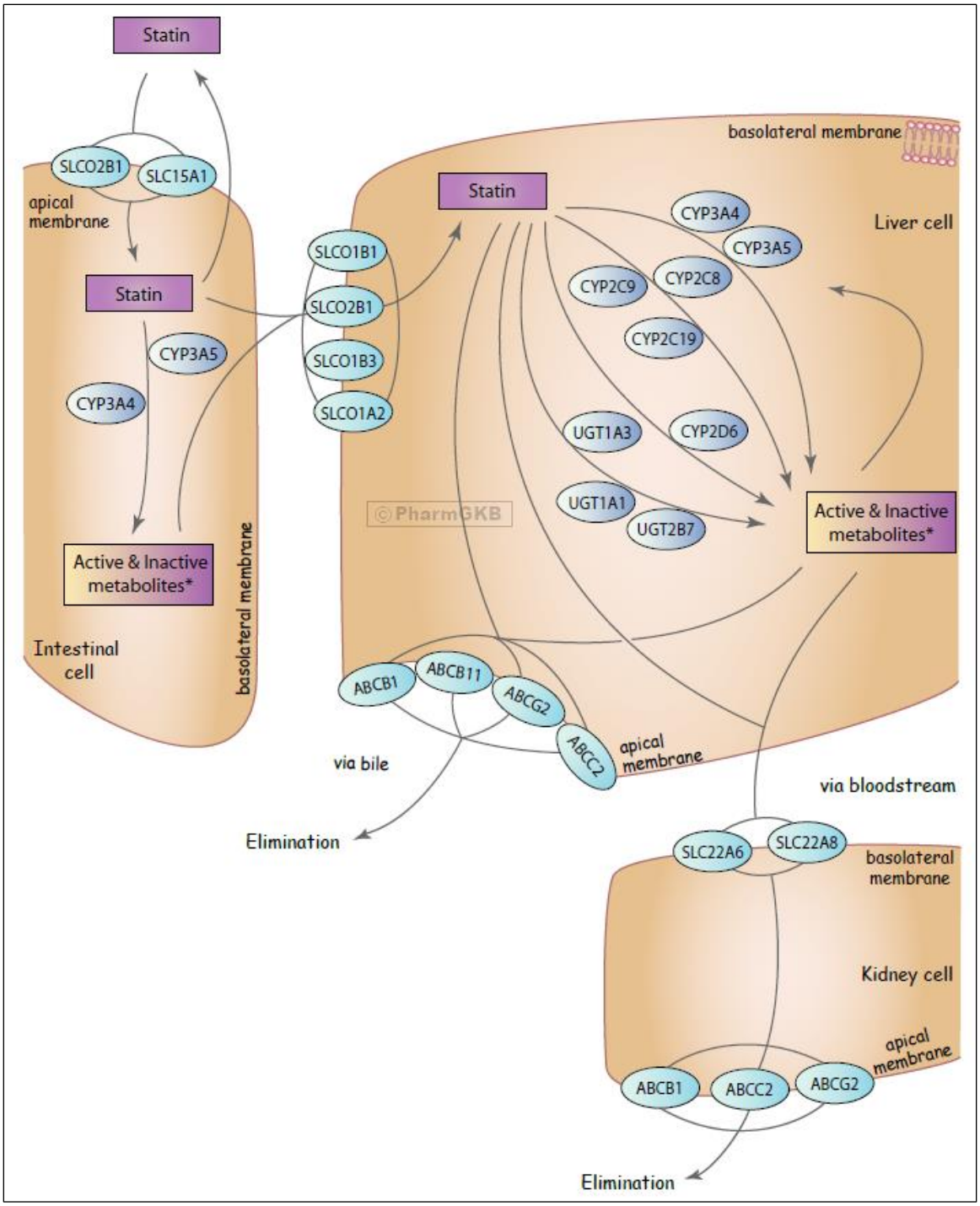


Figure 1.3 Representation of a superset of all genes involved in the transport, metabolism and clearance of statin class drugs (image from PharmGKB https://www.pharmgkb.org/pathway/PA145011108#PGG)

Metabolism of statins

Statins are predominantly metabolized by the CYPP450 (CYP450) family of enzymes, which contains over 30 isoenzymes (26). The specific gene enzyme involved in the metabolism of each statin is presented in Table 1.2 (5).

Table 1.2 Statins and genes involved in their first pass metabolism

Statin drug	Enzyme	Lipophilic
Simvastatin	CYP3A4	Yes
Atorvastatin	CYP3A4	Yes
Pravastatin	CYP3A4	No
Fluvastatin	CYP2C9	No
Rosuvastatin	CYP2C9 *	No
Cerivastatin	CYP3A4, CYP2C8	Yes

**Lack of strong evidence*

Simvastatin and atorvastatin are predominantly metabolized by CYP3A4, and a portion of the circulating inhibitory effect of these agents for HMG-CoA reductase is attributable to active metabolites (26,27). For atorvastatin the major active metabolites are 2-hydroxy and 4-hydroxy-atorvastatin acid, while for simvastatin the β-hydroxy acid and its 6'-hydroxy, 6'-hydroxymethyl and 6'-exomethylene derivatives are the major active metabolites (28–30). Fluvastatin is mainly metabolized by the CYP2C9 isoenzyme, while pravastatin and rosuvastatin do not undergo substantial metabolism by CYP450 pathways (5,26). Lipophilic drugs are known to be much more susceptible to oxidative metabolism by the CYP450 system (5). It is recognized that statins metabolized by the CYP450 system are more likely to produce muscle toxicity because of the risk of interactions with the many drugs that inhibit CYP450, notably the CYP3A4 isoform (31,32). Drug interactions increase plasma levels of statins, with an increased risk of toxic side effects (5).

The predominant route of elimination for the majority of statins is via the bile after hepatic metabolism (33). Consequently, hepatic dysfunction is an important risk factor for statin-induced toxicity (14); the British National Formulary recommends caution when prescribing statins to someone with a history of liver disease (10). Pravastatin is eliminated mostly as unchanged drug by both renal and hepatic routes (5,32). However, in patients with hepatic dysfunction pravastatin pharmacokinetics are altered (5,34). Rosuvastatin is also eliminated, largely unchanged, by both the kidneys and liver and its pharmacokinetic properties remain unaltered in patients with mild to moderate hepatic impairment (35).

1.5 Statin intolerance: causes, symptoms and consequences

Mechanisms

The excessive inhibition of cholesterol synthesis by statins may impair the integrity of neuronal cell membranes. Lipophilic statins that cross the blood brain barrier (such as simvastatin and atorvastatin) may also have adverse effects (36). Conversely, there is evidence of mechanisms which suggest that statins might have a beneficial effect on cognition such as improved endothelial function, reduction in free radicals formation, reduction in inflammation(37).

In a 2006 study Draeger et al. showed that statin-induced cholesterol lowering contributed to myocyte damage. They based their findings on biopsies of skeletal muscle fibers from statin users and non-users and reported substantial structural damage amongst statin users. Extracting cholesterol from muscle cells *in vitro* reproduced these characteristic abnormalities. They concluded that specific lipid-protein organization of the skeletal muscle is what renders it most vulnerable to statin exposure (38).

However, previously conducted *in vitro* studies provide contradictory findings. Using squalene synthase inhibitors, Flint *et al*, blocked cholesterol synthesis (in the mevalonate pathway Figure 1.2). They reported that this did not produce myotoxicity in *in vitro* models, suggesting that other compounds produced by HMG-CoA reductase activity might be responsible for the damage seen in skeletal muscles of statin users (39). They postulate that reduction in small guanosine triphosphate (GTP) -binding proteins contributes to the myotoxicity of statins; indeed pravastatin reduces protein synthesis in neonatal rat myocytes (39). This effect is reversed by adding farnesol and geranylgeraniol to the cultures; whereas reducing cholesterol levels with squalene synthase inhibitors

produces only minimal cytotoxicity. Such results suggest that depletion of the mevalonate metabolites (farnesol and geranylgeraniol) and not cholesterol contribute to statin-induced myotoxicity (39).

The hypothesized effect of farnesyl and geranylgeranyl pyrophosphate is via their role in the activation of certain regulatory proteins through prenylation. These regulatory proteins such as Ras, Rac and Rho promote cell maintenance and growth and attenuate apoptosis (40–43). Apoptosis, or programmed cell death, is a crucial mechanism designed to assist in the remodeling and maintenance of tissue structure. However, inappropriately activated apoptosis can produce pathological conditions.

Atorvastatin and simvastatin produce a dose-dependent increase in apoptosis in vascular smooth muscle cells (VSMCs). The effect is reversed by the addition of mevalonate, farnesyl pyrophosphate or geranylgeranyl pyrophosphate, but not ubiquinone or squalene. This also points to the role of statins in potentially enhancing apoptosis in VSMCs in a pathway not related to cholesterol depletion (44,45).

Raised plasma concentration of statins increase the exposure of skeletal muscle to statins. Therefore, compounds that inhibit CYP450 activity, such as fibrates, are known to increase the risk of statin-associated muscle symptoms (SAMS). Further, genetic factors affecting the metabolism and transport of statins (discussed in sections 0and 0) are also thought to be associated with the development of statin intolerance (SI) and SAMS.

The fact that statin-induced skeletal muscle injury involves inhibition of pathways that activate GTP might explain why exercise often appears to unmask the negative effects of statin therapy. Exercise has been shown to activate signaling pathways, specifically the

mitogen-activated protein kinase pathways, that are important in skeletal muscle cellular response to exercise-induced stress (43,46,47). These pathways are regulated by GTP binding proteins. Therefore, statins might impair the muscles ability to recover from physical exercise, resulting in skeletal muscle damage (43).

Symptoms

While statins are generally well-tolerated, adverse reactions can be neurological (48), gastro-intestinal or muscle-based (49,50). Muscle-based adverse reactions to statins range from non-specific complaints of myalgia (see in 29% of statin users), through symptomatic myopathy with accompanying increases in serum CK levels, to rhabdomyolysis where the accumulation of CK due to muscle breakdown leads to kidney damage and death (occurring in 4 per 100,000 person years) (51,52). There is evidence that adverse drug reactions (ADR) to statins are dose dependent (53,54).

Neurological adverse effects of statins are rare. They are highlighted by the United States Food and Drug Administration (FDA), who amended the statin product label in 2012, to state that some users might experiences “ill-defined memory loss” and “confusion”. Evidence suggesting statins might have neurological effects stem mainly from small randomized trials and case reports. Two large randomized controlled trials (RCT), the PROSPER (The Prospective Study of Pravastatin in the Elderly) trial and HPS (Heart Protection Study) found no significant association between use statins and cognitive decline (55–57). Some studies report an increased risk for the deterioration of front-executive function and working memory (58,59), attention (60), processing speed and motor speed (58–61). Conversely, statins also reduce the risk of clinical atherosclerotic disease, which is a known risk factor for vascular dementia (62). However

there is a lack of conclusive evidence in a RCT or large cohort study that shows a temporally valid relationship between statin usage and cognitive decline.

In 20 RCTs, the incidence of myalgia was reported to be 190 per 100,000 person years (63), whereas in the ambulatory setting it was reported to occur in 9% of statin users (64). Rhabdomyolysis was reported to occur 1.6 times per 100,000 person years in clinical trials. In the ambulatory setting, the incidence of hospitalization due to rhabdomyolysis was shown to be 4 per 100,000 person years on statin monotherapy (51,65).

There are several approaches to defining SI. A study by Alfirevic *et al.* on behalf of the PREDICTION-ADR consortium, described the clinical presentations of statin-induced muscle toxicity as being muscle symptoms such as muscle fatigue, pain, weakness, cramps or tenderness (66). In addition, plasma CK elevations, rhabdomyolysis, and the presence of autoantibodies that recognize HMG-CoA reductase (HMGCR) – seen in some cases that develop statin-induced autoimmune myopathies. This variety of clinical manifestations also highlights the different etiologies of statin-induced muscle damage. They conclude, with consensus from a panel of experts that a CK elevation greater than 4 times the upper limit of normal with or without clinical symptoms can be defined as myopathy.

However, increased risk of ADRs are not related to LDL lowering efficacy of a drug. For example, cerivastatin was discontinued in 2001 for being associated with a higher risk of developing rhabdomyolysis among users (65,67) but is not more efficient at lowering LDLs than other statins (*see* Table 1.1) (12,68).

Consequences of intolerance

ADR contribute to a significant health care burden, economic cost and a burden on disability-adjusted life years (DALYs). Drug adherence patterns are affected by ADRs; a large proportion of those suffering from severe ADR discontinue statin use, putting them at an increased risk of having a cardiac event. The high incidence of statin usage, combined with intolerance to them calls for perhaps a more individualized management of CVD, that takes into account the patient's probability of developing ADR and adjusting treatment plans accordingly. In this vein, it is essential to analyze variables that might affect the risk of ADR and take those into account, such as the possible role of genetics, concurrent therapy with other drugs, and interaction between these two and any other risk factors such as age, gender and ethnicity.

1.6 Pharmacogenetics of statin intolerance

With an ageing population and higher incidences of comorbidities for which prophylactic treatment with statins is recommended, such as diabetes, the number of people on statins is on the rise across the globe. In properly selected patients, statins decrease cardiovascular disease mortality by 25% (69); thus tens of millions of patients worldwide now receive statins for hypercholesterolemia. Surprisingly, more than 40% of those eligible for statin use are not currently taking them. While one reason is affordability, the other is intolerance to statins due to myopathy (17,70).

There has been a lack of consensus amongst researchers about the exact definition of a statin-induced myopathy (SIM). This is likely caused by a lack of understanding about the underlying mechanisms. Several studies investigating the underlying mechanisms of SIM have focused on the genes linked to the pharmacokinetics of statin metabolism, since the exposure to a particular statin is much higher for an individual with a poor metabolizer genotype for certain CYP450 enzymes, compared to an intermediate, extensive or rapid metabolizers (71). For example, the pharmacokinetics of fluvastatin depend on the *CYP2C9* genotype, with a three-fold difference in the active enantiomer and an even greater difference in the inactive enantiomer (72). A case report in 2004, Ishikawa *et al*, report that rhabdomyolysis associated with cerivastatin was caused by genetic variability within the *CYP2C8* gene (71,73,74). The relationship between *CYP3A4* and *CYP3A5* gene polymorphisms and atorvastatin-induced muscle damage was investigated in a case-control study with 68 cases and 69 controls. Serum CK elevation was used as a proxy for muscle damage, and was found to be highest in patients on atorvastatin treatment who were homozygous for the CYP3A5*3 allele (75). Another study examined the

relationship between 388 common single nucleotide polymorphisms (SNPs; most of which were within the CYP-coding genes) and elevated CK or myalgia in 136 patients taking either atorvastatin or simvastatin and 296 controls. They reported an association between the CYP2D6*4 allele and atorvastatin-induced myopathy. Interestingly these findings extended to myopathy in patients taking simvastatin which is not known to be metabolized by CYP2D6 (71,76). This brings into question both the validity of these findings and whether or not the effect was confounded or mediated by an unknown variable. However, the biological plausibility of these findings as they pertain to atorvastatin metabolism suggests otherwise.

Also of interest are drug transporters that mediate the uptake and elimination of statins. The hepatocellular influx transporter OAT1B1 (encoded by *SLCO1B1*) and intestinal and hepatocellular efflux transporters ABCB1 (encoded by *ABCB1*) and ABCG2 (encoded by *ABCG2*) have been shown to affect the pharmacokinetics of statins (24,71,77,78).

The pharmacokinetically different profile of statins between *SLCO1B1* genotypes has also been shown to affect the risk of myopathy. The Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine (SEARCH) collaborative group conducted a prospective cohort study conducting a genome-wide association study in 85 myopathy cases and 90 controls all of whom were taking 80 mg of simvastatin once daily (53). Only a non-coding SNP (rs4363657) within the *SLCO1B1* showed a strong association with myopathy. This SNP is in strong linkage disequilibrium with the non-synonymous Val174Ala (rs4149056) variant which had been previously association with altered statin pharmacogenetics (77,78). For each copy of the variant allele there was approximately a four times greater risk of myopathy. Importantly, this finding was

replicated in a trial with subjects treated with 40 mg simvastatin once daily in the SEARCH trial.

In a crucial replication study performed at the University of Dundee, Donnelly *et al.* examined 4196 individuals in the Genetics of Diabetes Audit and Research, Tayside Scotland (GoDARTS) cohort and concluded that two functional variants; the loss-of function Val174Ala and gain-of function Asp130Asn of the *SLCO1B1* gene encoding OATP-C/OATP1B1 are associated with general statin intolerance. The study produced a diplotype risk score for intolerance. Since this is a large retrospective cohort study the findings have external validity (54).

The Statin Response Examined by Genetic Haplotype Markers (STRENGTH) study investigated the genetics of four *CYP* genes and the *SLCO1B1* gene in relation to SIM (use of simvastatin, atorvastatin and pravastatin was included). Not only did the study confirm the findings from the SEARCH study, it also reported an association between *SLCO1B1* risk allele and myalgia symptoms without CK elevation for simvastatin and atorvastatin (weaker) but not for pravastatin treatment (79).

An additional potentially important, but very rare, SNP in *SLCO1B1* is 1628T>G. This novel variant was discovered by a Japanese group in a patient with pravastatin-induced myopathy (79), and was shown to reduce transporter activity of OATP1B1 (80). In another study, the TTT (or TAT) haplotype of ABCB1 1236C>T, 2677G>A/T, or 3435C>T polymorphisms was more frequently seen in the simvastatin treated group without myalgia (81). The 421AA variant of efflux transporter *ABCG2* has been shown to increase plasm concentrations of both rosuvastatin and atorvastatin (82). However, the findings for *ABCB1* and *ABCG2* are preliminary, and lack replication (22).

Statin related myopathy is usually managed by switching agents, use of fluvastatin, low-dose rosuvastatin, nondaily dosing, and ezetimibe or bile acid-binding resins (69).

Since statins were first introduced, their effect on liver enzymes has been documented, albeit in a small proportion of the population taking statins. Specifically, alanine and aspartate transaminase levels show abnormalities (4). With standard doses, little or no effect is seen on gamma glutamyl transferase, alkaline phosphatase or bilirubin (83). These increases in transaminase levels with statins are seen within the first six months of commencement of treatment, and while these effects are reversible with the cessation of statins, the main concern is hepatotoxicity or some hepatic reaction to reduction of lipid levels.

Other cholesterol lowering agents such as fibrates, resins, niacin and ezetimibe all increase liver enzymes, which suggests these changes could be a hepatic response to lipid lowering rather than hepatotoxicity (67,84–86).

Thus far, attempts to identify an association between a specific genetic biomarker and the development of ADR with the use of statins have consistently found an association with the *SLCO1B1* gene (53,54,87,88). This provides us with a robust candidate gene to examine using genetic association studies. However, the barrier to clinical applicability for these findings is a common pattern of pharmacogenomic associations that are not replicable in other studies (89). Often the reasons for inconsistent findings between studies include differences in study population, poorly defined phenotypes (different definitions of statin intolerance), statistical power issues and chance findings due to testing too many genetic variables (type 1 error). Many studies have been published that examine genetic markers for the prediction of statin efficacy. However, very few genetic

interactions have been confirmed, and the utility of these effects in clinical settings is very low, because the differential efficacy caused by genes is small (90). However, the association with the Val174Ala variant in *SLCO1B1* has been replicated in multiple studies and confirmed by pharmacokinetic studies (91).

It is important for statin-related studies to focus on pharmacokinetic enzymes that might be of importance to one or more statin. It would be preferable if the complete metabolic route for all statins were completely understood, however, this is not the case – the pathway is very complex and includes numerous transporters, and metabolizers (*see* Figure 1.3) (21,71).

Studies into statin-intolerance have pinpointed genetic factors involved in the pharmacokinetics of statins. However, so far no studies have explored factors associated with susceptibility to muscle damage or immune factors that could predispose a statin user to intolerance.

1.7 Genetics of statin intolerance

Outwith pharmacogenetic factors associated with general statin intolerance, there are known genetic factors that are associated with autoimmune-mediated statin-induced myopathies. This pathology is seen in statin-exposed individuals who produce autoantibodies that recognize 3-hydroxy-3-methylglutaryl-coenzyme A reductase, which is the pharmacological target of statins (92,93). This form of myopathy is characterized by progressive weakness in the muscles, muscle enzyme elevations, necrotizing myopathy when a biopsy is taken and the presence of antibodies of HMGCR. The lack of resolution on dechallenge is another distinguishing feature of this form on statin intolerance (66,94).

A study undertaken by Mammen *et al.* compared 28 cases of anti-HMGCR-associated autoimmune myopathy, with 654 statin tolerant controls and 51 cases of self-limited statin intolerance. Variants in the *HLA D* gene were found to be associated with this form of statin-induced myopathy. DR1*11:01, a HLA class II allele was found to be a risk factor for anti-HMGCR, while alleles DQA1 and DQB6 appear to be protective (95).

However, this form of statin intolerance is rare, there are fewer than 2 cases per million statin users (94,96).

1.8 PREDICTION-ADR: a background to the current project

The PREDICTION of Adverse Drug Reactions (PREDICTION-ADR) consortium is a collaboration between four academic institutions, the University of Dundee, the University of Liverpool, Universiteit Utrecht and Uppsala Universitet and two private companies, PGXIS and Asper Biotech. The project aims to discover the genetic factors predisposing patients to ADR from drugs used to treat CVD, specifically statins and angiotensin converting enzyme-inhibitors (ACE-Is). By pooling the clinically adjudicated cases of ADR collected at these different centers, we hope the findings will be clinically meaningful.

For the project arm investigating the genetics of ADR to statins, the main work packages for the project include: discovery of sequence variants predisposing to statin myopathy, exome sequencing, statistical genomics, and diagnostic modeling, replication and population based validation, functional genomics and dissemination of findings. The PREDICTION-ADR project has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) Under Grant Agreement no. 602108.

On a population level, the incidence of statin-induced myopathy is rare, occurring an estimated 5 per 100,000 person years (17,66). Other forms of intolerance such as non-specific complaints of muscle aches, often not accompanied by big elevations in CK are much more prevalent (97). Estimates vary drastically between 1.5 to 26% of statin users experiences such a reaction (53,66,98–100). The development of ADRs to statins lead to problems with drug compliance, dose changes, and can lead to discontinuation which puts the user at higher risk for a CVD-related event (97,101,102).

Due to the burden on population health, this study is focused on more prevalent forms of intolerance, which preclude statin users from adherence to therapy, and predispose them to poor outcomes such as CVD.

This project is driven by clinical necessity and the overall goal is to provide a robust diagnostic algorithm that can be used by clinicians while prescribing statins to ensure best outcome.

1.9 The GoDARTS study and JUPITER trial cohorts

The GoDARTS study is a longitudinal cohort study of over 18,000 participants. Originally created to study the genetics of type 2 diabetes. GoDARTS is a rich source of data, combining complete electronic medical records including prescription information, all laboratory results from clinical visits (e.g. biochemistry, immunology), and genetic data from a cohort of 18,190 individuals in Tayside, Scotland. GoDARTS has been previously used to perform crucial pharmacogenetic studies on drugs such as statins (54), metformin (103,104), sulfonylureas (105,106), thiazolidinedione (107) and others. GoDARTS was also used to study the genes associated with lipid lowering efficacy such as apolipoprotein E *(APOE)* (108), lipoprotein A (*LPA)* and *HMG-CoA* (109).

GoDARTS also has contributed to the discovery genome-wide association study (GWAS) of statin efficacy with the Genomic Investigation of Statin Therapy (GIST) consortium leading to the discovery of the sortilin 1 (*SORT1)* locus as a marker of statin efficacy (110). The maintenance and use of this cohort is described in detail in the *Methods* chapter of this thesis. GoDARTS was used as the underlying cohort to perform a nested case-control study to examine statin intolerance in this thesis.

Longitudinal cohort studies provide us with the unique opportunity to place the purported "exposure" and "outcome" in a temporally valid fashion, a distinct advantage over a cross-sectional study design (111). Another substantial advantage is the option to study many possible outcomes stemming from the same exposure. In this study the exposure was the use of statins, and the outcomes examined were different biochemical and prescribing patterns that would be suggestive of intolerance to statins, and the occurrence of events such as myocardial infarctions, ischaemic strokes and coronary artery disease that would be indicative of a failure of statin therapy. Another advantage of the cohort study is the ability to estimate not only odds and risk ratios but also the hazards of the exposure causing an outcome over time; analyses that cannot be performed when using cross-sectional data sets (112).

Creating an ideal counterfactual in case-control studies is always a challenge, but is easier to accomplish in a cohort study. Unlike a cross-sectional or simple case-control study, access to longitudinal datasets gives us the opportunity to not only precisely define the occurrence of the outcome of interest, but also to more conclusively define what the lack of the outcome might be (113). This is especially important when examining the effects of a drug for chronic conditions, like statins which constitute a continuous

exposure over time (112). For the purpose of this thesis, chronological trends in the usage of medication can quite conclusively indicate an individual's adherence to a drug, while hospital records and biochemistry test results can point to their tolerance of the same.

The financial and human capital involved in the creation of such cohorts is considerable. Therefore cohort studies generally suffer from a lack of reproducibility of findings, since access to similar longitudinal datasets is limited. In order to improve the validity of the results presented in this thesis, the verification of findings was important.

Randomized controlled trials (RCTs) are considered to be the gold standard for epidemiologic studies (111). Due to randomization, causal inference most robustly detected from RCTs compared to any other study design. However, due to the clinical nature of a RCT, various biases are introduced. Most notably reporting bias, where, due to the participant's awareness of the clinical trial, they are more likely to report any events they believe might be associated with the exposure and modify their own behaviour. Clinicians too, behave differently than they would in the ambulatory setting, as they are more likely to document and treat reported events. Due to blinding, reporting bias does not necessarily impact the validity of findings. Selection biases are introduced due to the inclusion criteria of a study and due to fundamental differences in the baseline population that impact their willingness to participate in a RCT, such as their general health. Randomization is key in reducing the impact of the reporting biases. It provides a robust baseline for statistical testing, and it allows the investigators to assess the impact of the placebo effect.

The JUPITER trial was conducted to determine the efficacy of Rosuvastatin in reducing CVD in people with raised C-reactive protein (CRP) levels, a known risk factor (114).

The cohort of approximately 8,000 individuals with genotype data were comprised of half of whom were randomized to receive Rosuvastatin and the remaining placebo (115).

By using the JUPITER trial as the replication cohort, we also gain insight into the statin independent impact of the genetic variants under consideration. The main weakness of cohort studies, unmeasured confounding factors, are controlled for in a RCT and this strengthens the overall finding. Meanwhile, the modified behaviours of participants and physicians in RCTs that might impact outcomes are counterbalanced in the GoDARTS study.

The results presented in this thesis have been examined in two large population-based studies, with different designs and independent strengths.

1.10 Gaps in knowledge addressed by this thesis

While criteria for the classification of statin intolerance abound, there remains a lacuna in the application of the most robust definitions of intolerance to large longitudinal population-based data sets. The use of CK however, remains the cornerstone in assessing the extent of statin-induced muscle damage, or indeed muscle damage in general. Paradoxically, there are cases of reported statin intolerance and myalgia, adjudicated by physicians that do not show the expected elevations in CK levels. We attempt to understand why this might occur. While the impact of poor adherence on outcomes for those on statin therapy have been examined, we attempt to uncover how intolerance to the drug impacts outcomes.

Centrally, this thesis explores the role of members of an immune system related gene family, the Leukocyte Immunoglobulin-like Receptor (*LILR*) in the development of statin intolerance. Members of this family have been shown to be associated with serum CK

and lactose dehydrogenase (LDH) levels, T-cell tolerance and inflammation (116–119). The findings presented in this thesis are novel in their exploration of the immune system as a contributor to the developing intolerance to statins – whether mild or severe.

1.11 Thesis aims

1. Create and apply definition of statin intolerance in GoDARTS EMR data
2. Validate these definitions against real-world outcomes
3. Examine genetic factors associated with creatine kinase, the main biomarker of SI, and seek replication
4. Examine the association between factors associated with CK and SI, and seek replication
5. Examine the role of immune system gene family *LILR* with SI

2 Methods

2.1 Clinical data – description of databases

The Health Informatics Centre (HIC) in partnership with the University of Dundee (UoD), National Health Service (NHS) Tayside and the information services division of national services provides researchers and other with information derived from person-specific datasets. These datasets are mainly derived from data held by the UoD and NHS and are anonymized in accordance with the Standard Operating Procedures approved by the Caldicott Guardians. In Scotland, every person registered with a medical practitioner is assigned a Community Health Index number (CHI). This is a unique 10 digit identification number that is lined to information on address such as postcode, medical practitioner registration status and their status at the General Registry Office (GRO) and tracks births and deaths. This information is collated and held by the Tayside Health Board for the entire Tayside population. Tayside also uses this number as the patient identifier in all health care activities from primary to tertiary care, thus allowing for the record-linkage of datasets

DARTS

Diabetes Audit and Research Tayside Scotland (DARTS) database includes information of all patients with diabetes in Tayside. Individuals with diabetes were identified from hospital records. The database as validated against general practice records and was confirmed to be robust. The methodology used was shown to be more sensitive than general practice alone at identifying individuals with diabetes. There are record linked data available for individuals within the DARTS cohort (120).

GoDARTS

Genetics of Diabetes Audit and Research Tayside Scotland (GoDARTS) comprises 18,190 participants enrolled between December 1998 and August 2012 in which there are 9829 with type 2 diabetes (T2D) and 8361 non-diabetics. Participants with T2D were identified for enrollment through DARTS. Age and sex matched diabetes-free participants were identified in populations within the Tayside region from general practice records (121). Relevant clinical data for all GoDARTS participants are drawn from electronic records of hospital admissions (Scottish Morbidity Register, SMR01), deaths (GRO), biochemical tests and dispensed drug prescriptions, available for the Tayside region. Data are available from 1980 until present for the SMR hospital admissions data; from 1998 for deaths from the GRO, from 1993 until June 2014 for biochemical tests.

The GODARTS study was approved by the Tayside Committee for Medical Research Ethics and written, informed consent was obtained from each participants. A single sample of blood was collected for DNA extraction and genotyping, and the participant was assigned a unique anonymized system identifier. Baseline characteristics were recorded at time of recruitment for all participants (121).

CHI master index

This demography database contains one entry per individual in the study enlisting details such as their date of birth, ethnicity, sex and date of recruitment to the GoDARTS study.

Scottish Morbidity Register (SMR01)

This is a record of acute hospital admissions in Tayside and Fife, Scotland. The database consists of one line per patient admission that includes a date of admission, one principal diagnostic field and five additional diagnostic fields. This database also includes admissions for hospital procedures that include one principal procedure field and eight additional procedure fields. The hospital admissions are classified according to the International Classification of Diseases (ICD) 9th and 10th versions (122). Procedures are classified according to the Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures' 3rd and 4th revisions.

General Registry Office (GRO) death certification

The GRO death certification database is a record of the date and cause of death. Deaths in Tayside have been electronically recorded since 1989 and the database includes a principal cause of death field and ten additional related causes of death fields. Cause of death is classified according to ICD9 and ICD10 codes (122).

Laboratory data

The Tayside Laboratory systems record all tests performed in surgeries, clinics and hospitals that have been sent to the Tayside laboratories for processing. Clinical laboratory data are available from 1992. The database contains biochemistry, hematology, microbiology, virology and serology laboratory results and reports.

Prescribing data

HIC provides complete data for prescriptions dispensed in Tayside since 1993. Prescriptions dispensed between 1993 and 2004 were recorded as scanned paper prescriptions analyzed with purpose written software. Since late2004, all prescriptions

were obtained in electronic format from the Practitioner Services Division (PSD). The PSD are responsible for the processing and pricing of all prescriptions in Scotland. The prescriptions recorded include all those dispensed in community pharmacies, dispensing doctors, and a small number of specialist appliance suppliers. Hospital prescriptions are included if they were dispensed in the community.

Individual drug prescriptions are linked to an individual CHI number and state the name of the drug, date of prescription, amount dispensed, as well as dosing instructions. Drugs are identified by name and individual drug codes linked to the British National Formulary (10).

2.2 Data handling and manipulation

The electronic medical records (EMR) are supplied as flat text files from which relevant data are extracted and combined in forms suitable for statistical analyses. Phenotypes used in this thesis are derived from prescribing, biochemistry, SMR01 and GRO data sets. All management, cleaning and processing of data was done using SAS 9.3 (SAS Institute, Cary, North Carolina).

Statistical analyses

All statistical analyses were performed in SAS 9.3 (SAS Institute, Cary, North Carolina). Specific statistical tests performed are described, along with the SAS command in each result chapter. The main methods used to determine genetic associations are binary logistic regression, linear regression and cox proportional hazards models.

2.2.1.1 Binary logistic regression

This model is employed when the dependent variable is binary such as a case – control study and the predictor or explanatory variables that are linear or categorical. Binary logistic regressions were used to determine associations with statin intolerance (treated as 1 – intolerant and 0 – tolerant). The assumptions of a logistic regression are:

1. The observations are independent of each other

2. The parameters are approximately normal (usually with a large sample size)

3. The model is correctly specified (categorical outcome, more observations than parameters).

$$\text{Logit (pi)} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$$

Equation 2.1 Formula for logistic regression (predictors with two levels)

$$\text{Logit (pi)} = \alpha + \beta_1 X_1 + \beta_2 X^1{}_2 + + \beta_3 X^2{}_2 + \beta_3 X_3 \ldots + \beta_n X_n$$

Equation 2.2 Formula for logistic regression (predictor X_2 has three levels)

Interpretability: The logistic regression co-efficient describes the size of the contribution of each subsequent sub-category compared to the reference sub-category of a categorical variable in modifying the probability of the dependent variable's occurrence (123).

2.2.1.2 Linear regression

Multiple linear regression is used to determine the relationship between a linear dependent variable and linear or quantitative independent variables.

1. That the residuals in the model are normally distributed
2. Error terms are constant and do not depend on the value of the independent variable
3. There is no collinearity between independent variables.

Residuals are calculated by subtracting the observed values of the dependent variable from those that are predicted by the model.

$$y = \beta_1 X_1 + \beta_2 X_2 + + \beta_3 X_3 \ldots + \varepsilon$$

Equation 2.3 Formula for Linear regression

Interpretability: The beta estimates produced by the linear regression can be interpreted as the expected change in the dependent variable for a unit change in the independent variable if all other independent variables are held constant (124).

2.2.1.3 Cox proportional hazards model

This model is used in epidemiological studies to model the independent variables that determine the dependent variable which in this case is time to failure. In observational studies, an observation time is defined, usually time from an index or start date until the occurrence an outcome (however, it may be defined) or until the study censor date in lieu of outcome. Each individual has a numerical value for the observation time and a binary measure for outcome (1 or 0). To calculate the survival function the survival time is broken into intervals and for each time interval the proportion of individuals that have not failed and go on to enter the next time interval is measured. The number of cases that have had an outcome can be measured and the number of individuals that were censored for that time interval can also be computed.

1. The model rests of the proportional hazards assumption which is that, the estimation of the hazard function from independent variables does not depend on the time and i.e. there is a constant relative hazard across the two strata over time.
2. Equation: where x is the variable, and β is the estimate generated in a linear model

$$HR = \frac{h(t|x_2)}{h(t|x_1)} = \frac{h_0(t)exp(x_2\beta_x)}{h_0(t)exp(x_1\beta_x)}$$

3. The baseline hazard rate is cancelled out, and the hazard rate does not depend on time.

4. $$HR = exp(\beta_x(x_2 - x_1))$$

Equation 2.4 Formula used to compute hazard ratios

5. Interpretability: Hazard ratios are used to represent instantaneous risk over the study period or a subset of the study period. They are interpreted as the chance of an event occurring in the "treatment" arm divided by the chance of the event occurring in the "control" arm of a study. Kaplan-Meier plots usually accompany survival analyses and represent the resolution of such endpoints (125).

2.3 Genetic data

Genetic data used in this thesis were generated on either high density arrays (such as Affymetrix 6.0 or Illumina Omni-express arrays) or using the TAQMAN procedure. Samples which were genotyped on the Genome-Wide Human SNP Array 6.0 were processed at Affymetrix's service laboratory for all samples passing Affymetrix's laboratory quality control; raw intensities were renormalized within collections using CelQuantileNorm. Their normalized intensities were used to call genotypes with an updated version of the Chiamo software adapted for Affymetrix 6.0 SNP data.

DNA was prepared on the QIAsympohony (QIAGEN, Hildon, Germany). The process for extraction is described in Figure 2.1. The quality of genomic DNA was validated using the Sequenom iPLEX assay designed to genotype four gender SNPs and 26 SNPs present on the Illumina Beadchips. DNA concentrations were quantified using a PicoGreen assay (Invitrogen) and an aliquot assayed by agarose gel electrophoresis. A DNA sample was considered to pass quality control if the DNA concentration was greater than or equal to 50 ng/μL, the DNA was not degraded, gender assignment from the

iPLEX assay matched those provided in the patient manifest and genotypes were obtained for at least two thirds of the SNPs on the iPLEX.

DNA preparation and genotyping were performed by members of the Palmer laboratory staff, principally, Ms. Fiona Carr, Dr. Roger Tavendale and Ms. Karen Wilson.

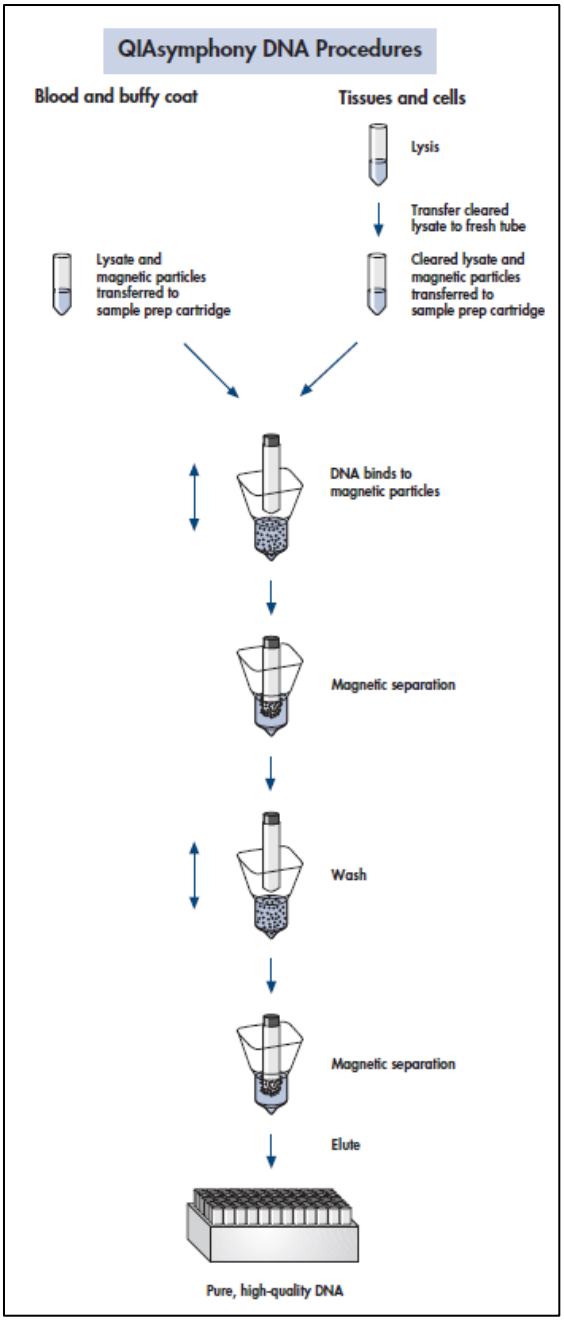


Figure 2.1 DNA extraction procedure (QIAsymphony DNA handbook, QIAGEN, Hilden, Germany)

Affymetrix 6.0 genotyping array

4000 diabetic individuals were genotyped on the Affymetrix 6.0 SNP genotyping array that includes 1 million SNPs (Affymetrix, Santa Clara, USA). These individuals were specifically chosen for genotyping as they had all gone onto receive statins after recruitment to GoDARTS.

Illumina Omni-express array

4000 diabetic cases were genotyped on the Illumina Omni-express array which consists of ~ 700 thousand SNPs selected from HapMap 1-3 for SNPs with a MAF greater than 5% (Illumina, San Diego, USA) (126). The array was designed by selecting tag SNPs to serve as a proxy for a number of other SNPs across the genome. This approach allows for the broadest selection of maximally informative markers, resulting in genome-wide coverage of both common and rare variants.

Exome Chip

This Illumina Infinium Exome-24 v1.0 BeadChip was designed to genotype functional exonic variants selected from over 12,000 individual human exome and whole-genome sequences. Marker selection was performed with the goal of developing an extensive catalog of exome variants. Exonic content consists of over 240,000 markers representing diverse populations, for a range of conditions such as type 2 diabetes, cancer, metabolic and psychiatric disorders.

TaqMan

Direct typing of individual SNPs was performed using TaqMan allelic discriminations assays as supplied by Applied Biosystems (Carlsbad, CA) as "Assays on Demand", or "Assays by Design". All typing was performed in 384 well plate format using 10-20 ng of DNA in 2μL reaction volumes using Universal TaqMan master mix (Applied Biosystems, Carlsbad, CA). Assays were plated using a DEERAC Equator GX microdispenser (Labcyte, Sunnyvale, CA) and thermal cycling was performed in H20BIT high throughput thermal cycler (KBiosystems, Basildon, Essex). End point

fluorescence was measured and genotypes were called using an ABI7900HT sequence detection system (Applied Biosystems, Carlsbad, CA).

Genotype imputation

Genotype imputation is conducted in two stages: the first is estimation of the haplotypes from the study population and the second is the imputation of genotypes by comparing study haplotypes to reference panel haplotypes.

2.3.1.1 Haplotype inference

The segmented haplotype estimation and imputation tool (SHAPE-IT, http://www.shapeit.fr) (127) was used to estimate haplotypes for downstream imputation with IMPUTEv2 (128). The method is highly accurate and computationally light when compared with other available methods and is particularly suited to populations that contain high linkage disequilibrium in their genomes such as Caucasians. The inference of haplotypes is computed in a similar way to Phase v2 where all possible haplotypes are estimated from the available genotype set with an associated probability. Since the number of haplotypes increases exponentially with the addition of more genotypes, there need to be methods to reduce the haplotype set to the most likely one.

2.3.1.2 Genotype imputation

Genotype imputation was performed using the IMPUTEv2 program (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (126) using the estimated haplotypes from SHAPE-IT to impute genotypes from the haplotype set. Impute 2 compares the study population haplotypes with up to two reference panels for the imputation of missing genotypes in the study population. Alleles are imputed in the study population by running a forward-backward algorithm to impute missing alleles with a certain probability. The two sets of haplotypes are compared to each other and missing

alleles are imputed into the study panel from the reference panel with certain probability. The two sets of haplotypes are compared to each other and missing alleles are imputed into the study panel from the reference panel with certain probability. Certain SNPs will be found in haplotypes together so if one SNP is not present in the study panel but the haplotypes match we can impute an allele for that SNP with a certain probability given the alleles that are present in the reference panel and the alleles provided by the study panel. If there is good coverage of a haplotype on a particular chip and the haplotypes are sparsely covered or SNPs are not in linkage disequilibrium with any other SNPs to form haplotypes then the alleles may be imputed with low confidence in their accuracy or may be missing all together. Given that we assume both sets of haplotypes are sampled from population in Hardy-Weinberg equilibrium the allelic probabilities can be converted to genotypic probabilities (126).

Genotype imputation for all datasets used were performed by Mr. Phillip Appleby.

Categorization of dominant or recessive traits

Genotypes of SNPs were coded as 0 – no rare alleles present, 1 – one rare allele present (heterozygous) and 2 – two rare alleles. Statistical analyses were first run using additive models, where the risk of each additional rare allele was calculated successively.

Dominant traits were defined as those where the presence of one or two rare alleles made no statistically significant difference to the probability of the outcome, while both were significantly different from those who had no rare alleles (1 and 2 v. 0). Recessive traits were the opposite, in that only carriers of two rare alleles were able to show a statistically significant difference in the probability of the outcome (2 v. 1 or 0).

2.4 Epidemiological Study Design

While GoDARTS is a longitudinal cohort study, this thesis focusses on the statin using subpopulation that are either tolerant or intolerant to their therapy, and therefore forms a nested case-control study (NCC). A NCC study is a retrospective design, using data already collected to determine if a study participant can be categorized as having the outcome, i.e. being classified as a case or a control. Since a participant's documented clinical records are used to ascertain their exposure, such a study is less prone to recall bias compared to a case-control study. Since cases and controls are selected from a database and no further participation is required, the NCC design is more representative of the baseline population than a traditional case control study and is less prone to selection bias. Figure 2.2 shows the overall design of a NCC.

The other advantage is that in studies of the long-term impact of drug exposure, randomized clinical trials, which are the gold standard for assessing causality, are poorly powered. In such an instance, a cohort of a large number of individuals with an exposure of interest and longitudinal follow-up is the more appropriate study design as it has the power to detect risk factors due to the person-years of exposure data available.

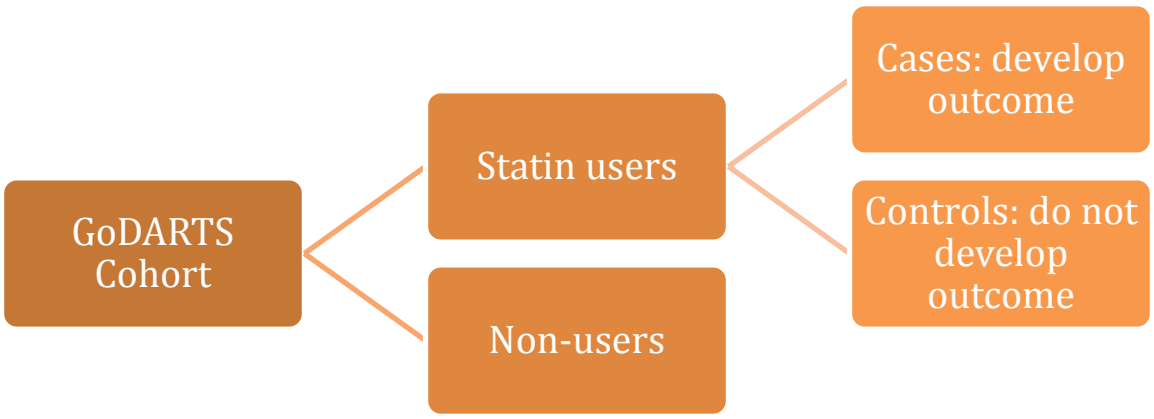


Figure 2.2 Design of this nested case-control study

3 Characterizing statin intolerance in the GoDARTS population

3.1 Introduction

The GoDARTS (Genetics of Diabetes Audit and Research, Tayside) cohort is comprised of 18,190 Scottish Caucasian individuals. The data set is maintained by the Health Informatics Centre at the University of Dundee and contains record-linked electronic medical records, biochemical, and clinical, prescribing and demographic information for all participants. The study was approved by the Tayside Medical Ethics Committee. The cohort is comprised of 9439 type 2 diabetics and 8187 non-diabetics.

GoDARTS has been previously used to establish the pharmacogenetic associations of genes such as the influx transporter: *SLCO1B1* and statin intolerance (54), and genes associated with lipid lowering efficacy such as *APOE* (108), *LPA* and *HMG-CoA* (109). GoDARTS has also contributed to the discovery GWAS of statin efficacy with the GIST consortium, leading to the discovery of the *SORT1* locus as a marker of statin efficacy (110). In addition, GoDARTS has been used extensively in understanding medication use, adherence and efficacy, for statins (129), metformin (103) and sulfonylureas (106), and to study the heritability of metformin response (104).

Statins are the most effective drug for the reduction of LDL-cholesterol. They are recommended for type 2 diabetics over the age of 40 years, and for younger diabetics with poor glycemic control (HBA_{1c} greater than 9%), low-HDL cholesterol and raised triglyceride concentration, hypertension or a family history of premature cardiovascular disease. Statins are also advised where total cholesterol concentration to HDL-cholesterol ratio exceeds 6 (10). For these reasons the use of statins is widespread in the adult population, which is reflected in GoDARTS.

While they are generally well-tolerated; neurological (48), gastro-intestinal or muscle-based (49,50) reactions to statins have been noted. Muscle-based adverse reactions reported by statin users range from non-specific complaints of myalgia, through symptomatic myopathy with accompanying increases in serum CK levels, to rhabdomyolysis where the accumulation of CK due to muscle breakdown leads to kidney damage (4 per 100,000 person years) (51,52,66). While these severe myopathies are attributable to statins, the incidence of milder muscle-based reactions to statins is highly debated. However, the co-existence of elevated CK with muscle symptoms makes a causal relationship between statin exposure and muscle symptoms more likely.

In order the study the genetic underpinnings of statin intolerance it is crucial to first successfully define the phenotype in a population-based study using EMRs. Due to the lack of physician adjudicated diagnoses of statin-induced myalgia or other forms of intolerance, we must employ trends in drug use and biomarkers that best reflect the occurrence of an adverse reaction.

A phenotype standardization paper put forward by the PREDICTION-ADR consortium proposed that CK elevations of 4 or more times the upper limit of normal (ULN) should be considered for statin-induced myositis (66). In it, Alfirevic *et al.* state that this threshold of CK elevations in the presence or absence of clinical symptoms provides the right balance in preventing inclusion of patients with CK elevation due to normal variation or other causes, while simultaneously ensuring that we do not unnecessarily exclude valuable patients in studies investigating genetic factors predisposing to statin myotoxicity.

Adverse events are cited as the most common cause of discontinuation of therapy (102). Adverse reactions to statins are likely to manifest as elevated CK levels, usually in conjunction with poor adherence to statin treatment (66,102) which in turn is

associated with an increased risk of cardiovascular events compared with compliant users of statins (101).

We therefore chose to examine intolerance based on raised CK and on prescription encashment patterns representing poor adherence, such as statin discontinuation or switching, in the presence and absence of elevated CK levels in the GoDARTS study. More recently, a position paper by Banach *et al.* sought to provide a unified definition of statin intolerance – a more general counterpart to statin-associated muscle symptoms. The authors state statin intolerance affects a wider population of statin users (10-15%), unlike rhabdomyolysis, myositis and myopathy that affect less than 1 % of statin users (130). They suggest new, slightly altered parameters to define statin intolerance from prescription trends; namely the individual should be intolerant to at least two different statin therapies and they should have been on the lowest starting dose of at least one of those therapies. We therefore applied this criteria as well, to create a dose intolerant phenotype.

In order to examine the internal validity of our definitions, we examined the hazards of statin failure for those classified as intolerant, as those who are unable to take their therapy regularly and as prescribed should be at an increased risk of having a CVD-related outcome. Further, we examined the association between the phenotype groups created and a known gene risk score in the influx transporter, OATP1B1-encoding *SDLCO1B1* gene (54,131), which has been previously validated as being associated with non-autoimmune mediated statin intolerance and statin-induced myopathies.

It is also necessary to validate the use of poor adherence to reflect the statin user experiencing adverse effects on therapy. Using physician adjudicated accounts of myalgia and non-compliance to therapy in the clinical trial setting, we attempt to address the external validity of using poor adherence to reflect intolerance.

The first documented statin used in the GoDARTS cohort was in the year 1989, which reflects the beginning of statin therapy as it was made available in the United Kingdom. As of June 2014 the GoDARTS cohort contained a total of 11,566 statin users who contribute approximately 98,400 person years of statin exposure, thus providing an ideal template for a study of the long-term trends in statin usage.

3.2 Methods

Data handling

EMRs in GoDARTS are held in separate files such as biochemistry, prescribing, baseline demography, hospital emergency admissions, regular hospital admissions, outpatient hospital records etc. as described in Chapter 2. These files were cleaned and merged as shown in the consort diagram Figure 3.1. The processing and handling of each file is described below.

Statin types, dose conversions, switching, and discontinuation

Individuals with at least two prescriptions of statins were identified, there were 11,566 such statin users. Medications are categorized under the code provided by the British National Formulary (BNF), statins are classified under "2.12 – Lipid-regulating drugs" (10).

The types of statins made available by NHS Scotland, and therefore observed in the GoDARTS dataset are Simvastatin, Cerivastatin, Atorvastatin, Pravastatin, Fluvastatin and Rosuvastatin. Their usage distributions were examined and elaborated on in the results section.

3.2.1.1 Average daily dose

In order to compare dosages of statins – each statin was converted to its' equivalent dose of simvastatin in terms of LDL-lowering efficacy (14) Table 3.1. The dose conversions are described in the table below. Average daily dose was calculated by

dividing the sum of the product of the usage directions and the strength of the individual tablets for each prescription by the total number of days of statin usage, Equation 3.1. The average daily dose was 32 mg (of simvastatin or equivalency of other statins), therefore 35 mg was used as the threshold for above average daily dose tolerance.

Table 3.1 Dose conversions used for different statin types

Statin	Equivalence in simvastatin
Simvastatin	-
Rosuvastatin	4x
Atorvastatin	2x
Pravastatin	0.5X
Cerivastatin	50X
Fluvastatin	0.25X

$$\text{Average daily dose} = \{\frac{\Sigma\,(Prescription\ directions)*(strength\ of\ tablet)}{Total\ number\ of\ days\ of\ statin\ coverage}\}$$

Equation 3.1 Average daily dose

3.2.1.2 Statin discontinuation

Discontinuation was defined as there being no statin prescription in the 9 months preceding an individual's date of death, the study censor date (30th July 2013) or the date they left the data catchment area. This is to account for any periods of hospitalization that might occur, since medications dispensed at Ninewells Hospital are not included in the GoDARTS prescribing data set.

3.2.1.3 Statin switching

Physicians often switch a user's statin if they complain of adverse reactions. This is especially common between lipophilic statins (atorvastatin and simvastatin) and hydrophilic statins (pravastatin and rosuvastatin) (5), this pattern is also observed in GoDARTS. Statin switching was ascertained in our data by first classifying all the brand name drugs into their generic formulations i.e. simvastatin, atorvastatin etc. and then looking for changes in the statins consumed over time. A systemic switch was noted from Simvastatin to Atorvastatin when the latter came off patent in the UK, since this

was independent of intolerance, individuals were not excluded from being considered statin tolerant if they had such a switch. It was noted that less than 10% of statin users in GoDARTS had switched 3 or more times, therefore this threshold was applied to account for any re-challenges to statin therapy.

3.2.1.4 Percent Daily Coverage

Percent Daily Coverage (PDC) was computed by calculating the time between the first and last documented statin prescription, calculating the number of days of coverage each prescription provided depending on directions (e.g. 1/day or 2/day) and the quantity of pills dispensed. The PDC was then calculated by comparing the days of coverage provided by the prescription and the total time spent on statin treatment, Equation 3.2. The average PDC in the study was 89%, therefore 90% was used at the threshold for high adherence.

The prescription patterns indicating intolerance used in this study are similar to those used by Donnelly et al. to establish the association between statin intolerance and *SLCO1B1* genotypes in the GoDARTS study (54).

$$\text{PDC} = \{\frac{\Sigma\ (days\ of\ coverage\ provided\ by\ statin\ prescriptions)/(prescription\ directions\ )}{Total\ number\ of\ days\ of\ statin\ coverage}\}\ \text{x}\ 100$$

Equation 3.2 Percent daily coverage

3.2.1.5 Co-medications

We examined the data for commonly prescribed co-medications. The most frequently used co-medications were Amlodipine (38% of statin users), Omeprazole (32%) and Clopidogrel (17%), Clarithromycin (11%), Fluconazole (9%), and Fibrates (5%). Several of these medications are substrates for the CYP3A enzymes, and could potentially interact with statins. Known medications that inhibit CYP enzymes such as azole anti-fungals, anti-biotics, anti-retrovirals, anti-diuretics (e.g. conivaptan), amiodarone taken for arrhythmias, and verapamil, a calcium channel blocker were

categorized as potential interacting medications. Fibrates, especially gemfibrozil are contra-indicated for use with statins, as they increase the risk of rhabdomyolysis (10). All these drugs were classified as potentially interacting co-medications.

Creatine kinase

CK tests from wards such as Accidents & Emergencies (A&E) or Emergency Rooms (ER), Cardiac Care, Stroke, surgical wards and high dependency units were excluded. We excluded individuals who had a history of thyroid disease, or those that had suffered a myocardial infarction, kidney disorder, or had a hospital admission associated with an accident in the 6 months preceding the CK test result. The usable test results were categorized into normal and above the upper limit of normal, 120 IU/L for women and 180 IU/L for men (NHS Tayside Biochemistry Meta-Data). The first high CK test result while on statins or within 3 months of their last statin was used to define intolerance.

CONSORT flow chart

The processing of data files and their subsequent use in producing the final phenotypes are described in the consort diagram below Figure 3.1.

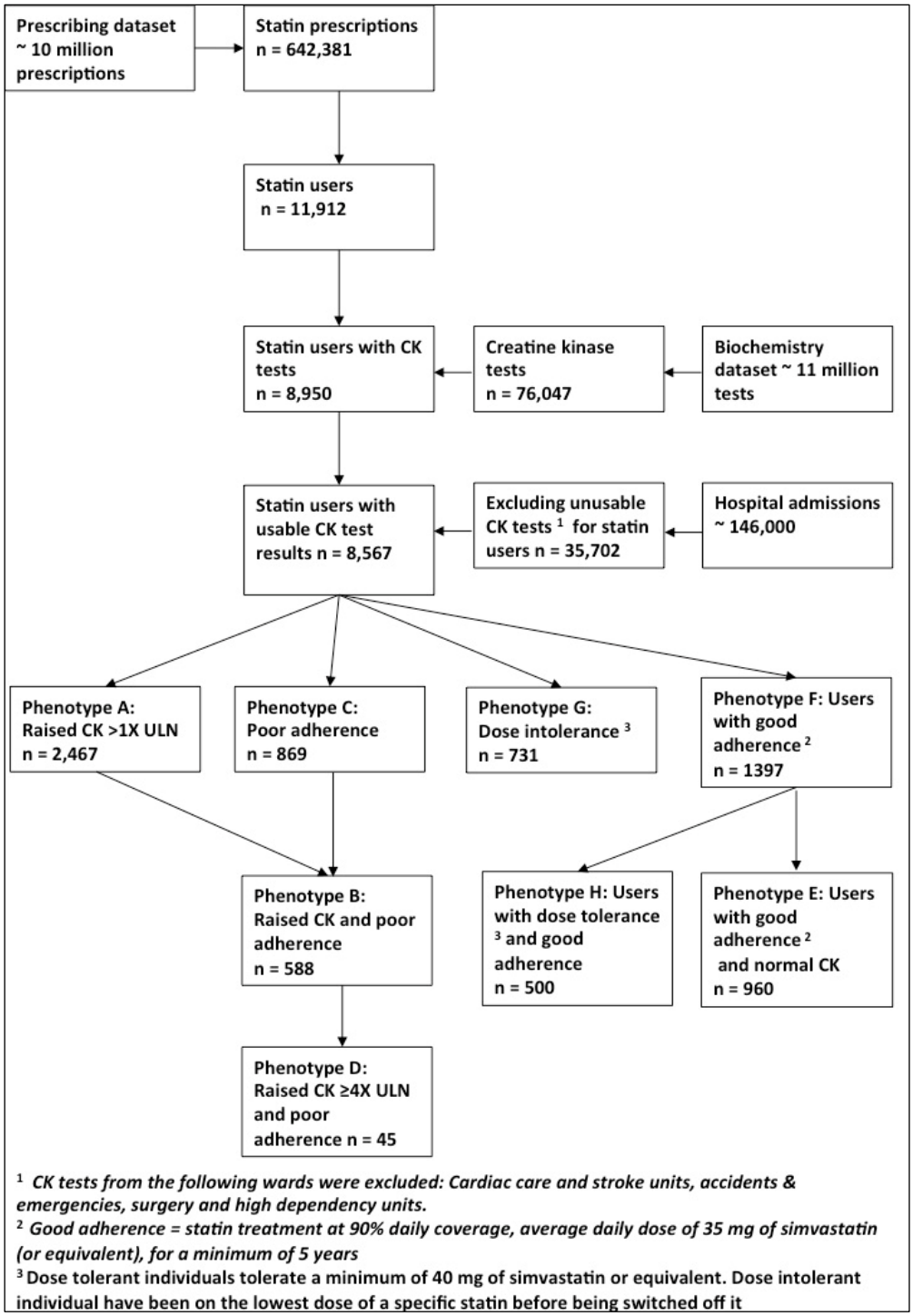


Figure 3.1 Consort diagram of the derivation of statin intolerant individuals in the GoDARTS study

In the consort diagram, phenotype groups A, B, C, D and G are classified as statin intolerant, while groups E and H are statin tolerant controls.

In Table 3.2 the phenotypes are specified along with the number of individuals meeting the criteria for each group. This number reflects those available from the

EMRs, however when examining genetic associations, the numbers available for analyses are lower to due to limited genotyping.

Table 3.2 Criteria used to create phenotypes and the number of individuals meeting the criteria in GoDARTS

Phenotypes	Elevated CK		≥3 Switches or Discontinuation	Lowest Dose Before Switch	Total
	≥ 1X ULN	≥ 4X ULN			
A: Raised CK	Yes	N/A	N/A	N/A	2467
B: Raised CK + Non-adherent	Yes	N/A	Yes	N/A	588
C: Non-adherent	N/A		Yes	N/A	869
D: Statin-induced myositis		Yes	Yes	N/A	45
G: Dose intolerance	N/A		Yes	Yes	731
E: Normal CK + Adherent	No	No	No	N/A	960
F: Adherent	N/A		No	N/A	1397
H: Dose tolerant	N/A		No	N/A	500

Compliance and myalgia in the JUPITER trial

In the JUPITER trial compliance to assigned therapy was adjudicated by pills counts. 1435 participants were found to be non-compliant to therapy, while 7046 were compliant. The trial had a placebo run-in period where participants were selected for having good compliance (115,114). Myalgia was ascertained by physicians who were blinded to the treatment status of the participant (132).

Defining phenotypes

3.2.1.6 Phenotypes of statin intolerance

We defined multiple phenotypes of intolerance, to account for the dependency of the definition of adverse reactions to statins on CK levels, and the modulation of CK by genetic variants put forward by Dubé *et al*. The effect of this variant on CK levels is the focus of Chapter 4. First we categorized statin users who had raised CKs (> 1X ULN) associated with their statin therapy (Case Phenotype A). We did this by excluding individuals who might have co-morbidities that could lead to high CK levels as described in the methods. Statin intolerance (Case Phenotype B) was defined as statin users who had a raised CK (>1X ULN) while on statins and showed patterns of poor adherence by either discontinuing their statins or switching their statins three or more times. Individuals who showed the same poor adherence patterns described above, regardless of their CK levels were classified as having general intolerance (Case Phenotype C). Individuals who were non-adherent and who had CKs raised (>4X ULN) were treated as having statin-induced myopathy or myositis (Case Phenotype D) based on parameters put forward by Alfirevic *et al.* (66). Finally, in a recent position paper, Banach *et al.* suggest that statin intolerance should be dose-related; in that an individual must have been placed on the lowest approved starting dose of the specific statin before being switched to another statin before being classified as "intolerant" (133). This specification was added to those who had discontinued therapy or switched therapy three or more times, irrespective of CK levels (Phenotype G). The lowest approved starting doses are 5 mg for Rosuvastatin, 10 mg Simvastatin, 40 mg Pravastatin, 40 mg Fluvastatin and 0.1 mg of Cerivastatin (133).

3.2.1.7 Phenotypes of statin tolerance

In order to be classified as statin tolerant, a statin user had to have no recorded high CK test results while on statins, be on a minimum average daily dose of 35 mg (calculated as equivalence of simvastatin), be on statins for a minimum of 5 years, have

a minimum of 90% daily coverage, with no discontinuation and no more than one switch between simvastatin and atorvastatin (to account for the mass switch that occurred when the latter came off patent in 2012) (Control Phenotype E). We also created a group of individuals who met the same prescribing criteria as phenotype E, irrespective of their CK levels (Control Phenotype F). This second group was created in order to form an appropriate comparison group for Case Phenotype C. Phenotype E was then further refined for a higher average dose tolerance of 40 mg (calculated as equivalence of simvastatin) which is generally the medium dose of statin therapy prescribed, in addition to the other criteria and irrespective of CK levels (Phenotype H).

The normal CK and adherent phenotype E was used for comparison with statin intolerance phenotypes (A and B), while the adherent, irrespective of CK phenotype (F) was used for comparison with the non-adherent phenotype (C) to ensure that the difference detected was not just from inducible CK levels. Finally, we compared the dose intolerance phenotype (G) to dose tolerant individuals (phenotype H).

A summary of the phenotypes and the criteria used to define them is provided in Table 3.2 Criteria used to create phenotypes and the number of individuals meeting the criteria in GoDARTS.

3.2.1.8 Baseline co-variables

Important risk factors for statin intolerance include age and female gender, use of fibrates and other interacting co-medications as listed in the British National Formulary and the use of higher dose of statins (10,66)

Additionally, we examined baseline variables such as sex, age and LDL level at time of starting statin therapy and diabetic status that could affect the probability of developing statin intolerance. Statin users commencing treatment after the occurrence of a CV event were classified as being on therapy the secondary prevention of disease, while statin users who began commenced use without any noted CV events were

classified as being on therapy for the primary prevention of disease. This is an important co-factor impacting adherence to therapy.

LDL measures within a 1-year window prior to the commencement of statin treatment were considered as a baseline LDL measurement. Since biochemistry data was only available from Fife after 2005, we do not have baseline LDL levels from individuals who were resident in Fife at the time of starting statin therapy as seen in Table 3.3.

Table 3.3 LDL testing data Tayside v. Fife

LDL testing prior to statin treatment	Tayside	Fife
Yes	16%	3.4%
No	77%	4%

Statins are prescribed to Type 2 Diabetics (T2D) upon diagnosis, since T2D is considered a major risk factor for CVD, irrespective of LDL levels. Therefore, testing for LDL prior to commencing statin therapy is low in GoDARTS. However, differences in LDL levels across phenotype comparison groups and the number of tests used are quoted in the results.

Association of phenotypes with statin failure

In order to assess the validity of the phenotypes developed, we checked their association with a coronary artery disease or ischemic event while on statin therapy, classified as "statin failure". The event could result in a hospitalization or in death. Hospitalization (SMR01) and General Register Office of Scotland (GRO) records were checked for ICD (International Classification of Diseases) (http://www.who.int/classifications/icd/en/) codes that corresponded to such an event. Adapted from Bijlsma *et al.* (134) we used the codes presented in Table 3.4 below.

Table 3.4 ICD codes used to classify statin failure (SF)

Outcome	ICD 9	ICD 10
Ischemic stroke	433 - 436	I63, I64, I67, G45
Coronary artery disease	410 – 414, 427, 789	I20-I25, I46

Association of phenotypes with the *SLCO1B1* gene risk score

The score was developed using data from two variants rs4149056 (Val174Ala) and rs2306283 (Asp130Asn) and the model provided by Donnelly *et al.* (54).

Statistical analysis

All data cleaning and statistical analyses were undertaken in SAS 9.3. Kaplan-Meier survival plots were also produced using SAS 9.3 (SAS Institute, Cary, North Carolina). All other graphs were produced in R studio (135). Statistical analyses for the JUPITER trial were performed by Dr. Dan Chasman using R (136).

Categorical variables are presented as percentages. Associations between categorical variables (sex, statin doses, statin types etc.) were assessed using the Chi square test (proc freq/chisq measures). Continuous variables (such as age, LDL measures, CK response etc.) are presented as mean, median and range (proc means). The association with continuous variables were assessed using a T tests (proc ttest in SAS) or linear regression to produce beta estimates (proc reg). Association with statin failure was assessed using Cox proportional hazards model (proc phreg). Statin failure was treated as binary outcome and intolerance status was the binary predictor. This was done for each phenotype.

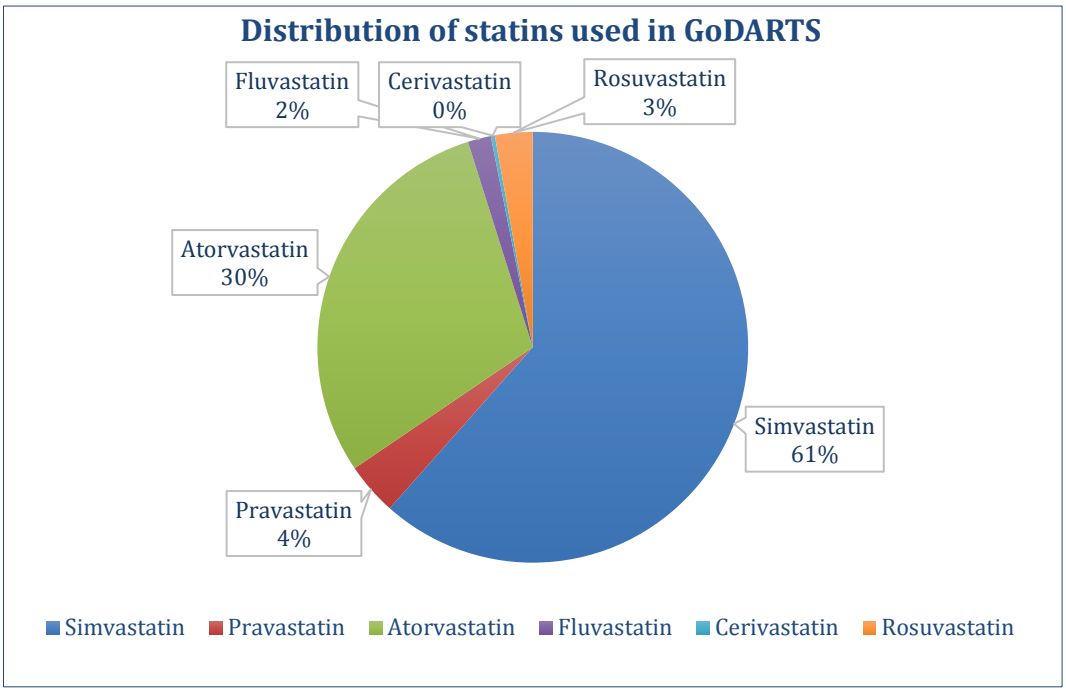
All p values reported are two-sided and values below 1×10^{-4} are represented as < 0.0001.

3.3 Results

Statin types, dose conversions, switching, and discontinuation

The vast majority of statins used in the GoDARTS were simvastatin (61%), followed by atorvastatin (30%) and the least frequently used was fluvastatin (2%). Since cerivastatin was discontinued due to reports of toxicity leading to rhabdomyolysis (11,137) the usage is below 1%. A representation of the usage is presented in Figure 3.2

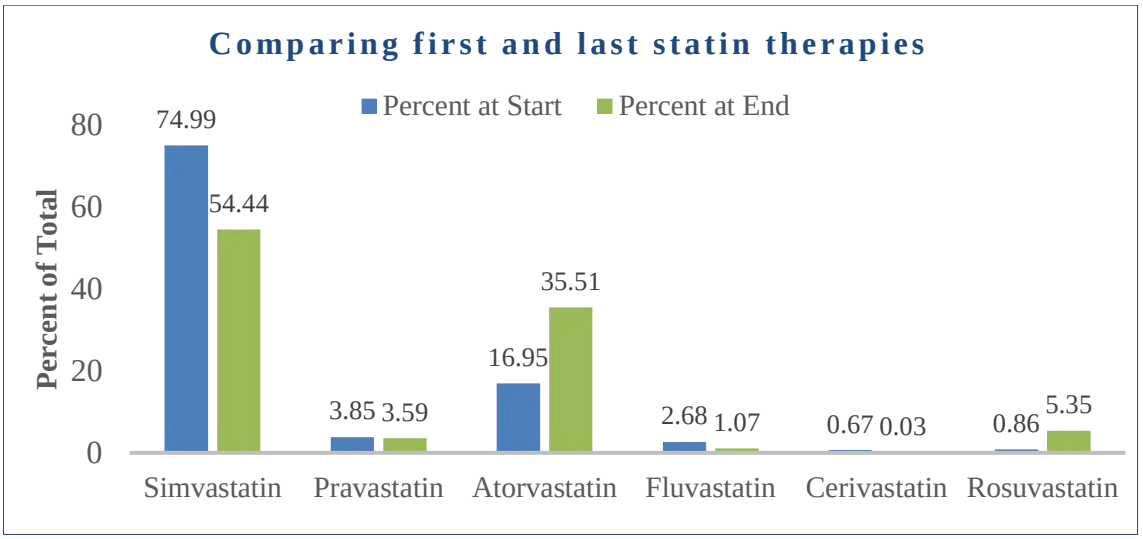
Figure 3.2 Distribution of statins used in GoDARTS



However, it is important to note that while simvastatin was most widely used, approximately 20% of statin users switched from the use of simvastatin to atorvastatin when their first and last statin treatments were compared. Conversely, there is a commensurate increase in the use of atorvastatin. This is mostly driven by the expiration of Pfizer's patent on Lipitor (brand name atorvastatin) in 2012, and the mass switch from simvastatin to atorvastatin that occurred subsequently. Rosuvastatin was the only other statin that that shows a higher percent of users at the end of treatment. A

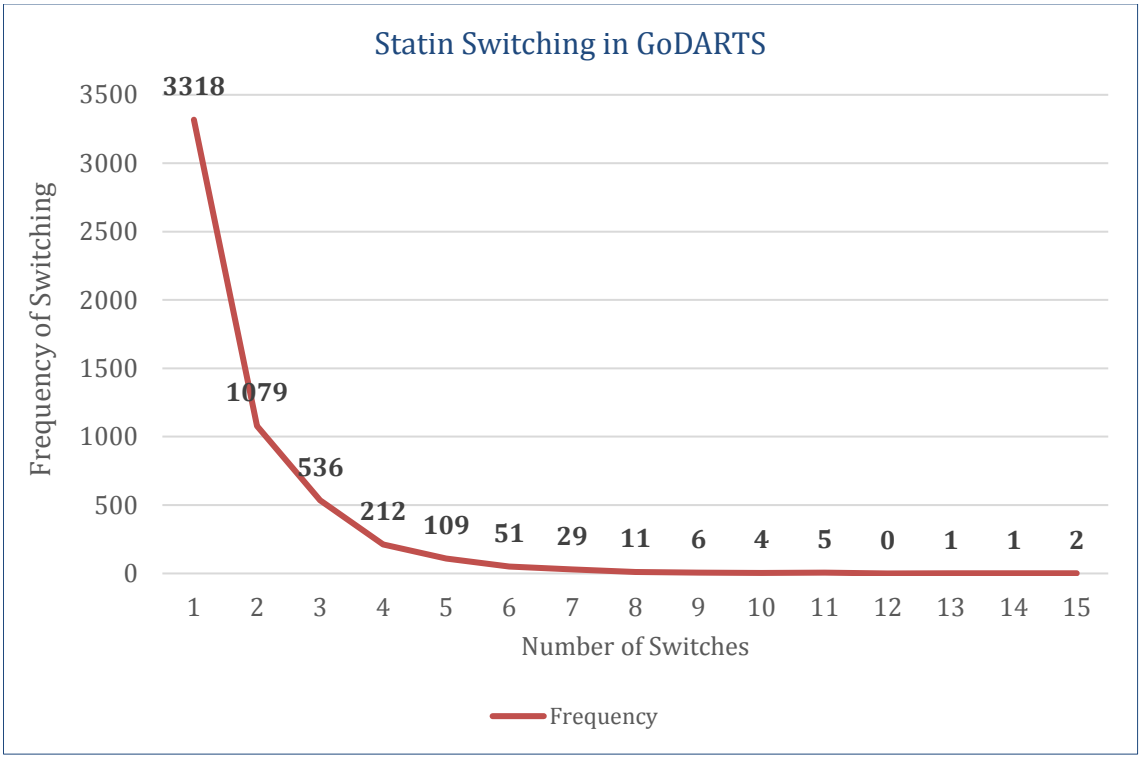
complete summary of the percentage of users of specific statins at the beginning and end of treatment is provided in Figure 3.3 below.

Figure 3.3 Transitions from first statin used to the last statin prescribed



Switching from one statin to another often occurs when a patient is unable to either show the necessary reduction in LDL levels or is unable to tolerate the drug. Therefore, statin switching is an important trend in an EMR to determine rates of intolerance. Figure 3.4 below represents the number of switches noted for statin users and frequency of their occurrence.

Figure 3.4 Number and frequency of statin switching in GoDARTS



Using the criteria described in the methods section to define discontinuation of statin treatment, there were 1721 individuals (15% of statin users) who had discontinued treatment for 9 months preceding the date of their death or leaving the data catchment area of the study censor date.

Creatine kinase

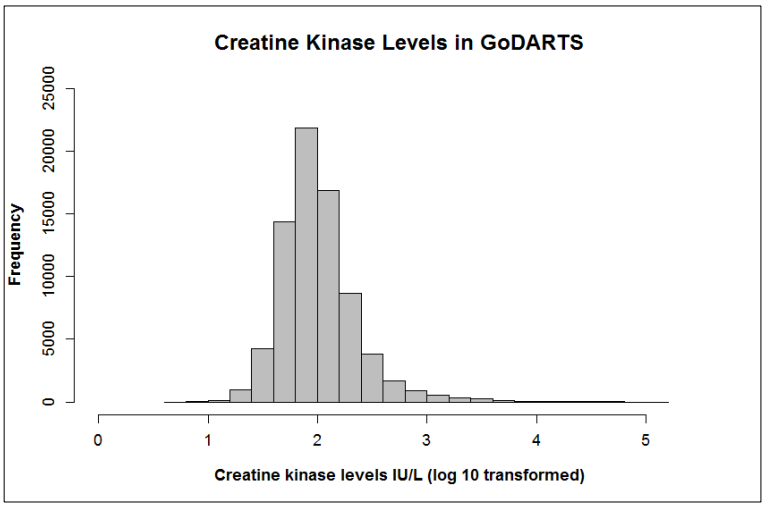


Figure 3.5 Creatine kinase (originally in IU/L) distribution in the GoDARTS population

The highest recorded CK level in GoDARTS was 129,825 IU/L. The CK measures were log10 transformed to create the histogram in Figure 3.5 above.

Defining phenotypes

Using various combinations described in the methods section phenotypes of statin intolerance (SI) and tolerance (ST) were created. They are presented in Table 3.2.

Baseline characteristics

5041 (43.6%) statin users were female. The average age for starting statin therapy was 62 (+/-11) years. Additional baseline characteristics pertaining to statin intolerance are discussed in the subsequent tables that are set up to contrast between intolerance and tolerance.

3.3.1.1 Phenotype A (Raised CK) v. E (Normal CK and Adherent)

Phenotype A (raised CK while on statin treatment) is compared to Phenotype E (those who are adherent to their therapy and have always shown normal CK while on statins) seen in Table 3.5 Phenotype A (Raised CK) v. E.

We find that while the age at which statin therapy commenced was not significantly different across the two groups, those who were intolerant were likely to be older at the end of follow-up period, indicating that they had been on statin therapy for longer. There was no significant difference in the diabetes status across the two groups. Statin tolerant individuals were more likely to have begun therapy on simvastatin. Intolerant individuals were more likely to be on simvastatin at the end of the study period. SI individuals began therapy on lower doses and were more likely to end on lower doses than those who were able to tolerate their statins. The use of potentially interacting co-medications was 48% amongst those with raised CK compared to 44% amongst tolerant controls, the difference was not statistically significant. 21% of those classified as intolerant were on statins for the secondary prevention of CVD contrasted with 23% of tolerant controls, although the difference was not statistically significant. CK levels differed between the two groups, by definition of the phenotype, where the median CK for those who were SI was 175 IU/L and 74 IU/L for those who were ST. Although the numbers available for analysis were reduced – there was no significant difference is baseline LDL levels between the groups.

Table 3.5 Phenotype A (Raised CK) v. E (Normal CK and Adherent)

Variables	Statin intolerant (Phenotype A)	Statin tolerant (Phenotype E)	Test of Association
	n = 2467	n = 960	
Mean age starting statin therapy (SD)	62 (11)	62 (10)	T value = 1.23 P value = 0.22
Mean age at censor date (SD)	73 (10)	71 (10)	T value = 4.7 P value <0.0001
Sex (% Females)	45	43	$X^2 = 1.1$ P value = 0.3
Diabetes (%)	76	79	$X^2 = 2.1$ P value = 0.15
First Statin as Simvastatin (%)	71	78	$X^2 = 47.5$ P value <0.0001
Last Statin as Simvastatin (%)	47	41	$X^2 = 120$ P value <0.0001
Starting dose as "low" (< 20mg/day) (%)	80	53	$X^2 = 165$ P value <0.0001
Ending dose as "high" (≥80 mg/day) (%)	21	36	$X^2 = 376$ P value <0.0001
CYP inhibitors or substrates (Yes v. No) (%)	48	44	$X^2 = 3.13$ P value = 0.08
Statin use for secondary prevention of CVD (%)	21	23	$X^2 = 0.4$ P value = 0.50
CK levels (IU/L)			
Median	175	74	T* value = 15 P value <0.0001
Minimum	8	11	
Maximum	21,214	7,457	
Mean	282	97	
LDL at baseline (mmol/L)	n= 192	n= 219	
Median	3.1	3.1	T value = 0.22 P value = 0.82
Minimum	0.8	0.4	
Maximum	6.3	8.7	
Mean	3.2	3.2	

T: unequal variances*

3.3.1.2 Phenotype B (Raised CK + Non-Adherent) v. E (Normal CK + Adherent)

Next we compared individuals who had raised CK while on statins and who also either discontinued their treatment entirely or switched their statin 3 or more times (Phenotype B) to the same statin tolerant group as before (Phenotype E) seen in Table 3.6 Phenotype B (Raised CK + Non-adherent) v. E (Normal CK + Adherent).

SI individuals were more likely to start their treatment at an earlier age, but were older at end of follow-up. More women were likely to be SI. There was no significant in the diabetes status across the two groups. More ST individuals began treatment on simvastatin and were also more likely to be on simvastatin at the end of the study period. SI individuals more likely to begin and end treatment on a lower doses. The use of potentially interacting co-medications was 52% amongst those classified as SI compared to 44% amongst tolerant controls, the difference was statistically significant. 27% of those classified as intolerant were on statins for the secondary prevention of CVD contrasted with 23% of tolerant controls, the difference was not statistically significant. Again, by definition criteria SI individuals had higher CK levels, median CK was 200 IU/L and 76 IU/L for ST individuals. Median baseline LDL levels were slightly higher for SI individuals (3.5 mmol/L) compared to those who were ST (3.2 mmol/L).

Table 3.6 Phenotype B (Raised CK + Non-adherent) v. E (Normal CK + Adherent)

Variables	Statin intolerant (Phenotype B)	Statin tolerant (Phenotype E)	Test of Association
	n = 588	n = 960	
Mean age starting statin therapy (SD)	60 (10)	62 (10)	T value = 2.8 P value= 0.005
Mean age at censor date (SD)	72 (10)	71 (10)	T value = 2.6 P value= 0.01
Sex (% Females)	50	43	X^2 = 7.3 P value = 0.007
Diabetics (%)	78	79	X^2 = 0.17 P value = 0.68
First Statin as Simvastatin (%)	64	71	X^2 = 87.4 P value < 0.0001
Last Statin as Simvastatin (%)	31	41	X^2 = 186.5 P value < 0.0001
Starting dose as "low" (< 20mg/day) (%)	85	53	X^2 = 160.6 P value < 0.0001
Ending dose as "high" (≥80 mg/day) (%)	22	36	X^2 = 335.8 P value < 0.0001
CYP inhibitors or substrates (Yes v. No) (%)	52	44	X^2 = 9.1 P value = 0.0025
Statin use for secondary prevention of CVD (%)	27	23	X^2 = 1.8 P value = 0.18
CK levels			
Median	200	76	T value = 40.4 P value < 0.0001
Minimum	120	17	
Maximum	12,700	179	
Mean	306	81	
LDL at baseline	n= 87	n= 219	
Median	3.5	3.2	T* value = 1.83 P value = 0.07
Minimum	1.1	0.4	
Maximum	5.5	8.7	
Mean	3.5	3.2	

*T * represents unequal variances*

3.3.1.3 Phenotype C (Poor adherence, irrespective of CK) v. F (Adherent, irrespective of CK)

We then compared individuals who were unable to stay on their statin therapy, evidenced by 3 or more switches or discontinuing therapy (phenotype C to individuals highly adherent to their statins, irrespective of their CK elevations (phenotype F) seen in Table 3.7 .

SI individuals were younger at the time of starting therapy, they were not older at the end of follow-up and were more likely to be female. They were less likely to start or end therapy on simvastatin. They were more likely to start therapy on a low dose (<20 mg) and less likely to be on a high dose (>80 mg) at the end of study follow-up period. The use of potentially interacting co-medications was significantly higher (59%) amongst those classified as SI compared to amongst tolerant controls (43%). Of those classified as intolerant, 31% were on statins for the secondary prevention of CVD contrasted with 22% of tolerant controls, the difference was statistically significant. In spite of not being factored into the phenotypic definition, CK levels were significantly different across the two groups. SI individuals had a median CK of 116 IU/L, compared to 102 IU/L for individuals who were tolerant to statin therapy, but both were with the normal range. Median LDL levels were not significantly different across the two groups.

Table 3.7 Phenotype C (Poor adherence, irrespective of CK) v. F (Adherent, irrespective of CK)

Variables	Statin intolerant (Phenotype C)	Statin tolerant (Phenotype F)	Test of Association
	n = 869	n = 1430	
Mean age starting statin therapy (SD)	58 (9)	61 (10)	T value = 7.6 P value < 0.0001
Mean age at censor date (SD)	72 (9)	71 (10)	T value = 1.4 P value= 0.16
Sex (% Females)	47	44	X^2 = 2.5 P value = 0.11
Diabetes (%)	82	78	X^2 = 6.4 P value = 0.01
First Statin as Simvastatin (%)	59	70	X^2 = 163 P value <0.0001
Last Statin as Simvastatin (%)	17	41	X^2 = 306 P value <0.0001
Starting dose as "low" (< 20mg/day) (%)	90	55	X^2 =249 P value <0.0001
Ending dose as "high" (≥80 mg/day) (%)	34	38	X^2 =245 P value <0.0001
CYP inhibitors or substrates (Yes v. No) (%)	59	43	X^2 = 48 P value < 0.0001
Statin use for secondary prevention of CVD (%)	31	22	X^2 = 11.6 P value = 0.0006
CK levels			
Median	116	102	T value = 2.6 P value = 0.0095
Minimum	21	17	
Maximum	3,271	5,217	
Mean	173	144	
LDL at baseline	n= 64	n= 338	
Median	3.3	3.2	T value = 1.18 P value = 0.24
Minimum	1.5	0.4	
Maximum	6.4	8.7	
Mean	3.5	3.4	

3.3.1.4 Phenotype D (Statin-induced myositis) v. E (Normal CK + Adherent)

Next we compared a group of individuals who showed CK elevations 4 or more times the upper limit of normal while on statins and had either switched statin 3 or more times or discontinued treatment (Phenotype D) to highly adherent individuals with normal CKs (Phenotype E) seen in Table 3.8 Phenotype D (Statin-induced myositis) v. E. SI individuals started therapy at a younger age, although there was no difference in the ages across the groups at the end of follow-up period. SI individuals were less likely to start or end therapy on simvastatin. They were more likely to begin on a low dose and also less likely to be on a higher dose at the end of follow-up period. The use of potentially interacting co-medications was 60% amongst those classified as SI compared to 44% amongst tolerant controls, the difference was statistically significant. 31% of those classified as intolerant were on statins for the secondary prevention of CVD contrasted with 23% of tolerant controls, the difference was marked by did not achieve statistical significance. By phenotype definition CK levels varied drastically across the groups, the median in the SI group was 903 IU/L and 76 in the ST group. LDL levels did not vary significantly across the groups, although the numbers in the analyses are quite low.

Table 3.8 Phenotype D (Statin-induced myositis) v. E (Normal CK + Adherent)

Variables	Statin intolerant (Phenotype D)	Statin tolerant (Phenotype E)	Test of Association
	n = 45	n = 960	
Mean age starting statin therapy (SD)	56 (10)	62 (10)	T *value = 3.35 P value= 0.0004
Mean age at censor date (SD)	71 (10)	71 (10)	T value = 0.29 P value= 0.77
Sex (% Females)	36	43	$X^2 = 0.96$ P value = 0.3
Diabetes (%)	80	79	$X^2 = 0.03$ P value = 0.87
First Statin as Simvastatin (%)	53	71	$X^2 = 36$ P value <0.0001
Last Statin as Simvastatin (%)	18	41	$X^2 = 134$ P value <0.001
Starting dose as "low" (< 20mg/day) (%)	82	53	$X^2 = 15$ P value = 0.0005
Ending dose as "high" (≥80 mg/day) (%)	31	36	$X^2 = 98$ P value <0.0001
CYP inhibitors or substrates (Yes v. No) (%)	60	44	$X^2 = 4.7$ P value = 0.03
Statin use for secondary prevention of CVD (%)	31	23	$X^2 = 0.9$ P value = 0.34
CK levels (IU/L)			
Median	903	76	T value = 42 P value < 0.0001
Minimum	483	17	
Maximum	19,156	179	
Mean	1604	81	
LDL at baseline (mmol/L)	8	549	
Median	3.6	3.1	T value = 1.3 P value = 0.2
Minimum	3.1	0.4	
Maximum	4.7	8.7	
Mean	3.7	3.2	

3.3.1.5 Phenotype G (Dose-related Intolerant) v. H (Dose tolerant + Adherent)

We defined this phenotype based on criteria proposed by Banach et al. that in order to be classified as intolerant, an individual must have been on the lowest prescribed daily dose of a specific statin before being switched to another statin. Therefore, individuals who had been on the lowest approved dose of a specific statin before being switched to another, or discontinuing it completely were compared to individuals with a tolerant phenotype where users had been on a minimum of 40 mg of simvastatin (or equivalent), for a minimum of 5 years, with a 90% daily coverage rate with no switches or discontinuations. This is presented in Table 3.9 Phenotype G (Dose-related intolerant) v. H (Dose tolerant + Adherent).

These groups started statin therapy at the same age, but SI individuals had used statins for 6 months longer at the end of follow-up. SI individuals were less likely to have started or ended therapy on simvastatin or atorvastatin compared to tolerant individuals. SI individuals were more likely to have begun therapy on a dose lower than 20 mg of simvastatin (or equivalent). Statin tolerant individuals were more likely to be on a therapy of 80 mg or higher at end of follow-up. However, this is an artefact of the definition of tolerance that involves a high average daily dose. The use of potentially interacting co-medications was significantly higher, (51%) amongst those classified as SI compared to tolerant controls (42%). 28% of those classified as intolerant were on statins for the secondary prevention of CVD contrasted with 25% of tolerant controls, the difference was statistically significant. Even though CK levels were not included in the definitions, we see significantly higher median CK levels amongst those who were intolerant (98 IU/L v. 85 IU/L), however both were within the normal range.

A main concern with a phenotype that is dose dependent is that differences might arise from the LDL cholesterol reduction required between tolerant and intolerant individuals. However, we find that there is no significant difference in LDL levels prior to starting therapy between the two groups.

Table 3.9 Phenotype G (Dose-related intolerant) v. H (Dose tolerant + Adherent)

Variables	Statin intolerant (Phenotype G)	Statin tolerant (Phenotype H)	Test of Association
	n = 731	n = 443	
Mean age at start of statin therapy (SD)	60 (10)	60 (10)	T value = 0.15 P value= 0.9
Years on statin therapy (SD)	10 (5)	9.5 (3)	T value = 2.7 P value= 0.007
Sex (% Females)	48	46	$X^2 = 2$ P value = 0.16
Diabetics (%)	92	90	$X^2 = 0.97$ P value = 0.33
First Statin as			$X^2 = 69$ P value = 2 x 10^{-15}
Simvastatin (%)	59	65	
Atorvastatin (%)	21	30	
Last Statin as			$X^2 = 129$ P value = 3 x 10^{-34}
Simvastatin (%)	31	36	
Atorvastatin (%)	44	61	
Starting dose as			$X^2 = 303$ P value < 0.0001
< 20 mg/day (%)	94	37	
20-80 mg/day (%)	5.7	57	
≥80 mg/day (%)	0.3	6	
Ending dose as			$X^2 = 249$ P value < 0.0001
< 20 mg/day (%)	43	0.40	
20-80 mg/day (%)	34	57.4	
≥80 mg/day (%)	23	50	
CYP inhibitors or substrates (Yes v. No)	51%	42%	$X^2 = 18$ P value < 0.0001
Statin use for secondary prevention of CVD	28%	25%	$X^2 = 1.1$ P value = 0.3
CK levels (IU/L)			
Median	98	85	T* value = 4.3 P value < 0.0001
Minimum	13	19	
Maximum	12735	1369	
Mean	170	107	
LDL levels at baseline (mmol/L)	n = 149	n = 285	
Median	3.2	3.1	T* value = 0.10 P value 0.38
Minimum	1.1	0.5	
Maximum	6.4	8.7	
Mean	3.1	3.2	

Association between non-compliance and reports of myalgia in a RCT

Since, we do not have accounts of physician diagnosed or patient-reports of myalgia in the GoDARTS study, we sought to validate the use of poor adherence as a proxy. Due to the primary nature of data collected for myalgia and non-compliance in the clinical trial setting, we sought to examine this in the JUPITER trial. We observed that those having myalgia had 2.2 times the odds of being non-compliant with their therapy (p value = 1×10^{-10}).

Association of phenotypes with statin failure in GoDARTS

In order to determine the validity of each of the phenotypes developed, the association with the main negative outcome of statin intolerance – failure of statin therapy (SF), was tested. Of the 11,912 who were ever on statin treatment, 3123 (26%) had an ischemic or coronary artery event after commencing therapy. We performed a Cox proportional hazards regression for each comparison group in a 10 year follow-up period.

The hazards of SF were 1.5 times higher for those who had raised CK (A) while on therapy and 1.9 times higher for those who had raised CK and poor adherence (B) compared to those with consistently normal CK and good adherence. For those with poor adherence (C) the hazards were 1.7 times higher compared to those with good adherence, irrespective of CK levels (F). Those meeting our criteria for statin-induced myositis had 2 times the hazards of SF compared to those with normal CK and good adherence (E). Finally, those showing intolerance to even the lowest dose of specific statin therapies (G) had 1.9 times the hazards of SF compared to dose-tolerant controls (H). All models were adjusted for age at time of starting therapy, sex and whether the statin was prescribed after a CV-event (secondary prevention) or not (primary prevention). Analyses stratified by primary or secondary prevention and related risks of

statin intolerance are also presented. Table 3.10 presents the number of events of "statin failure" that occurred in each phenotypic group.

Table 3.10 Number of events of statin failure in phenotype groups

Phenotype groups	Number of events	Percentage of events (%)
Case: A	213	40
Case: B	134	41
Case: C	208	45
Case: D	15	58
Case: G	199	42
Control: E	152	30
Control: F	224	30
Control: G	104	31

Before undertaking any statistical analyses it is clear that statin intolerant "case" groups had a much higher percent of events compared to tolerant groups. A complete summary with hazard ratios, standard errors and p values in each category are presented in Table 3.11.

Table 3.11 Hazards of statin failure for each phenotype of intolerance. All main effects models were adjusted for age at start of therapy, sex and whether the therapy was for the primary or secondary prevention of CVD. Stratified results are presented for those who were on statins for primary or secondary prevention of CVD.

Phenotypes	Hazards of Statin Failure	Standard Error	P value
A v. E: Raised CK	1.5	0.12	0.0008
Primary	1.70	0.16	0.0003
Secondary	1.14	0.20	0.5
B v. E: Raised CK + non-adherent	1.9	0.14	<0.0001
Primary	2.30	0.19	<0.0001
Secondary	1.50	0.22	0.07
C v. F: Non-adherent	1.7	0.11	<0.0001
Primary	2.6	0.15	<0.0001
Secondary	1.0	0.17	0.92
D v. E: Statin-induced myositis	2.1	0.31	0.015
Primary	2.0	0.38	0.076
Secondary	2.4	0.54	0.12
G v. H: Dose intolerance	1.9	0.14	<0.0001
Primary	2.1	0.2	0.0001
Secondary	1.6	0.2	0.028

3.3.1.6 Hazards of statin failure for raised CK stratified by primary or secondary prevention

For primary prevention, the hazards of statin failure for those classified as having raised CK while on statins were 1.7 times that compared to tolerant controls, where 64% of all individuals in the analysis had the outcome. However, due to the reduced numbers in the secondary prevention group, and the high proportion of individuals who had another CV event (90%) the association was not significant. Results seen in Table 3.11and Figure 3.6.

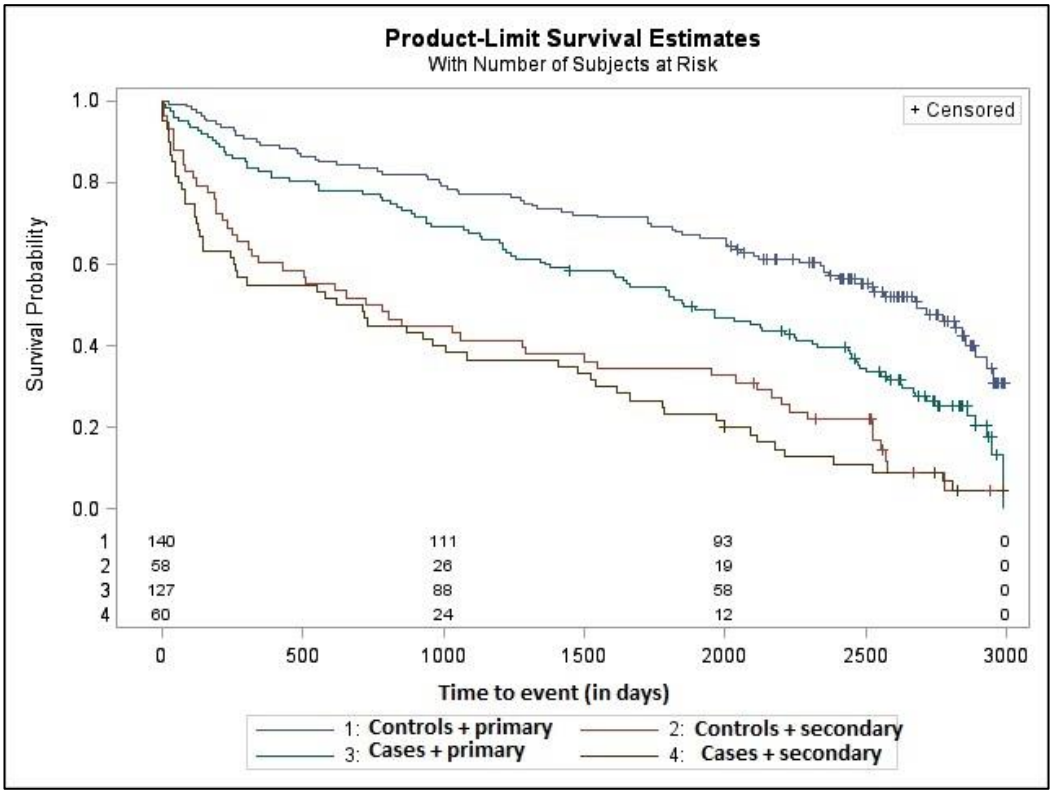


Figure 3.6 Raised CK: Kaplan-Meier of hazards of statin failure by statin intolerance status stratified by whether therapy was for primary of secondary prevention of CVD. Primary refers to primary prevention, while secondary refers to secondary prevention.

3.3.1.7 Hazards of statin failure for raised CK and non-adherence stratified by primary or secondary prevention

For primary prevention, the hazards of statin failure for those classified as having raised CK and being non-adherent to their statins were 2.3 times higher compared to tolerant controls, where 61% of the individuals in the analysis had the outcome. Once again due to the reduced number of individuals receiving statins for secondary prevention, and the accompanying high proportion that had another CV event (92%), the association is not statistically significant. Results seen in Figure 3.7 and Table 3.11.

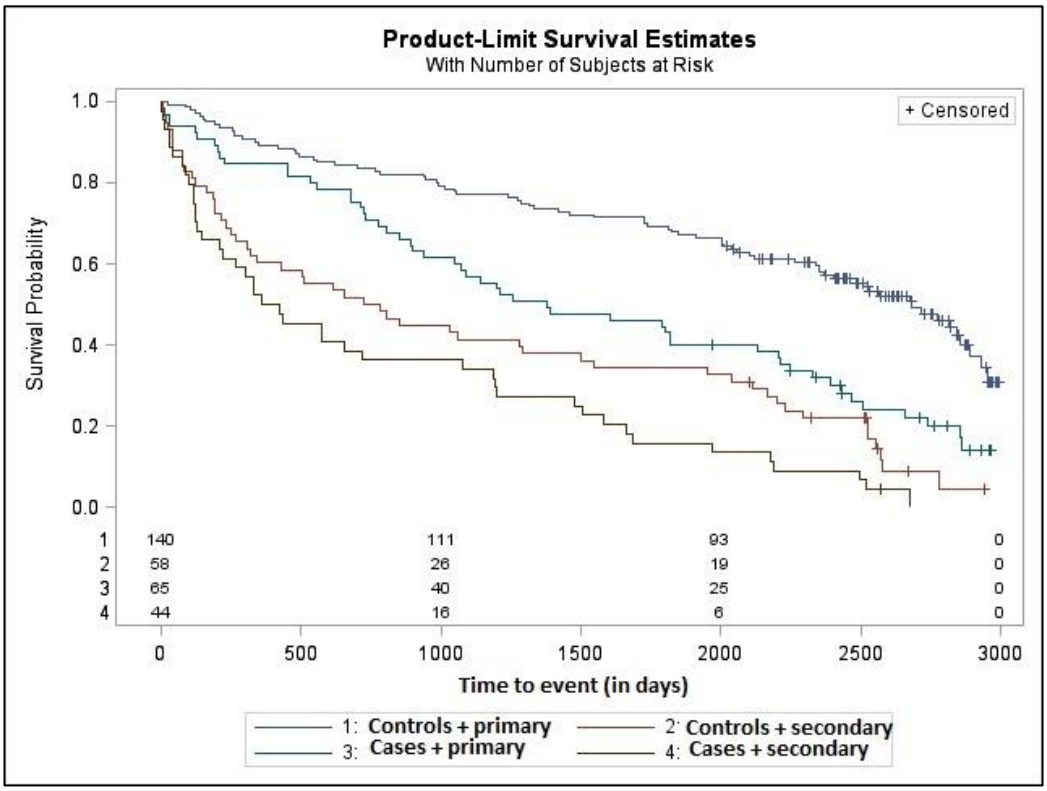


Figure 3.7 Raised CK and non-adherent: Kaplan-Meier of hazards of statin failure by statin intolerance status stratified by whether therapy was for primary of secondary prevention of CVD. Primary refers to primary prevention, while secondary refers to secondary prevention.

3.3.1.8 Hazards of statin failure for non-adherence stratified by primary or secondary prevention

For primary prevention, the hazards of statin failure for those classified as non-adherent to their statins were 2.6 times higher compared to adherent controls, where 62% of the individuals in the analysis had the outcome. Once again due to the reduced number of individuals receiving statins for secondary prevention, and the accompanying high proportion that had another CV event (90%), the association between adherence and statin failure is not statistically significant. Results seen in Figure 3.8 and Table 3.11.

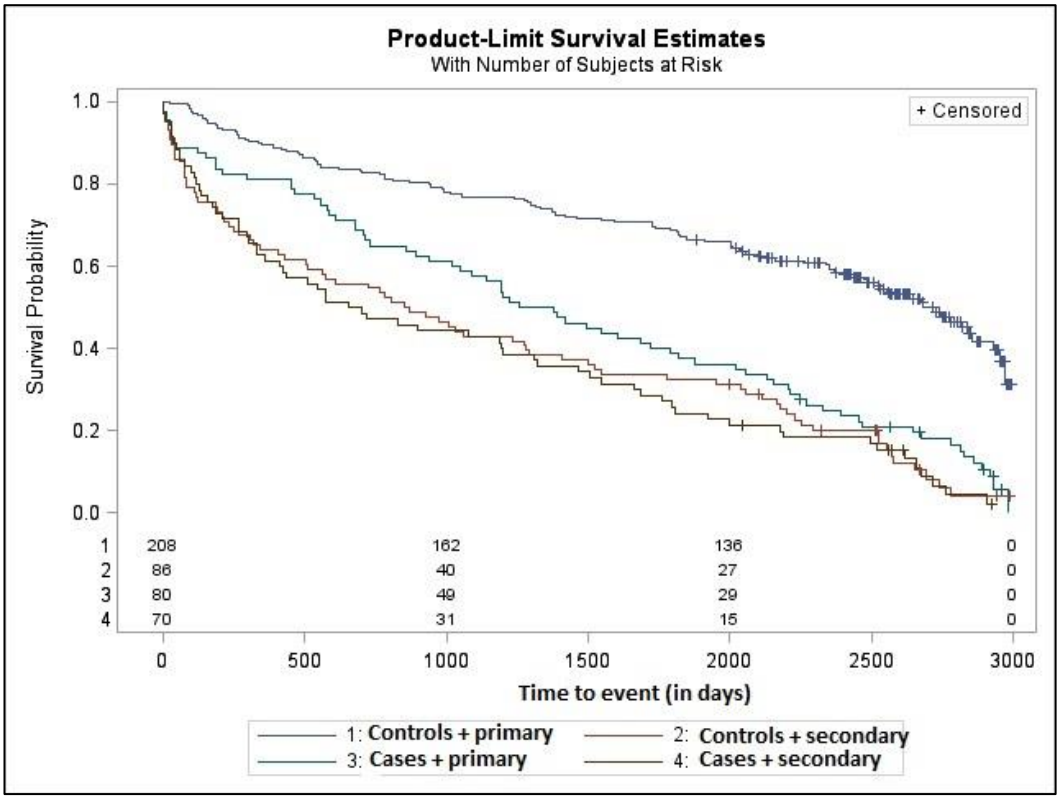


Figure 3.8 Non-adherent: Kaplan-Meier of hazards of statin failure by statin intolerance status stratified by whether therapy was for primary of secondary prevention of CVD. Primary refers to primary prevention, while secondary refers to secondary prevention.

3.3.1.9 Hazards of statin failure for statin-induced myositis stratified by primary or secondary prevention

For primary prevention, the hazards of statin failure for those classified as having statin-induced myositis were 2 times higher compared to tolerant controls, where 55% of the individuals in the analysis had the outcome. For the secondary prevention, the hazards of statin failure were 2.4 times higher for those who developed statin-induced myositis compared to tolerant controls. However, due to low numbers in both primary and secondary prevention arms of this analysis, the associations were not statistically significant. Results seen in Figure 3.9 and Table 3.11.

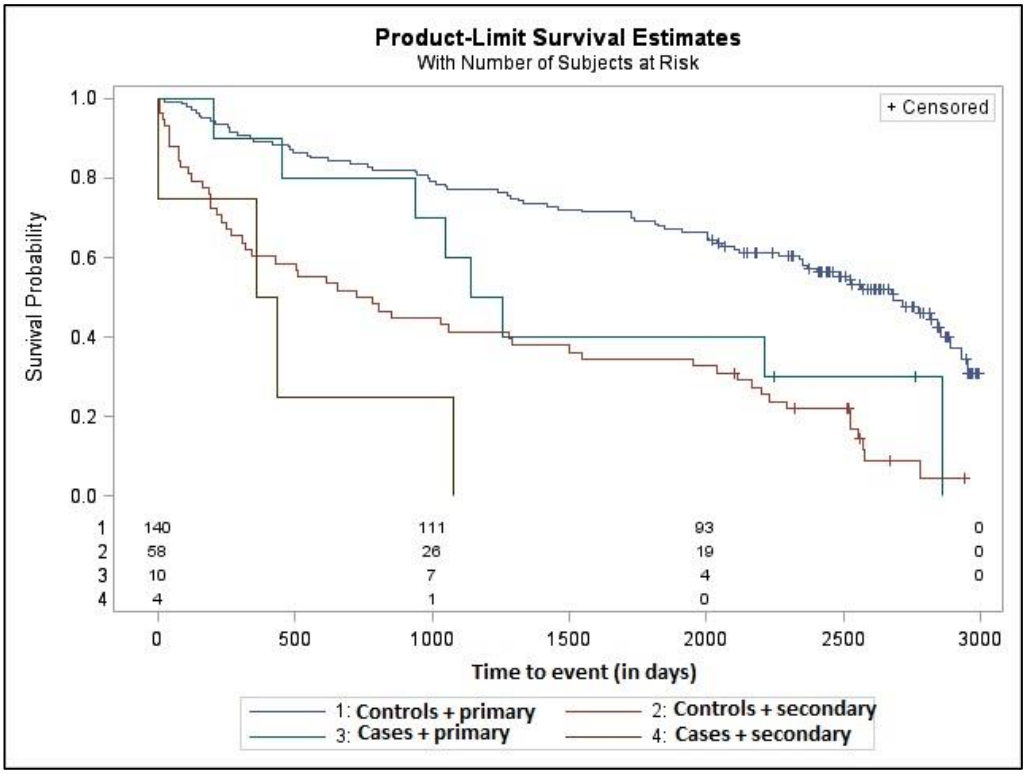


Figure 3.9 Statin-induced myositis: Kaplan-Meier of hazards of statin failure by statin intolerance status stratified by whether therapy was for primary of secondary prevention of CVD. Primary refers to primary prevention, while secondary refers to secondary prevention.

3.3.1.10 Hazards of statin failure for dose intolerance stratified by primary or secondary prevention

For primary prevention, the hazards of statin failure for those classified as dose intolerant were 2.1 times higher compared to dose tolerant controls, where 69% of the individuals in the analysis had the outcome. For the secondary prevention of CV events, the hazards for those who were dose intolerant were 1.6 times that of those who could tolerate higher doses. In this arm of the analysis 93% had another CV event. Results seen in Figure 3.10 and Table 3.11.

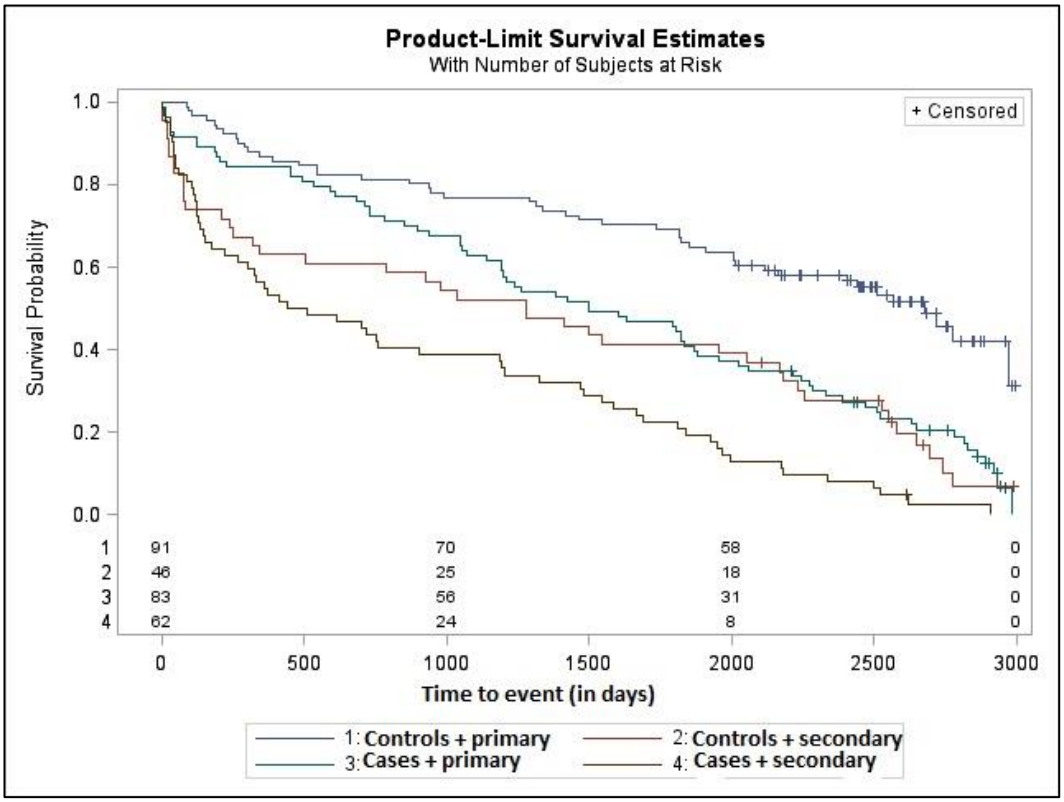


Figure 3.10 Dose intolerance: Kaplan-Meier of hazards of statin failure by statin intolerance status stratified by whether therapy was for primary of secondary prevention of CVD. Primary refers to primary prevention, while secondary refers to secondary prevention.

Association between phenotypes and *SLCO1B1* gene risk score

Using the gene risk score created by Donnelly *et al.* that was shown to be associated with SI, we attempt to validate our phenotype. The results are presented in Table 3.12 and show that all the groups created are strongly associated with the gene risk score, where those classified as cases appear to carry the deleterious genotypes.

Table 3.12 Phenotype validation using *SLCO1B1* gene risk score

Phenotype	*SLCO1B1* gene risk score	
	Beta (SE)	**P value**
A: Raised CK	0.43 (0.14)	0.0016
B: Raised CK + non-adherence	0.48 (0.18)	0.009
C: Non-adherence	0.43 (0.19)	0.02
G: Dose intolerance	0.56 (0.17)	0.0012

3.4 Discussion & conclusions

The aim of this chapter was to describe trends in the population level usage of statins that reflected an inability to tolerate the medication, using the most up-to-date knowledge in field. Defining statin intolerance using EMR presents several challenges, most notably the necessary use of surrogates to define the occurrence of an adverse event such as patient reports of myalgia etc. Therefore, our findings had to be examined for both internal validity (did those who were intolerant have increased risk of suffering a CV event) and external validity (are the trends of poor adherence associated with increased reports of adverse events).

We observed, individuals in the statin tolerant groups were more likely to have started therapy at a later age (phenotypes B, C and D), however, this difference in age was not noted in the raised CK (phenotype A) and dose-related intolerance groups (phenotype G). The younger age points toward a potential difference in whether therapy was commenced for the primary or secondary prevention of CVDs. We observe that those

who began therapy at a younger age were more likely to be intolerant to their medications, and to be on therapy for secondary prevention compared to tolerant controls. The majority of individuals in this study were on statin therapy for primary prevention of CVD. Consistent with reported risk factors, females were more likely to have raised CK and be non-adherent to their medication (phenotype B), but there was no significant difference in genders across other phenotypes. Those who were non-adherent to therapy (phenotype C) were more likely to have type 2 diabetes (T2D), however diabetic status did not differ significantly across other phenotypes. While reports of statin use leading to the development of T2D have gained traction (138,139), we did not assess the temporal relationship between the two in this analysis.

We observed that those classified as intolerant were less likely to have started therapy on simvastatin (across all groups), at the end of follow up phenotype B, C, D and G were less likely to be on simvastatin. This effect is blunted for those with raised CKs (phenotype A) as some of the individuals included would have had asymptomatic increases in CK, and their statins would likely remain unchanged. As such, phenotype A is most prone to detecting false positives. Due to experimental design individuals classified as tolerant were on higher starting and ending doses. Interestingly, those classified as intolerant began therapy on daily doses lesser than 20 mg of simvastatin (or equivalent of other statins), contrary to reports that intolerance might develop in response to initiation at high dose, here it appears that it might occur in response to gradual dose escalation, as seen in phenotype G – where merely 0.3% of intolerant individuals began on a daily dose ≥80 mg, while 23% ended on such a dose. The concurrent use of interacting medications was significantly higher amongst those who became intolerant to their therapy, confirming the use of CYP inhibitors or substrates as risk factors for statin intolerance and statin-induced myopathy.

We observe that while CK measures differ significantly across the comparison groups, by experimental design for phenotypes A, B, C and D; for phenotypes C and G where intolerance was adjudicated irrespective of CK elevations, levels were higher amongst those who were intolerant. This indicates that there is some muscle damage in response to statin usage for those unable to tolerate the drug, even if the change in CK is not clinically significant.

Crucially, there was no significant difference in baseline LDL measures, reflecting that it is not merely LDL cholesterol levels at baseline that determine the type of statin and dose an individual is prescribed. This indicates that the overall risk of CVD that factors in age, BMI, family history and co-morbidities also inform the statin treatment an individual receives.

Clinically the attribution of muscle-based side effects to statin exposure is strengthened by an individual having on-treatment circulating levels of CK higher than a standard reference range. However, since many patients do not have recorded pre-treatment CK levels in GoDARTS, it is not possible to assess relative increases in CK levels.

We report, in a RCT-setting, that those with documented myalgia had twice the odds of non-compliance to therapy in a clinical trial setting, where both outcomes were adjudicated. This serves as the external validation of our use of non-compliance parameters, and conversely, our use of compliance or adherence as markers of tolerance to therapy.

The phenotypes of intolerance developed in this study appear to increase the risk of CV-related events. This study is especially powered to detect these effects amongst those seeking primary prevention of CVD as the vast majority of statin users in the GoDARTS cohort are placed on statins for primary prevention. We represent the increased risk by intolerance status stratified by primary or secondary prevention in

Kaplan-Meier survival plots – where being classified as intolerant shows a higher risk in both arms. Behaviorally, it seems plausible that those on primary prevention are more likely to discontinue therapy at the onset of symptoms of intolerance, than those who are using statins for secondary prevention.

Further, our validation of the association between the gene-risk score in *SLCO1B1* put forward by Donnelly *et al.* across all the phenotype of statin intolerance created shows that the definitions applied are robust. Additionally it implies that the outcomes seen are affected by the rate of efflux of statins that is regulated by the gene.

While our calculations of adherence are necessarily a proxy for whether the individual actually consumed the medicine, it is unlikely that many subsequent prescriptions would be encashed if the patient were not taking the medication with some degree of regularity.

Advantages: The thresholds used for phenotypes in this study have been previously validated and some are more stringent than used before. The usage of statin switching to define statin intolerance in a population study is unusual, but is more likely to reflect intolerance than dose reductions which could reflect better LDLc response in those with dose reductions.

Future directions: The GoDARTS study provides a unique opportunity to study a population's health and drug usage trends, with data available from over two decades. The uninterrupted prescribing history, juxtaposed with medical records and genotype data can provide insights into genetic factors associated with drug use patterns and outcomes in a real-world setting. Subsequent analyses can examine the association of genetic variants with these phenotypes of statin intolerance, and arrive at valid conclusions.

4 The association of variants in *CKM* and *LILRB5* with creatine phosphokinase levels

4.1 Introduction

Creatine phosphokinase (CK) is an enzyme, 381 amino acids in length (140). It catalyzes the reversible reaction that utilizes creatine to produce phosphocreatine and ADP by the dephosphorylation of ATP. This is an exergonic reaction, and is important in the maintenance of energy homeostasis in all muscle tissues. CK subunits are present in the brain (CK-B) and skeletal muscles (CK-M), and an isozyme is present in the heart (CK-MB) (141).

The enzyme creatine phosphokinase or CK is used as a marker for tissue damage, muscle breakdown, muscular dystrophy, infection, pulmonary infarction, acute kidney failure, heart attacks, rheumatoid arthritis and some liver diseases. Notably, it is used a marker of muscle damage or myopathy in adverse reactions to statins. CK is a routine biochemical test performed in the clinical setting with widespread applicability.

In 2014 Dubé *et al.* performed an original study with 3412 Caucasian statin users recruited in Quebec, Canada (119). The 1262 cases had statin-related muscle symptoms based on a clinician's assessment that resolved upon withdrawal or dose reduction. 2150 controls were treated with a moderate or higher dose of statin, for a minimum of 3 months. However, 954 past sufferers of statin-induced myotoxicity were removed from the study, considerably impacting the power to detect the effect of these variants on CK levels in statin intolerant individuals. They applied exclusion criteria such as known renal impairment, hereditary or acquired muscle disease, liver disease etc. Serum CK measures and statins were noted at the time of recruitment into the study. The study found two SNPs that passed the genome-wide significance threshold: a synonymous variant, rs2361797 in the *LILRB5* gene and a nonsynonymous variant rs11559024: Glu83Gly in

the *CKM* gene for an association with serum CK levels. They found heterozygous carriers of the rare allele at *CKM* had a mean CK level of 68.13 (SD 35.57) U/L compared to 119.32 (SD 84.74) for homozygous carriers of the common allele (119).

A replication study was undertaken using the Montreal Heart Institute Biobank, where the association was tested in 3389 statin users and 1941 non-users. Only Caucasian participants were selected for the study. Replication confirmed the findings for rs11559024 Glu83Gly in the *CKM* gene and rs12975366 Asp247Gly (BP 19:54255498) and rs2361797 (BP 19:54249685) in *LILRB5* with associations in statin users and non-users – showing that this genetic effect is not modulated by statin use (119). The results from the original study are presented in Figure 4.1.

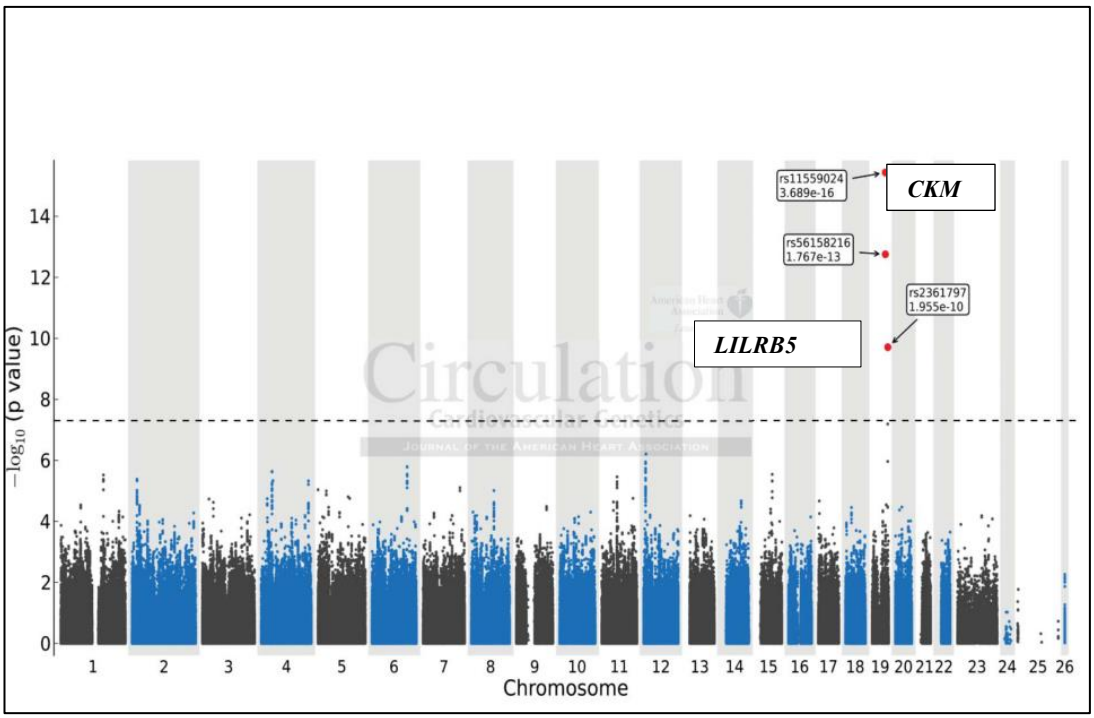


Figure 4.1 Results from GWAS performed by Dube *et al.* signals from chromosome 19 appeared strongly associated with serum CK levels. The variants were found in genes *CKM* and *LILRB5.*

The Canadian study described two SNPs in tight LD i.e. rs2361797 and rs12975366 as being associated with CK levels. The SNP in *LILRB5* primarily used by our study, rs12975366 (Asp247Gly) has a D' of .92 with rs2361797 (142) in the GoDARTS cohort

and similar association results were found with both SNPs in our cohort. These signals are therefore not independent and it is as yet unclear if either are causal variants.

In 2016, Kristjansson et al. replicated these findings in a GWAS performed on 63,159 Icelanders with CK measurements. They report the main effect of the Glu83Gly variant ($\beta = -0.446$, p value= 1.8×10^{-115}) and for the Asp247Gly variant ($\beta = -0.08$, p value= 6.5×10^{-44}) (118).

We attempt to test this association in the GoDARTS population. Furthermore, since the study has access to longitudinal health records we are able to assess the impact of these variants on intra-individual variability or inducibility of creatine kinase levels. Inducibility of CK in response to appropriate stressor is the clinically significant feature, as CK is usually measured when tissue damage is suspected, and is no longer part of a routine biochemistry panel. The findings of this analysis could impact on the viability of CK as a reliable biomarker, especially for statin intolerance.

4.2 Methods:

CK testing

Creatine kinase measures were gathered per the methods described in Chapter 3.

The CK measures used for replication in the JUPITER trial were tested at baseline, when the population was treatment-naïve.

The CK enzyme assays were performed at the Ninewells Hospital Central Biochemistry Laboratories and follow the protocols summarized subsequently.

Enzyme assay kits for CK produced by abcam (Cambridge, United Kingdom) are designed to detect the products of the reaction catalyzed by creatine phosphokinase. CK catalyzes the dephosphorylation of ATP to produce phosphocreatine and ADP. The generated phosphocreatine and ADP then react with the CK enzyme mix to form an intermediate, which reduces a colourless probe to a coloured product with strong

absorbance (λ = 450 nm). The assay is quite sensitive and can detect activity less than 1 mU. Enzyme assay kits produced by Sigma-Aldrich (St. Louis, United States of America) detect the reverse. In the reaction, phosphocreatine and ADP are converted to creatine and ATP. The generated ATP is used by hexokinase to phosphorylate glucose resulting in glucose-6-phosphate, which is oxidized by NADH in the presence of glucose-6-phosphate dehydrogenase to produce NADPH and 6-phospho-D-gluconate. The assay detects NADPH at 340 nm.

In both assays, one unit of CK is the amount of enzyme that will catalyze the conversion of 1.0 µmole of substrate per minute at pH 6.0 at 37°C. Both options are used widely, and rely on the enzyme's ability to act on substrate, measuring products on either side of the reversible reaction as seen in Figure 4.2.

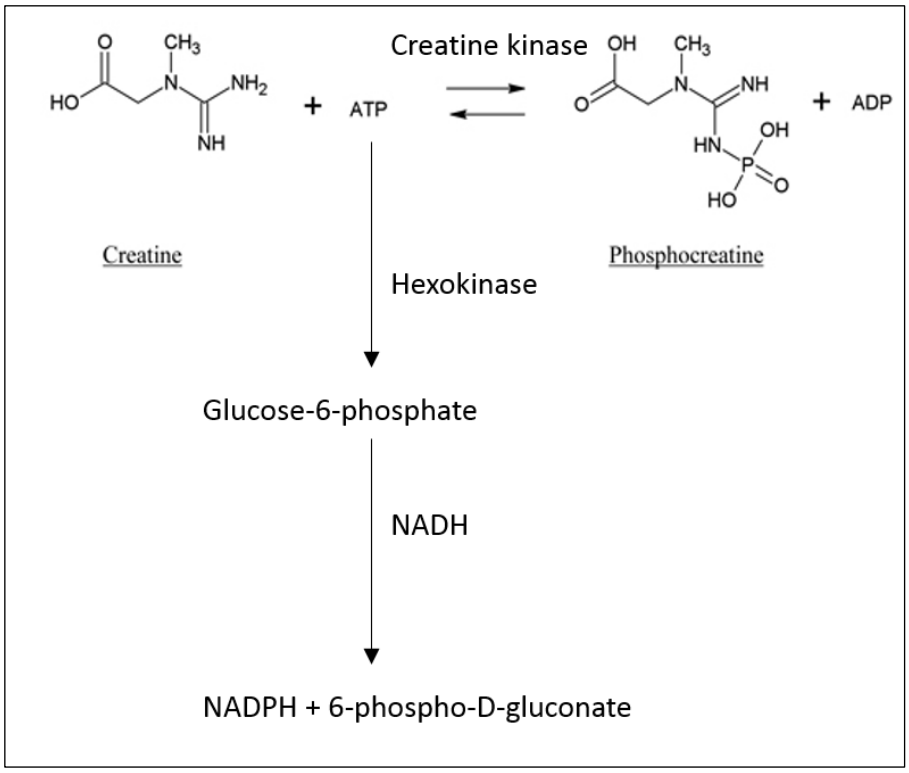


Figure 4.2 Reaction catalyzed by CK and end products detected by the two assay methods.

The end products of the assays helps us understand what aspect of CK activity is being measured and what an abnormal CK results might suggest about the altered functionality or structure of the molecule.

Genotype data

Genotype data for the *LILRB5* variant (rs12975366: Asp247Gly) was available for 5785 individuals. Out of those 2747 were imputed from the Illumina HumanOmni Express -12VI platform (Illumina, San Diego) and 2451 from the Affymetrix 6.0 platform (Affymetrix, Santa Clara) with an imputation quality of 86.8% and 81.7% respectively. An additional 587 were genotyped using TAQMAN. Imputation was performed against the 1000G Phase I V3 reference panel using Impute2 (126).

Genotype data for the *CKM* variant (rs11559024:Glu83Gly) was available for 6271 individuals, of whom 4578 were genotyped using the Human Exome -12 VI_A_chip and 1693 using TAQMAN.

Genotyping techniques for the GoDARTS cohort are described in Chapter 2.

Genotyping for 8749 JUPITER trial participants was performed on the Omni1-Quad platform (Illumina, San Diego). The imputation quality of the *LILRB5* Asp247Gly variant, rs12975366 was 94%, R^2=0.90 and MAF = 0.40. The *CKM* Glu83Gly variant, rs11559024 was directly typed with R^2=0.99 and MAF = 0.018 (115).

The variants were in Hardy-Weinberg equilibrium.

Statistical analysis

This analysis was not restricted to statin users. Logarithmic transformations were applied to all CK levels to normalize their distribution. Genotypes were treated as numerical variables, 0 represented homozygous carriers of the wild type (T/T) alleles, 1 represents heterozygotic carriers (T/C) and 2 represents homozygous carriers of the rare allele (C/C). For the association of "baseline" CK, each individual's first test result in an ambulatory setting was used. To ascertain the intra-individual variation in CK, the standard deviations were calculated (proc means noprint) procedure and specifying the by function to represent the participant ID. The associations were tested using linear regression (proc reg). The beta, standard error and R^2 are reported.

All analyses was performed in SAS 9.3 (SAS Institute, Cary NC). Plots were generated in R studio (135). Meta-analyses were performed using the metafor package in the R studio environment (135,136,143).

Statistical analyses in the JUPITER trial were performed by Dr. Dan Chasman using R (136).

4.3 Results: population distributions, MAF association and case report

Creatine kinase

The effect of the Glu83Gly variant in *CKM* was assessed using SWISSPROT(144–146) and the change in protein structure is presented in Figure 4.3. The thymine to cytosine switch, resulting in the substitution of glycine for glutamic acid occurs at the 83rd position in the 381 amino acid long enzyme. As seen in the protein model, the mutation does not fall in the active site of the enzyme.

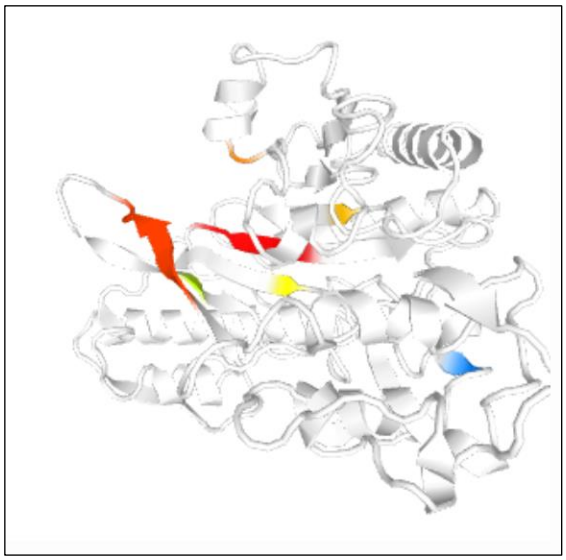


Figure 4.3 Tertiary structure of the enzyme CK. ATP-binding sites of the enzyme are highlighted in red, orange, amber, yellow and fluorescent green. Location of the Glu83Gly variant is highlighted in blue.

Mean CK levels were significantly different across genders, with males having significantly higher CK levels (β= 0.13, p value 2 x 10^{-16}). Gender accounts for 7% of the variation seen in CK levels (*see* Figure 4.4).

Although significantly associated with CK (p value = 0.015), age did not appear to be a significant predictor of CK levels; explaining lesser than 0.0004% of the variation, (*see* Figure 4.5).

In this analysis, 87% of individuals tested for the association with *CKM* Glu83Gly and 93% of those tested for *LILRB5* Asp247Gly were statin users.

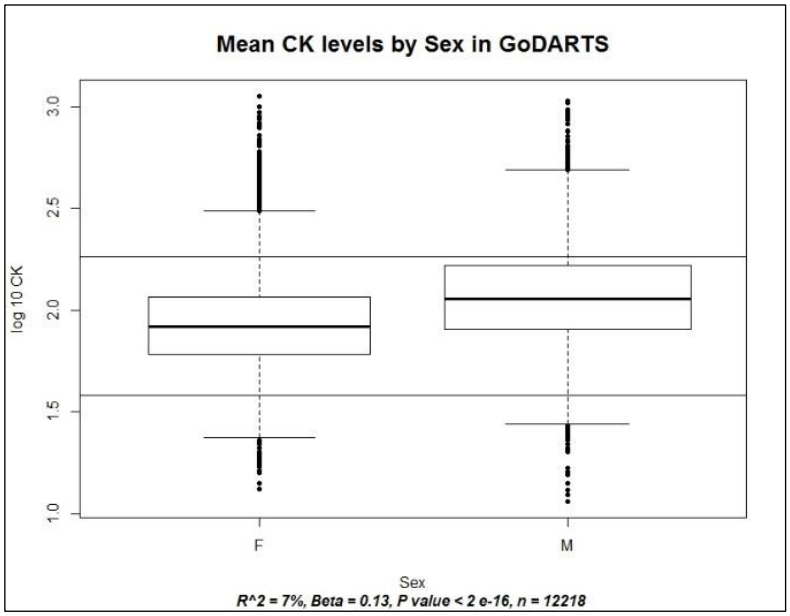


Figure 4.4 Boxplot of log 10 transformed CK levels stratified by sex in the GoDARTS study. The upper band corresponds to the upper limit of normal CK levels i.e. 180 IU/L, while the lower band corresponds to the lower limit of normal CK levels i.e. IU/L. 45% were female.

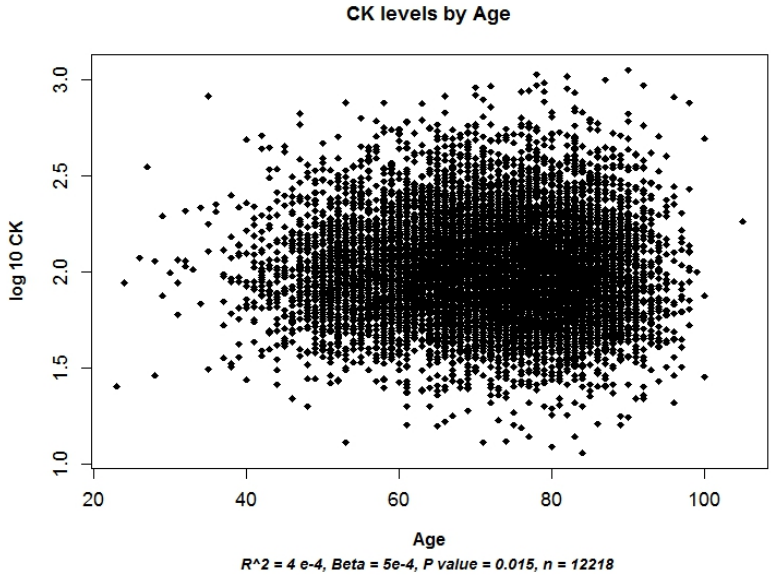


Figure 4.5 Scatterplot of the association between log 10 transformed CK levels and Age. Correlation: r = -0.022, p value = 0.015

Genotype data

Genotype information for the Glu83Gly variant in *CKM* was obtained for 6271 individuals. We observed 6064 were homozygous for the reference allele (T/T), 207 were heterozygous carriers of the mutation (T/C) and one individual was homozygous for the *CKM* variant (C/C), the MAF was 0.02.

Genotype information for the Asp247Gly variant in *LILRB5* was obtained for 5785 individuals in GoDARTS. We observed that 2295 were homozygous for the reference allele (T/T), 2715 were heterozygotes (T/C) and 790 (C/C) were homozygous carriers of the minor allele. The minor allele frequency MAF was 0.37.

In the JUPITER trial, genotyping for the trial participants was performed on the Omni1-Quad platform (Illumina, San Diego). The *CKM* variant (rs11559024) was directly typed, and had a MAF of 0.018. The imputation quality of *LILRB5* Asp247 variant, rs12975366 was 94% and had a MAF of 0.40 (115).

Association of the variants with CK levels

Creatine kinase levels were significantly associated with the *CKM* Glu83Gly variant (n = 4599, p value = 2 x10 $^{-16}$). Heterozygous carriers of the *CKM* 83Gly variant (T/C) had mean CK of 86 (+/-68) compared to 126 (+/-82) for homozygotes of the ancestral allele Table 4.1.

CK levels were also significantly associated with the *LILRB5* Asp247Gly variant (n = 5020, p value = 4 x 10 $^{-7}$). Homozygous carriers of the variant (C/C) had mean CK of 100 (+/-84) compared to 124 (+/-83) for homozygotes of the ancestral allele.

Box plots of log transformed CK levels stratified by the genotypes are provided in Figure 4.6 and Figure 4.7.

Table 4.1. Creatine kinase levels by genotype

SNP	Mean CK	SD CK	Median CK	Minimum CK	Maximum CK	N
***CKM:* rs11559024 (Glu83Gly)**						
T/T: Glu83	126	82	102	16	934	4447
T/C: 83Gly	86	68	61	14	420	152
***LILRB5:*rs12975366 (Asp247Gly)**						
T/T: Asp247	124	83	100	18	574	1972
T/C: 247Gly	119	82	96	12	931	2361
C/C: 247Gly	100	84	87	17	900	688

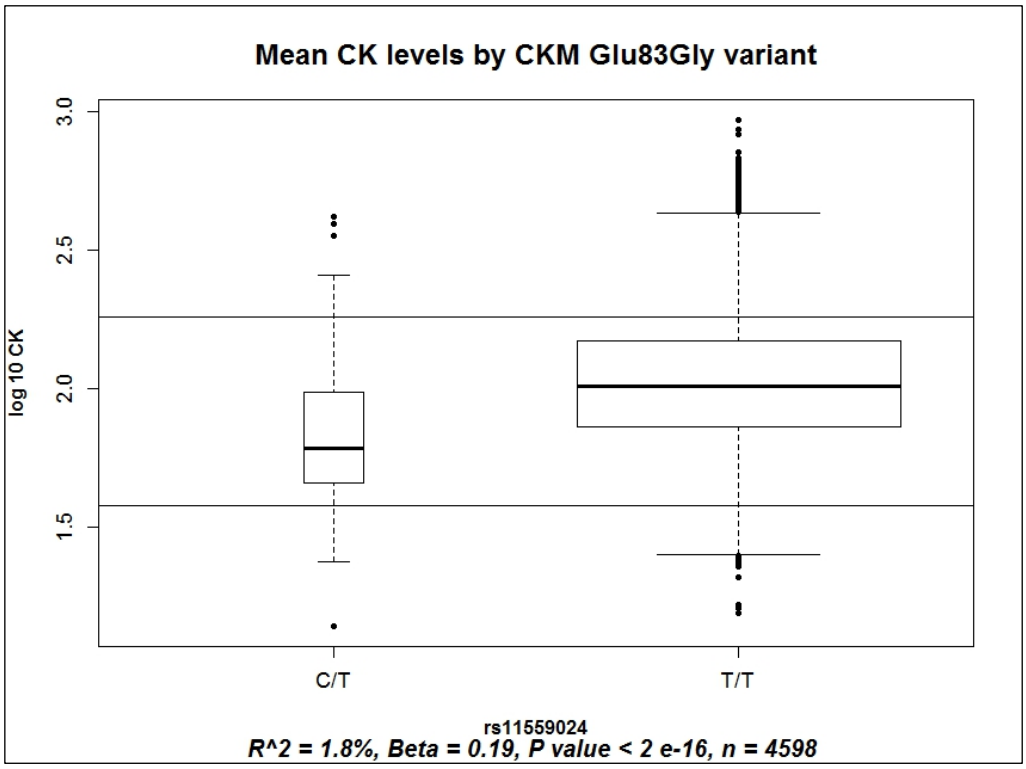


Figure 4.6 Boxplot of log 10 transformed CK levels stratified by *CKM* genotypes in the GoDARTS study. The upper band corresponds to the upper limit of normal CK levels i.e. 180 IU/L, while the lower band corresponds to the lower limit of normal CK levels i.e. 38 IU/L

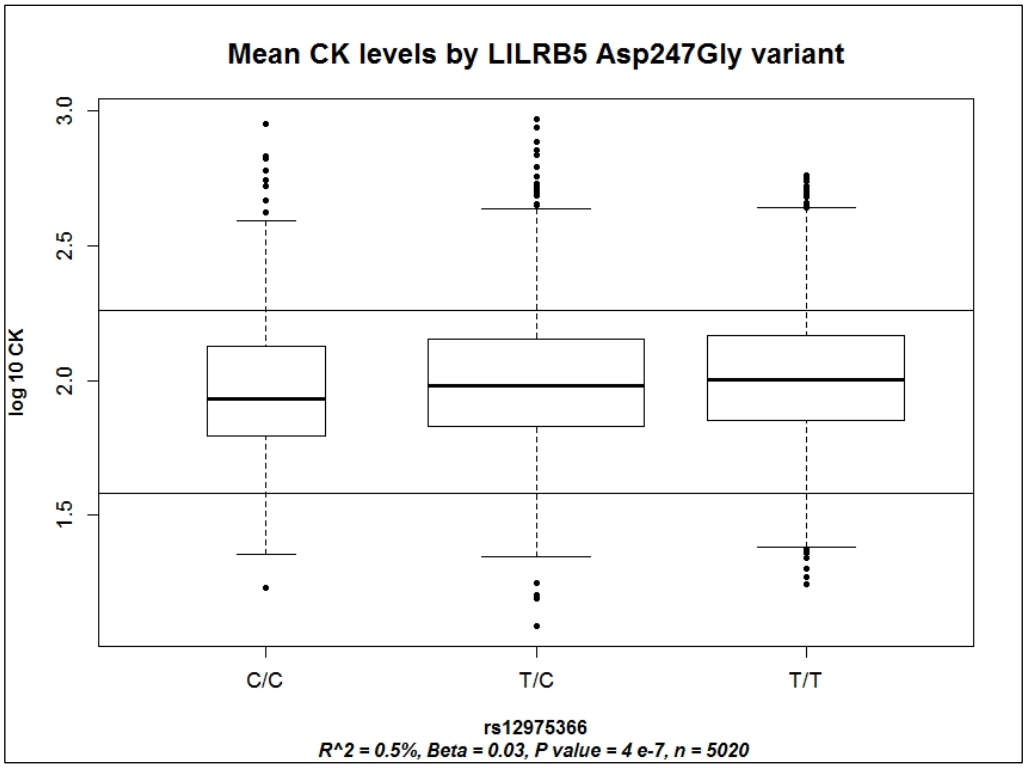


Figure 4.7 Boxplot of log 10 transformed CK levels stratified by *LILRB5* genotypes in the GoDARTS study. The upper band corresponds to the upper limit of normal CK levels i.e. 180 IU/L, while the lower band corresponds to the lower limit of normal CK levels i.e. 38 IU/L

The association with the *CKM* Glu83Gly variant was replicated in the JUPITER trial (n = 8745, p value <2 x 10^{-16}). A meta-analysis with the GoDARTS cohort showed a highly robust association (β = -0.18, p value = 1 x 10^{-63}). A forest plot of the association is presented in Figure 4.8.

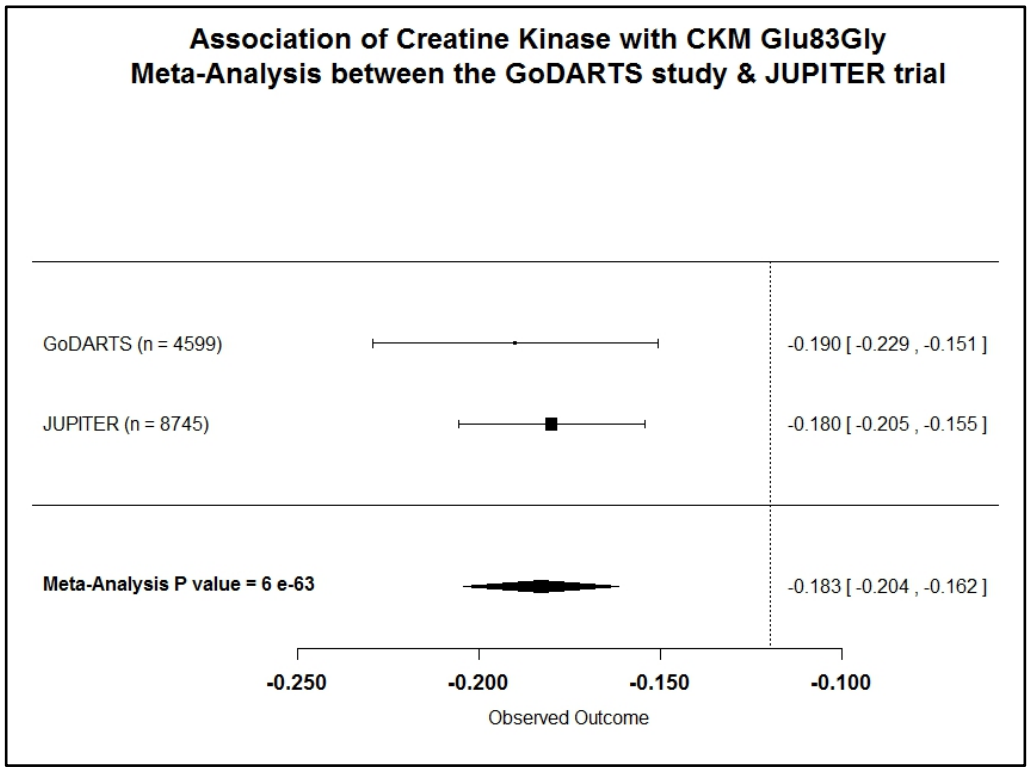


Figure 4.8 Forest plot of the meta-analysis of the association of creatine kinase with the *CKM* Glu83Gly variant in the GoDARTS study (p value = 1 x 10^{-16}) and the JUPITER trial (p value = 1 x 10^{-16}).

The association with the *LILRB5* Asp247Gly variant was replicated in the JUPITER trial (n = 8745, p value = 2.8 x 10^{-16}). A meta-analysis of the effect of the *LILRB5* Asp247Gly variant across the two studies showed a robust association (beta = -0.029, p value 3.5 x 10^{-17}). A forest plot of the association is presented in Figure 4.9.

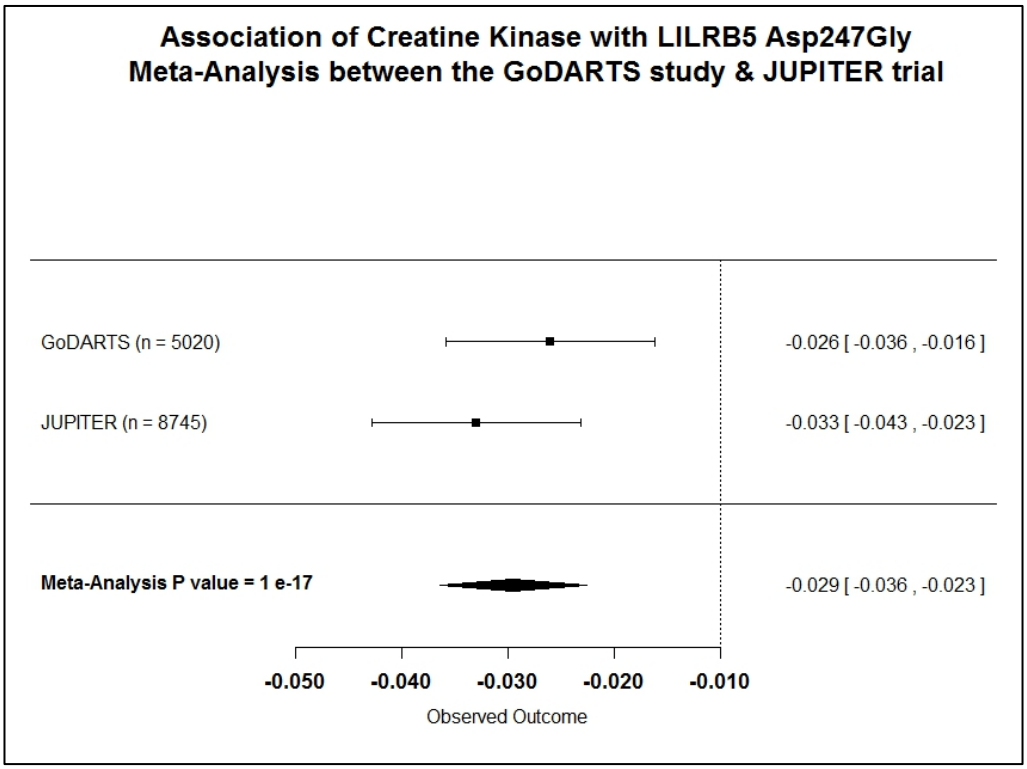


Figure 4.9. Forest plot of the meta-analysis of the association of creatine kinase with the LILRB5 Asp247Gly variant in the GoDARTS study (p value = 1 x 10^{-7}) and the JUPITER trial (p value = 1 x 10^{-16}).

We conclude that "resting or un-induced" CK levels are strongly associated with the variants in *CKM* and *LILRB5*. From the meta-analyses of the GoDARTS and JUPITER populations we conclude that heterozygous carriers of the Glu83Gly variant in *CKM* there was a 0.18 log decrease, equivalent to an average decrease of 1.20 IU/L serum CK levels compared to non-carriers. Meanwhile, the addition of each rare allele for the Asp247Gly variant in *LILRB5* results in a 0.03 log decrease in enzyme levels, which is equivalent to an average reduction of 1.03 IU/L of serum CK per C allele.

Association of the variants with CK variability

With access to an average of 13 years of follow-up and a median of 9 measures of CK per individual in GoDARTS we decided to examine the impact of these variants on the intra-individual variation in CK levels. We restricted this analysis to individuals with three or more measures of CK. Per individual standard deviations were calculated.

The standard deviation of an individual's CK test results were stratified by the genotype to create Figure 4.10 and Figure 4.11. We observe that *CKM* Glu83Gly variant exerts a strong effect on the inducibility of CK measures in an individual (beta = -0.24, p value 2 x 10-5). There is a 1.27 IU/L reduction in the standard deviation of their intra-individual serum CK levels. However, the *LILRB5* Asp247Gly variant does not have a similar effect (p value = 0.8), this effect was also absent when the genotype was treated as a dominant trait Figure 4.12.

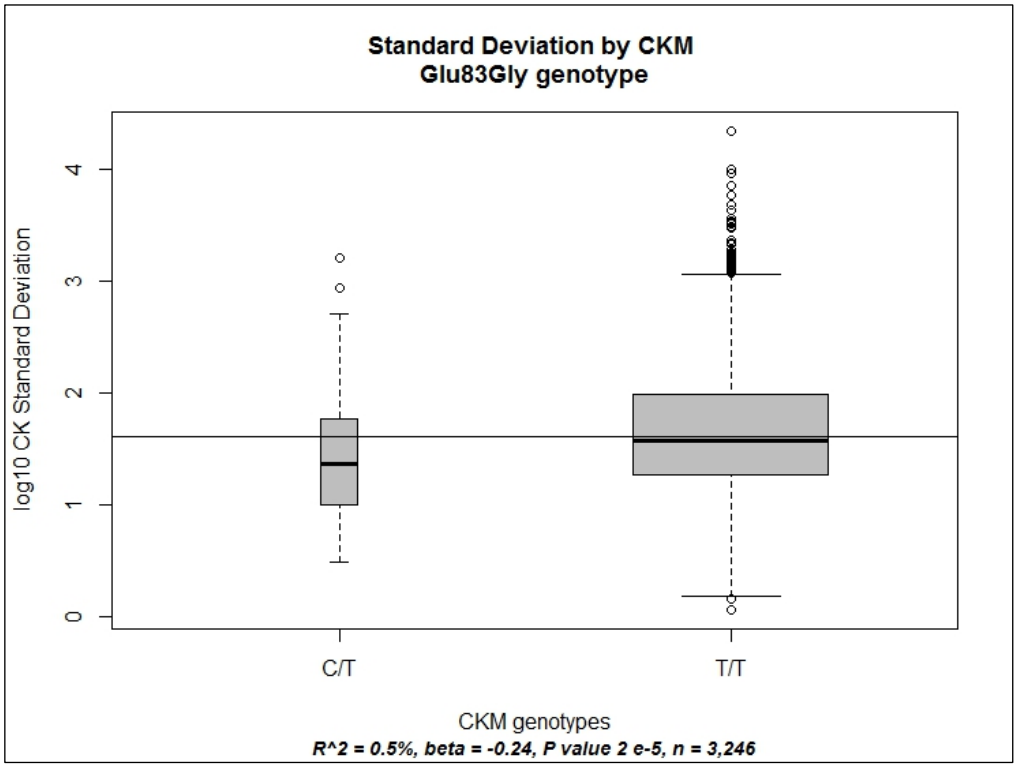


Figure 4.10 Boxplot showing intra-individual variability demonstrated as standard deviation by *CKM* Glu83Gly genotype. The reference line indicates the mean standard deviation in the population (40 IU/L).

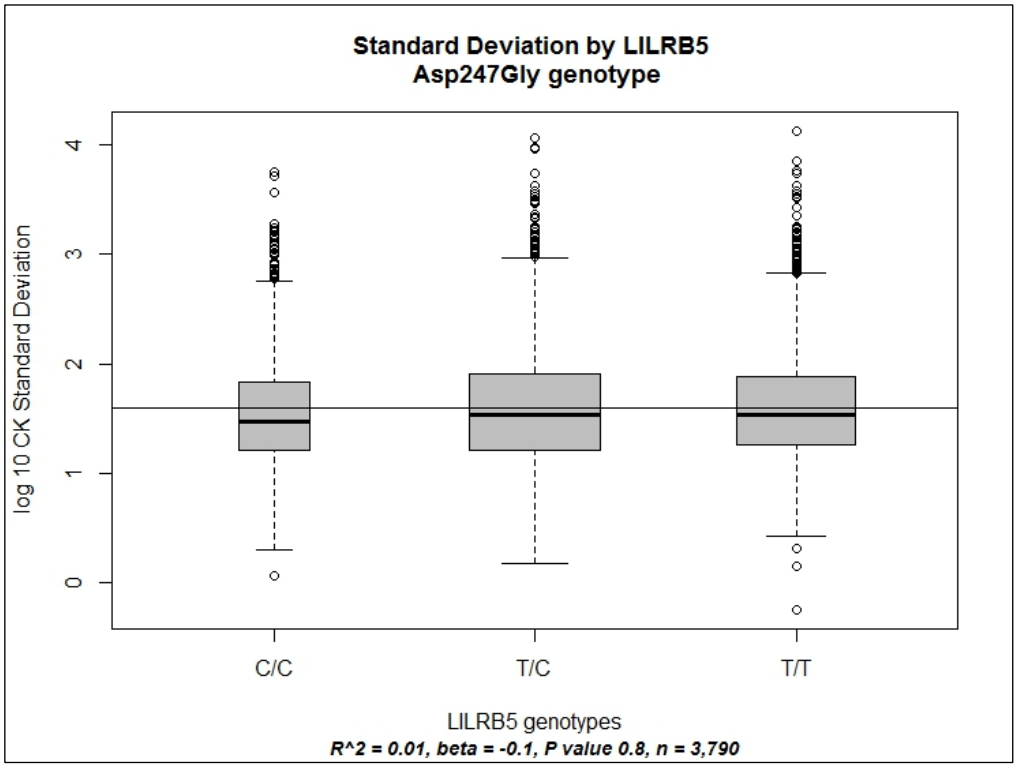


Figure 4.11 Boxplot showing intra-individual variability demonstrated as standard deviation by *LILRB5* Asp247Gly genotype. The reference line indicates the mean standard deviation in the population (40 IU/L).

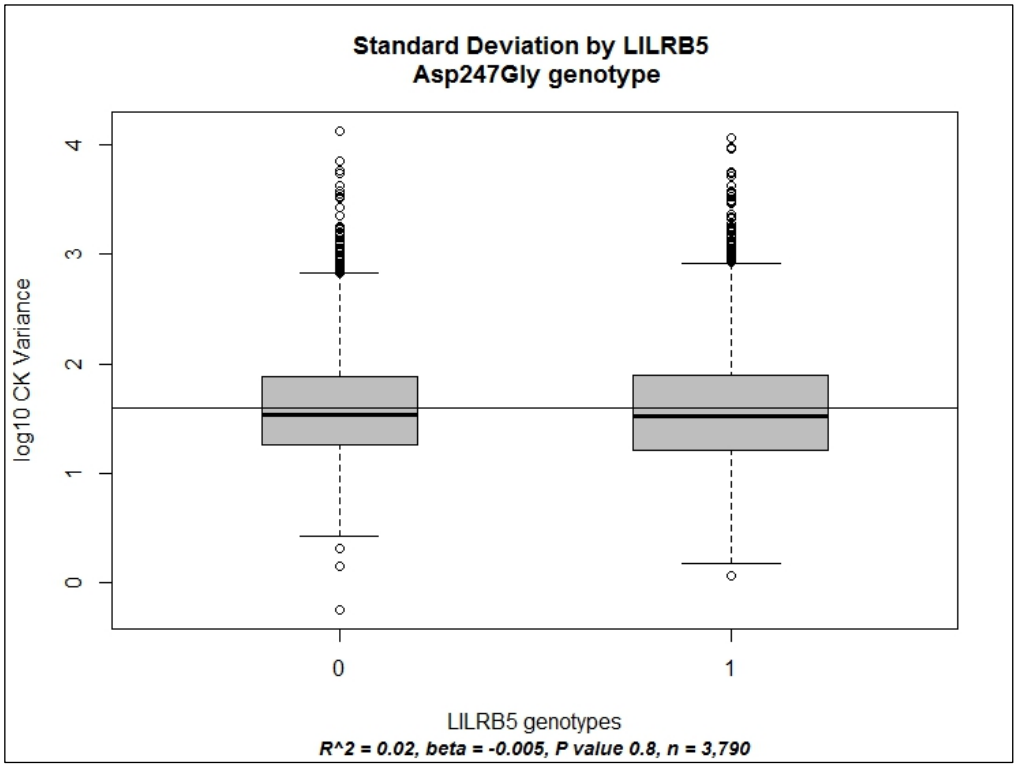


Figure 4.12 Boxplot showing the intra-individual variability in CK levels stratified by *LILRB5* Asp247Gly genotype treated as a dominant trait. C/C and C/T = 1 and T/T = 0

We observe that when considering intra-individual variability the *LILRB5* Asp247Gly variant shows no significant effect.

In the absence of a clear additive effect for the *LILRB5* variant on CK levels, subsequent analyses in this thesis employed a dominant model; where individuals homozygous for the ancestral allele (T/T: Asp247) were compared to carriers of the minor allele (T/C and C/C: 247Gly).

The impact of the Glu83Gly variant in *CKM* on CK inducibility is further highlighted in the single individual in the GoDARTS cohort who was homozygous for the variant (C/C). In Figure 4.13, we see the patient's CK levels during hospitalization for necrotizing fasciitis, a condition during which there is aggressive infection of the tissue and where CK levels could rise to > 600IU/L (147), while the patient's CK levels did not exceed 15 IU/L/. In response to the subsequent development of gangrene the patient underwent a debridement procedure, post-operatively CK levels were at a maximum of 28 IU/L. Later, the patient underwent a hemicolectomy for bowel cancer. The patient's pre and post-operative CK levels remained relatively unchanged and in fact, seemed lower (34 IU/L

and 25 IU/L respectively). The findings are presented by Wallace *et al* as a case report. Notably, the patient was a statin user who had undergone 4 switches in statin therapy between 3 types of statins, namely pravastatin, atorvastatin and simvastatin. This switching is attributed to complaints of intolerance however her CK measurements had been deemed normal by clinicians looking for evidence of statin-induced myositis (148).

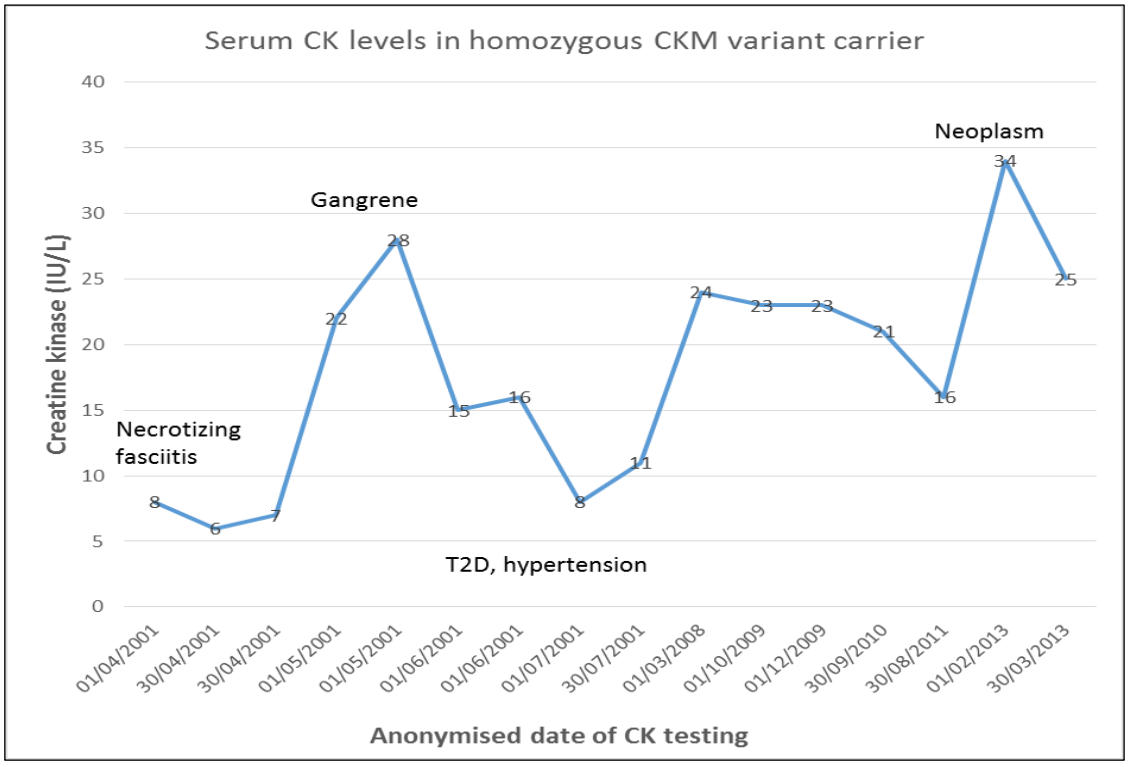


Figure 4.13 Serum CK levels of homozygous CKM Glu83Gly carrier

4.4 Discussion

We report the replication of the association between two variants, one in *CKM* and the other in *LILRB5* with serum CK levels, first reported by Dubé *et al.* using data obtained in a prospective cohort study (GoDARTS) and in the treatment-naïve population of a clinical trial (JUPITER). Further, we show that the variant in *CKM*, Glu83Gly is also associated with the inducibility of CK levels. The use of standard deviations to calculate the inducibility gives us a conservative estimate of the effect, as the impact of the variant could be assessed by contrasting the highest and lowest CK recorded for an individual. In the case of the homozygous carrier of the variant we see the lack of CK response in

conditions that would normally cause extremely high CK levels, such as severe tissue infection and surgical trauma (148). We conclude that carriers of this *CKM* Glu83Gly variant are less likely to produce large quantities of measureable, and therefore, functioning CK in response to tissue damage.

Since the variant inhibits measureable CK levels from rising, potentially obfuscating the correct diagnosis, it might be essential to factor in the genotype of the individual before determining the validity or normalcy of the result.

The MAF for the *CKM* Glu83Gly variant in the GoDARTS population was 0.02, in JUPITER it was 0.018, similar to the European population group in the 1000 Genome project (0.022). Similarly the MAF for the *LILRB5* Asp247Gly variant was 0.37 and 0.40 in the GoDARTS and JUPITER populations respectively, comparable to the 1000 Genome project MAF of 0.43 amongst Caucasians. This indicates that the populations under study were not suffering from participant selection bias. The frequency of the *CKM* variant in the Kenyan Masai population is 0.223 (149). This significant difference in the allele frequencies in the Masai is striking and warrants further investigation in other populations and for association with features of muscle or athletic performance.

The diminished quantity of functional of CK in the serum, would have to be compensated by other mechanisms in the body to maintain energy homeostasis. One potential hypothesis revolves around switch between aerobic and anaerobic metabolism. An oxygen debt created when there is insufficient phosphocreatine to make enough ATP needed during periods of high exertion. The process of energy generation then shifts from muscles to the liver (where ATP is generated anaerobically via glycolysis and the Cori cycle, which delays the oxygen consumption process). Therefore, those with low CK activity (*CKM* carriers) might make this switch sooner, to maintain homeostasis.

All definitions of statin intolerance hinge on the elevation of CK and, in some instances on the resolution of muscle-associated CK elevations upon the discontinuation of statin

therapy. The most established threshold for a "clinically relevant" CK elevation is 4 or more times the upper limit of normal (4 X ULN) and 10 X ULN (66). However, as this study suggests, these outcomes are less likely to occur amongst carriers of the *CKM* variants.

We therefore posit that the employment of the traditional classifications would lead to the artificial enrichment of non-carriers of the *CKM* variant being classified as having statin-induced myalgia, myositis or rhabdomyolysis. It would be interesting to examine if the *CKM* Glu83Gly variant is associated with reduced reports of myalgia or muscle-aches, as it is associated with lower CK. If the variant confers no protective effect, this might provide a novel mechanistic rationale for a sub population of individuals presenting with statin intolerance or myalgia without raised CK levels (97).

While the *LILRB5* variant shows a significant trend toward having lower CK levels, carriers do still seem to retain the ability to induce high CK levels. This finding, along with the reported association of this variant with serum lactose dehydrogenase (LDH) levels (118), potentially implicates it as a baseline muscle-susceptibility marker, with carriers being less prone to muscle breakdown or being better able to repair muscle damage.

5 Are the missense variants in *CKM* and *LILRB5* associated with statin intolerance?

5.1 Introduction

In Chapter 4, we provide evidence that un-induced or baseline CK levels, and the intra-individual variation in CK levels is confounded by the Glu83Gly variant. We provide evidence to suggest that CK levels must be considered in light of the genotypes of individuals at Glu83Gly and Asp247Gly before their validity can be determined. These conclusions point to a potential genetic underpinning for the phenomenon of muscle-based symptoms that are not reflected in commensurate increases in serum enzyme levels of CK (97). Therefore, in Chapter 3, we created phenotypes of statin intolerance that accounted for this confounding effect, by virtue of their being independent of CK elevations, but strongly reflecting patterns of statin consumption that indicate intolerance. These phenotypes also showed increased risk of failure of statin therapy.

Since carriers of the *CKM* 83Gly variant do not produce as much functional CK, we investigated the data to assess if this variant is also associated with lower reported instances of muscle aches and non-compliance (in the JUPITER trial), with the phenotypes of intolerance created in Chapter 3, and with clinically adjudicated cases of SIM. It would be clinically interesting to known if the *CKM* Glu83Gly variant is associated with reduced reports of myalgia or muscle-aches. If the variant confers no protective effect, this might provide a novel mechanistic rationale for a sub population of individuals presenting with statin intolerance or myalgia without raised CK levels (97).

The finding by Kristjansson *et al.* of the association of variants in *LILRB5* including Asp247Gly with serum LDH levels suggests the variant might impart a statin

independent susceptibility to muscle-based events, as LDH is often used in conjunction with CK as a marker of tissue damage. The variant shows the same direction of effect, i.e. carriers of 247Gly appear to have lower LDH levels. This makes it a viable marker for susceptibility to the commonly noted muscle-based symptoms that are attributed to statin intolerance.

Therefore, we tested the association between the phenotypes created *a priori* with the *LILRB5* Asp247Gly variant to check if the variant predisposed to statin intolerance. We hypothesize that carriers who appear to have lower muscle enzyme activity levels (CK and LDH) will also be protected from forms of statin intolerance that are independent of CK levels.

We then attempted to replicate the findings presented here in a clinical trial setting that might be less prone to the confounding factors associated with a cohort study such as GoDARTS. The main outcomes from the JUPITER trial being examined in this chapter are compliance and the development of myalgia. Widely understood to play a role in CVD outcomes for statin users, compliance to therapy is often impacted when an individual experiences an adverse reaction to a medication (102), whether or not that is reflected in commensurate increases in biomarkers.

Finally, we sought to replicate our findings using clinically adjudicated cases of statin-induced myositis (SIMs) from the PREDICTION-ADR consortium. The aims and work packages for the PREDICTION-ADR project are described in Chapter 2.

5.2 Methods

Methods and development of phenotypes covariables and genotype data used in the analyses are detailed in Chapter 3. The JUPITER cohort was introduced in Chapter 2, and Chapters 3 and 4 describe how the genetic data and patient-level data were collected. The outcomes studied in JUPITER are contextually re-introduced here. The novel cohort introduced in this chapter is the PREDICTION-ADR cases of SIM and matched controls.

JUPITER trial

The replication effort in JUPITER focused on 4381 study participants randomized to receive statin treatment and 4368 who received the placebo. The JUPITER trial ascertained compliance on the basis of a questionnaire, pill counts and non-trial statin use (114,115). Myalgia was ascertained by physicians who were blinded to treatment (132).

Since the trial focused on the role of low-grade underlying inflammation (evidenced by high C-reactive protein levels), patients with inflammatory conditions such as severe rheumatoid arthritis, lupus, or inflammatory bowel disease were excluded, as were patients using immunosuppressant agents (114).

Genotyping for 8749 JUPITER trial participants was described in Chapter 4 (115).

All analyses for JUPITER trial data were performed, upon request, by Dr. Daniel I Chasman.

Statin-induced myositis (SIM) in PREDICTION-ADR

Cases and controls for SIM were contributed by the consortium's study centers in Dundee, Liverpool, Uppsala and Utrecht.

5.2.1.1 Definition and adjudication of cases and controls

Cases of statin-induced myopathy were classified according to the criteria cited by Alfirevic *et al.* (13); of CK raised ≥ 4 X ULN when using population-based databases such as GoDARTS or CPRD (Clinical Practice Research Datalink) where identification

is dependent on electronic medical records (14,15). The CK elevations must not be attributable to any other clinical circumstances that would result in such drastic increases.

Individuals so identified were then subjected to clinical adjudication by a panel of physicians and specialists. Factors considered for adjudication were the resolution of CK after de-challenge, post-event prescribing changes (such as switching or total discontinuation), patient's medical history of kidney disease, trauma, falls, MI, thyroid disease and tests for HMGCR antibodies, muscle biopsy and physical activity, if available.

Additional cases were identified from CVD clinics, General Practitioners (GP) practices, and clinics specialized for muscle diseases where the individuals were adjudicated directly by the physician as having statin-induced myopathy. A total of 249 adjudicated cases of statin-induced myositis were age, gender and starting statin therapy matched to 246 adjudicated controls. Statin tolerant controls had been on therapy for a minimum of 1 year and had shown no CK elevations, while maintaining regular therapy without any modifications.

5.2.1.2 Sequencing for SIM cases and controls

Sequencing was performed by the PREDICTION-ADR consortium (refer Chapter 1 and 2). Sequencing of 495 statin samples (249 cases and 246 controls) was performed using exome-enriched sequence data. SureSelect QXT reagents (Agilent Technologies, Wokingham, UK) were used to perform fragmentation, end-repair, A-addition and adaptor ligation reactions to generate Illumina-compatible sequencing libraries. Hybridization capture enrichment of whole genome libraries was performed using the SureSelect v5 all-exon probe set, following manufacturer's recommendations throughout (http://www.agilent.com/cs/library/usermanuals/Public/G9681-90000.pdf). Equimolar aliquots of 12 post-enrichment libraries (6 cases and their 6 matching controls) were pooled before sequencing using version 2 TruSeq chemistry on a

Nextseq500 (Illumina Inc., San Diego, CA, USA). Paired-end 150-bp sequence reads were then analyzed through the Basespace app BWA enrichment v2.1.0 (Illumina) which gathers demultiplexing, alignment, duplicate removal, variant discovery and annotation. The core algorithm in this workflow is the alignment to indexed reference genome (hg19) using BWA and indel realignment, base quality score recalibration as well as variant discovery were performed using Genome Analysis Toolkit (GATK) v.1.6 UnifiedGenotyper.

For the PREDICTION-ADR cohort, selection of cases and controls was carried out by the candidate for samples contributed by the University of Dundee. Selection for the Liverpool and Uppsala cohorts was undertaken by Dr. Ana Alfirevic and Dr. Mia Wadelius respectively. Exome sequencing, post-sequencing processing were performed by Dr. Cyrielle Maroteau who extracted genotypes of interested upon request.

Statistical analysis

All statistical analyses on GoDARTS and PREDICTION-ADR data were performed in SAS 9.3 (SAS Institute, Cary, North Carolina). Association testing between the variants and each phenotype of intolerance as well as SIM were done using a binary logistic model using the "proc logistic" command.

Covariables known to be associated with intolerance such as gender, age, co-medication usage, type of statin, dose of statin were added to each model, as discussed in Chapter 3. A backwards step-wise approach was used to eliminate covariables that were not significant predictors in each model.

Statistical analyses in the JUPITER trial were performed by Dr. Dan Chasman upon request using R (136). A binary logistic regression was performed to test the association between compliance and the Asp247Gly variant using the "glm" function with the "family = 'binomial'" specification. Survival analyses were performed using the

"coxph" function. Stratified survival analyses were performed using the "strata" function in the relevant command.

5.3 Results of association tests with both variants in GoDARTS

Distribution of *CKM* Glu83Gly across phenotypes

The distribution of Glu83 genotypes across each phenotype is presented in Table 5.1. Due to the infrequency of occurrence of the rare allele, the vast majority of each phenotype is populated by non-carriers. However there doesn't seem be any notable difference is frequency of T/T homozygotes (Glu83) between cases and controls, this will be formally tested in the next section.

Table 5.1 Distribution of *CKM* Glu83Gly genotypes across phenotype groups

Case/ Control	Phenotypes	*CKM*		
		Glu83	83Gly	Total
Case	A: Raised CK	1064 (99%)	16	1080
Case	B: Raised CK + Non-adherent	288 (98%)	5	293
Case	C: Non-adherent	318 (96%)	12	330
Case	D: Statin-induced myositis	20 (91%)	2	22
Case	G: Dose Intolerant	286 (95%)	14	300
Control (A,B,D)	E: Normal CK+ Adherent	419 (97%)	15	434
Control (C)	F: Adherent	671 (97%)	19	690
Control (G)	H: Dose tolerant	269 (97%)	8	277

Distribution of *LILRB5* Asp247Gly across phenotypes

The distribution of Asp247Gly genotypes across each phenotype group is presented in Table 5.2. Of note is the difference in Asp247 (T/T) allele frequency between the groups classified as cases and those are controls. Cases appear to have higher percentage of individuals encoding Asp247.

Table 5.2 Distribution of *LILRB5* Asp247Gly genotypes across phenotype groups

Case/ Control	Phenotypes	*LILRB5*			
		Asp247	247Gly	247Gly	Total
Case	A: Raised CK	279 (46%)	263	63	605
Case	B: Raised CK + Non-adherent	176 (47%)	160	37	373
Case	C: Non-adherent	198 (43%)	195	64	457
Case	D: Statin-induced myositis	11 (39%)	12	5	28
Case	G: Dose Intolerant	270 (42%)	272	98	640
Control (A,B,D)	E: Normal CK+ Adherent	198 (35%)	281	79	558
Control (C)	F: Adherent	237 (36%)	328	89	654
Control (G)	H: Dose tolerant	177 (35%)	256	73	506

The number of samples available for analyses differs across the two genotypes as the *CKM* variant was genotyped on the exome chip and through TAQMAN, while the *LILRB5* variant is imputed from the Illumina and Affymetrix platforms as well as directly typed using TAQMAN.

Next, the odds ratios for the development of each definition of intolerance by genotypes are presented in Table 5.3 and Table 5.4 and discussed subsequently.

Table 5.3 Association of phenotypes of statin intolerance with Glu83Gly

Phenotype Groups	**Phenotype Description**	***CKM* Glu83Gly**			
		Unadjusted model		Adjusted model	
		83Gly v. Glu83: OR (95% CI)	**P value**	**83Gly v. Glu83: OR (95% CI)**	**P value**
A v. E	Raised CK	0.42 (0.21, 0.87)	0.017	1.61 (0.66, 3.92)	0.55
B v. E	Raised CK + Non-adherence	0.44 (0.16, 1.18)	0.10	-	
C v. F	Non-adherence	1.33 (0.64, 2.78)	0.44	-	
D v. E	Statin-induced myositis	2.52 (0.55, 11.55)	0.24	-	
G v. F	Dose intolerance	1.64 (0.68 3.96)	0.28	-	

All models adjusted as stated in the methods section, with variables found to be significantly associated with each phenotype in Chapter 3. Subsequently a backwards step-wise process was used to create a model with only significant predictors.

Table 5.4 Association of phenotypes of statin intolerance with Asp247Gly. Here the odds of intolerance for those with the ancestral allele are being contrasted to carriers

Phenotype Groups	**Phenotype Description**	***LILRB5* Asp247Gly**			
		Unadjusted model		Adjusted model	
		Asp247 v. 247Gly: OR (95% CI)	**P value**	**Asp247 v. 247Gly: OR (95% CI)**	**P value**
A v. E	Raised CK	1.36 (1.10, 1.70)	3×10^{-3}	1.82 (1.25, 2.63)	1×10^{-3}
B v. E	Raised CK + Non-adherence	1.62 (1.24, 2.12)	4×10^{-4}	1.96 (1.25, 3.07)	3×10^{-3}
C v. F	Non-adherence	1.35 (1.06, 1.73)	0.015	2.07 (1.17, 3.67)	1×10^{-2}
D v. E	Statin-induced myositis	1.18 (0.54, 2.57)	0.68	1.92 (0.61, 6.11)	0.27
G v. F	Dose intolerance	1.36 (1.07, 1.73)	0.013	2.00 (1.27, 3.15)	2×10^{-3}

All models adjusted as stated in the methods section, with variables found to be significantly associated with each phenotype in Chapter 3. Subsequently a backwards step-wise process was used to create a model with only significant predictors.

Association of raised CK while on statins (A v. E)

Consistent with our prior knowledge of the modulation of CK levels by the *CKM* variant, individuals with the reduced CK; 83Gly variant had much lower odds of having raised CK while on statin therapy, compared to non-carriers (OR = 0.42, p=0.017). However, this protective effect was no longer significant in a model adjusted for covariables such as starting and ending statin therapies, starting dose and the use of interacting co-medications as seen in Table 5.3.

Also, as expected, based on the association of the *LILRB5* Asp247 variant with CK levels, we find this phenotype is associated with the variant. Individuals homozygous for Asp247 have 1.82 times the odds of having CK raised above the normal while on statin therapy compared to carriers of the variant, in a model adjusted for the starting and ending statin therapies, dose at which therapy was started, the use of interacting co-medications and age they were first prescribed statin as seen in Table 5.4

Association with raised CK and non-adherence while on statins (B v. E)

The *CKM* Gly83 variant was not significantly associated with having raised CK and being non-adherent to statin therapy, although the point estimate still demonstrates a marked protection (OR =0.4; 95% CI: 0.16 1.18) as seen in Table 5.3. This could be due to the smaller sample size and, a power analysis reveals that given the MAF of the variant, we would need 505 cases matched at least 1:1 with controls to have 80% power to detect an effect. At present with a power of 67% the analysis is slightly underpowered (150).

When considering the *LILRB5* 247Gly variant, in spite of a smaller sample size compared to the previous analyses, this genotype shows a significant association with this definition of intolerance. Those homozygous for Asp247 had 1.96 times the odds of having raised CK and being non-adherent to therapy compared to carriers of the variant, in a model adjusted for the first statin on therapy and it's dose, the age and sex

of the individual and their concurrent use of interacting medications as seen in Table 5.4. This analysis has over 99% power to detect an effect of this magnitude (150).

Association with non-adherence, irrespective of CK (C v. F)

The *CKM* 83Gly variant was not associated with non-adherence to statin therapy as seen in Table 5.3. However, a power analysis reveals that given the MAF of the variant, the power to detect an odds ratio between 1.1 and 1.5 for this analysis was under 60% (150).

For the *LILRB5* Asp247 genotype, individuals who were homozygous carriers of Asp247 had 2.07 times the odds of being non-adherent to their treatment compared to carriers of the variant, in a model adjusted for the first statin on treatment, its dose, the last statin prescribed and their concurrent use of interacting medications as seen in Table 5.4. This analysis had over 99% power to detect an effect of this magnitude (150).

Association with statin-induced myositis (D v. E)

Given the MAF of the *CKM* variant it was impossible to draw any conclusions from the data.

For the *LILRB5* Asp247 variant, individuals homozygous for Asp247 showed 1.92 times the odds of developing statin-induced myositis, in a model adjusted for the concurrent use of interacting co-medications, however, this is merely a trend and the finding was not statistically significant, seen in Table 5.4 . The analysis had 37% power to detect an effect of the magnitude.

Association with dose-related intolerance (G v. H)

The *CKM* 83Gly variant was not associated with dose-related intolerance to statins, as seen in Table 5.3. However, a power analysis reveals that given the MAF of the variant, at present the power to detect an odds ratio between 1.1 and 1.5 for this analysis was under 60% (150).

For the *LILRB5* Asp247 genotype individuals homozygous for Asp247 had 2 times the odds of being intolerant to the lowest dose of a statin compared to carriers of the variant, in a model adjusted for the first and last statin on treatment and the concurrent use of interacting medications, as seen in Table 5.4. This analysis had over 99% power to detect an effect of this magnitude (150).

5.4 Results of replication studies

Replication in the JUPITER trial

The two outcomes studied in the JUPITER trial population were compliance to therapy and the incidence of myalgia.

5.4.1.1 Association of variants with compliance in JUPITER

In the JUPITER trial 1435 participants were found to be non-compliant with their therapy, while 7046 were compliant. There was no association between the Glu83Gly variant and non-compliance (OR = 1.04, SE = 1.37, p value = 0.80).

However, individuals homozygous for Asp247 had 1.15 times the odds of being non-compliant (95% CI 1.03, 1.27) p value = 0.02 compared with carriers of 247Gly.

5.4.1.2 Association of variants with myalgia in JUPITER

Myalgia was observed in 837 trial participants and showed no association with their assigned therapy (132).

Table 5.5 shows the incidence of myalgia by the *CKM* Glu83Gly genotype. The Chi-square statistic for this association is 1.2 and the p value 0.27, indicating there was no statistically significant overall association between 83Gly and incident myalgia in the trial.

Table 5.5 Incidence of myalgia by Glu83Gly genotype

Myalgia	Glu83 (T/T)	83Gly (T/C)	Total
No	7627	285	7912
Yes	813	24	837
Total	8440	309	8749

Table 5.6 shows the incidence of myalgia by *LILRB5* Asp247Gly genotype. The Chi-square statistic for this association was 1.6 and the p value 0.2, indicating there was no statistically significant association between Asp247 and incident myalgia in JUPITER.

Table 5.6 Incidence of myalgia by Asp247Gly genotype

Myalgia	Asp247 (T/T)	247Gly (T/C or C/C)	Total
No	2674	5238	7912
Yes	301	536	837
Total	2975	5774	8749

5.4.1.3 Association with myalgia stratified by treatment arms (statin v. placebo)

Next we examined the association of the variants with the development of myalgia stratified by the two arms of the trial. For the *CKM* Glu83Gly variant, as seen in Table 5.7, there was no association in either arm of the trial, nor was myalgia significantly associated with the interaction between the treatment allocation and genotype.

Table 5.7 Association of Glu83Gly with myalgia stratified by trial allocation arms (modelled for carriage of the variant)

Trial arm	Myalgia events	OR (95% CI)	P value	Interaction OR (95% CI)	P value
Rosuvastatin	439	0.64 (0.35,1.19)	0.16	0.65(0.29,1.47)	0.31
Placebo	398	0.96 (0.93,1.75)	0.91		

The interaction term is between trial arm and genotype

For the *LILRB5* Asp247Gly variant, as seen in Table 5.8, we observed a significant effect only in the placebo arm of the study, with individuals homozygous for the ancestral (T/T) allele having 1.27 times the odds of developing myalgia compared to carriers of the 247Gly variant. The interaction between statin allocation and the Asp247 genotype appeared to trend toward significant (p = 0.085).

Table 5.8 Association of Asp247Gly with myalgia stratified by trial allocation arms (modelled for carriage for ancestral allele)

Trial arm	Myalgia events	OR (95% CI)	P value	Interaction OR (95% CI)	P value
Rosuvastatin	439	0.97 (0.79,1.19)	0.76	0.77(0.57,1.03)	0.085
Placebo	398	1.27 (1.01,1.56)	0.036		

The interaction terms is between trial therapy-allocation arm and the genotype

From the stratified analyses we conclude that there is an association between the Asp247 variant in *LILRB5* and the development of myalgia in the placebo arm. To understand the differentiation of risk across genotype and treatment allocation groups we conducted a stratified survival analysis.

5.4.1.4 Incidence of myalgia and the *LILRB5* Asp247Gly variant

This association is explored further in a survival analysis performed on the study participants with myalgia as the outcome as seen in Figure 5.1 in a 5 year follow-up period. The Kaplan-Meier showed that individuals homozygous for Asp247 regardless of whether they were allocated to receive statin or placebo, were at most risk of developing myalgia. Those most protected from the development of myalgia were carriers of the 247Gly variant who were randomized to receive placebo.

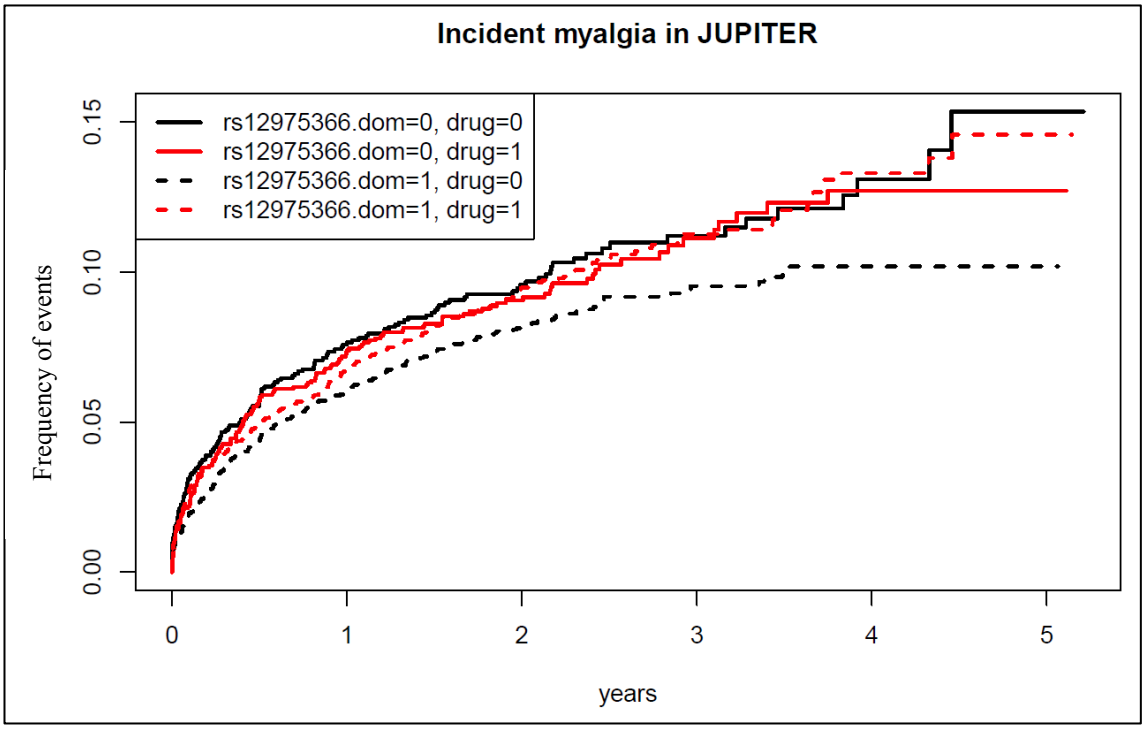


Figure 5.1 Kaplan-Meier of the association of myalgia with Asp247Gly. Plot produced by Dr. Dan Chasman

(rs12975366.dom = 0: Asp247 or T/T, rs12975366.dom = 1: 247Gly or C/T or C.C). Drug = 1: Rosuvastatin arm, drug = 0: Placebo arm)

Detailed output from the survival analysis are presented in Table 5.9.

Table 5.9 Hazard ratios of survival to myalgia

Variables	Hazard Ratio (95% CI)	P value
Asp247Gly	1.24 (1.02, 1.52)	0.039
Rosuvastatin	0.95 (0.725, 3.24)	0.63
Interaction	0.79 (0.60, 1.04)	0.10

5.4.1.5 Association of statin use with myalgia stratified by *LILRB5* Asp247Gly genotypes

The protective effect of the *LILRB5* 247Gly variant, is not seen in the statin treated group, where the levels of myalgia are similar to the Asp247 homozygous group. This reveals a genotypic subgroup that is apparently protected from statin independent myalgia, but is selectively susceptible to statin induced myalgia. This data agrees strongly with the cross sectional analysis presented in Table 5.8.

Table 5.10 Association of Rosuvastatin treatment with myalgia stratified by genotype

	Hazards Ratio (95% CI)	P value
Asp247 (T/T)		
Rosuvastatin v. placebo	0.95 (0.76, 1.19)	0.64
247Gly (C/T or C/C)		
Rosuvastatin v. placebo	1.2 (1.011, 1.42)	0.037

Replication in the PREDICTION-ADR cohort

In this cohort, clinically-adjudicated cases of statin-induced myositis (SIM) were examined. Due to the size of the cohort and MAF of the *CKM* 83Gly variant we were underpowered to detect an association between the variant and SIM.

For the *LILRB5* Asp247 variant, we observed consistent with all our findings so far, that those with the Asp247 genotype had 1.5 times the odds of developing SIM (95% CI: 1.03, 2.16, p value = 0.033). The cases and controls were matched for sex, type of statin at time of event and study center.

A cross-tabulation of the cases and controls in this cohort by their genotype is presented in Table 5.11 below.

Table 5.11 Distribution of LILRB5 Asp247Gly genotypes by clinically adjudicated SIM status

Group	Asp247	247Gly	Total
Statin-induced myositis	102	147	249
Statin tolerant controls	78	168	246

5.5 Conclusion and Discussion

We find a consistent relationship between various definitions of statin intolerance and created using EMRs and the *LILRB5* Asp247 variant in the GoDARTS population. There is a congruent association with non-compliance in the JUPITER trial, which has a median follow-up period of 1.9 years (with a maximum of 5) and with clinically adjudicated cases of statin-induced myopathy.

However, the variant in *CKM* appears to be associated purely with constitutive serum CK levels. As hypothesized in Chapter 4, this divergence between measurable/functional CK and physician documented myalgia points to a hypothesis for the occurrence of muscle-based symptoms with no accompanying rise in CK levels. It also calls into question the validity of using CK as the cornerstone for classifications of statin intolerance, especially the exclusive use of extremely high thresholds that carriers of the *CKM* 83Gly variant might not be able to achieve.

The association of *LILRB5* Asp247 with the development of myalgia in the JUPITER trial suggests a more complex gene-drug interaction. The Asp247 variant which is robustly associated with higher CK levels is also associated with greater statin intolerance in GoDARTS, with evidence of higher levels of statin switching and discontinuation. The JUPITER study allows us to look at these effects in the absence of statins and this reveals that Asp247 is associated with greater frequency of myalgia, over the study period. This would support that concept that *LILRB5* modulates CK levels through statin independent muscle damage. This would give rise to increased apparent intolerance in a statin treated population where the causality in not evident. Even more intriguing, however is the finding that the protected group, the 247Gly carriers, display a higher risk of statin induced myalgia. This suggests that there is a subpopulation of individuals who are inherently protected from muscle pain, but are susceptible to true “statin induced” pain.

We have opted to differentiate between an observational study and a RCT by the use of adherence and compliance in the respective populations. The term adherence suggests a "persistence in practice" which is more applicable to a real-world setting of medication consumption as opposed to compliance which implies "acting in accordance with" the demands of a RCT (151). Participants in the JUPITER trial had been pre-selected for good compliance in the trial placebo run-in period. Therefore, the reported finding of non-compliance is more likely to be associated with intolerance, rather than inter-individual variability in compliance to medication.

Certain co-medications are known to interact with statin use. Fibrates, especially gemfibrozil, are inhibitors of the hepatic uptake transporter OATP1B1 which transports statins into the liver for metabolism (152). Therefore, concurrent use of fibrates could increase un-metabolized statin concentration in circulating blood (153). Fibrates are prescribed to Type 2 diabetics to control triglyceride levels, and are generally not recommended as a co-medication with statins for primary lipid control (10). However, since GoDARTS is primarily a population of T2 diabetics, fibrates are widely prescribed. Additionally, other drugs that are known inhibitors or substrates of CYP3A4 enzymes are contra-indicated for concurrent use with statins. These co-medications are included in analyses of statin intolerance. We find these drugs appear to consistently increase the risk of statin intolerance, but the association is independent of the effect of the Asp247 variant as seen in the adjusted models.

This is the first evidence from JUPITER (or any other RCT) of any role of statins in myalgia, as there is absolutely no difference in the incidence of any measure of muscle pain or myalgia between the placebo and rosuvastatin arm, but the incidence of these phenotypes in both arms is similar to the "intolerance" observed in populations. The complete lack of association of statins with muscle pains in RCTs has led to a debate regarding the existence of statin related muscle symptoms. Indeed the validity of this

conclusion is supported by the recently concluded Goal Achievement after Utilizing an anti-PCSK9 Antibody in Statin-Intolerant Subjects -3 (GAUSS-3) trial, that was designed to examine the efficacy of non-statin therapies in lowering LDL for individuals with uncontrolled LDL, who were also statin intolerant (154). The trial had a randomized, blinded crossover run-in period, where participants were given either placebo or 20 mg atorvastatin followed by a washout period and the alternate therapy. This run-in was designed to eliminate those whose complaints of myalgia were not statin specific. The results revealed that 37% of participants reported non-specific myalgia (either in response to both therapies or just to placebo), while 43% had atorvastatin-specific complaints.

The *LILRB5* 247Gly genotype presents a unique opportunity to probe this phenomena of muscle pain specific to statins compared to "constitutive' muscle pain that appears in the Asp247 individuals.

6 Statin intolerance and the *LILR* gene family

6.1 Introduction – LILR gene family in GWAS

In Chapter 5, an association between a missense variant in *LILRB5* and statin intolerance and adherence was demonstrated and replicated. The variant was also observed to have a statin-independent effect on myalgia. The novel implication of the immune system in statin intolerance warrants further probing of the functional role of the gene itself, and its gene family. A candidate gene approach was employed with an aim to gain a more complete understanding of how and why these effects were being observed.

The LILR family

The *LILR* are located in the Leukocyte Receptor Complex (LRC) (117) on chromosome 19. The LRC encodes the Ig-superfamily proteins that regulate the function of various hematopoietic cell types and other genes also associated with immunoregulatory functions; including Killer Immunoglobulin-Receptors (*KIR*) and Leukocyte Associated Inhibitory Receptors (*LAIR*) and the Leukocyte Immunoglobulin-like Receptor family (*LILR)* (116). The *LILR* are a family of 11 genes; five activating (*LILRA1, 2, 4*-6), five inhibitory (*LILRB1*-5) and one has soluble (*LILRA3*) form. They are expressed on cells of myeloid and lymphoid lines, indicating that they are involved in both the innate and adaptive arms of the immune system. Several members of the LILR family recognize Human Leukocyte Antigen (HLA) class 1 molecules (116,117,155,156). LILR activity influences the antigen presenting properties of macrophages and dendritic cells, and may thus play a role in T cell tolerance (116). The range of effects of LILR signaling on immune cell activity indicates that they might be involved in a wide range of clinical situations.

6.1.1.1 Gene organization

From the 5' end (positive strand) in the centromeric cluster the genes are in the following order: *LILRB3, LILRA6, LILRB5, LILRB2, LILRA3, LILRA5,* and *LILRA4*. In the telomeric cluster they are: *LILRA2*, *LILRA1*, *LILRB1* and *LILRB4* as seen in Figure 6.1. This study focuses on the centromeric cluster of genes.

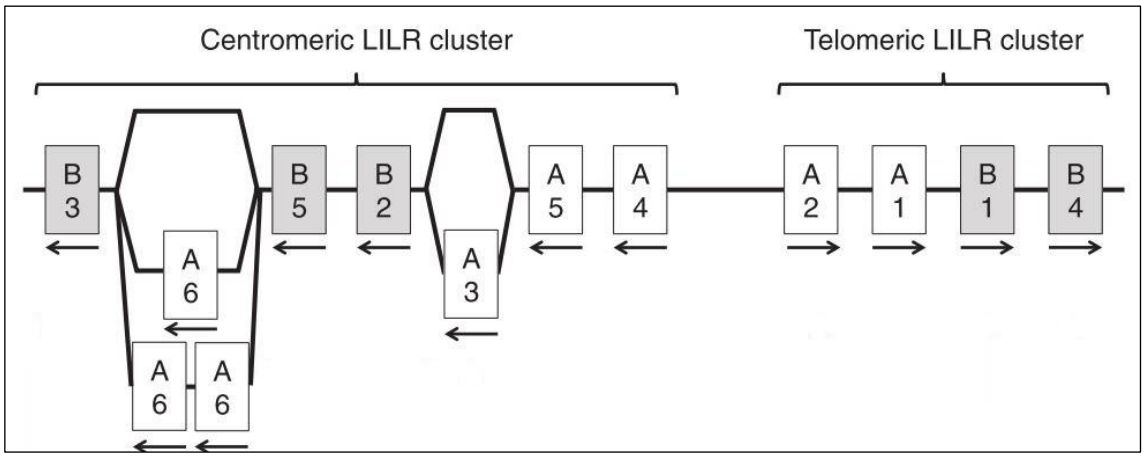


Figure 6.1 Gene organization of the *LILR* adapted from Hirayasu *et al.* © 2015 Nature Publishing Group, Hirayasu, K *et al*., Functional and genetic diversity of leukocyte immunoglobulin-like receptor and implication for disease associations. Jour of Hum Gen , 6, 64 (2015). All rights reserved.

6.1.1.2 Structure of the LILR receptors

Activating (LILRAs) or inhibitory (LILRBs) receptor isoforms are defined by residues found within their transmembrane or cytoplasmic domains as seen in Figure 6.2. Activating LILRs have a short cytoplasmic tail and an arginine residue located within the transmembrane domain. Protein sequence motifs; immunoreceptor tyrosine-activating motifs (ITAMs) and immunoreceptor tyrosine-inhibitory motifs (ITIMs) are responsible for the activating and inhibitory signals transmitted by *LILR* respectively (116).

Inhibitory LILRs do not express a charged arginine residue in the transmembrane region, they instead have a long cytoplasmic tail with two or four ITIM domains. The ITIM motifs are phosphorylated upon cell activation and receptor ligation and inhibit leukocyte activation through SHP (Src-homology 2-domain-cotaining protein tyrosine)

phosphatase recruitment (156). Stimulatory receptors have a shorter tail and interact with ITAM incorporating adaptor molecules to activate immune cells.

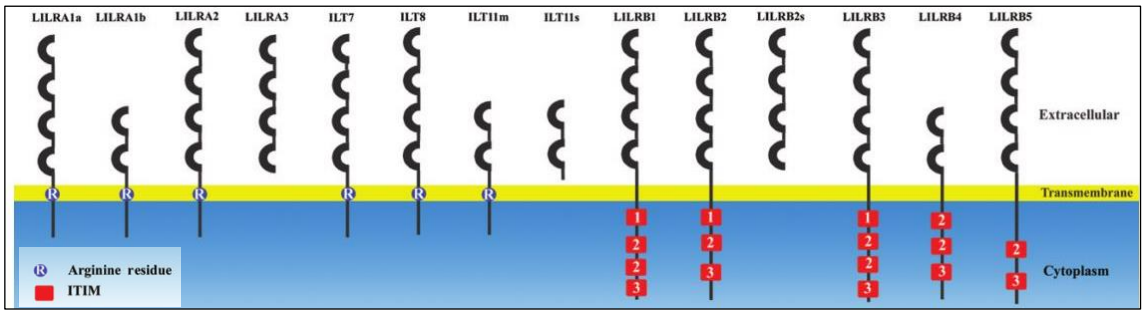


Figure 6.2 Schematic diagram of the LILR family adapted from Brown *et al.* showing their specific cytoplasmic and transmembrane features. © 2004 Blackwell Munksgaard, Brown, D *et al*., The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens 5, 64 (2004). All rights reserved.

Both inhibitory and stimulatory LILRs have 2-4 extracellular Ig domains which are responsible for ligand binding at the cell surface. The 2-4 extracellular Ig domains are termed D1, D2, D3 and D4 (as seen in Figure 6.3). The membrane distal D1 and D2 domains form the sites for binding to HLA class I ligands for receptors such as LILRB1 and LILRB2. The D3 and D4 domains form a stalk region, enabling some LILR receptors such as LILRB2 to bind to class I ligands both in *cis* (on the same cell) and in *trans* orientations (157). Recently it has been shown that both the D1 and D4 domains of LILRB2 are necessary for binding to the non-HLA ligand angiopoietin-like protein (Angtpl2) (156,158).

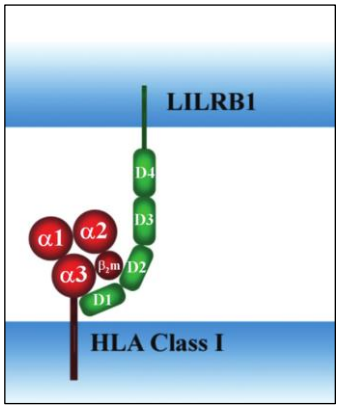


Figure 6.3 LILRB1 interacts with HLA class I molecules at two interfaces from *Brown et al.* © 2004 Blackwell Munksgaard, Brown, D *et al*., The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens 5, 64 (2004). All rights reserved.

Comparisons between human and chimpanzee *LILR* genes showed that the *LILR* genes are rapidly evolving and exhibiting greater interspecies differences than the

genome average (159). These differences make it difficult to elucidate the physiological function of LILR using an animal model. There are limited studies in humans but they have shown that *LILR* expressions are upregulated in peripheral blood leukocytes after endotoxin challenge (160). Since there is a LILR that recognizes viral proteins and since there is high interspecies variability, Hirayasu et al, hypothesize that *LILR* genes might have co-evolved with pathogens (117).

6.1.1.3 Genetic variation of the LILR family

There are two haplotypes in the gene cluster one at the centromeric cluster and one at the telomeric end as seen in Figure 6.1 (117). *LILRs* consist of 2 pseudogenes and 11 functional genes encoding five activating (*LILRA1, A2, A4, A5 and A6*), five inhibitory (*LILRB1-5*) and one soluble (*LILRA3*) form. The gene cluster is separated into centromeric and telomeric clusters. The centromeric and telomeric clusters are transcribed in opposite directions, from telomere to centromere and the reverse (117). The number of *LILR* genes is conserved among individuals, except for *LILRA3* and *LILRA6*, which show variation in copy number (CNV). *LILRA3* exhibits a presence or absence in variation due to a large deletion that encompasses almost the entire gene. The gene also shows the most extreme differences in copy number in the HapMap populations (117,161,162). Northeast Asians have a high frequency (up to 84%) of the *LILRA3*-deletion compared with that in European (17%) (163), and the deletion is uncharacterized in African populations. Common variants in the gene show similarly opposite frequencies in populations, and Sub-Saharan Africans have frequencies similar to Caucasians. In addition, it was found that in Northeast Asians the *LILRA3* deletion was in strong LD with variants in adjacent *LILRB2* and the region shows strong evidence of positive selection (162). Additionally, nonfunctional alleles containing premature stop codons have been detected in East Asians. The nonfunctional alleles are estimated to have been maintained for a long time in humans – suggesting that balancing

selection has been acting on this locus. There has been no characterization of the *LILRA3* deletion in a Scottish/British Caucasian population. In contrast, the copy number of *LILRA6* observed in an individual ranges from one to six, which can be explained by the deletion of duplication of this gene (164,165). The bulk of increased diversity seen in the African population is explained by the CNVs of *LILRA6* (166).

6.1.1.4 Functional role of LILR members

Information on the mechanistic mode of action for the genes is sparse. Zhang *et al.* performed tetramer staining studies to determine the ligands for *LILR* gene family members (156). They conclude, with confirmation from co-immunoprecipitation that HLA class I free heavy chains are ligands for *LILRB5*. They further hypothesize that this unique binding specificity is due to differences in the D1 and D2 Ig-like binding domains, which are distinct from other LILR genes; many of which bind to the β2m-associated HLA-class 1. This suggests that the roles of individual genes in this region could be independent, especially as pertains to an immune-mediated response to statin therapy and therefore their associations should be examined both independently and conditionally.

Information on the impact of these genes on human physiology implicates low-expressing *LILRB1* and non-deleted allele in *LILRA3* in the development of rheumatoid arthritis (RA) (155,163,167). A variant associated with low expression of *LILRB3* was associated with Takayasu's arteritis in a GWAS (168). Variants in *LILRB4* (169) and *LILRA2* were found to be associated with Systemic Lupus Erythematosus (SLE) (170). The same variant in *LILRA2* is associated with polyangiitis (170). It has also been speculated that *LILRA2* plays a role in leprosy – as elevated expression of the gene has been observed in lesions of lepromatous patients (171) and the mean numbers of *LILRA2*-expressing cells in synovial tissue are associated with the severity of RA (172). The deletion polymorphism in *LILRA3* has been found to be associated with several

autoimmune diseases. Since this deletion polymorphism shows significantly different frequencies across ethnicities (162) the association of diseases or phenotypes with the deletion polymorphism are highly specific to the population group. The deletion was found to be associated with multiple sclerosis in German and Spanish cohorts, however this effect was not replicated in a Polish population (163,173,174). The functional form of the gene was associated with Sjögren's syndrome in Chinese (167,175). Homozygosity of the non-deletion increased susceptibility to RA and SLE (167,176). A GWAS in a Chinese cohort found a SNP in *LILRA3* to be associated with prostate cancer, the SNP, rs103294, is in strong LD with the deletion polymorphism (177). A SNP located between *LILRB2* and *LILRA3* was found to be associated with plasma levels of high-density lipoprotein cholesterol (HDL) in a GWAS (178). A summary of the associations and the cells on which the receptors are expressed is presented in Table 6.1.

Table 6.1 LILR receptors, expression and disease associations

Gene	Expression (from Hirayasu *et al.*)	Disease
LILRB1	Subsets of T and NK cells, B cells, Mo, Mac, DC	RA (155,163,167)
LILRB2	Mo, Mac, DC, HSC, neuron	Unknown
LILRB3	Mo, DC, G	Takayasu's arteritis (168)
LILRB4	Mo, Mac, DC, plasmablast cells	SLE (169,170)
LILRB5	NK cells, Mo, mast cell granules	Serum CK levels (118,119), serum LHD levels (118) and statin intolerance
LILRA1	Mo, Mac, DC	Unknown
LILRA2	Subsets of T and NK cells, Mo, Mac, DC, G	SLE (167,176) + microscopic polyangiitis (170).
LILRA3	Subsets of T – cells, B cells, Mo, Mac, DC	MS (163,173,174), SS (167,175), SLE & SLE (167,176), prostate cancer (177)., HDL-C (178)
LILRA4	pDC	Unknown
LILRA5	Mo, PMN	Unknown
LILRA6	Mo	Unknown

NK: natural killer, MO: monocytes, MAC: macrophages, DC: dendritic cells, HSC: hematopoietic stem cells, G: granulocytes, pDC: plasmacytoid DC, PMN: poly morphonuclear neutrophil. SLE: Systemic lupus erythematosus, CK: creatine phosphokinase, RA: Rheumatoid arthritis, MS: multiple sclerosis, SS: Sjögren's syndrome, HDL-c: high density lipoprotein cholesterol

Study aims

The role of the immune system in drug intolerances was relatively unexplored with the exception of the Human Leukocyte Antigen (HLA) system (179–181). Evidence for the role of the immune system in statin intolerance is minimal, with the exception of the extremely rare, statin-association autoimmune myopathy noted in the presence of anti-HMGcR antibodies. There have been reports of the immunomodulatory effects of statins (182,183). This novel exploration of the *LILR* gene family could shed light on the potential role played by the immune system in drug adherence and intolerance.

Study plan

While designing the candidate gene approach, external replication of findings was sought from the University of Liverpool (UoL), a partner in the PREDICTION-ADR consortium. The cohort analyzed in UoL contains cases and controls from the Clinical Practice Research Datalink (CPRD) and the Wellcome Trust Case Control Consortium (WTCCC). A meta-analysis across the studies in UoD and UoL was conducted for the genomic region of interest. Hits replicated across these studies would potentially have external validity to be genetic variations predisposing to statin intolerance.

Performing a meta-analysis of the gene region could result in many signals that are proximately located. Alleles of SNPs located in close proximity to another are not necessarily inherited independently, these alleles can be highly correlated such that many individuals could share the same haplotype (or combination of consecutive alleles on a single chromosome). Therefore tagSNPs can be selected, that best represent a haplotype block (184).

The scope of this chapter is to examine the signals from the gene family, determine their independence, and the strength of their associations with statin intolerance. In determining the independence of the signals, it was necessary to effectively characterize the deletion polymorphism in *LILRA3*. Therefore we used data from exome sequencing

performed in the PREDICTION-ADR consortium, to assess the association of the deletion polymorphism with the index SNP selected from our study and with the development of SIM.

6.2 Methods

Cohorts used

6.2.1.1 The GoDARTS study

The development of phenotypes is described in detail in Chapter 3 "Defining Statin Intolerance in GoDARTS". For this chapter, analyses were performed using phenotypes B – statin intolerance defined as raised CK after commencing statin therapy and poor adherence to therapy and C – poor adherence to statin therapy. Their genetic differences were contrasted with those categorized as phenotype E – statin tolerant defined as individuals who had normal CK measures while on statin therapy and who were highly adherent and phenotype F – statin adherent. The *LILR* region is sparsely typed on the genetic chips available in GoDARTS and we therefore used a combination of directly typed and imputed data available for the Affymetrix 6.0 and Illumina HumanOmni Express -12V1 platforms. Imputation was performed against 1000G Phase I V3 reference panel using Impute2 (126). Samples with imputation quality below 80% were not included in the analysis.

6.2.1.2 University of Liverpool cohort

The Liverpool cohort was composed of statin intolerant individuals from the Wellcome Trust Case Control Consortium (WTCCC) and the Clinical Practice Research Datalink (CPRD) (185) and statin tolerant individuals from CPRD (186). GWAS data was available for 585 statin exposed controls and 128 statin intolerant cases (187). Individuals were classified as cases if they had CK measures raised 4 or more times the upper limit of normal while on statin treatment; and as controls if they had used statins for at least 1 year and had either no recorded CK tests, normal CK tests and had no record of rhabdomyolysis. WTCCC genotyping was performed using the Affymetrix v 6.0 (Affymetrix, Santa Clara, USA) and Illumina 1.2 M (Illumina, San Diego, USA) chips. Imputation was performed against the 1000G Phase I V3 reference

panel using Impute2 (126). CPRD was used to replicate signals found in the WTCCC cohort. Genotyping for CPRD cases was performed on the Illumina OmniExpress Exome v1.0 Beadchip (Illumina, San Diego), while controls were genotyped for significant hits (p value < 1 x 10^{-6} in a primary GWAS) using Sequenom MassArray (Sequenom, San Diego) and TAQMAN (187). Data for meta-analysis was received as SNPTEST (188) output from a GWAS conducted by Dr. Dan Carr at the University of Liverpool, upon request. This cohort was used to perform a meta-analyses of the *LILR* gene region.

6.2.1.3 The JUPITER trial

The population of the trial are described in Chapter 5. Genotyping for the trial participants was performed on the Omni1-Quad platform (Illumina, San Diego). The outcomes being used in this study is change in CK from baseline. This was assessed as the absolute difference in CK levels collected when participants were treatment-naïve and CK tested from post-treatment samples. There was no specific time point for the collection of the subsequent sample in the study; therefore the time interval between the two samples differs by participant. All analyses in the JUPITER trial data were performed by Dr. Daniel I Chasman upon request. This cohort was used to validate signals reported from the meta-analysis on the *LILR* gene region. Due to the low frequency of non-compliance and myalgia, coupled with the low MAF of the variant identified through the meta-analysis and the intention to perform stratified/conditional analyses, we were underpowered to test these associations in the trial population.

6.2.1.4 PREDICTION-ADR study

The sequencing data used to characterize the *LILRA3* deletion polymorphism was obtained from 175 individuals in the Dundee arm of the PREDICTION-ADR study (99 cases of clinically adjudicated SIM and 76 controls). The selection of cases and controls is detailed in Chapter 1 and 5. All sequencing experiments were performed by Dr.

Cyrielle Maroteau upon request. Linkage testing between the sequenced variants and deletion polymorphism were performed by the candidate.

Regional GWAS

We selected the region on chromosome 19, between base positions 54720147 to 54850421 in the Dundee and Liverpool studies.

The regional GWAS in GoDARTS was conducted to contrast phenotype B: those who had raised CK and were non-adherent to their statin therapy (phenotype B) with those who had normal CK and were adherent to therapy (phenotype E). The GWAS was performed using SNPTEST (188), as a logistic model (binary outcome of case and control), using the frequentist association test method, and the score method to deal with genotypic uncertainty. The GWAS was run on separately on each platform (Affymetrix and Illumina) as described in Chapter 2. The model was adjusted for all the covariates associated with each phenotype discussed in Chapter 5, namely the age at which CK was tested, sex, interacting co-medications, the first daily dose and the first statin on treatment.

Meta-analysis

The meta-analysis of results from the GWA studies conducted in GoDARTS and the Liverpool cohort was conducted using the GWAMA (Genome-Wide Association Meta-Analysis) software (189). Results from GWAMA output file were then pruned to only include hits coming from the same direction from both studies and all three platforms (GoDARTS Affymetrix, GoDARTS Illumina and Liverpool) and those with p value lesser than 0.05.

This list was then annotated with the functional roles of each SNP and whether or not they were an eQTL for the gene they were located in, or the flanking genes. This information is presented in Table 6.5 and Table 6.6.

Selecting index SNPs from the meta-analysis

A TagSNP, as defined by Pettersson *et al.* is a SNP in a region of the genome featuring high LD, which is a proxy for others in close proximity and which can be used to genotype individuals at a reduced cost, while maintaining power (184). Index SNPs were selected using these criteria, additionally we selected based on imputation quality in the GoDARTS cohort so they could be subsequently used for other analyses, and for their functional roles. The index SNP thus selected from the *LILRB2-A3* block was rs383369 in *LILRB2*, (*see* Table 6.5); the trend of the variant with the surrounding SNPs is seen in both the effect allele frequencies and the direction of effects. Figure 6.7 shows the LD between the variant and the other signals in the genomic region. The only unlinked SNPs are from *LILRB5*. Additionally, rs383369 is a coding variant (His20Arg) and is an eQTL for *LILRB2* and *LILRA3* gene expression. These factors together make it the ideal index SNP. Further, the LD between the top hits was tested using the - - ld command on the imputed file for chromosome 19 in Plink 1.9 (190,191) and are presented in the results.

Conditional analysis on index SNPs in the GoDARTS study

The selected variant in *LILRB2*, His20Arg (rs383369) was then tested in regression models predicting creatine kinase levels, statin intolerance (phenotype B) and statin adherence (phenotype C); phenotypes described in Chapters 3 and 4. Models were built using the backwards step-wise method, including the *LILRB5* variant, Asp247Gly (rs12975366), *LILRB2* His20Arg, and other covariables known to be associated with the phenotypes. Associations for phenotype D (CK elevations 4 or more times the upper limit of normal) were not tested as we are insufficiently powered to detect effects due to the small sample size.

A linear regression for log transformed CK levels was performed and the results are presented in Table 6.7 as point estimates, standard error and p values. Logistic

regression analyses were performed for the association with phenotypes B and C. The results reflect Maximum Likelihood Estimates. Percent variability (R^2) analyses for phenotypes B and C were performed using linear regression, results for which are presented in Table 6.8 and Table 6.9.

Characterizing the *LILRA3* deletion polymorphism in the PREDICTION-ADR cohort

Exome sequencing methods are described in Chapter 5. The deletion was characterized in 175 samples. Per base coverage analysis was performed on the *LILRA3* region (NG_034046.1, NCBI) using BEDTools (192). Visual representations of a random selection of individuals' coverage enabled us to fix discrimination thresholds for homozygotes and heterozygotes (*see* Figure 6.4). Subsequently, using those thresholds a custom AWK script (193) was used to assign a deletion status to each individual.

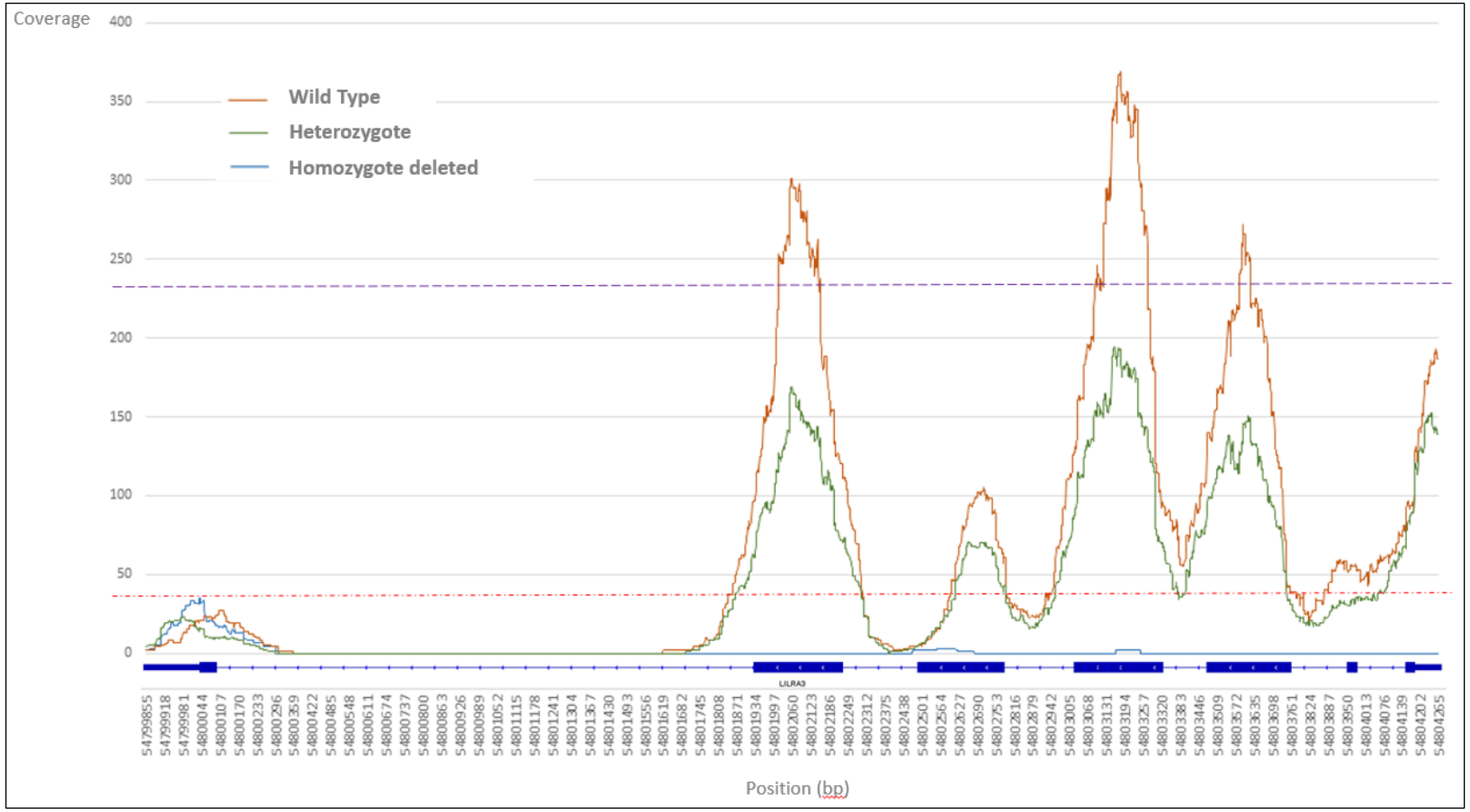


Figure 6.4 Graphical representation of the *LILRA3* deletion on chromosome 19.

The blue boxes at the base represent the exons, numbered 7 - 1(from left to right). The deletion is characterized by a drop in coverage as seen in the contrast between those homozygous for the ancestral allele (in orange), and those carrying the deletion heterozygously (in green). Coverage for homozygous carriers of the deletion is in blue and is only seen in exon 7. Thresholds for coverage are represented in purple for homozygous carriers of the ancestral allele (set at 240X and above) and in red for homozygous carriers of the deletion (set at 40X and below) for exon 4 and 6. Intermediate coverage represents heterozygotes. Figure produced by Dr. Cyrielle Maroteau.

6.2.1.5 Linkage between variants in LILRB2 and the *LILRA3* deletion polymorphism

The presence of absence of the deletion was treated as "0" for no deletion, "1" for heterozygous carriers of the deletion polymorphism and "2" for homozygous carriers. A similar dosage was applied for applied for the number of rare alleles carried at LILRB2 His20Arg. A Pearson correlation coefficient and the squared correlation coefficient were computed and reported. Diplotypes were used to calculate D and D' as highlighted in Equation 6.1

SNP1		**SNP2**	
Allele	**Frequency**	**Allele**	**Frequency**
G	p1	A	q1
C	p2	T	q2

Table 6.2 Representation of allele frequencies

		SNP2	
	Allele	**A**	**T**
SNP1	**G**	GA (p11)	GT (q12)
	C	CA (p21)	CT (q22)

Table 6.3 Representation of diplotype frequencies

Equation 6.1 Formula for relative measure of disequilibrium D'

$$D' = \frac{D}{D\ max}$$

Where D = p11*p22 – p12*p21, and D max = min [p1q2 or p2q1] when D is positive or D max = min [(p1q1) or (p2q2)] when D is negative (*see* Table 6.2 and Table 6.3) (194,195).

6.3 Results

The results are presented in the following order: results from the regional GWAS, linkage between the top hits and picking index SNPs, testing their association with CK levels and non-adherence in the GoDARTS population, performing conditional analyses with the *LILRB5* Asp247Gly and the index SNPs, replication of stratified effect in the JUPITER trial, eQTL effects of the main variants identified, and finally characterizing the LILRA3 deletion in exome sequenced samples and testing its linkage with the index SNP.

LILR Family regional GWAS analysis

6.3.1.1 Regional GWAS and meta-analysis

The LILR region GWAS was performed using SNPtest. The numbers available for the meta-analysis are shown in Table 6.4. The number available for this analysis include only samples for whom GWAS data was available on the Affymetrix and Illumina platforms.

Table 6.4 Number of individuals included in the meta-analysis of statin intolerance

Study	Number of cases	Number of controls	Total
Dundee*	181	645	826
Liverpool	128	585	713

**Dundee samples are those available on the GWAS platforms of Affymetrix and Illumina, which was necessary in order to perform a genome-wide meta-analysis.*

The results of the individual GWAS were meta-analyzed using GWAMA and the results are presented graphically in Figure 6.5 as a regional association plot using LocusZoom (196). Subsequently, the top hits are presented, sorted by their base position on chromosome 19 and annotated to represent the gene in which they are located, their functional role, and their association with statin intolerance in the meta-analysis. This information is presented in Table 6.5 and Table 6.6.

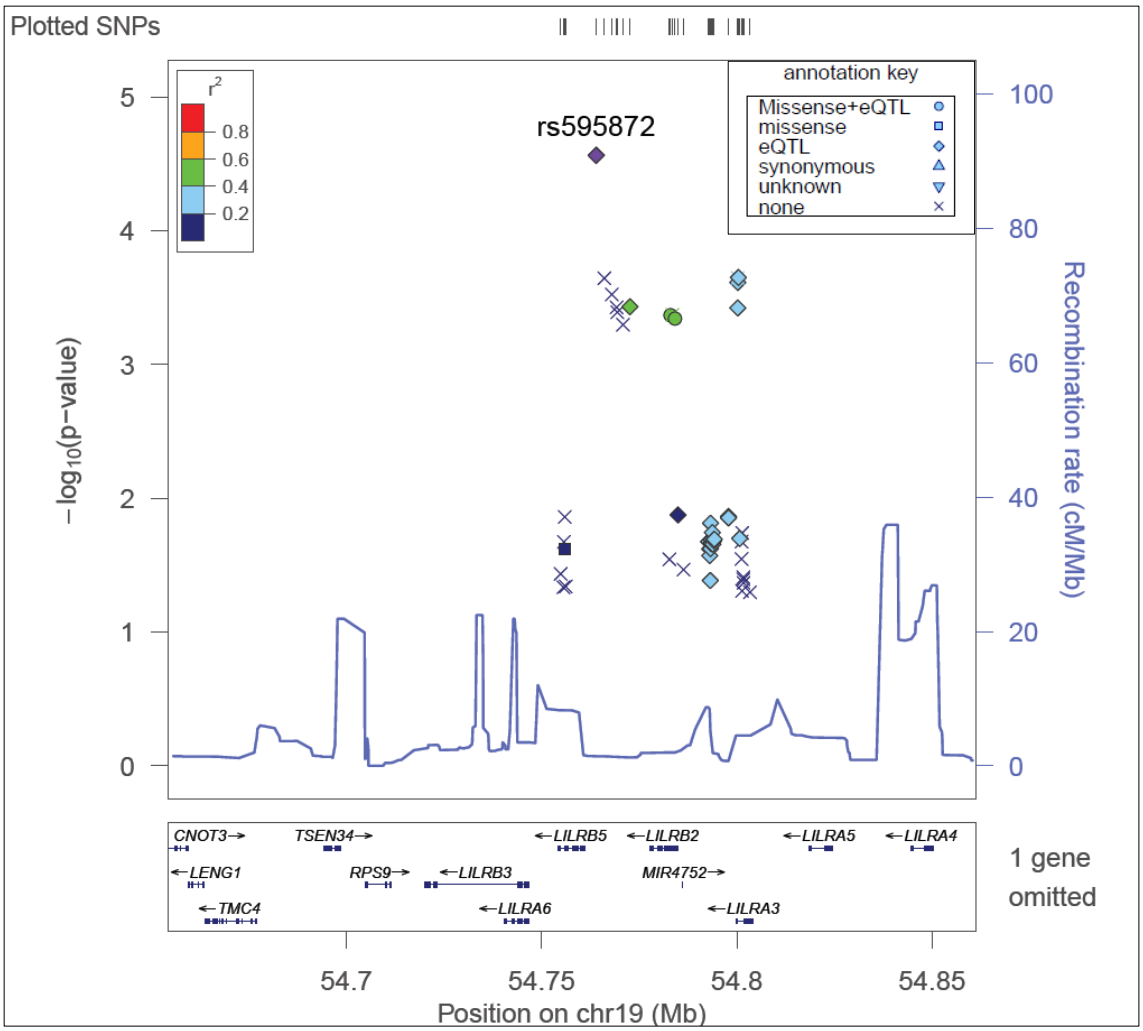


Figure 6.5 Results from the *LILR* region GWAS.

Results represented are only those that showed the same direction of effect in all three studies (the two platforms in Dundee and Liverpool). The shapes for each SNP correspond to a functional role, as annotated. The colours represent the r^2 between the SNPs and the index SNP rs595872

Table 6.5 Meta-analysis output for SNPs with a deleterious effect on statin intolerance

Position	Role	SNP	Substitution	eQTL	Gene	Reference Allele	Other Allele	EAF	OR	OR_95L	OR_95U	P Value
54754946	Intron	rs146755009		Yes	LILRB3	A	G	0.26	1.44	1.14	1.81	0.001993
54755636	Intron	rs624315		Yes	LILRB3	C	T	0.26	1.41	1.12	1.78	0.003886
54763969	Unknown	rs595872		Yes	LILRB2	A	G	0.23	1.59	1.28	1.97	2.72E-05
54764322	Unknown	rs389096		Yes	LILRB2	G	A	0.31	1.42	1.17	1.74	0.000527
54766423	Unknown	rs377681		Yes	LILRA3	G	A	0.17	1.51	1.19	1.92	0.00065
54772618	Unknown	rs4083825		Yes	LILRB2	C	A	0.18	1.53	1.21	1.93	0.000369
54775349	Unknown	rs419304		Yes	LILRB2	G	C	0.18	1.51	1.20	1.90	0.000518
54779101	Unknown	rs443501		Yes	LILRB2	T	C	0.16	1.50	1.16	1.94	0.002716
54779389	Intron	rs443874		Yes	LILRB2	T	C	0.16	1.47	1.15	1.89	0.002214
54779455	Intron	rs444004			LILRB2	T	C	0.79	1.42	1.13	1.79	0.002345
54781078	Intron	rs452717		Yes	LILRB2	T	G	0.16	1.55	1.21	1.98	0.000608
54781541	Intron	rs450937		Yes	LILRB2	T	C	0.13	1.45	1.10	1.92	0.009031
54781554	Intron	rs451000		Yes	LILRB2	T	C	0.13	1.44	1.09	1.91	0.010064
54781557	Intron	rs400942		Yes	LILRB2	C	T	0.13	1.44	1.09	1.91	0.010081
54782919	Missense*	rs386056	Val235Met	Yes	LILRB2	T	C	0.21	1.49	1.19	1.85	0.000423
54783375	Missense	rs373032	Asp161Glu		LILRB2	A	T	0.23	1.41	1.12	1.77	0.003142
54783521	Unknown	rs383925		Yes	LILRA3	T	C	0.21	1.48	1.19	1.85	0.000424
54783923	Synonymous	rs366337			LILRB2	G	A	0.94	1.95	1.28	2.96	0.001727
54784130	Missense *	rs383369	His20Arg	Yes	LILRA3	C	T	0.21	1.48	1.19	1.85	0.00045
54784920	5' UTR	rs448083		Yes	LILRA3	T	C	0.21	1.47	1.18	1.84	0.000507
54784936	Unknown	rs200926686			LILRA3	T	C	0.22	1.47	1.18	1.83	0.000526
54784936	Unknown	rs448092		Yes	LILRB2/ MIR475	T	C	0.22	1.47	1.18	1.83	0.000833
54785593	Unknown	rs380573			LILRB2	G	A	0.80	1.46	1.16	1.84	0.001342
54786546	Upstream	rs436911			LILRB2	G	C	0.80	1.44	1.15	1.82	0.001859
54786922	Unknown					A	A	0.80	1.38	1.10	1.73	0.006058
54786925	Upstream	rs367720231	in-del		LILRB2	G	G	0.80	1.38	1.10	1.73	0.006063

54792079	Unknown	rs431420		Yes		T	G	0.21	1.45	1.17	1.80	0.000651
54792761	Unknown	rs386000		Yes		C	G	0.23	1.27	1.04	1.57	0.021027
54792769	Unknown	rs386003		Yes		T	G	0.23	1.27	1.04	1.57	0.021041
54793038	Unknown	rs398217		Yes		G	A	0.23	1.27	1.03	1.56	0.023945
54793048	Unknown	rs398227		Yes		C	A	0.23	1.26	1.03	1.55	0.02681
54793188	Unknown	rs798887		Yes		G	A	0.23	1.24	1.01	1.52	0.041111
54793250	Unknown	rs798889		Yes		T	G	0.23	1.29	1.05	1.58	0.015337
54793273	Unknown	rs57827784		Yes		A	T	0.24	1.30	1.06	1.59	0.013254
54793280	Unknown	rs56883673		Yes		C	T	0.24	1.31	1.07	1.61	0.010387
54793357	Unknown	rs57906249		Yes		A	G	0.24	1.28	1.04	1.57	0.020315
54793360	Unknown	rs59605183		Yes		T	C	0.24	1.27	1.04	1.57	0.021566
54793415	Unknown	rs58609643		Yes		T	C	0.23	1.27	1.04	1.56	0.021776
54793505	Unknown	rs399657		Yes		G	A	0.23	1.28	1.04	1.57	0.0204
54793721	Unknown	rs61703366		Yes		A	G	0.23	1.28	1.04	1.58	0.017989
54793830	Unknown	rs798893		Yes		C	G	0.23	1.27	1.04	1.56	0.022028
54794098	Unknown	rs397558		Yes		C	G	0.23	1.28	1.04	1.57	0.020733
54794205	Unknown	rs419772		Yes		C	T	0.23	1.28	1.04	1.57	0.020206
54795299	Unknown	rs427366			LILRA6	C	T	0.71	1.39	1.14	1.71	0.001396
54796630	Unknown	rs103294		Yes	LILRA3	A	G	0.23	1.29	1.05	1.58	0.013994
54796719	Unknown	rs384116		Yes		G	C	0.24	1.32	1.08	1.61	0.007175
54799083	Unknown	rs380267		Yes		G	A	0.23	1.31	1.07	1.61	0.009691
54799692	Unknown					G	G	0.24	1.38	1.12	1.69	0.001963
54800222	Unkown	rs368178		Yes	LILRA3	G	A	0.33	1.44	1.18	1.75	0.000243
54800225	MNV	rs71302140	CCTC/TCTT		LILRA3	G	A	0.33	1.44	1.19	1.75	0.000225
54800225	Unknown	rs368177		Yes	LILRA3	G	A	0.33	1.44	1.19	1.75	0.000377
54800371	Unknown	rs410852		Yes	LILRA3	C	T	0.33	1.44	1.19	1.74	0.000223
54800500	Unknown	rs367070		Yes	LILRA3	G	A	0.23	1.33	1.08	1.63	0.006749
54800679	Unknown	rs798895		Yes	LILRA3	C	T	0.29	1.42	1.14	1.76	0.001657
54800703	Intron	rs798896		Yes	LILRA3	C	T	0.17	1.44	1.06	1.95	0.019957
54800810	Unknown				LILRA3	C	C	0.32	1.38	1.13	1.69	0.001851
54800810	Unknown	No info				C	C	0.32	1.38	1.13	1.69	0.002695

54800856	Unknown				LILRA3	G	G	0.22	1.33	1.08	1.64	0.007459
54801469	Unknown	rs7245916			LILRA3	G	T	0.13	1.70	1.24	2.33	0.000998
54801475	Unknown	rs7245918			LILRA3	G	T	0.13	1.64	1.20	2.24	0.001805
54802440	Unknown	rs7508470			LILRA3	G	T	0.13	1.68	1.23	2.31	0.001273
54803504	Missense	rs6509862	Leu107Arg		LILRA3	C	A	0.13	1.73	1.26	2.38	0.000746

*Table legend: SNPs labelled in bold are lead to amino acid substitutions, * represents eQTL effect on the same gene.*

Table 6.6 Meta-analysis output for SNPs with protective effect on statin intolerance

Position	Role	SNP	Substitution	eQTL	Gene	Reference Allele	Other Allele	EAF	OR	OR_95L	OR_95U	P Value
54754385	3' UTR	rs2361796		Yes	LILRB5	G	T	0.32	0.78	0.64	0.95	0.013302
54754865	Intron	rs62133136			LILRB5	G	A	0.16	0.76	0.58	0.98	0.036724
54755572	Intron	rs111901459			LILRB5	T	C	0.16	0.77	0.59	1.00	0.046198
54755709	Intron	rs12986034			LILRB5	A	G	0.22	0.76	0.60	0.96	0.021144
54755911	Intron	rs12977057			LILRB5	T	C	0.22	0.74	0.59	0.94	0.013765
54756155	Intron	rs45446093			LILRB5	C	T	0.16	0.76	0.59	0.99	0.045662
54759361	Missense	rs12975366	Asp247Gly		LILRB5	C	T	0.39	0.72	0.59	0.87	0.000896
54759666	Intron	rs10405357			LILRB5	C	T	0.44	0.73	0.60	0.88	0.001041
54760691	Intron	rs11879136			LILRB5	T	C	0.25	0.69	0.55	0.86	0.001054
54762408	Unknown	rs6509859			LILRB2	T	C	0.25	0.70	0.56	0.88	0.001968
54766055	Unknown	rs34450379			LILRB2	C	T	0.22	0.65	0.52	0.82	0.000226
54767965	Unknown	rs12974390			LILRB2	A	G	0.22	0.66	0.53	0.83	0.000299
54769205	Unknown	rs12984962			LILRB3	C	T	0.31	0.69	0.57	0.85	0.000373
54769366	Unknown	rs12984029			LILRB2	A	G	0.22	0.67	0.53	0.83	0.000406
54770827	Unknown	rs4090914			LILRB2	G	T	0.22	0.67	0.54	0.84	0.000503

Table legend: SNPs labelled in bold are lead to amino acid substitutions

6.3.1.2 Linkage with index SNP in *LILRB5*

The Asp247Gly variant in *LILRB5* was selected as the index SNP to represent the signal arising from the gene. Using criteria explained in 0 Selecting index SNPs from the meta-analysis. Additionally, the association with this variant has been studied in-depth in Chapter 5. The linkage between Asp247Gly and the other significantly associated SNPs in the meta-analysis was checked. The results are presented in Figure 6.6. The signals from *LILRB5* appear to be independent of others in the region, specifically signals from *LILRB2* and *LILRA3*. Therefore the variant can be used to represent a signal from *LILRB5* while predicting effects on the phenotypes discussed previously in this chapter.

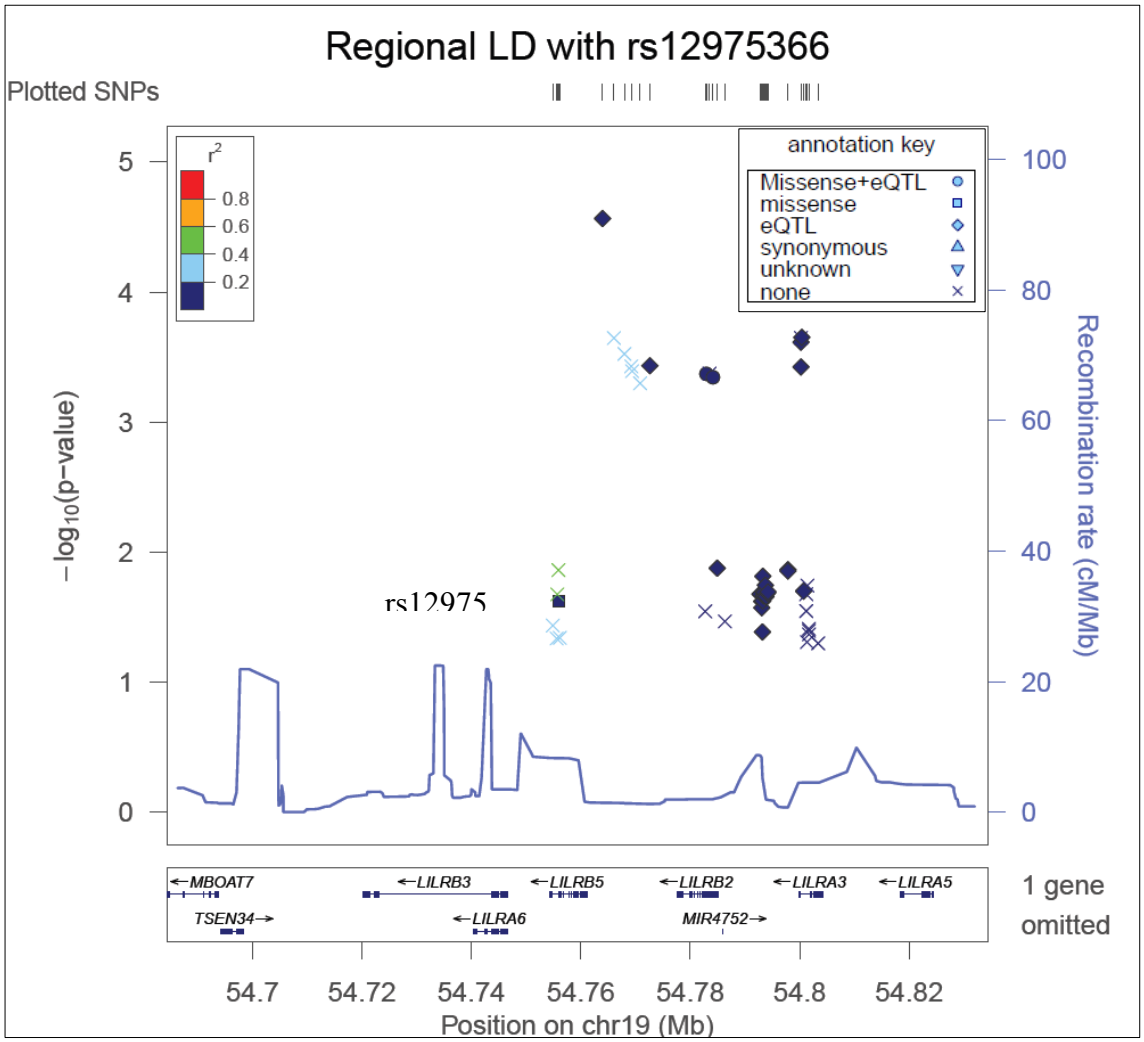


Figure 6.6 Linkage with *LILRB5* Asp247Gly (rs12975366)

Results represented are only those that showed the same direction of effect in the GWAS of all three studies (the two platforms in Dundee and Liverpool). The shapes for each SNP correspond to a functional role, as annotated. The colours represent the r2 between the SNPs and the LILRB5 variant rs12975366 (Asp247Gly)

6.3.1.3 Linkage with index SNP in *LILRB2*

The His20Arg variant in *LILRB2* was selected as the index SNP to represent the signal arising from the gene. Using criteria explained in 0 Selecting index SNPs from the meta-analysis. There were two functionally comparable variants in *LILRB2* (His20Arg and Val235Met), that were also eQTLs for gene expression. We compared their imputation quality in the GoDARTS population, and His20Arg was found to be better imputed (average information score 0.89 compared to 0.84). In Figure 6.7, we examined the linkage between *LILRB2* His20Arg (in purple) and the other signals from the meta-analysis.

The variant was in strong LD with other signals from the gene, especially the other missense + eQTL variant, Val235Met in *LILRB2*. Using the imputed chromosome 19 variant calling file in Plink 1.9, the r^2 and D' between the two variants rs386056 (Val235Met) and rs383369 (His20Arg) was 0.99 and 1 respectively (190,191). Furthermore, congruous signals from abutting SNPs indicate they belong to a haplotype, and that the His20Arg variant could be used to represent the effect. The *LILRB2* His20Arg variant appears linked to SNPs in *LILRA3* as well, albeit not as strongly. However, we were unable to test the linkage with the deletion polymorphism in *LIRA3* using the chip-based genotyping and imputation methods available to us. This linkage is further examined in section 0 "Linkage between the *LILRA3* deletion polymorphism and variants in *LILRB2*". Therefore, going forward, we employ *LILRB2* His20Arg to be the index SNP representing signals from the gene.

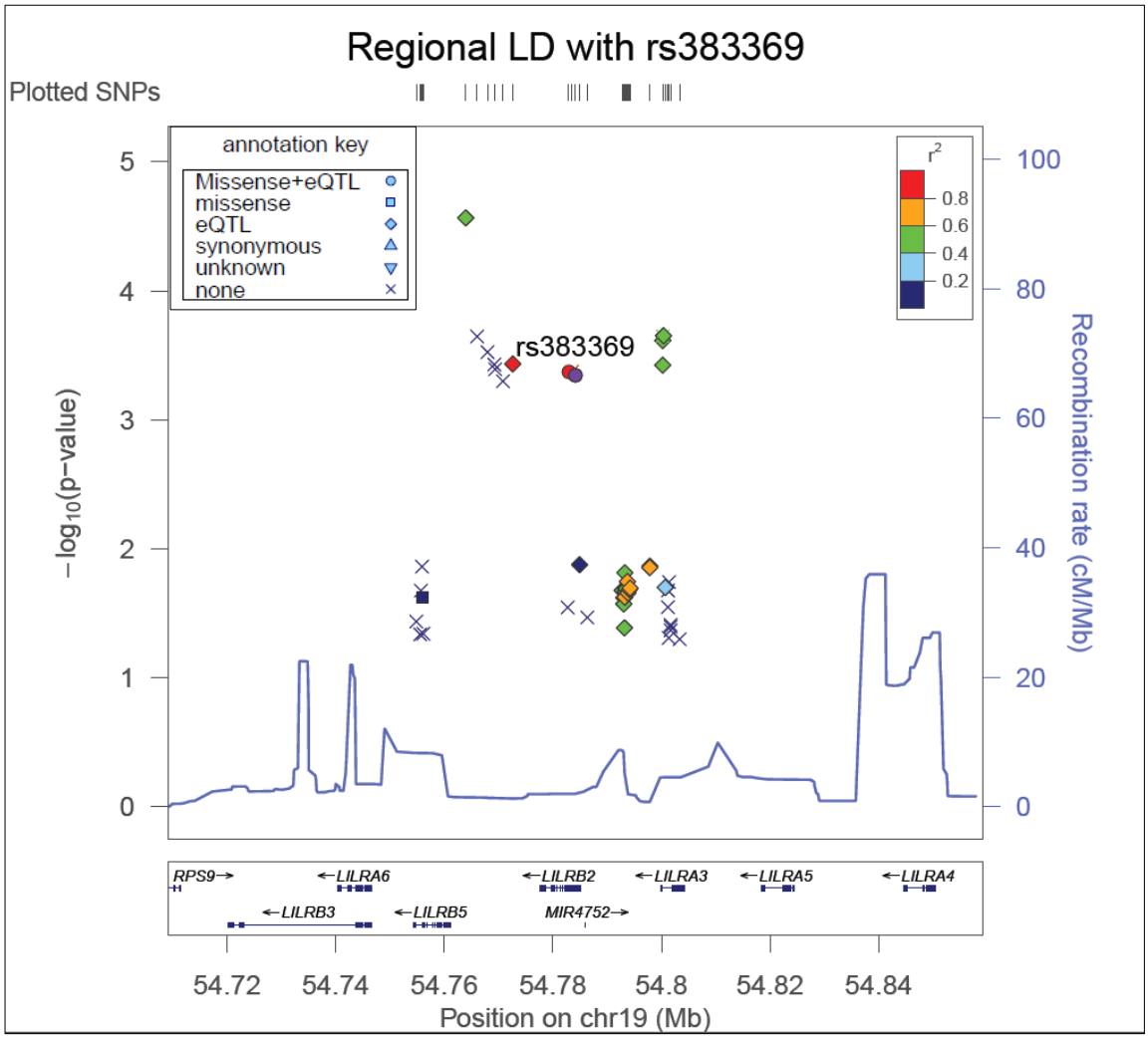


Figure 6.7 Linkage with *LILRB2* His20Arg (rs383369)

Results presented are only those that showed the same direction of effect in the GWAS of all three studies (the two platforms in Dundee and Liverpool). The shapes for each SNP correspond to a functional role, as annotated. The colours represent the r2 between the SNPs and the LILRB2 variant rs383369 (His20Arg)

The index SNPs selected in *LILRB5* and *LILRB2* were thus both missense variants, and they were not in linkage disequilibrium with each other. Using the imputed chromosome 19 variant calling file in Plink 1.9, the r^2 and D' between the two variants *LILRB5* Asp247Gly and *LILRB2* His20Arg was 0.04 and 0.54 respectively (190,191).

Independent and Conditional analyses using index SNPs

Finally, conditional analyses were performed for the two SNPs selected from the meta-analyses; rs12975366 (Asp247Gly in *LILRB5*) and rs383369 (His20Arg in *LILRB2)*. The *LILRB2* variant, His20Arg was tested for association with CK levels, statin intolerance (Phenotype B) and statin adherence (Phenotype C). The two variants were then tested together for independence of association in predictive models. The main effects models adjusted for the variants one at-a-time then both together. Finally they were tested in full models containing phenotypic variables that were found to be significantly associated with the specific phenotypes in Chapter 3.

6.3.1.4 Association of index SNPs with CK levels

The association with (log transformed) creatine kinase levels was conducted in a sample of 4,224 individuals for whom CK, as well as genotype information was available. As seen in Chapter 4, carriers of the *LILRB5* Asp247Gly variant had significantly lower CK levels; the variant predicted 0.5% of the variation in CK.

While the *LILRB2* His20 shows a slightly weaker association, carriers of the variant had higher CK levels and the variant predicted 0.1% of the variation in CK.

The full model, including gender and age can predict 13.3% of the variation CK levels. Results are presented in Table 6.7.

Table 6.7 Association with CK levels

Log Transformed Creatine Kinase (IU/L) (n = 4224)				
Variables	**Beta**	**SE**	**P Value**	**R^2 (%)**
Univariate models				
rs12975366	-0.03	0.005	<0.0001	0.5
rs383369	0.014	0.006	0.025	0.1
Gender (F v. M)	-0.14	0.007	<0.0001	7.4
Age	-0.004	7.4 e-4	<0.0001	2.3
Full model				
rs12975366	-.02	0.005	<0.0001	
rs383369	0.005	0.006	0.4	13.3
Gender (F v. M)	-0.13	0.007	<0.0001	
Age	-0.004	8.6 e-4	<0.0001	

SNP annotation: rs12975366 = LILRB5 Asp247Gly and rs383369 = LILRB2 His20Arg

6.3.1.5 Association of index SNPs with statin intolerance: raised CK and non-adherence to therapy

Carriers of the *LILRB5* Asp247Gly variant were observed to be protected from developing raised CK and being non-adherent to statin therapy as seen in Chapter 5. *LILRB2* His20 was also associated with statin intolerance, with carriers of the variant having higher risk. Independently, the variant predicted 0.5% of the variance in the trait. However, when both variants were added into the same model, the association with *LILRB2* His20 was substantially weakened. In order to examine if the associations were independent or if there was an interaction between the two, we stratified the association by *LILRB5* Asp247 genotype (T/C or C/C) or non-carriers (T/T). We found that the deleterious effect of the *LILRB2* His20 variant was only significant for those who did not carry the variant at *LILRB5* Asp247 ancestral (T/T). This stratification of effect was also observed in the full predictive model containing covariables known to be associated with the outcome (OR 1.4, p value = 0.049). In the fully adjusted model, the R^2 is lower in the group with the *LILRB5* Asp247 (T/T) genotype than in the 247Gly variant carriers (T/C or C/C). This might be due to diminished sample size in that group. Further, for this analyses only individuals with genotype data available for both SNPs were used in order to maintain a consistent set of individuals through the analyses. The results are presented in Table 6.8.

Table 6.8 Association with raised CK and non-adherence (B v. E)

Phenotype B: (n = 660: n cases = 250, n controls = 410)				
Variables	**Beta**	**SE**	**P Value**	**R2 (%)**
Univariate models				
rs12975366	-0.103	0.025	<0.0001	2.0
rs383369	0.06	0.03	0.04	0.5
rs12975366 + rs383369	-0.09 0.06	0.03 0.03	0.0016 0.07	2.5
Main Effect Stratified by LILRB5 Asp247Gly (rs12975366) **247Gly (variant carriers) n = 466, n cases = 156, n controls 310**				
rs383369	0.16	0.18	0.36	
Asp247rs12975366 = T/T (ancestral allele) n = 318, n cases = 145, n controls = 173				
rs383369	0.32	0.18	0.076	
Full Models				
Full model rs12975366 (dom) rs383369 Starting statin (Oth v. S) Starting statin (A v. S) Starting dose Gender Age	 -0.134 0.155 0.72 -0.30 -0.009 0.32 -0.02	 0.037 0.08 0.20 0.16 0.001 0.15 0.007	 0.0003 0.05 0.0003 0.053 <0.0001 0.03 0.003	12.5
Full Model Stratified by LILRB5 Asp247Gly (rs12975366) **247Gly rs12975366 = T/C or C/C (variant carriers) n = 466, n cases = 156, n controls = 310**				
Full model rs383369 Starting statin (Oth v. S) Starting statin (A v. S) Starting dose Gender Age	 0.16 1.3 -0.96 -2.1 0.33 -0.027	 0.18 0.23 0.20 0.32 0.20 0.009	 0.37 <0.0001 <0.0001 <0.0001 0.095 0.006	19
Asp247rs12975366 = T/T (ancestral allele) n = 318, n cases = 145, n controls = 173				
Full model rs383369 Starting statin (Oth v. S) Starting statin (A v. S) Starting dose Gender (F v. M) Age	 0.36 0.99 -0.52 -0.98 0.29 -.05	 0.18 0.29 0.20 0.26 0.24 0.011	 0.049 0.0005 0.01 0.0001 0.22 0.20	6.7

Table legend: Oth: refers to all other statins, A: Atorvastatin S: Simvastatin. Gender comparisons are Female v. Male. SNP annotation: rs12975366 = LILRB5 Asp247Gly (dom = dominant model) and rs383369 = LILRB2 His20Arg

6.3.1.6 Association of index SNPs with statin intolerance: non-adherence to therapy, irrespective of CK

Carriers of the *LILRB5* Asp247Gly variant were observed to be protected from being non-adherent to statin therapy as seen in Chapter 5. The *LILRB2* His20Arg variant showed no significant association with statin non-adherence in a main effects model, however when stratified by *LILRB5* Asp247 genotype, a deleterious effect was noted. Carriers of the 20Arg variant had increased odds of being non-adherent to statin therapy if they did not carry the *LILRB5* Asp247 variant, predicting 1.08% of the variance in the trait. Consistent, with the previous analysis, this effect was strengthened in a fully adjusted model stratified by the Asp247 genotype. Again, the R^2 of the fully adjusted model in this genotype group is lower than those with the T/C or C/C genotypes, as stated previously, this might be due to a smaller sample size. Further, for this analyses only individuals with genotype data available for both SNPs were used in order to maintain a consistent set of individuals through the analyses. The results are presented in Table 6.9.

Table 6.9 Association with non-adherence, irrespective of CK (C v. F)

Phenotype C: (n = 1043: n cases = 369, n controls = 674)				
Variables	Beta	SE	P Value	**R^2 (%)**
rs12975366 (dom)	-0.06	0.03	0.03	0.4
rs383369	0.06	0.09	0.52	0.2
rs12975366 (dom) + rs383369	-0.13 0.075	0.12 0.11	0.11 0.50	0.55
Main Effect Stratified by LILRB5 Asp247Gly (rs12975366) 247Gly rs12975366 = T/C or C/C (variant carriers) n = 629, n cases = 209, n controls 420				
rs383369	-0.20	0.17	0.23	0.23
Asp247 rs12975366 = T/T (ancestral allele) n = 414, n cases = 160, n controls = 254				
rs383369	0.33	0.16	0.035	1.08
Full Model Stratified by LILRB5 Asp247Gly (rs12975366) 247Gly rs12975366 = T/C or C/C (variant carriers) n = 629, n cases = 209, n controls 420				
Full model rs383369 Starting statin (Oth v. S) Starting statin (A v. S) Last statin (Oth v. S) Last statin (A v. S) Starting dose Ending dose Gender Age	 -0.27 1.3 -0.75 1.54 -0.25 -2.12 -0.6 0.001 -0.04	 0.21 0.20 0.19 0.24 0.18 0.34 0.19 0.22 0.01	 0.21 <0.0001 <0.0001 <0.0001 0.16 <0.0001 0.0016 0.99 0.0004	21.7
Asp247 rs12975366 = T/T (ancestral allele) n = 414, n cases = 160, n controls = 254				
Full model rs383369 Starting statin (Oth v. S) Starting statin (A v. S) Last statin (Oth v. S) Last statin (A v. S) Starting dose Ending dose Gender Age	 0.41 0.79 -0.37 1.58 -0.32 -1.1 -1.12 0.41 -0.036	 0.19 0.23 0.21 0.33 0.21 0.29 0.22 0.24 0.013	 0.019 0.0007 0.07 <0.0001 0.13 0.0002 <0.0001 0.09 0.0061	19.6

Table legend: Oth: refers to all other statins, A: Atorvastatin S: Simvastatin. Gender comparisons are Female v. Male. SNP annotation: rs12975366 = LILRB5 Asp247Gly (dom = dominant model) and rs383369 = LILRB2 His20Arg

Replication in the JUPITER trial

We sought to replicate our findings of the stratification of the effect of the *LILRB2* His20 variant stratified by the *LILRB5* Asp247Gly variant in the trial setting. The imputation quality of *LILRB2* His20Arg variant, rs383369 was 98%, r^2=0.94 JUPITER cohort (115) in the. The MAF was 0.18.

6.3.1.7 Association of *LILRB2* variant with baseline and induced CK levels

The variant was not significantly associated with baseline CK levels (β = -0.0002, SE = 0.005, p value = 0.95). An analysis stratified by the *LILRB5* Asp247Gly variant also showed no significant association. This indicates that the variant has no relationship with CK levels in the trial and that its association with intolerance might occur from a mechanism independent to that of the *LILRB5* variant.

6.3.1.8 Association of *LILRB2* variant with changes in CK levels stratified by *LILRB5* variant

The *LILRB2* His20Arg variant was not associated with a change in CK from baseline to final measurement in a main effects model (β = 1.23, SE = 0.89, p value = 0.16.

However, when stratified by *LILRB5* Asp247 genotypes a significant association was observed among individuals who did not carry the variant, consistent with our findings so far. Individuals with the *LILRB2* 20Arg variant have increased risk of having raised CK after starting statin therapy if they did not carry the *LILRB5* Asp247 variant (β = 3.02, SE = 1.39, p value = 0.03, as seen in Table 6.10.

Table 6.10 Association of *LILRB2* His20Arg with changes in CK stratified by *LILRB5* Asp247 genotypes

***LILRB2* His20Arg**	**Beta**	**SE**	**P Value**
247Gly rs12975366 = T/C or C/C (variant carriers) n = 5679			
rs383369	-0.38	1.18	0.74
Asp247 rs12975366 = T/T (ancestral allele) n = 2403			
rs383369	3.02	1.39	0.03

Next, we examine how these variants impact on the expression of their own and neighboring genes, since the *LILRB2* variant is a known eQTL.

Expression quantitative trait loci

Gene expression in whole blood and skeletal muscle were examined for the variants in *LILRB5* and *LILRB2* in order to explore their functional roles using the Genotype-Tissue Expression (GTEx) Project data set (197–199).

6.3.1.9 *Cis*-eQTL effects of *LILRB5* variant on *LILRB2* gene expression (might be trans – regulatory)

The Asp247Gly variant was a significant eQTL for *LILRB2* expression in whole blood ($\beta = -0.10$, $p = 0.03$), and trended toward significance for skeletal muscle cells ($\beta = -0.11$, $p = 0.057$). Indicating that carriers of the 247Gly variant had lower expression of *LILRB2* in whole blood seen in Figure 6.8. However, the variant was only weakly associated with the expression of *LILRB5* in the spleen, but in no other tissues. It was not associated with the expression of any other members of the *LILR* family.

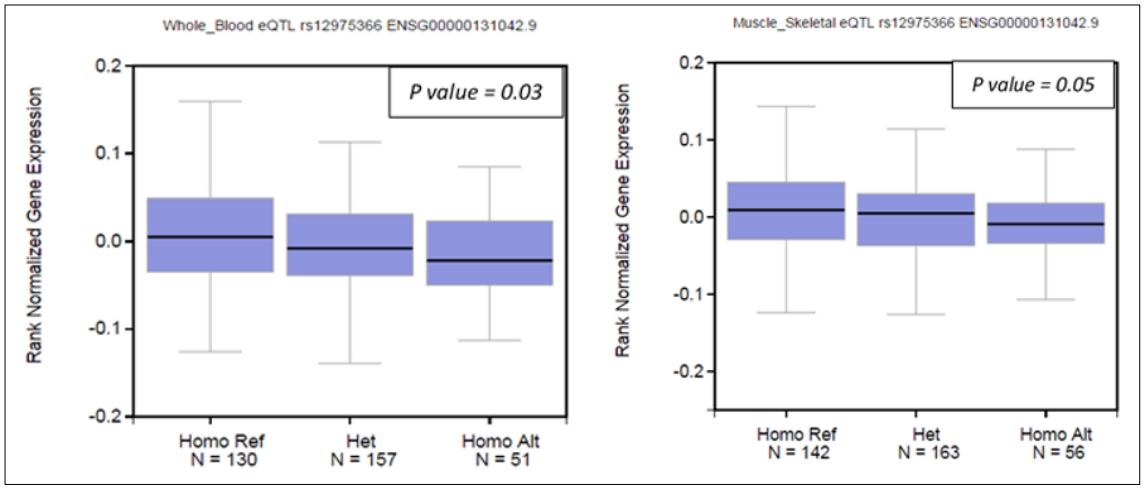


Figure 6.8 eQTL of *LILRB5* variant rs12975366 (Asp247Gly) on *LILRB2* in whole blood (on left) and skeletal muscle (right) (Box plots generated on GTEx portal)

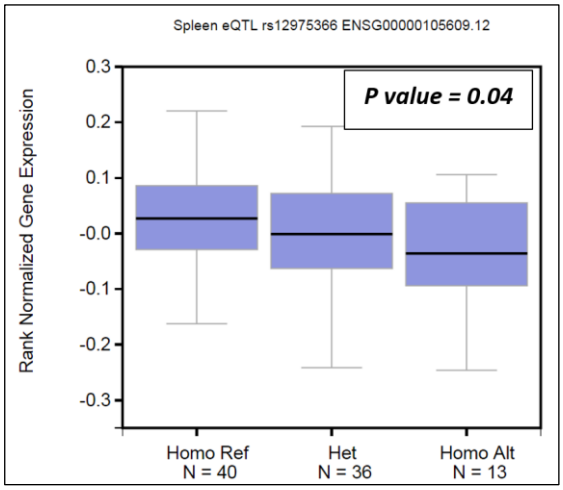


Figure 6.9 eQTL of *LILRB5* variant rs12975366 (Asp247Gly) on *LILRB5* expression in the spleen (Box plots generated on GTEx portal)

6.3.1.10 *Cis*-eQTL effects of *LILRB2* variant on LILRB2 and *LILRA3* gene expression

The His20Arg variant was a significant eQTL for *LILRB2* and *LILRA3* in whole blood and skeletal muscle cells. In whole blood, carriers of the 20Arg variant had an increased expression of *LILRB2* ($\beta = 0.28$, $p = 3 \times 10^{-6}$) (*see* Figure 6.10) and a decreased expression of *LILRA3* ($\beta = -0.93$, $p = 4 \times 10^{-28}$) (*see* Figure 6.11). Similarly, in skeletal muscles carriers of the 20Arg variant had increased expression of *LILRB2* ($\beta = 0.23$, $p = 0.003$) (*see* Figure 6.10) and a decreased expression of *LILRA3* ($\beta = -0.36$, $p = 1 \times 10^{-4}$ (*see* Figure 6.11).

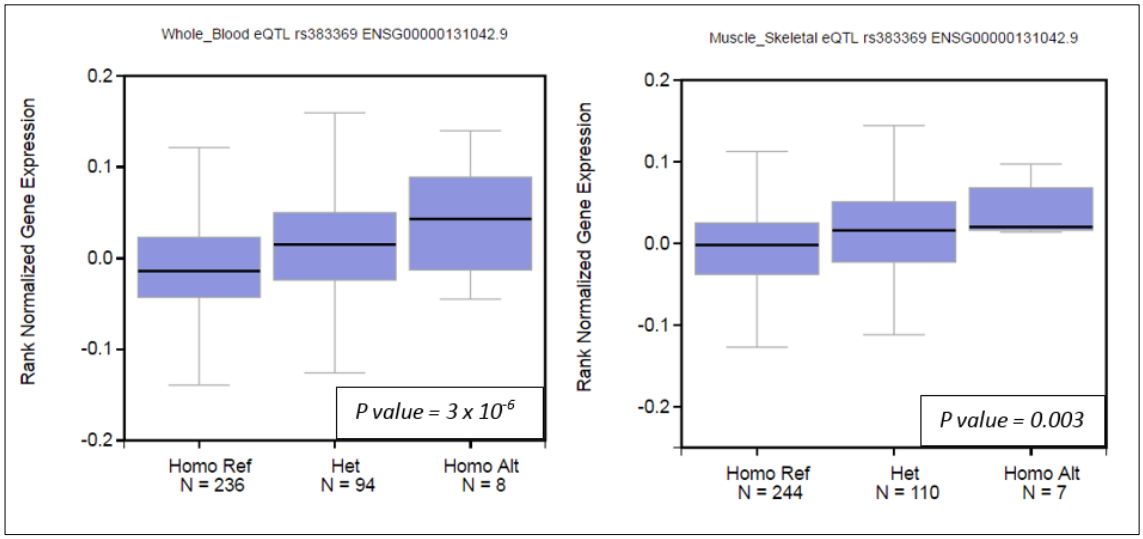


Figure 6.10 eQTL of *LILRB2* variant rs383369 (His20Arg) on *LILRB2* in whole blood (on left) and skeletal muscle (right) (Box plots generated on GTEx portal)

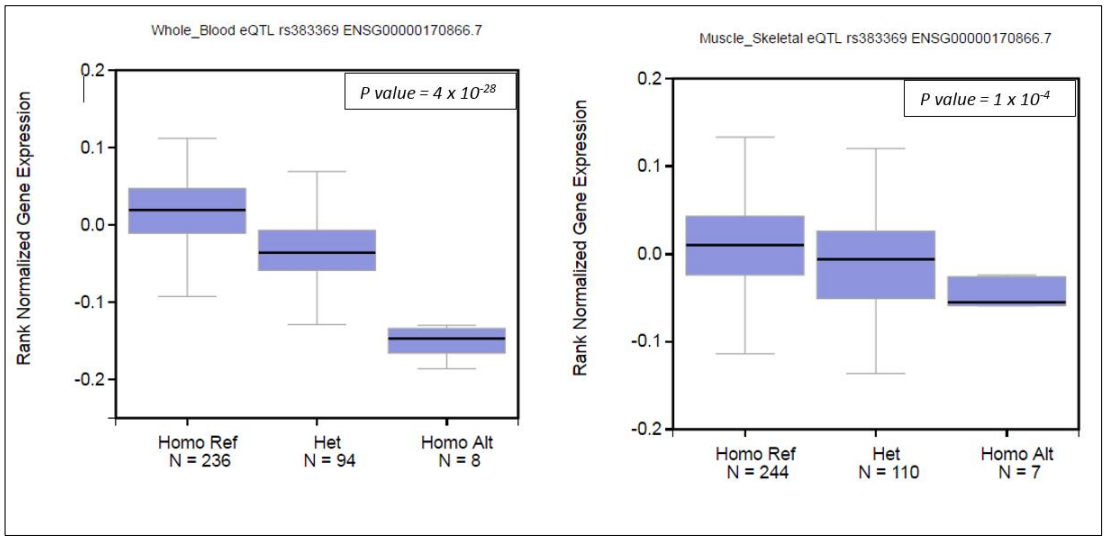


Figure 6.11 eQTL of *LILRB2* variant rs383369 (His20Arg) on *LILRA3* in whole blood (on left) and skeletal muscle (right) (Box plots generated on GTEx portal)

In summary, those with the risky Asp247 genotype in *LILRB5* would express more LILRB2, as would those carrying the deleterious His20Arg variant in *LILRB2*.

6.3.1.11 *Trans*-eQTL effects of *LILRB2* variant

LILRB2 is known to interact with certain transfectants of *HLA-A*, *-B*, *-C*, *-F* and *–G*. Due to their highly interlinked roles in the immune system, we checked if the 20Arg variant was a trans-eQTL for the expression of any of these *HLA* sub-types.

As seen in Figure 6.12, His20Arg is an eQTL for the expression of *HLA-C* in both spleen cells ($\beta = -0.54$, $p = 0.008$) and in whole blood cells ($\beta = -0.21$, $p = 0.02$), where each copy of the rare allele appears to decrease the expression of *HLA-C*. The variant appears to have the inverse effect on expression of *HLA-G* in spleen cells ($\beta = 0.44$, $p = 0.035$) and in whole blood cells ($\beta = 0.19$, $p = 0.073$), the rare allele appears to increase the expression of *HLA-G*.

It is worth noting that the eQTL effects are identical for the *LILRB2* Val235Met variant, which was in strong LD with His20Arg, indicating that the haplotype block is crucial in the expression of *LILRA3, LILRB2, HLA-C* and *HLA-G*.

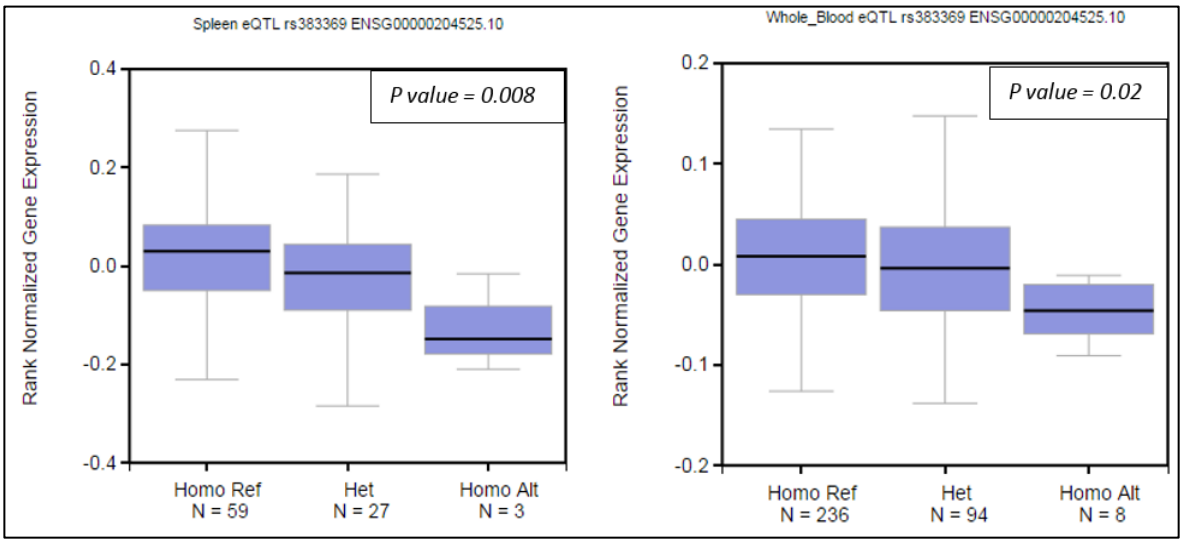


Figure 6.12 eQTL of *LILRB2* variant rs383369 on *HLA-C* in the spleen (on left) and whole blood (right) (Box plots generated on GTEx portal)

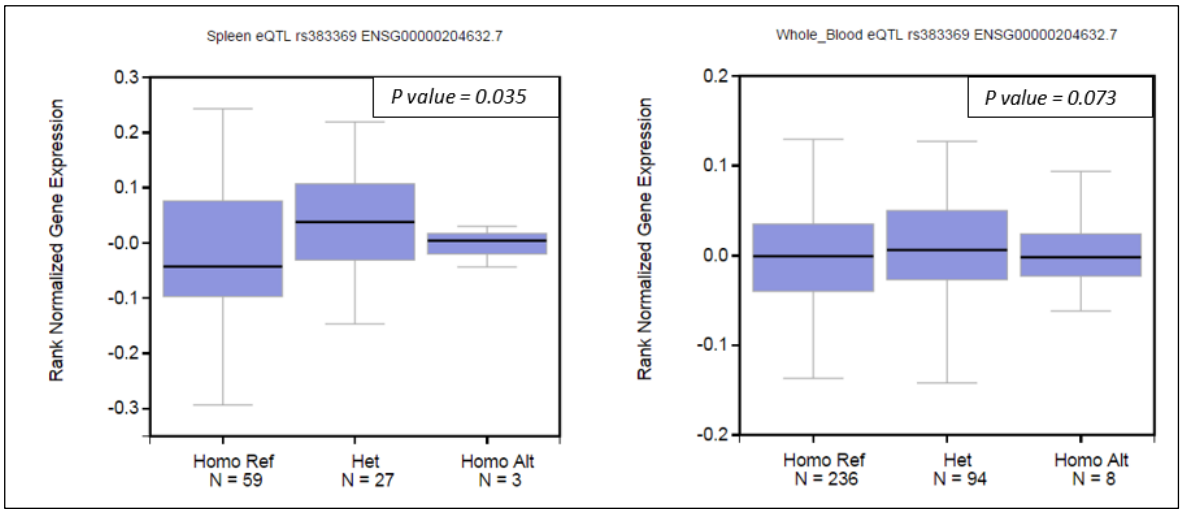


Figure 6.13 eQTL of *LILRB2* variant rs383369 on *HLA-G* in the spleen (on left) and whole blood (right) (Box plots generated on GTEx portal)

The *LILRB5* variant did not show an eQTL effect for the expression of any *HLA*-class I subtypes.

Linkage between the *LILRA3* deletion polymorphism and variants in *LILRB2*

Based on the strong association of the *LILRB2* (His20Arg and Val235Met) variants on reduced *LILRA3* expression, we wanted to examine if this was due to linkage with the deletion polymorphism in *LILRA3,* and to test whether reduction in *LILRA3* was associated with the development of statin intolerance.

The deletion was successfully characterized from 150 exome sequenced samples. The *LILRB2* Val235Met variant was successfully extracted from the data and used as a proxy for the haplotype containing both *LILRB2* variants of interest, since it had better coverage. The deletion polymorphism was observed as homozygous in 10 individuals and heterozygous in 80. There were 8 homozygous carriers of the *LILRB2* variant and 64 heterozygous carriers. The correlation between the *LILRB2* Val235Met variant and deletion polymorphism was 50% (p value < 0.0001) and the $r^2 = 0.25$.

The occurrence of the deletion polymorphism is cross-tabulated with carriers of the variant in *LILRB2* Val235Met. The calculated D' was 0.5.

Table 6.11 Cross-tabulation of *LILRA3* deletion polymorphism and *LILRB2* variant

***LILRA3* deletion**	***LILRB2* Val235Met**	
	Val235	**235Met**
Carriers of deletion	35	55
Non-carriers (homozygous ancestral allele)	68	17

The *LILRB2* variant and *LIRA3* deletion polymorphism do not show genetically significant linkage and the D' indicates that there is significant recombination occurring between them.

However, there was no significant association between the deletion polymorphism or the *LILRB2* Val235 variant and SIM in the PREDICTION-ADR cohort tested. However, we are underpowered to determine an association with SIM in this analysis (150).

6.4 Discussion

These analyses were undertaken to probe the role of the the *LILR* family in statin intolerance, which is a novel hypothesis. The *LILR* gene family belongs to the immunoglobulin (Ig) superfamily that plays an essential roles as antigen receptors, costimulatory proteins, adhesion molecules and immunomodulatory receptors. The *LILR* are involved in both the innate and adaptive arms of the immune system.

This study concludes that an association exists between SNPs in LD in *LILRB2* and statin intolerance represented by the index SNP, His20Arg in *LILRB2*. The effect of this variant was modulated by the Asp247Gly variant in *LILRB5*, which was independently reported to be associated with intolerance and myalgia in Chapter 5. Those with the risky genotype in *LILRB5* (Asp247) showed the deleterious effects of the *LILRB2* His20 variant on raised CK and non-adherence to statin treatment, and to non-adherence to treatment irrespective of CK. The association was replicated in the JUPITER trial where those with the *LILRB5* Asp247 genotype and *LILRB2* 20Arg variant were more likely to have raised CK after commencement of therapy, which is a traditional statin intolerance diagnosis paradigm.

The lower percent of variation explained in the *LILRB5* Asp247 (T/T) genotype group for statin intolerance seen in the fully adjusted model in Table 6.8 might also point to the fact that there are perhaps other factors that render an individual unable to tolerate their statin therapy that have not been captured in our analysis. Due to the stratified analyses, we were underpowered to include co-medications which we know are a significant, but independent risk factor as they were included as a covariables in the original meta-analysis.

The MAF in GoDARTS was 0.21 and 0.18 in the JUPITER trial, in the 1000Genomes for Caucasian population the MAF was 0.17 indicating these results are likely to be extrapolatable to Caucasian populations in the western world. A drawback of the study is

lack of overlap in case control definitions across the Dundee-Liverpool cohorts. The relatively relaxed definition of statin tolerance for the Liverpool cohort might have blunted the possible effects of the variants. Sustained exposure to medium to high dose of statin therapy is what is understood to be a risk factor for the development of an adverse reaction (97,154) .

Based on eQTL analyses it appears that *LILRB2* expression could be related to statin intolerance (SI) and myalgia. *LILRB5* Asp247Gly variants decrease *LILRB2* expression, however, since variant carriers are protected from SI and myalgia (as seen in Chapter 5), higher expression of *LILRB2* is associated with both outcomes. Furthermore, carriers of the *LILRB2* His20Arg variant also express more *LILRB2*, and since carriers are also more likely to develop SI and myalgia, we can support our previous hypothesis that *LILRB2* expression is associated with both outcomes.

In the receptor, the *LILRB2 v*ariant, His20Arg encodes the third domain of the extracellular domain (162) and since these are the sites of interaction with the HLA-class 1 molecules this points to a potential role in the adaptive arm of the immune system (116). The eQTL effect on *HLA-C* and *–G* expression by the variant in *LILRB2* suggests a more closely connected association and one worthy of further exploration in gene expression and response studies. However, trans-eQTL effects could be prone to type 1 error, and must be examined in a larger population than presented in these results.

In our eQTL analyses we find that the *LILRB2* His20Arg variant carriers have lower *LILRA3* expression. A study by Hirayasu *et al.* showed that the His20Arg variant is in strong LD with the deletion polymorphism (D' = 0.96) in Northeast Asians (162). Our characterization of the deletion polymorphism in *LILRA3* in the Caucasian population is novel. We demonstrate that in the Caucasian population, the variant in *LILRB2* is not linked

with the *LILRA3* deletion polymorphism. This provides further evidence for population selection in this region. However, we are insufficiently powered to test the association between the deletion polymorphism and statin-induced myositis. The lower *LILRA3* expression seen in those at risk of developing statin intolerance (*LILRB2* variant carriers) could be an independent effect of the gene. Identifying a variant that is in strong linkage with the deletion (from the exome sequenced samples) could help answer this question in genotyped populations, such as GoDARTS.

This study presents a statin intolerance susceptibility marker in the immune system and suggests that the development of a gene risk score for the *LILR* region, initially perhaps predicting *LILRB2* expression levels, would help categorize risk levels for muscle-associated susceptibility. The selection of this index SNP will also simplify future replication studies, as it did in the JUPITER trial.

7 General Discussion

7.1 Summary of findings:

Perspectives on statin intolerance

This thesis presents findings contrasting two central concepts of statin intolerance. Symptoms of statin intolerance, specifically myalgia (resulting from muscle damage), has been reported to occur in between 7% and 29% of statin users (97) and is thought to be caused by statin use. Furthermore, a diagnosis of statin-induced muscle damage relied heavily on changes in levels of a biomarker of muscle breakdown – creatine phosphokinase (CK). Studies present various cut-offs for CK elevations, to reflect the severity of statin-induced muscle damage. At the same time, there are numerous reports of a lack of extreme CK elevations in statin users who complain of myalgia, and whose symptoms resolve upon discontinuation of therapy.

First, we provide evidence to suggest that there might be genetic factors that determine an individual's constitutive and inducible CK levels. The Glu83Gly variant in *CKM* had previously been shown to be associated with CK levels under normal conditions. We show that this variant impairs an individual's ability to induce CK levels in response to tissue trauma. This finding provides an example of how the phenomenon of myalgia with no commensurate CK elevations may occur. The stratification of CK response by genotype calls into question the use of CK as the end-all be-all biomarker of statin-induced muscle damage.

Second, we provide evidence to suggest that there might be one or more genetic factors that increase an individual's predisposition to myalgia, irrespective of statin use. We observe carriers of the 247Gly variant in *LILRB5* are much better able to tolerate their statin

therapy, carriers of the same variant are also shown to have lower baseline CK levels and other studies have shown they have lower lactose dehydrogenase (LDH) levels as well. LDH is an enzyme found ubiquitously in cells, and used predominantly to help guide diagnoses of tissue (including muscle) damage. This evidence suggests that carriers of the *LILRB5* 247Gly variant might have improved capacity to repair muscle, irrespective of statin exposure. The modulation of this effect by statin use was seen clearly is the RCT setting, where we saw evidence congruent with our findings; carriers of the variant were less likely to have myalgia in general, representing a baseline susceptibility for non-carriers. However, the true statin-induced effect was seen only amongst carriers of the putatively protective variant. These observations coincide with the general observation of similar levels of reported myalgia in placebo and statin treated arms of the published clinical trials. This suggests that the trends of intolerance seen in populations unable to tolerate their statin therapy are not entirely statin-specific, but could be due to a genetic susceptibility to myalgia.

Randomized clinical trials including JUPITER and the GAUSS-3 trial have reported clear instances of non-specific myalgia. In the JUPITER trial, 48% of reported myalgia occurred in participants receiving placebo (114,115). In the run-in to the GAUSS-3 trial, designed to select statin-specific myalgia, 37% of the reported instances were non-specific (154). Therefore, if those with non-specific myalgia were placed on statin therapy, they would likely complain of muscle-based symptoms, which might be exacerbated by the structural disruption occurring in myocytes due to statin therapy (38). These individuals are then at a much higher risk of being non-adherent to their statin therapy, which in-turn raises their risk of suffering from CVD-related event.

Mechanisms of intolerance mediated by the innate immune system

Statins not only reduce circulating LDL but also increase HDL (200,201). A SNP in *LILRA3*, (rs386000) was reported in our meta-analysis for the association between the *LILR* gene family and statin intolerance (*see* Chapter 6). This variant was also reported to be associated with plasma HDL levels (0.83 m/dL increase in HDL for variant carriers, p value = 4×10^{-16}) (178). The variant was also is a strong eQTL for decreased *LILRA3* expression and increased *LILRB2* expression (197–199).These expression profiles are consistent with those noted in association with intolerance. The variant in *LILRA3* has an r^2 of 0.75, and D' of 0.92 with the the index SNP in our analysis of statin intolerance. While the association with the *LILRA3* variant might have been driven through its linkage with the index SNP, it does suggest that the haplotype could be related to HDL levels.

HDL plays an active role in the innate immune system by reducing inflammation. HDL's ability to modulate the bioavailability of cholesterol in lipid rafts in the plasma membranes of cells is evolutionarily conserved and affects the properties of cells involved in both innate and adaptive immune responses; tuning inflammatory responses and antigen presentation functions of macrophages, as well as B and T cell activation. B and T cell receptors are located in these lipid rafts, and their activity is modulated by changes in lipid raft composition and structure. Higher HDL levels would transport more cholesterol out of the lipid rafts. Therefore, the modulation of HDL by statins and *LILRA3* could be responsible for the immunomodulatory effects that result in statin intolerance. Interestingly, *LILR* gene expressions are upregulated in peripheral blood leukocytes after endotoxin challenge (160). Again, highlighting the potentially interlinked roles of the gene family and HDL in the innate immune system.

The well-known interaction between lipids and inflammation might also help shed some light on the specific types of immune reactions occurring in those who are intolerant to

statin treatment. It might also help explain the increased risk of CVD-events for statin intolerant individuals as seen in Chapter 3. A study by Sokolowska *et al.* showed that hyaluronan (HA) which is an essential component of extracellular matrix is broken down during inflammation (202). Fragmented HA polarizes macrophages toward a proinflammatory MI phenotype with a unique eicosanoid profile (202). This mechanism could help explain a proportion of the increased risk of CVD-events for those intolerant to their statins seen in Chapter 3.

However, while the mechanism by which these variants modulate HDL levels is not known, it is apparent that the association is being driven by a haplotype that contains two missense variants in the extracellular domain of the *LILRB2* receptor and also is a high expressing haplotype of *LILRB2* mRNA in several tissues. It is unclear if all of these features are required for the observed phenotypes due to the tight linkage in this region. The amino acid substitutions could affect the binding affinity for HLA ligands, and indeed we observe that the variants are *trans*-eQTLs for *HLA-C* and *HLA-G* gene expression as well. A variant affecting both a gene in *cis* and in trans suggests that the *cis* gene is functionally linked to the trans gene expression (203). Therefore, it seems plausible that there is a strong functional relationship between *LILRB2* and the proteins encoded by the *HLA* subtypes. *LILRB2* receptors are present on monocytes which are adaptive and can differentiate into macrophages and dendritic cells (DC). When activated DC can generate mediators of the innate immune system, which leads to the expression of co-stimulatory proteins that in-turn induce an adaptive immune response. The versatility of monocytes places the impact of the *LILRB2* variants in either or both arms of the immune system.

Mechanisms of intolerance mediated by the adaptive immune system

The genetic variant reported in Chapter 5 *LILRB5* Asp247Gly is involved in the immune system rather than in statin metabolism or statin pharmacokinetics. This makes the conclusion that the drug is not central to the observed phenotype more valid. Since the variant is not an eQTL for *LILRB5* gene expression, its effect most likely arises from the amino-acid substitution occurring in the extracellular domain. HLA class I heavy chains are reported ligands for *LILRB5* (156). It may be that the *LILRB5* Asp247Gly variant is modulating muscle pain in response to a wide range of common environmental agents, or may be involved in the general day to day repair homeostasis of muscles, perhaps in conjunction with the observed role of T regulator (Treg) cells.

Kuswanto *et al.* report that the presence and rapid accumulation of Treg cells in healthy and young skeletal muscles is associated with swifter repair and regeneration. This feature is lost with age. Treg cells are regulated by FOXP3, ST2 (the IL-33 receptor). Sphingosine-1-phosphate receptor 1 (S1P1) expression on T cells determines the migration of lymphocytes from lymphoid organs (204,205), and can lead to poor accumulation of Treg cells, even if the splenic populations are normal (206) (*see* Figure 7.1).

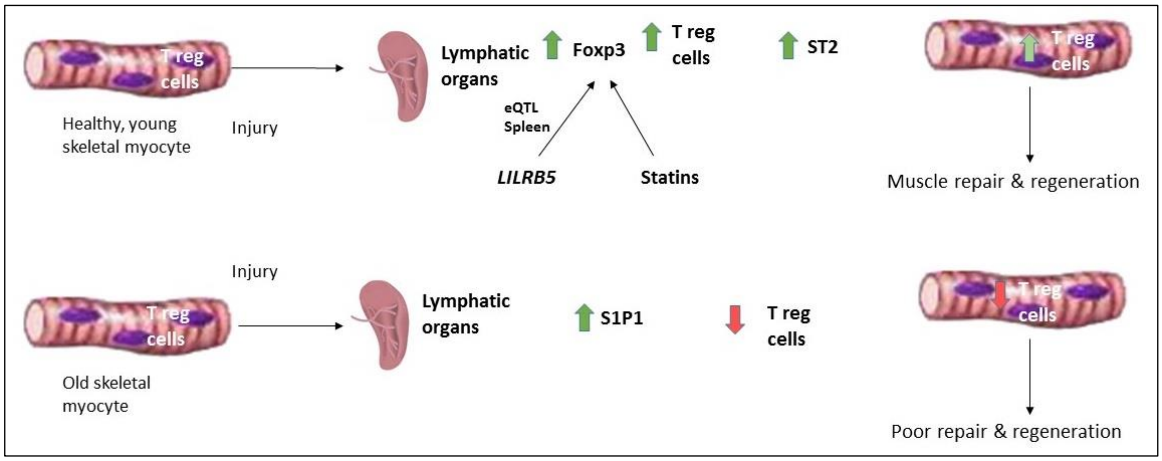


Figure 7.1 Outline of factors influencing T reg cell accumulation in skeletal muscle cells. Statins and *LILRB5* modulate Foxp3 expression.

T reg: T regulator, Foxp3: Forkhead box P3, ST2: Suppression of Tumorigenicity (IL-33 receptor), IL-33: Interleukin-33, S1P1: Sphingosine-1-phosphate receptor

Rodriguez-Perea *et al.* report that statins increase both the number and suppressive function of $CD4^+$ $Foxp3^+$ Treg cells in inflammatory conditions (*see* Figure 7.1) (182). They conducted the study on normocholesterolemic male subjects who were placed on statin therapy for 45 days; and found that after an initial increase in Treg and Foxp3 mRNA levels, Treg cells numbers returned to normal. However, Foxp3 expression levels remained high.

Treg cells are characterized by their expression of Foxp3, a transcription factor that is the master-regulator of Treg immune-suppressive activity. The mechanisms that induce Foxp3 expression and sustain Treg activity are therefore of great interest. The clear importance of Foxp3 in T reg cell function and muscle homeostasis, and the apparent association between myalgia and the *LILRB5* variant, led us to test the eQTL effect of the Asp247 variant in *LILRB5* on Foxp3 expression in the spleen. The Asp247 variant in *LILRB5* is a *cis*-eQTL for *LILRB5* expression (*see* Chapter 6) and a *trans*-eQTL for Foxp3 expression ($\beta = 0.36$, p value = 0.009) in the spleen. This eQTL effect localized in the spleen suggests that perhaps the roles of the receptor and transcription factor are interlinked.

IL-33 receptor (ST2) expression was found to be crucial in Treg cells accumulation post-injury (206). Interestingly, IL-33 is produced mainly by fibro-adipogenic progenitor cells, which are found abundantly in regions surrounding muscles and are often associated with nerve structures. This has led to a hypothesized role for IL-33 in pain perception via the relaying of signals between nervous and immune systems in muscle (206), which could be linked with the inter-individual variation in myalgia-associated pain.

Treg frequency was found to be positively correlated with plasma HDL levels, suggesting a possible role of HDL cholesterol in T reg homeostasis (182) and a common thread through the various etiologies of statin intolerance – HDL.

The pleiotropic effects of statins on human physiology are of increasing interest; reports of their immunomodulatory effects (182,207) come on the heels of reports of the association between statins and new-onset diabetes (138,208,209). It has been reported recently that fat-resident Treg cells accumulate in adipose tissue with age, and their selective depletion increases adipose tissue insulin sensitivity (210). The Treg cell-mediated failure to preserve an optimal immune state in the aged adipose tissue may contribute directly to metabolic disorders such as insulin resistance and age-associated diabetes. This non-obesity dependent, Treg cell-driven metabolic disease of the elderly has been designated Type IV diabetes, and it is plausible that the immunomodulatory role of statins in increasing Treg cells frequency might be causative (208).

Therefore, while Treg cells might be beneficial in regulating muscle regeneration and recovery, and tempering immune reactions, they can also interfere with immune response.

7.2 Comments on study design:

Causal inference is a major aim of epidemiological studies. Therefore, in order to assess the validity of the findings presented in this thesis, we examined factors that compromise causal inference (111).

- Construct validity: This arises when the exposure and outcomes measured do not reflect the intended exposure and outcomes. These are addressed in Chapter 3. Measurement of the exposure, which is statin use is robust and relies on repeated encashment of prescriptions. Errors in the encashment data are possible, but unlikely to be systematic. A second factor that could be considered an "exposure" is genotype. Individuals were genotyped on widely used chip-based arrays, TAQMAN genotyping, exome sequencing, or imputation. These techniques are widely used. Further evidence for consistent genotyping or imputation in our datasets arises from the similarity of MAFs seen in the JUPITER trial population and on the NCBI data base. Finally, characterisation of the

LILRA3 deletion polymorphism was performed using a novel method to determine deletions in gene regions using sequencing output. The limitations and necessary modifications to improve this methodology will become apparent with more experimentation. Outcomes include CK test results and changes in prescribing patterns that are suggestive of intolerance. CK assay results are produced with kits that are used universally; the results are used in the Tayside and Fife area for the treatment and diagnoses of patients. They are all measured centrally and are therefore unlikely to suffer from site-based drift in results. Classical definitions of statin-induced myopathy or myositis include CK elevations 4 or 10 times above the normal. However, while these elevations represent an extreme reaction to the drug, they are not normally observed in statin users who complain of muscle aches, a fact noted by Banach *et al*. when detailing the limitations of myalgia classifications (133). Our definitions of intolerance (switching or discontinuation) are robust, in that they predict a pattern that results in poor outcomes for statin users, as seen in Chapter 3 (validation against statin failure) and are patterns seen more commonly amongst statin users. The inability to tolerate more than two or three therapies is considered sufficient criteria even for recruitment into clinical trials (154). Studying the genetic underpinnings of such traits is crucial to improving patient compliance and adherence to long-term statin therapy.

- Statistical validity: This is generally driven by a lack of power in the analysis. Due to the size of the GoDARTS cohort this is generally not a problem, but for outcomes that were rare, the results of power calculations are provided in the relevant results sections. The expected frequency of outcomes was determined from the literature. The analyses presented in this thesis have employed the appropriate tests to measure the outcomes of interest, and the interpretation of data has been congruent to the test performed.
- Internal validity: This stems from an information or selection bias, or the occurrence of unmeasured or untested confounding factors. This threat to the causal inference of a study

can be mitigated by evaluating the effects of systematic errors in the design, conduct and analysis of a study. Limitations of intolerance phenotypes dependent on raised CK could lead to false positives, as CK can be raised in a variety of non-clinical conditions such as exercise, or prolonged sitting. The switching or discontinuation of statins could be for reasons other than statin-induced intolerance. Perhaps those discontinuing therapy are less careful about their health, or these could be intrinsic effects, such as a susceptibility to myalgia that is independent of statin use. This would appear to be consistent with reports of myalgia in patients treated with placebo in the JUPITER trial. However, in the absence of patient-reported or physician-documented myalgia, using prescribing changes in conjunction with changes in biomarkers is the closest approximation of statin intolerance and has been used successfully to find variants in efflux transporter encoding *SLCO1B1* in the GoDARTS study (54) and in the CPRD (187).

- External validity: This evaluates what populations, geographic locations etc. these findings apply to. With replication from the American Caucasian cohort (JUPITER), the British Caucasia cohort (CPRD and WTCCC) and the combined Dutch, English, Swedish and Scottish Caucasian populations in the PREDICTION-ADR consortium, in addition to the Scottish Caucasian population in the GoDARTS study, our findings are applicable to the Western Caucasian population. Since we do not observe any environmental factors predisposing to statin intolerance or myalgia, we have no reason to expect these genotypes effects will not follow the same dosage-associated risk in other ethnic groups as well.

7.3 Future directions

The impact of the *CKM* Glu83Gly variant on CK brings into focus the population selection that has occurred, resulting in the MAF in the Caucasian population being ten times lower than Kenyan Masai population's. In sustained exercise the transition from fast twitch muscle to slow-twitch is crucial to maintain energy homeostasis. Population studies

using common variants in *CKM* have reported no association with elite endurance performance, however these variants have no known association with CK levels. CK plays an important role in this transition, a role that has been explored in the past (211). Studies have found that over more prolonged exercise, CK might contribute to the fatigue process by increasing myoplasmic concentration of inorganic phosphate (212). The low levels of CK produced as a result of the *CKM* 83Gly genotype seen widely in the Masai population might explain their enhanced ability to withstand prolonged exercise such as marathon running. The mechanism by which the compensation occurs in the body is worth exploring, perhaps with the aid of metabolomic studies.

The question raised by this thesis, regarding what specific immunogenetic factors are associated with statin-induced myalgia and how they interact with ligands in the *HLA* to produce an immune response might best be answered on a population scale in large cohort studies by running well-powered conditional analyses using associated variants in interacting both *HLA* class I and II ligands. It is plausible that some of the genetic factors in HLA D, that predispose to auto-immune mediated myopathies might be involved in the etiology of general statin intolerance, albeit to a lesser extent.

Contrasting the metabolic by-products of those who develop intolerance to those able to tolerate statins therapy might shed light on the changes that occur in those who become intolerant. A lipidomic profile (e.g. eicosanoids, sphingolipids etc.) might be especially helpful in understanding the role of lipids in inflammation-mediated statin intolerance and intolerance-related adverse effects. Perhaps this would help clarify the etiologies of statin-intolerance, as they appear to straddle both the innate and adaptive arms of the immune system.

The challenge remains untangling statin-induced and statin-independent muscle pain. A question that can best be answered in a RCT with Mendelian randomization of genotype. Perhaps, genotyping in a trial such as GAUSS-3 might reveal if those with non-specific myalgia also have the Asp247 genotype in *LILRB5*. With the localization of IL-33 receptors in the nerves, examining NSAIDs or muscle relaxants use by *LILRB5*/*B2* genotypes might help answer if those with risky genotypes are actually experiencing more muscle pain and being medicated for it in the ambulatory setting.

Finally, recruitment by genotype might allow investigators to more specifically contrast such outcomes as time-to-development of muscle-based symptoms, metabolites found in serum, changes in muscle biopsy results, changes in lipid profiles. Studying muscle biopsies raises its own challenges; due to the invasive nature of the procedure and the lack of concrete proof that muscle biopsies provide "best evidence" of statin-induced damage, makes it harder to justify as a study design. Furthermore, since we are not yet certain what specific changes might occur in smooth muscle cells, it might be advisable to wait until more specific information exists about muscle cell transformation by *LILRB5/B2* genotype before muscle biopsies are sought to understand the impact of statin use on myocytes.

7.4 Clinical impact of the findings

Based on our findings with the *CKM* variant, we would expect that approximately 4% of individuals might not show expected CK elevations with tissue damage. While this is a rare genotype, and CK diagnoses rarely affect the outcomes of acute therapy, it is a clear effect and statin users complaining of myalgia, but showing no CK elevations might carry this genotype.

The British Heart Foundation states that over half of all adults in the United Kingdom have LDL cholesterol raised above normal. With revised National Institute for Health and

Care Excellence (NICE) guidelines an estimated 12 million people will be prescribed statins by 2020 (16,17). The findings of this thesis suggest between 5 - 7% of statin users will not adhere to their therapy, that is, they will be switched across various statin types, or discontinue therapy in spite of medical necessity. Based on the survival analyses performed in Chapter 3, where non-adherent individuals showed significantly increased risk of CVD-events, it appears that the healthcare and economic burden of those unable to sustain statin therapy will be enormous.

This thesis provides insights into the reasons for their inability to adhere to therapy. Further, we provide evidence to suggest that the main side effect of statin treatment might not be solely statin-mediated, but instead be mediated by genetic factors in the immune system. The MAF of the *LILRB5* Asp247Gly variant, associated with myalgia is 0.37, meaning that approximately 60 out of every 100 people carries the protective variant, that has a dominant effect. Therefore, for the remaining 40% are at risk of having muscle-based symptoms. The symptoms are unrelated to statin-type but are dependent on dose, which can be modulated for those at risk.

We present possible genetic variants that might be interacting or acting independently to predispose to myalgia. With the involvement of immune cells in the maintenance and repair of skeletal muscles, the role of the *LILRB5* and *LILRB2* variants in statin-induced myalgia is more plausible. This thesis presents the first reports of variants in the immune system predisposing to non-autoimmune-mediated statin intolerance and myalgia.

Bibliography

1. WHO| Cardiovascular diseases (CVD) [Internet]. 2016 [cited 2016 Jul 18]. Available from: http://www.who.int/mediacentre/factsheets/fs317/en/
2. National Health Service Coronary Heart Disease [Internet]. 2016 [cited 2016 Jul 18]. Available from: http://www.nhs.uk/conditions/Cardiovascular-disease/Pages/Introduction.aspx
3. Choudhury RP, Fuster V, Fayad ZA. Molecular, cellular and functional imaging of atherothrombosis. Nat Rev Drug Discov. England; 2004 Nov;3(11):913–25.
4. Grundy SM. HMG-CoA reductase inhibitors for treatment of hypercholesterolemia. N Engl J Med [Internet]. 1988;319(1):24–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3288867
5. Schachter M. Chemical, pharmacokinetic and pharmacodynamic properties of statins: an update. Fundam Clin Pharmacol. England; 2005 Feb;19(1):117–25.
6. Endo A, Kuroda M, Tanzawa K. Competitive inhibition of 3-hydroxy-3-methylglutaryl coenzyme A reductase by ML-236A and ML-236B fungal metabolites, having hypocholesterolemic activity. FEBS Lett. Netherlands; 1976 Dec;72(2):323–6.
7. Endo A. The discovery and development of HMG-CoA reductase inhibitors. J Lipid Res. UNITED STATES; 1992 Nov;33(11):1569–82.
8. Endo A, Kuroda M, Tsujita Y. ML-236A, ML-236B, and ML-236C, new inhibitors of cholesterogenesis produced by Penicillium citrinium. J Antibiot (Tokyo). JAPAN; 1976 Dec;29(12):1346–8.
9. Corsini A, Mazzotti M, Raiteri M, Soma MR, Gabbiani G, Fumagalli R, et al. Relationship between mevalonate pathway and arterial myocyte proliferation: in vitro studies with inhibitors of HMG-CoA reductase. Atherosclerosis. IRELAND; 1993 Jun;101(1):117–25.
10. Joint Formulary Committee. British National Formulary [Internet]. London: BMJ Group and Pharmaceutical Press. 2014 [cited 2015 Nov 25]. Available from: http://www.medicinescomplete.com
11. Staffa JA, Chang J, Green L. Cerivastatin and reports of fatal rhabdomyolysis. The New England journal of medicine. United States; 2002. p. 539–40.
12. Furberg CD, Pitt B. Withdrawal of cerivastatin from the world market. Curr Control Trials Cardiovasc Med [Internet]. 2001;2(5):205–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11806796
13. Roberts WC. The rule of 5 and the rule of 7 in lipid-lowering by statin drugs. The American journal of cardiology. UNITED STATES; 1997. p. 106–7.
14. Maron DJ, Fazio S, Linton MF. Current perspectives on statins. Circulation [Internet]. 2000;101(2):207–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10637210
15. O´Neill P, Sussex J. International Comparison of Medicines Usage : Quantitative Analysis. Off Heal Econ. 2014;
16. Sathasivam S. Statin induced myotoxicity. Eur J Intern Med [Internet]. 2012;23(4):317–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22560377
17. Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol [Internet].

2006;97(8A):52C–60C. Available from: http://www.sciencedirect.com/science/article/pii/S0002914905021430
18. OECD. Health at a Glance 2013 OECD INDICATORS [Internet]. 2013. Available from: http://dx.doi.org/10.1787/health_glance-2013-en
19. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart. England; 2008 Jan;94(1):34–9.
20. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet [Internet]. 2005;366(9493):1267–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16214597\nhttp://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)67394-1/fulltext
21. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther. United States; 2012 Oct;92(4):414–7.
22. Canestaro WJ, Austin MA, Thummel KE. Genetic factors affecting statin concentrations and subsequent myopathy: a HuGENet systematic review. Genet Med [Internet]. American College of Medical Genetics and Genomics; 2014; Available from: http://dx.doi.org/10.1038/gim.2014.41
23. Oshiro C, Mangravite L, Klein T, Altman R. PharmGKB very important pharmacogene: SLCO1B1. Pharmacogenet Genomics. United States; 2010 Mar;20(3):211–6.
24. Keskitalo JE, Kurkinen KJ, Neuvoneni PJ, Niemi M. ABCB1 haplotypes differentially affect the pharmacokinetics of the acid and lactone forms of simvastatin and atorvastatin. Clin Pharmacol Ther [Internet]. 2008;84(4):457–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19238649
25. Fiegenbaum M, da Silveira FR, Van der Sand CR, Van der Sand LC, Ferreira ME, Pires RC, et al. The role of common variants of ABCB1, CYP3A4, and CYP3A5 genes in lipid-lowering efficacy and safety of simvastatin treatment. Clin Pharmacol Ther. 2005/12/03. 2005;78(5):551–8.
26. Bottorff M, Hansten P. Long-term safety of hepatic hydroxymethyl glutaryl coenzyme A reductase inhibitors: the role of metabolism-monograph for physicians. Arch Intern Med. UNITED STATES; 2000 Aug;160(15):2273–80.
27. Michalets EL. Update: clinically significant cytochrome P-450 drug interactions. Pharmacotherapy. UNITED STATES; 1998;18(1):84–112.
28. Jacobsen W, Kuhn B, Soldner A, Kirchner G, Sewing KF, Kollman PA, et al. Lactonization is the critical first step in the disposition of the 3-hydroxy-3-methylglutaryl-CoA reductase inhibitor atorvastatin. Drug Metab Dispos. UNITED STATES; 2000 Nov;28(11):1369–78.
29. Vickers S, Duncan CA, Vyas KP, Kari PH, Arison B, Prakash SR, et al. In vitro and in vivo biotransformation of simvastatin, an inhibitor of HMG CoA reductase. Drug Metab Dispos. UNITED STATES; 1990;18(4):476–83.
30. Lennernas H, Fager G. Pharmacodynamics and pharmacokinetics of the HMG-CoA reductase inhibitors. Similarities and differences. Clin Pharmacokinet. NEW

ZEALAND; 1997 May;32(5):403–25.
31. Sica DA, Gehr TWB. Rhabdomyolysis and statin therapy: relevance to the elderly. Am J Geriatr Cardiol. United States; 2002;11(1):48–55.
32. Muscari A, Puddu GM, Puddu P. Lipid-lowering drugs: are adverse effects predictable and reversible? Cardiology. Switzerland; 2002;97(3):115–21.
33. Knopp RH. Drug treatment of lipid disorders. N Engl J Med. UNITED STATES; 1999 Aug;341(7):498–511.
34. Garnett WR. Interactions with hydroxymethylglutaryl-coenzyme A reductase inhibitors. Am J Health Syst Pharm. UNITED STATES; 1995 Aug;52(15):1639–45.
35. Martin PD, Warwick MJ, Dane AL, Hill SJ, Giles PB, Phillips PJ, et al. Metabolism, excretion, and pharmacokinetics of rosuvastatin in healthy adult male volunteers. Clin Ther. United States; 2003 Nov;25(11):2822–35.
36. Wagstaff LR, Mitton MW, Arvik BM, Doraiswamy PM. Statin-associated memory loss: analysis of 60 case reports and review of the literature. Pharmacotherapy. United States; 2003 Jul;23(7):871–80.
37. Cucchiara B, Kasner SE. Use of statins in CNS disorders. J Neurol Sci. Netherlands; 2001 Jun;187(1–2):81–9.
38. Draeger A, Monastyrskaya K, Mohaupt M, Hoppeler H, Savolainen H, Allemann C, et al. Statin therapy induces ultrastructural damage in skeletal muscle in patients without myalgia. J Pathol. England; 2006 Sep;210(1):94–102.
39. Flint OP, Masters BA, Gregg RE, Durham SK. Inhibition of cholesterol synthesis by squalene synthase inhibitors does not induce myotoxicity in vitro. Toxicol Appl Pharmacol. UNITED STATES; 1997 Jul;145(1):91–8.
40. Macaluso M, Russo G, Cinti C, Bazan V, Gebbia N, Russo A. Ras family genes: an interesting link between cell cycle and cancer. J Cell Physiol. United States; 2002 Aug;192(2):125–30.
41. Coleman ML, Marshall CJ, Olson MF. RAS and RHO GTPases in G1-phase cell-cycle regulation. Nat Rev Mol Cell Biol. England; 2004 May;5(5):355–66.
42. Olson MF, Ashworth A, Hall A. An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. Science. UNITED STATES; 1995 Sep;269(5228):1270–2.
43. PD T, P C, RH K. STatin-associated myopathy. JAMA [Internet]. 2003;289(13):1681–90. Available from: +
44. Guijarro C, Blanco-Colio LM, Ortego M, Alonso C, Ortiz A, Plaza JJ, et al. 3-Hydroxy-3-methylglutaryl coenzyme a reductase and isoprenylation inhibitors induce apoptosis of vascular smooth muscle cells in culture. Circ Res. UNITED STATES; 1998 Sep;83(5):490–500.
45. Knapp AC, Huang J, Starling G, Kiener PA. Inhibitors of HMG-CoA reductase sensitize human smooth muscle cells to Fas-ligand and cytokine-induced cell death. Atherosclerosis. IRELAND; 2000 Sep;152(1):217–27.
46. Widegren U, Wretman C, Lionikas A, Hedin G, Henriksson J. Influence of exercise intensity on ERK/MAP kinase signalling in human skeletal muscle. Pflugers Arch. Germany; 2000 Dec;441(2–3):317–22.
47. Wretman C, Lionikas A, Widegren U, Lannergren J, Westerblad H, Henriksson J. Effects of concentric and eccentric contractions on phosphorylation of

MAPK(erk1/2) and MAPK(p38) in isolated rat skeletal muscle. J Physiol. England; 2001 Aug;535(Pt 1):155–64.

48. Gregoor PJHS. Atorvastatin may cause nightmares. BMJ [Internet]. 2006 Apr 20;332(7547):950. Available from: http://www.bmj.com/content/332/7547/950.abstract
49. Lansberg PJ, Mitchel YB, Shapiro D, Kastelein JJP, Altman R, Jerums G, et al. Long-Term Efficacy and Tolerability of Simvastatin in a Large Cohort of Elderly Hypercholesterolemic Patients. Atherosclerosis [Internet]. 1995;116(2):153–62. Available from: <Go to ISI>://WOS:A1995RU33200001
50. van Dam M, Zwart M, de Beer F, Smelt AHM, Prins MH, Trip MD, et al. Long term efficacy and safety of atorvastatin in the treatment of severe type III and combined dyslipidaemia. Heart. 2002;88(3):234–8.
51. Pasternak RC, Smith SC, Bairey-merz CN, Grundy SM, Cleeman JI, Lenfant C, et al. ACC / AHA / NHLBI Clinical Advisory on Statins ACC / AHA / NHLBI Clinical Advisory on the Use and Safety of Statins. J Am Coll Cardiol. 2002;40(3):1024–8.
52. Floyd JS, Heckbert SR, Weiss NS, Carrell DS, Psaty BM. The use of administrative data and natural language processing to estimate the incidence of statin-related rhabdomyolysis. In: Epidemiology and Prevention/Nutrition, Physical Activity and Metabolism 2012 Scientific Sessions. 2012.
53. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy--a genomewide study. N Engl J Med [Internet]. 2008;359(8):789–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18650507\nhttp://www.nejm.org/doi/full/10.1056/NEJMoa0801936
54. Donnelly LA, Doney AS, Tavendale R, Lang CC, Pearson ER, Colhoun HM, 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 [Internet]. 2011;89(2):210–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21178985
55. Ford I, Blauw GJ, Murphy MB, Shepherd J, Cobbe SM, Bollen ELEM, et al. A Prospective Study of Pravastatin in the Elderly at Risk (PROSPER): Screening Experience and Baseline Characteristics. Curr Control Trials Cardiovasc Med. 2002 May;3(1):8.
56. Trompet S, van Vliet P, de Craen AJM, Jolles J, Buckley BM, Murphy MB, et al. Pravastatin and cognitive function in the elderly. Results of the PROSPER study. J Neurol. Germany; 2010 Jan;257(1):85–90.
57. Richardson K, Schoen M, French B, Umscheid CA, Mitchell MD, Arnold SE, et al. Statins and cognitive function: a systematic review. Ann Intern Med. United States; 2013 Nov;159(10):688–97.
58. Roth T, Richardson GR, Sullivan JP, Lee RM, Merlotti L, Roehrs T. Comparative effects of pravastatin and lovastatin on nighttime sleep and daytime performance. Clin Cardiol. UNITED STATES; 1992 Jun;15(6):426–32.
59. Muldoon MF, Ryan CM, Sereika SM, Flory JD, Manuck SB. Randomized trial of the effects of simvastatin on cognitive functioning in hypercholesterolemic adults. Am J Med. United States; 2004 Dec;117(11):823–9.
60. Muldoon MF, Barger SD, Ryan CM, Flory JD, Lehoczky JP, Matthews KA, et al.

Effects of lovastatin on cognitive function and psychological well-being. Am J Med. UNITED STATES; 2000 May;108(7):538–46.

61. Carlsson CM, Gleason CE, Hess TM, Moreland KA, Blazel HM, Koscik RL, et al. Effects of simvastatin on cerebrospinal fluid biomarkers and cognition in middle-aged adults at risk for Alzheimer's disease. J Alzheimers Dis. Netherlands; 2008 Mar;13(2):187–97.
62. Mihaylova B, Emberson J, Blackwell L, Keech A, Simes J, Barnes EH, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. Lancet (London, England). England; 2012 Aug;380(9841):581–90.
63. Bruckert E, Hayem G, Dejager S, Yau C, Begaud B. Mild to moderate muscular symptoms with high-dosage statin therapy in hyperlipidemic patients--the PRIMO study. Cardiovasc Drugs Ther. United States; 2005 Dec;19(6):403–14.
64. Nichols GA, Koro CE. Does statin therapy initiation increase the risk for myopathy? An observational study of 32,225 diabetic and nondiabetic patients. Clin Ther. United States; 2007 Aug;29(8):1761–70.
65. Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol [Internet]. 2006;97(8A):52C–60C. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16581329
66. Alfirevic A, Neely D, Armitage J, Chinoy H, Cooper RG, Laaksonen R, et al. Phenotype Standardization for Statin-Induced Myotoxicity. Clin Pharmacol Ther [Internet]. 2014; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24897241
67. Armitage J. The safety of statins in clinical practice. Lancet [Internet]. 370(9601):1781–90. Available from: http://www.sciencedirect.com/science/article/pii/S0140673607607168
68. Psaty BM, Furberg CD, Ray WA, Weiss NS. Potential for conflict of interest in the evaluation of suspected adverse drug reactions: use of cerivastatin and risk of rhabdomyolysis. JAMA [Internet]. 2004;292(21):2622–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15572720
69. Joy TR, Hegele RA. Narrative review: statin-related myopathy. Ann Intern Med [Internet]. 2009;150(12):858–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19528564
70. Mitka M. Expanding statin use to help more at-risk patients is causing financial heartburn. JAMA [Internet]. 2003;290(17):2243–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14600168
71. Peters BJ, Klungel OH, Visseren FL, de Boer A, Maitland-van der Zee AH. Pharmacogenomic insights into treatment and management of statin-induced myopathy. Genome Med [Internet]. 2009;1(12):120. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20090898
72. Kirchheiner J, Kudlicz D, Meisel C, Bauer S, Meineke I, Roots I, et al. Influence of CYP2C9 polymorphisms on the pharmacokinetics and cholesterol-lowering activity of (-)-3S,5R-fluvastatin and (+)-3R,5S-fluvastatin in healthy volunteers. Clin Pharmacol Ther [Internet]. 2003;74(2):186–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12891229
73. Ozaki H, Ishikawa CT, Ishii T, Toyoda A, Murano T, Miyashita Y, et al. Clearance rates of cerivastatin metabolites in a patient with cerivastatin-induced

rhabdomyolysis. J Clin Pharm Ther [Internet]. 2005;30(2):189–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15811174

74. Ishikawa C, Ozaki H, Nakajima T, Ishii T, Kanai S, Anjo S, et al. A frameshift variant of CYP2C8 was identified in a patient who suffered from rhabdomyolysis after administration of cerivastatin. J Hum Genet [Internet]. 2004;49(10):582–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15365880
75. Wilke RA, Moore JH, Burmester JK. Relative impact of CYP3A genotype and concomitant medication on the severity of atorvastatin-induced muscle damage. Pharmacogenet Genomics [Internet]. 2005;15(6):415–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15900215
76. Frudakis TN, Thomas MJ, Ginjupalli SN, Handelin B, Gabriel R, Gomez HJ. CYP2D6*4 polymorphism is associated with statin-induced muscle effects. Pharmacogenet Genomics [Internet]. 2007;17(9):695–707. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17700359
77. Niemi M. Role of OATP transporters in the disposition of drugs. Pharmacogenomics [Internet]. 2007;8(7):787–802. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18240907
78. Pasanen MK, Backman JT, Neuvonen PJ, Niemi M. Frequencies of single nucleotide polymorphisms and haplotypes of organic anion transporting polypeptide 1B1 SLCO1B1 gene in a Finnish population. Eur J Clin Pharmacol [Internet]. 2006;62(6):409–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16758257
79. Morimoto K, Oishi T, Ueda S, Ueda M, Hosokawa M, Chiba K. A novel variant allele of OATP-C (SLCO1B1) found in a Japanese patient with pravastatin-induced myopathy. Drug Metab Pharmacokinet [Internet]. 2004;19(6):453–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15681900
80. Furihata T, Satoh N, Ohishi T, Ugajin M, Kameyama Y, Morimoto K, et al. Functional analysis of a mutation in the SLCO1B1 gene (c.1628T>G) identified in a Japanese patient with pravastatin-induced myopathy. Pharmacogenomics J [Internet]. 2009;9(3):185–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19238167
81. Fiegenbaum M, da Silveira FR, Van der Sand CR, Van der Sand LC, Ferreira ME, Pires RC, et al. The role of common variants of ABCB1, CYP3A4, and CYP3A5 genes in lipid-lowering efficacy and safety of simvastatin treatment. Clin Pharmacol Ther [Internet]. 2005;78(5):551–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16321621
82. Keskitalo JE, Zolk O, Fromm MF, Kurkinen KJ, Neuvonen PJ, Niemi M. ABCG2 polymorphism markedly affects the pharmacokinetics of atorvastatin and rosuvastatin. Clin Pharmacol Ther. United States; 2009 Aug;86(2):197–203.
83. Keech A, Collins R, MacMahon S, Armitage J, Lawson A, Wallendszus K, et al. Three-year follow-up of the Oxford Cholesterol Study: assessment of the efficacy and safety of simvastatin in preparation for a large mortality study. Eur Hear J [Internet]. 1994;15(2):255–69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8005129
84. Bays H. Ezetimibe. Expert Opin Investig Drugs [Internet]. 2002;11(11):1587–604. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12437505

85. Capuzzi DM, Guyton JR, Morgan JM, Goldberg AC, Kreisberg RA, Brusco OA, et al. Efficacy and safety of an extended-release niacin (Niaspan): a long-term study. Am J Cardiol [Internet]. 1998;82(12A):74U–81U; discussion 85U–86U. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9915666
86. Tolman KG. The liver and lovastatin. Am J Cardiol [Internet]. 2002;89(12):1374–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12062731
87. Carr DF, O'Meara H, Jorgensen AL, Campbell J, Hobbs M, McCann G, et al. SLCO1B1 Genetic Variant Associated With Statin-Induced Myopathy: A Proof-of-Concept Study Using the Clinical Practice Research Datalink. Clin Pharmacol Ther [Internet]. 2013;94(6):695–701. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23942138
88. de Keyser CE, Peters BJM, Becker ML, Visser LE, Uitterlinden AG, Klungel OH, et al. The SLCO1B1 c.521T>C polymorphism is associated with dose decrease or switching during statin therapy in the Rotterdam Study. Pharmacogenet Genomics [Internet]. 2014;24(1):43–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24263182
89. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med [Internet]. 2002;4(2):45–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11882781
90. Peters BJ, Klungel OH, de Boer A, Maitland-van der Zee AH. Genetic determinants of response to statins. Expert Rev Cardiovasc Ther [Internet]. 2009;7(8):977–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19673675
91. Pasanen MK, Neuvonen M, Neuvonen PJ, Niemi M. SLCO1B1 polymorphism markedly affects the pharmacokinetics of simvastatin acid. Pharmacogenet Genomics [Internet]. 2006;16(12):873–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17108811
92. Basharat P, Christopher-Stine L. Immune-Mediated Necrotizing Myopathy: Update on Diagnosis and Management. Curr Rheumatol Rep. United States; 2015 Dec;17(12):72.
93. Basharat P, Lahouti AH, Paik JJ, Albayda J, Pinal-Fernandez I, Bichile T, et al. Statin-Induced Anti-HMGCR-Associated Myopathy. Journal of the American College of Cardiology. United States; 2016. p. 234–5.
94. Mohassel P, Mammen AL. Statin-associated autoimmune myopathy and anti-HMGCR autoantibodies. Muscle Nerve. United States; 2013 Oct;48(4):477–83.
95. Mammen AL, Gaudet D, Brisson D, Christopher-Stine L, Lloyd TE, Leffell MS, et al. Increased frequency of DRB1*11:01 in anti-hydroxymethylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Arthritis Care Res (Hoboken). United States; 2012 Aug;64(8):1233–7.
96. Alshehri A, Choksi R, Bucelli R, Pestronk A. Myopathy with anti-HMGCR antibodies: Perimysium and myofiber pathology. Neurol Neuroimmunol Neuroinflammation [Internet]. Hagerstown, MD: Lippincott Williams & Wilkins; 2015 Aug 4;2(4):e124. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4459046/
97. Stroes ES, Thompson PD, Corsini A, Vladutiu GD, Raal FJ, Ray KK, et al. Statin-associated muscle symptoms: impact on statin therapy-European Atherosclerosis Society Consensus Panel Statement on Assessment, Aetiology and Management.

Eur Heart J. England; 2015 May;36(17):1012–22.
98. Kashani A, Phillips CO, Foody JM, Wang Y, Mangalmurti S, Ko DT, et al. Risks associated with statin therapy: a systematic overview of randomized clinical trials. Circulation. United States; 2006 Dec;114(25):2788–97.
99. Blaier O, Lishner M, Elis A. Managing statin-induced muscle toxicity in a lipid clinic. J Clin Pharm Ther. England; 2011 Jun;36(3):336–41.
100. Bays H. Statin safety: an overview and assessment of the data--2005. Am J Cardiol. United States; 2006 Apr;97(8A):6C–26C.
101. Rannanheimo PK, Tiittanen P, Hartikainen J, Helin-Salmivaara A, Huupponen R, Vahtera J, et al. Impact of Statin Adherence on Cardiovascular Morbidity and All-Cause Mortality in the Primary Prevention of Cardiovascular Disease: A Population-Based Cohort Study in Finland. Value Heal J Int Soc Pharmacoeconomics Outcomes Res. United States; 2015 Sep;18(6):896–905.
102. Maningat P, Gordon BR, Breslow JL. How Do We Improve Patient Compliance and Adherence to Long-Term Statin Therapy? Curr Atheroscler Rep [Internet]. 2013 Jan;15(1):291. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3534845/
103. Dujic T, Zhou K, Donnelly LA, Tavendale R, Palmer CNA, Pearson ER. Association of Organic Cation Transporter 1 With Intolerance to Metformin in Type 2 Diabetes: A GoDARTS Study. Diabetes. United States; 2015 May;64(5):1786–93.
104. Zhou K, Donnelly L, Yang J, Li M, Deshmukh H, Van Zuydam N, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. lancet Diabetes Endocrinol. England; 2014 Jun;2(6):481–7.
105. Zhou K, Donnelly L, Burch L, Tavendale R, Doney AS, Leese G, et al. Loss-of-function CYP2C9 variants improve therapeutic response to sulfonylureas in type 2 diabetes: a Go-DARTS study. Clin Pharmacol Ther [Internet]. 2010;87(1):52–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19794412
106. Pearson ER, Donnelly LA, Kimber C, Whitley A, Doney ASF, McCarthy MI, et al. Variation in TCF7L2 influences therapeutic response to sulfonylureas: a GoDARTs study. Diabetes. United States; 2007 Aug;56(8):2178–82.
107. Dawed AY, Donnelly L, Tavendale R, Carr F, Leese G, Palmer CNA, et al. CYP2C8 and SLCO1B1 Variants and Therapeutic Response to Thiazolidinediones in Patients With Type 2 Diabetes. Diabetes Care. 2016 Jun;
108. Donnelly LA, Palmer CN, Whitley AL, Lang CC, Doney AS, Morris AD, et al. Apolipoprotein E genotypes are associated with lipid-lowering responses to statin treatment in diabetes: a Go-DARTS study. Pharmacogenet Genomics [Internet]. 2008;18(4):279–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18334912
109. Donnelly L a, Doney ASF, Dannfald J, Whitley AL, Lang CC, Morris AD, et al. A paucimorphic variant in the HMG-CoA reductase gene is associated with lipid-lowering response to statin treatment in diabetes: a GoDARTS study. Pharmacogenet Genomics [Internet]. 2008;18(12):1021–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18815589
110. Postmus I, Trompet S, Deshmukh HA, Barnes MR, Li X, Warren HR, et al. Pharmacogenetic meta-analysis of genome-wide association studies of LDL

cholesterol response to statins. Nat Commun [Internet]. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2014 Oct 28;5. Available from: http://dx.doi.org/10.1038/ncomms6068

111. Rothman, K. J., Greenland, S., & Lash TL. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.
112. Pearce N. Classification of epidemiological study designs. Int J Epidemiol [Internet]. 2012 Apr 1;41(2):393–7. Available from: http://ije.oxfordjournals.org/content/41/2/393.short
113. S. S. Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate [Internet]. Hernandez LM, Blazer DG E, editor. Washington (DC): National Academies Press (US); 2006. Available from: http://www.ncbi.nlm.nih.gov/books/NBK19918/
114. Ridker PM, Danielson E, Fonseca FAH, Genest J, Gotto AM, Kastelein JJP, et al. Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein. N Engl J Med [Internet]. Massachusetts Medical Society; 2008 Nov 20;359(21):2195–207. Available from: http://dx.doi.org/10.1056/NEJMoa0807646
115. Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. Circ Cardiovasc Genet. United States; 2012 Apr;5(2):257–64.
116. Brown D, Trowsdale J, Allen R. The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens. Denmark; 2004 Sep;64(3):215–25.
117. Hirayasu K, Arase H. Functional and genetic diversity of leukocyte immunoglobulin-like receptor and implication for disease associations. J Hum Genet [Internet]. The Japan Society of Human Genetics; 2015 Jun 4; Available from: http://dx.doi.org/10.1038/jhg.2015.64
118. Kristjansson RP, Oddsson A, Helgason H, Sveinbjornsson G, Arnadottir GA, Jensson BO, et al. Common and rare variants associating with serum levels of creatine kinase and lactate dehydrogenase. Nat Commun. England; 2016;7:10572.
119. Dube MP, Zetler R, Barhdadi A, Brown AM, Mongrain I, Normand V, et al. CKM and LILRB5 are associated with serum levels of creatine kinase. Circ Cardiovasc Genet [Internet]. 2014;7(6):880–6. Available from: http://circgenetics.ahajournals.org/content/7/6/880.long
120. Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. BMJ. ENGLAND; 1997 Aug;315(7107):524–8.
121. Doney ASF, Dannfald J, Kimber CH, Donnelly LA, Pearson E, Morris AD, et al. The FTO gene is associated with an atherogenic lipid profile and myocardial infarction in patients with type 2 diabetes: a Genetics of Diabetes Audit and Research Study in Tayside Scotland (Go-DARTS) study. Circ Cardiovasc Genet. United States; 2009 Jun;2(3):255–9.
122. International Classification of Disease (ICD) [Internet]. World Health Organization.

1992. Available from: http://apps.who.int/classifications/icd10/browse/2016/en
123. Agresti A. An Introduction to Categorical Data Analysis. 2nd ed. Wiley-Blackwell;
124. Weisberg S. Applied Linear Regression. 4th ed. Inc JW& S, editor. Hoboken, New Jersey; 2014.
125. Hosmer D, Lemeshow S MS. Applied Survival Analysis Regression Modeling of Time-to-event data. 2nd ed. John Wiley & Sons I, editor. 2008.
126. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. United States; 2009 Jun;5(6):e1000529.
127. Delaneau O, Coulonges C, Zagury J-F. Shape-IT: new rapid and accurate algorithm for haplotype inference. BMC Bioinformatics. England; 2008;9:540.
128. Browning BL, Browning SR. Haplotypic analysis of Wellcome Trust Case Control Consortium data. Hum Genet. Germany; 2008 Apr;123(3):273–80.
129. Donnelly L a., Doney a. SF, Morrist a. D, Palmer CN a, Donnan PT. Long-term adherence to statin treatment in diabetes. Diabet Med. 2008;25(7):850–5.
130. Banach M, Rizzo M, Toth PP, Farnier M, Davidson MH, Al-Rasadi K, et al. Statin intolerance - an attempt at a unified definition. Position paper from an International Lipid Expert Panel. Expert Opin Drug Saf. England; 2015 Jun;14(6):935–55.
131. SLCO1B1 Variants and Statin-Induced Myopathy — A Genomewide Study. N Engl J Med [Internet]. 2008;359(8):789–99. Available from: http://www.nejm.org/doi/full/10.1056/NEJMoa0801936
132. Hsia J, MacFadyen JG, Monyak J, Ridker PM. Cardiovascular event reduction and adverse events among subjects attaining low-density lipoprotein cholesterol <50 mg/dl with rosuvastatin: The JUPITER trial (justification for the use of statins in prevention: An intervention trial evaluating rosuvastati. J Am Coll Cardiol. United States; 2011 Apr;57(16):1666–75.
133. Banach M, Rizzo M, Toth PP, Farnier M, Davidson MH, Al-Rasadi K, et al. Statin intolerance - an attempt at a unified definition. Position paper from an International Lipid Expert Panel. Arch Med Sci. Poland; 2015 Mar;11(1):1–23.
134. Bijlsma MJ, Vansteelandt S, Janssen F, Hak E. The effect of adherence to statin therapy on cardiovascular mortality: quantification of unmeasured bias using falsification end-points. BMC Public Health. England; 2016;16(1):303.
135. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA; 2015. Available from: http://www.rstudio.com/
136. R Development Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2008. Available from: http://www.r-project.org
137. Graham DJ, Staffa JA, Shatin D, et al. INcidence of hospitalized rhabdomyolysis in patients treated with lipid-lowering drugs. JAMA [Internet]. 2004;292(21):2585–90. Available from: http://dx.doi.org/10.1001/jama.292.21.2585
138. Sattar N, Preiss D, Murray HM, Welsh P, Buckley BM, de Craen AJM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. Lancet (London, England). England; 2010 Feb;375(9716):735–42.
139. Rajpathak SN, Kumbhani DJ, Crandall J, Barzilai N, Alderman M, Ridker PM. Statin Therapy and Risk of Developing Type 2 Diabetes: A Meta-Analysis. Diabetes Care [Internet]. 2009 Sep 30;32(10):1924–9. Available from:

http://care.diabetesjournals.org/content/32/10/1924.abstract
140. NCBI - creatine kinase [Homo sapiens] [Internet]. [cited 2016 Jan 28]. Available from: http://www.ncbi.nlm.nih.gov/protein/AAA96609.1
141. Wyss M, Kaddurah-Daouk R. Creatine and creatinine metabolism. Physiol Rev. UNITED STATES; 2000 Jul;80(3):1107–213.
142. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PIW. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics [Internet]. Oxford University Press; 2008 Dec 15;24(24):2938–9. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2720775/
143. Wolfgang Viechtbauer. Conducting meta-analyses in R with the metafor package. J Stat Softw [Internet]. 2010;36(3):1--48. Available from: http://www.jstatsoft.org/v36/i03/
144. UniProt: a hub for protein information. Nucleic Acids Res. England; 2015 Jan;43(Database issue):D204-12.
145. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. England; 2014 Jul;42(Web Server issue):W252-8.
146. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Research. Oxford, UK; 2000. p. 45–8.
147. Anaya DA, Dellinger EP. Necrotizing soft-tissue infection: diagnosis and management. Clin Infect Dis. United States; 2007 Mar;44(5):705–10.
148. Wallace B, Siddiqui MK, Palmer CNA, George J. Common Creatine Kinase gene mutation results in falsely reassuring CK levels in muscle disorders. QJM [Internet]. 2015 Nov 14; Available from: http://qjmed.oxfordjournals.org/content/early/2015/11/13/qjmed.hcv215.abstract
149. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature [Internet]. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2012 Nov 1;491(7422):56–65. Available from: http://dx.doi.org/10.1038/nature11632
150. Gauderman W MJ. QUANTO documentation. (Technical report no. 157). Department of Preventive Medicine, University of Southern California; 2001.
151. Aronson JK. Compliance, concordance, adherence. Br J Clin Pharmacol [Internet]. Blackwell Publishing Ltd; 2007 Apr;63(4):383–4. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2203247/
152. Kimoto E, Li R, Scialis RJ, Lai Y, Varma MVS. Hepatic Disposition of Gemfibrozil and Its Major Metabolite Gemfibrozil 1-O-beta-Glucuronide. Mol Pharm. United States; 2015 Nov;12(11):3943–52.
153. Bellosta S, Paoletti R, Corsini A. Safety of statins: focus on clinical pharmacokinetics and drug interactions. Circulation. United States; 2004 Jun;109(23 Suppl 1):III50-7.
154. Nissen SE, Stroes E, Dent-Acosta RE, Rosenson RS, Lehman SJ, Sattar N, et al. Efficacy and Tolerability of Evolocumab vs Ezetimibe in Patients With Muscle-Related Statin Intolerance: The GAUSS-3 Randomized Clinical Trial. JAMA. United States; 2016 Apr;315(15):1580–90.
155. Kuroki K, Tsuchiya N, Shiroishi M, Rasubala L, Yamashita Y, Matsuta K, et al.

Extensive polymorphisms of LILRB1 (ILT2, LIR1) and their association with HLA-DRB1 shared epitope negative rheumatoid arthritis. Hum Mol Genet. England; 2005 Aug;14(16):2469–80.

156. Zhang Z, Hatano H, Shaw J, Olde Nordkamp M, Jiang G, Li D, et al. The Leukocyte Immunoglobulin-Like Receptor Family Member LILRB5 Binds to HLA-Class I Heavy Chains. PLoS One. United States; 2015;10(6):e0129063.
157. Masuda A, Nakamura A, Maeda T, Sakamoto Y, Takai T. Cis binding between inhibitory receptors and MHC class I can regulate mast cell activation. J Exp Med [Internet]. The Rockefeller University Press; 2007 Apr 16;204(4):907–20. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2118540/
158. Deng M, Lu Z, Zheng J, Wan X, Chen X, Hirayasu K, et al. A motif in LILRB2 critical for Angptl2 binding and activation. Blood. United States; 2014 Aug;124(6):924–35.
159. Canavez F, Young NT, Guethlein LA, Rajalingam R, Khakoo SI, Shum BP, et al. Comparison of chimpanzee and human leukocyte Ig-like receptor genes reveals framework and rapidly evolving genes. J Immunol. United States; 2001 Nov;167(10):5786–94.
160. Talwar S, Munson PJ, Barb J, Fiuza C, Cintron AP, Logun C, et al. Gene expression profiles of peripheral blood leukocytes after endotoxin challenge in humans. Physiol Genomics. United States; 2006 Apr;25(2):203–15.
161. Torkar M, Haude A, Milne S, Beck S, Trowsdale J, Wilson MJ. Arrangement of the ILT gene cluster: a common null allele of the ILT6 gene results from a 6.7-kbp deletion. Eur J Immunol. GERMANY; 2000 Dec;30(12):3655–62.
162. Hirayasu K, Ohashi J, Tanaka H, Kashiwase K, Ogawa A, Takanashi M, et al. Evidence for Natural Selection on Leukocyte Immunoglobulin-like Receptors for HLA Class I in Northeast Asians. Am J Hum Genet [Internet]. American Society of Human Genetics; 2008 May 9;82(5):1075–83. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2427302/
163. Wisniewski A, Wagner M, Nowak I, Bilinska M, Pokryszko-Dragan A, Jasek M, et al. 6.7-kbp deletion in LILRA3 (ILT6) gene is associated with later onset of the multiple sclerosis in a Polish population. Hum Immunol. United States; 2013 Mar;74(3):353–7.
164. Lopez-Alvarez MR, Jones DC, Jiang W, Traherne JA, Trowsdale J. Copy number and nucleotide variation of the LILR family of myelomonocytic cell activating and inhibitory receptors. Immunogenetics. United States; 2014 Feb;66(2):73–83.
165. Bashirova AA, Apps R, Vince N, Mochalova Y, Yu XG, Carrington M. Diversity of the human LILRB3/A6 locus encoding a myeloid inhibitory and activating receptor pair. Immunogenetics. United States; 2014 Jan;66(1):1–8.
166. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, et al. Diversity of human copy number variation and multicopy genes. Science. United States; 2010 Oct;330(6004):641–6.
167. Du Y, Cui Y, Liu X, Hu F, Yang Y, Wu X, et al. Contribution of functional LILRA3, but not nonfunctional LILRA3, to sex bias in susceptibility and severity of anti-citrullinated protein antibody-positive rheumatoid arthritis. Arthritis Rheumatol (Hoboken, NJ). United States; 2014 Apr;66(4):822–30.
168. Renauer PA, Saruhan-Direskeneli G, Coit P, Adler A, Aksu K, Keser G, et al.

Identification of Susceptibility Loci in IL6, RPS9/LILRB3, and an Intergenic Locus on Chromosome 21q22 in Takayasu Arteritis in a Genome-Wide Association Study. Arthritis Rheumatol (Hoboken, NJ). United States; 2015 May;67(5):1361–8.

169. Jensen MA, Patterson KC, Kumar AA, Kumabe M, Franek BS, Niewold TB. Functional genetic polymorphisms in ILT3 are associated with decreased surface expression on dendritic cells and increased serum cytokines in lupus patients. Ann Rheum Dis. England; 2013 Apr;72(4):596–601.
170. Mamegano K, Kuroki K, Miyashita R, Kusaoi M, Kobayashi S, Matsuta K, et al. Association of LILRA2 (ILT1, LIR7) splice site polymorphism with systemic lupus erythematosus and microscopic polyangiitis. Genes Immun. England; 2008 Apr;9(3):214–23.
171. Bleharski JR, Li H, Meinken C, Graeber TG, Ochoa M-T, Yamamura M, et al. Use of genetic profiling in leprosy to discriminate clinical forms of the disease. Science. United States; 2003 Sep;301(5639):1527–30.
172. Tedla N, An H, Borges L, Vollmer-Conna U, Bryant K, Geczy C, et al. Expression of activating and inhibitory leukocyte immunoglobulin-like receptors in rheumatoid synovium: correlations to disease activity. Tissue Antigens. Denmark; 2011 Apr;77(4):305–16.
173. Koch S, Goedde R, Nigmatova V, Epplen JT, Muller N, de Seze J, et al. Association of multiple sclerosis with ILT6 deficiency. Genes Immun. England; 2005 Aug;6(5):445–7.
174. Ordonez D, Sanchez AJ, Martinez-Rodriguez JE, Cisneros E, Ramil E, Romo N, et al. Multiple sclerosis associates with LILRA3 deletion in Spanish patients. Genes Immun. England; 2009 Sep;10(6):579–85.
175. Kabalak G, Dobberstein SB, Matthias T, Reuter S, The Y-H, Dorner T, et al. Association of immunoglobulin-like transcript 6 deficiency with Sjogren's syndrome. Arthritis Rheum. United States; 2009 Oct;60(10):2923–5.
176. Du Y, Su Y, He J, Yang Y, Shi Y, Cui Y, et al. Impact of the leucocyte immunoglobulin-like receptor A3 (LILRA3) on susceptibility and subphenotypes of systemic lupus erythematosus and Sjogren's syndrome. Ann Rheum Dis. England; 2015 Nov;74(11):2070–5.
177. Xu J, Mo Z, Ye D, Wang M, Liu F, Jin G, et al. Genome-wide association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4. Nat Genet. United States; 2012 Nov;44(11):1231–5.
178. Teslovich TM, Musunuru K, Smith A V, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. England; 2010 Aug;466(7307):707–13.
179. Ye Y-M, Hur G-Y, Kim S-H, Ban G-Y, Jee Y-K, Naisbitt DJ, et al. Drug-specific CD4+ T-cell immune responses are responsible for antituberculosis drug-induced maculopapular exanthema and DRESS. Br J Dermatol. 2016 Jul;
180. Nguyen D Van, Vidal C, Li J, Fulton RB, Fernando SL. Validation of a rapid test for HLA-B*58:01/57:01 allele screening to detect individuals at risk for drug-induced hypersensitivity. Pharmacogenomics. England; 2016 Apr;17(5):473–80.
181. Mugwagwa AN, Fischer R, Zailan I. HLA-B*5801: a genetic susceptibility to allopurinol-induced DRESS. Med J Aust. Australia; 2016 Mar;204(4):159–60.
182. Rodriguez-Perea AL, Montoya CJ, Olek S, Chougnet CA, Velilla PA. Statins

increase the frequency of circulating CD4+ FOXP3+ regulatory T cells in healthy individuals. J Immunol Res. Egypt; 2015;2015:762506.
183. Coward WR, Marei A, Yang A, Vasa-Nicotera MM, Chow SC. Statin-induced proinflammatory response in mitogen-activated peripheral blood mononuclear cells through the activation of caspase-1 and IL-18 secretion in monocytes. J Immunol. United States; 2006 May;176(9):5284–92.
184. Pettersson FH, Anderson CA, Clarke GM, Barrett JC, Cardon LR, Morris AP, et al. Marker selection for genetic case-control association studies. Nat Protoc [Internet]. 2009;4(5):743–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19390530
185. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. England; 2007 Jun;447(7145):661–78.
186. Williams T, 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. England; 2012 Apr;3(2):89–99.
187. O'Meara H, Carr DF, Evely J, Hobbs M, McCann G, van Staa T, et al. Electronic Health Records For Biological Sample Collection: Feasibility Study Of Statin-Induced Myopathy Using The Clinical Practice Research Datalink. Br J Clin Pharmacol [Internet]. 2013; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24308359
188. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. United States; 2007 Jul;39(7):906–13.
189. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics. England; 2010;11:288.
190. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. England; 2015;4:7.
191. 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 [Internet]. The American Society of Human Genetics; 2007 Sep 25;81(3):559–75. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838/
192. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. England; 2010 Mar;26(6):841–2.
193. Aho A.V., Kernighan B WPJ. The AWK programming language. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.; 1987.
194. Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics. UNITED STATES; 1995 Sep;29(2):311–22.
195. Lewontin RC. On measures of gametic disequilibrium. Genetics. UNITED STATES; 1988 Nov;120(3):849–52.
196. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. England; 2010 Sep;26(18):2336–7.
197. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. Biopreservation and biobanking. United States; 2015. p. 307–8.

198. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. Biopreserv Biobank. United States; 2015 Oct;13(5):311–9.
199. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. United States; 2015 May;348(6235):648–60.
200. McTaggart F, Jones P. Effects of Statins on High-Density Lipoproteins: A Potential Contribution to Cardiovascular Benefit. Cardiovasc Drugs Ther [Internet]. Boston: Springer US; 2008 Aug 14;22(4):321–38. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2493531/
201. Barter PJ, Brandrup-Wognsen G, Palmer MK, Nicholls SJ. Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database. J Lipid Res [Internet]. The American Society for Biochemistry and Molecular Biology; 2010 Jun 1;51(6):1546–53. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3035518/
202. Sokolowska M, Chen L-Y, Eberlein M, Martinez-Anton A, Liu Y, Alsaaty S, et al. Low molecular weight hyaluronan activates cytosolic phospholipase A2alpha and eicosanoid production in monocytes and macrophages. J Biol Chem. United States; 2014 Feb;289(7):4470–88.
203. Bryois J, Buil A, Evans DM, Kemp JP, Montgomery SB, Conrad DF, et al. Cis and trans effects of human genomic variants on gene expression. PLoS Genet. United States; 2014 Jul;10(7):e1004461.
204. Matloubian M, Lo CG, Cinamon G, Lesneski MJ, Xu Y, Brinkmann V, et al. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. Nature. England; 2004 Jan;427(6972):355–60.
205. Allende ML, Dreier JL, Mandala S, Proia RL. Expression of the sphingosine 1-phosphate receptor, S1P1, on T-cells controls thymic emigration. J Biol Chem. United States; 2004 Apr;279(15):15396–401.
206. Kuswanto W, Burzyn D, Panduro M, Wang KK, Jang YC, Wagers AJ, et al. Poor Repair of Skeletal Muscle in Aging Mice Reflects a Defect in Local, Interleukin-33-Dependent Accumulation of Regulatory T Cells. Immunity. United States; 2016 Feb;44(2):355–67.
207. Goldstein MR, Mascitelli L, Pezzetta F. The double-edged sword of statin immunomodulation. International journal of cardiology. Netherlands; 2009. p. 128–30.
208. Mascitelli L, Goldstein MR. Statins, regulatory T cells and diabetes. Diabetic medicine : a journal of the British Diabetic Association. England; 2016. p. 1152–3.
209. Goldstein MR, Mascitelli L. Do statins cause diabetes? Curr Diab Rep. United States; 2013 Jun;13(3):381–90.
210. Bapat SP, Myoung Suh J, Fang S, Liu S, Zhang Y, Cheng A, et al. Depletion of fat-resident Treg cells prevents age-associated insulin resistance. Nature. England; 2015 Dec;528(7580):137–41.
211. Doring F, Onur S, Kurbitz C, Boulay MR, Perusse L, Rankinen T, et al. Single nucleotide polymorphisms in the myostatin (MSTN) and muscle creatine kinase (CKM) genes are not associated with elite endurance performance. Scand J Med Sci Sports. Denmark; 2011 Dec;21(6):841–5.

212. Dahlstedt AJ, Katz A, Wieringa B, Westerblad H. Is creatine kinase responsible for fatigue? Studies of isolated skeletal muscle deficient in creatine kinase. Faseb J [Internet]. 2000;14(7):982–90. Available from: http://www.fasebj.org/content/14/7/982.full.pdf