Statin Safety in Prevention of Cardiovascular Diseases: Causal Inference and Risk Prediction

Thesis submitted for the degree of Doctor of Philosophy

Ting Cai

Nuffield Department of Primary Care Health Sciences

Wolfson College

University of Oxford

Trinity Term, 2023

This thesis is dedicated to my parents,

for their unconditional love, support and sacrifices.

Acknowledgement

I would like to express my sincere gratitude to my supervisors: Associate Professor James Sheppard, Professor Richard Hobbs, Dr Constantinos Koshiaris, Dr Jennifer Hirst, Associate Professor Brian Nicholson, and Dr Sarah Lay-Flurrie, for all their support, guidance, and valuable advice during my research for this thesis. My deep appreciation is especially given to Associate Professor James Sheppard and Dr Constantinos Koshiaris, who have invested considerable time in helping and mentoring me throughout.

I am also grateful to Professor Anthony Harnden and Professor Rafael Perera, the former and the current Director of Graduate Studies in the Department of Primary Care Health Sciences, for their listening and advice during my study. I would like to acknowledge Professor Richard Stevens and Professor Jeffrey Aronson for providing their expertise for part of the work in this thesis. My thanks should also go to Professor Richard McManus, Ms Lucy Curtin, and other members of the Hypertension Team within the department, all of whom create a supportive and friendly environment for my study.

I would like to mention my fellow DPhil students in the department, with special thanks to Rik van der Veen, Sara Rotenberg, Kome Gbinigie, and Mona Koshkouei, for their accompany and sharing.

Finally, I am deeply indebted to my parents, without whom this journey would not have been possible.

This thesis presents independent research funded by the British Heart Foundation (BHF) and the National Institute for Health and Care Research (NIHR) School for Primary Care Research (SPCR). The views expressed are those of the author and not necessarily those of the National Health Service (NHS), the NIHR, or the UK Department of Health.

Table of Contents

Abstract

Background

The widespread concerns about statin safety have resulted in low uptake of and poor adherence to statin treatment for prevention of cardiovascular diseases. The use of statins for primary prevention has been particularly challenging due to the controversy about the balance between benefits and harms of treatment. Personalised clinical decision-making and stratified treatment strategies that take into account the risk of adverse events are potential approaches towards better use of statins.

Methods

A systematic review of randomised controlled trials was conducted, with pair-wise, network, and dose-response meta-analyses, to assess the associations between statins and common adverse events and explore the variations by drug type and dose in primary prevention patients. A prognostic model (StatinMD) was derived and externally validated to predict the personalised risk of serious muscle disorders in individuals eligible for statin treatment, using a competing risk model with data from electronic healthcare records.

Results

Statins were associated with a small increase in the risk of muscle symptoms, liver dysfunction, renal insufficiency, and eye conditions, but not with muscle disorders or diabetes. There was little evidence of the difference between statin drugs or the dose-response relationships of their adverse effects. The StatinMD model included 22 predictors to predict the risk of serious muscle disorders in 1, 5, and 10 years. The model showed overall good discrimination and calibration in the majority of the population.

Conclusions

The overall balance between benefits and harms of statins supports their use for primary prevention of cardiovascular diseases. The StatinMD model provides a reliable predicted risk of serious muscle disorders for most individuals to assist clinical decision-making on statin treatment.

List of Abbreviations

ACC	American College of Cardiology
AHA	American Heart Association
ALT	Alanine aminotransferase
ARD	Absolute risk difference
AST	Aspartate aminotransferase
BMI	Body mass index
BNF	British National Formulary
CI	Confidence interval
CIF	Cumulative incidence function
CK	Creatine kinase
COPD	Chronic Obstructive Pulmonary Disease
CPRD	Clinical Practice Research Datalink
CrI	Credible intervals
CTT	Cholesterol Treatment Trialists' Collaboration
CVD	Cardiovascular diseases
CYP450	Cytochrome P450 enzymes
EHR	Electronic healthcare records
EPP	Number of events per predictor coefficient
FI	Frailty index
FP	Factional polynomial function
GP	General practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation

HbA1c	Glycated haemoglobin A1c
HES	Hospital Episode Statistics
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
ICD	International Classification of Diseases
IMD	Index of multiple deprivation
IPD	Individual participant data
IQR	Interquartile range
LDL-C	Low-density lipoprotein cholesterol
LOESS	Local polynomial regression
LR-	Negative likelihood ratio
LR+	Positive likelihood ratio
MAR	Missing at random
MBMA	Model-based meta-analysis
MBNMA	Model-based network meta-analysis
MCAR	Missing completely at random
MD	Muscle disorders
MICE	Multivariate imputation by chained equations
MNAR	Missing not at random
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NPV	Negative predictive value
ONS	Office of National Statistics
OR	Odds ratio
PI	Prognostic index

PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
RCT	Randomised controlled trial
RR	Risk ratio
SD	Standard deviation
SE	Standard error
SHR	Subdistribution hazards ratio
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis
ULN	Upper limit of normal

Chapter 1 Introduction and Objectives

1.1 Importance of understanding statin safety

1.1.1 Crucial role of statins in prevention of cardiovascular diseases

Cardiovascular diseases (CVD) are the leading cause of global mortality, taking around 17.9 million lives each year and accounting for one third of all deaths worldwide.[1] One important risk factor of CVD is high plasma cholesterol, especially low-density lipoprotein cholesterol (LDL-C).[2] There is strong evidence that lowering LDL-C level reduces the incidence of CVD and related deaths.[3, 4] Two major strategies for cholesterol management are lifestyle intervention and lipid-lowering medication.[5-7] Statins, a class of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, are the cornerstone of lipid-lowering therapy.[8, 9]

Over the past half century, the evidence of statins' effects on lowering LDL-C level and reducing CVD risk has been well established.[10, 11] Large clinical trials showed that lowering LDL-C per 1 mmol/L with statins reduced major vascular events by about 21%, including coronary deaths, myocardial infarctions, strokes, and coronary revascularisations.[11] According to this treatment effect and the CVD incidence in the population, it has been estimated that taking a statin regimen for 5 years to lower LDL-C by 2 mmol/L on average would prevent 1,000 cases of major vascular events in 10,000 persons with a history of CVD or 500 cases in those without CVD.[12] Based on the evidence, statins have been recommended in clinical guidelines as the front-line therapy for prevention of CVD and are widely prescribed around the world.[5-7] In 2018, an estimated 145.8 million people used statins in 83 countries.[13] In the UK, statins are one of the most commonly prescribed drugs, with about 1 in 8 adults taking statins.[14] Without

a doubt, statins play a crucial role in CVD prevention, with life-saving benefits for a large number of people.

1.1.2 Concerns about statin safety and the consequences

With the broad utilisation of statins, concerns about their safety have arisen from the reports of adverse events in statin users.[15, 16] The concerns were amplified by the cases of fatal rhabdomyolysis with the use of cerivastatin.[17] Following a total of 52 reported deaths from rhabdomyolysis, cerivastatin was withdrawn from the world market in 2001.[18] This led to cautions about the safety of this drug class and provoked re-examination of the safety data on all statins.[17, 18] Although the reports of fatal rhabdomyolysis with the use of other statins were found exceptionally rare, myopathy and rhabdomyolysis have been labelled on all marketed statins as potential adverse events since then.[17, 19] Continuing post-marketing surveillance has been carried out and many research efforts have been put into understanding the muscle-related adverse events of statins.[19, 20]

Like many other drugs that are primarily metabolised via the liver, hepatotoxicity is another potential adverse effect of statins and it has been recommended to perform liver function tests before the initiation of statin treatment.[21, 22] A newly emerged concern is about the increased risk of type 2 diabetes in statin users reported in clinical trials, which has led to a warning message added to the product safety labels of statins.[23, 24] Other reported adverse events from statin users, including kidney injury, cataracts, cognitive impairments, and haemorrhagic stroke, have also caught attention, although most of their causal associations with statins remain under exploration.[25, 26]

Since the withdrawal of cerivastatin, even though the authorities and professional societies have tried to reassure patients about the safety of other statins, negative stories about statins' adverse events have been widely covered in the mass media.[27, 28] This has

exacerbated the concerns in the public and led to an increasing rate of statin treatment discontinuation in many countries.[29, 30] Consequently, the incidence of CVD and related deaths has increased in patients with poor adherence to statin treatment.[30, 31]

In the UK, public attention on statin safety peaked in 2013, following publications in a top medical journal that were perceived as critical of statins.[32, 33] A subsequent study found that discontinuation of statin treatment significantly increased after the intense media coverage of the publications and the following public discussion.[34] This study estimated that about 219,000 patients might have discontinued statins after the media coverage of the criticism on statins, which might have resulted in an excess of up to 6,400 cases of avoidable CVD. These consequences highlight the urgency of tackling public concerns and improving the understanding of statin safety, given the fact that the rate of statin uptake in eligible patients is currently very low in most countries.[35]

1.2 Challenges in the use of statins for primary prevention

1.2.1 Controversy over the widened use for primary prevention

Although the use of statins for secondary prevention in patients with established CVD is well supported, their use for primary prevention in patients without previous cardiovascular events has been more controversial, particularly in patients at low risk of CVD.[36, 37] The debate was intensified by the change to the threshold of CVD risk for statin treatment in clinical guidelines. In 2013, the American College of Cardiology (ACC) and the American Heart Association (AHA) guideline was updated to recommend prescribing statins to almost all patients aged 40-75 years with a 10-year CVD risk above 7.5%.[38] The National Institute for Health and Care Excellence (NICE) in the UK has also progressively lowered the 10-year CVD risk threshold for prescribing statins from 20%

to 10% in the 2014 guideline and to below 10% in the draft guideline expected to be published later in 2023.[5, 39] Many other guidelines have similarly reduced the risk threshold for statin initiation.[40]

These changes have aroused criticism from those who argued that the benefits of statin treatment for patients at low CVD risk are limited.[32, 41, 42] Such criticisms were made despite the evidence from systematic reviews of randomised controlled trials (RCT), which showed significant reductions of major vascular events and all-cause mortality by statins in primary prevention patients, including low-risk patients.[43, 44] The arguments have focused on the small amount of absolute reduction in the risk of vascular events and mortality, given the lower baseline risk in this population compared to the population of secondary prevention.[32, 41, 45]

More importantly, concerns about statin safety have been highlighted in this debate. It was estimated that lowering the risk threshold from 20% to 10% would make an additional 4.5 million people in the UK eligible for statin treatment, including 83% of all men older than 50 years and 56% of all women older than 60 years.[46, 47] This means that a large number of people at low CVD risk may be exposed to the risk of adverse events with statins. However, compared to the well-established benefits of statins, more uncertainty remains in their potential harms.[36] Although some known adverse events, such as rhabdomyolysis, are believed to be rare, given the smaller benefits of statins in the primary prevention population, the centre of the controversy is around whether the potential harms outweigh the benefits in this population.[32, 41]

1.2.2 Unproven practices of tailoring statin regimens for safety

Currently, there are seven types of statins in clinical use around the world: atorvastatin, fluvastatin, lovastatin, pitavastatin, pravastatin, rosuvastatin, and simvastatin, each with

multiple doses available.[48, 49] Given the smaller benefits of statin treatment in primary prevention patients, in clinical practice, one may attempt to tailor statin regimens by choosing a particular drug type or dose to minimise the risk of adverse events in these patients.[50-52]

In current clinical guidelines, the recommendations on the choice of statin regimens are only based on their effects on lowering LDL-C and the goal of treatment.[5, 7, 10] However, statin drugs differ not only in their lipid-lowering potency but also potentially in their safety profiles, given their different pharmacokinetic properties.[53] The difference in their efficacy may not be in accordance with the difference in their safety, due to the separate mechanisms of their lipid-lowering effect and potential adverse effects. Choosing a low-potency statin drug with the expectation of a lower risk of adverse events is currently an unproven practice and further evidence is needed to support the choice of statin drugs for safety concerns in clinical practice.

Similarly, the dose-response relationships of statins' effects on LDL-C may not be applicable to their adverse effects, given the separate mechanisms that have different pharmacodynamic processes.[48, 53] The current stratification of the intensity of statin doses is based on the percentage of LDL-C reduction, and moderate to high-intensity doses (e.g. atorvastatin 20mg) are recommended in clinical guidelines as initial statin treatment for primary prevention. However, high-intensity dosing of statins has been questioned for producing little further reduction of coronary death or total mortality. The use of high-intensity doses of statins may alter the benefit-risk ratio of treatment, particularly in primary prevention, if their beneficial effects plateau while adverse effects continue escalating. Understanding the dose-response relationships of adverse effects could be useful for determining the optimal therapeutic dose range of statins for primary

prevention, avoiding doses that provide little extra benefit but may increase the risk of potential harm.[49, 54]

1.3 Potential approaches towards better use of statins

1.3.1 Personalised treatment decision-making based on risk prediction

General treatment recommendations in clinical guidelines are usually based on the relative effects at a population level. However, in clinical practice, treatment decisions are made based on individual cases and the absolute gain or loss from treatment in an individual is more relevant when making the decision.[55, 56] Although a relative effect of a treatment is often believed to be consistent across the population, even if it is true, the absolute increase or reduction in the risk of the treatment outcome may vary greatly by individual.[57] This is because the baseline risk of the outcome event for each individual may be different due to different personal characteristics.[58] For example, with a relative effect measured by risk ratio (RR) being 1.5, the absolute increase in the outcome risk would be 25% when the baseline risk is 50% and 5% when the baseline risk is 10%. This may lead to completely different treatment decisions. Such personalised information is required regarding both the beneficial and harmful outcomes of a treatment in order to make a well-informed decision.

In terms of statin treatment, the main benefit is to reduce CVD, of which the personalised baseline risk can be calculated by the existing risk prediction tools, such as the QRISK2, the SCORE2, and the Framingham models.[59-61] The absolute reduction of the CVD risk in an individual can be subsequently obtained by applying the relative effect of statins on CVD demonstrated in clinical trials to the predicted baseline risk. However, for the potential harms of statins, very few tools are available to predict the personalised

baseline risk of adverse events.[62] This makes it difficult to understand the balance between benefits and harms of statin treatment when making a treatment decision for an individual. It is therefore warranted to develop risk prediction tools for the adverse events of statins to assist personalised decision-making on statin treatment.

1.3.2 Stratified treatment strategy with risk stratification

As an important strategy to improve healthcare in the population, stratified treatment aims to target a treatment at the patient group who could have the most benefits or the least harms from the treatment.[63, 64] One approach to stratifying the use of a treatment in the population is to consider the absolute gain or loss from the treatment, which is essentially based on the baseline risk of the treatment outcome when the treatment effect is believed to be consistent, as explained above.[65] When considering the efficacy, it may be more worthwhile treating individuals with a higher baseline risk of the outcome as they could have substantial absolute benefits from treatment. On the contrary, when considering safety, one may wish to be more cautious in treating individuals with a higher baseline risk of the outcome since they may suffer more harms from treatment. A threshold of the baseline risk of the outcome may be considered to determine the stratifying strategy of treatment.

The current stratifying strategy of statin treatment is determined only by the threshold of the baseline CVD risk.[40] However, as discussed above, it is crucial to take into account the risk of adverse events when considering the stratified treatment strategy in the population. This could help better target statin treatment in the population with an ideal balance between benefits and harms and prevent potential harms from statins in high-risk individuals. Such stratified treatment strategy could be determined upon the availability of prediction tools for the adverse events of statins.

1.4 Adverse events of interest in this thesis

The specific outcomes of interest in this thesis are the common adverse events that have resulted in amends of safety labelling on statin products or caused significant concerns for the clinical use of statins. The associations of these adverse events with statin treatment have not been fully understood or explicitly quantified, with continuing controversy and requiring further research, as described below.

1.4.1 Muscle-related adverse events

Most of the concerns about statin safety come from the adverse events related to muscles, given the notorious history of cerivastatin.[17, 20] Muscle-related problems have become the most commonly reported adverse event in statin users and a major cause of poor adherence to statin treatment.[66, 67] They are therefore the focus of much of the research into statin safety, including in this thesis. The muscle-related adverse events observed in statin users have a wide spectrum of clinical presentations, including myalgia, myopathy, myositis, and rhabdomyolysis as the common types.[68] The definitions and distinctions of these specific types are inconsistent in studies, but they are often considered as the following two groups based on their severity.[20, 67]

- **Serious muscle disorders**

Serious muscle disorders are muscle conditions presenting with clinically significant elevation in blood creatine kinase (CK), mainly including myopathy, myositis, and rhabdomyolysis.[20, 67] Myopathy is considered with CK over 10 times the upper limit of normal (ULN), often but not necessarily accompanied by muscle symptoms, especially muscle weakness.[26, 67] Myositis is similar to myopathy but associated with muscle

inflammation.[20] Rhabdomyolysis is the most serious type that could present a very high CK >40ULN and lead to renal impairment or myoglobinuria.[26, 67]

Although such serious muscle disorders are generally rare in the population, their associations with statin treatment have been widely recognised based on evidence from clinical trials, large observational studies, and pharmacovigilance surveillance data. Several theories about the underlying mechanisms of statins causing muscle injuries also have been proposed.[12, 20, 67] The magnitude of this adverse effect of statins remains undetermined, but some studies have suggested that the relative treatment effect could be large.[12, 25] Moreover, these serious conditions usually require clinical actions, such as pharmacotherapy, physiotherapy, or hospitalisation, and in the worst cases may end up in death.[69, 70] As such, they may be given more weight in the trade-off between benefits and harms of statins and could have a great influence on statin treatment decision-making.

- **Mild muscle problems**

On the contrary, mild muscle problems seen in statin users, mostly referred to as myalgia, are muscle symptoms, typically muscle pain and tenderness, with normal or slightly raised CK.[67] These conditions are common in the general population, either with or without the use of statins.[71] There are other common causes of myalgia, such as intensive exercise, injury, or viral infection.[67] Their attribution to statin treatment has been particularly controversial, with discrepancy in the evidence from clinical trials and observational studies.[12] Due to the potential 'nocebo effect' in observational studies that the reporting of self-perceived muscle symptoms may be subjected to patients' pre-treatment awareness of muscle-related adverse effects, it has been argued that most of these mild problems are not caused by statins.[12, 72] However, the tolerance of these mild

symptoms varies by individuals and the discomfort could lead to poor adherence to or discontinuation of statin treatment.[67]

1.4.2 Other concerning adverse events

- **Liver dysfunction/injury**

Drug-induced liver injury is a major complication with medications and among the most common causes of discontinuation of drug development and drug withdrawal from the market.[73] Statins, as many other drugs that are mainly metabolised via the liver cytochrome P450 (CYP450) enzymes, have been under close watch for potential hepatotoxicity since their first introduction.[21, 22] Early trials showed that patients on statin treatment experienced an elevation in the liver enzymes alanine aminotransferase (ALT) or aspartate aminotransferase (AST) and the statin labels used to recommend periodic liver enzyme tests to monitor patients' liver function during statin treatment.[22, 74] As the evidence accumulates along with the increasing use of statins, the attribution of the common mild elevation of liver enzymes to statins has been questioned and the requirement for routine monitoring of liver function was removed from the safety labels in 2012.[75, 76] However, large studies of post-marketing surveillance data showed that statin-related liver injury could be severe, even though the cases are rare.[77, 78] With the remaining uncertainty, it is still recommended to perform liver enzyme tests before starting statin treatment and decompensated cirrhosis and acute liver failure are labelled as contraindications for statin use.[22]

- **Type 2 diabetes**

The risk of diabetes with statins became the centre of attention following the publication of a large clinical trial in 2008, which found a higher average concentration of glycated haemoglobin (HbA1c) and a modest increase in the rate of newly diagnosed

type 2 diabetes among patients on statin treatment.[79] This motivated the re-examination of the data from previous clinical trials, which yielded similar results.[80, 81] Based on the evidence, a statement indicating potential elevation in HbA1c and fasting glucose levels was added to the statin labels in 2012.[23] Considering that diabetes could become a life-long chronic disease with serious complications and substantial disease burden, discussions have focused on whether the increased risk of diabetes or more probably the earlier presentation of diabetes outweighs the risk reduction of CVD by statin treatment.[24]

- **Renal insufficiency/failure**

Concerns about the potential harms of statins to the kidney were raised shortly after rosuvastatin was first approved in 2003 in the US. The call for the withdrawal of rosuvastatin from the market was made based on evidence that trial participants taking rosuvastatin, particularly with a dose of 80 mg/d, had a higher rate of persistent proteinuria or haematuria, and cases of renal insufficiency or failure had been observed in rosuvastatin users since marketing began.[82] As a consequence, 80 mg/d rosuvastatin was no longer approved, but the lower doses remain on the market with even increasing use.[83, 84] This also provoked questions about the risk of kidney injury with other statins, especially in light of the statin-related rhabdomyolysis and diabetes which may lead to renal failure and microvascular kidney diseases.[12] Interestingly, some attention has been drawn to the anti-inflammation effect of statins and the potential benefits for improving kidney function in patients with chronic kidney disease.[85, 86]

- **Cataracts and other eye conditions**

In the early stage of statin development, cataractogenic effects of statins were observed in animal studies but this did not receive much attention.[87] Subsequent clinical trials that examined lens opacities and other vision-related outcomes did not identify an increased

risk in statin users.[88, 89] However, serious concern has been provoked by several large observational studies since 2010, which found a considerable increase in the risk of developing cataracts and having cataract surgery among statin users compared to non-users.[90] This concern was intensified by the publication of a large clinical trial in 2016 that showed a small excess of cataract surgery in patients on statin treatment.[91] In the meantime, similar to the case of kidney-related adverse effects, studies exploring the potential protective effect of statins on the progression of age-related macular degeneration have emerged, contributing to the heterogeneous evidence base of statins' effect on eyes.[92, 93]

1.5 Objectives of this thesis

1.5.1 Understanding statin safety in primary prevention of CVD

As introduced above, with the widespread concerns about the safety of statins, continuing controversy surrounds the use of statins in primary prevention of CVD. Given the smaller benefits of statin treatment for primary prevention, it becomes critical to address the remaining uncertainty about the common adverse events in this population, as the potential harms are more likely to shift the benefit-risk ratio of treatment.

Therefore, this thesis first aims to investigate the safety profile of statins in the population of primary prevention, in order to better understand the balance between benefits and harms of statin treatment in this population. This thesis also aims to explore the rationales of tailoring statin regimens by drug type or dose, which may be considered as a potential approach to minimise the harm from statins when treating primary prevention patients.

These can be achieved by leveraging existing data from clinical trials to perform the following specific tasks:

1) To assess the associations between statins and the adverse events of interest listed above in patients without a history of CVD;
2) To examine the major efficacy outcomes and compare the benefits and harms of statin treatment in these patients;
3) To compare the potential adverse effects between statin drug types;
4) To examine the dose-response relationships in the potential adverse effects of each statin drug.

1.5.2 Predicting personalised risk of serious muscle disorders

The second aim of this thesis is to develop risk prediction tools for the adverse events of statins, in order to assist personalised clinical decision-making on statin treatment and support potential stratified statin treatment strategy. This will be beneficial for patients of both primary and secondary prevention, whenever a decision on statin treatment needs to be made and safety concerns come into play.

Although all of the common adverse events introduced above could be concerns for patients, some may be more influential than others in the decision-making on statin treatment. As discussed above, serious muscle disorders could have a great influence on statin treatment decisions, given their widely-recognised association with statins and the potential poor prognosis. Therefore, as a priority for the time-consuming work of prediction modelling, this thesis focuses on developing one prediction model for serious muscle disorders. To facilitate the comparison with the CVD risk to understand the benefit-harm balance of statin treatment in clinical practice, the prediction model primarily aims to predict the 10-year risk of serious muscle disorders. Considering that

shorter-term risks may be also of concern for patients, the model is expected to also provide predictions for 1-year and 5-year risks.

The model development includes two studies:

1) To derive a model predicting the risk of serious muscle disorders based on personal characteristics in patients eligible for statin treatment for both primary and secondary prevention of CVD;
2) To validate the prediction model with the assessment of model performance and potential clinical utility for the purpose of prediction.

Chapter 2 Design and Methodology

2.1 Overarching design of this thesis

In line with the objectives set in the first chapter, the overarching design of this thesis comprises two paths towards the purposes of making causal inference and risk prediction respectively, as illustrated in **Figure 2.1**.

For making causal inferences about the potential adverse effects of statins in primary prevention, a systematic review of existing randomised controlled trials (RCT) was conducted. Pair-wise meta-analysis comparing statin users and non-users was performed to assess the overall associations with adverse events and the balance between potential benefits and harms for the drug class of statins. Network meta-analysis was used to further compare the potential adverse effects between different statin drug types. E_{max} model-based meta-analysis was applied to examine the dose-response relationships in the potential adverse effects for each type of statin.

To predict the risk of serious muscle disorders for statin treatment decision-making, a prediction model was developed in the population potentially eligible for statin treatment. The model was derived to estimate the personalised risk of serious muscle disorders in 1, 5, and 10 years based on patient characteristics, medical history, and treatment prescription, using data from electronic healthcare records (EHR) in the UK. The model was internally and externally validated by assessing the prediction performance of the model and the potential clinical utility for the purpose of risk prediction.

The rationale and important methodological considerations for the key methods used in the three studies are described below. The technical and execution details are further reported in the following chapters for each study.

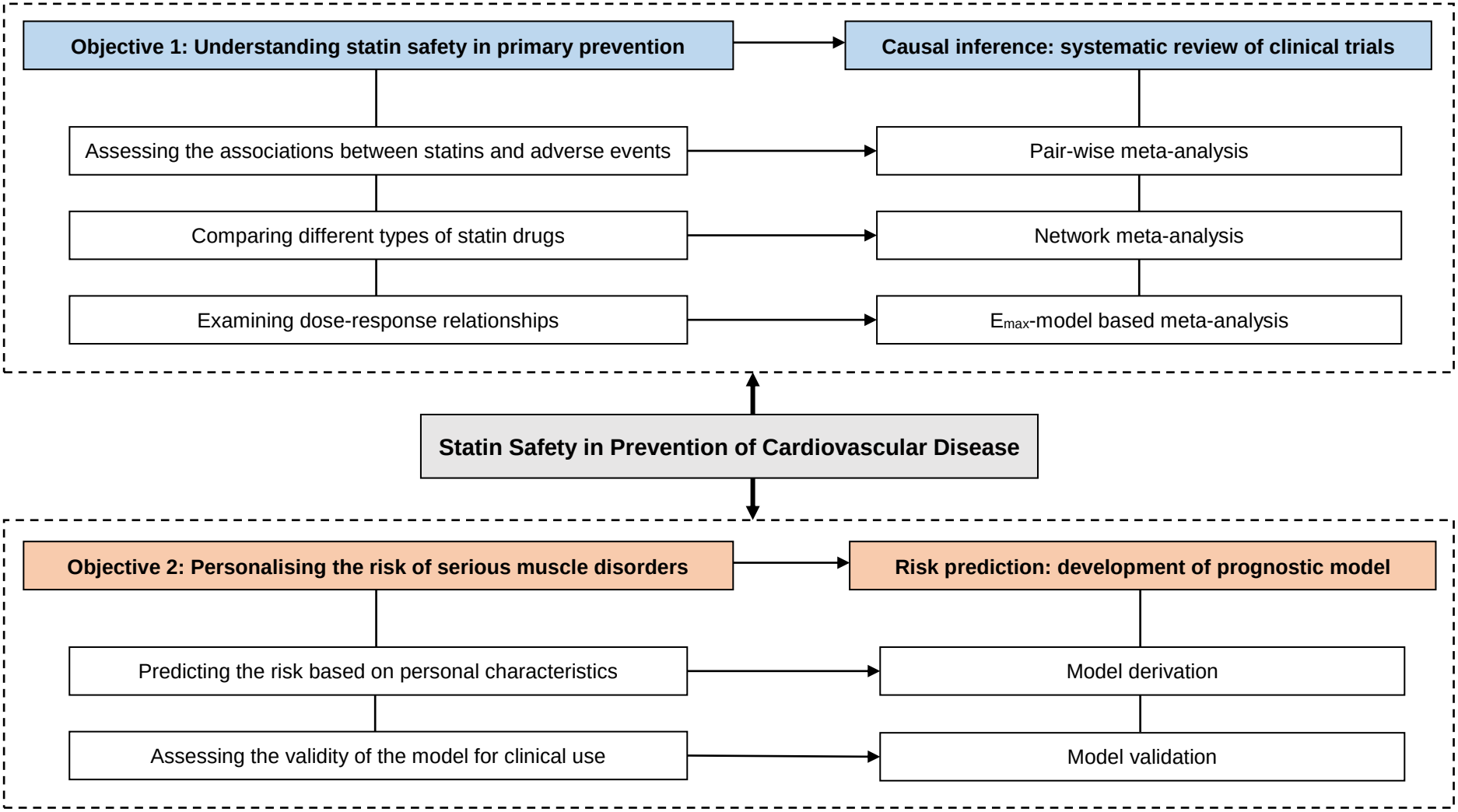


Figure 2.1 Overarching design of this thesis

2.2 Systematic review and meta-analysis

2.2.1 Strengths of a systematic review of randomised controlled trials

In this thesis, the associations between statins and adverse events were assessed through a systematic review of RCTs, which is regarded as the strongest evidence for making causal inferences about treatment effects.[94, 95] It is especially useful for examining subjective outcomes, such as muscle pain with statin treatment, which could be biased by the 'nocebo effect' in the absence of randomisation or control, as individuals on treatment are more likely to perceive an event given their pre-treatment awareness of the potential treatment outcomes.[72] Although there have been a considerable number of RCTs for statin treatment, not all RCTs were conducted with as much rigour as others and the results on safety outcomes from these RCTs were conflicting.[96, 97] A systematic review provides an approach to address some of the limitations of an individual RCT and the discrepancies among different trials, through synthesis and scrutiny of the results from all relevant studies.[94]

The systematic review in this thesis was conducted following the standard process recommended in the Cochrane Handbook, involving a series of strategies to reduce potential bias from selective inclusion of studies or subjective summary of study findings.[94, 98] These strategies included comprehensive searches for all relevant studies, the use of pre-specified explicit criteria in study selection, critical appraisal of the included studies, and reproducible integration of study results.[98] Through such a systematic approach, it would not only ensure the reliability of the synthesised evidence for making causal inferences about the adverse effects of statins but also help identify the gap in the current trial evidence regarding statin safety, which would be useful for considering future studies.[94, 99]

2.2.2 Different types of meta-analyses in this thesis

Meta-analysis is the statistical method in a systematic review to integrate results from the included primary studies.[100, 101] The main strengths of this statistical integration are to obtain a pooled result that coordinates the discrepancies among primary studies and to improve the power of statistical tests or the precision of estimation based on a combined sample.[102] With the widening application of systematic reviews, different types of meta-analyses have emerged to integrate studies of various designs and to serve particular purposes of analysis.[103] In this thesis, pair-wise meta-analysis, network meta-analysis, and model-based meta-analysis were conducted for the specific objectives of the systematic review.

2.2.2.1 Pair-wise meta-analysis comparing a treatment to a control

A typical pair-wise meta-analysis was performed to compare statin treatment with non-statin control, combining data from trials that included these two groups.[98, 100, 104] Pair-wise meta-analysis is a two-stage process, where a desired measurement of the treatment effect, such as an odds ratio (OR) or risk ratio (RR), is first estimated in each trial using the same method and a pooled estimate of the measurement is then calculated as a weighted average of the individual estimates.[98, 104]

Depending on the heterogeneity among the included studies, different methods can be used to calculate the pooled estimate.[105] The sources of heterogeneity include clinical heterogeneity due to differences in participants, treatments, or outcomes and methodological heterogeneity due to differences in study design, outcome measurement, or risk of bias. As a result, the treatment effects being evaluated in individual studies may present statistical heterogeneity, which can be tested or quantified by some statistics when pooling the data from individual studies.[106] When the heterogeneity is believed to be low,

one can use a fixed-effects model with the assumption that all included studies are estimating the same underlying true treatment effect and the pooled estimate incorporates only the variation across the participants within each study. When the heterogeneity appears to be considerable, a random-effects model can be used to incorporate variations within and between the included studies into the pooled estimate.[107] In certain circumstances, it may not be appropriate to pool individual studies if the heterogeneity is too high.

2.2.2.2 Network meta-analysis for comparing multiple treatment options

Network meta-analysis (NMA) was used to compare different statin drug types, making use of trials that included multiple treatment groups of different statin drugs, in addition to those trials involved in the pair-wise meta-analysis. NMA was invented to support evidence-based healthcare policy-making or clinical decision-making among multiple available treatment options, in the lack of direct comparisons among some of the treatments. The idea of NMA is to build up indirect comparisons between two treatments through one or more common controls that both treatments have been respectively compared within different studies.[108] It subsequently integrates all indirect comparisons, as well as direct comparisons if they exist, to obtain a pooled estimate of the comparative effect between the two treatments.[109, 110] Although such indirect evidence may not be as powerful as evidence from a direct comparison, NMA provides an opportunity to efficiently compare multiple treatments based on existing trial data, which is particularly useful when it is difficult or ethically inappropriate to conduct head-to-head trials.[111]

NMA shares the fundamental assumptions and methods with pair-wise meta-analysis, including the heterogeneity concept and the fixed-effects or random-effects models.[112] In addition, it takes account of the consistency (or coherence) across different comparisons

between two treatments, which is the assumption about whether the comparative effect between two treatments in their direct and indirect comparisons is expected to be identical.[113] Consistency model and inconsistency model can be used to combine direct and indirect comparisons under different assumptions. The consistency can be further assessed after pooling the data from all comparisons.[114, 115]

2.2.2.3 Model-based meta-analysis for examining dose-response relationships

A model-based meta-analysis, specifically E_{max} model-based network meta-analysis, was used to examine the dose-response relationships of the potential adverse effects of statins. Clinical trials for examining dose-response relationships of treatment effects require more complicated designs and are less common.[116] Meta-analysis provides an opportunity to make fuller use of the data from existing dose-response trials to help understand the relationships.[117] This needs a combination of meta-analysis approach and mathematical modelling of dose-response relationships, which is the type of model-based meta-analysis (MBMA).[118]

Early MBMA methods for dose-response data use simple regression models, such as linear regression, which ignore pharmacological plausibility and are not clinically interpretable.[119] Recent methods have evolved to combine pharmacological models that express the underlying pharmacokinetics and pharmacodynamics of drug effects, such as the E_{max} model,[120] into the meta-analysis process, but only allow for examining the dose-response relationship for one drug at a time.[121, 122] A model-based network meta-analysis (MBNMA) method has been newly developed, combining the E_{max} model or other pharmacological models with the NMA framework.[123] This makes it possible to simultaneously examine the dose-response relationships for several drugs, such as different statin drugs in this thesis. This method has the theoretical strength of the

pharmacological models and the advantages of the NMA approach to leverage the dose-response data that are generally sparse.[123]

2.3 Development of a risk prediction model

2.3.1 Principles and key steps in prognostic prediction modelling

2.3.1.1 Definition and purpose of a prognostic prediction model

Prognosis refers to a clinical outcome that individuals could potentially develop in the future, given their current health status.[124, 125] As a health outcome, it is often associated with multiple factors by nature, including patient characteristics and medical history, which are called prognostic factors.[126] A prognostic prediction model is a statistical combination of a set of prognostic factors, usually a multivariable regression model, from which the risk of developing the outcome of interest can be calculated.[126]

Although multivariable regression models are also commonly used in etiologic studies for causal inference, the interpretation of a prognostic prediction model is different from that of an etiologic model. In an etiologic model, the interest is the coefficient of a particular risk factor that indicates the relative risk of the outcome with or without the risk factor, taking into account other factors (confounders) that may confound the association between the risk factor and the outcome.[127] In contrast, the interpretation of a prognostic model focuses on the predicted probability of the outcome calculated by the model, which is an absolute risk of the outcome based on the combination of all included prognostic factors (predictors) with their coefficients.[128] This combination is also known as the ‘prognostic index’ or the ‘prognostic score’, which personalises the predicted risk of the prognostic outcome for an individual based on a shared baseline risk of the outcome in the population.[129, 130] The coefficient of a single predictor reflects the contribution of

this predictor to the prognostic index when combined with other included predictors but does not necessarily indicate its causal association with the outcome.[127]

2.3.1.2 Derivation of a prognostic prediction model

Depending on the type of outcome for prediction, the derivation of a prognostic prediction model starts from choosing an appropriate statistical model. In this thesis, the outcome for prediction (serious muscle disorders in 1, 5, and 10 years) is a time-to-event outcome, which is defined not only by whether the outcome occurs or not but also by the time of the outcome occurrence.[131] A series of models for survival analysis, such as the Cox model and several extended versions, can be used to derive a prediction model for a time-to-event outcome. In order to use a specific model, it is essential to check whether the underlying assumptions of the model are fulfilled in the study data.[132] For survival models, common assumptions include the proportionality of the outcome hazard and the linearity of the associations between the outcome hazard and continuous predictor variables.

In addition to specifying the statistical model, two key steps in prediction model derivation include the identification of candidate predictors for the outcome and the selection of the predictors to be included in the final model. Candidate predictors usually can be identified from previous studies of the prognostic factors of the outcome or based on clinical knowledge about the factors that are believed to be clinically relevant.[133] These factors do not necessarily cause the outcome but should be assumed to be predictive of the outcome occurrence.[127] Currently there is no consensus on the best method for selecting the final predictors. Some statistical selection methods based on the significance tests have been widely used but also criticised for the purely data-driven selection procedure that may result in overfitting of the model and elimination of some predictors

of clinical concerns.[134, 135] For clinical prediction models like the model for serious muscle disorders in this thesis, it is suggested that the predictor selection should consider clinical knowledge and previous studies rather than solely relying on statistical tests.[136]

2.3.1.3 Validation of a prognostic prediction model

If a model has not been derived properly in the above-mentioned process, for example, the statistical model used is inappropriate or some important predictors are omitted, the model may not be able to predict the outcome accurately.[137] When a model is applied to unseen patients in clinical practice, the patient characteristics and clinical settings are often different from those the model derivation was based on and the model may not work well in this situation.[138] For these reasons, before a prediction model can be used in clinical practice, evidence is required that the model's performance is satisfactory for the purpose of prediction in the target population.[139] The process of assessing model performance is referred to as 'model validation'.[137] This sometimes also includes the evaluation of the potential clinical utility of the model to understand the consequences of using the model for clinical decision-making.

The strategies of model validation are classified as 'internal validation' and 'external validation'.[140] Internal validation is to assess the model performance using a part or the whole sample that has been used for the model derivation.[141] When using all the derivation data for internal validation, it is also called apparent validation. Other common types of internal validation include the split-sample approach, cross-validation, and bootstrapping resampling techniques, which are more useful when the sample size is limited.[141] It can inspect the robustness of the model derivation process and demonstrate the reproducibility of the derived model, which could be informative for corrections or adjustments of the model when necessary.[142, 143] However, the model performance

measured in internal validation is likely to be over-optimistic, as the model has been fitted to the same or similar data.[141] But it can provide a reference level for external validation to help understand the model performance.

External validation is to assess the model performance in another sample that is different from the model derivation sample, in terms of their population characteristics, geographic locations, clinical settings, or data collection approaches.[144] This enables evaluation of the generalisability, also called the transportability, of the model, which could justify the application of the model to unseen individuals in clinical practice.[142] Since different data are used, the model performance measured through external validation can avoid over-optimism in internal validation.[144] It could also demonstrate the robustness and reproducibility of the model, as a model with poor internal validity is unlikely to perform well in an external sample.[145] Therefore, external validation is believed to be a stronger proof of model performance and an indispensable step in the development of a prediction model.[139, 143]

2.3.2 Competing risk in predicting time-to-event outcomes

2.3.2.1 Competing risk and its impact on predicting time-to-event outcomes

For the time-to-event outcome in this thesis, a common model to predict the outcome risk is the Cox proportional hazards model, which regresses the hazard of the outcome, an instantaneous rate of the outcome occurrence at a specific time point, on multiple predictor variables.[146, 147] The cumulative incidence (the risk) of the outcome can be directly obtained through a transformation of the modelled hazard, based on the assumption of independent censoring.[148] Censoring is when the follow-up of a subject ends before the outcome is observed due to some reasons, such as a subject's death, losing track of a subject, or the end of the desired length for follow-up.[131] Independent censoring

assumes that the distributions of the time to censoring and the time to outcome event are independent and the probability of the outcome occurrence in the subjects who are censored, if they continue to be observed, is the same as in the subjects who do not experience the censoring events and remain for follow-up.[149]

However, this assumption of independence may not be true in reality because there are some circumstances where the future risk of the outcome in the censored subjects, if they remain for follow-up, is different from that in the uncensored subjects. When the distribution of the time to censoring also provides information about the distribution of the time to outcome event, it is called informative censoring.[149] An important type of informative censoring is competing events, which are events that prohibit the occurrence of the outcome.[150] In this thesis, the outcome for prediction is muscle disorders and a possible competing event is deaths from causes other than muscle disorders that have happened before the potential onset of muscle disorders. This is because the subjects who have died without developing muscle disorders will never develop this condition if they hypothetically continue to be observed, meaning their future risk of muscle disorders is different from those who remain to be observed. The cumulative incidence of a competing event is called the competing risk.[150]

With the presence of competing risk, using the Cox model with the incorrect assumption of independent censoring will lead to an overestimation of the outcome risk.[151] This is because the subjects who have experienced a competing event and are no longer at risk of the outcome are treated as if they could still develop the outcome. This overestimation could be considerable when the competing risk in the population is high.[152, 153] In this thesis, the target population for predicting muscle disorders are those potentially eligible for statin treatment. This population could be relatively old and is more likely to be multi-morbid or frail, who may have a high risk of death from causes

other than muscle disorders. In this case, ignoring the competing risk could result in a significant overestimation of the risk of muscle disorders, which may exaggerate the concerns about statin safety and lead to inappropriate treatment decisions. Therefore, it is important to take into account the competing risk in the development of the risk prediction model in this thesis.

2.3.2.2 The Fine-Gray model to handle competing risk

To take account of the competing risk and predict the outcome risk more accurately, the Fine-Gray proportional subdistribution hazards model was used in this thesis to develop the risk prediction model for muscle disorders. The Fine-Gray model is a well-accepted approach with sound theoretical and statistical footing to model the risk of a time-to-event outcome in the presence of competing events.[154] It has been shown to reduce overestimation of the outcome risk and improve prediction accuracy compared to the Cox model when competing events exist.[155]

In the presence of competing risk, the outcome risk can no longer be directly transformed from the hazard of the outcome, as it now depends on both the hazards of the outcome and the competing event.[156] The Fine-Gray model defines a subdistribution hazard as the theoretical hazard of the outcome in the presence of competing events, which can be regressed on multiple predictor variables like in the Cox model.[157] The difference between this subdistribution hazard and the hazard in the Cox model is the risk set for calculating the instantaneous rate of the outcome occurrence.[158] For the subdistribution hazard, the risk set includes not only the subjects who have survived without any events but also those who have experienced the competing event but not the outcome at the time point for the calculation.[153] Hence the subjects censored by the competing event hypothetically remain to be observed as part of the risk set for the

outcome but with the hazard of the outcome being zero. The proportion of these censored subjects in the risk set depends on the competing risk in the population.[158] The calculated subdistribution hazard of the outcome, therefore, incorporates both the hazards of the outcome and the competing event and can be directly transformed into the outcome risk.[156]

Another common approach to handling competing risk is the cause-specific model, which is more suitable for an etiological model rather than a prognostic prediction model.[159] It requires creating separate models for the outcome and the competing event, which adds complexity to the calculation of the outcome risk and makes it difficult to apply to clinical practice as the predictors for the outcome and the competing event are often different.[156, 159, 160] On the contrary, the Fine-Gray approach incorporates the competing risk into one model for the outcome, which is more efficient for modelling and easier to use in clinical practice.

2.3.3 Handling missing data with multiple imputation

2.3.3.1 Presence of missing data and common approaches

Missing data is a common issue in multivariable regression analyses, because often not all of the variables are measured or recorded for each subject in a study.[161] This problem is particularly evident in studies using data from routine healthcare records, as these data are not measured or collected for the specific purposes of a study.[162, 163] In this thesis, the prediction model for muscle disorders is a multivariable model using data from primary care records, specifically the Clinical Practice Research Datalink (CPRD) in the UK. Missing data is therefore expected, as shown in the CPRD data profile, and needs to be handled properly.

One straightforward approach to handling missing data is to exclude subjects with missing data and perform a complete-case analysis. This approach often considerably reduces the sample size of the study and the statistical power of analysis and is only appropriate when data are missing completely at random (MCAR), which is rarely true in reality. Missing indicator is another common approach that uses an extra variable to indicate the missingness of a variable.[164] Although this method allows the use of the whole sample for analysis, it essentially treats all missing values as the same, which is an unrealistic assumption and ignores the variation in the missing values.[161, 165] Similarly, some simple imputation methods without modelling that replace all missing values of a variable with one value defined by a certain rule, such as 'last observation carried forward' and 'simple mean imputation',[165, 166] are based on assumptions not scientifically justified and artificially reduce the variation in the imputed variables.[161, 165]

Compared to these approaches, the model-based imputation approach, which imputes the missing values of a variable with the observed data of other variables through an imputation model, has the advantage of leveraging all information in the data and taking account of the variation in the missing values. Such an imputation approach is based on the assumption of missing at random (MAR), meaning that any potential differences between the missing values and the observed values of a variable can be explained by other variables observed in the study.[167] Although missing data could be potentially missing not at random (MNAR), there is currently no way to distinguish between the patterns of MAR and MNAR. The MAR assumption is widely believed to be the most common data missing mechanism in clinical and epidemiological studies and is more reasonable than assuming one identical value for all missing data in other approaches above.[165, 168] It has also been argued that MAR is an assumption established to facilitate the use of the imputation methods rather than to demonstrate a true property of the data.[167]

Early model-based imputation methods, such as 'conditional mean imputation', perform only single imputation to replace each missing value with one plausible value estimated by the imputation model.[165] A major drawback of single imputation is that it can't incorporate the potential error and uncertainty in the imputation by only one estimation.[161] Multiple imputation was further developed to overcome this limitation and improve the results of data imputation. This method has been widely used and validated in observational studies with data from healthcare records, including many previous studies using data from the CPRD.

2.3.3.2 Key steps and considerations in multiple imputation

Multiple imputation aims to create multiple completed versions of the original incomplete data to enable analyses of the whole sample and allow for uncertainty about the missing values.[167] The basic procedure includes three main steps: imputation of missing values, analysis of completed data, and combination of analysis results.[169]

In the first step, each missing value in a variable is replaced by multiple plausible values that are drawn from a specified distribution for the variable. The distribution of the variable can be specified by a model with the observed data of other variables, which is called the imputation model. In this way, several completed versions of the original dataset are created, which contain the same observed data and the different imputed data of the missing values. In the following step, the same analyses are performed in each completed dataset to obtain estimates of the desired parameters. These estimates from each dataset will be different, as the imputed data in each dataset are different. The final step is to combine the different estimates from all datasets, usually following the Rubin's rules that calculate the average of the estimates and take account of the variations within and between the datasets.[167] For certain parameters, the estimates need to be transformed

before using the Rubin's rules or other robust combination approaches could be used, such as obtaining the median and the interquartile range (IQR) of the estimates.[170]

Specification of an imputation model with necessary variables is crucial for a proper imputation. For a time-to-event outcome in this thesis, it has been recommended to include a variable for the cumulative hazard of the outcome calculated by the Nelson-Aalen estimator to reduce potential bias in the imputation.[171] When using data from healthcare records, there is often more than one variable with missing data.[172] The multivariate imputation by chained equations (MICE) approach imputes the missing values on a variable-by-variable basis, with different imputation models specified for each variable.[173] This allows more flexibility in building the imputation models, which may lead to better validity of the imputed data.[174] The MICE approach has been shown to outperform other approaches of imputing multiple variables, such as the joint modelling, and is more robust if the imputation models are misspecified.[175, 176]

Chapter 3 Associations between Statins and Adverse Events in Primary Prevention of Cardiovascular Disease: Systematic Review and Meta-analyses

3.1 Introduction

3.1.1 Importance of understanding statin safety in primary prevention

The use of statins for primary prevention of cardiovascular disease (CVD) in patients without previous cardiovascular events has been controversial, with the debate focusing on the balance between benefits and harms of statin treatment in these patients.[37, 41] Compared to patients of secondary prevention, the baseline risk of CVD is generally lower in patients of primary prevention, leading to a smaller absolute reduction in the risk of cardiovascular events by statin treatment that is more likely to be outweighed by the increase in the risk of potential adverse events.[32] Unlike the well-studied benefits of statins, more uncertainty remains in their potential harms, which needs particular examination in patients of primary prevention in order to better understand the controversial benefit to harm balance of statin treatment in this population.[36]

In clinical practice, given the smaller benefits of treating patients for primary prevention, one may try to minimise the risk of adverse events by tailoring statin regimens with a specific drug type and dose.[51] However, there is a lack of clinical guidance on this, and the current recommendations on the choice of statin regimens are only based on their lipid-lowering effects and the goal of treatment, which is probably not applicable when considering safety.[5, 7] The recommendations for high or moderate-intensity statin regimens in primary prevention have been particularly controversial due to the concerns

about safety.[177, 178] A better understanding about the potential variations of statins' adverse effects by drug type and dose would help make a choice of statin regimens for primary prevention, taking into account both the goal of treatment and the risk of potential harms.

3.1.2 Previous research and current evidence gap

Through the long history of statins, there have been a large number of clinical trials examining outcomes of statin treatment and systematic reviews have also been conducted to synthesise the evidence from statin trials.[12, 179] However, most of the previous systematic reviews focused on the efficacy of statins for lowering low-density lipoprotein cholesterol (LDL-C) and reducing cardiovascular events, mainly based on the population of secondary prevention.[11, 97, 180] This leaves the safety of statins in the population of primary prevention not well understood.

Only a few reviews examined statin safety exclusively in patients of primary prevention, which provided conflicting results about the associations between statins and adverse events based on different criteria of study inclusion and inconsistent definitions of the adverse events.[44, 181-183] In particular, the muscle-related adverse events examined in these reviews were poorly defined, with some lacking clear definitions of the outcomes and some combining a wide range of muscle-related conditions of varying severity.[44, 182] This makes the results difficult to interpret and less meaningful for clinical implication.

Very few reviews compared adverse effects between different statin drug types and the only one focusing on primary prevention did not include all statin types and omitted some trials.[97, 182, 184] Results from these reviews contained considerable uncertainty. Some reviews compared a higher dose level of statins to a lower dose level, where the stratification of dose level was based on statins' effect on LDL-C, and showed that the

higher doses may be associated with a higher risk of adverse events.[185, 186] No previous studies or reviews, to my knowledge, have examined the pharmacological dose-response relationships of the potential adverse effects of statins.

3.1.3 Objectives of this chapter

In order to better understand the benefit-to-harm balance of statin treatment for primary prevention of CVD, the main objective of this chapter was to assess the associations between statins and adverse events by systematically scrutinising and synthesising data from existing clinical trials in patients without previous cardiovascular events. This chapter also aimed to compare the potential adverse effects of different statin drug types and examine the dose-response relationships in the adverse effects, to support further development of clinical guidance on tailoring statin regimens that take into account both the goal of treatment and the risk of potential harms.

3.2 Methods

The systematic review in this chapter is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).[187] The study protocol was registered on PROSPERO, the international prospective register of systematic reviews (registration number: CRD42020169955).

3.2.1 Literature search

There are a large number of clinical trials of statins conducted over half a century and several previous systematic reviews exist. To avoid duplicate efforts and maximise the efficiency of identifying relevant studies, statin trials were first sought from six previous large-scale systematic reviews.[44, 181, 182, 184, 188, 189] These reviews were identified by systematically searching the literature database PubMed for reviews of statin treatment

published in the past ten years (the search strategy is provided in **Appendix 1**), and the reviews with broadest criteria of study inclusion were selected. The reference lists of the selected six reviews were cross-checked to ensure the completeness of the identification of potential studies.

Among the six reviews, the most recent one included trials published up to January 2018 but its selection criteria were narrower than the current review, while the broadest review covered trials published only up to March 2013.[182, 184] To supplement the coverage of these previous reviews and update with recently published studies, three literature databases: PubMed, Embase, and the Cochrane CENTRAL, were searched for statin trials published from 1 January 2013 to 1 August 2020 (the latest date when this review was updated). Systematic search strategies were created through consultation with a health sciences librarian and are provided in **Appendix 2**.

3.2.2 Study selection

Eligible studies for this review were randomised controlled trials (RCT) in adults (>18 years) without previous cardiovascular events, which compared statins with non-statin controls or compared different drug types or doses of statins. Statin treatment could be monotherapy or add-on treatment to non-pharmacological interventions, but combination therapies of statins with other medications were ineligible. Non-statin controls included placebo, usual care, and no treatment. Studies were considered eligible as long as 70% or more of the participants did not have previous cardiovascular events, to avoid excluding large trials with a small proportion of CVD patients and limit the loss of information in primary prevention patients. Studies that enrolled fewer than 100 participants or lasted for shorter than four weeks were excluded, to avoid early-phase trials designed for mechanistic research.

As the primary outcomes for this review, the included studies must report at least one of the following adverse events of interest, as introduced in Chapter 1: muscle-related problems, liver dysfunction/injury, renal insufficiency/failure, diabetes, and cataract/eye conditions. To reconcile the inconsistent definitions of muscle-related problems in statin trials and better distinguish their clinical significance, the muscle-related problems reported in the included studies were classified as muscle symptoms and muscle disorders, which were examined separately. Each outcome was specifically defined based on clinical guidelines and how they were measured in the clinical trials, as the following:

- **Muscle symptoms**

Myalgia (muscle pain), muscle weakness, muscle spasms, or other non-specific muscle discomforts perceived by the trial participants, with normal blood creatine kinase (CK) level or slightly raised CK below 10 times the upper limit of normal (ULN).[67]

- **Muscle disorders**

Myopathy, myositis, or rhabdomyolysis as diagnosed in the clinical trials, or clinically significant elevation of CK to over 10ULN.[67]

- **Liver dysfunction/injury**

Elevation of liver enzymes aspartate transaminase (AST) or alanine transaminase (ALT) to over 3ULN, or liver diseases as diagnosed in the trials.[22]

- **Renal insufficiency/failure**

Proteinuria, renal function decline, or renal diseases as diagnosed in the trials.[190]

- **Diabetes**

Type 2 diabetes as diagnosed in the trials.[24]

- **Eye conditions**

Cataracts or cataract surgery and other eye conditions as diagnosed in the trials.[90]

To compare the potential harms with the benefits of statin treatment in the same population, three major cardiovascular events as secondary outcomes were also examined for the efficacy of statins: myocardial infarction, stroke, and death from CVD. These outcomes were defined as diagnosed in the clinical trials.

The detailed eligibility criteria are provided in **Appendix 3**. According to the eligibility criteria, titles and abstracts of the identified references were screened for potential studies and the full texts were further assessed to include eligible studies by two reviewers independently. Discrepancies between the reviewers were resolved through discussion.

3.2.3 Data extraction

Information about study designs, participant characteristics, intervention allocation, and outcome ascertainment were extracted from each of the included studies using a designed data extraction table. Study designs included study location, year of study, recruitment setting, sample size, and participant enrolment approach. Participant characteristics included mean age, the proportion of females, the proportions of ethnic groups, the proportion of CVD patients, and common comorbidities among the participants. Intervention allocation involved the specific treatments (drug type and dose) and controls, the number of participants allocated to each group, and the allocation method. Outcome ascertainment involved the duration of follow-up, the definition and measurement of each outcome, and the number of events observed in each group for each outcome.

For studies identified from previous reviews, data were extracted from the original studies by one reviewer and confirmed by the data reported in the previous reviews. For studies identified from database searching, data were extracted in duplicate by two reviewers and cross-checked to ensure accuracy.

3.2.4 Quality assessment

The risk of bias in the included studies was assessed using the Cochrane Risk of Bias Tool.[191] This assessment involved six aspects of a trial: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting. The risk of bias in each aspect was ranked as high risk, low risk, or unclear risk for each study.

The overall quality of the evidence for each outcome was assessed according to the GRADE (Grading of Recommendations Assessment, Development and Evaluation) criteria.[192] The criteria took into account five aspects of the evidence: risk of bias, inconsistency, indirectness, imprecision, and publication bias, which were further combined into the grade of the overall certainty of evidence as high, moderate, or low. The GRADE assessment was performed for the evidence of each primary and secondary outcome from the pair-wise meta-analyses and the evidence of each significant comparison between statin drug types from the network meta-analyses.[193]

Two reviewers conducted these assessments independently and discrepancies were resolved through discussion.

3.2.5 Statistical analysis

3.2.5.1 Pair-wise meta-analysis

Pair-wise meta-analysis was performed in R with the 'meta' package to assess the association between statins and each of the primary and secondary outcomes.[194] The odds ratio (OR) with the 95% confidence interval (CI) of each outcome comparing statin treatment to non-statin control was first calculated within individual studies with the extracted data.[98] In multi-arm trials, statin groups with different statin types or doses were combined into one statin group. As a pragmatic approach, the fixed-effects and random-effects meta-analyses were both performed in this review. A pooled OR with 95% CI across individual studies for each outcome was estimated by the fixed-effects model with the Mantel-Haenszel method and by the random-effects model with the DerSimonian-Laird method.[107] The Mantel-Haenszel method is the default method for the fixed-effects model recommended by the Cochrane handbook,[98] given its better performance over other methods when the event rate is low or the sample sizes of the compared groups are unequal,[195] which was the expected case in this review. In order to better estimate the absolute risk difference described below, studies with zero events in one or two of the compared groups remained in the analyses, with continuity correction by the Yates method applied to these studies.[196]

Results from the fixed-effects model were presented as the primary analysis when there was no substantial statistical heterogeneity,[105, 197] because the pooled effect from a fixed-effects model could be regarded as the estimate of the desired 'typical intervention effect' and the model requires fewer assumptions for the estimation.[98, 198] The Mantel-Haenszel method for a fixed-effects model also usually works better and leads to less bias than the DerSimonian-Laird method for a random-effects model when the event rate is low.[195] If

substantial statistical heterogeneity existed, results from the random-effects model were presented as the primary analysis since the model could better reflect the observed effects in the included studies in this case.[107] Statistical heterogeneity among individual studies was assessed using the Q test and quantified by the I^2 statistic and substantial heterogeneity was defined by p <0.05 from the Q test and I^2 >50%.[106, 199] Depending on the model presented in the primary analysis, results from the alternative model were also presented as sensitivity analysis.

Absolute risk difference (ARD) by the number of events per 10,000 person-years between the statin group and the control group for each outcome was estimated based on the pooled OR and the overall outcome incidence across all control groups in the included studies reporting the outcome, using **Equation 3.1** and **Equation 3.2**, where I_s is the overall incidence in the statin group and I_c is the overall incidence in the control group:

$$OR = \frac{I_s/(1-I_s)}{I_c/(1-I_c)} \quad \textbf{(Equation 3.1)}$$

$$ARD = I_s - I_c \quad \textbf{(Equation 3.2)}$$

In order to compare the ARD for the safety and efficacy outcomes, before the calculation of ARD above, the overall (cumulative) incidences in the control groups over different durations in the trials for different outcomes were transformed into annual incidences (incidence rate per 10,000 person-years) that were comparable among all outcomes. This transformation was based on the general relationship between cumulative incidence and incidence rate,[200] as illustrated in **Equation 3.3**, where $cumI$ is the cumulative incidence, IR is the incidence rate (i.e. annual incidence), and T is the time period that is the average study duration of the trials for each outcome:

$$cumI = 1 - e^{(-IR \times T)} \quad \textbf{(Equation 3.3)}$$

Publication bias was examined by the Harbord test of the symmetry of funnel plots when at least ten studies were included in the analysis of an outcome.[201] When publication bias was presented, defined by the asymmetry of the funnel plot detected by $p < 0.1$ from the Harbord test, sensitivity analysis was performed by excluding small studies. The robustness of the pooled results was further assessed by leave-one-out influence analysis that inspected whether the pooled results significantly relied on a single study.[202] Additional sensitivity analysis was performed by excluding the studies that involved a small proportion of CVD patients.

3.2.5.2 Network meta-analysis

Network meta-analysis (NMA) with the frequentist method based on the electrical networks and graph theory was performed in R using the 'network' package to assess the association between each type of statin drug and each adverse event and compare the potential adverse effects between different statin types.[203] Statin groups with the same drug type but different doses in a study were combined into one group. Heterogeneity among individual studies and global inconsistency across different designs of treatment comparisons for each outcome were both assessed using a generalised Q test, with a significance level of $p < 0.05$ for the test.[204] A fixed-effects consistency model was first assumed and used to calculate a pooled OR with 95% CI for each outcome and each pair of treatment comparisons, as this model is more robust when the outcome data are rather sparse and requires fewer assumptions for the calculation.[203] A random-effects consistency model was also performed as sensitivity analysis. The potential local inconsistency between direct and indirect evidence for each treatment comparison was further examined by node-splitting analysis.[115]

3.2.5.3 E_{max} model-based meta-analysis

E_{max} model-based network meta-analysis (MBNMA) was performed in R with the 'MBNMAdose' package to examine the dose-response relationships of the potential adverse effects of each statin drug.[123] With this method, the dose-specific effects of each statin drug on an adverse event estimated from an underlying network meta-analysis were fitted to an E_{max} dose-response model.[205] The E_{max} model was specified by three key parameters E_0, E_{max}, and ED_{50}, which respectively denoted the basal effect when the drug dose is zero, the asymptotic maximum drug effect, and the dose that produces half of the maximum effect.[120] Posterior means with 95% credible intervals (CrI) of the model parameters of interest (E_{max} and ED_{50}) were estimated through a Bayesian approach with the package.[123] The maximum effect E_{max} was reported as the maximum odds ratio (OR_{max}) of each statin compared to non-statin control (i.e. the dose of the statin being 0). Model convergence was assessed by the potential scale reduction factor $\hat{R}$ with convergent level defined at $\hat{R}$<1.1.[206] Dose-response curves were drawn based on the established models.

3.3 Results

3.3.1 Study inclusion and characteristics

3.3.1.1 Study inclusion

A total of 7,555 references of potentially relevant studies were identified, including 308 from the previous reviews and 7,247 from the database searches. After removing duplicates, screening by title and abstract, and further selecting by full-text, 62 eligible studies were included in this systematic review.[79, 91, 207-266] The selection process is illustrated in **Figure 3.1**.

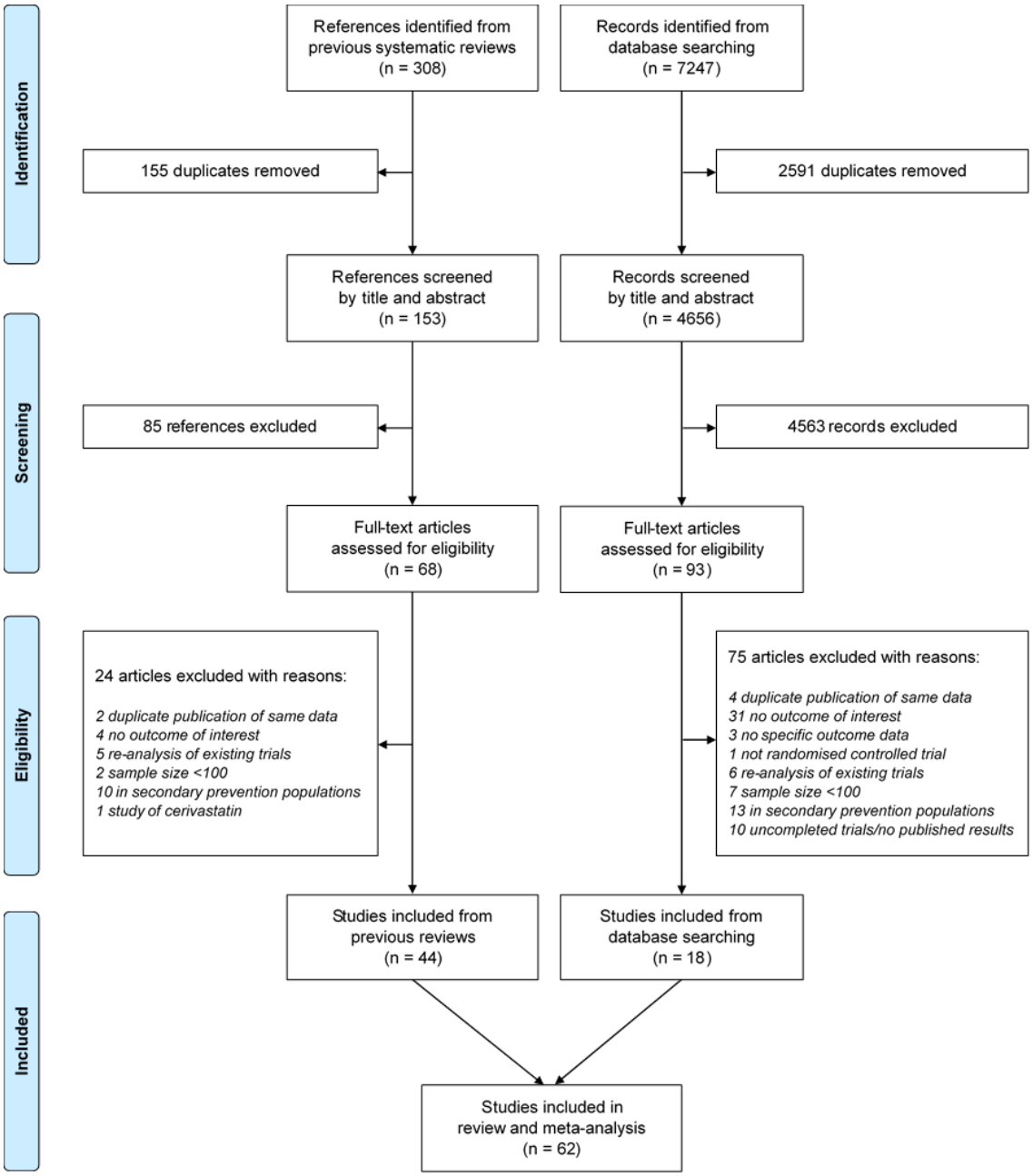


Figure 3.1 Flow diagram of study selection

3.3.1.2 Study characteristics

Characteristics of individual studies are presented in **Table 3.1**. The included studies overall enrolled 120,456 participants followed up for an average of 3.9 years. The mean age of all participants was 61 years and 40% were women. Most studies (60 studies) enrolled participants with hyperlipidaemia or dyslipidaemia and the common

comorbidities in these participants were diabetes (11 studies), asymptomatic atherosclerosis (9 studies), and hypertension (4 studies). Twenty studies included some participants with previous cardiovascular events, which comprised only 6% of the total participants.

Thirty-eight studies compared statin treatment with non-statin controls, which included placebo (35 studies), usual care (2 studies), and no treatment (1 study). Twenty-three studies compared different types of statins and 11 studies compared different doses of statins, with or without a group of non-statin controls. Seven types of statins were examined among all the included studies: atorvastatin (29 studies), fluvastatin (2 studies), lovastatin (5 studies), pitavastatin (9 studies), pravastatin (21 studies), rosuvastatin (18 studies), and simvastatin (9 studies).

The most commonly reported adverse event in these studies was muscle disorders (42 studies), mainly including diagnosis of myopathy or CK elevation to over 10ULN (35 studies). The second was muscle symptoms (40 studies), mainly including myalgia or musculoskeletal pain (34 studies). Another commonly reported adverse event was liver dysfunction (38 studies), mainly including an elevation of liver enzymes to over 3ULN (36 studies). Renal insufficiency (16 studies), diabetes (10 studies), and eye conditions (6 studies) were less commonly reported. Renal insufficiency mainly included proteinuria (9 studies) and eye conditions mainly included eye inflammation or other non-specified eye-related conditions (5 studies).

Table 3.1 Characteristics of included studies

Study ID	Study Reference	No. of Participants	Study Duration	Study Population			Statin Treatment (dose, mg/d)	Comparator
				Comorbidities with Hyperlipidaemia or Dyslipidaemia	Mean Age, year	Female, %		
1	EXCEL, 1991	8245	1y	None	56	41	Lovastatin (20/40/80)	Placebo, Statin doses
2	ACAPS, 1994	919	3y	Carotid atherosclerosis	62	48	Lovastatin (20)	Placebo
3	PMSG-D, 1994	325	4m	Diabetes	58	49	Pravastatin (10)	Placebo
4	Jacobsen et al, 1995	245	3m	None	57	31	Pravastatin (20)	Placebo
5	KAPS, 1995	447	3y	None	57	0	Pravastatin (40)	Placebo
6	WOSCOPS, 1995	6595	5y	None	55	0	Pravastatin (40)	Placebo
7	CAIUS, 1996	305	3y	Carotid atherosclerosis	55	47	Pravastatin (40)	Placebo
8	Bertolini et al, 1997	305	1.1y	None	56	55	Atorvastatin (10), Pravastatin (20)	Statin types
9	AFCAPS/TexCAPS, 1998	6605	5.2y	None	58	15	Lovastatin (20)	Placebo
10	Bak et al, 1998	215	6m	None	54	0	Pravastatin (20)	Placebo
11	Gentile et al, 2000	409	6m	Diabetes	59	32	Atorvastatin (10), Lovastatin (20), Pravastatin (20), Simvastatin (10)	Placebo, Statin types
12	ALLHAT-LLT, 2002	10355	4.8y	Hypertension	66	49	Pravastatin (40)	Usual Care
13	ALERT, 2003	2102	5.1y	Renal transplantation	50	34	Fluvastatin (40)	Placebo
14	ASCOT-LLA, 2003	10305	3.3y	Hypertension	63	19	Atorvastatin (10)	Placebo
15	ESG-L, 2003	548	3m	None	56	58	Lovastatin (10)	Placebo

16	ESG-P, 2003	538	3m	None	55	56	Pravastatin (10)	Placebo
17	Mohler et al, 2003	354	1y	Peripheral arterial disease	68	23	Atorvastatin (10/80)	Placebo, Statin doses
18	STELLAR, 2003	2431	1.5m	None	58	51	Atorvastatin (10/20/40/80), Pravastatin (10/20/40), Rosuvastatin (10/20/40/80), Simvastatin (10/20/40/80)	Statin types, Statin doses
19	CARDS, 2004	2838	3.9y	Diabetes	62	32	Atorvastatin (10)	Placebo
20	DISCOVERY, 2004	1024	3m	None	61	45	Atorvastatin (10), Rosuvastatin (10)	Statin types
21	ESG-S, 2004	1528	3m	None	56	52	Simvastatin (10)	Placebo
22	Muldoon et al, 2004	308	6m	None	54	52	Simvastatin (10/40)	Placebo, Statin doses
23	PHYLLIS, 2004	508	2.6y	Hypertension, Carotid atherosclerosis	58	60	Pravastatin (40)	Placebo
24	PREVEND-IT, 2004	864	3.8y	Microalbuminuria	51	35	Pravastatin (40)	Placebo
25	BELLES, 2005	614	1y	None	64	100	Atorvastatin (80), Pravastatin (40)	Statin types
26	COMETS, 2005	396	1.5m	Metabolic syndrome	58	36	Atorvastatin (10), Rosuvastatin (10)	Placebo, Statin types
27	CORALL, 2005	263	4.5m	Diabetes	60	54	Atorvastatin (20), Rosuvastatin (10)	Statin types
28	HYRIM, 2005	568	4y	Hypertension	57	0	Fluvastatin (40)	Placebo
29	URANUS, 2005	469	4m	Diabetes	64	43	Atorvastatin (10), Rosuvastatin (10)	Statin types
30	ARIES, 2006	774	1.5m	None	55	65	Atorvastatin (10/20), Rosuvastatin (10/20)	Statin types, Statin doses
31	ASPEN-Primary, 2006	1905	4y	Diabetes	61	38	Atorvastatin (10)	Placebo
32	ATOROS, 2006	120	6m	None	53	44	Atorvastatin (20), Rosuvastatin (10)	Statin types

33	MEGA, 2006	7832	5.3y	None	58	68	Pravastatin (10)	No Treatment
34	Schmermund et al, 2006	467	1y	Coronary atherosclerosis	62	59	Atorvastatin (10/80)	Statin doses
35	ANDROMEDA, 2007	509	4m	Diabetes	62	39	Atorvastatin (10), Rosuvastatin (10)	Statin types
36	Bone et al, 2007	604	1.1y	None	59	100	Atorvastatin (10/20/40/80)	Placebo, Statin doses
37	Lewis et al, 2007	326	9m	Chronic liver disease	50	48	Pravastatin (80)	Placebo
38	METEOR, 2007	981	2y	Atherosclerosis	57	40	Rosuvastatin (40)	Placebo
39	JUPITER, 2008	17802	5y	Elevated C-reactive protein	66	38	Rosuvastatin (20)	Placebo
40	RCASS, 2009	227	2y	Cerebral atherosclerosis	63	67	Simvastatin (20)	Placebo
41	ASTRONOMER, 2010	269	3.5y	Aortic stenosis	58	38	Rosuvastatin (40)	Placebo
42	Eriksson et al, 2011	352	3m	None	60	32	Pitavastatin (4), Simvastatin (40)	Statin types
43	PATROL, 2011	302	4m	None	62	65	Atorvastatin (10), Pitavastatin (2), Rosuvastatin (2.5)	Statin types
44	Ghia et al, 2013	119	3m	None	54	43	Atorvastatin (10/20)	Statin doses
45	Stender et al, 2013	942	3m	None	70	56	Pitavastatin (1/2/4), Pravastatin (10/20/40)	Statin types, Statin doses
46	STOMP, 2013	420	6m	None [#]	44	51	Atorvastatin (80)	Placebo
47	J-PREDICT, 2014	1269	5y	Impaired glucose tolerance [#]	56	38	Pitavastatin (1)	Placebo
48	LISTEN, 2014	1018	1y	Concurrent diabetes	66	54	Atorvastatin (10), Rosuvastatin (5)	Statin types
49	Sponseller et al, 2014	328	3m	None	58	50	Pitavastatin (4), Pravastatin (40)	Statin types

50	Nakagomi et al, 2015	146	1y	None	66	47	Atorvastatin (5), Pitavastatin (1)	Statin types
51	HOPE-3, 2016	12705	5.6y	None	66	46	Rosuvastatin (10)	Placebo
52	Patil et al, 2016	100	2m	None	60	49	Atorvastatin (20), Pitavastatin (4)	Statin types
53	Chen et al, 2018	180	6m	Cerebral atherosclerosis	61	45	Atorvastatin (20), Pravastatin (20), Rosuvastatin (10), Simvastatin (40)	Statin types
54	EMPATHY, 2018	5042	5y	Diabetic retinopathy	63	52	Atorvastatin (7.6/13.1) *	Statin doses
55	INTREPID, 2018	252	1.1y	HIV infection	50	14	Pitavastatin (4), Pravastatin (40)	Statin types
56	Liu et al, 2018	180	6m	Atherosclerosis	51	45	Atorvastatin (20), Pravastatin (20), Simvastatin (20)	Usual Care, Statin types
57	BALANCE, 2019	193	6m	Diabetes	56	50	Rosuvastatin (5)	Placebo
58	METEOR-China, 2019	540	2y	Atherosclerosis	60	56	Rosuvastatin (20)	Placebo
59	Peng et al, 2019	150	1y	Renal artery atherosclerosis	64	38	Rosuvastatin (5/10)	Statin doses
60	TRACE RA, 2019	3002	5y	Rheumatoid arthritis	61	74	Atorvastatin (40)	Placebo
61	Moroi et al, 2020	622	5y	None	65	46	Atorvastatin (10), Pitavastatin (2)	Statin types
62	Thongtang et al, 2020	150	3m	Diabetes	59	72	Atorvastatin (40), Simvastatin (20)	Statin types

ID: identification; y: year; m: month; mg/d: milligram per day.

STOMP and J-PREDICT trials did not use hyperlipidaemia or dyslipidaemia as a criterion of patient enrolment.

* A mixture of different statin types and dosages was used in the trial, which was equivalent to atorvastatin 7.6 mg/d and 13.1 mg/d in the standard and intensive arms.

3.3.1.3 Risk of bias and quality of evidence

In the six aspects of risk of bias, most of the included studies had a low or unclear risk, as shown in **Figure 3.2**. The specific assessment of each study is presented in **Appendix 4**. Seventeen studies were judged with an overall high risk of bias across the six aspects, mainly due to a high risk of bias in the blinding of participants and researchers or the blinding of outcome assessment. Most of these studies were comparisons between different statin regimens or reported clinically-confirmed outcomes. In contrast, among the 21 studies that compared statins with non-statin controls for self-perceived muscle symptoms, which might be more susceptible to the bias in blinding, only one small study with usual care control presented an unclear risk of bias in blinding,[259] while the other studies were all placebo-controlled trials with a low risk of bias in blinding.

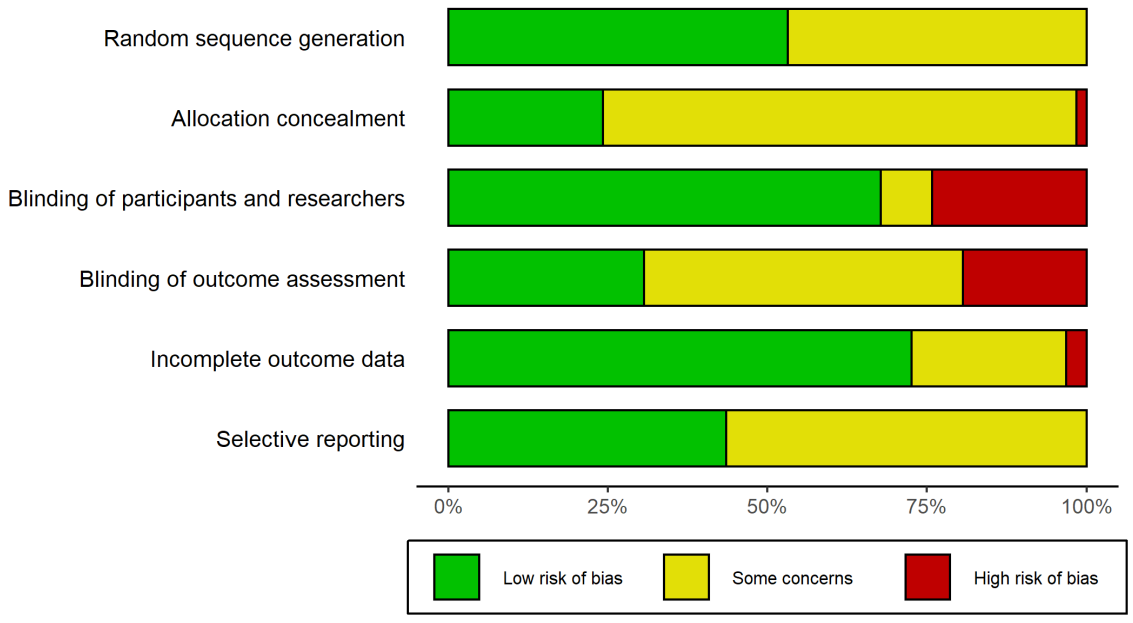


Figure 3.2 Risk of bias across all included studies

For the comparisons between statins and non-statin controls in the pair-wise meta-analysis, the overall quality of evidence for muscle symptoms, liver dysfunction, and the three major cardiovascular events (secondary outcomes) was rated as high. While the quality of evidence was moderate for clinically-confirmed muscle disorders due to the imprecision in the effect estimate and for renal insufficiency, diabetes, and eye conditions all due to the inconsistency in the outcome measurement. The specific assessment for each outcome is reported in **Table 3.2**.

For the comparisons that were found with significant differences in the network meta-analysis, the quality of evidence for the comparisons between individual statin drugs and non-statin controls was moderate or high, mainly based on direct evidence. While for the comparisons between different statin drugs, the quality of evidence was low, due to the lack of direct evidence and low quality of indirect evidence. The GRADE assessment of these comparisons is reported in **Appendix 5**.

Table 3.2 GRADE assessment of the quality of evidence for each outcome in the pair-wise meta-analyses

Outcome	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Overall Certainty of Evidence
Primary (Safety) Outcomes						
Muscle Symptoms	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
Muscle Disorders	Not serious	Not serious	Not serious	Serious †	None	⊕⊕⊕◯ Moderate
Liver Dysfunction	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
Renal Insufficiency	Not serious	Serious #	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
Type 2 Diabetes	Not serious	Serious Δ	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
Eye Conditions	Not serious	Serious §	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
Secondary (Efficacy) Outcomes						
Myocardial Infarction	Not serious	Not serious	Not serious	Not serious	None ‡	⊕⊕⊕⊕ High
Stroke	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
Death from CVD	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High

CVD: cardiovascular disease

† The analysis was underpowered to detect the between-group difference, given the very low incidences in both groups.

Four studies reported the presence of proteinuria while the other four studies reported non-specific renal insufficiency.

Δ One study was conducted in patients with impaired glucose tolerance, which resulted in statistical heterogeneity among the included studies.

§ One study reported cataracts, one reported diminished visual acuity, one reported eye inflammation, two reported non-specific eye and adnexa disorders, and the other one reported composition of different eye conditions.

‡ Publication bias was detected by examining the asymmetry of the funnel plot for the pair-wise meta-analysis of myocardial infarction, but the pooled results did not change after excluding the small studies that caused the detected bias.

3.3.2 Overall safety of statins

Thirty-eight studies that compared statins with non-statin controls were included in the pair-wise meta-analyses. The number of studies and participants involved in the analysis of each outcome varied greatly (**Table 3.3**). No substantial statistical heterogeneity was found among the studies for each outcome ($p > 0.05$ from the Q test and $I^2 < 50\%$), except for diabetes ($p = 0.04$ from the Q test and $I^2 = 50\%$ [95% CI: 0% – 77%]). The results from the random-effects model were therefore presented as the primary analysis for diabetes while the results from the fixed-effects model were used as the primary analysis for other outcomes.

Table 3.3 Included studies and the heterogeneity assessment for each outcome in the pair-wise meta-analyses

Outcome	No. of Studies	Total No. of Participants	Control Group		Treatment Group		Heterogeneity	
			No. of Events	No. of Participants	No. of Events	No. of Participants	Q Test (*p* value)	I^2 (95% CI)
Primary (Safety) Outcomes								
Muscle Symptoms	21	65304	2785	29278	3459	36026	0.45	1% (0% – 47%)
Muscle Disorders	25	85740	55	38994	70	46746	0.99	0% (0% – 0%)
Liver Dysfunction	21	54803	217	23498	406	31305	0.84	0% (0% – 23%)
Renal Insufficiency	8	32001	520	15143	597	16858	0.89	0% (0% – 23%)
Type 2 Diabetes	9	58629	1161	29311	1190	29318	0.04	50% (0% – 77%)
Eye Conditions	6	25328	234	10046	321	15282	0.85	0% (0% – 36%)
Secondary (Efficacy) Outcomes								
Myocardial Infarction	22	95148	1316	45055	996	50093	0.07	33% (0% – 60%)
Stroke	17	78473	786	39133	634	39340	0.22	20% (0% – 55%)
Death from CVD	22	95959	979	44954	836	51005	0.12	27% (0% – 57%)

CVD: cardiovascular disease; CI: confidence interval

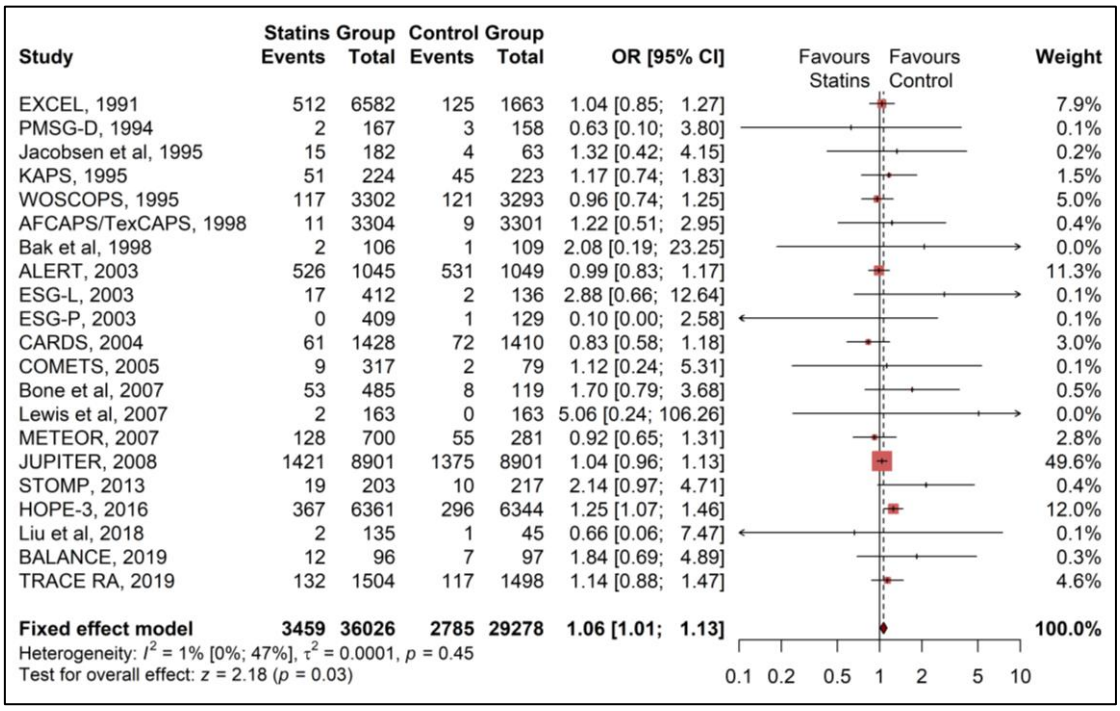
3.3.2.1 Associations of statins with adverse events

The forest plots of the pair-wise meta-analysis of the primary (safety) outcomes are presented in **Figure 3.3**. No significant publication bias was detected for any of the primary outcomes (**Appendix 6**).

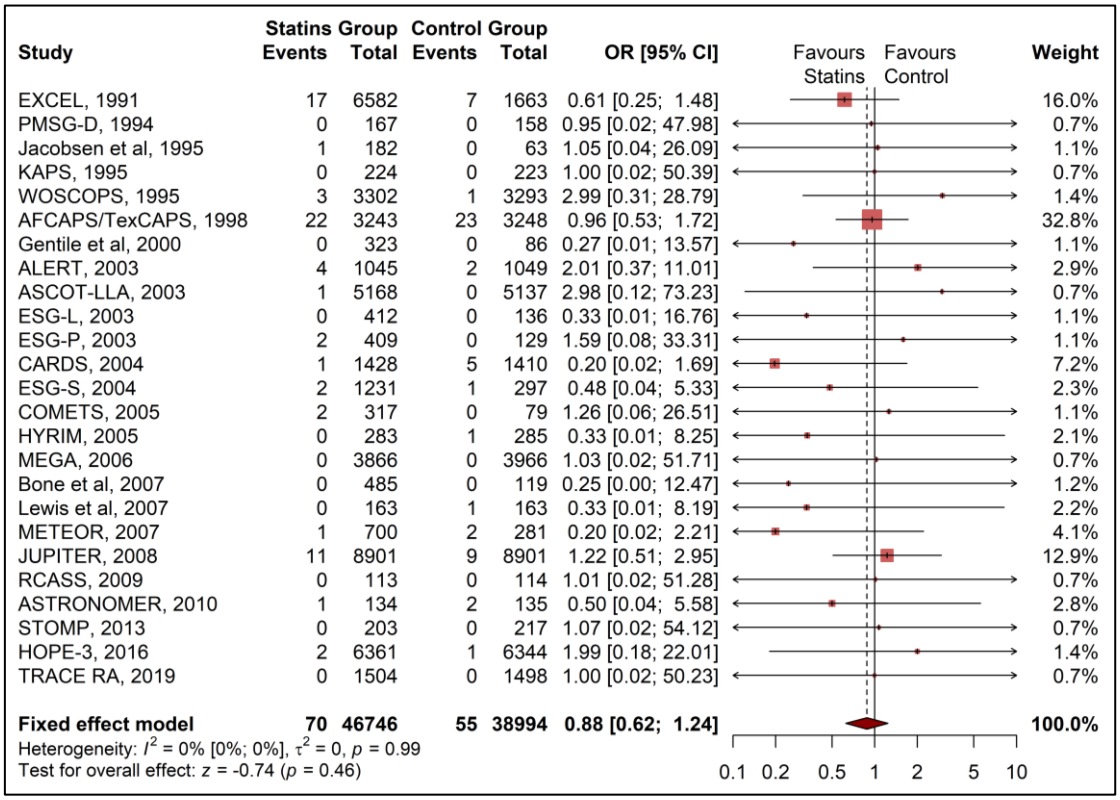
Statins were associated with a slightly increased risk of muscle symptoms (21 studies, OR = 1.06 [95% CI: 1.01 – 1.13]). As shown in the leave-one-out influence analyses (**Appendix 7**), this association was largely determined by the double-blind placebo-controlled trial HOPE-3, while the study with usual-care control and a high risk of bias in blinding had little influence on the pooled result.[91, 259] The association of statins with muscle disorders was not found significant in the included studies (25 studies, OR = 0.88 [95% CI: 0.62 – 1.24]), showing considerable uncertainty in the pooled estimate based on a small number of events observed in these studies.

Statins were also associated with an increased risk of liver dysfunction (21 studies, OR = 1.33 [95% CI: 1.12 – 1.58]), renal insufficiency (8 studies, OR = 1.14 [95% CI: 1.01 – 1.28]), and eye conditions (6 studies, OR = 1.23 [95% CI: 1.04 – 1.47]). Specifically, liver dysfunction was defined as raised liver enzymes in all the involved studies. In the influence analyses (**Appendix 7**), the association with renal insufficiency was primarily determined by the JUPITER trial,[79] which examined non-specified renal disorders, and the association with eye conditions was determined by the HOPE-3 trial,[91] which specifically examined cataract surgery. The association of statins with diabetes was not found significant in the included studies (9 studies, OR = 1.01 [95% CI: 0.88 – 1.16]), with moderate uncertainty in the pooled estimate.

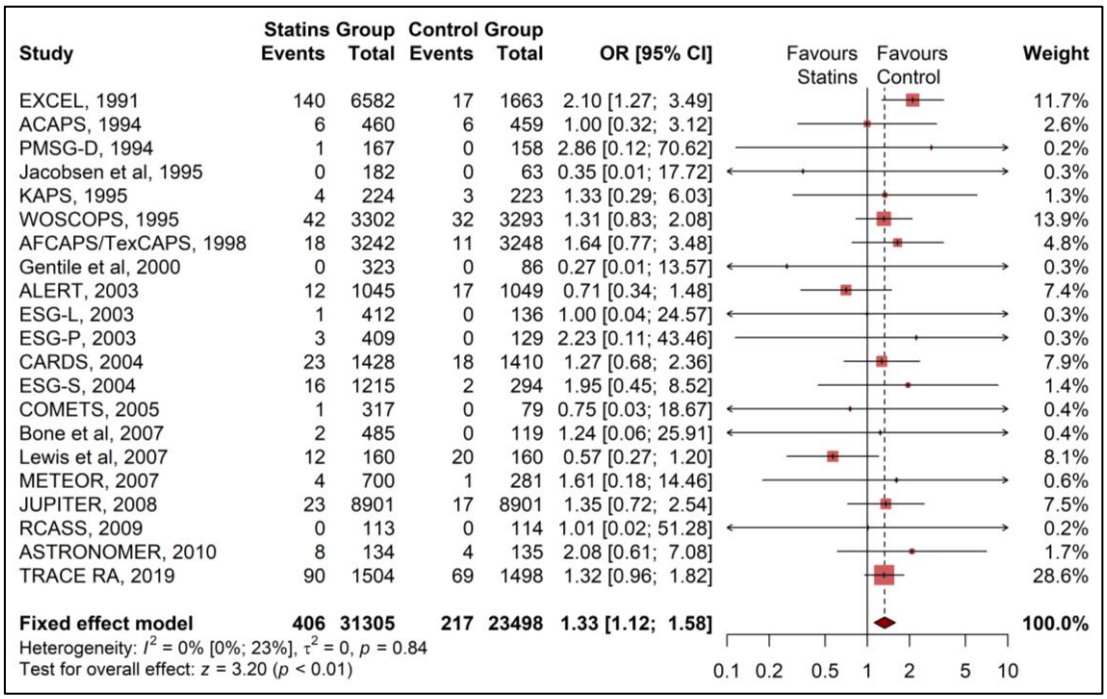
A. Muscle Symptoms



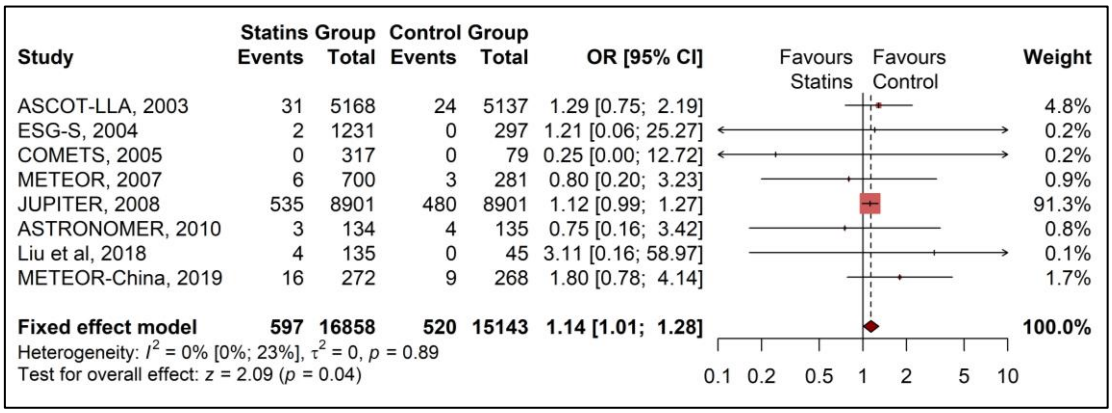
B. Muscle Disorders



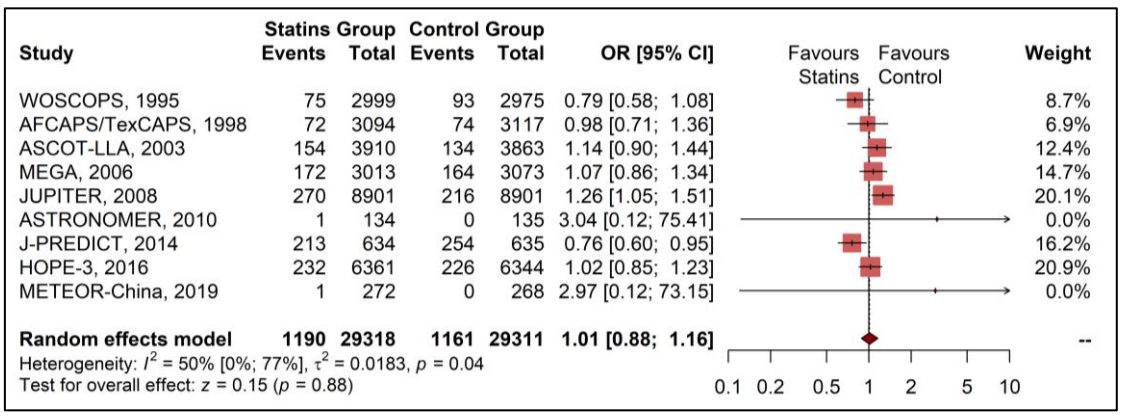
C. Liver Dysfunction



D. Renal Insufficiency



E. Type 2 Diabetes



F. Eye Conditions

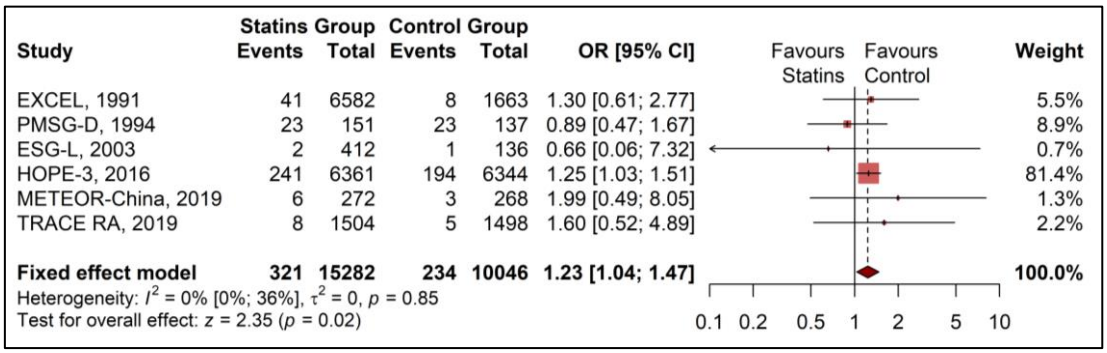


Figure 3.3 Forest plots of pair-wise meta-analyses of each primary outcome

The red squares denote the OR estimates of each primary (safety) outcome in individual studies, with their sizes corresponding to the weights of the individual estimates contributed to the pooled estimates. The horizontal lines through the squares denote the 95% CIs of the individual OR estimates, some of which with arrows at the ends indicating the lower or upper limits of the 95% CIs beyond the range of the presented axis (0.1 – 10). The diamonds denote the pooled OR estimates across the individual studies for each outcome, with the vertical dashed lines through the diamonds indicating the point estimates and the widths of the diamonds indicating the 95% CIs of the estimates. The vertical solid lines denote OR = 1, indicating a null effect. To the left of this vertical line, 'favours statins' means a lower risk in the statin group (OR<1); to the right of the line, 'favours control' means a higher risk in the statin group (OR>1). OR: odds ratio; CI: confidence interval.

In the sensitivity analyses (**Table 3.4**), the pooled results from an alternative meta-analysis model (fixed-effects model for diabetes and random-effects model for the other outcomes) were similar to the primary analyses presented above. No substantial changes were seen after excluding the studies with some CVD patients.

Table 3.4 Sensitivity analyses for pair-wise meta-analyses of the primary outcomes

Outcome	Primary Model	Alternative Model	No CVD Patients
Muscle Symptoms	1.06 (1.01 – 1.13)	1.06 (1.00 – 1.13)	1.08 (1.01 – 1.16)
Muscle Disorders	0.88 (0.62 – 1.24)	0.88 (0.62 – 1.24)	0.83 (0.55 – 1.24)
Liver Dysfunction	1.33 (1.12 – 1.58)	1.31 (1.09 – 1.56)	1.26 (1.02 – 1.57)
Renal Insufficiency	1.14 (1.01 – 1.28)	1.14 (1.01 – 1.28)	1.13 (1.00 – 1.28)
Type 2 Diabetes	1.01 (0.88 – 1.16)	1.03 (0.94 – 1.12)	1.02 (0.87 – 1.20)
Eye Conditions	1.23 (1.04 – 1.47)	1.23 (1.03 – 1.47)	1.23 (1.03 – 1.48)

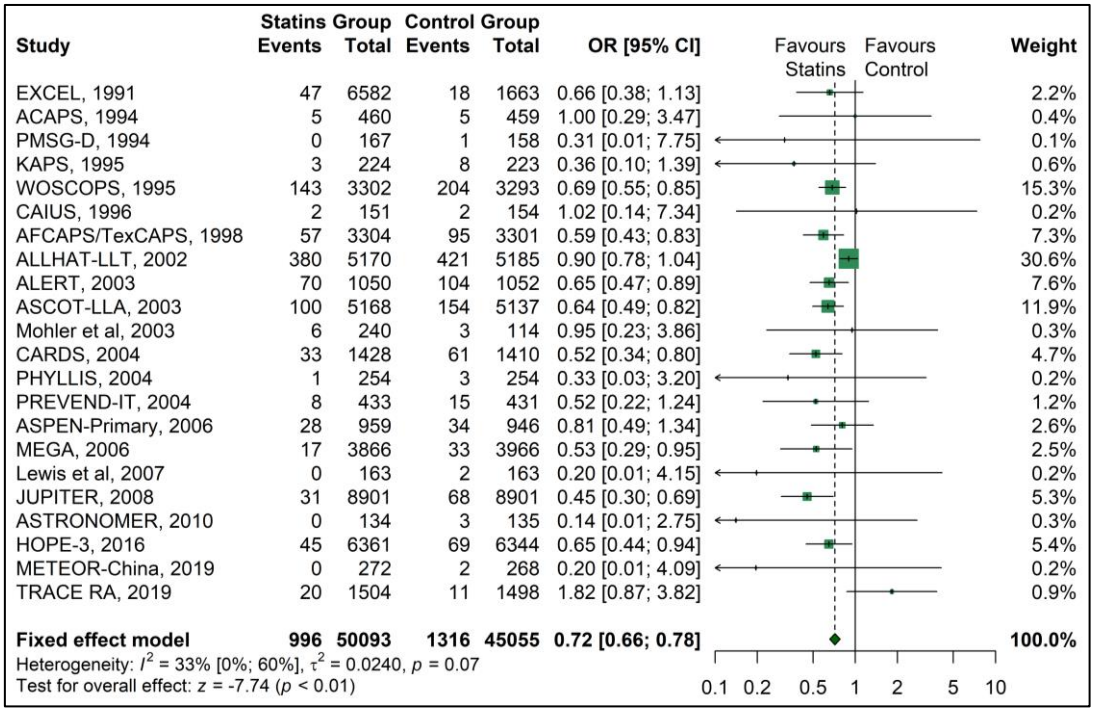
CVD: cardiovascular disease

3.3.2.2 Comparison between the beneficial and adverse effects of statins

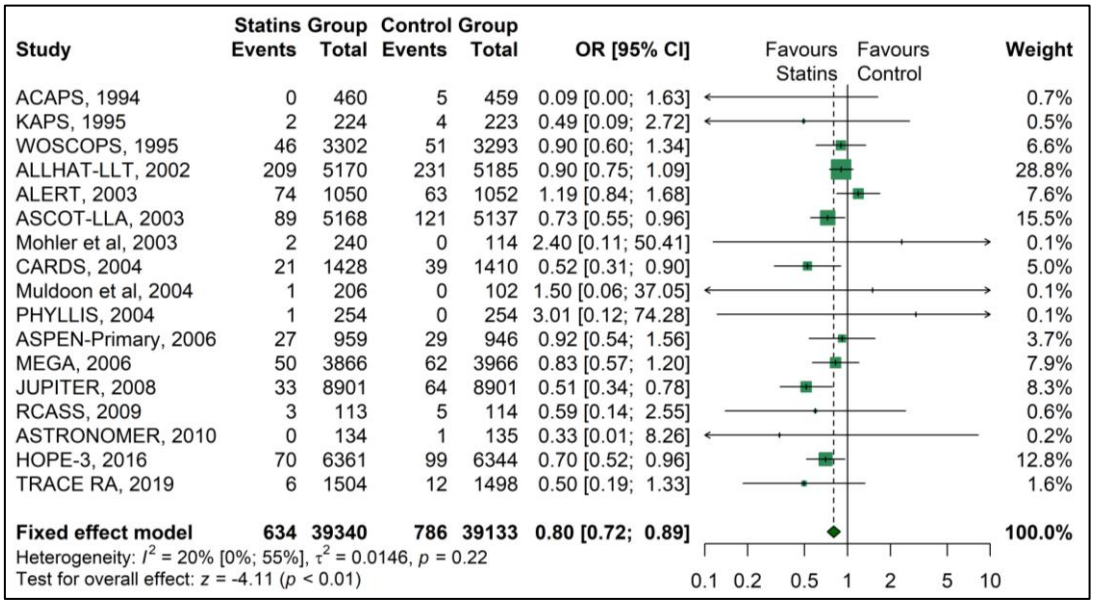
The forest plots of the pair-wise meta-analyses of the secondary (efficacy) outcomes are presented in **Figure 3.4**. Statins were found to reduce the risks of myocardial infarction (22 studies, OR = 0.72 [95% CI: 0.66 – 0.78]), stroke (17 studies, OR = 0.80 [95% CI: 0.72 – 0.89]), and death from CVD (22 studies, OR = 0.83 [95% CI: 0.76 – 0.91]). The influence analysis (**Appendix 7**) showed that the pooled results did not rely on any single study.

Publication bias was detected only in the analysis of myocardial infarction (**Appendix 6**) and the results from the sensitivity analysis (**Appendix 8**) by excluding small studies were similar to the main analysis. No substantial differences were seen in the sensitivity analysis (**Appendix 8**) by the alternative meta-analysis model (random-effects model) or by excluding the studies with CVD patients.

A. Myocardial Infarction



B. Stroke



C. Death from CVD

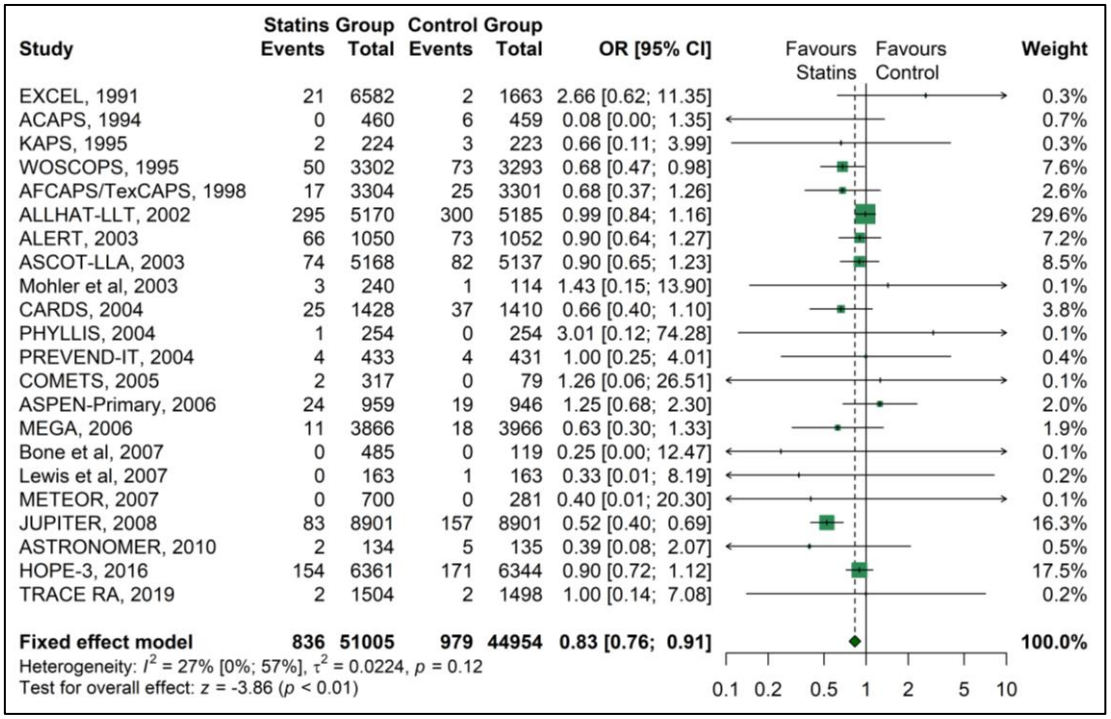


Figure 3.4 Forest plots of pair-wise meta-analyses of the secondary outcomes

The green squares denote the OR estimates of each secondary (efficacy) outcome in individual studies, with their sizes corresponding to the weights of the individual estimates contributed to the pooled estimates. The horizontal lines through the squares denote the 95% CIs of the individual OR estimates, some of which with arrows at the ends indicating the lower or upper limits of the 95% CIs beyond the range of the presented axis (0.1 – 10). The diamonds denote the pooled OR estimates across the individual studies for each outcome, with the vertical dashed lines through the diamonds indicating the point estimates and the widths of the diamonds indicating the 95% CIs of the estimates. The vertical solid lines denote OR = 1, indicating a null effect. To the left of this vertical line, 'favours statins' means a lower risk in the statin group (OR<1); to the right of the line, 'favours control' means a higher risk in the statin group (OR>1). OR: odds ratio; CI: confidence interval; CVD: cardiovascular disease.

Based on the pooled effects of statins on each safety and efficacy outcome reported above and the overall baseline incidences of these outcomes across all the control groups, the calculated ARDs (**Table 3.5**) indicated that statins could cause an excess of 15 (95% CI: 1 – 29) muscle symptoms, 8 (95% CI: 3 – 14) liver dysfunctions, 12 (95% CI: 1 – 24) renal insufficiencies, and 14 (95% CI: 2 – 29) eye conditions per 10,000 person-years. As a comparison, statins would prevent 19 (95% CI: 15 – 23) myocardial infarctions, 9 (95% CI: 5 – 12) strokes, and 8 (95% CI: 4 – 12) deaths from CVD per 10,000 person-years.

Table 3.5 Comparison between the beneficial and the adverse effects of statins

Outcome	Average Follow-up (year)	Baseline Incidence in Control Group (per 10,000 persons)		Statin Treatment Effect	
		Cumulative Incidence*	Annual Incidence#	OR (95% CI)	ARDΔ (95% CI)
Primary (Safety) Outcomes					
Muscle Symptoms	4.3	951	232	1.06 (1.01 – 1.13)	15 (1 – 29)
Muscle Disorders	4.2	14	3	0.88 (0.62 – 1.24)	0 (-1 – 1)
Liver Dysfunction	3.8	92	24	1.33 (1.12 – 1.58)	8 (3 – 14)
Renal Insufficiency	4.0	343	87	1.14 (1.01 – 1.28)	12 (1 – 24)
Type 2 Diabetes	4.9	396	82	1.01 (0.88 – 1.16)	1 (-10 – 13)
Eye Conditions	3.8	233	62	1.23 (1.04 – 1.47)	14 (2 – 29)
Secondary (Efficacy) Outcomes					
Myocardial Infarction	4.4	292	67	0.72 (0.66 – 0.78)	-19 (-23 – -15)
Stroke	4.7	201	43	0.80 (0.72 – 0.89)	-9 (-12 – -5)
Death from CVD	4.4	218	50	0.83 (0.76 – 0.91)	-8 (-12 – -4)

OR: odds ratio; ARD: absolute risk difference; CI: confidence interval; CVD: cardiovascular disease

* Cumulative incidence was the overall event rate per 10,000 persons over the average years of follow-up.

Annual incidence was the event rate per 10,000 person-years (also called incidence rate), calculated from the cumulative incidence as described in the methods.

Δ ARD was presented as the number of events increased or reduced per 10,000 person-years, calculated based on the annual incidence and the OR.

3.3.3 Differences in the adverse effects between statin drug types

A total of 58 studies were included in the comparisons among individual statin drugs and non-statin controls, forming the networks of treatment comparisons for each safety outcome illustrated in **Appendix 9**. No significant heterogeneity among the studies or inconsistency across different designs of treatment comparison for each outcome was detected, both with $p > 0.05$ from the generalised Q tests (**Appendix 10**).

Comparing individual statin types to non-statin controls (**Figure 3.5**), atorvastatin (17 studies, OR = 1.41 [95% CI: 1.08 – 1.85]) and lovastatin (5 studies, OR = 1.81 [95% CI: 1.23 – 2.66]) were both associated with an increased risk of liver dysfunction. Rosuvastatin was associated with an increased risk of muscle symptoms (13 studies, OR = 1.09 [95% CI: 1.01 – 1.16]), renal insufficiency (11 studies, OR = 1.13 [95% CI: 1.00 – 1.28]), diabetes (4 studies, OR = 1.14 [95% CI: 1.00 – 1.30]), and eye conditions (2 studies, OR = 1.26 [95% CI: 1.04 – 1.52]).

In the comparisons between different statin types (**Table 3.6**), lovastatin showed a higher risk of liver dysfunction than fluvastatin and pravastatin, and atorvastatin and rosuvastatin both showed a higher risk of diabetes than pitavastatin. No other significant differences were found between statin drug types. Most of the comparative effects were estimated with wide CIs, based on few data on certain statins and outcomes (**Figure 3.5**).

In the sensitivity analyses, the results from the random-effects model (**Appendix 11**) were similar to those from the fixed-effects model presented above. No significant inconsistency between direct and indirect evidence for each treatment comparison was detected for any of the outcomes, as shown in the node-splitting analyses (**Appendix 12**).

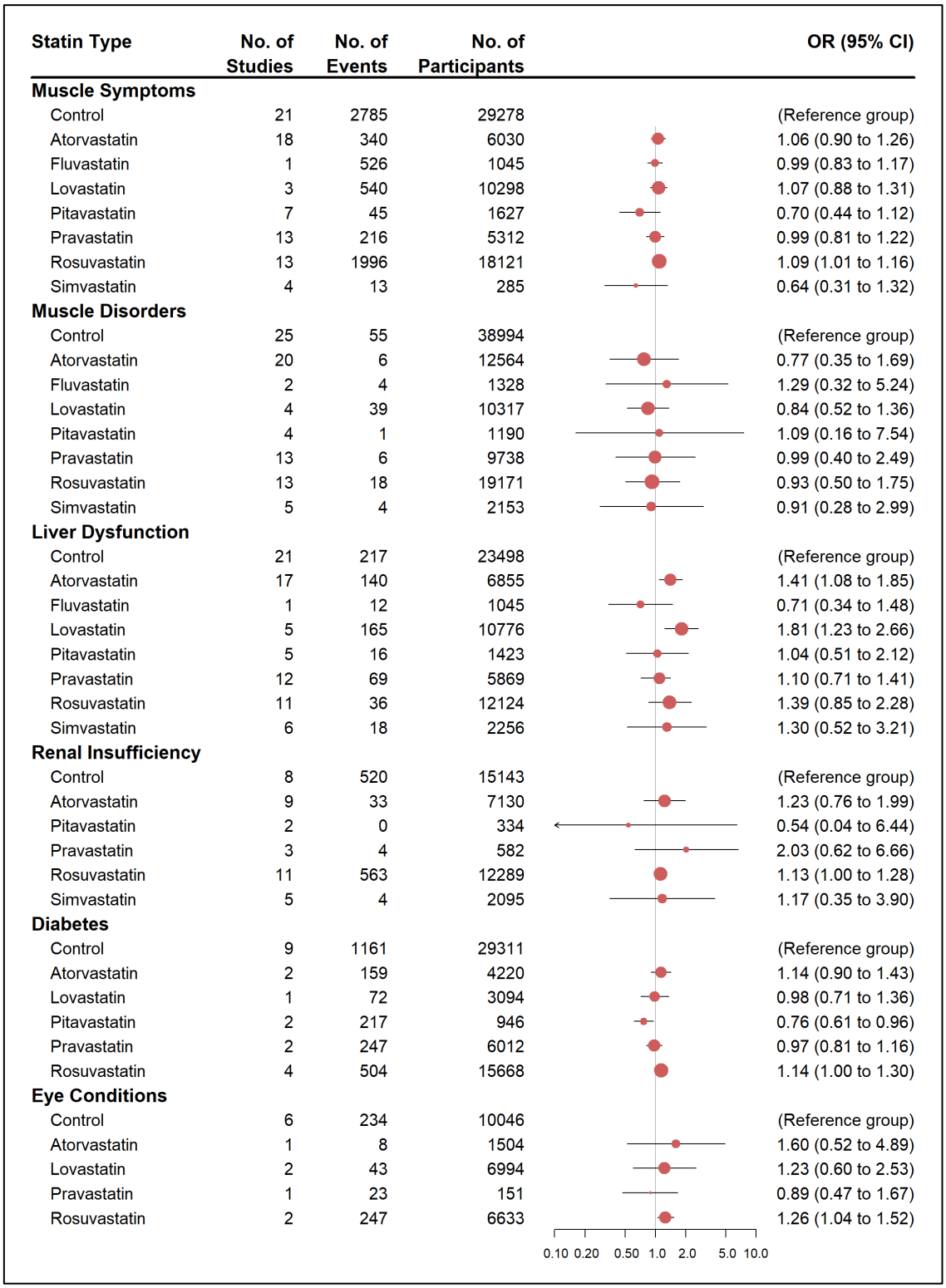


Figure 3.5 Associations of individual statins with the adverse events from network meta-analyses

The red dots denote the pooled estimates of the ORs comparing individual statins with the control for each outcome, with their sizes corresponding to the number of participants allocated to individual statins. The horizontal lines through the dots denote the 95% CIs of the pooled OR estimates, one of which with an arrow at the left end indicating the lower limit of the 95% CI beyond the range of the presented axis (0.1 – 10). The vertical grey line denotes OR = 1, indicating a null effect. OR: odds ratio; CI: confidence interval.

Table 3.6 Comparative effects on the adverse events between different statin types from network meta-analyses

Muscle Symptoms						
Atorvastatin	1.07 (0.84, 1.37)	0.99 (0.76, 1.28)	1.52 (0.94, 2.44)	1.07 (0.83, 1.39)	0.98 (0.82, 1.17)	1.66 (0.81, 3.39)
0.93 (0.73, 1.19)	**Fluvastatin**	0.92 (0.71, 1.20)	1.41 (0.86, 2.32)	1.00 (0.76, 1.31)	0.91 (0.76, 1.10)	1.55 (0.74, 3.24)
1.01 (0.78, 1.31)	1.09 (0.84, 1.41)	**Lovastatin**	1.53 (0.92, 2.54)	1.08 (0.82, 1.44)	0.99 (0.80, 1.22)	1.68 (0.79, 3.54)
0.66 (0.41, 1.06)	0.71 (0.43, 1.16)	0.65 (0.39, 1.08)	**Pitavastatin**	0.71 (0.45, 1.11)	0.64 (0.40, 1.03)	1.09 (0.51, 2.33)
0.93 (0.72, 1.21)	1.00 (0.77, 1.31)	0.92 (0.69, 1.23)	1.41 (0.90, 2.22)	**Pravastatin**	0.91 (0.73, 1.13)	1.55 (0.74, 3.22)
1.02 (0.86, 1.22)	1.10 (0.91, 1.32)	1.01 (0.82, 1.25)	1.55 (0.97, 2.48)	1.10 (0.88, 1.36)	**Rosuvastatin**	1.70 (0.82, 3.50)
0.60 (0.30, 1.23)	0.65 (0.31, 1.36)	0.60 (0.28, 1.26)	0.91 (0.43, 1.95)	0.65 (0.31, 1.35)	0.59 (0.29, 1.21)	**Simvastatin**
Muscle Disorders						
Atorvastatin	0.59 (0.12, 2.96)	0.92 (0.37, 2.30)	0.71 (0.11, 4.60)	0.77 (0.27, 2.24)	0.83 (0.36, 1.89)	0.85 (0.24, 2.99)
1.68 (0.34, 8.38)	**Fluvastatin**	1.55 (0.35, 6.80)	1.19 (0.11, 12.93)	1.30 (0.24, 6.95)	1.39 (0.30, 6.44)	1.42 (0.23, 8.94)
1.09 (0.43, 2.72)	0.65 (0.15, 2.84)	**Lovastatin**	0.77 (0.10, 5.63)	0.84 (0.30, 2.36)	0.90 (0.41, 1.98)	0.92 (0.26, 3.30)
1.42 (0.22, 9.23)	0.84 (0.08, 9.17)	1.30 (0.18, 9.54)	**Pitavastatin**	1.10 (0.16, 7.46)	1.17 (0.16, 8.42)	1.20 (0.13, 10.66)
1.29 (0.45, 3.74)	0.77 (0.14, 4.09)	1.19 (0.42, 3.33)	0.91 (0.13, 6.20)	**Pravastatin**	1.07 (0.37, 3.06)	1.09 (0.27, 4.41)
1.21 (0.53, 2.77)	0.72 (0.16, 3.34)	1.11 (0.51, 2.46)	0.86 (0.12, 6.16)	0.94 (0.33, 2.70)	**Rosuvastatin**	1.02 (0.30, 3.51)
1.18 (0.33, 4.18)	0.70 (0.11, 4.42)	1.09 (0.30, 3.90)	0.83 (0.09, 7.42)	0.92 (0.23, 3.70)	0.98 (0.28, 3.34)	**Simvastatin**
Liver Dysfunction						
Atorvastatin	2.00 (0.91, 4.42)	0.78 (0.49, 1.24)	1.36 (0.67, 2.75)	1.41 (0.93, 2.15)	1.02 (0.59, 1.75)	1.09 (0.44, 2.72)
0.50 (0.23, 1.10)	**Fluvastatin**	0.39 (0.17, 0.90)	0.68 (0.24, 1.90)	0.71 (0.31, 1.60)	0.51 (0.21, 1.24)	0.54 (0.17, 1.76)
1.28 (0.80, 2.04)	**2.57 (1.11, 5.93)**	**Lovastatin**	1.74 (0.78, 3.90)	**1.81 (1.08, 3.03)**	1.30 (0.70, 2.43)	1.40 (0.52, 3.73)
0.74 (0.36, 1.49)	1.48 (0.53, 4.13)	0.57 (0.26, 1.29)	**Pitavastatin**	1.04 (0.51, 2.12)	0.75 (0.32, 1.76)	0.80 (0.26, 2.45)
0.71 (0.47, 1.08)	1.42 (0.62, 3.22)	0.55 (0.33, 0.92)	0.96 (0.47, 1.96)	**Pravastatin**	0.72 (0.40, 1.31)	0.77 (0.30, 2.01)
0.98 (0.57, 1.69)	1.97 (0.81, 4.81)	0.77 (0.41, 1.43)	1.34 (0.57, 3.15)	1.39 (0.76, 2.53)	**Rosuvastatin**	1.07 (0.39, 2.98)
0.92 (0.37, 2.29)	1.84 (0.57, 5.94)	0.72 (0.27, 1.91)	1.24 (0.41, 3.79)	1.30 (0.50, 3.38)	0.93 (0.34, 2.58)	**Simvastatin**

Renal Insufficiency				
Atorvastatin	2.29 (0.19, 28.04)	0.61 (0.18, 2.02)	1.09 (0.66, 1.78)	1.05 (0.31, 3.58)
0.44 (0.04, 5.34)	**Pitavastatin**	0.26 (0.02, 3.80)	0.47 (0.04, 5.69)	0.46 (0.04, 5.83)
1.65 (0.50, 5.50)	3.79 (0.26, 54.47)	**Pravastatin**	1.80 (0.55, 5.88)	1.73 (0.47, 6.42)
0.92 (0.56, 1.51)	2.11 (0.18, 25.28)	0.56 (0.17, 1.82)	**Rosuvastatin**	0.97 (0.29, 3.21)
0.95 (0.28, 3.25)	2.18 (0.17, 27.82)	0.58 (0.16, 2.14)	1.04 (0.31, 3.44)	**Simvastatin**
Type 2 Diabetes				
Atorvastatin	1.16 (0.77, 1.73)	**1.49 (1.08, 2.05)**	1.17 (0.87, 1.57)	0.99 (0.76, 1.30)
0.86 (0.58, 1.29)	**Lovastatin**	1.28 (0.86, 1.91)	1.01 (0.70, 1.47)	0.86 (0.60, 1.22)
0.67 (0.49, 0.92)	0.78 (0.52, 1.16)	**Pitavastatin**	0.79 (0.59, 1.05)	0.67 (0.51, 0.87)
0.85 (0.64, 1.15)	0.99 (0.68, 1.44)	1.27 (0.95, 1.70)	**Pravastatin**	0.85 (0.68, 1.06)
1.01 (0.77, 1.31)	1.17 (0.82, 1.66)	**1.50 (1.16, 1.94)**	1.18 (0.95, 1.47)	**Rosuvastatin**
Eye Conditions				
Atorvastatin	1.30 (0.34, 4.92)	1.79 (0.50, 6.48)	1.27 (0.41, 3.95)	
0.77 (0.20, 2.92)	**Lovastatin**	1.38 (0.53, 3.60)	0.98 (0.46, 2.06)	
0.56 (0.15, 2.02)	0.72 (0.28, 1.89)	**Pravastatin**	0.71 (0.37, 1.37)	
0.79 (0.25, 2.46)	1.02 (0.49, 2.16)	1.41 (0.73, 2.73)	**Rosuvastatin**	

Each cell is the odds ratio (95% confidence interval) of the treatment in the row compared to the treatment in the column for the risk of the outcome. Significant results are highlighted in red.

3.3.4 Dose-response relationships in the adverse effects of statins

All 62 studies were included in the dose-response meta-analyses. Only the effect of atorvastatin on liver dysfunction showed a significant E_{max} dose-response relationship, with the estimated maximum effect (OR_{max}) being 2.03 (95% CrI: 1.03 – 12.64) and the dose to produce half of this maximum effect (ED_{50}) being 3.84mg (95% CrI: 0.08 – 237.33mg). For other statins or other adverse effects, no significant E_{max} dose-response relationships (OR_{max} was non-significant) were identified (**Table 3.7**), for which the ED_{50} were meaningless and could not be estimated precisely as the dose-response relationships were not established.

Table 3.7 Maximum effects of each statin on the adverse events from E_{max} model-based meta-analyses

	Muscle Symptoms	Muscle Disorders	Liver Dysfunction	Renal Insufficiency	Type 2 Diabetes	Eye Conditions
Atorvastatin	1.35 (0.92 – 4.46)	0.89 (0.47 – 1.67)	2.03 (1.03 – 12.64)	1.35 (0.61 – 12.83)	1.18 (0.37 – 4.39)	1.32 (0.36 – 6.10)
Fluvastatin	1.10 (0.61 – 5.12)	1.12 (0.49 – 2.79)	1.02 (0.34 – 33.99)	No data	No data	No data
Lovastatin	1.25 (0.79 – 5.00)	0.85 (0.46 – 1.46)	3.29 (0.95 – 40.24)	No data	0.98 (0.14 – 5.63)	1.09 (0.38 – 2.83)
Pitavastatin	0.85 (0.43 – 2.01)	0.94 (0.27 – 2.62)	1.50 (0.41 – 100.85)	0.94 (0.27 – 4.70)	0.74 (0.19 – 3.34)	No data
Pravastatin	1.18 (0.79 – 4.84)	1.02 (0.58 – 1.76)	1.13 (0.59 – 14.72)	1.33 (0.60 – 4.57)	0.95 (0.25 – 3.77)	0.91 (0.26 – 3.33)
Rosuvastatin	1.26 (0.97 – 2.82)	0.92 (0.45 – 1.63)	1.61 (0.75 – 16.82)	1.39 (0.84 – 11.03)	1.16 (0.22 – 4.33)	1.36 (0.56 – 4.22)
Simvastatin	0.83 (0.44 – 2.58)	0.94 (0.44 – 2.06)	1.59 (0.60 – 21.95)	1.17 (0.48 – 5.20)	No data	No data

Each cell presents the maximum odds ratio (OR_{max}) with 95% credible interval (CrI) derived from the E_{max} model, which indicates the maximum effect of each statin on the adverse event compared with non-statin controls (i.e. the dose of the statin being 0).

The E_{max} models were derived from the dose-specific effects of individual statin drugs on each adverse event, of which the data were scarce and the estimates presented wide 95% CrIs, as seen in the dose-response curves. An example of the dose-response curves for the effects of each statin on liver dysfunction, where the only significant dose-response relationship was found on atorvastatin, is shown in **Figure 3.6**. The curves for the effects on other adverse events are presented in **Appendix 13**.

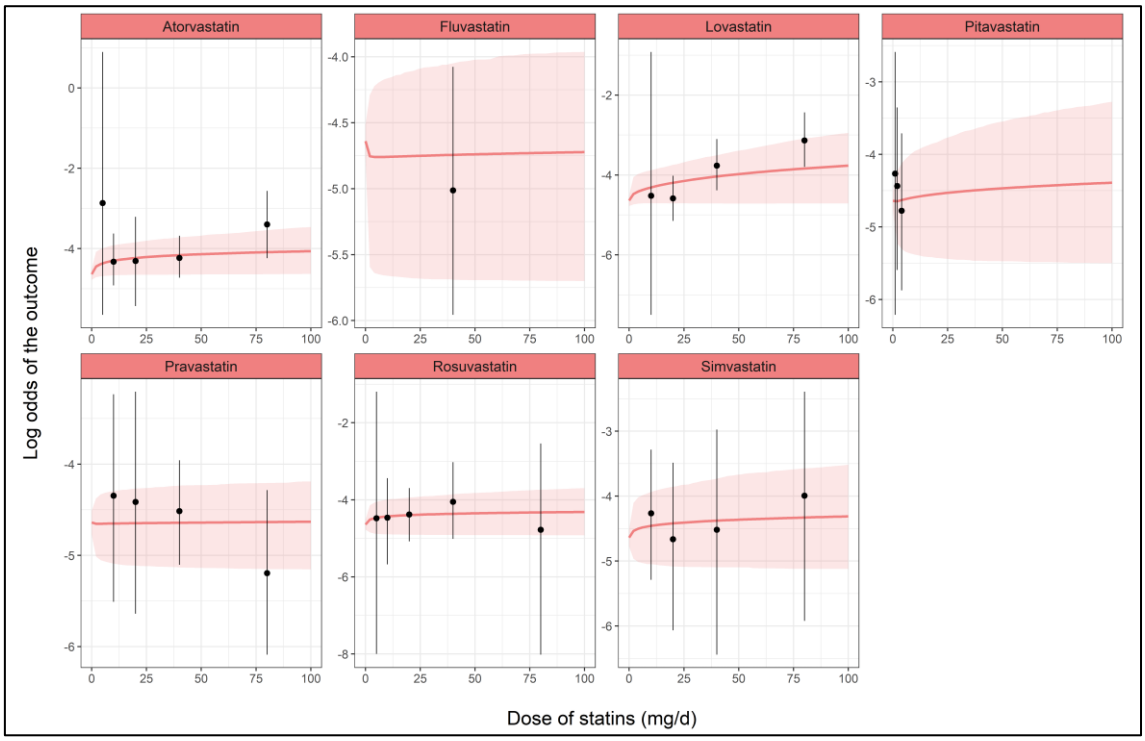


Figure 3.6 Dose-response curves for the effects of each statin on liver dysfunction from E_{max} model-based meta-analyses

The black dots denote the estimates of the dose-specific effects of each statin on liver dysfunction, in the log form of the odds comparing different doses to dose 0. The vertical lines through the dots denote the 95% credible intervals (CrIs) of the estimates. The red curves are the dose-response curves predicted by the E_{max} models of each statin fitted to their dose-specific effects. The pink areas around the curves cover the 95% CrIs of the predicted responses from the E_{max} models across the dose range. mg/d: milligram per day.

3.4 Discussion

3.4.1 Summary of key findings

This systematic review included 62 randomised controlled trials in a total of 120,456 patients of primary prevention of CVD. Statin treatment was found to be associated with a slightly increased risk of muscle symptoms, but not found to be associated with muscle disorders. Statins were also associated with higher risks of liver dysfunction, renal insufficiency, and eye conditions, but were not associated with diabetes. The absolute increase in the risk of these associated adverse events was small and did not outweigh the reduction in the risk of major cardiovascular events by statin treatment.

Analyses by drug type showed that atorvastatin and lovastatin were both associated with liver dysfunction and rosuvastatin was associated with muscle symptoms, renal insufficiency, diabetes, and eye conditions. Other statins were not found to be individually associated with any adverse events. Few significant differences were found between statin drug types, with limited data for certain statins. A potential ascending dose-response relationship was identified in the effect of atorvastatin on liver dysfunction, but the pharmacological parameters of the relationship were uncertain. No dose-response relationships for other statins and other adverse effects were identified, based on scarce data of specific doses for each statin.

3.4.2 Comparison with previous studies

3.4.2.1 Associations of statins with adverse events

Muscle-related adverse events

Most previous reviews of statin treatment in patients of primary prevention did not find an association of statins with myalgia, myopathy, or rhabdomyolysis, which defined these

specific conditions inconsistently and omitted some landmark trials due to the constrained selection criteria or being outdated.[44, 181, 183] One review found an association of statins with a broadly defined muscle-related outcome that included a wide range of muscle conditions with different severities and incidences, which made the clinical interpretation of the association not explicit.[182] In the current review, the analyses of muscle-related adverse events were based on a more comprehensive inclusion of statin trials, including those omitted in previous reviews. More importantly, the muscle-related adverse events were classified as muscle symptoms and muscle disorders based on their clinical significance and the two defined outcomes were examined separately. The findings, therefore, provided better clarification and revealed that a small increase in the risk of muscle symptoms may be attributable to statins in primary prevention, while the association between statins and muscle disorders was not specified in these patients.

It is worth noting that the attribution of muscle symptoms to statin treatment was originally identified in observational studies and has been controversial.[67] A main argument is that the higher incidence of muscle symptoms observed among statin users in clinical practice has been biased by the 'nocebo effect'.[72] This is because, unlike more serious muscle disorders and other adverse events, muscle symptoms are a self-perceived subjective outcome, which is more likely to be reported by patients when they are aware of the potential adverse effect. The current review of blind, randomised, placebo-controlled trials showed a smaller proportion of muscle symptoms attributable to statins than that found in observational studies, supporting the view that most muscle symptoms reported by statin users are not caused by statins.[12, 71] A recently published individual participant data (IPD) meta-analysis by the Cholesterol Treatment Trialists' (CTT) Collaboration also found that the risk of muscle pain or weakness increased with statin

treatment by a similarly small proportion, based on large-scale, double-blind, randomised trials in both primary and secondary prevention patients.[267]

Concerning more serious muscle disorders, although the association was not found in primary prevention patients in this review, previous reviews that included trials for secondary prevention detected an association between statins and myopathy or rhabdomyolysis based on a larger number of participants.[15, 181, 268] However, the magnitude of statins' effect on muscle disorders remains uncertain, given the small numbers of observed events and the limited sample sizes in clinical trials that have led to underpowered estimation. This association was better established in observational studies based on data from millions of patients in clinical practice, where the analyses had more statistical power to examine a small increase in the risk of a rare outcome.[25] Some of these studies, including one study with a validated design in comparison with randomised trials, suggested that the relative effect of statins on more serious muscle disorders could be considerable, although the absolute risk increase was small.[269, 270]

Other adverse events

The findings in this review supported the association of statins with a higher risk of liver dysfunction found in previous reviews for both primary and secondary prevention.[182, 186] As reported in the included clinical trials in this review, the associated liver dysfunction was mainly referred to raised liver enzymes of over 3ULN, while liver injuries or other serious liver outcomes were rarely reported.

The associations between statins and higher risks of renal insufficiency and eye conditions found in this review were also reported in previous reviews for primary prevention.[181, 182] The definitions and measurements of renal and eye-related adverse events in the included trials were highly inconsistent. The influence analyses in this

review further indicated that the associations may be limited to non-specified renal disorders and surgery-required cataracts. In contrast, previous reviews that included trials in secondary prevention patients showed inverse or non-significant associations of statins with renal function decline and cataracts.[90, 190] Overall, the current trial evidence on specific renal or eye-related adverse events of statins is too limited to draw any firm conclusions.

This review and previous reviews in primary prevention did not find an association between statins and diabetes.[90, 181] However, reviews that included trials in secondary prevention patients showed a significantly increased risk of diabetes with statin treatment and this effect increased with a higher dose of statins.[81, 185] This may be owing to different patient characteristics in primary and secondary prevention, which has led to an overall higher incidence of diabetes in secondary prevention patients and consequently a larger risk difference between the treatment and the control groups to be detected in these patients. It may be also because the trials in secondary prevention patients have examined statin regimens of higher doses that potentially have a greater effect on diabetes if the dose-responsiveness exists.

3.4.2.2 Comparison between benefits and harms of statins for primary prevention

In this review, the absolute increases in the risks of the associated adverse events did not exceed the reductions in the risks of major cardiovascular events with statin treatment, based on the incidences of these outcomes in the included trials. However, the incidences of adverse events were generally lower than those reported in observational studies in the population from clinical practice.[25, 270] Such lower incidences may be owing to some limitations of the included trials discussed below, such as the generally younger and healthier participants or the limited sample size and length of follow-up. These lower

incidences may therefore result in the smaller absolute increases in the risks of the adverse events estimated from the trial data. Nevertheless, the risk increases of most adverse events in observational studies were still small.[25] When taking into account the estimates from both clinical trials and observational studies, a previous benefit-risk assessment by modelling suggested that the benefits of statins for primary prevention outweigh their potential harms in most patients eligible for treatment.[271]

3.4.2.3 Differences in the adverse effects among statin drug types

The network meta-analysis in this review found individual associations of atorvastatin and lovastatin with liver dysfunction, which were also reported in a previous network meta-analysis based on trials in both primary and secondary prevention patients.[182, 186] These associations may be attributed to the fact that atorvastatin and lovastatin were primarily metabolised in the liver and,[272] compared to other statins, there were more data on their liver effects in the included trials. Rosuvastatin was found individually associated with muscle symptoms, renal insufficiency, diabetes, and eye conditions, of which only the association with renal insufficiency was also seen in the previous network meta-analysis of primary prevention trials.[182] These significant associations may be because the potency of rosuvastatin is stronger than other statins,[10] the doses of rosuvastatin examined in the included trials were higher, or the data on rosuvastatin were more sufficient to establish the associations. However, this did not mean that the associations with these adverse events were limited to rosuvastatin, as the data on other statins, especially fluvastatin, pitavastatin, and simvastatin, were overall insufficient to examine the individual associations.

Similarly, owing to the insufficient data on several statins, most of the estimated comparative effects between different statin drugs were imprecise. Such uncertainty was

also seen in the between-statin comparisons in previous network meta-analyses that included trials for both primary and secondary prevention.[182, 184]

3.4.2.4 Dose-response relationships of the adverse effects of statins

The only dose-response relationship identified in this review was in the effect of atorvastatin on liver dysfunction, indicating that the risk of liver dysfunction may increase with increasing doses of atorvastatin. This relationship concurred with the finding from a previous review, which showed that a higher dose of statins was associated with a higher risk of liver dysfunction, compared to a lower dose.[186] This dose-response relationship may also support the individual association of atorvastatin with liver dysfunction found in the network meta-analysis in this review. Based on the established relationship, the estimated ED_{50} for this effect was only slightly higher than the ED_{50} for atorvastatin's effect on lowering LDL-C estimated in previous studies.[49] This may result in a narrow therapeutic window for effective and safe dosing of atorvastatin, which would be much lower than the currently licenced doses of atorvastatin. However, the estimate of the ED_{50} in this review was imprecise, leaving considerable uncertainty about the pharmacological attributes of the dose-response relationship, and therefore could not be used to determine the specific therapeutic dose range of atorvastatin.

For other statins and adverse effects, no dose-response relationships were identified in this review, partially due to the scarce data on each statin of specific doses. Previous trials and reviews that included patients of secondary prevention showed a higher risk of myopathy and diabetes associated with a higher dose of statins through subgroup comparison between different dose levels, suggesting these adverse effects may be also dose-responsive.[185, 268] But the specific dose-response relationships for these effects and other potential adverse effects of statins have not been examined before.

3.4.3 Strengths and limitations

3.4.3.1 Strengths

This review was the largest systematic review of clinical trials of statin treatment in patients without a history of CVD to date. Compared to previous reviews of statin trials in primary prevention patients,[44, 181, 183, 271] this review included more complete and up-to-date trials and focused on examining the safety outcomes, which could help address the ongoing safety concerns and controversies over the use of statins for primary prevention.

By classifying the muscle-related adverse events as muscle symptoms and muscle disorders based on their clinical significance, this review better reconciled the inconsistent definitions of specific muscle conditions in statin trials and considered the clinical variety of these muscle-related outcomes. Therefore the findings clarified the risks of muscle symptoms and muscle disorders separately, which could be more helpful for addressing patients' concerns and misperceptions about potential muscle-related adverse effects of statins in clinical practice.

To my knowledge, this review was the first to examine the pharmacological dose-response relationships of the potential adverse effects of statins using data from clinical trials. Compared to the linear model that has been used to examine the dose-response relationship of statins' lipid-lowering effect,[10] the E_{max} model used in this review could better reflect the fundamental pharmacodynamics of statins and the model parameters were more clinically interpretable.[54]

3.4.3.2 Limitations

- **Limitations inherited from the included trials**

This review may have underestimated the risk of adverse events due to the designs of the included trials. The participant selection criteria in some of the trials excluded individuals who were more likely to suffer from adverse events. For example, the CARDS and METEOR trials both excluded patients with previous high serum creatinine.[227, 243] Some trials were designed with a run-in period, during which individuals with poor compliance to treatment, often due to adverse events, were excluded. For example, the ALLHAT-LLT trial excluded individuals who were found to be intolerant to statins.[218] These could result in lower incidences of adverse events compared to the incidences observed in clinical practice, particularly in the statin treatment group, and may bias the estimation of the treatment effects. The study duration of some trials was no more than 6 months, which may also lead to under-observed incidences of adverse events, especially those serious late-onset cases.

Potential bias may exist as a result of the trial conduction. Some of the trials presented a high risk of bias due to unclear methods of blinding. This may be mitigated by the fact that most of the adverse events in this review were defined by clinical tests or diagnoses. For the most subjective outcome, muscle symptoms, its association with statins was examined mainly based on the trials with a low risk of bias from blinding. However, since the included trials were primarily designed for examining treatment efficacy rather than safety, the reported adverse events in most of the trials were not pre-specified, which may be subject to incomplete or selective reporting.

The risks of the adverse events estimated in this review may not be explicit for the patients of CVD primary prevention. Some of the trials enrolled a small number of patients with established CVD, although these patients accounted for only 6% of the total participants in this review and the inclusion of these trials had little impact on the pooled results. The average age of the participants in many of the included trials was below 60

years, which may be younger than the population taking statins in clinical practice. As discussed above, individuals vulnerable to adverse events were excluded from some of the trials and these individuals could be eligible patients for statin treatment in clinical practice. Therefore, the participants included in this review may not be fully representative of the population eligible for statin treatment for primary prevention.

- **Limitations in the data analyses**

The choice of the meta-analysis model presented as the primary analysis in this review was determined by the statistical examination of the heterogeneity among the included studies, and most of the outcomes ended up being primarily analysed by the fixed-effects model. Although using statistical tests or measurements of heterogeneity to decide on the data synthesis strategy has been widely used in previous systematic reviews,[105, 197] the statistical tests and measurements of heterogeneity have limitations and may not necessarily reflect the actual heterogeneity.[106, 273] Others have advocated the importance of taking into account clinical and methodological heterogeneity among individual studies and suggested that the random-effects model is based on a more realistic assumption.[107] In this review, the clinical and methodological heterogeneity among the included studies was seen in their different participant characteristics, examined statin regimens, and trial designs and conduction. This indicates that the random-effects model may be more appropriate to pool these studies as the primary analysis for all outcomes, although the heterogeneity was not manifest in the statistical examinations and the results from both models turned out similar. It may be also a better practice to pre-specify the meta-analysis model based on the clinical assumption of heterogeneity, instead of determining the model based on the statistical examination, which is essentially a post hoc decision and may be biased by subjective choice.[107] However, in this study, the primary analysis model was determined still by objective criteria, although it was after

pooling the data, and this did not bias the conclusion because the results from both models were almost the same.

The estimation and calculation in the meta-analyses contained uncertainty and may be subject to potential bias. Given the low incidences of adverse events and the limited sample sizes of the included trials even though combined, most of the meta-analyses lacked sufficient statistical power to detect the small difference in the incidence between the treatment and the control groups. This resulted in low precision of the estimates of some adverse effects and left uncertainty about their associations with statins, especially for muscle disorders that were the rarest outcome. Continuity correction was applied to many studies that had zero outcome events in at least one group, which may have introduced bias when the sample sizes of the treatment and the control groups were unequal, leading to overestimation of the adverse effects (biasing the estimates away from 1 when the true OR >1) and underestimation of the beneficial effects (biasing the estimates towards 1 when the true OR <1).[274] In the absence of time-to-event data, the incidences of the outcomes in this review were calculated from the observed event rates and the average study duration across all involved trials, which may be inaccurate and lead to uncertainty about the comparison of the absolute risk difference between the efficacy and safety outcomes.

Uncertainty of estimation was even more prominent in the network meta-analyses and the dose-response meta-analyses, due to the scarce data on specific drug types and doses. The few significant differences between drug types found in the network meta-analyses may be subject to a false discovery rate caused by the multiple tests performed within the network of treatment comparisons.[275] Given the insufficient data, the E_{max} model in the dose-response meta-analyses was simplified by assuming a common Hill coefficient (the

steepness of the dose-response curve) of 1, which may not properly reflect the sensitivity of the dose-responsiveness of the effects.[276]

3.4.4 Clinical implications

3.4.4.1 Advising on concerns about adverse events

In this review, the risk of muscle symptoms was found to be slightly increased with statin treatment by only 6%, suggesting that most of the muscle symptoms that have been commonly reported during statin treatment are unlikely to be caused by statins. This should help alleviate some concerns from statin users who have experienced muscle symptoms. In clinical practice, doctors may want to investigate other possible causes of patients' muscle symptoms and carefully differentiate from the small proportion of statin-induced symptoms, following some existing clinical guidance.[52, 67] This could help manage patients' intolerance to statins and avoid inappropriate treatment withdrawals.

For the more serious muscle disorders, the association with statins was inconclusive in patients of primary prevention based on the current trial data, but the observed incidences with or without statins were both very low in this review. This suggests that concerns about muscle disorders should not deter the initiation of statin treatment in eligible patients in most cases. For some patients who may misperceive serious muscle disorders as being common with statin treatment, perhaps due to the confusion with muscle symptoms and the misleading media coverage, it may be helpful for doctors to clarify such misperception with the current evidence to alleviate some concerns. Especially for patients who have experienced muscle symptoms, it is important to inform patients that the risk of developing more serious muscle disorders is very low for most people before making any decisions.

The increased risk of liver dysfunction with statins shown in this review supports the current clinical recommendation for performing tests of liver function before the initiation of statin treatment.[22] However, serious liver injuries were not reported in the included trials, indicating that the mild liver dysfunction defined by raised liver enzymes does not necessarily cause damage to the liver. This supports the current recommendation that statins are generally safe to be used in patients with non-critical liver diseases.[12, 22]

Unlike in secondary prevention patients, statins were not found to increase the risk of diabetes in primary prevention patients in this review, which may be owing to the different patient characteristics or the lower doses of statins used in primary prevention. The effects of statins on renal function and eye conditions remain unclear based on the current inconsistent evidence, but serious outcomes were rarely reported in clinical trials. As such, the worries about diabetes, renal insufficiency, or eye conditions probably should not hinder the use of statin in most patients for primary prevention.

3.4.4.2 Informing the trade-off between benefits and harms of treatment

Given the absolute risk increases of the adverse events and the risk reduction of the major cardiovascular events with statin treatment estimated in this review, the potential harms of statins overall do not outweigh their benefits in patients of primary prevention. As discussed above, the incidences of these adverse events are generally low and there is remaining uncertainty about their associations with statins due to the low statistical power of the analyses and the heterogeneity in the evidence. Compared to the major cardiovascular events, most of the reported adverse events are milder conditions with better prognoses and less disease burden.

Although patients' tolerance to these adverse events could vary by individual and adjusting statin treatment may be necessary depending on patients' experience, it is

important to inform patients about the current evidence, as presented in this review and other studies, for patients to better consider the trade-off between benefits and harms before making any decisions. In particular, for patients who hesitate to take up statins for primary prevention due to concerns about statin safety, it could be helpful to reassure them with the presented evidence that the benefit-harm balance of statins for primary prevention is overall favourable for initiating the treatment, in order to avoid missing the opportunity to reduce the risk of cardiovascular events.

3.4.4.3 Considerations for tailoring statin regimens

As shown in this review, currently there is no evidence to support tailoring statin regimens by drug type or dose to reduce the risk of adverse events when initiating statin treatment for primary prevention. Cautions may be taken for prescribing atorvastatin at a higher dose when considering the risk of liver dysfunction. Overall, the absence of established explicit dose-response relationships for most of the adverse effects may alleviate some concerns about the current recommendations for the use of moderate to high-intensity statin regimens in primary prevention that have aimed at the treatment target.[5, 177] The choice of which statin regimen to initiate for primary prevention may follow the clinical guidelines that take into account the goal of treatment.[5, 7] Other existing guidance may help consider switching statin drugs or lowering the dose when patients have experienced adverse events,[51, 67] but the effectiveness of such practices requires further understanding.

3.4.5 Future research

3.4.5.1 Observational studies of the associations between statins and serious adverse events

Given the limitations of the evidence from clinical trials discussed above, observational studies could provide supplementary evidence to help better understand the associations between statins and adverse events. As the incidences of most adverse events of statins are low, observational studies could be designed with a large sample size and a long-term follow-up by leveraging data from routine healthcare records, in order to have sufficient statistical power to examine any small differences in the risks between statin users and non-users and therefore more precisely estimate the effects of statins on the adverse events.[277] This would be particularly useful for examining the serious but rare adverse events, such as serious muscle disorders, liver injury, or renal failure, which lacked sufficient data in clinical trials as seen in this review.[278] These serious outcomes that are usually diagnosed with objective clinical tests and examinations are also less likely to be biased by the potential 'nocebo effects' in observational studies. Compared to the highly selective participants in clinical trials, participants in observational studies are generally more representative of the target treatment population in clinical practice.[279] These participants usually involve a wider range of patient groups with different characteristics, which provides an opportunity to explore potential variations of the adverse effects of statins by patient groups that may be useful for better-targeting statin treatment.[277]

3.4.5.2 Meta-analysis of individual participant data focusing on safety outcomes

Another approach to a finer examination of the adverse effects of statins and their potential variations by patient characteristics is the IPD meta-analysis of clinical trials.[280] More comprehensive and detailed data on safety outcomes may be sought from the

original records of the trials, which may partially address some bias due to selective reporting in publications and clarify unclear definitions of the adverse events in statin trials. With individual-level data, this approach also enables assessing the adverse effects in patients with certain characteristics.[281] The above-mentioned CTT Collaboration, as the biggest research group for IPD meta-analysis of statin trials, has been focusing on analysing the efficacy outcomes and only very recently examined the data on muscle symptoms.[267, 268] It is probably worth further efforts to examine the data on other adverse events based on the established collaboration.

3.4.5.3 Exploring the variation of adverse effects by statin regimens

The significant lack of trial data for drug comparison or dose-response relationship of statins' adverse effects presented in this review poses the need for further evidence to understand the safety profiles of different statin regimens. One possible approach may be observational studies using rich data from prescribing records in clinical practice to compare different statin drugs.[270] Previous studies have also used the data from healthcare records to examine the dose-response relationship of statins' lipid-lowering effect, which may be also applied to examining the adverse effects.[282] However, prescription data from routine practice may not be standardised or directly analysable, which may require advanced data management and data mining techniques. Such analyses may be also subject to confounding in observational data and caution is needed when interpreting the results.[279] The consequences of switching statin drugs or adjusting statin doses to reduce the risk of adverse events and manage treatment intolerance when patients have experienced suspected symptoms also need to be further studied. This may be explored by pragmatic trials, such as n-of-1 trials, conducted in clinical settings to understand the effectiveness of such practices and identify the best strategy.[283]

3.5 Conclusion

Based on current evidence from randomised controlled trials, statins were associated with small increases in the risk of self-perceived muscle symptoms, minor liver dysfunction, mild renal insufficiency, and eye conditions in patients without previous cardiovascular events, but their associations with muscle disorders or diabetes were not confirmed in these patients. The risk increases of the associated adverse events did not outweigh the risk reductions of major cardiovascular events associated with statin treatment, suggesting that the benefit-to-harm balance of statins for primary prevention of CVD is overall favourable for treatment. There was little evidence that these adverse effects varied by statin drug type or dose, and therefore tailoring statin regimens to address safety concerns when initiating statins for primary prevention is not currently supported.

Chapter 4 Predicting Personalised Risk of Serious Muscle Disorders in Patients Eligible for Statin Treatment: Derivation of the StatinMD Model

4.1 Introduction

4.1.1 Importance of personalising the risk of serious muscle disorders for statin treatment decision-making

Knowing an individual's baseline risk of treatment outcomes could help personalise clinical treatment decisions and support the implementation of stratified treatment strategies, which are two potential approaches towards better use of statins.[56, 63] Currently, initiation of statin treatment is determined by an individual's baseline risk of cardiovascular disease (CVD), with little consideration for the risk of potential adverse events.[5-7] Given the widespread concerns about statin safety, providing information about the personalised risk of adverse events, in conjunction with the CVD risk, could help patients and doctors better consider the trade-off between benefits and harms of treatment and avoid inappropriate treatment decisions.[30] In the meantime, stratifying statin treatment not only by the risk of CVD but also by the risk of adverse events could help better target statin treatment at the population with the most to gain and the minimum risk to suffer from the potential harm of treatment.[63]

As introduced in Chapter 1, muscle-related problems are the most commonly reported adverse event with statins and have become a major cause of poor adherence to statin treatment.[20, 66] These muscle-related adverse events could have various clinical manifestations from mild muscle symptoms with normal or slightly increased blood

creatine kinase (CK) to serious muscle disorders with significantly increased CK of over 10 times the upper limit of normal (ULN).[20, 67] As shown in the systematic review in Chapter 3 and a recent large meta-analysis of individual participant data, most of the muscle symptoms are not caused by statins.[267] In contrast, for muscle disorders, although the trial evidence in patients of primary prevention was not sufficient, as presented in Chapter 3, their associations with statins in the general population including both primary and secondary prevention patients is clearer based on evidence from other clinical trials and large observational studies.[12, 25, 284] These serious conditions usually require clinical actions or hospitalisation and in the worst cases may end up in death, which makes them a major concern for patients and doctors in statin treatment decision-making.[70, 284] Some studies have suggested that the relative effect of statins on muscle disorders could be large,[12, 25] in which circumstance an individual's baseline risk of muscle disorders becomes crucial in determining the absolute risk increase by treatment. Therefore, serious muscle disorders are considered the priority adverse event for risk personalisation in this thesis to assist statin treatment decision-making.

4.1.2 Previous research and unfulfilled needs

Unlike for the risk of CVD, there are very few tools available for predicting the personalised risk of potential adverse events of statins. A previous study developed a series of prediction models (the QStatin models) for four potential adverse events of statins, including myopathic events, liver dysfunction, acute renal failure, and cataracts.[285] In the model for myopathic events, the defined outcome included milder muscle conditions with raised CK of over 4 times the ULN, which may not be necessarily the serious muscle disorders of concern and could be more common in the population, leading to a higher predicted risk. The population where this model was derived was not considered for eligibility for statin treatment, which may limit the applicability of this

model to the individuals who are considered for statin treatment in clinical practice. This model only predicted the 5-year risk of the outcome while the CVD risk is commonly predicted for 10 years,[59, 60] which makes it difficult to directly compare the benefits to harms of statin treatment for an individual.

Another study created a risk score for myopathy with CK of over 10 times the ULN based on patients who received simvastatin in three clinical trials.[286] Since this risk score was derived from current statin users and considered the use of simvastatin only, it may not apply to assessing the baseline risk of myopathy for an individual who is under consideration for the initiation of statin treatment (potential new user) or the use of other commonly prescribed statins in clinical practice, such as atorvastatin. This risk score could be used to predict a short-term risk of myopathy in 6 months and a longer-term risk in 2 years only, which is also not directly comparable to the 10-year risk of CVD.

More importantly, these two models did not take into account the competing risk in their prediction. As explained in Chapter 2, ignoring the competing risk could lead to an overestimation of the outcome risk.[151] When predicting the adverse events of statins, the competing risk could exist because of deaths from causes other than the adverse events, which have happened before the onset of the adverse events. This competing risk could be high in the population eligible for statin treatment, who are relatively old and more likely to be multi-morbid.[59] As such, the risk of adverse events may be largely overestimated if the competing risk is ignored. Overestimation of the risk of the adverse events may further exaggerate the widespread concerns about statin safety and lead to inappropriate treatment hesitation or withdrawal, which needs to be avoided given the generally low uptake of and poor adherence to statins.[30, 35, 66]

4.1.3 Objectives of this chapter

The objective of this chapter was to derive a prognostic model (the StatinMD model) to predict the risk of serious muscle disorders in 10 years (primarily), as well as in 1 and 5 years, based on personal characteristics and medical history in the population potentially eligible for statin treatment. The model development aimed to improve the prediction accuracy by taking into account the competing risk in the risk prediction. The expected use of the StatinMD model was to better inform personalised clinical decision-making on statin treatment and support the implementation of potential stratified statin treatment strategies.

4.2 Methods

4.2.1 Data source

The development of the prediction model in this thesis was based on electronic healthcare records (EHR) from the Clinical Practice Research Datalink (CPRD), an ongoing collection of primary care data from a network of general practitioners (GP) across the UK, with linkage to other national health-related databases. These data have been demonstrated to be representative of the UK population.[287] The model derivation in this chapter specifically used the primary care data from the CPRD GOLD database, which is contributed by the GPs who use the Version EHR system. The primary care data contain information on patients' demographics, clinical diagnoses, laboratory tests, and treatment prescriptions. These data are linked by individual patient identification to the hospital episode statistics (HES) data from the National Health Service (NHS), the death registration data from the Office of National Statistics (ONS), and the index of multiple deprivation (IMD) data from the UK government administration database.

All the CPRD data used in this thesis are anonymised with encrypted patient identification and the CPRD holds ethics approval from the Health Research Authority to support research using these data.

4.2.2 Study population and cohort design

An open cohort was designed in individuals who were registered to a CPRD "up-to-standard" GP in England since 1st January 1998 (the earliest date of the required data linkage and therefore the start date of this study) and were trackable in the linked databases. Statin treatment is currently recommended for individuals with a 10-year CVD risk over 10%, 7.5%, and 5% in the UK, the US, and the European clinical guidelines respectively, based on different risk prediction models.[5-7] Statins could also be used in lower-risk patients with a high level of cholesterol and further lowering of the CVD risk threshold has been proposed for the coming UK guideline.[6, 39] Therefore, in order to consider potentially wider statin users in clinical practice, this study included males over 50 years old and females over 60, who were broadly corresponding to the population with 10-year CVD risk over 5% based on the QRISK2 model that was commonly used in the UK during the study period.[59] Participants entered the cohort on the start date of this study or the date when they turned the eligible age if it was later.

Follow-up of the participants started from their index date, defined as 12 months after they entered the cohort, to allow baseline measurement of medication predictors described below. Participants exited the cohort on the date of death, being transferred out of their GP surgery, the GP's latest upload of data, the end of the intended 10-year follow-up, or 31st December 2018 (the latest date of the available data at the time of designing this study and therefore the end date of this study), whichever occurred first. Individuals

who exited the cohort before or on their index date were excluded due to the lack of follow-up data. The cohort design is illustrated in **Figure 4.1**.

4.2.3 Outcome and competing event

The outcome to be predicted was hospitalisation or death with a diagnosis of serious muscle disorders. These diagnoses were identified according to a list (**Appendix 14**) of the International Classification of Diseases (ICD) codes for serious muscle disorders that was created by searching the ICD code dictionary and consulting with clinicians. The first event of the outcome that occurred after the index date (the start of follow-up) was ascertained from the HES inpatient data and the ONS death registration data using the ICD code list. Follow-up of the participants who experienced the outcome ended on the date of their first outcome event.

The competing event to the outcome was defined as death from other causes (without a diagnosis of muscle disorders). These deaths were ascertained from either the ONS death registration or the CPRD GOLD primary care records. Follow-up of the participants who experienced the competing event ended on the date of their death. If a participant's death was recorded in both the ONS and the GOLD data with different dates, the ONS record was used.

For other participants without the outcome or the competing event, follow-up was censored when they exited the cohort on the earliest date of the situations defined above: being transferred out of the GP, the latest data upload, the end of the 10-year follow-up, or the end of this study.

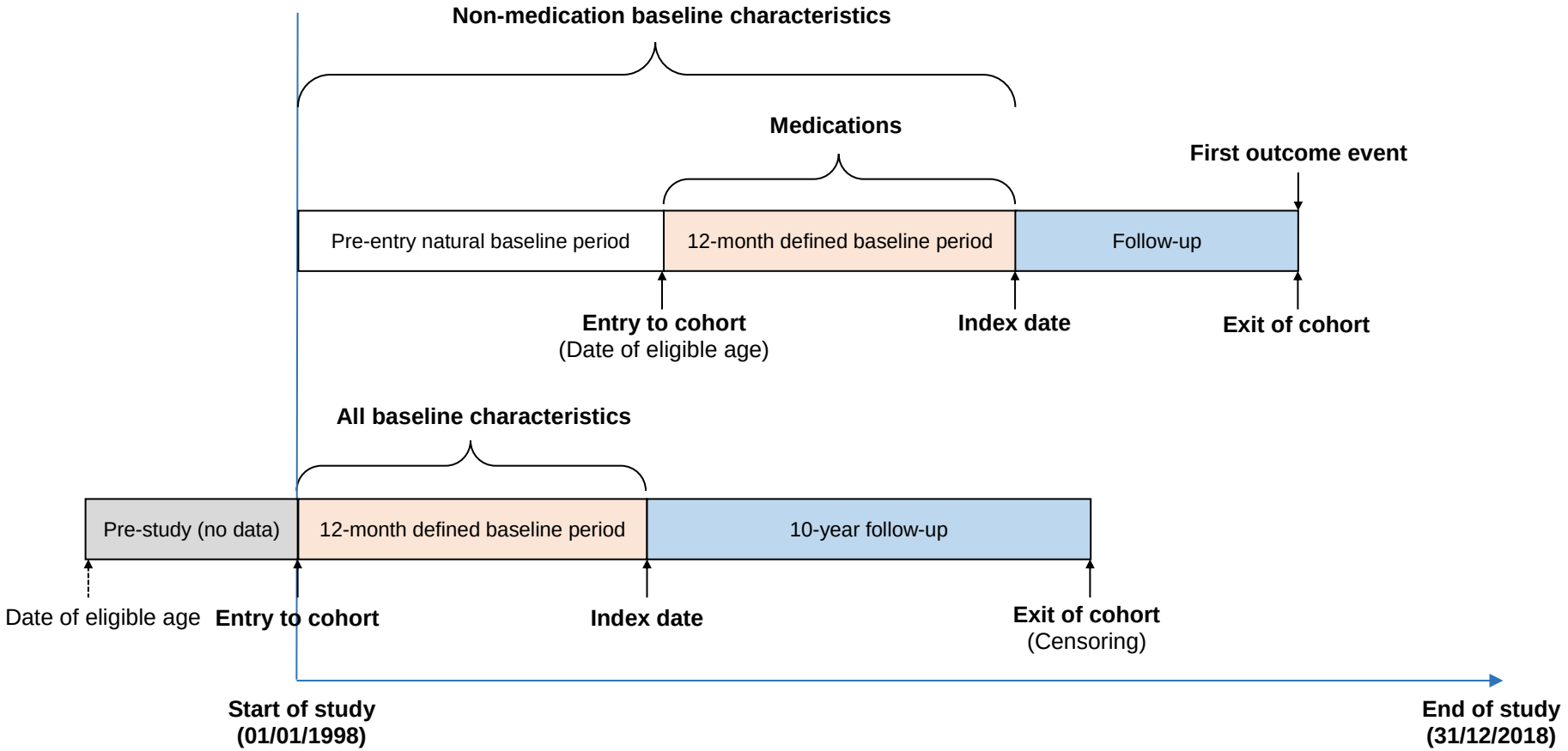


Figure 4.2 Design of the study cohort

4.2.4 Candidate predictors

4.2.4.1 Identification of candidate predictors

As a known risk factor for muscle disorders and for the purpose of assisting treatment decision-making in prospective and current statin users, the use of statins was included as one predictor in the model. Other candidate predictors in the model, including demographics, health status, comorbidities, and concomitant medications that have been reported as risk factors for muscle disorders, were identified and assessed based on the evidence of their associations with muscle disorders from the literature, their clinical relevance and importance evaluated by clinicians, and the data quality of the relevant recording in the CPRD databases.

- **Literature review**

Systematic reviews, narrative reviews, and clinical guidelines published by the end of 2020 that reported one or more risk factors of muscle disorders were sought by searching the MEDLINE/PubMed database. The titles, abstracts, and full texts of the retrieved literature were screened to select those that provided a statistical measurement, a plausible biological mechanism, or a clinical empirical explanation of the potential association between a risk factor and muscle disorders. The potential risk factors were identified from the selected literature and the relevant information about the associations was extracted. For each identified risk factor, the overall strength of the evidence supporting their associations with muscle disorders was graded as strong, moderate, or weak, based on the assessment of the adequacy of the evidence and the likelihood of the association described in the evidence. The literature searching strategy, selection process, identified risk factors from each included review, criteria for evidence assessment, and the assessment results are provided in **Appendix 15 – 19**.

- **Clinical evaluation**

The risk factors identified in the literature were further evaluated by two GPs for their clinical relevance and importance in primary care, the setting in which this prediction model would be used. The evaluation was conducted using a structured assessment form and the discrepancies between the GPs were resolved through discussion. Additional risk factors suggested by the GPs, which were not identified in the literature but believed to be clinically relevant to muscle disorders, were also considered as candidate predictors.

For comorbidities, their prevalence in the population was also taken into account, to exclude rare conditions of which few events may be observed and the data is likely to be insufficient for statistical modelling. The prevalence data were sought from the literature and disease registries. For medications, two groups were distinguished by their different ways of associating with muscle disorders, either through interaction with statins or through independent myotoxicity, based on the literature.[288, 289] Potential statin-interactive drugs were sought from the British National Formulary (BNF) clinical guidance and the myotoxic drugs were identified from the literature. A short list of drugs with a high likelihood of interaction with statins and a list of those with strong or moderate strength of the evidence supporting their association with muscle disorders were further selected, based on the BNF guidance, the relevant literature, and the opinions from the pharmacists. The clinical usage of each drug was assessed according to the prescription data from the OpenSafely database (https://www.opensafely.org/), to exclude drugs that were barely used in primary care. The specific criteria for clinical evaluation and the results are provided in **Appendix 20 – 21**.

- **Data quality assessment**

For each risk factor identified in the literature review, the quality of the relevant data in the CPRD databases was assessed in terms of availability of the relevant measurement, proportion

of missing data, and accuracy of data recording. The data quality was regarded as unacceptable if 1) the desired measurement was unavailable in the CPRD databases, 2) a large proportion of patients did not have the measurement, 3) the presence of the risk factor was not well captured in primary care, or 4) the records of the measurement were generally not accurate or reliable. This was determined based on the description in the CPRD data specifics handbook, reports of data quality from previous studies using CPRD data, exploration of a sample CPRD dataset, opinions from the CPRD data scientists, and clinical experience of recording primary care data from the GPs. The assessment results are provided in **Appendix 22**.

- **Assessment of overall eligibility as candidate predictors**

The overall assessment of the eligibility of each risk factor to be included in the model was based on the results from the literature review, clinical evaluation, and data quality assessment above. The eligible risk factors were those with strong or moderate strength of the evidence supporting their associations with muscle disorders or those with high or moderate clinical importance in primary care. Twenty-four candidate predictors were selected, including four demographics, four indicators for general health status, thirteen comorbidities, two groups of concomitant medications, and statin treatment. The statin-interactive medication included 14 drugs and the myotoxic medication included 15 drugs. The full criteria for the overall selection and the results are provided in **Appendix 23 – 24**.

4.2.4.2 Definition and ascertainment of predictor variables

The definition of the model variables for the candidate predictors and the categorisation of those multinomial variables were based on previous studies, clinical recommendations, and considerations for the statistical modelling, as described in **Table 4.1**.

Measurements of the non-medication predictors for an individual were ascertained from the CPRD GOLD primary care records, the linked HES inpatient records (supplementary data for

ethnicity only), and the IMD records (for deprivation level) before the start of the follow-up (the index date) for the individual. For ethnicity, alcohol consumption level, smoking status, and BMI level, the value was determined by the latest record before the index date and no record was regarded as missing data. For comorbidities, the presence was determined by the presence of any relevant records before the index date and no record was regarded as the absence of the comorbidities. The frailty index (FI) was calculated for each individual based on the presence of 36 deficits (comorbidities) as defined in the eFI model. Lists of medical codes that were used in the CPRD GOLD database for the relevant measurements of these predictors were compiled by searching the CPRD GOLD medical code dictionary and adopting previously published CPRD code lists from the CALIBER platform (now in the HDR UK Phenotype Library https://phenotypes.healthdatagateway.org/) and the CPRD Cambridge group (https://www.phpc.cam.ac.uk/pcu/research/research-groups/crmh/cprd_cam/codelists/). The lists were further reviewed by a GP.

The use of statins and the two groups of concomitant medications was determined by recent prescriptions ascertained from the CPRD GOLD primary care records within the 12 months before the index date. No record of the relevant prescriptions during the 12 months was regarded as no use of the medications. Code lists for these mediations were created by searching the CPRD GOLD product code dictionary and based on the BNF chapters (https://bnf.nice.org.uk/). These lists were further reviewed by a pharmacist.

For the use of statins, the specific drug was determined by the latest prescription in the 12 months. Five statin drugs are currently in clinical use in the UK, including atorvastatin, fluvastatin, pravastatin, rosuvastatin, and simvastatin. Due to a very small proportion of the cohort population prescribed fluvastatin, the number of outcome events observed in fluvastatin users was insufficient for estimating the predictor parameter for the use of fluvastatin. In consultation with pharmacists, a post hoc categorisation of statin drugs was applied to the

model derivation, combining fluvastatin with pravastatin. This was based on their similar pharmacokinetic properties (both are not significantly metabolised by CYP450 3A enzymes and therefore potentially have a lower risk of drug interactions than other statins) and clinical administration (both are short-acting statins given at bedtime).[48, 53]

Table 4.1 Definitions of candidate predictor variables

Predictor	Description	Definition
Demographics		
Age	Age on the date of cohort entry	Years between the date of cohort entry and the year of birth
Gender	Two categories: Male Female	As coded in CPRD GOLD data
Ethnicity	Four categories: White Black South Asian Mixed & Other	As defined in the 2001 UK census[290]
Deprivation	Five categories: Level 1 (the least deprived) Level 2 Level 3 Level 4 Level 5 (the most deprived)	As coded in the IMD data and defined by the population quintiles of patient-level IMD value
Health Status Indicators		
Alcohol Consumption	Three categories: Non-drinker General drinker High-risk drinker	Non-drinker: CPRD-coded non-drinker and ex-drinker General drinker: CPRD-coded trivial drinker, light drinker and other non-specified drinkers, or with a recorded weekly drinking unit ≤ 14 High-risk drinker: CPRD-coded moderate drinker and heavy drinker, or with a recorded weekly drinking unit > 14 (as defined in the NHS recommendation[291]), or with alcoholism-related diagnoses

Smoking Status	Three categories: Non-smoker Ex-smoker Current smoker	As coded in the CPRD GOLD data
Body Mass Index (BMI) Level	Five categories: Normal Underweight Overweight Obese Morbidly obese	Normal: with a recorded BMI of 18.5 – 25 or CPRD-coded normal BMI level Underweight: with a recorded BMI < 18.5 or CPRD-coded underweight conditions Overweight: with a recorded BMI of 25 – 30 or CPRD-coded overweight conditions Obese: with a recorded BMI of 30 – 40 or CPRD-coded obese conditions Morbidly obese: with a recorded BMI > 40 or CPRD-coded morbidly obese conditions (BMI levels were defined as in the NHS recommendation[292])
Frailty Index (FI)	A score of 0 – 1 describing the severity of frailty	Calculated based on 36 deficits as defined in the eFI model[293]
Comorbidities		
Cardiovascular Disease (CVD)	Presence or absence of a history of CVD	Presence: a diagnosis of coronary heart diseases, peripheral vascular diseases, heart failure, stroke, or transient ischaemic attack, or a history of cardiac surgery before the index date Absence: no record of the diagnoses defined above
Hypertension	Presence or absence of a history of hypertension	Presence: a diagnosis of primary or secondary hypertension before the index date Absence: no record of the diagnoses defined above
Diabetes	Presence or absence of a history of diabetes	Presence: a diagnosis of type 1, type 2, or secondary diabetes before the index date Absence: no record of the diagnoses defined above

Mild Liver Disease	Presence or absence of a history of mild liver disease	Presence: a diagnosis of mild alcoholic liver diseases, non-alcoholic fatty liver diseases, or viral hepatitis before the index date Absence: no record of the diagnoses defined above
Kidney Disease	Presence or absence of a history of kidney disease	Presence: a diagnosis of chronic kidney diseases or acute kidney injury, or a history of dialysis before the index date Absence: no record of the diagnoses defined above
Rheumatic Arthritis	Presence or absence of a history of rheumatic arthritis	Presence: a diagnosis of rheumatic arthritis or related complications before the index date Absence: no record of the diagnoses defined above
Previous Muscle Problems	Presence or absence of a history of muscle problems	Presence: a diagnosis of myalgia, myositis, myopathy, or rhabdomyolysis before the index date Absence: no record of the diagnoses defined above
Degenerative Joint Disorder	Presence or absence of a history of degenerative joint disorder	Presence: a diagnosis of osteoarthritis or spondylosis before the index date Absence: no record of the diagnoses defined above
Chronic Obstructive Pulmonary Disease (COPD)	Presence or absence of a history of COPD	Presence: a diagnosis of COPD or COPD-related complications before the index date Absence: no record of the diagnoses defined above
Hyperuricaemia	Presence or absence of a history of hyperuricaemia	Presence: a diagnosis of hyperuricaemia, uric acid-related calculus, uric acid-related nephropathy, or gout, or a test result of serum uric acid > 420umol/L[294] before the index date Absence: no record of the diagnoses defined above
Hypothyroidism	Presence or absence of a history of hypothyroidism	Presence: a diagnosis of symptomatic or subclinical hypothyroidism, or a test result of serum thyroid stimulating hormone > 10mU/L[295] before the index date Absence: no record of the diagnoses defined above

Vitamin D Deficiency	Presence or absence of a history of vitamin D deficiency	Presence: a diagnosis of vitamin D deficiency or related complications, or a test result of serum total vitamin D/25-hydroxy vitamin D < 25nmol/L[296] before the index date Absence: no record of the diagnoses defined above
Vitamin B_{12} Deficiency	Presence or absence of a history of vitamin B_{12} deficiency	Presence: a diagnosis of vitamin B_{12} deficiency or related complications, or a test result of serum vitamin B_{12} < 200ng/L[297] before the index date Absence: no record of the diagnoses defined above
Medications		
Statins	Five categories: No statins Atorvastatin Rosuvastatin Simvastatin Fluvastatin/Pravastatin	No statins: no record of prescription of any statins within the 12 months before the index date Other statin groups: defined by the latest prescription of statins within the 12 months before the index date
Statin-interactive Drugs	Presence or absence of the use of statin-interactive drugs	Presence: prescription of any of the selected statin-interactive drugs (see Appendix 24) Absence: no record of prescription of the drugs defined above
Myotoxic Drugs	Presence of absence of the use of myotoxic drugs	Presence: prescription of any of the selected myotoxic drugs (see Appendix 24) Absence: no record of prescription of the drugs defined above

4.2.5 Sample size calculation

The minimum sample size required for the model derivation was calculated following the steps and the criteria recommended for developing a multivariable prediction model with a time-to-event outcome.[298] The parameters required for the sample size calculation were chosen based on the information from the previous similar models and other studies.[270, 285]

The number of predictor coefficients in the model was expected to be 40. According to the definition and categorisation of the predictor variables above, a total of 38 coefficients for these variables need to be estimated, including one coefficient for each continuous or dichotomous variable and one coefficient for each category of each multinomial variable (except the reference category). An allowance of 2 more coefficients was considered for potential transformation terms of the two continuous variables. Given the statistics of the QStatin model performance,[285] the adjusted Cox-Snell R^2 statistic (R^2_{CS}) of the model to be derived was expected to be around 0.0004 and the maximum apparent R^2_{CS} was expected to be 0.0109. Based on a previous observational study in two million primary care patients in the UK, the incidence rate of muscle disorders in this study was estimated to be 1 per 10,000 person-years, with an expected average follow-up of 7 years.[270]

Based on this information, the minimum sample size was calculated to be around 900,000 individuals. This sample size ensured a small optimism in the estimation of predictor effects, defined by a global shrinkage factor (S) of the model bigger than 0.9 and a small absolute difference (less than 0.0005) between the model's apparent and adjusted Nagelkerke's R^2 statistic (R^2_N). Both of these would reduce the potential overfitting of the model that are generally caused by an insufficient sample size. The minimum sample size also fulfilled the criterion for precise prediction of the overall risk

of the outcome, with an expected absolute margin of error less than 0.0001. This sample size would yield at least 630 events of the outcome, based on the expected incidence rate and the average follow-up. As such, the number of events per predictor coefficient (EPP) in the model would be 15.7, which was larger than the commonly suggested 10 EPP as a simple rule of thumb.[299]

The actual sample size of the study cohort was much larger than this required sample size, with more observed events of the outcome.

4.2.6 Statistical analysis

4.2.6.1 Descriptive analysis

- **Baseline characteristics and follow-up**

The distribution of each candidate predictor in the cohort before the start of follow-up was reported to outline the baseline characteristics of the study population. The mean and the standard deviation (SD) were used to describe the continuous predictor variables. The frequency and the corresponding percentage were used for the categorical variables. The median and the interquartile range (IQR) were used to summarise the observed follow-up time across all individuals in the cohort.

- **Outcome incidence**

The crude incidence of the outcome (serious muscle disorders), without considering the censorship, was reported as the proportion of the total outcome events observed in the cohort population throughout the 10-year follow-up. The average incidence rate over the 10 years was reported as the number of events per 10,000 person-years.

With consideration of the censorship, the cumulative incidence of the outcome for a time period of interest was calculated using the Aalen-Johansen estimator.[300] Compared

to the classic Kaplan-Meier estimator for a single cause of failure in survival analysis,[301] the Aalen-Johansen estimator can be used when multiple causes of failure exist, such as a competing event, and avoid overestimation of the outcome incidence. This estimation was conducted in R using the package 'survival'. The estimated cumulative incidences of the outcome, with their 95% confidence intervals (CI), for 1, 5, and 10 years were reported and the cumulative incidence curve over 10 years was plotted based on the estimates.

- **Competing risk**

The crude incidence, incidence rate, and cumulative incidence for 1, 5, and 10 years of the competing event were calculated and reported through the same approaches for the outcome, as described above. To investigate the impact of the competing risk, the cumulative incidence of the outcome was re-estimated using the Kaplan-Meier method, assuming no competing events. The re-estimated incidences and the 10-year cumulative incidence curve were compared to those by the Aalen-Johansen method.

4.2.6.2 Model derivation

4.2.6.2.1 The Fine-Gray model and the model parameters to be estimated

As introduced in Chapter 2, the Fine-Gray model was chosen to derive the prognostic prediction model in this thesis, in order to reduce overestimation and directly estimate the outcome incidence in the presence of competing risk.[154] The Fine-Gray model regressed the subdistribution hazard ($sh(t)$), an instantaneous rate of the outcome occurrence at a specific time point (t) with the presence of competing events, on a set of predictors (x_i), constructed as:

$$sh(t) = sh_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i) \quad \textbf{(Equation 4.1)}$$

The model parameters to be estimated were the baseline subdistribution hazard ($sh_0(t)$) and the coefficients (β_i) of the predictor variables. The baseline subdistribution hazard is the subdistribution hazard of the outcome when all the predictor variables are set to their baseline values which are usually zero or the mean values in the population.[147] The coefficient of each predictor variable determines how the predictor contributes to the variation of the subdistribution hazard of the outcome.[132] Its exponential form is the subdistribution hazards ratio (SHR) that indicates the change of the subdistribution hazard when the value of the predictor variable is changed by 1 unit. The linear function that combines all predictor variables and their coefficients ($\sum_{i=1} \beta_i x_i$) is called the linear predictor or the prognostic index.[129]

This modelled subdistribution hazard integrated over time is the cumulative subdistribution hazard ($SH(t)$) of the outcome, which can be transformed into the cumulative incidence function ($CIF(t)$) that gives the desired absolute risk of the outcome over a time period (t).[156] Upon the transformation, an individual's predicted risk of the outcome can be expressed by the following equation, incorporating the individual's prognostic index (PI):

$$CIF(t) = 1 - \left(1 - CIF_0(t)\right)^{\exp(PI)} \quad \textbf{(Equation 4.2)}$$

The baseline cumulative incidence function ($CIF_0(t)$) is the baseline risk of the outcome when an individual has all the predictors equivalent to the baseline values (zero or the mean values) in the population.

4.2.6.2.2 Model diagnosis and specification

To more efficiently check whether the prerequisite assumptions for the Fine-Gray model are fulfilled with the data and specify any necessary transformation of the model variables, the model was first fitted with the complete-case data in the cohort participants

who had no missing data in any predictors, given the smaller sample size compared to the full cohort and no need for imputing missing data. To assist the model convergence and interpretation, the frailty index was scaled up ten times from the range of 0-1 to the range of 0-10 before fitting the model.

- **Diagnosis of model assumption: proportionality of subdistribution hazards**

The Fine-Gray model assumes proportional subdistribution hazards, which enables the coefficients of the predictor variables to be constant over time.[302] This assumption was examined using the Schoenfeld residuals plot.[303] The Schoenfeld residual is the difference between the observed value of a predictor on an individual and the individual's expected value of the predictor from the model, calculated at the time point when the individual experiences the outcome event. The Schoenfeld residuals for each predictor were plotted against the time and a smoothing line was fitted to the plot to inspect the pattern. A roughly horizontal line centred at 0 indicates that the proportionality of subdistribution hazards is established.

To obtain the Schoenfeld residuals, the model was fitted to the complete-case data using R with the package 'cmprsk' that can calculate and report the residuals from the fitted competing risk model. The complete-case sample covered the whole period of follow-up in the cohort and was believed to be sufficient for identifying any significant violation of the proportionality assumption.

- **Diagnosis of model assumption: linearity of continuous variables**

Another essential assumption in the Fine-Gray model is the linearity of the relationship between each continuous predictor variable and the logarithm (log) of the subdistribution hazard of the outcome, which allows the prognostic index in the model to be a simple linear function of all predictor variables and linked to the subdistribution hazard of the

outcome through an exponential function. This linear relationship can be inspected graphically by plotting the log SHRs, which denote the magnitudes of the change in the log subdistribution hazard, against the values or magnitudes of a predictor variable.[304] A roughly straight line suggests that the assumption of linearity is held.

In order to obtain the log SHRs for each continuous variable across the range of their values in a practical manner, the continuous variables in this study (age and frailty index) were categorised by the twenty percentiles of their values and included in the model as categorised variables. The model was fitted to the complete-case data using the package 'fastcmprsk', which does not provide the residuals as the 'cmprsk' does but accelerates the intensive computing for the competing risk model in a large dataset.[305] The log SHRs for each variable from the model were then plotted in the order of the twenty categories and a smoothing line was laid on each plot.

- **Model specification: transformation of continuous variables**

When a non-linear association between a continuous variable and the log subdistribution hazard of the outcome hazard was observed, the factional polynomial (FP) function was used to transform the continuous variable to rebuild a linear relationship with the outcome.[306] The FP function allows for multiple power transformation terms, where the power in each transformation is chosen from a set of suggested numbers ($\text{power} = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, where 0 denotes the log transformation). The FP function that fits the data best was determined by comparing the models with the more complex second-degree FP functions (includes two transformation terms of the variable) to those with the simpler first-degree FP functions (includes only one transformation term), through a closed test procedure with a significance level of $\alpha \leq 0.05$.[307]

Due to the lack of an efficient R package to implement the FP function selection procedure for a competing risk model, the selection of the optimal FP function in this study was performed using the 'mfp' package based on the Cox model that shares similar model assumptions and statistical fundaments as the Fine-Gray model. This process was also based on the complete-case data. The transformed terms of the continuous variable using the selected FP function, instead of the original form of the variable, were used in the following model estimation.

4.2.6.2.3 Model estimation and finalisation

Upon the fulfilment of the prerequisite assumptions and with the specified transformation, the model was fitted to the data of the whole cohort to estimate the model parameters. The missing data in the whole cohort were imputed using multiple imputation,[167] as introduced in Chapter 2. According to the estimates, the candidate predictors were further reviewed to finalise the inclusion of the predictors in the model. The specific steps are described below:

- **Imputation of missing data**

In this study, the values of the variable 'deprivation level' were missing in less than 1% of the cohort population and, for practical purposes, these missing data were simply imputed by replacing the missing values with the most common level of deprivation in the cohort. Four other variables (ethnicity, alcohol consumption, smoking status, and BMI level) presented larger proportions of missing data and the multivariate imputation by chained equations (MICE) was used to impute the missing data.[174]

As explained in Chapter 2, the multiple imputation of the four variables was based on the assumption of missing at random (MAR), which assumes that the potential differences between their missing values and observed values could be explained by other variables

observed in this study.[167] For ethnicity, a previous study showed that individuals with more frequent medical encounters were more likely to have their ethnicity recorded and the frequency of medical encounters may be influenced by an individual's health status, such as frailty index and comorbidities that were also recorded in this study.[308] Alcohol consumption, smoking status, and BMI level also have been shown to be more commonly recorded in individuals with a higher degree of multi-morbidity.[163] For example, alcohol consumption may be better recorded in individuals with liver diseases, smoking status may be better recorded in individuals with COPD, and BMI level may be better recorded in individuals with diabetes or CVD. Therefore it is reasonable to assume that the four variables were missing at random, conditional on other variables in this study, and the missing values could be imputed based on the observed data in the cohort.

A multinomial logistic regression model was used to impute the unordered categorical variable (ethnicity) and an ordered logistic regression model was used for each of the other three variables (alcohol consumption, smoking status, BMI level). Each imputation model included all the candidate predictors measured in this study, two variables indicating the occurrence of the outcome and the competing event, as well as two variables of the cumulative hazards of the outcome and the competing event estimated by the Nelson-Aalen estimator, in order to reduce the potential bias from imputation.[171]

Ten imputations were conducted, which has been suggested to be a sufficient number of imputations needed in most cases.[167] The imputations were performed in R using the 'mice' package that implements the MICE approach to impute multiple variables at one time.[309] To inspect the appropriateness and robustness of the imputation results, histograms were plotted to compare the distribution of the observed data and the imputed data from each imputation.

- **Model fitting in each imputed dataset**

The model with all the candidate predictors was fitted in each of the ten imputed datasets using the R package 'fastcmprsk'. The estimate and the standard error (SE) of the coefficient of each predictor as well as the corresponding SHR with 95% CI from the fitted model were reported. The standard error of the coefficient estimate was obtained using the bootstrapping technique with 100 times resampling, assisted by parallel computing with multiple cores.

Before fitting the model, the transformed variables of age were centred on their mean value in the cohort population to derive a more meaningful baseline incidence of the outcome from the model, which represented the outcome risk when an individual is at the age corresponding to the average level in the population instead of a senseless age of zero. When fitting the model in each dataset, the baseline cumulative subdistribution hazards ($SH_0(t)$) of the outcome in 1, 5, and 10 years were estimated by the Breslow method, which were further transformed into the baseline cumulative incidence functions of the outcome by the following equation:

$$CIF_0(t) = 1 - \exp\big(-SH_0(t)\big) \quad \textbf{(Equation 4.3)}$$

- **Combination of estimates across imputed datasets**

The estimates of each predictor coefficient from the ten imputed datasets were combined into a final estimate using the Rubin's rules.[167] The pooled estimate was the mean of the ten estimates from the imputed datasets and the total variance for the pooled estimate took into account both the variance within each imputed dataset and the variance between the datasets. The corresponding SHR with 95% CI for each predictor was also reported based on the pooled coefficient. The estimates of the baseline cumulative

incidence were also combined using the Rubin's rules, after a complementary log-log transformation.[170]

- **Final model: selection of final predictors**

As highlighted in Chapter 2, the inclusion of the predictors in a clinical prediction model should consider clinical relevance and purpose rather than solely rely on statistical significance, to avoid eliminating clinically meaningful predictors.[134, 135] In line with this principle, this study did not employ any statistical selection algorithm when fitting the model and, instead, the candidate predictors were further reviewed after the model was fitted. Candidate predictors with non-significant coefficients (significance level $\alpha = 0.05$) were regarded as having insufficient predicting capability for the outcome and excluded from the final model. However, those predictors that were considered clinically important and relevant in treatment decision-making could remain in the final model regardless of the statistical significance of their coefficients.

With the selected final predictors, the model was re-fitted in each imputed dataset and the estimates of the model parameters were combined across the datasets, through the same process described above.

4.3 Results

4.3.1 Population characteristics

A total of 1,785,207 individuals were included in the cohort, with a median follow-up of 6.6 years (IQR: 2.8 – 10). The mean age of all participants was 64 (SD: 11.5) and 785,238 (44%) of them were women. The majority (95%) of the participants with recorded ethnicity were identified as white. A large proportion (61%) of the participants with recorded alcohol consumption were general or high-risk drinkers, while only 18%

of those with recorded smoking status were current smokers. For the participants with recorded BMI, 62% were above the normal level, including 22% with obesity or morbid obesity. The most common comorbidities were hypertension (26%), degenerative joint disorders (19%), and cardiovascular disease (15%). A considerable proportion (20%) of the participants were prescribed at least one of the defined myotoxic drugs (other than statins). A total of 14% of the participants were recently prescribed statins, with simvastatin (8%) as the most commonly prescribed followed by atorvastatin (4%).

Excluding the participants with missing data, the complete-case subcohort included 914,457 participants, whose baseline characteristics were similar to the whole cohort. The detailed baseline characteristics of the cohort participants are presented in **Table 4.2**.

Table 4.2 Baseline characteristics of the cohort participants

Baseline Characteristics	**The Whole Cohort**	**Complete-Case Subcohort**
Total number of patients	1,785,207	914,457
Years of follow-up (median and IQR)	6.6 (2.8 – 10.0)	7.2 (3.1-10.0)
Demographics		
Age (mean and SD)	64.0 (11.5)	63.6 (10.4)
Gender		
Male	999,969 (56.0%)	488,112 (53.4%)
Female	785,238 (44.0%)	426,345 (46.6%)
Ethnicity		
White	1,193,371 (66.9%)	860,304 (94.1%)
Black	18,238 (1.0%)	14,913 (1.6%)
South Asian	22,619 (1.3%)	18,244 (2.0%)
Mixed and other	27,186 (1.5%)	20,996 (2.3%)
Missing	523,793 (29.3%)	/
Deprivation		
Level 1 (least deprived)	423,260 (23.7%)	206,265 (22.6%)

Level 2	409,934 (23.0%)	208,051 (22.8%)
Level 3	383,727 (21.5%)	194,187 (21.2%)
Level 4	314,029 (17.6%)	167,245 (18.3%)
Level 5 (most deprived)	252,628 (14.2%)	138,709 (15.2%)
Missing	1,629 (0.1%)	/
Health Status Indicators		
Alcohol consumption		
Non/Ex-drinker	279,213 (15.6%)	197,319 (21.6%)
General drinker	973,783 (54.6%)	644,619 (70.5%)
High-risk drinker	119,158 (6.7%)	72,519 (7.9%)
Missing	413,053 (23.1%)	/
Smoking		
Non-smoker	850,660 (47.7%)	499,560 (54.6%)
Ex-smoker	359,202 (20.1%)	231,468 (25.3%)
Current smoker	320,494 (18.0%)	183,429 (20.1%)
Missing	254,851 (14.3%)	/
BMI level		
Normal	495,609 (27.8%)	319,871 (35.0%)
Underweight	23,949 (1.3%)	15,344 (1.7%)
Overweight	542,577 (30.4%)	363,409 (39.7%)
Obese	279,006 (15.6%)	196,242 (21.5%)
Morbidly obese	26,274 (1.5%)	19,591 (2.1%)
Missing	417,792 (23.4%)	
Frailty index (mean and SD)	0.062 (0.064)	0.072 (0.067)
Comorbidities		
CVD	267,710 (15.0%)	150,764 (16.5%)
Hypertension	470,050 (26.3%)	292,097 (31.9%)
Diabetes	136,048 (7.6%)	99,731 (10.9%)
Mild liver diseases	4,308 (0.2%)	2,924 (0.3%)
Kidney diseases	67,540 (3.8%)	45,519 (5.0%)

Rheumatic arthritis	22,325 (1.3%)	12,990 (1.4%)
Previous muscle problems	85,581 (4.8%)	52,926 (5.8%)
Degenerative joint disorders	343,640 (19.3%)	206,491 (22.6%)
COPD	62,410 (3.5%)	36,189 (4.0%)
Hyperuricaemia	74,502 (4.2%)	46,330 (5.1%)
Hypothyroidism	87,452 (4.9%)	54,075 (5.9%)
Vitamin D deficiency	7,739 (0.4%)	5,999 (0.7%)
Vitamin B12 deficiency	26,224 (1.5%)	17,474 (1.9%)
Concomitant Medications		
Statin-interactive drugs	156,335 (8.8%)	95,576 (10.5%)
Myotoxic drugs	362,595 (20.3%)	221,062 (24.2%)
Statin Treatment		
No statins	1,537,894 (86.2%)	742,017 (81.1%)
Atorvastatin	74,350 (4.2%)	52,091 (5.7%)
Fluvastatin	3,518 (0.2%)	2,292 (0.3%)
Pravastatin	13,564 (0.8%)	9,372 (1.0%)
Rosuvastatin	6,555 (0.4%)	4,889 (0.5%)
Simvastatin	149,326 (8.4%)	103,796 (11.4%)

IQR: interquartile range; SD: standard deviation; BMI: body mass index; CVD: cardiovascular disease; COPD: chronic obstructive pulmonary disease.

4.3.2 Outcome incidence and competing risk

A total of 5,723 cases of defined serious muscle disorders were observed in the cohort over the 10-year follow-up, with a crude incidence of 0.3% and an incidence rate of 5 per 10,000 person-years. A total of 431,320 competing events (deaths from causes other than muscle disorders) occurred during the follow-up, with a crude incidence of 24.2% and an incidence rate of 391 per 10,000 person-years. The estimated cumulative incidences of serious muscle disorders and the competing events for 1, 5, and 10 years in the cohort are reported in **Table 4.3**.

Table 4.3 Cumulative incidences of the outcome and the competing event

Cumulative Incidence (95% CI)	Muscle Disorder		Competing Event
	Unadjusted	Adjusted By Competing Risk	
1-year	0.040% (0.037% - 0.044%)	0.040% (0.037% - 0.042%)	3.49% (3.47% - 3.52%)
5-year	0.22% (0.21% - 0.22%)	0.20% (0.19% - 0.20%)	16.73% (16.67% - 16.79%)
10-year	0.56% (0.54% - 0.58%)	0.45% (0.44% - 0.47%)	33.28% (33.20% - 33.37%)

CI: confidence interval

After adjusting for the competing risk, the estimated 10-year cumulative incidence of serious muscle disorders reduced from 0.56% (95% CI: 0.54% - 0.58%) to 0.45% (95% CI: 0.44% - 0.47%), as illustrated in **Figure 4.2**. The 1-year and 5-year incidences also reduced after the adjustment by competing risk.

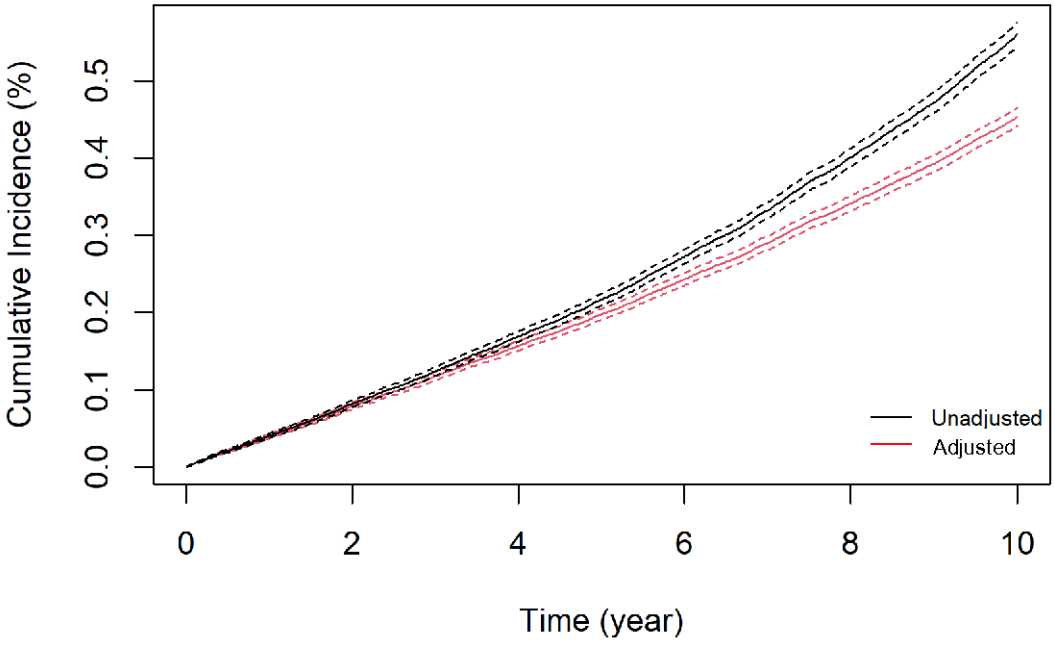


Figure 4.2 Cumulative incidence curves of serious muscle disorders in the StatinMD model derivation cohort with and without adjustment by the competing risk

4.3.3 Derivation of the StatinMD model

The model assumption of the proportional subdistribution hazards was fulfilled for all the candidate predictors, as shown in the Schoenfeld residuals plots (**Appendix 25**) where a horizontal line centred at around 0 appeared in each plot. The continuous variable frailty index was found to be linearly associated with the subdistribution hazards of the outcome, while age presented a non-linear relationship with the outcome, as shown in the log SHR plots (**Appendix 26**). A second-degree FP function was therefore selected to transform age into two terms in the model:

$$Age_{FP1} = \left(\frac{age}{100}\right)^{-2}$$

$$Age_{FP2} = \left(\frac{age}{100}\right)^{-2} \times \ln\left(\frac{age}{100}\right)$$

As shown in **Appendix 27**, the imputed data for model estimation showed a similar distribution as the observed data and consistency across the ten imputed datasets. Based on the imputed data, estimates of the coefficients of all the candidate predictors in the initial model are reported in **Appendix 28**. The candidate predictors COPD (SHR: 0.90, 95% CI: 0.79 – 1.02) and hyperuricaemia (SHR: 1.07, 95% CI: 0.96 – 1.20) were both found with non-significant coefficients and excluded from the final model. Although ethnicity was also found to have little predictive capability, it was considered an essential patient characteristic in clinical practice and remained in the final model.

Estimates of the coefficients of the included predictors in the final StatinMD model from each imputed dataset are provided in **Appendix 29** and the pooled estimates across all imputed datasets are presented in **Table 4.4**. All the statin treatments were predictive

for an increased risk of serious muscle disorders, compared to no statin treatment (atorvastatin: SHR = 1.77 [95% CI: 1.58 – 1.97], rosuvastatin: SHR = 2.04 [95% CI: 1.58 – 2.63], simvastatin: SHR = 1.58 [95% CI: 1.45 – 1.71], fluvastatin/pravastatin: SHR = 1.38 [95% CI: 1.14 – 1.68]). Other strong predictors for an increased risk of serious muscle disorders were previous muscle problems (SHR = 6.24 [95% CI: 5.87 – 6.64]), vitamin D deficiency (SHR = 2.65 [95% CI: 2.12 – 3.30]), and the use of myotoxic drugs (SHR = 2.04 [95% CI: 1.93 – 2.71]).

Table 4.4 Predictor coefficients in the StatinMD model

Predictor	Coefficient (SE)	SHR (95% CI)
Demographics		
Age		
Age $_{FP1}$	3.790 (0.173)	44.24 (31.52-62.09)
Age $_{FP2}$	3.253 (0.177)	25.87 (18.27-36.63)
Gender		
Male	Reference level	Reference level
Female	0.588 (0.034)	1.80 (1.69-1.92)
Ethnicity		
White	Reference level	Reference level
Black	0.033 (0.159)	1.03 (0.76-1.41)
South Asian	-0.039 (0.124)	0.96 (0.76-1.23)
Mixed and other	0.155 (0.109)	1.17 (0.94-1.45)
Deprivation		
Level 1 (Least deprived)	Reference level	Reference level
Level 2	0.103 (0.043)	1.11 (1.02-1.21)
Level 3	0.137 (0.042)	1.15 (1.06-1.25)
Level 4	0.176 (0.040)	1.19 (1.10-1.29)
Level 5 (Most deprived)	0.159 (0.045)	1.17 (1.07-1.28)
Health Status Indicators		
Alcohol consumption		
Non-drinker/Ex-drinker	Reference level	Reference level
General drinker	-0.171 (0.032)	0.84 (0.79-0.90)
High-risk drinker	-0.424 (0.075)	0.65 (0.56-0.76)
Smoking		
Non-smoker	Reference level	Reference level
Ex-smoker	0.283 (0.033)	1.33 (1.24-1.41)
Current smoker	0.195 (0.040)	1.22 (1.12-1.31)

BMI level		
Normal	Reference level	Reference level
Underweight	0.090 (0.103)	1.09 (0.89-1.34)
Overweight	0.139 (0.033)	1.15 (1.08-1.23)
Obese	0.396 (0.043)	1.49 (1.37-1.62)
Morbidly obese	0.489 (0.068)	1.63 (1.43-1.86)
Frailty index	0.361 (0.026)	1.43 (1.36-1.51)
Comorbidities		
CVD	-0.383 (0.043)	0.68 (0.63-0.74)
Hypertension	-0.122 (0.033)	0.88 (0.83-0.94)
Diabetes	-0.214 (0.047)	0.81 (0.74-0.89)
Mild liver diseases	0.688 (0.181)	1.99 (1.39-2.84)
Kidney diseases	0.302 (0.051)	1.35 (1.22-1.49)
Rheumatic arthritis	0.455 (0.068)	1.58 (1.38-1.80)
Previous muscle problems	1.832 (0.032)	6.24 (5.87-6.64)
Degenerative joint disorders	0.314 (0.032)	1.37 (1.29-1.46)
Hypothyroidism	0.257 (0.045)	1.29 (1.18-1.41)
Vitamin D deficiency	0.974 (0.113)	2.65 (2.12-3.30)
Vitamin B12 deficiency	0.506 (0.072)	1.66 (1.44-1.91)
Concomitant Medications		
Statin-interactive drugs	0.216 (0.037)	1.24 (1.15-1.33)
Myotoxic drugs	0.714 (0.030)	2.04 (1.93-2.17)
Statin Treatment		
No statins	Reference level	Reference level
Atorvastatin	0.569 (0.056)	1.77 (1.58-1.97)
Rosuvastatin	0.713 (0.129)	2.04 (1.58-2.63)
Simvastatin	0.454 (0.042)	1.58 (1.45-1.71)
Fluvastatin/Pravastatin	0.326 (0.099)	1.38 (1.14-1.68)

SE: standard error; SHR: subdistribution hazard ratio; CI: confidence interval;
FP: fractional polynomial function; BMI: body mass index; CVD: cardiovascular disease.

With these estimated predictor coefficients, an individual's prognostic index (PI) for serious muscle disorders can be expressed as:

$$PI = \left(\left(\frac{age}{100}\right)^{-2} - 2.672\right) \times 3.790 + \left(\left(\left(\frac{age}{100}\right)^{-2} \times \ln\left(\frac{age}{100}\right)\right) - (-1.386)\right) \times 3.253 +$$

$$\begin{cases} 0 \; if \; male \\ 0.588 \; if \; female \end{cases} + \begin{cases} 0 \; if \; white \; ethnicity \\ 0.033 \; if \; black \; ethnicity \\ -0.039 \; if \; south \; asian \; ethnicity \\ 0.155 \; if \; mixed \; or \; other \; ethnicity \end{cases} +$$

$$\begin{cases} 0 \; if \; deprivation \; level \; 1 \\ 0.103 \; if \; deprivation \; level \; 2 \\ 0.137 \; if \; deprivation \; level \; 3 \\ 0.176 \; if \; deprivation \; level \; 4 \\ 0.159 \; if \; deprivation \; level \; 5 \end{cases} + \begin{cases} 0 \; if \; non - drinker \\ -0.171 \; if \; general \; drinker \\ -0.424 \; if \; high - risk \; drinker \end{cases} +$$

$$\begin{cases} 0 \; if \; non - smoker \\ 0.283 \; if \; ex - smoker \\ 0.195 \; if \; current \; smoker \end{cases} + \begin{cases} 0 \; if \; normal \; BMI \\ 0.090 \; if \; underweight \; BMI \\ 0.139 \; if \; overweight \; BMI \\ 0.396 \; if \; obese \; BMI \\ 0.489 \; if \; morbidly \; obese \; BMI \end{cases} +$$

$$(FI \times 10 \times 0.361) + (-0.383 \; if \; cardiovascular \; diseases) +$$

$$(-0.122 \; if \; hypertension) + (-0.214 \; if \; diabetes) +$$

$$(0.688 \; if \; mild \; liver \; diseases) + (0.302 \; if \; kidney \; diseases) +$$

$$(0.455 \; if \; rheumatic \; arthritis) + (1.832 \; if \; previous \; muscle \; problems) +$$

$$(0.314 \; if \; degenerative \; joint \; disorders) + (0.257 \; if \; hypothyroidism) +$$

$$(0.974 \; if \; vitamin \; D \; decifiency) + (0.506 \; if \; vitamin \; B12 \; deficiency) +$$

$$(0.216 \; if \; statin - interactive \; drugs) + (0.714 \; if \; myotoxic \; drugs) +$$

$$\begin{cases} 0 \; if \; no \; statins \\ 0.569 \; if \; atorvastatin \\ 0.713 \; if \; rosuvastatin \\ 0.454 \; if \; simvastatin \\ 0.326 \; if \; fluvastatin \; or \; pravastatin \end{cases} \quad \textbf{(Equation 4.4)}$$

Based on the personalised prognostic index and the estimated baseline cumulative incidence (CIF_0) in the population, an individual's risk of serious muscle disorders in 1, 5, and 10 years can be calculated by the following equations:

$$1 - Year\ Risk = 1 - \left(1 - CIF_0(1)\right)^{\exp(PI)} \quad \textbf{(Equation 4.5a)}$$

$$5 - Year\ Risk = 1 - \left(1 - CIF_0(5)\right)^{\exp(PI)} \quad \textbf{(Equation 4.5b)}$$

$$10 - Year\ Risk = 1 - \left(1 - CIF_0(10)\right)^{\exp(PI)} \quad \textbf{(Equation 4.5c)}$$

With these estimates and equations, a web-based risk calculator **(Figure 4.3)** was created to implement the derived StatinMD model, which allows personal data input and reports the predicted risk of serious muscle disorders. For intellectual property protection, the risk calculator with the full algorithms (including the CIF_0 values) for the StatinMD model is freely available from the Oxford University Innovation platform (https://process.innovation.ox.ac.uk/software/) for academic and research use on request.

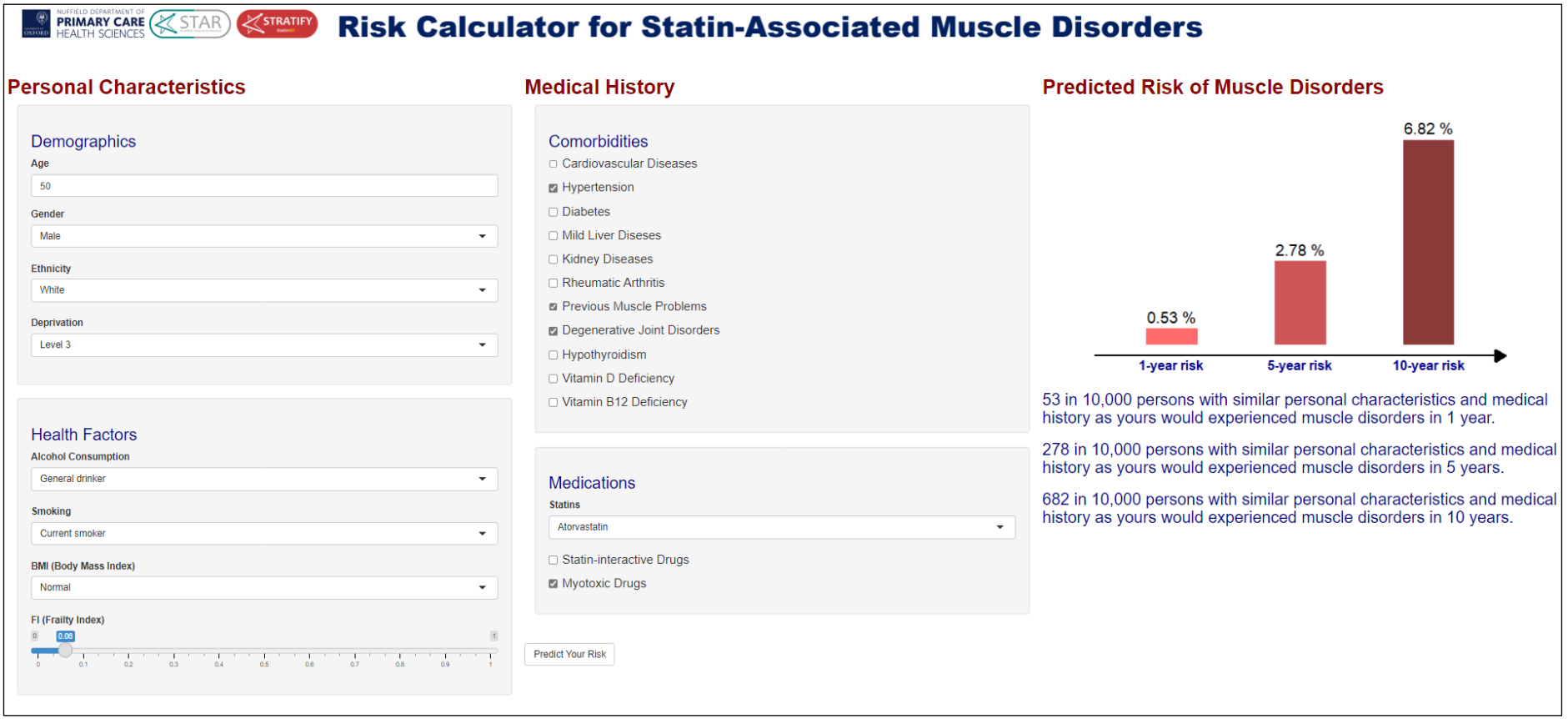


Figure 4.3 A web-based risk calculator based on the StatinMD model (beta version)

4.4 Discussion

4.4.1 Summary of key findings

Based on around 1.8 million individuals in primary care in the UK, who were potentially eligible for statin treatment, a prognostic prediction model (StatinMD) was derived to predict the personalised risk of serious muscle disorders, taking into account the competing risk of death from other causes. The model incorporated 22 predictors, including essential demographics, indicators for general health status, comorbidities, and medications, which were routinely recorded in clinical practice. Statin treatments were found to be predictive of an increased risk of serious muscle disorders in this model. Other strong predictors contributing to an increased risk included previous muscle problems, vitamin D deficiency, and the use of myotoxic drugs. This model provided the personalised prognostic index and the population baseline incidences of serious muscle disorders, which can be used to calculate an individual's absolute risk in 1, 5, and 10 years when deciding on statin treatment.

4.4.2 Comparison with previous studies

4.4.2.1 Definition and ascertainment of muscle-related outcomes

The previous QStatin model predicted myopathic events that included clinically diagnosed myopathy and rhabdomyolysis as well as raised CK of over 4ULN.[285] However, clinical diagnosis of myopathy or more serious muscle disorders is generally considered with at least moderate CK elevation of over 10ULN.[284] Minor CK elevation could be caused by many reasons, such as intensive physical activity, which are much more common than diagnosed myopathy and rhabdomyolysis but less likely to be attributable to statins.[12, 67] With the inclusion of these milder conditions, the predicted risk from the

QStatin model is probably higher than the actual risk of those muscle disorders of concern, which may exaggerate patients' worries about statin safety and result in inappropriate treatment rejection or withdrawal.

The other previous prediction model for myopathy with simvastatin defined the outcome as unexplained muscle pain or weakness with raised CK of over 10ULN.[286] Although this is a general indication for underlying muscle disorders, it does not necessarily lead to clinical diagnosis of myopathy or other serious muscle conditions. The diagnostic criteria of myopathy usually include other clinical examinations and tests, such as standardised muscle strength examination, metabolic tests, and confirmation by electromyography.[20] More serious conditions, such as rhabdomyolysis, often present a much higher CK level of over 40ULN and their diagnosis requires further examinations and tests including urinary myoglobin test and muscle biopsy.[20, 67] Only relying on the criteria of moderate CK elevation, most of the ascertained outcome events for developing the simvastatin model may not be the clinically important myopathy and other serious muscle disorders.

For the StatinMD model in this study, the outcome was defined by clinical diagnoses of serious muscle disorders limited to myopathy, myositis, and rhabdomyolysis, which were associated with hospitalisation or death. Although the clinical diagnoses in practice may be made based on various examinations or tests and using inconsistent criteria, this definition probably better limited the outcome events to those with clinically-confirmed severity, compared to solely relying on CK elevation.

4.4.2.2 Inclusion and definition of predictors

The selection of the predictors for both the QStatin model and the simvastatin model relied on the statistical significance of the predictor coefficients, which is likely to be

subject to the testing methods and the sample size and has resulted in the elimination of some clinically important predictors.[132, 136] Specifically, the QStatin model excluded common health risk factors, such as drinking and smoking, and the simvastatin model excluded all comorbidities except diabetes.[285, 286]

In contrast, the predictor selection in the current StatinMD model was determined based on the statistical predictive ability as well as the clinical importance of the predictors, which has led to the inclusion of both the strong predictors for the outcome and the predictors of clinical concerns. As a result, the predictors in the StatinMD model have covered most of the predictors in the previous models and included additional predictors that were previously omitted but clinically relevant to muscle disorders, such as a history of muscle problems, degenerative joint disorders, and vitamin D deficiency.

The simvastatin model did not include the use of statins as a predictor, since this model was derived based on current statin users only.[286] As such, the simvastatin model may not be appropriate to be used in individuals who are eligible for but not currently on statin treatment. The QStatin model included statin use as a predictor but, by definition, the statin users included only new users without previous use of statins and the non-users were those who never used statins.[285] However, in clinical practice, statin users would include a large number of patients who have been taking statins for a long time and the current non-users would include those who previously took but stopped taking statins. The risk of muscle disorders for these patients may be different from the new users and the never-users, given the history of statin treatment or other different patient characteristics. This may therefore limit the applicability of the QStatin model to statin users who are deciding on continuation of statin treatment or current non-users who are considering the re-uptake of statins in clinical practice.

For the StatinMD model in the current study, statin users were defined as all individuals with current statin use in the recent 12 months, regardless of previous use, and current non-users also included those with previous uses of statins. Such a definition makes the StatinMD model more widely applicable to individuals considered for initiation, continuation, or re-uptake of statin treatment, for all of whom the predicted risk of muscle disorders may be needed. In addition, the current StatinMD model took into account the specific drug types of statins, which may provide greater precision in the prediction, while the QStatin model did not.

4.4.3 Strengths and limitations

4.4.3.1 Strengths

- **Improving prediction accuracy by addressing competing risk**

A major strength of the StatinMD model in this study is improving the prediction accuracy by addressing the competing risk. As explained in Chapter 2, the presence of competing risk, if not addressed properly, is likely to result in overestimation of the outcome risk, because individuals who have experienced competing events and are no longer at risk of the outcome are treated as they could still develop the outcome.[151] This overestimation could be large when the competing risk in the population is high. In the population eligible for statin treatment in this study, who were generally older and with multiple comorbidities, the competing risk of death from causes other than muscle disorders was high (33% in 10 years) and showed an impact on the estimate of muscle disorder risk.

In the context of statin treatment decision-making, given the widespread concern about statin safety that has in part led to poor adherence to and low uptake of statins in the eligible population,[30, 35] overestimating the risk of a potential adverse event is particularly

unwanted. However, neither of the previous models for muscle-related adverse events has taken into account the competing risk.[285, 286, 310] In the current study, by using the Fine-Gray approach to adjust the competing risk in the model derivation, the StatinMD model should reduce potential overestimation of the risk of muscle disorders. This may help avoid inappropriate statin treatment withdrawal or hesitation in individuals at low risk of muscle disorders and, on the other hand, justify the necessary discontinuation of statins when an individual's risk is high.

- **Potential strengths in the study design and modelling process**

As discussed above, by using the clinical diagnoses in hospitalisation or death records to define the outcome, this StatinMD model better targets the risk of serious muscle disorders that are more of concern in statin treatment decision-making than those milder conditions. Compared to the previous models, the selection of the predictors in this model not only relied on the statistical significance but also considered the clinical importance of the predictors, which helped avoid eliminating predictors that are of clinical concern for statin treatment. The definition of statin use should make this model widely applicable to all individuals who would potentially need the prediction of muscle disorders to assist clinical decision-making for initiation, continuation, or re-uptake of statins.

In addition, while the previous models predicted either a short-term risk in 6 months or 2 years (the simvastatin model) or a longer-term risk in 5 years (the QStatin model),[285, 286] the current StatinMD model was designed to predict the risk of muscle disorders in 1, 5, and 10 years. This gives a risk directly comparable to the CVD risk that is commonly predicted for 10 years and also provides the flexibility to take into account both short-term and long-term risks depending on personal considerations in statin treatment decision-making.

4.4.3.2 Limitations

- **Limitations in the study designs**

The target population for predicting the risk of serious muscle disorders in this study was the individuals eligible for statin treatment for prevention of CVD. This eligibility should be determined by an individual's 10-year CVD risk predicted by existing tools, according to the clinical guidelines.[5, 6] However, the predicted CVD risk is poorly recorded in the CPRD primary care data, missing in more than 90% of the population as found in the initial data inspection. To construct the study cohort practically, this study used age as a proxy of the predicted CVD risk to determine the eligibility of study participants. The study cohort included men over 50 years old and women over 60 years old, who roughly represent the population with a predicted 10-year CVD risk above 5% based on the QRISK2 model that was commonly used in the UK during the study period.[59] However, using the cut-off age instead of the CVD risk to determine study participants will have missed some younger individuals who have a CVD risk above 5% due to the presence of other risk factors for CVD. Although such individuals may make up only a small proportion of the target population, as most of the risk factors for CVD are probably more common in older individuals, this design may diminish the representativeness of the study cohort for the target population and impact the validity of the derived StatinMD model in these younger individuals.

The desired outcome for prediction was serious muscle disorders that resulted in hospitalisation or death. However, the outcome in this study was defined by a hospital admission or death record with a diagnosis of serious muscle disorders. The hospitalisation or death may not be necessarily caused by the diagnosed muscle disorders and it is difficult to confirm the direct cause based on the EHR data.[311] Moreover, this

study considered all associated diagnoses for hospital admission and all recorded causes for death, instead of relying on the record of 'primary diagnosis' or 'primary cause'. Although this was to increase the sensitivity to avoid missing any occurrences of clinically-confirmed muscle disorders that are generally rare in the population, it may lead to a lower specificity to identify the serious cases. Consequently, the predicted risk from the StatinMD model may be higher than the risk of serious muscle disorders that could actually cause hospitalisation or death.

As predictors for the outcome, the use of the medications was determined by a recent prescription within a 12-month baseline exposure period after the eligible participants entered the cohort. This design treated the individuals with previous use of the medications before the baseline exposure period as the same as those who never used the medications. The risk of serious muscle disorders for these two groups of individuals may be different, particularly if the previous users were on the medications for a long time and just recently stopped.[312] However, if the study cohort was designed without a baseline exposure period, the use of the medications may be defined by a prescription years ago, which may have little predictive ability for the outcome. More importantly, this defined period enabled measuring all baseline characteristics for the individuals without data before they entered the cohort (i.e. they reached the eligible age before the start date of this study).

- **Difficulties and compromises in the modelling and computation**

When deriving the StatinMD model, the predictor BMI was included as a categorical variable instead of a continuous variable, in order to make use of the data from the records of categorised BMI levels defined by the CPRD medical codes as a supplement to the records of BMI values that contained considerable missing data. Due to the low incidence

of the outcome, several categorical variables (ethnicity, alcohol consumption level, statin drug type) were re-categorised by combining some of the originally defined groups, for example, fluvastatin and pravastatin were combined into one group, in order to have a sufficient number of outcome events in each category to facilitate model convergence. Such categorisation of a continuous variable or combination of categories may both lead to loss of information in the data and reduce the statistical power of model estimation,[313] although this may be only by a small degree given the huge sample size in this study.

Due to the intensive and time-consuming computation for the competing risk model, the model assumptions were checked in the complete-case data only, which benefited from the smaller sample size than the whole cohort and improved the analysis efficiency. Although the sample size of the complete-case data was believed to be sufficient to detect any violations of the model assumptions in the study data, there may be potential differences between the complete-case sample and the whole cohort. It would have been more reassuring to confirm the fulfilment of the model assumptions with the full-cohort data if there was a more efficient computation algorithm to do so.

Similarly, there is currently no efficient algorithm in R to perform the selection of FP function for the competing risk model, due to the complexity of the model and the intensive computation in the selection procedure. Therefore, in this study, the optimal FP function to transform age had to be selected using the available R package 'mfp' that implements the Cox model for a time-to-event outcome. Although the Cox model and the competing risk model share similar model assumptions and statistical fundaments,[156] the selected FP function based on the Cox model may not still be optimal for the competing risk model. In addition, the FP function was selected also based on the complete-case data only, which may be subject to the potential differences between the complete-case sample and the whole cohort.

Another compromise to the intensive computation was that only 10 imputations were able to be performed to impute the missing data, given the huge sample size. Although this was a reasonable number of imputations for a balance between statistical power and efficiency, more imputations may further improve the imputation results given the considerable proportion of missing data in this study.[314] Moreover, the imputation models for the four imputed variables included the same set of explanatory variables that consisted of all candidate predictors for the StatinMD model, the prediction outcome, the competing event, and their hazards. However, no auxiliary variables were particularly sought for each imputed variable to further explain the missingness, due to the practical purpose of imputing multiple variables at once and the complexity of this imputation procedure. Although the constructed imputation models performed well, as indicated in the comparison between the imputed data and the observed data, the missingness of the four imputed variables may not be fully explained by the explanatory variables in the imputation models and the imputation results may be improved by including auxiliary variables.[315]

- **Pitfalls with the use of electronic healthcare records**

The StatinMD model was derived using the EHR data that were collected in routine clinical practice. As a common drawback of using routine healthcare records,[311] there were missing data in four predictor variables (ethnicity, alcohol consumption, smoking status, BMI level), with a considerable missing proportion ranging from 14% to 30%. These missing data were imputed using the multiple imputation approach, for which some technical compromises had to be made due to the intensive computation and the huge sample size, as discussed above. In addition, even if the computation capacity allowed more imputations and the inclusion of auxiliary variables, a more fundamental pitfall is that the assumption of missing at random for performing multiple imputation may not be

always true.[162] For example, ethnicity may be more likely to be recorded if an individual is from a minority group, which may result in missing not at random in the data. However, there is currently no way to distinguish between missing at random and missing not at random.[162] Although the missing-at-random assumption has been widely accepted in epidemiological studies using healthcare records and this assumption for the four imputed variables in this study is probably reasonable as justified in the methods session, it should be acknowledged that the fulfilment of this assumption is not confirmed and there may be potential bias introduced due to the data imputation.[316]

Another pitfall of using EHR data is the potential misclassification of the presence of outcomes or exposures because it is impossible to distinguish between a missing record and the actual absence of a condition.[162, 317] In this study, given potential missed outcome events that were not captured in the hospitalisation or death records, the overall incidence of the outcome observed in the study cohort may be lower than the actual incidence in the population and consequently the personalised risk predicted from the derived model may be underestimated. Moreover, such misclassification may be subject to potential recording bias.[311] For example, in this study, muscle disorders may be more likely to be recorded in statin users since clinicians are aware of this potential adverse event and therefore more outcome events may be missed in statin non-users. This may result in biased estimates of the coefficients for some baseline characteristics, like the use of statins, probably leading to a larger underestimation of the personalised outcome risk in individuals without the presence of the characteristics than those with the characteristics. In this study, this concern may be alleviated to some degree because the outcome events were serious conditions that were unlikely to be ignored in the hospitalisation and death recording.

Similarly, the presence of the comorbidities at baseline may be also misclassified and under-recorded in the study cohort, especially those non-serious but common conditions in the population, such as degenerative joint disorders and vitamin D deficiency.[311] In addition, unlike the ICD diagnosis codes used to identify the outcome from the hospitalisation and death records, the CPRD medical codes used to identify the baseline comorbidities from the primary care records are less structured or explicit, making it difficult to compile a complete code list for each comorbidity. The potential incompleteness of the codes may result in further under-capture of the comorbidities. In contrast, the baseline use of the medications is less likely to be under-recorded, as most of the baseline medications in this study are only available through a GP's prescription and the codes for identifying these medications were based on the BNF chapter and more definite. However, misclassification bias may still exist the other way around, because a prescription of a medication does not necessarily mean the actual uptake of the medication but it is difficult to confirm the medication uptake based on the EHR data. Nevertheless, the prevalence of the comorbidities and the utility of the medications in the study cohort were similar to those observed in national surveys and other population-based studies in the UK,[270, 318, 319] which suggests that the potential misclassification of the baseline characteristics may not be concerning.

4.4.4 Clinical implications

4.4.4.1 Supporting personalised clinical decision-making on statin treatment

The personalised risk of serious muscle disorders predicted by the StatinMD model derived in this study, in conjunction with the personalised CVD risk predicted from other tools such as the commonly used QRISK2 model in the UK, could help doctors and

patients better understand the balance between potential benefits and harms of statin treatment, in order to make well-informed patient-centred clinical decisions.

This personalised assessment of the benefit-harm balance is through the calculation of the absolute risk reduction of CVD and the risk increase of serious muscle disorders by statin treatment, using the predicted personalised baseline risks without statins and the relative effects of statins on CVD and muscle disorders. The relative effects could be obtained from previous reliable research evidence, ideally from randomised controlled trials or large observational studies if trial evidence is unavailable. In the calculation examples below, the effect of statins on CVD (RR = 0.79, 95% CI: 0.77-0.81) is from a meta-analysis of individual patient data from 21 randomised controlled trials.[268] Since the trial evidence on adverse effects is less generalizable (due to the exclusion of vulnerable individuals) and the effect of statins on muscle disorders hasn't been estimated precisely with trial data (due to the very low incidence of muscle disorders and the limited sample size of clinical trials),[12, 279] the estimate of this effect (OR = 2.63, 95% CI: 1.50-4.61) is from a meta-analysis of 6 large observational studies in a total of nearly 3 million participants.[25] The post-treatment risks of CVD and serious muscle disorders can be subsequently estimated by applying the relative effects to the predicted baseline risks. For current statin users, the 'post-treatment' risks are the risks when continuing statin treatment.

According to this personalised assessment, different decisions on statin treatment may be made for individuals based on the comparison between potential benefits and harms. For individuals whose risk of serious muscle disorders is low and the potential risk reduction of CVD by statin treatment outweighs the risk increase of serious muscle disorders, statin treatment should be initiated or continued. It would be particularly helpful to use the prediction from the StatinMD model to inform statin users who have

experienced muscle symptoms or prospective users who have concerns about muscle-related adverse events that their risk of developing serious muscle disorders is low, in order to avoid inappropriate treatment withdrawal and improve statin uptake.

The comparison between benefits and harms should be more carefully considered for individuals with a higher risk of serious muscle disorders. For these individuals, if their CVD risk is high and could have a considerable reduction by statins, it may be still appropriate to initiate or continue statin treatment but cautions are warranted. Other clinical characteristics of the individuals may need to be taken into account to help make the decision, such as the cholesterol level, and clinical monitoring of CK level or muscle functions could be considered while taking statins. However, for individuals whose risk of serious muscle disorders with statins is expected to be very high, it may not be appropriate to initiate or continue statin treatment. In this circumstance, other lipid-lowering medications or non-pharmaceutical interventions could be considered for CVD prevention.

Nevertheless, the final treatment decision needs to be made through transparent communication and discussion between doctors and patients.[320] The personalised risk information provided by the StatinMD model could assist such communication and discussion. Patients' understanding of the risk information and personal preferences in the trade-off between benefits and harms are also important and need to be taken into account in the discussion, in order to achieve personalised clinical decision-making for statin treatment.[321]

Examples are given below to demonstrate the personalised assessment of the benefit-harm balance and decision-making for statin treatment in different clinical scenarios.

Example 1:

*For a 55-year-old White man at a median deprivation level (level 3), who is a (moderate) current smoker, with a BMI of 27kg/m^2, a frailty index of 0 (with no any of the defined deficits), without any other health indicators, comorbidities, or medications included in the QRisk2 model or the StatinMD model, the 10-year baseline CVD risk predicted by the QRisk2 is **11%**, which would be reduced to **8.7%** by taking statins, and the 10-year baseline risk of serious muscle disorders predicted by the StatinMD model is **0.27%**, which would be increased to **0.71%** by statins. In this case, the risk of serious muscle disorders with statins is negligible and statin treatment should be initiated to help reduce the CVD risk.*

Example 2:

*For a 55-year-old White man at a median deprivation level (level 3), who is a (moderate) current smoker, with a BMI of 27kg/m^2, a frailty index of 0.08 (with 3 of the 36 deficits, such as dizziness, skin ulcer, and sleep disturbance), a history of previous muscle problems, degenerative joint disorders, vitamin D deficiency, and currently taking Omeprazole (one of the myotoxic drug included in the current model) to treat heartburn, the 10-year baseline CVD risk from the QRisk2 model is also **11%**, which would be reduced to **8.7%** by statins, but the 10-year baseline risk of serious muscle disorders from the StatinMD model is **15.5%**, which would be increased to **40.8%** by statins. In this case, the initiation of statin treatment should be cautious and the individual's other clinical characteristics, such as the cholesterol level, may be taken into account to make the decision. If decides to initiate statin treatment, clinical monitoring of the CK level or muscle functions may be considered. Otherwise, alternative medications or lifestyle*

interventions could be considered for CVD prevention, depending on the discussion with the individual.

Example 3:

For a 60-year-old White man at a median deprivation level (level 3), with a BMI of 27kg/m^2, a frailty index of 0.06 (the average in the study population), a cholesterol/HDL ratio of 5, a family history of heart attack, a personal history of atrial fibrillation, and currently taking atorvastatin, the present 10-year CVD risk from the QRisk2 model is ***25%****, which would be reduced to* ***19%*** *if statin treatment is continued, and the 10-year risk of serious muscle disorders from the StatinMD model is* ***0.24%****, which would be increased to* ***0.63%*** *with continuing statin treatment. In this case, the individual should be reassured that he is unlikely to develop serious muscle disorders with statins and should be encouraged to continue statin treatment for CVD prevention.*

4.4.4.2 Potential stratified statin treatment strategies

With the prediction from the StatinMD model, clinical recommendations on statin treatment in the population could be stratified not only by the risk of CVD but also by the risk of serious muscle disorders. According to current clinical guidelines, decisions on statin treatment should be first based on an individual's CVD risk and the recommended risk threshold for statin treatment. Individuals considered eligible for statin treatment could be further stratified into two groups with a high or low risk of serious muscle disorders, using the prediction from the StatinMD model.

Based on the risk stratification, different statin treatment recommendations may be implemented in different risk groups of serious muscle disorders, which is the stratified statin treatment strategy. For individuals in the low-risk group, whose benefit from the risk reduction of CVD overweights the potential harm from the risk increase of serious

muscle disorders, statin treatment should be generally recommended for CVD prevention. For individuals in the high-risk group, cautions for statin treatment may need to be taken and some clinical actions to address concerns about serious muscle disorders may be considered. Such clinical actions may include performing a test of CK before initiating statin treatment, monitoring CK and muscle functions during the treatment, a decision not to initiate statin treatment, or withdrawing statins when necessary. As discussed above, a decision on such clinical actions should depend on a full personalised assessment of the benefit-to-harm balance, as well as patients' preferences and doctors' clinical opinions. As such, by targeting these clinical actions at the high-risk individuals identified by the StatinMD model, the appropriateness, efficiency, and cost-effectiveness of these clinical actions may be improved.

A sensible risk threshold to stratify individuals of low and high risk of serious muscle disorders will need to be determined based on many aspects, including the prediction performance of the StatinMD model and the potential clinical utility of the model with different risk thresholds, which are examined and discussed in the next chapter. Nevertheless, the potential stratified statin treatment recommendations that take into account the widespread public concerns about statin safety could better embody the principle of developing clinical guidelines, which is to ensure patients' needs and priorities are reflected.[322, 323]

4.4.5 Future research

4.4.5.1 Validation of the StatinMD model

As explained in Chapter 2, even if a model has been derived following proper methods and processes, it is not guaranteed that the model would work well for the purpose of prediction. Therefore, before the derived model can be used in clinical practice, it is

necessary to validate the model, which includes an assessment of its prediction performance and potential clinical utility.[137] Model performance is how well the model can distinguish individuals who will and will not develop the outcome and how accurate the predicted risk would be.[140] It is important to validate the model in an external population different from the population where the model has been derived, as eventually the model will be used to predict the risk for unseen patients in clinical practice.[144] The validation of the StatinMD model is reported in the next chapter.

4.4.5.2 Application of the StatinMD model

In order to apply the StatinMD model to clinical decision-making, a practical prediction tool needs to be created based on the model algorithm. This could be a web-based risk calculator, as shown in **Figure 4.3** above. Other examples of such prediction tools are the QRISK3 calculator for CVD risk and the PREDICT calculator for breast cancer risk.[324, 325] However, there are many issues to be taken into account when developing a prediction tool, including a user-friendly design of the risk calculator, the most effective presentation of the risk information, the workload of using the prediction tool in clinical decision-making, and the regulatory issue of such a potential medical device.[326-328] Further research is therefore warranted for applying the StatinMD model to clinical practice properly.

4.4.5.3 Development of prediction models for other potential adverse events

This thesis has prioritised serious muscle disorders for risk prediction, as they have caused major concern in the public that impacts clinical decisions on statin treatment. However, other potential adverse events may be also concerning for patients and clinicians when making a treatment decision. Predicting the risks of other adverse events could help make a more comprehensive comparison between the benefits and harms of statin treatment, leading to better-informed decision-making. In particular, similar to

muscle-related adverse events, liver-related adverse events of statins also involve common but mild liver dysfunction and rare but severe liver injury.[22] In this case, it would be useful to identify individuals at high risk of severe liver injury from those with mild dysfunction, in order to decide further clinical actions on the right patients.

4.5 Conclusion

A prognostic prediction model StatinMD has been successfully derived to predict the risk of serious muscle disorders in 1, 5, and 10 years for individuals potentially eligible for statin treatment, based on personal characteristics and medical history. This model has several advantages over previous similar models, including taking into account the competing risk. This model could help assist well-informed, personalised decision-making regarding statin treatment in clinical practice, reassuring eligible patients with a low risk of serious muscle disorders or determining clinical actions for patients with a high risk. The prediction from the model could be used to develop potential stratified statin treatment strategies that better address patients' concerns and needs in clinical recommendations and accelerate clinical decision-making.

Chapter 5 Predicting Personalised Risk of Serious Muscle Disorders in Patients Eligible for Statin Treatment: Validation of the StatinMD Model

5.1 Introduction

5.1.1 Importance of model validation

A prediction model may not work well for the purpose that it has been designed for, due to the limitations of the model derivation process or the heterogeneity of the populations and clinical settings that it is applied to.[137, 140] For the StatinMD model derived in the preceding chapter, it may cause unnecessary concerns and inappropriate treatment withdrawals if the risk of serious muscle disorders is overestimated, or it may leave high-risk individuals to the potential harms of serious muscle disorders if the risk is underestimated. Therefore, it is crucial to validate the StatinMD model before it can be used in clinical practice to provide reliable risk information to support statin treatment decision-making.

As introduced in Chapter 2, model validation is the process of assessing the prediction performance of a model and the validation strategies include internal and external validation.[140] Internal validation is to assess the model performance using a part or all of the data that have been used for model derivation.[141] When using all the derivation data for internal validation, it is also called apparent validation. Internal validation could be used to inspect the robustness of the model derivation process and demonstrate the reproducibility of the derived model.[142] However, the model performance measured in internal validation is likely to be over-optimistic, as the model has been fitted to the same

data.[141] In contrast, external validation is to assess the model performance using data that are not involved in model derivation, which can avoid over-optimism in measuring the model performance.[144] More importantly, since the external validation data are usually from a different population, geographic location, or clinical setting, it could demonstrate the generalisability of the model.[142] This is important for a clinical prediction model, like the StatinMD model, which will eventually be applied to unseen individuals from the general population in clinical practice.

5.1.2 Key aspects of model validation

In model validation, the prediction performance of a model is commonly assessed on two essential aspects: discrimination and calibration, as recommended in the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) guidance.[143] For clinical prediction models, it is also recommended to evaluate the potential clinical utility for the purpose of prediction.

Discrimination, also known as 'prognostic separation' for a prognostic model, is the ability of a model to distinguish between individuals who will and will not develop the outcome of interest.[329, 330] If a model has good discriminative ability, individuals with a higher predicted risk from the model should have a higher observed incidence of the outcome than those with a lower predicted risk.[331] Poor discrimination can be a major failure of a prognostic model that is designed for identifying high-risk individuals who need clinical actions.[140] It may be caused by some issues in the model derivation process, such as important predictors being omitted from the model.[132]

Calibration reflects the agreement between the predicted risk and the observed risk.[140, 330] A model with satisfactory calibration should be able to predict the outcome risk accurately for most individuals in the target population.[137] The presence of miscalibration,

either underestimation or overestimation of the outcome risk, may lead to inappropriate clinical decisions based on the model. Significant miscalibration may suggest under-fitting or over-fitting of the model to the derivation data or underlying differences between the populations for model derivation and validation.[137]

Clinical utility of a model is broadly referred to as the clinical consequence of using the model to make a decision, usually to classify individuals by the predicted risk from the model for certain clinical actions, such as prescribing medications or performing clinical examinations.[330, 332] The clinical consequence of using the model could be compared to other available tools or common practices for the same decision-making to demonstrate the advantages of the model.[333] One can also measure the classification accuracy of the model with specific risk thresholds to understand if the model would work well for the purpose of risk classification.[330, 334] Clinical utility of a model depends on not only the model prediction performance but also the clinical settings (e.g. the overall outcome incidence in the population, the risk threshold for classification), but it should be part of the validation of a clinical prediction model to help support its clinical use.

5.1.3 Objectives of this chapter

This chapter aimed to validate the StatinMD model to support the clinical use of this model for personalised decision-making on statin treatment and the implementation of potential stratified statin treatment strategies in the population. The prediction performance of the model was first assessed through internal (apparent) validation based on the model derivation cohort to inspect the robustness of the model derivation process. The model performance was further assessed through external validation based on another cohort to evaluate the generalisability of the model.

5.2 Methods

5.2.1 Data source and cohort design

The internal (apparent) validation used the same data from the model derivation cohort, which included electronic healthcare records from the GOLD primary care database and the linked databases (hospital inpatient records, death registration, deprivation data) in the Clinical Practice Research Datalink (CPRD) in the UK,[287] as described in Chapter 4.

The external validation used data from the Aurum primary care database in the CPRD, which includes a population different from that in the GOLD database.[335] The key difference is that the patients in the Aurum database are registered with general practitioners (GP) who use the EMIS clinical system, while the patients in the GOLD database are registered with those using the Vision system. Similar to the GOLD database, the primary care data in the Aurum database are also linked to hospital inpatient records, death registration, and deprivation data. All data were anonymised and the use complied with the ethics approval obtained by the CPRD.

The external validation cohort was constructed from the Aurum population using the same design as the model derivation cohort, as shown in Chapter 4 (**Figure 4.1**). Briefly, the cohort included a population potentially eligible for statin treatment, who were measured for a set of personal characteristics (predictors) at the baseline and followed up for 10 years to observe the occurrence of serious muscle disorders.

In terms of the consideration for sample size, it has been suggested that external validation of a prognostic model requires a minimum of 100 outcome events and ideally 200 (or more) events.[336] The constructed external cohort in this study included many more outcome events than this requirement.

5.2.2 Model variables

For the internal validation cohort, the variables for the outcome and the predictors in the StatinMD model were created in the model derivation process. As described in Chapter 4, the outcome variables were defined by the first occurrence of serious muscle disorders, the occurrence of the competing events (deaths from causes other than muscle disorders), or the occasions of censoring, which were identified from the hospitalisation records or the death registration. The variables for the 22 predictors, including age, gender, ethnicity, deprivation level, alcohol consumption, smoking status, body mass index (BMI) level, frailty index, the presence of 11 comorbidities, and the use of 3 groups of medications, were measured using the data from the primary care records and the deprivation data.

For the external validation cohort, all these variables were extracted from the Aurum database and the linked databases based on the same definitions as for the model derivation cohort. The International Classification of Diseases (ICD) codes to identify the outcome events were the same as those used for the derivation cohort. The medical and product code lists to extract the predictor variables from the Aurum database were created by mapping to the code lists for the GOLD database used in the model derivation and searching the Aurum code dictionaries for complementary codes.

5.2.3 Statistical analysis

5.2.3.1 Comparison of derivation and external validation cohorts

In model validation studies that involve external validation, it is recommended to assess the relatedness between the samples for model derivation and external validation, to help interpret external validation results and understand the difference between the internal and external performance of the model.[145] A common approach to assessing the sample

relatedness is to compare their case mix, which is the distribution of a specific characteristic in the sample, including the baseline characteristics and the outcome, or the overall distribution of all the characteristics of relevance.[145]

5.2.3.1.1 Comparison of specific case mix

- **Baseline characteristics**

The baseline characteristics of the external validation cohort, which were the predictor variables in the StatinMD model, were described by the mean and standard deviation (SD) for continuous variables and by the frequency and percentage for categorical variables. These characteristics were compared to that of the model derivation cohort reported in Chapter 4.

- **Outcome and competing event**

The crude incidence of the outcome in the external validation cohort was calculated as the proportion (percentage) of the individuals who experienced the outcome throughout the 10-year follow-up. The average incidence rate was reported as the number of outcome events per 10,000 person-years. The cumulative incidences of the outcome for 1, 5, and 10 years were calculated by the Aalen-Johan estimator,[300] with adjustment by the competing risk, and the cumulative incidence curve was plotted over 10 years. The occurrence of the competing event was measured in the same way. These measurements were compared to those in the derivation cohort.

5.2.3.1.2 Comparison of overall case mix

The overall case mix in the derivation and external validation cohorts was compared by the distribution of their participants' prognostic index that incorporates all the baseline characteristics (predictors).[145] Each individual's prognostic index in both cohorts was

calculated with the individual's values of the predictor variables in the StatinMD model, using the derived model formula (**Equation 4.3**) in Chapter 4. The distribution of the calculated prognostic index in each cohort was summarised by mean and SD and compared to each other using a density plot.[337]

5.2.3.2 Model validation

The internal (apparent) and external validations were conducted following the same process and methods described below to assess the model performance in the two cohorts.

5.2.3.2.1 Model prediction

The risks of the outcome in 1, 5, and 10 years for each individual in the derivation and external validation cohorts were predicted with the individual's prognostic index calculated above, using the StatinMD model formula (**Equation 4.5**) that incorporates the baseline cumulative incidence functions of the outcome for the three time periods. These predicted individuals' risks were used for the following assessment of the model performance.

5.2.3.2.2 Assessment of model discrimination

The discrimination of the model was assessed by the *c*-index and the D statistic.

- ***c*-index**

The *c*-index, also called the concordance index, measures the overall concordance between the predicted and the observed outcome occurrence.[338] It takes a value between 0.5 and 1, which is the probability that given a randomly selected pair of individuals, one with the outcome and the other without, the model will predict a higher risk for the one with the outcome.[339] A *c*-index closer to 1 indicates a better discriminative ability of the model, while a *c*-index of 0.5 indicates that the model is incapable of distinguishing

between individuals who will and will not develop the outcome.[340] The *c*-index of the StatinMD model was calculated using the function 'cindex' from the R package 'pec'.

- **D statistic**

The D statistic, also called the prognostic separation index, has been developed to particularly assess the discriminative ability of a prognostic model with a time-to-event outcome.[329] It can be regarded as an equivalent of the relative risk of the outcome (e.g. subdistribution hazard ratio for the Fine-Gray model) between two equal-sized risk groups divided by the median of individuals' prognostic index calculated from the prognostic model in a population.[341] A larger estimate of the D statistic indicates a better prognostic separation based on the prediction from the model.[329] Since there was no available function in R for calculating the D statistic of a competing risk model at the time of this study, I developed an R function 'royston.cpr' to calculate the D statistic of the StatinMD model, based on the relevant methodology papers and the existing function 'royston' from the R package 'survival'.

5.2.3.2.3 Assessment of model calibration

The calibration of the model was assessed by the ratio of the overall expected and observed outcome risk (E/O), the calibration slope, and the calibration plot (curve).

- **E/O ratio**

The E/O ratio reflects the agreement between the predicted and the observed outcome risk on average in the population.[139] A ratio bigger than 1 indicates an overall overestimation of the risk by the model, while a ratio smaller than 1 indicates an overall underestimation.[342] In this study, the E/O ratio was calculated with the mean of the predicted outcome risk from the StatinMD model across all individuals and the observed

cumulative incidence of the outcome in the cohort calculated by the Aalen-Johan estimator that took into account the competing risk.[300]

- **Calibration slope**

Calibration slope is the coefficient of the prognostic index obtained by fitting a Fine-Gray model to the cohort data with the prognostic index from the StatinMD model as the only predictor for the outcome.[343] Calibration slope should be interpreted in conjunction with the overall calibration indicated by the E/O ratio. When overall good calibration is indicated by an E/O ratio close to 1, a calibration slope around 1 means the good calibration is consistent across the population, otherwise, it indicates there may be some subgroups with poorer calibration.[344] A calibration slope away from 1 also suggests there may be a need to re-calibrate the prognostic index (the predictor coefficients).[338]

- **Calibration plot**

Calibration plot is a scatter plot presenting the observed outcome risk against the predicted risk, with a calibration curve fitted to the plotted data points, which is usually a curve smoothed by the local polynomial regression (LOESS) model.[342, 345] The position of the curve against the diagonal of the plot indicates the calibration of the prediction model. The diagonal represents the complete agreement between the observed and the predicted risks (i.e. perfect calibration), while the area under the diagonal represents an overestimation of the outcome risk by the model and the area above represents an underestimation.[342]

As a common approach, the calibration plot in this study was first drawn based on ten risk groups defined by the deciles of the individuals' prognostic index in the cohort.[144] The observed incidence of the outcome in each risk group was calculated by the Aalen-Johan estimator,[300] and the predicted risk in each group was the mean of the individuals'

predicted risks from the StatinMD model. The ten data points were plotted, with a fitted LOESS curve, to illustrate the calibration of the StatinMD model at the group level.

To further understand the model calibration across the full range of possible predicted risks in the population, a calibration plot was drawn with the individuals' pseudo observations of the outcome occurrence and the individuals' predicted risks from the StatinMD model in the cohort. An individual's pseudo observation can be viewed as the contribution of the individual to the estimate of the outcome incidence in the population, obtained by comparing the estimates with and without the individual.[346, 347] This enables having a pseudo value of the outcome risk for individuals who are not observed with the outcome event during the study period (i.e. being censored).[347] These pseudo values were calculated using the Jackknife resampling method, in conjunction with the Aalen-Johan estimator.[300, 348] A LOESS curve was also fitted to these individual data points to show the model calibration at the individual level.

5.2.3.2.4 Assessment of the goodness-of-fit

In addition to discrimination and calibration, it is also common to assess the goodness-of-fit of a model to understand how well the included predictors explain the variation of the outcome.[349] In this study, the goodness-of-fit of the StatinMD model was assessed by the Royston and Sauerbrei's R_D^2 statistic based on the model's D statistic, which takes a value between 0 and 1 measuring the proportion of the prognostic separation explained by the predictors in the model.[329, 338] A R_D^2 statistic closer to 1 indicates that the predictors better explain the prognostic separation in the cohort. Although the goodness-of-fit does not necessarily reflect the prediction performance of a model, it may justify the appropriateness of the inclusion of the explanatory predictors. The R_D^2 statistic of the

StatinMD model was calculated also using the self-developed R function 'royston.cpr' described above.

5.2.3.2.5 Assessment of clinical utility

As discussed in Chapter 4, the StatinMD model could be used to classify individuals at low or high risk of serious muscle disorders, to implement a stratified statin treatment strategy that involves addressing the concern about serious muscle disorders in high-risk individuals who are considered for statin treatment. The potential clinical utility of this stratified strategy with the StatinMD model was explored through decision curve analysis, measuring classification accuracy, and inspecting risk distribution in the population.[330] Since the eventual clinical use of the stratified strategy with the StatinMD model would be in unknown individuals who are not involved in the model derivation, the clinical utility was therefore assessed in the external validation cohort only.

- **Decision curve analysis**

In clinical practice, when making a decision on statin treatment, some doctors may be cautious about potential adverse events and tend to take clinical actions (e.g. test of blood creatine kinase, examination of muscle strength) to address the concern about serious muscle disorders in all individuals, which is a strategy of 'intervention for all'. While others may be bolder to prescribe statins for individuals eligible for treatment, regardless of the concern about serious muscle disorders, which is a strategy of 'intervention for none'. Compared to these one-size-fits-all strategies, the potential clinical benefit of the stratified strategy using the StatinMD model to determine these 'interventions' (clinical actions) was assessed by the decision curve analysis.[333]

The decision curve analysis calculates the net benefit of a decision-making strategy, which is the proportion of the targeted cases in the population that are correctly identified

and receive a relevant intervention (true positives), offset by the non-cases that are misclassified to receive the intervention (false positives).[333, 350] For an 'intervention for none' strategy, the numbers of true positives and false positives are both zero and therefore there is no net benefit. For an 'intervention for all' strategy, the net benefit varies by the offset ratio that depends on the clinical perspective about the trade-off between the benefit from treating a true case and the harm from treating a false case, and the maximum net benefit equals to the observed prevalence of the targeted cases when the offset ratio is zero. For a stratified strategy with a classification model, the net benefit also varies by the offset ratio, which is believed to correspond to the clinical preference for the risk threshold used for the classification. For example, if the clinical perspective is that the benefit from treating one true case is so important that it could be achieved at the cost of the harm from treating 9 false cases, the offset ratio is 1:9 and the risk threshold is 10% to classify low and high individuals for different decisions on the intervention. As such, the calculated net benefit of a strategy can be plotted against the varying risk threshold (equivalently the offset ratio used in the calculation), which forms the decision curve for the strategy.

In this study, the decision curve for the stratified strategy with the StatinMD model was drawn based on the net benefits calculated from the individuals' predicted risk from the StatinMD model and the observed cases of serious muscle disorders in the cohort, in comparison to the curves for the 'intervention for all' and 'intervention for none' strategies described above, to demonstrate the potential advantage of the stratified strategy. The 'intervention' (clinical actions) for addressing concerns about serious muscle disorders may be more often considered in current statin users and such 'intervention' may include withdrawal of statin treatment. In these individuals, a common strategy to determine the 'intervention' could be based on patients' experience of

suspected muscle problems. However, experiencing muscle problems, particularly those common mild muscle symptoms, does not indicate serious muscle disorders and the 'intervention' may be unnecessary. To further understand the clinical benefits of the stratified strategy with the StatinMD model in this decision-making scenario, the decision curve of the stratified strategy in the current statin users in the cohort was compared to the decision curve of the strategy based on patients' experience of muscle problems. To better interpret the potential advantage of the stratified strategy with the StatinMD model over other strategies, the net benefit was also transformed to the number of avoided interventions, which emphasises the relative number of false positives reduced, compared to the 'intervention for all' strategy.[351] This is particularly useful in the decision-making scenario in current statin users, where the stratified strategy with the StatinMD model may help avoid some unnecessary statin withdrawals in individuals with experience of muscle symptoms but at low risk of serious muscle disorders.

- **Classification accuracy**

The accuracy of using the StatinMD model to classify individuals at low or high risk of serious muscle disorders with specific risk thresholds was measured by a set of relevant statistics, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR-).[334, 352, 353] The specific risk thresholds for assessment were chosen within the range where the stratified treatment strategy with the StatinMD model had a positive net benefit, as shown in the decision curve analysis described above.

Among the measurements, sensitivity and specificity are useful for comparing the overall accuracy of different classification tests or the accuracy of a classification test with different risk thresholds at a population level, but in terms of the accuracy of a

classification result for an individual, they are informative only when the value is very high.[352] A positive result from a classification test with a very high specificity could confidently rule in a true case, and a negative result from a test with a very high sensitivity could rule out a non-case. PPV and NPV are more intuitive for understanding the accuracy of a test result for an individual, as they indicate the probability of an individual being a true case given a positive result or being a non-case given a negative result.[334] However, their values could be significantly influenced by the outcome incidence in the population as they essentially reflect the post-test probability of the outcome and PPV could be low when the cases are overall rare in the population, such as serious muscle disorders in this study. LR+ and LR- are more robust when the outcome incidence is extreme as they indicate the odds rather than the probability of a true case given a positive or negative test result.[354] It has been suggested that a test with an LR+ >10 is generally good at ruling in outcome cases and a test with an LR- <0.1 is good at ruling out non-cases, but LR+ may need to be much larger for ruling in the cases of a rare outcome.[353]

All of these statistics are linked and need to be interpreted together. For the StatinMD model, the main purpose of using it for classification is to identify (rule in) high-risk individuals who will develop serious muscle disorders and the outcome is rare. In this case, high specificity and large LR+ would be expected to satisfy the purpose of classification.

The accuracy statistics were estimated using the inverse probability of censoring weighting approach by the 'timeROC' package in R.

- **Overall utility of the risk stratification in the population**

To understand the overall utility of the risk stratification with the StatinMD model in the population, the distribution of individuals' predicted 10-year risk of serious muscle

disorders and CVD in the population eligible for statin treatment according to the current clinical guideline in the UK (i.e. with 10-year CVD risk over 10%) was inspected in a scatter plot. Individuals' CVD risk was predicted from the QRISK2 model that was commonly implemented in the UK during the study period, as recorded in the CPRD primary care data, and those individuals without a record of the CVD risk in the cohort were excluded from this inspection. The proportion of the individuals who would be classified as high-risk for serious muscle disorders by the StatinMD model with potential risk thresholds was calculated.

5.2.3.3 Handling missing data

The internal validation used the imputed data of the derivation cohort created in Chapter 4. The missing data in the external validation was handled by the same approaches as in the model derivation. Briefly, for the variable (deprivation) with a small proportion of missing data, the missing values were imputed using single imputation by the mode in the observed values. For the four variables (ethnicity, alcohol consumption, smoking status, BMI level) with more missing data, the missing values were imputed using the multivariate imputation by chained equations (MICE) with ten imputations.[174]

All the assessments described above were conducted in each imputed dataset and the results were pooled across the ten datasets using the recommended methods.[170] Specifically, the individuals' prognostic index, the D statistic, and the calibration slope were pooled using the Rubin's rules.[167] The *c*-index, the E/O ratio, and the R_D^2 statistic were summarised by median with interquartile range (IQR) across the ten imputed datasets. The calibration plots and the decision curves were created based on the pooled prognostic index.

5.3 Results

5.3.1 Comparison of the derivation and external validation cohorts

5.3.1.1 Baseline characteristics

The baseline characteristics of the model derivation and external validation cohorts are reported in **Table 5.1**. The external validation cohort included a total of 3,889,504 individuals followed up for a median of 8.5 years (IQR: 3.5 – 10.0 years), which was around two years longer than the median length of follow-up in the derivation cohort. The demographics were similar to that of the derivation cohort. The prevalence of some health status indicators or comorbidities, including general or high-risk alcohol consumption, above-normal BMI level, CVD, hypertension, previous muscle problems, and degenerative joint disorders, was lower than that in the derivation cohort. The use of the statin-interactive drugs and the myotoxic drugs was less common than in the derivation cohort, but the proportion of individuals on statin treatment was similar in both cohorts.

Table 5.1 Baseline characteristics of the derivation and external validation cohorts

Baseline Characteristics	Derivation Cohort	Validation Cohort
Total number of patients	1,785,207	3,889,504
Years of follow-up (median with IQR)	6.6 (2.7-10.0)	8.5 (3.5-10.0)
Demographics		
Age (mean with SD)	64.0 (11.5)	62.0 (11.2)
Gender		
Male	999,969 (56.0%)	2,224,742 (57.2%)
Female	785,238 (44.0%)	1,664,762 (42.8%)
Ethnicity		
White	1,193,371 (66.9%)	2,241,421 (57.6%)
Black	18,238 (1.0%)	81,044 (2.1%)

South Asian	22,619 (1.3%)	75,740 (1.9%)
Mixed and other	27,186 (1.5%)	77,224 (2.0%)
Missing	523,793 (29.3%)	1,414,075 (36.4%)
Deprivation		
Level 1 (least deprived)	423,260 (23.7%)	889,783 (22.9%)
Level 2	409,934 (23.0%)	830,618 (21.4%)
Level 3	383,727 (21.5%)	773,322 (19.9%)
Level 4	314,029 (17.6%)	717,981 (18.4%)
Level 5 (most deprived)	252,628 (14.1%)	673,054 (17.3%)
Missing	1,629 (0.1%)	4,746 (0.1%)
Health Status Indicators		
Alcohol consumption		
Non-drinker and ex-drinker	279,213 (15.6%)	816,140 (21.0%)
General drinker	973,783 (54.6%)	1,469,858 (37.8%)
High-risk drinker	119,158 (6.7%)	479,618 (12.3%)
Missing	413,053 (23.1%)	1,123,888 (28.9%)
Smoking		
Non-smoker	850,660 (47.7%)	1,649,606 (42.4%)
Ex-smoker	359,202 (20.1%)	729,622 (18.8%)
Current smoker	320,494 (17.9%)	786,981 (20.2%)
Missing	254,851 (14.3%)	723,295 (18.6%)
BMI level		
Normal	495,609 (27.8%)	966,648 (24.9%)
Underweight	23,949 (1.3%)	62,235 (1.6%)
Overweight	542,577 (30.4%)	1,030,612 (26.5%)
Obese	279,006 (15.6%)	635,987 (16.3%)
Morbidly obese	26,274 (1.5%)	61,387 (1.6%)
Missing	417,792 (23.4%)	1,132,635 (29.1%)
Frailty index (mean with SD)	0.06 (0.06)	0.05 (0.06)

Comorbidities		
Cardiovascular diseases	267,710 (15.0%)	452,040 (11.6%)
Hypertension	470,050 (26.3%)	881,681 (22.7%)
Diabetes	136,048 (7.6%)	290,777 (7.5%)
Mild liver diseases	4,308 (0.2%)	11,051 (0.3%)
Kidney diseases	67,540 (3.8%)	135,112 (3.5%)
Rheumatic arthritis	22,325 (1.3%)	41,117 (1.1%)
Previous muscle problems	85,581 (4.8%)	119,043 (3.1%)
Degenerative joint disorders	343,640 (19.3%)	590,163 (15.2%)
Hypothyroidism	87,452 (4.9%)	159,357 (4.1%)
Vitamin D deficiency	7,739 (0.4%)	41,447 (1.1%)
Vitamin B12 deficiency	26,224 (1.5%)	62,772 (1.6%)
Concomitant Medications		
Statin-interactive drugs	156,335 (8.8%)	249,831 (6.4%)
Myotoxic drugs	362,595 (20.3%)	702,280 (18.1%)
Statin Treatment		
No statins	1,537,894 (86.1%)	3,415,527 (87.8%)
Atorvastatin	74,350 (4.2%)	162,847 (4.2%)
Rosuvastatin	6,555 (0.4%)	13,401 (0.3%)
Simvastatin	149,326 (8.4%)	266,818 (6.9%)
Fluvastatin/Pravastatin	17,082 (0.9%)	30,911 (0.8%)

IQR: interquartile range; SD: standard deviation; BMI: body mass index; CVD: cardiovascular disease.

5.3.1.2 Outcome incidence and competing risk

In the external validation cohort, 13,549 outcome events (serious muscle disorders) were observed over the 10-year follow-up, leading to a crude event rate of 0.35% and an incidence rate of 5 per 10,000 person-years. A total of 701,686 competing events (deaths from causes other than serious muscle disorders) were observed in the 10 years with a crude event rate of 18% and an incidence rate of 264 per 10,000 person-years. The 10-

year cumulative incidence of the outcome was estimated to be 0.44% (95% CI: 0.44% – 0.45%), after adjusting by the competing risk. This was similar to that in the derivation cohort, which was 0.45% (0.44% – 0.47%). In contrast, the 1-year and the 5-year cumulative incidences of the outcome were slightly higher than that in the derivation cohort. For the competing event, the 10-year cumulative incidence was estimated to be 23.4% (95% CI: 23.3% – 23.4%), and it was lower than that in the derivation cohort, which was 33.2% (95% CI: 33.1% – 33.3%). Similarly, the 1-year and the 5-year cumulative incidences of the competing event were also lower than that in the derivation cohort. The estimated cumulative incidences of the outcome and the competing event in both cohorts are reported in **Table 5.2**.

Table 5.2 Cumulative incidences of the outcome and the competing event

Cumulative Incidence (95% CI)	Derivation Cohort	Validation Cohort
Outcome event		
1-year	0.04% (0.04% - 0.04%)	0.07% (0.06% - 0.07%)
5-year	0.20% (0.19% - 0.20%)	0.24% (0.24% - 0.25%)
10-year	0.45% (0.44% - 0.47%)	0.44% (0.44% - 0.45%)
Competing event		
1-year	3.5% (3.5% - 3.5%)	2.7% (2.7% - 2.7%)
5-year	16.7% (16.7% - 16.8%)	12.1% (12.0% - 12.1%)
10-year	33.2% (33.1% - 33.3%)	23.4% (23.3% - 23.4%)

CI: confidence interval

5.3.1.3 Prognostic index

The distributions of the individuals' prognostic index in the derivation and the external validation cohorts are presented in **Figure 5.1**. The average level and the variation of the individuals' prognostic index in the external validation cohort (mean = 1.000, SD = 0.864) were similar to that in the derivation cohort (mean = 0.998, SD = 0.915).

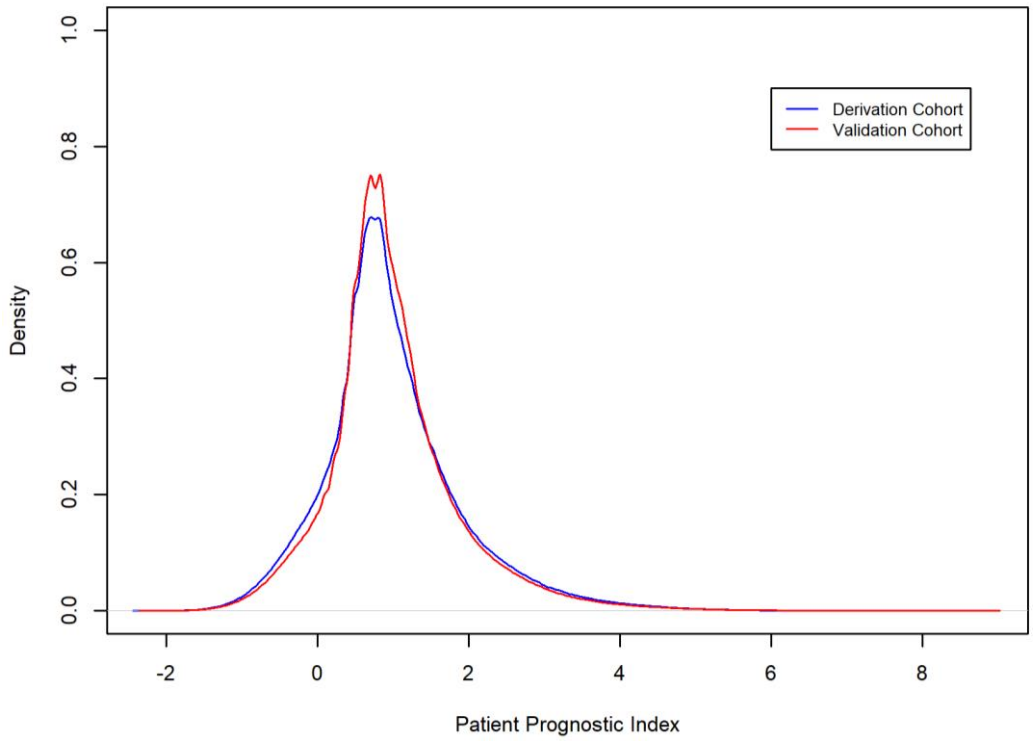


Figure 5.1 Distribution of prognostic index in the derivation and validation cohorts

5.3.2 Prediction performance of the StatinMD model

The measurements of the prediction performance of the StatinMD model are reported in **Table 5.3** and the estimates in each imputed dataset are provided in **Appendix 30**.

5.3.2.1 Model discrimination

The StatinMD model showed good discrimination for the 1, 5, and 10-year predictions in both the internal (apparent) and external validations. In the external validation, for the 10-year prediction, the *c*-index (0.782) indicated that the model had a probability of about 78% to correctly distinguish between individuals with and without the outcome, and the D statistic (2.176) indicated that an individual with a higher predicted risk from the model was more than 2 times more likely to develop the outcome than an individual with a lower predicted risk. The discrimination was better for the 1-year (*c*-index = 0.840, D statistic = 2.659) and 5-year predictions (*c*-index = 0.806, D statistic = 2.419).

Table 5.3 Prediction performance of the StatinMD model in the internal (apparent) and external validations

Performance Measurements*		1-Year Prediction		5-Year Prediction		10-Year Prediction	
		Internal	**External**	**Internal**	**External**	**Internal**	**External**
Discrimination	*c*-index	0.790 (0.789 - 0.790)	0.840 (0.839 - 0.840)	0.788 (0.788 - 0.789)	0.806 (0.806 - 0.806)	0.775 (0.775 - 0.775)	0.782 (0.782 - 0.782)
	D statistic	2.234 (2.232 - 2.242)	2.659 (2.578 - 2.740)	2.200 (2.197 - 2.202)	2.419 (2.374 - 2.464)	2.013 (2.012 - 2.014)	2.176 (2.139 - 2.212)
Calibration	E/O ratio	1.016 (0.946 – 1.099)	0.596 (0.574 – 0.620)	1.078 (1.041 – 1.119)	0.853 (0.836 – 0.872)	1.160 (1.130 – 1.191)	1.151 (1.131 – 1.171)
	calibration slope	1.070 (1.026 - 1.113)	1.146 (1.126 – 1.165)	1.067 (1.045 - 1.089)	1.113 (1.102 – 1.125)	1.008 (0.994 - 1.023)	1.063 (1.052 – 1.073)
Goodness-of-fit	R_D^2	0.544 (0.543 - 0.546)	0.628 (0.628 - 0.628)	0.536 (0.535 - 0.536)	0.583 (0.583 - 0.583)	0.492 (0.491 - 0.492)	0.530 (0.530 - 0.531)

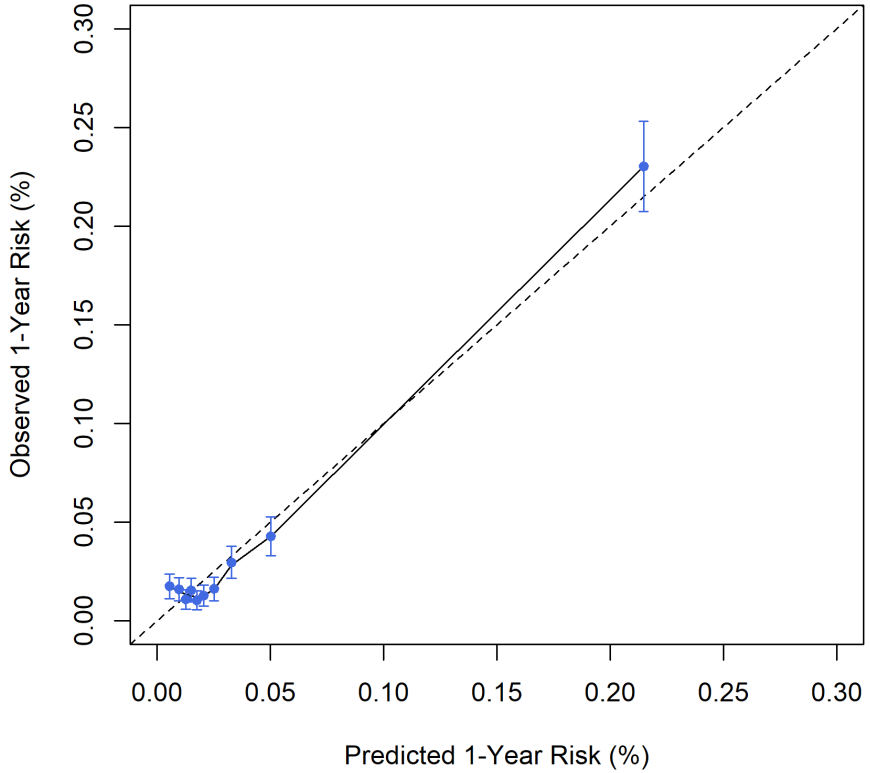
* The estimate of the c-index, E/O ratio, and R_D^2 was the median with interquartile range across the ten imputed datasets; the estimate of the D statistic and calibration slope was the mean with 95% confidence interval across the ten imputed datasets.

5.3.2.2 Model calibration

As presented in **Table 5.3**, in the external validation, the StatinMD model showed a small overestimation of the 10-year risk on average in the whole cohort (E/O = 1.151, calibration slope = 1.063), which was similar to that shown in the internal validation. In contrast, the model showed an overall underestimation of the 1-year risk (E/O = 0.596, calibration slope = 1.146) and the 5-year risk (E/O = 0.853, calibration slope = 1.113), which was different from the overall good calibration for the 1-year and 5-year predictions in the internal validation.

The calibration plots by risk group are presented in **Figure 5.2 – 5.4** and the risks in each group are reported in **Appendix 31**. In the external validation, the calibration plot for the 10-year prediction showed good agreement between the averaged predicted risk and the observed outcome incidence in each of the ten risk groups. This was similar to the internal validation. For the 1-year and 5-year predictions, the averaged predicted risk was also close to the observed incidence in most risk groups, but there was some underestimation in the highest-risk group. In contrast, the calibration in the internal validation was shown to be good in all risk groups for the 1-year and 5-year predictions.

A. Internal (apparent) validation



B. External validation

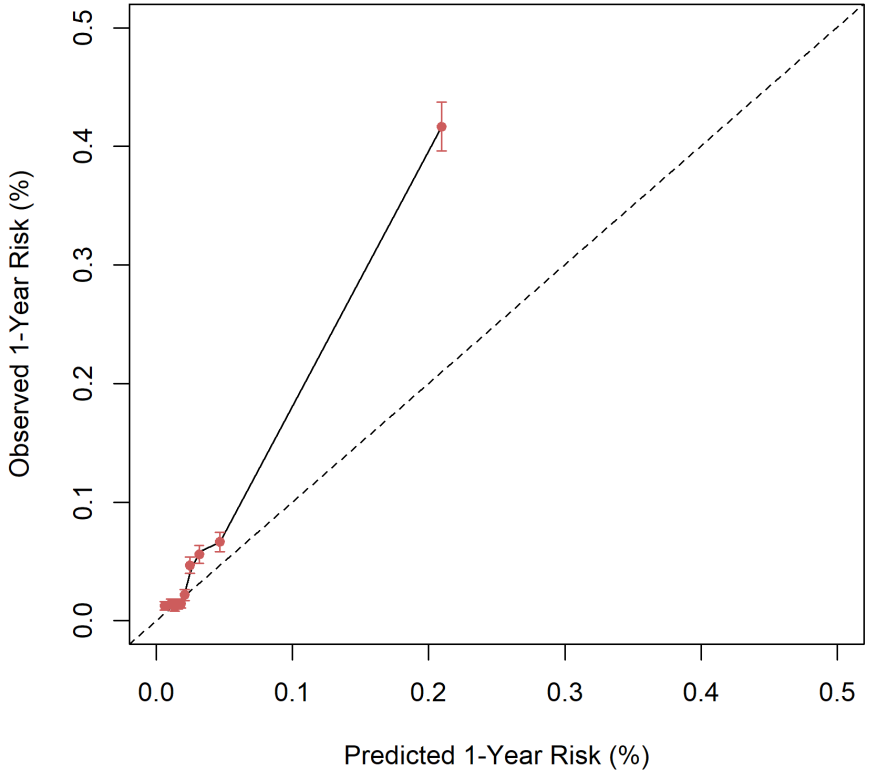
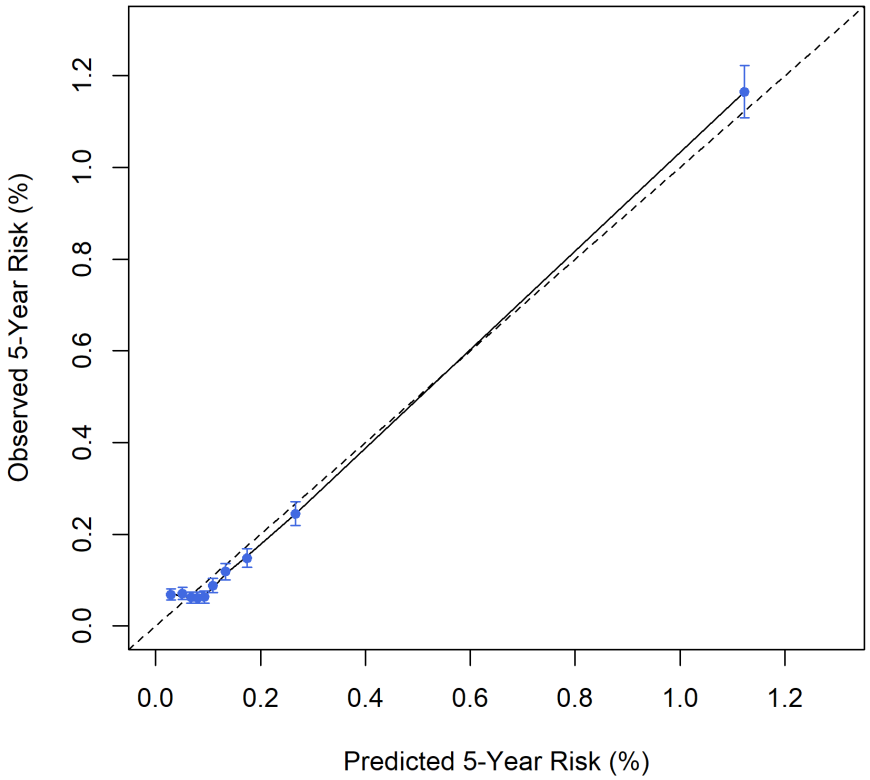


Figure 5.2 Calibration plots by risk group for 1-year prediction by the StatinMD model in the internal and external validations

The blue and red dots represent the ten risk groups in the internal (apparent) and external validation cohorts, defined by the deciles of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 1-year risk of serious muscle disorders from the StatinMD model and the observed incidence calculated by the Aalen-Johan estimator in each risk group. The vertical bar through each dot denotes the 95% confidence interval of the Aalen-Johan estimate of the observed incidence.

A. Internal (apparent) validation



B. External validation

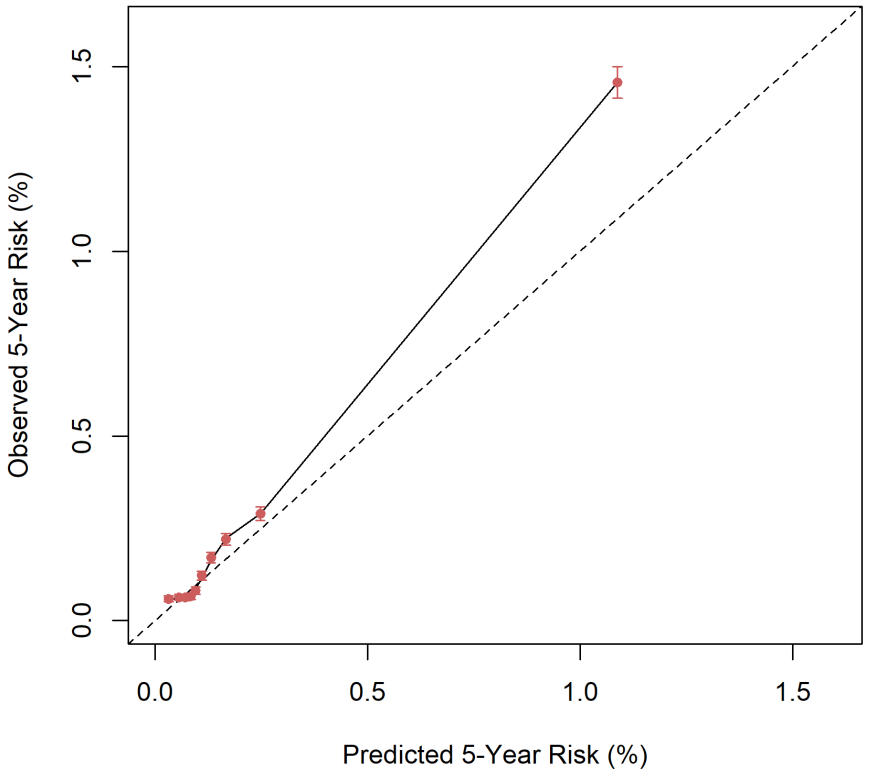
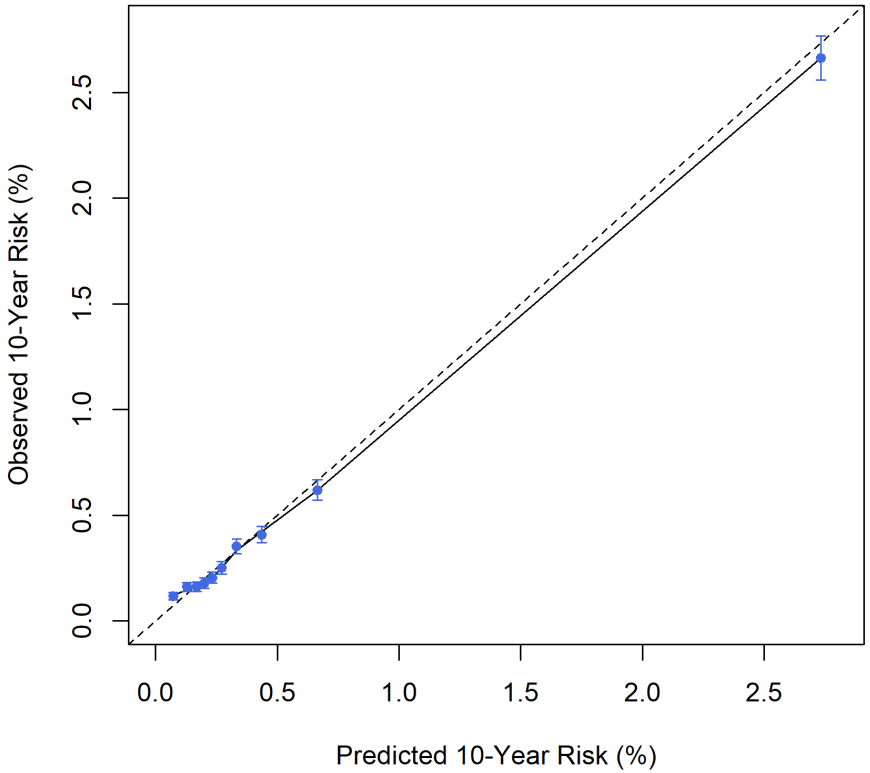


Figure 5.3 Calibration plots by risk group for 5-year prediction by the StatinMD model in the internal and external validations

The blue and red dots represent the ten risk groups in the internal (apparent) and external validation cohorts, defined by the deciles of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 5-year risk of serious muscle disorders from the StatinMD model and the observed incidence calculated by the Aalen-Johan estimator in each risk group. The vertical bar through each dot denotes the 95% confidence interval of the Aalen-Johan estimate of the observed incidence.

A. Internal (apparent) validation



B. External validation

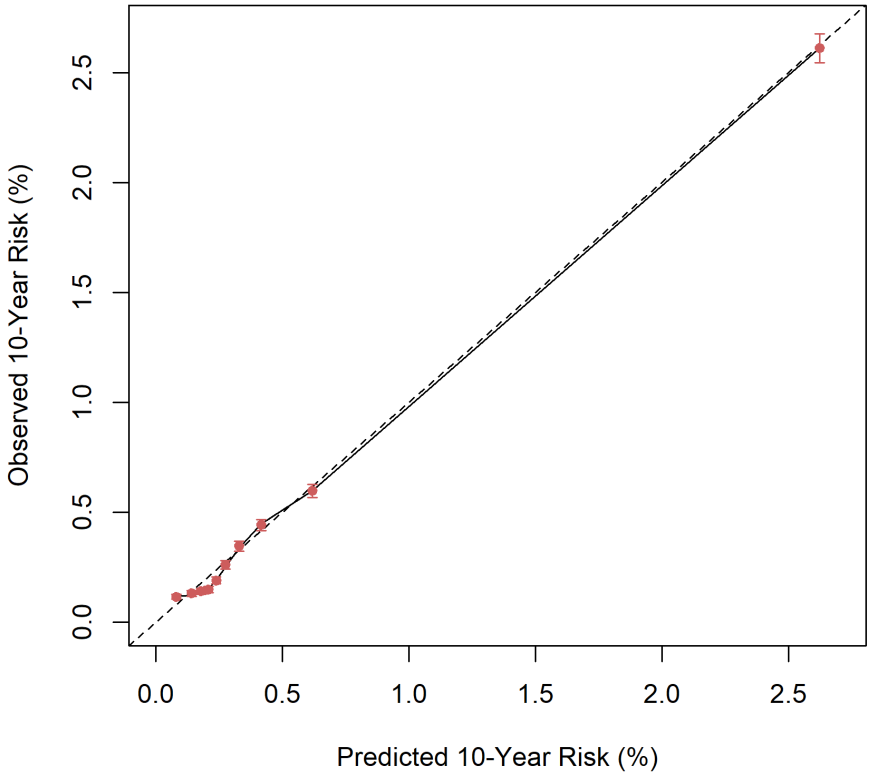


Figure 5.4 Calibration plots by risk group for 10-year prediction by the StatinMD model in the internal and external validations

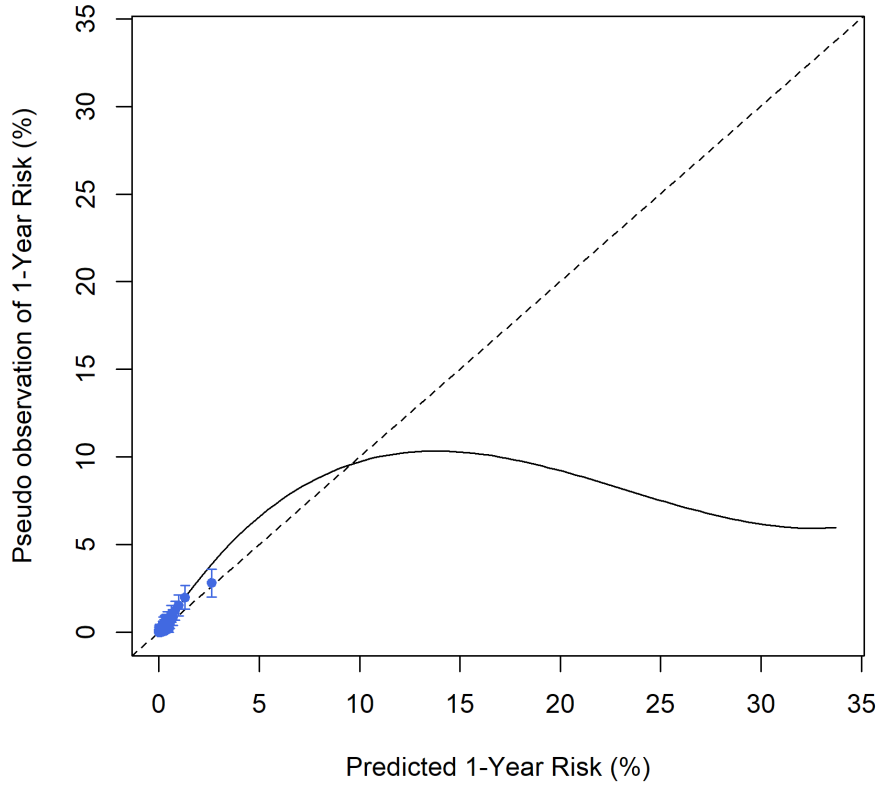
The blue and red dots represent the ten risk groups in the internal (apparent) and external validation cohorts, defined by the deciles of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 10-year risk of serious muscle disorders from the StatinMD model and the observed incidence calculated by the Aalen-Johan estimator in each risk group. The vertical bar through each dot denotes the 95% confidence interval of the Aalen-Johan estimate of the observed incidence.

The calibration plots by individual are presented in **Figure 5.5 – 5.7**. In the external validation, the calibration plots for the 1-year, 5-year, and 10-year predictions presented a similar distribution and all showed good agreement between the predicted and the observed risks in individuals with a low predicted risk, who made up the majority of the cohort. This confirmed the good calibration in most risk groups shown in the calibration plots at the group level. There was a small underestimation in individuals with a moderate predicted risk and a potentially significant overestimation in individuals with a high predicted risk. This revealed the unseen miscalibration in the highest risk group in the group-level calibration plots, where the underestimated or overestimated risks of these individuals were averaged together with others' risks in the group. Potentially significant overestimation was shown in those individuals with a 1-year predicted risk over around 15%, a 5-year risk over 25%, or a 10-year risk over 35%, who respectively accounted for about 0.01%, 0.02%, and 0.05% of the cohort. The individual-level calibration plots in the internal validation were similar but the degree of miscalibration shown in the plots was slightly smaller than in the external validation.

5.3.2.3 Model goodness-of-fit

As reported in **Table 5.3**, the model showed moderate goodness-of-fit in the external data, which was better than that in the derivation data. Specifically, the R_D^2 (0.530) indicated that the model explained 53% of the 10-year prognostic separation of the outcome in the external cohort. This explained proportion by the model was larger for the 1-year (R_D^2 = 0.628) and the 5-year prognostic separation (R_D^2 = 0.583).

A. Internal (apparent) validation



B. External validation

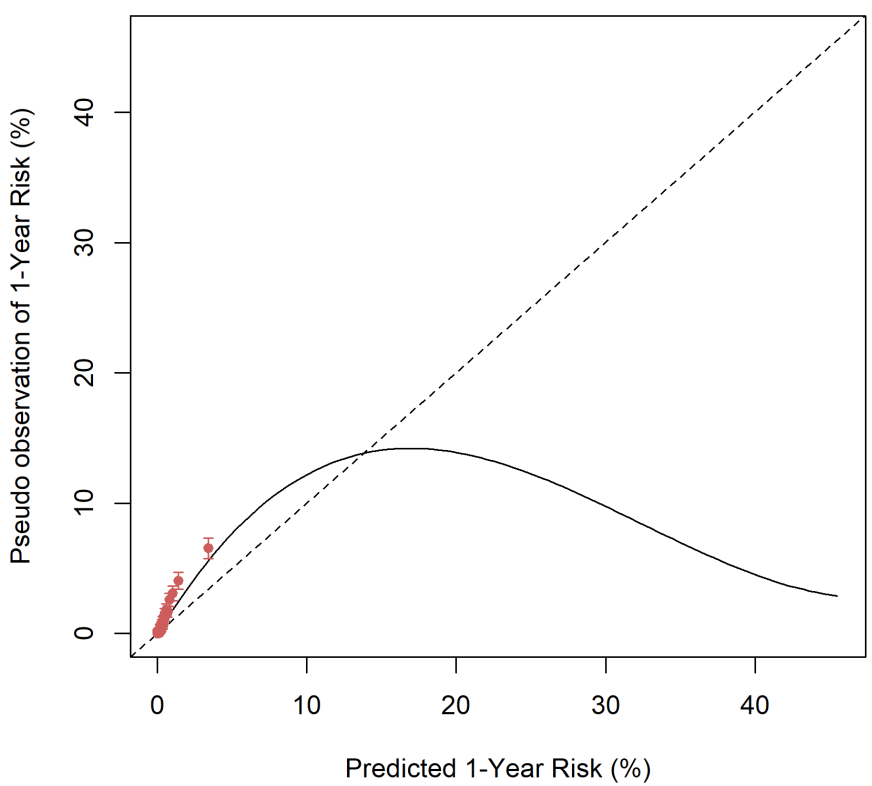


Figure 5.5 Calibration plots at the individual level for 1-year prediction by the StatinMD model in the internal and external validations

The calibration curves were fitted to the individuals' predicted 1-year risks from the StatinMD model and their pseudo values of the observed risk. The blue and red dots represent the 1,000 groups in the internal and external validation cohorts, defined by the 1,000th of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 1-year risk from the StatinMD model and the mean of the individuals' pseudo observations in each group. The vertical bar through each dot denotes the 95% confidence interval of the mean pseudo observation in each group. The tails of the curves were estimated based on sparse data on a few individuals with predicted risk above the mean level of the highest risk group, which may be inaccurate.

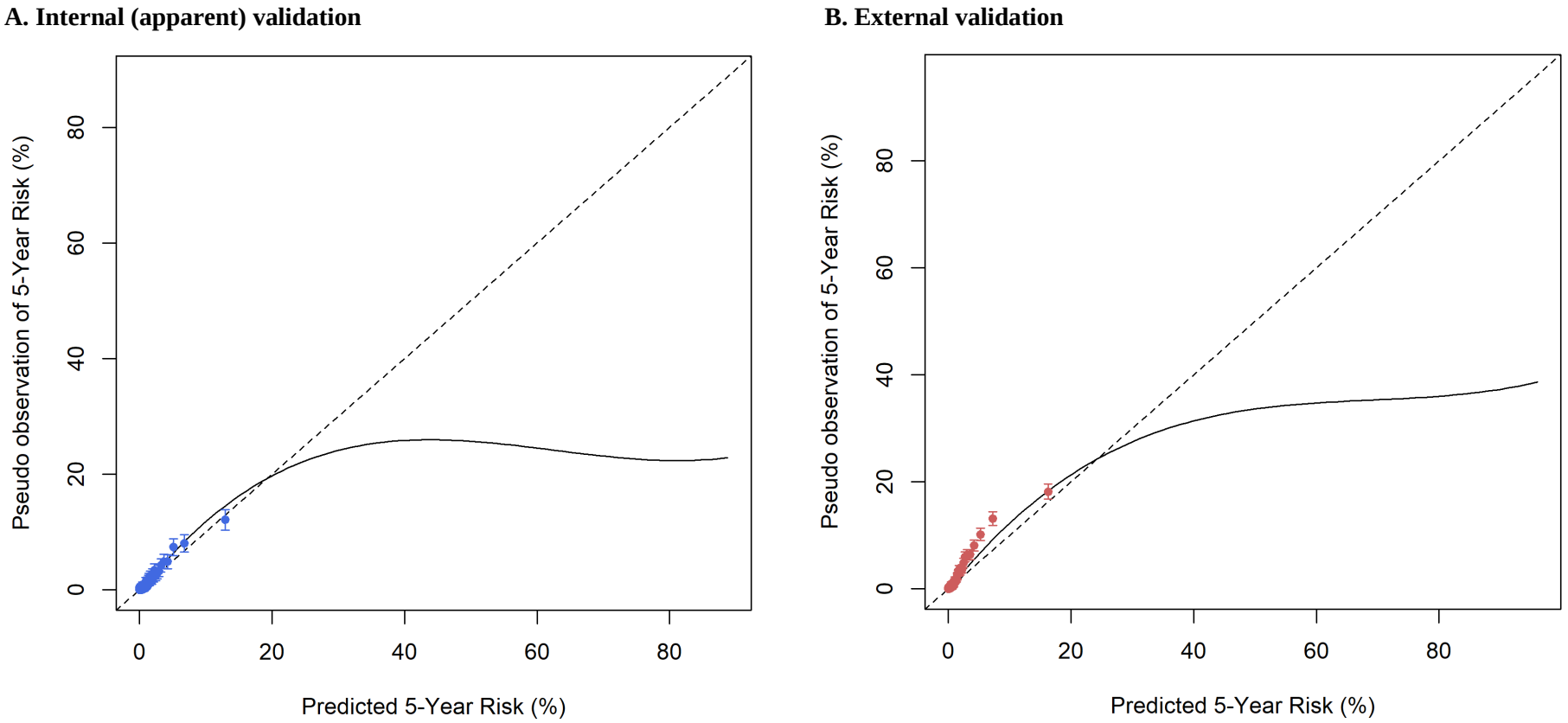
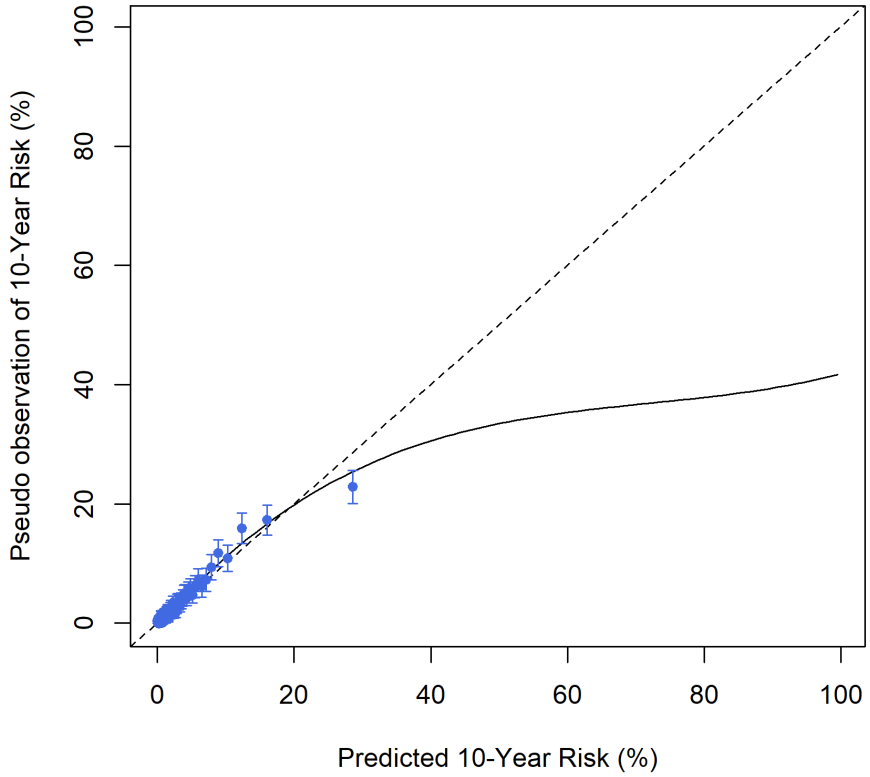


Figure 5.6 Calibration plots at the individual level for 5-year prediction by the StatinMD model in the internal and external validations

The calibration curves were fitted to the individuals' predicted 1-year risks from the StatinMD model and their pseudo values of the observed risk. The blue and red dots represent the 1,000 groups in the internal and external validation cohorts, defined by the 1,000th of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 5-year risk from the StatinMD model and the mean of the individuals' pseudo observations in each group. The vertical bar through each dot denotes the 95% confidence interval of the mean pseudo observation in each group. The tails of the curves were estimated based on sparse data on a few individuals with predicted risk above the mean level of the highest risk group, which may be inaccurate.

A. Internal (apparent) validation



B. External validation

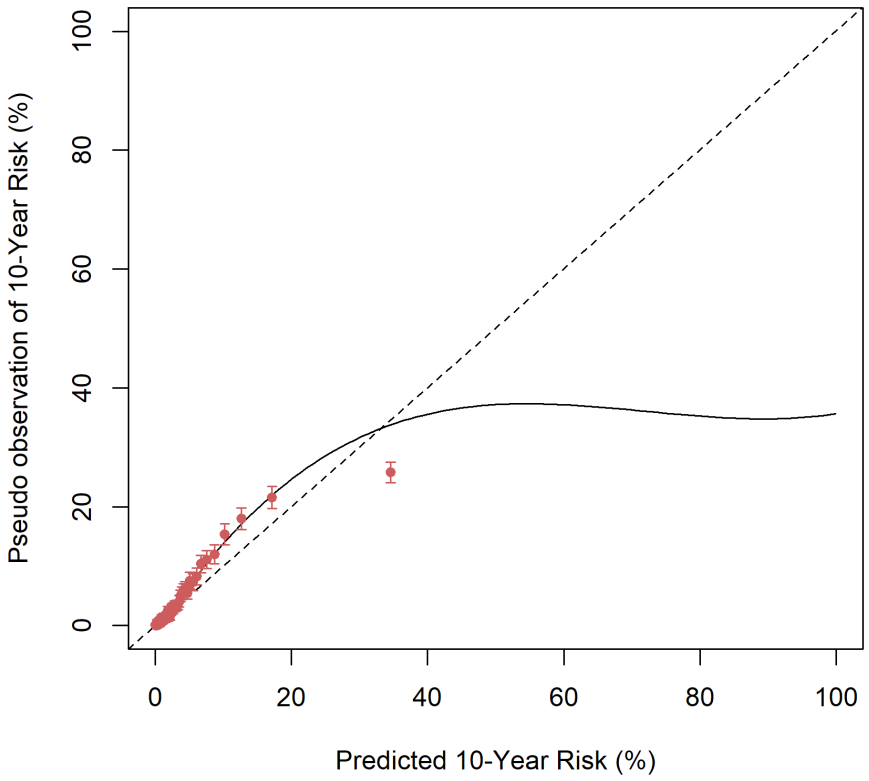


Figure 5.7 Calibration plots at the individual level for 10-year prediction by the StatinMD model in the internal and external validations

The calibration curves were fitted to the individuals' predicted 1-year risks from the StatinMD model and their pseudo values of the observed risk. The blue and red dots represent the 1,000 groups in the internal and external validation cohorts, defined by the 1,000th of the individual's prognostic index from the StatinMD model in each cohort. The coordinates of the dots correspond to the mean of the individuals' predicted 10-year risk from the StatinMD model and the mean of the individuals' pseudo observations in each group. The vertical bar through each dot denotes the 95% confidence interval of the mean pseudo observation in each group. The tails of the curves were estimated based on sparse data on a few individuals with predicted risk above the mean level of the highest risk group, which may be inaccurate.

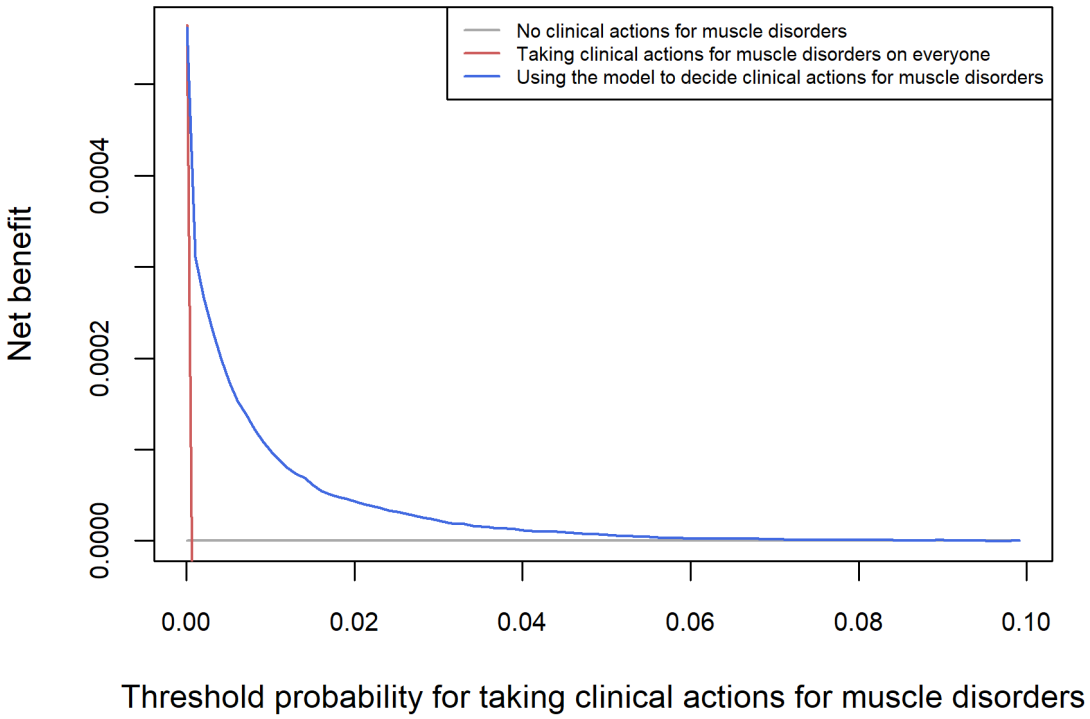
5.3.3 Potential clinical utility of the StatinMD model

5.3.3.1 Decision curve analysis

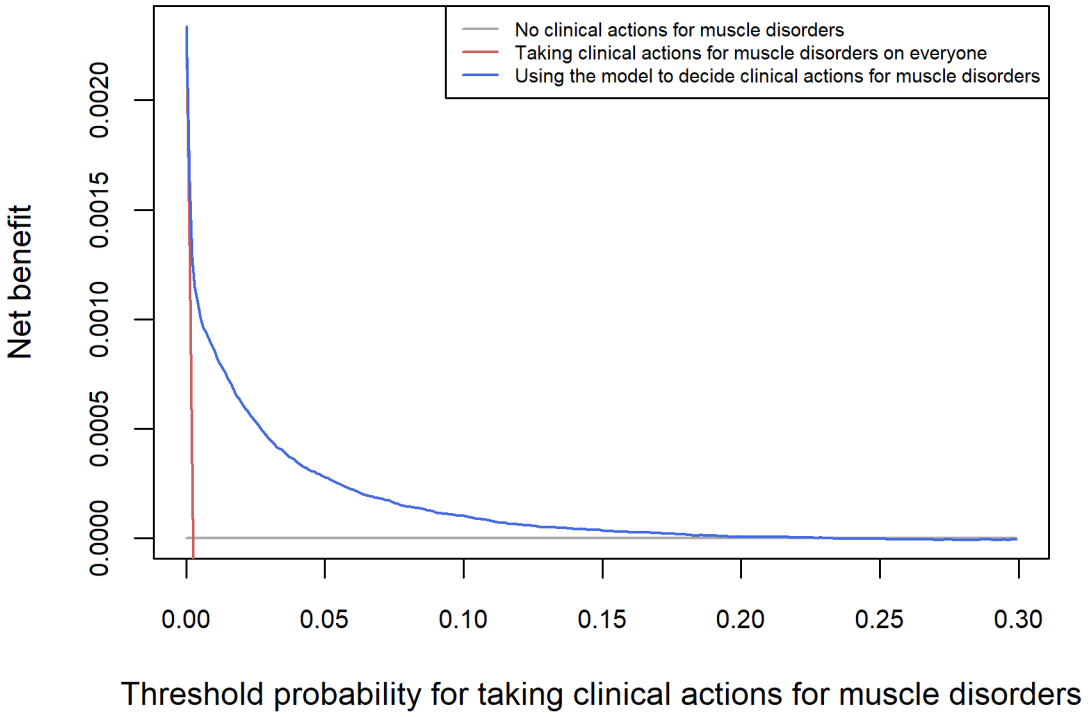
In the external validation cohort, the decision curves for the stratified decision strategy with the StatinMD model, which uses the predicted risk from the model to determine clinical actions for addressing the concern about serious muscle disorders when considering statin treatment, are presented in **Figure 5.8**. This stratified strategy was shown to have a positive net benefit when a risk threshold used to stratify the predicted 10-year risk of serious muscle disorders was within 30%. For instance, if the 10-year risk threshold was 10%, the net benefit of the stratified strategy shown in the decision curve was around 0.0005. This means that such risk stratification could correctly identify and decide on taking clinical actions in around 5 high-risk individuals who would develop serious muscle disorders in 10 years among 10,000 individuals eligible for statin treatment, after being offset by potential misclassification of individuals who would not develop muscle disorders. When considering the stratification of the 1-year and 5-year risks, the stratified strategy would have a positive net benefit if the threshold was within around 6% for the 1-year risk and 20% for the 5-year risk.

With a risk threshold within these ranges, the positive net benefit of the stratified strategy using the StatinMD model demonstrated its advantage over the 'intervention for none' strategy (with zero net benefits by definition) which is not to take clinical actions for potential serious muscle disorders in any individuals when considering statin treatment. The decision curves for the stratified strategy also lay above the curves for the 'intervention for all' strategy which is to take clinical actions in all individuals eligible for statin treatment, indicating that the net benefit of the stratified strategy was consistently larger than the 'intervention for all' strategy.

A. For 1-year risk stratification



B. For 5-year risk stratification



C. For 10-year risk stratification

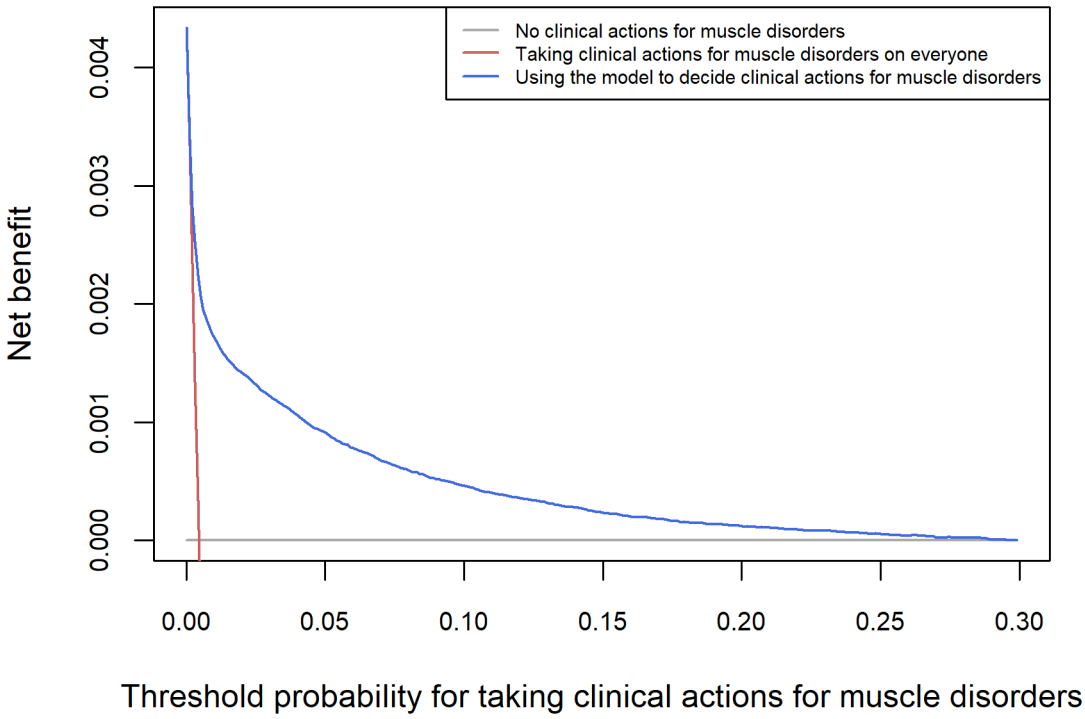


Figure 5.8 Decision curves of the stratified strategy using the StatinMD model in the external validation cohort

The advantage of the stratified strategy using the StatinMD model over the 'intervention for all' strategy was also shown by the number of avoided interventions (**Figure 5.9**). For example, with a 10% threshold for stratification of the 10-year risk, using the prediction from the StatinMD model to determine clinical actions for potential serious muscle disorders could avoid around 95 unnecessary interventions in 100 individuals eligible for statin treatment. These avoided interventions may include unnecessary pre-treatment tests of blood creatine kinase in individuals considered for statin initiation, inefficient clinical monitoring of muscle functions in current statin users, or inappropriate invasive examinations (e.g. muscle biopsy) and treatment withdrawals in statin users who may report suspected muscle symptoms. When considering the stratification of the 1-year and 5-year risks, the number of unnecessary interventions avoided by using the stratified strategy with the StatinMD model could be larger.

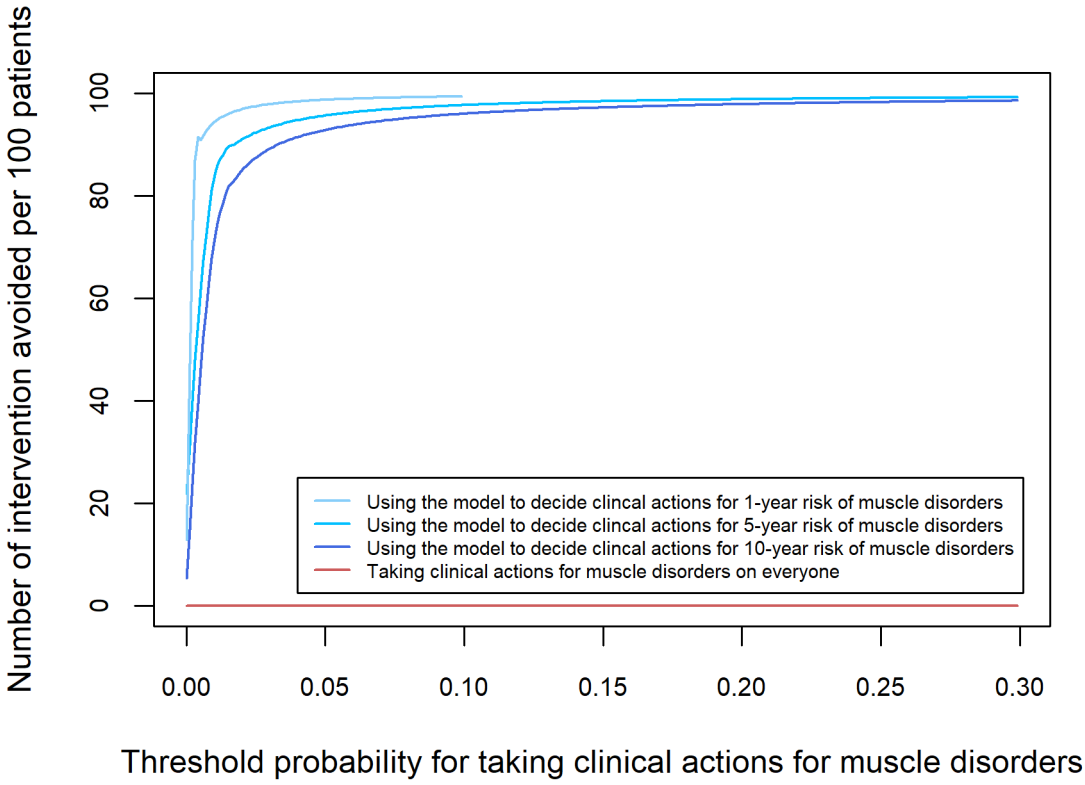


Figure 5.9 Number of avoided interventions by the stratified strategy using the StatinMD model in the external validation cohort

The decision curves for the stratified strategy with the StatinMD model among the statin users in the external validation cohort are presented in **Appendix 32**. These decision curves also illustrated the advantage of the stratified strategy over the 'intervention for none' and 'intervention for all' strategies. In addition, the stratified strategy was shown to have larger net benefits than the strategy that determines clinical actions for potential serious muscle disorders simply based on statin users' experience of muscle problems.

5.3.3.2 Classification accuracy

Within the range of the risk threshold where the stratified strategy with the StatinMD model could have a positive net benefit, as indicated in the decision curve analysis above, the classification accuracy of the StatinMD model at specific potential risk thresholds is reported in **Table 5.4**.

Table 5.4 Classification accuracy of the StatinMD model by potential risk thresholds

Risk Threshold	True Positive	False Positive	True Negative	False Negative	Sensitivity	Specificity	PPV	NPV	LR+	LR-
1-year risk										
1%	446	9,275	3,877,699	2,084	17.58%	99.79%	5.38%	99.95%	85.52	0.83
3%	121	1,406	3,885,568	2,409	4.76%	99.97%	9.83%	99.94%	163.94	0.95
5%	47	477	3,886,497	2,483	1.85%	99.99%	11.15%	99.93%	188.66	0.98
5-year risk										
1%	3,231	90,214	3,790,840	5,219	37.65%	98.44%	5.57%	99.85%	24.15	0.63
3%	1,735	20,575	3,860,479	6,715	20.06%	99.70%	14.04%	99.80%	66.85	0.80
5%	1,092	9,273	3,871,781	7,358	12.56%	99.88%	20.77%	99.79%	107.25	0.88
7%	739	5,253	3,875,801	7,711	8.47%	99.94%	25.91%	99.78%	143.06	0.92
10%	461	2,796	3,878,258	7,989	5.26%	99.97%	31.71%	99.77%	189.98	0.95
10-year risk										
1%	6,157	297,730	3,578,225	7,392	43.82%	95.52%	4.18%	99.74%	9.78	0.59
5%	2,998	34,706	3,841,249	10,551	20.69%	99.65%	20.88%	99.65%	59.23	0.80
10%	1,621	12,340	3,863,615	11,928	10.98%	99.90%	33.83%	99.60%	114.71	0.89
15%	978	6,226	3,869,729	12,571	6.50%	99.96%	43.42%	99.58%	172.16	0.94
20%	638	3,660	3,872,295	12,911	4.18%	99.98%	50.72%	99.57%	230.92	0.96

PPV: positive predictive value; NPV: negative predictive value; LR+: positive likelihood ratio; LR-: negative likelihood ratio.

The model showed low sensitivity and very high specificity with all the potential risk thresholds. This indicated the overall classification accuracy of the model in the population: only a small proportion of the individuals with serious muscle disorders could be identified by the model, while most of the individuals without serious muscle disorders could be correctly classified as low-risk by the model.

In terms of classification accuracy for an individual, the model showed a low to moderate PPV and a high NPV with most of the risk thresholds. This indicated that an individual classified as high-risk by the model (a positive result) had a low to moderate probability of having serious muscle disorders, which significantly raised the 'pre-test' probability (observed incidence) of serious muscle disorders in the cohort (1 year: 0.07%, 5 years: 0.24%, 10 years: 0.44%, see **Table 5.2**). In contrast, an individual classified as low-risk by the model (a negative result) had a high probability of being free of serious muscle disorders, which was similar to the 'pre-test' probability.

Also reflecting the classification accuracy of a given result, the model showed an LR+ much larger than 1 and an LR- close to 1 with most of the risk thresholds. This indicated that being classified as high-risk by the model (a positive result) is much more likely to happen in the individuals who would develop serious muscle disorders than in those who would not, while being classified as low-risk (a negative result) may equally happen in individuals who would and would not develop serious muscle disorders.

5.3.3.3 Overall utility of potential risk stratification in the population

Among the individuals eligible for statin treatment (i.e. with 10-year CVD risk over 10%), most of the individuals had a low predicted risk of serious muscle disorders and only a small proportion would be classified as high-risk individuals with the potential risk

thresholds (**Table 5.5**). For example, with a 10% risk threshold for the 10-year risk, only around 2% of the individuals would be classified as high-risk for serious muscle disorders.

Table 5.5 Proportion of high-risk individuals for serious muscle disorders classified by the StatinMD model prediction with potential risk thresholds

Risk Threshold	Proportion of High-risk Individuals
1-year risk	
1%	1.4%
3%	0.3%
5%	0.1%
5-year risk	
1%	9.5%
3%	2.9%
5%	1.5%
7%	1.0%
10%	0.5%
10-year risk	
1%	24.6%
5%	4.5%
10%	1.9%
15%	1.1%
20%	0.7%

The distribution of the predicted 10-year risk of serious muscle disorders and the 10-year CVD risk among these eligible patients is presented in **Figure 5.10**. This showed that most individuals at high risk of serious muscle disorders had a relatively low to moderate CVD risk and many of them had a risk of serious muscle disorders higher than their CVD risk. As the example in **Figure 5.10**, among the high-risk individuals classified

by a 10% threshold of the 10-year risk, 52.1% of them had a risk of serious muscle disorders higher than their CVD risk.

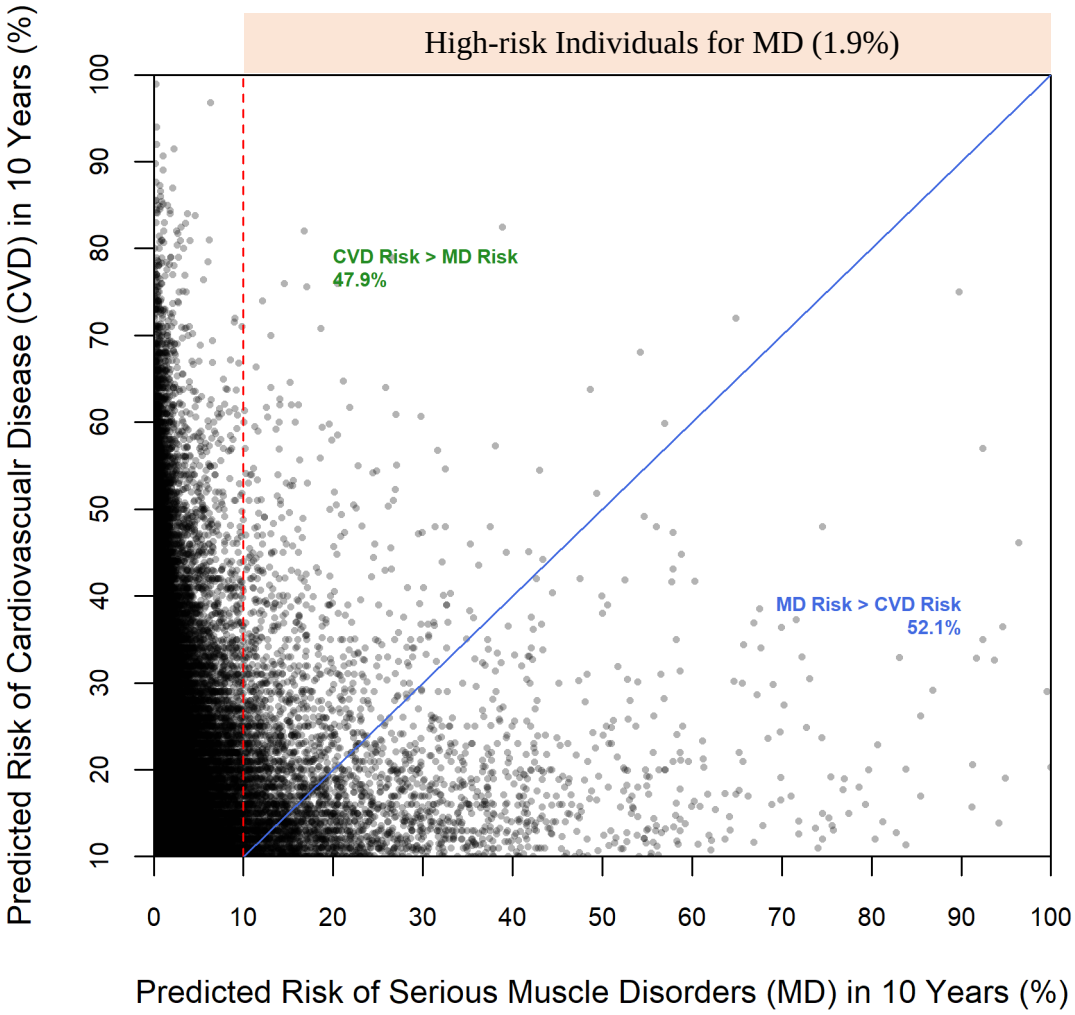


Figure 5.10 Distribution of individuals' 10-year risk of serious muscle disorders and CVD among those eligible for statin treatment in the external validation cohort

The 10-year risk of serious muscle disorders (MD) was predicted from the StatinMD model. The 10-year risk of cardiovascular disease (CVD) was predicted from the QRISK2 model, as recorded in the CPRD primary care data. The eligibility of these individuals for statin treatment is defined as the 10-year risk of CVD over 10% according to the NICE guideline. The red dashed line denotes the potential threshold to stratify the 10-year MD risk among these eligible patients. The blue line denotes that the MD risk equals the CVD risk for an individual. The area under the blue line is where the MD risk is higher than the CVD risk for an individual and the area above is where the CVD risk is higher than the MD risk. The percentages in blue and green are the proportion of the individuals in these two areas respectively among those individuals at a high MD risk above 10%.

5.4 Discussion

5.4.1 Summary of key findings

This study conducted an apparent validation of the StatinMD model for predicting the risk of serious muscle disorders in the model derivation cohort, and an external validation of the model in another cohort of 3,889,504 individuals who were potentially eligible for statin treatment. The model showed good apparent performance in the derivation cohort, demonstrating the appropriateness and robustness of the model derivation process. In the external validation, the model showed good discrimination for the 1, 5, and 10-year predictions (*c*-index: 0.840, 0.806, 0.782; D statistic: 2.659, 2.419, 2.176). The model presented an underestimation of the 1-year (E/O: 0.596, calibration slope: 1.146) and the 5-year risks (E/O: 0.853, calibration slope: 1.113) and a small overestimation of the 10-year risk (E/O: 1.151, calibration slope: 1.063) on average in the whole cohort. The calibration plots by risk group also showed an underestimation of the 1-year and the 5-year risks but good agreement between the predicted and the observed 10-year risks on average within each risk group. The calibration plots by individual further revealed that the model predictions in all three timeframes were accurate in most individuals with a low predicted risk and slightly underestimated the risk in some individuals with a moderate predicted risk. The model may overestimate the risk by a potentially large degree in a few individuals with a high predicted risk, who accounted for a very small proportion of the population.

Decision curve analysis showed that using the StatinMD model to stratify the risk of serious muscle disorders in the population and determine the corresponding clinical actions when considering statin treatment had a positive net benefit from correctly identifying high-risk individuals for necessary actions and could avoid a number of

unnecessary interventions in low-risk individuals. With specific potential thresholds for the risk stratification, the model presented high specificity, modest PPV higher than the population incidence of the outcome, and large LR+, indicating that the model was useful for identifying individuals who would develop serious muscle disorders from the population.

5.4.2 Comparison with previous studies

5.4.2.1 Performance of previous models for muscle-related adverse events

As introduced in Chapter 4, there were two known models previously developed to predict muscle-related adverse events in the context of statin treatment. The model for myopathy with simvastatin was claimed to be validated through internal cross-validation, but the validation process and model performance were not reported in the publication.[286] The QStatin model for the 5-year risk of myopathic events was assessed through internal and external validation.[62] Compared to the QStatin model, the StatinMD model in this study showed overall better performance for the 5-year prediction in both the internal and the external validation. The StatinMD model also provided predictions of the 1-year and 10-year risks, for which the model performance was also better than the performance of the QStatin model for its 5-year prediction. The following comparisons between the two models focused on their prediction performance in the same 5-year timeframe in the external validation.

Regarding discrimination, the *c*-index and the D statistic of the StatinMD model for the 5-year prediction (0.806 and 2.42) were both larger than that of the QStatin model (0.763 and 1.75 in women, 0.727 and 1.60 in men).[62] This suggests a better discriminative ability of the StatinMD model to distinguish between individuals who will and will not develop the defined outcome. This improvement in discrimination may be because the StatinMD

model included some strong predictors for muscle-related adverse events that were omitted in the QStatin model, such as previous muscle problems, degenerative joint disorders, and vitamin D deficiency, and therefore could better explain the differences between individuals with and without the outcome.

The calibration of the QStatin model was only graphically assessed by calibration plots based on ten risk groups. From the calibration plots, the current StatinMD model and the QStatin model both showed good agreement between the predicted and observed risks in most groups but some miscalibration in the highest-risk group.[62] The absolute amount of this miscalibration by the StatinMD model was smaller than that by the QStatin model. This may be because the StatinMD model focused on predicting more serious events and was also adjusted by the competing risk, both leading to a smaller predicted risk of the outcome on an absolute scale.

The proportion of the variation in the predicted outcome explained by the model, as indicated by the R^2 statistic, was larger for the current StatinMD model (58%) than the QStatin model (42% in women and 38% in men).[62] Although the R^2 statistic does not directly reflect the prediction performance, the larger value may suggest that the included predictors in the StatinMD model better explained the outcome occurrence.[355]

5.4.2.2 General performance of prognostic prediction models

Although measurement statistics and methods have been developed to assess the performance of prognostic prediction models, there are no definite criteria for good model performance based on these measurements.[330] The consideration for the model performance mainly depends on the purpose of prediction.[338]

For a model that would be used to identify high-risk individuals, the discriminative ability of the model to distinguish between individuals with and without the predicted

outcome is essential. For clinical prediction models, common values of the *c*-index for the discrimination are at least 0.65,[356] but the value is closer to 1 for some widely-accepted models. For example, the SCORE model for predicting fatal cardiovascular diseases (CVD) had a *c*-index ranging from 0.71 to 0.84 in different populations.[357] The QRISK2 model for predicting fatal and non-fatal CVD had a *c*-index of 0.82 in women and 0.79 in men.[59] As for the D statistic, which was more recently proposed and hasn't been reported in many published models so far, theoretically the larger the value is the better.[329, 356] One example is the QRISK2 model, which showed a D statistic of 1.79 in women and 1.62 in men.[59] A recently published model predicting falls in patients with an indication for antihypertensive treatment, using similar approaches as in this study, showed a D statistic ranging from 1.60 to 2.16 for its predictions in different timeframes.[358] Compared to these recognised models, the discrimination of the StatinMD model could be considered a good performance based on its *c*-index and D statistic.

For calibration, the E/O ratio could be influenced by the outcome incidence in the population, making it difficult to compare models for different outcomes. It has been suggested that an E/O ratio between 0.83 and 1.25 may indicate good overall calibration, using the example of the EuroSCORE model for predicting 30-day mortality after cardiac surgery which had an incidence of around 3%. However, for rare outcomes with a lower incidence, like serious muscle disorders in this study, a small difference between the predicted and the observed risks could result in an E/O ratio significantly departing from 1. The StatinMD model showed good overall calibration for the 5-year (E/O = 0.853) and the 10-year predictions (E/O = 1.151), but only modest calibration for the 1-year prediction (E/O = 0.593), which resulted from a small absolute difference between the averaged predicted risk (0.04%) and the observed risk (0.07%). Calibration slope does not by itself measure calibration and only has been used to help understand model

calibration recently, with inconsistent and inappropriate interpretations.[344] It should be interpreted together with the overall calibration indicated by the E/O ratio and is more useful when good overall calibration has been shown by an E/O ratio close to 1. In this study, the calibration slope did not add much about the model performance but the slope being close to 1 suggested there was no need to re-calibrate the prognostic index (predictor coefficients).[338, 359]

When the calibration is assessed graphically by a calibration plot, which is commonly drawn based on ten risk groups, it is even more difficult to define good performance.[140] In this study, the calibration plots of the StatinMD model showed good calibration across most risk groups but some miscalibration in the highest risk group. This was also seen in some well-recognised models such as the QRISK2.[59] However, the miscalibration in the highest risk group did not deter the wide use of the QRISK2 model in clinical practice,[5] since the absolute risk difference resulted from the miscalibration was small. For the StatinMD model, the absolute amount of miscalibration in the highest risk group was also small on average, but the calibration by individual further revealed that the overestimation in a few individuals with a very high predicted risk may be large. Such individual-level calibration was not used for the QRISK2 or other well-recognised models and has been only reported in prognostic modelling studies recently.[358, 360]

When a model was assessed through both internal and external validation, theoretically, a model's internal performance should be better than its external performance, since the model has been derived from the same data.[144] However, the current StatinMD model showed slightly better discrimination in the external validation cohort than in the derivation cohort. This was also seen in other prognostic models. For example, the SCORE model has been shown to have better discrimination in an external population from a low-risk region than in the original derivation population.[357] The discrimination

of the QStatin model for myopathic events was also slightly better in the external validation than in the internal validation.[62] In the current study, one explanation may be because the observed outcome risk had a wider spread in the external cohort, as shown in the risks across the ten risk groups in the two cohorts (**Appendix 30**), and therefore they were easier to be distinguished by the model.

5.4.3 Strengths and limitations

5.4.3.1 Strengths

- **Model calibration at the individual level**

As a common practice, model calibration is usually graphically assessed by a calibration plot based on ten risk groups in the population, where the averaged predicted and observed risks in each group are compared.[140] However, these group-level risks usually do not reveal the calibration across the full range of plausible probabilities for all individuals in the population. For example, in the calibration plot of the QRISK2 model for the 10-year CVD risk, the averaged risks in the ten risk groups ranged from around 1% to 25% in women and from 1% to 30% in men.[59] Although this range may represent the risk for the majority of the population, the model calibration in the individuals with a risk over 30% was left unknown. It is possible that the calibration in these individuals is not as good as in others, as the model is derived with often much fewer data in these 'outliers', which probably leads to less precise predictions in these individuals.

In this study, in addition to the calibration by risk group, the calibration of the StatinMD model based on individuals' pseudo observations was also presented, which has not been widely used in the validation of clinical prediction models.[361] This calibration covered the full range of plausible outcome risks in all individuals and provided useful supplementary information about the model calibration. The individual-level calibration

plots confirmed the good accuracy of the model predictions in individuals with a low predicted risk who represented the majority of the population and further revealed the miscalibration in some individuals in the highest risk group. This miscalibration was not seen in the calibration plots by risk group, where the predicted risks of these underestimated or overestimated individuals were averaged with others' risks in the highest risk group. More specifically, this revealed miscalibration showed that, within the highest risk group in the population, there was a small underestimation by the model in individuals with a moderate predicted risk but a potentially significant overestimation in those with a very high risk above a certain level. Knowing this additional information helps better understand the full picture of the model calibration across the whole population and the degree of potential miscalibration in individuals at different risk levels. This could direct where cautions need to be taken when using the predicted risk from the model to make a clinical decision for an individual and is also an important consideration when determining the threshold for risk stratification based on the model predictions, as discussed in the clinical implication session below.[362]

- **Comprehensive assessment of prediction performance and clinical utility**

A systematic review of external validation studies for prediction models has shown that the key aspects of model performance have been poorly reported in publications, with some important measurements often omitted, such as calibration measurements.[356] In the current study, a comprehensive assessment of the performance of the StatinMD model, including the discrimination, calibration, and goodness-of-fit, was performed, with specific measurements fully reported, following the TRIPOD guidance.[143] In addition, this study also assessed the potential clinical utility of the StatinMD model, which has been proposed but not yet widely taken up in the validation of clinical prediction models.[330] The decision curve analysis used in this study is a relatively new approach,

which demonstrated the advantage of using the StatinMD model for relevant decision-making over common practices and further established its validity for clinical use.[333]

5.4.3.2 Limitations

- **Methodological issues in the assessment of model performance**

Overall the StatinMD model underestimates the risk for low-risk individuals and overestimates the risk for high-risk individuals. This may be due to the sparsely distributed data of outcome occurrence that is likely to lead to the range of predicted risk across the population being too wide.[363] Some approaches could be used to adjust such miscalibration by shrinking the predictor coefficients to narrow down the range of predicted risk.[364] This study did not make such an adjustment because the amount of underestimation of a low risk in the majority of the population was small and the overestimation of a high risk would only happen in a very small proportion of the population, which is unlikely to invert the treatment decision in most cases. The estimates of calibration slope did not indicate a need for re-calibration either. Nevertheless, although the potential improvement may be limited, an attempt to re-calibrate the model may reduce the miscalibration to a certain degree.

The validation data in this study were from the primary care records collected in the GP practices across England and there may be some heterogeneity in the model performance among different GP practices due to potentially different case mix of patient characteristics or varying incidence of the outcome in these practices.[365] This potential heterogeneity was not taken into account when assessing the model performance in this study, due to the consideration for analysis efficiency. To address this heterogeneity, a more sophisticated approach would be assessing the model performance within each GP practice and combining the performance measurements across all practices using meta-

analysis.[365] By this approach, the prediction intervals for the estimates of performance measurements could be obtained to give an indication of expected model performance in a new GP practice, which would be wider than the confidence intervals shown in the current results.[358]

This study used the *c*-index as a measurement of model discrimination, which has been widely used in the validation of prognostic models.[356] The *c*-index reflects the average discriminative ability of the model across all possible risk thresholds in the full theoretical range from 0% to 100% for discrimination between high-risk and low-risk individuals.[339] Such an average value may be less meaningful when the model prediction in the population is limited to a small range or the model is to be used with a specific risk threshold in practice.[366] In this study, among 99% of the individuals in the external validation cohort, the predicted risk from the StatinMD model is within the range of 0% - 5%, where the model discrimination may be different from the average discrimination in the full range. The partial *c*-index has been proposed to measure discrimination in a probability range of interest, but it also has limitations and is less interpretable.[366, 367] The StatinMD model is expected to be used with a specific risk threshold to identify high-risk individuals, in which case the discrimination may be also different from the average discrimination indicated by the *c*-index.[368] Although several potential risk thresholds were picked in this study to assess the model discrimination (classification accuracy) in specific scenarios, the risk threshold that would be eventually used in clinical practice remains undetermined and the actual model discrimination is unknown.

Although the model calibration at the individual level is useful for better understanding the model performance across the whole population, as pointed out above, this approach may also introduce some uncertainty and potential bias from the generation of individuals' pseudo observations. The generation of the pseudo values was based on an assumption

that the time to censoring (the censoring other than by the competing event) is completely independent of the time to outcome event.[347] This assumption is more restrictive than the assumption established for a survival model, which allows the time to censoring to be conditionally independent of the time to outcome event, given a set of covariates in the model.[362] Although the censorship in this study was arguably independent of the outcome occurrence, according to the defined reasons for censoring (except the competing event), it is difficult to confirm or test this assumption. Approaches to addressing potential dependent censoring include computing pseudo values separately in the subsets stratified by categorical covariates that are suspected to be related to the dependent censoring mechanism, or modelling the censoring mechanism conditional on a set of covariates.[347] However, these approaches would add much complexity to the calculation of pseudo values, especially when competing risk exists.[369] In fact, without using these approaches, the pseudo value calculation in this study has already encountered computation issues, due to the huge sample size and the memory requirement for the intensive computation. The whole sample was therefore divided into fifty subsamples to compute pseudo values respectively. This compromised computation process may also introduce bias, in addition to the potential violation of independent censoring assumption, into the generated pseudo values and consequently the calibration assessment based on these values.

In addition, the calibration plots by individual showed a potentially significant overestimation in high-risk individuals for the predictions of all three timeframes in both the apparent and the external validation. These high-risk individuals accounted for a very small proportion of the population, for whom the model may not be able to predict the risk accurately based on the scarce data. But for the same reason, the large overestimation in these individuals shown by the calibration curves fitted to the scarce data on their

predicted risks and pseudo values may also be inaccurate and the degree of this overestimation remained uncertain.

Another potential methodological issue in the assessment of the model performance is that the calculation of the D statistic and the R_D^2 statistic was performed in R using the function 'royston.cpr' that was developed by myself based on the relevant methodology papers and the function 'royston' from the widely-used R package 'survival',[154, 329, 370] given there is currently no available package or function in R to calculate these two statistics for a competing risk model. However, this function has not been tested or validated in other studies.

- **Influence by the use of data from routine healthcare records**

As pointed out in the preceding chapter, using the data from routine healthcare records has some limitations, which may not only affect the model derivation but also influence the model validation. Missing data of certain predictor variables in both the derivation and the external validation cohorts, although imputed, may leave uncertainty in the measurements of the model prediction performance in the internal and external validation.[163] Potential misclassification of the outcome and the baseline characteristics in the cohorts may lead to some bias in the model performance assessment.[287] Moreover, since the primary care records in the derivation and the external validation cohorts were sought from different databases, the baseline characteristics were defined by different medical codes in the two cohorts, which needs to be taken into account when interpreting the validation results. The difference seen in the internal and external validation results may partially result from the inconsistent definitions and recording of the predictor variables, in addition to the actual variation of model performance in the different populations.

- **Sample similarity and non-independent validation**

The key purpose of external validation is to evaluate the generalisability of a model, in order to support its wider use beyond the population from where the model has been derived. In this study, although the data for external validation were from a different data source, the population characteristics in the external validation cohort were very similar to the derivation cohort. In this case, a good performance in the external validation further demonstrated the reproducibility of the StatinMD model, in addition to the internal validation, rather than its generalisability.[145] Since the external validation cohort, as well as the derivation cohort, was from the primary care population in the UK, the generalisability of the StatinMD model may be limited to this setting. The model may not perform as well in other settings, such as in other countries or the secondary care population with different characteristics.

Furthermore, an independent validation by investigators who are not involved in the model derivation is encouraged to reduce potential inflated findings and 'spin'.[143, 144] However, in this study, the StatinMD model was derived and validated by the same author, which may lead to better validation results than in an independent validation.[356]

5.4.4 Clinical implications

5.4.4.1 Considerations in the use of the StatinMD model for personalised decision-making on statin treatment

As suggested in Chapter 4, the StatinMD model could be used to support personalised clinical decision-making on statin treatment, through comparison between the predicted risks of CVD and serious muscle disorders and calculation of the absolute treatment effects by statins. For this purpose, the calibration of the StatinMD model is crucial, as it

reflects how close the predicted risk is to the actual risk for an individual and could influence the risk comparison and the treatment effect calculation.

This study showed that the StatinMD model could accurately predict the risk of serious muscle disorders in 1, 5, and 10 years for individuals with a low predicted risk, who represent the vast majority of the population eligible for statin treatment. The predicted risk for these individuals could be used to reassure them that their risk of serious muscle disorders is low, which may help encourage the uptake of statins or improve adherence to statin treatment in eligible patients. The model would slightly underestimate the risk for individuals with a moderate predicted risk but the absolute difference from their actual risk could be negligible. The predicted risk in these individuals is reliable for making a comparison with their CVD risk and calculating the potential risk increase of serious muscle disorders by statins, which could support the decision-making on statin treatment.

However, the model may overestimate the risk for individuals with a high predicted risk, potentially by a considerable degree. This would happen when the predicted 1-year risk is above around 15%, the 5-year risk is above 25%, or the 10-year risk is above 35%, which could occur respectively in 1, 2, or 5 persons per 10,000 people. In this case, it may not be appropriate to use the exact predicted risk of serious muscle disorders to compare with the CVD risk or calculate the absolute risk increase by statins. However, caution for statin treatment should still need to be taken for these individuals because their actual risk of serious muscle disorder is very likely to be higher than most people. For the individuals with an overestimated 10-year risk above 35%, their actual risk may be at least 25%, as shown in the individual-level calibration plots in this study.

5.4.4.2 Potential utility of the StatinMD model for stratified statin treatment

Another use of the StatinMD model is to stratify the risk of serious muscle disorders in the population to implement a potential stratified statin treatment strategy and determine corresponding clinical actions to address the concern about serious muscle disorders. For this purpose, the discriminative ability of the StatinMD model and the potential clinical utility of such risk stratification with the model are important considerations.

The StatinMD model showed a strong discriminative ability to distinguish between individuals who will and will not develop serious muscle disorders in 1, 5, or 10 years. A clear prognosis separation also demonstrated that individuals with a higher predicted risk from the StatinMD model are more likely to experience serious muscle disorders than those with a lower predicted risk.

The potential threshold to stratify the risk of serious muscle disorders in the population may lie in the range with an upper limit of around 6% for the 1-year risk, 20% for the 5-year risk, and 30% for the 10-year risk, as indicated by the decision curve analysis. With a risk threshold within these ranges, using the StatinMD model prediction to implement the stratified decision strategy would have clinical benefits from correctly targeting clinical actions at high-risk individuals who will develop serious muscle disorders. In the meantime, it could help avoid a number of unnecessary clinical actions in low-risk individuals, such as inefficient clinical monitoring of CK level during statin treatment or inappropriate treatment withdrawal in statin users due to the concern about muscle disorder. The benefits are more evident and important for the decision-making for statin users who have experienced suspected muscle symptoms, for whom these clinical actions are most likely to be considered.

With a specific risk threshold within the ranges above, the classification accuracy of the StatinMD model suggested that the model can help better identify individuals with serious muscle disorders from the population. For an individual classified as high-risk by the StatinMD model, clinicians can be confident that this individual is more likely to develop serious muscle disorders than those classified as low-risk and should consider the corresponding clinical actions described above. However, clinicians should be aware that an individual classified as low-risk by the model may also develop serious muscle disorders. Given that serious muscle disorders are rare in the population and the exaggerated concerns about statin safety have resulted in low uptake of statins in eligible patients, the number of serious muscle disorder cases missed out by the classification would be very small and this should be less concerning than giving false alarms to most individuals who actually will not develop serious muscle disorders.

Since the eligibility for statin treatment is determined based on the 10-year CVD risk, one may first consider the risk of serious muscle disorder predicted in the same timeframe when using the StatinMD model to assist the treatment decision-making. Considering the prediction performance of the StatinMD model in all aspects and the potential clinical utility of the stratified statin treatment strategy, a sensible threshold to stratify the predicted 10-year risk of serious muscle disorders may be around 10%. This threshold is within the range indicated by the decision curve analysis, where the risk stratification would have clinical benefits. Compared to a higher risk threshold (e.g. 15% or 20%), the classification accuracy of the StatinMD model with the 10% threshold is a better balance between a low rate of false alarms and a limited number of cases missed. The model calibration for a predicted risk of around 10% is also better than the calibration at a higher risk level, which is critical for correctly classifying individuals with a predicted risk around the threshold. However, since the threshold of CVD risk for statin treatment may

be lowered in the coming updated clinical guideline in the UK, or patients may have a strong preference to avoid serious muscle disorders, a lower threshold to stratify the predicted risk of serious muscle disorders could also be considered.

The overall clinical utility of such risk stratification may be limited to a small proportion of the population. As seen in this study, if the threshold for the 10-year risk is 10%, the classified high-risk individuals may account for only 2% of the population eligible for statin treatment. However, more than half of these high-risk individuals would have a risk of serious muscle disorders higher than their CVD risk, for whom the trade-off between benefits and harms is more likely to alter the treatment decision. Therefore, identifying these high-risk individuals is important even though their proportion is small.

5.4.5 Future research

5.4.5.1 Validation in other populations

Given the similarity between the derivation and the external validation cohorts in this study, the generalisability of the StatinMD model may need to be further extended or firmly established. This could be achieved by assessing the model performance in a population at different geographic locations (e.g. in other countries) or clinical settings (e.g. secondary care).[145] It could be also useful to validate the model in particular patient groups of concerns, such as the older population aged over 80 years for whom polypharmacy is common and the risk of adverse events needs to be carefully considered. Based on the validation results, the StatinMD model may be modified or updated to suit the use in other populations, without deriving a new model. A successful example of such extended validation and use is the Framingham risk score for CVD, of which several modified versions after validation have been recommended in the clinical guidelines in different countries.[371]

5.4.5.2 Further determination of the risk threshold

In addition to the considerations of the performance and clinical utility of the StatinMD model, a cost-effectiveness assessment of the potential stratified statin treatment in the population could be conducted to further determine the risk threshold for serious muscle disorders.[372] Like the health economic assessment determining the threshold of CVD risk for statin treatment, this may take into account lifetime clinical outcomes, quality-adjusted life-years (QALYs), costs, and incremental cost-effectiveness ratio (ICER) of using the model.[373]

Patients' experiences and preferences are always important in clinical decision-making.[374] A survey by questionnaire, focus group discussion, or individual interview may be designed to collect patients' beliefs and preferences about the beneficial and adverse effects of statins.[375] This could help understand how patients weigh the increase of serious muscle disorders against the reduction of CVD and further determine a risk threshold that is meaningful to patients' decisions.

5.4.5.3 Clinical impact study

The ultimate purpose of developing a prognostic prediction model is to positively impact clinical decision-making and consequently improve patient outcomes. To achieve this purpose, a well-validated model is fundamental. However, this does not guarantee the use of the model will actually change clinicians' or patients' behaviours in clinical practice and lead to an improved health outcome.[138] Prospective comparative studies, such as cluster-randomised trials, can be designed to evaluate the ultimate impact of a prediction model on patient outcomes, which are often referred to as 'impact studies'.[376] Impact studies often require substantial time and funding and have barely been conducted for the published prediction models, including those in current clinical use.[138]

Nevertheless, an impact study would be helpful for further understanding how the use of the StatinMD model could change clinical decisions on statin treatment and the eventual outcome of CVD prevention with statins in the population.

5.5 Conclusion

The StatinMD model has presented a satisfactory prediction performance to predict serious muscle disorders for individuals potentially eligible for statin treatment. The model has a strong discriminative ability to distinguish individuals who will and will not develop serious muscle disorders. It also has an overall good calibration to accurately predict the risk of serious muscle disorders for the majority of the population, which could help assist in personalised decision-making on statin treatment. The model showed some miscalibration in a small proportion of the population, which is unlikely to bias the decision-making on statin treatment. With a possible risk threshold to stratify the risk of serious muscle disorders in the population, the model could help better identify individuals who will develop muscle disorders. Implementing the stratified statin treatment strategy based on the risk stratification with the model could better target the clinical actions for addressing the concern about serious muscle disorders at high-risk individuals, while avoiding unnecessary actions or inappropriate treatment decisions.

Chapter 6 Discussion and Conclusions

6.1 Summary of key findings in this thesis

As introduced at the beginning of this thesis, the widespread concerns about statin safety have resulted in low uptake of statins and poor adherence to statin treatment for prevention of cardiovascular diseases (CVD).[34, 35, 66] The recommendation for wider use of statins in primary prevention has posed a particular challenge, with controversies focusing on the balance between benefits and harms of treatment.[32, 36] In this population, tailoring statin regimens by drug type or dose to address the safety concerns may be attempted in clinical practice but currently lacks guidance to inform decision-making.[51] Personalised treatment decision-making and stratified treatment strategy, taking into account the risk of adverse events, are the potential approaches towards better use of statins for prevention of CVD.[56, 63] These require prediction tools to provide personalised information. As the most commonly discussed adverse event of statins, serious muscle disorders are a major concern in the clinical decision-making and policy-making for statin treatment and a priority for developing a prediction tool.

This thesis aimed to better understand the safety of statins in primary prevention of CVD, by assessing the association between statins and common adverse events and exploring the evidence for tailoring statin regimens to address safety concerns. Another objective of this thesis was to derive and validate a prediction tool for serious muscle disorders in the population eligible for statin treatment.

6.1.1 Statin safety in primary prevention of CVD

A systematic review of 62 randomised controlled trials (RCT) on statin treatment in primary prevention patients without previous cardiovascular events was conducted in this

thesis, with a total of 120,456 participants followed up for an average of 3.9 years. The pair-wise meta-analysis comparing statins with placebo or usual care found that statin treatment was associated with an increased risk of self-perceived muscle symptoms, liver dysfunction, renal insufficiency, and eye conditions, but the association with clinically-confirmed muscle disorders or type 2 diabetes was not found in these patients. The absolute increase in the risk of the associated adverse events was small and did not exceed the reduction in the risk of major cardiovascular events by statin treatment.

The network meta-analysis of statin drug types showed that atorvastatin and lovastatin were associated with an increased risk of liver dysfunction and rosuvastatin was associated with a higher risk of muscle symptoms, renal insufficiency, type 2 diabetes, and eye conditions. Few significant differences between statin drug types were found in their associations with the adverse events, based on limited data for some statin types, particularly fluvastatin and pitavastatin. The E_{max} model-based meta-analysis identified a significant dose-response relationship suggesting a higher dose of atorvastatin was associated with a higher risk of liver dysfunction but with considerable uncertainty in the E_{max} model parameters to determine the explicit relationship. No dose-response relationships were found between other statins and adverse events, based on scarce data for individual statin types of specific doses.

6.1.2 Prediction of the risk of serious muscle disorders

A prediction model for the personalised risk of serious muscle disorders in the population potentially eligible for statin treatment (the StatinMD model) was developed in this thesis. Based on 1,785,207 individuals in primary care in the UK, the model was derived to predict the risk of serious muscle disorders associated with hospitalisation or death in 1, 5, or 10 years, taking into account the competing risk of death from other

causes. The risk prediction was based on 22 personal characteristics, including essential demographics, indicators of general health status, comorbidities, and use of statins and other medications, which were available in routine primary care records.

The StatinMD model presented a good prediction performance in the original population for model derivation, demonstrating the robustness and reproducibility of the model. The model was externally validated in another primary care population of 3,889,504 individuals in the UK. In this population, the model showed good discrimination to distinguish between individuals who would and would not develop serious muscle disorders. The model also showed good calibration in the majority of the population with low predicted risk. There was a small underestimation by the model in some individuals with a moderate predicted risk, and a potentially large overestimation in a few individuals with a high predicted risk who accounted for a very small proportion of the population. The decision curve analysis indicated that using the StatinMD model to stratify the risk of serious muscle disorders in the population had a positive net benefit of correctly targeting relevant clinical actions at high-risk individuals, while avoiding unnecessary interventions in low-risk individuals. With a specific potential threshold for risk stratification, the assessment of classification accuracy suggested that the StatinMD model was useful for identifying individuals who would develop serious muscle disorders from the population.

6.2 Clinical implications of the findings

6.2.1 Supporting evidence-based statin treatment for primary prevention of CVD

The findings from the systematic review in this thesis provide evidence that should, on top of previous knowledge, help address the controversies over the benefit-harm balance of statin treatment for primary prevention of CVD and support the recommendations for the use of statins in this population.

Firstly, the causal associations between statins and most adverse events are not well supported by current evidence, compared to their established beneficial effects on reducing cardiovascular events, particularly when used for primary prevention. According to the Bradford-Hill criteria,[377] the strength and consistency of the association and the biological plausibility and gradient of the effect are key considerations for causal inference. In the systematic review in this thesis, the associations between statins and the adverse events in primary prevention patients were found to be weak and their significance was not robust due to the low statistical power of the analyses. These associations were inconsistent among the individual trials included in this review and some of them were not found significant in previous studies in other populations or clinical settings.[44, 181, 189] Unlike the well-studied effect of statins on lowering low-density lipoprotein cholesterol (LDL-C), the underlying biological mechanisms for most of their potential adverse effects have not been well understood.[10, 12] The lack of clear dose-response relationships between statins and the adverse events, as shown in the review in this thesis, may also diminish the causal plausibility of their associations.[377]

Second, even if the causal associations exist, the potential increase of these adverse events is unlikely to exceed the reduction of cardiovascular events by statin treatment in primary prevention patients. In the systematic review in this thesis, the biggest potential adverse effect of statins was found on liver dysfunction, which increased the risk by 33% but resulted in an absolute increase of only 8 cases per 10,000 person-years. For self-perceived muscle symptoms, the most commonly reported adverse event, statins were found to increase the risk by 6%, which resulted in 15 extra cases per 10,000 person-years. For the more concerning clinically-confirmed muscle disorders, including myopathy and rhabdomyolysis, the annual incidence was very low (4 per 10,000 person-years) in the statin treatment group, without a significant difference from the incidence in the control group (3 per 10,000 person-years). As a comparison, statins reduced major cardiovascular events by 17%-28%, which led to an absolute reduction of these events up to 19 cases per 10,000 person-years. As discussed above, the increase in adverse events is not necessarily caused by statins given the uncertainty in their causal associations, while the reduction of cardiovascular events is more certain. Although it has been argued that the incidence of adverse events is generally lower in clinical trials as the participants are often selective and healthier, this is similar for cardiovascular events when considering the absolute treatment effects. The prediction modelling study in this thesis also showed a very low incidence of serious muscle disorders (4 per 10,000 person-years) based on a large population from primary care.

Finally, most of the adverse events are clinically and economically less concerning than the cardiovascular events that statins could prevent, in terms of their prognosis and disease burden.[12] As seen in the trials in this thesis, most of the reported adverse events were mild conditions that are usually tolerable and curable, such as self-perceived muscle symptoms (mostly unspecified muscle pain) without significant elevation of blood

creatine kinase, liver dysfunction with a one-time small elevation of liver enzymes, or renal insufficiency with trivial proteinuria. These conditions are unlikely to require hospitalisation and some of them may not need treatments. In contrast, cardiovascular events often end up in hospitalisation and are the leading cause of global mortality, which could significantly reduce the life quality of an individual and increase the disease burden for society.[1] Although patients' perspectives about these outcomes could vary by individual when considering statin treatment, the need to prevent CVD in the population should not be defeated by the concern about potential adverse events in policy-making.

All the presented evidence, showing the uncertain causal association with statins, the small potential risk increase, and the general mildness of the adverse events, may alleviate some concerns about statin safety in the public and help patients and doctors make evidence-based appropriate clinical decisions on statin treatment for CVD primary prevention. Regarding the consideration for tailoring statin regimens by drug type or dose to address safety concerns in these patients, the current evidence is unable to provide specific guidance on such practices, as explored in this thesis.

6.2.2 Towards personalised and stratified statin treatment

The StatinMD model developed in this thesis could be used to assist personalised decision-making on statin treatment for individuals in clinical practice and to implement a potential stratified statin treatment strategy in the population for prevention of CVD, as illustrated in the decision-making flowchart in **Figure 6.1**.

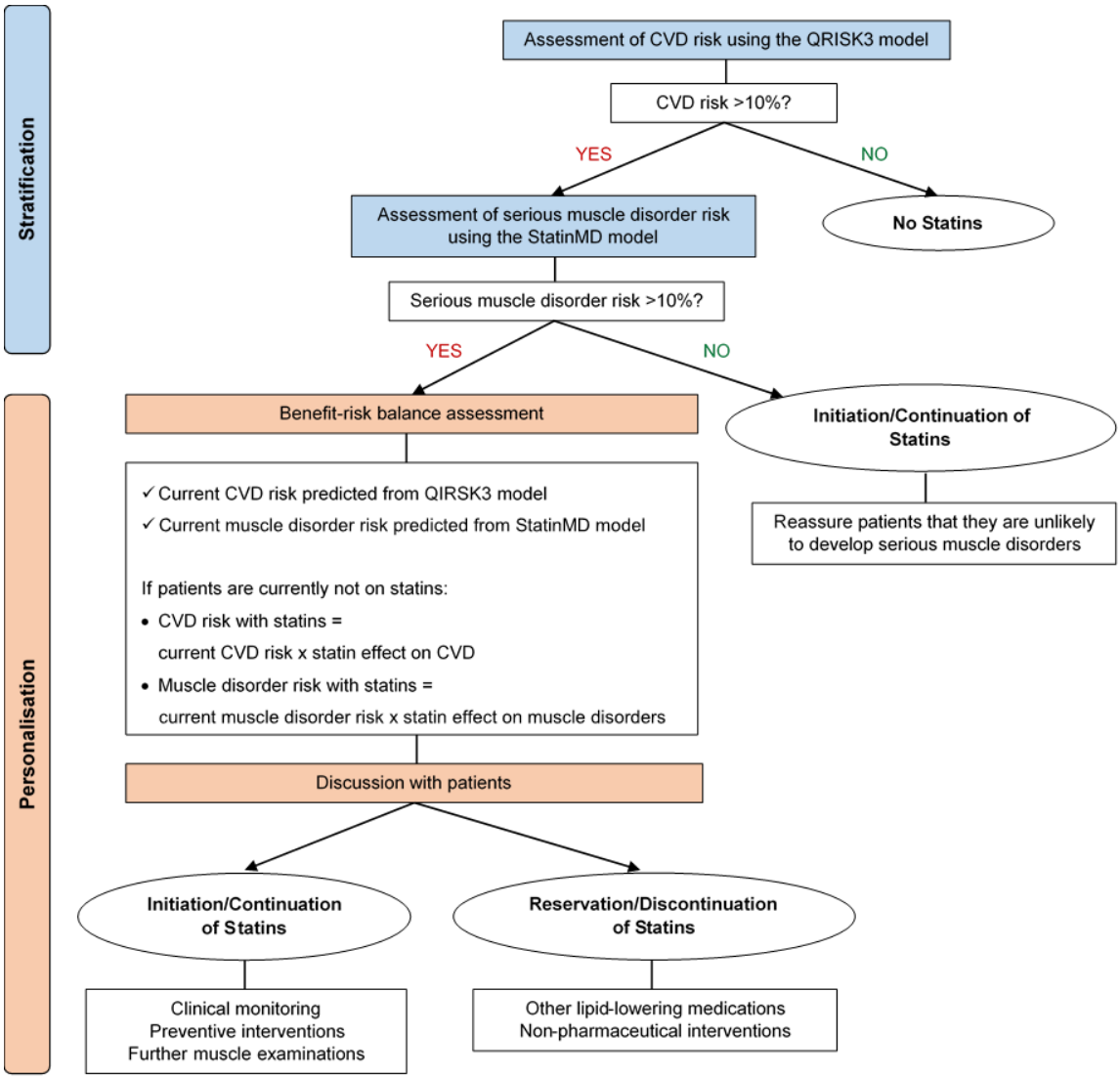


Figure 6.1 Potential stratified and personalised statin treatment decision-making process with the StatinMD model

According to the latest clinical guideline in the UK, statin treatment is recommended for individuals with a 10-year CVD risk above 10% predicted by the QRISK3 model.[39, 378] In this population, individuals could be further stratified by their risk of serious muscle disorders to take into account the concerns about safety in the treatment decision-making. Considering the prediction performance of the StatinMD model and the potential clinical utility of the risk stratification, a possible threshold to stratify the 10-year risk of serious muscle disorders predicted from the StatinMD model may be around 10%, as discussed in Chapter 5.

For individuals with a predicted risk of serious muscle disorders below 10%, the benefit from risk reduction of CVD with statins generally outweighs potential harm from risk increase of serious muscle disorders and statins should be prescribed according to the clinical guideline.[5, 6] These individuals represent the vast majority of the population eligible for statin treatment. The predicted risk from the StatinMD model could be used to reassure these individuals that they are unlikely to develop serious muscle disorders, in order to encourage uptake of statins or improve adherence to current statin treatment.

For individuals with a predicted risk of serious muscle disorders above 10%, the decision on statin treatment should be made based on a further personalised assessment of the benefit-harm balance of treatment and through a discussion between patients and doctors. These individuals account for around 2% of the population eligible for statin treatment. The further personalised assessment involves the calculation of the absolute treatment effects and the comparison between the risk of CVD and the risk of serious muscle disorders with statins for an individual, as described in Chapter 4.

Based on the personalised assessment, if an individual would benefit from a considerable reduction of CVD risk by statin treatment while the risk of developing serious muscle disorders with statins is acceptable, a decision may be made to initiate or continue statin treatment. In this circumstance, to address concerns about serious muscle disorders during the treatment, some clinical actions could be considered, such as clinical monitoring of relevant biochemical profiles (e.g. blood creatine kinase, urinary myoglobin), closer examinations of muscle functions (e.g. aerobic test, muscle biopsy), preventive interventions on modifiable risk factors (e.g. adjusting unnecessary concomitant medications that are potentially myotoxic or interact with statins, lowering the body mass index).[70] These are particularly relevant in current statin users who have

experienced muscle problems and could be based on existing guidance for managing statin intolerance in clinical practice.[20, 67]

If an individual has a significant increase in the risk of serious muscle disorders with statins that outweighs the reduction of CVD risk or the post-treatment risk of serious muscle disorders reaches a very high level, it may be inappropriate to initiate or continue statin treatment. In this circumstance, non-pharmaceutical interventions (e.g. exercise, diet) may be prioritised and the use of alternative lipid-lowering medications may be justified (e.g. ezetimibe, PCSK9 inhibitors) for prevention of CVD.[6, 7]

With risk information from the full assessment, the eventual decisions on statin treatment and corresponding clinical actions should be made through discussion between patients and doctors, taking into account patients' preferences and doctors' clinical opinions. The discussion may determine at which level the risk of serious muscle disorders is too high to initiate or continue statin treatment, which may vary by individuals' perspectives. In addition, since the latest (draft) clinical guideline in the UK has indicated a lower threshold of CVD risk below 10% for considering statin treatment, the threshold to stratify the risk of serious muscle disorders in this decision flowchart could also be lowered, as the benefit-harm balance of treatment in individuals at low CVD risk is more likely to be reversed by the risk of adverse events and requires careful assessment.

6.3 Limitations and challenges in this thesis

6.3.1 Limitations of the trial evidence and the analyses

Although the systematic review in this thesis indicated an overall benefit-harm balance favourable for the use of statins for CVD primary prevention, the explicit effects of statins on adverse events were not fully understood based on the current findings. As discussed

in Chapter 3, this is mainly due to the limitations regarding the external validity of the trial evidence and some deficiencies in the analyses of these trial data.

6.3.1.1 Limitations regarding the external validity of the trial evidence

Given the selective participants in the included trials, the incidence of adverse events in the systematic review may have been underestimated. The participant selection criteria in some of the trials excluded individuals who were more likely to develop adverse events, such as those with previous elevations in serum creatinine or liver enzymes.[79, 218, 227, 243] Some trials were designed with a run-in period, during which individuals with poor compliance to treatment, often due to adverse events, were excluded.[218] These could result in lower incidences of adverse events compared to the incidences observed in clinical practice, particularly in the statin treatment group. Moreover, since all of the trials were designed primarily for examining treatment efficacy rather than safety, the adverse events were often not pre-specified and their definitions in some trials were unclear. As such, the data on these adverse events may not be collected systematically and could be biased by incomplete or selective reporting.[379] Overall, the effects of statins on adverse events, particularly the absolute risk increase by statins, may be underestimated.

The participants included in this review may not be fully representative of the primary prevention patients in clinical practice. Some of the trials included a small number of patients with established CVD. The average age of the participants in many of the trials was below 60 years, which may be younger than most patients taking statins in clinical practice. As discussed above, those excluded individuals with certain comorbidities could be in fact eligible for statin treatment. Therefore, the estimated risks of adverse events with statins may not be explicit for the population of primary prevention, which may

consequently impact the comparison between the benefits and harms of statin treatment in this population.

6.3.1.2 Deficiencies in the analyses of the trial data

Given the low incidences of adverse events and the limited sample sizes of the included trials, most of the meta-analyses in this review lacked sufficient statistical power to examine the small difference in the incidence between the statin treatment and the control groups. This resulted in low precision of the estimates of some adverse effects and left uncertainty about their associations with statins, especially for muscle disorders that were the rarest outcome. The low precision of estimation was even more prominent in the network meta-analyses and the dose-response meta-analyses, due to the scarce data on specific drug types and doses.

Continuity correction was applied to many studies that had zero outcome events in at least one group. [274] This may have introduced bias when the sample sizes of the treatment and the control groups were unequal, leading to overestimation of the adverse effects (biasing the estimates away from 1 when the true OR >1) and underestimation of the beneficial effects (biasing the estimates towards 1 when the true OR <1).

In the absence of time-to-event data, the incidences of the outcomes in this review were calculated from the observed event rates and the average study duration across all involved trials. This is probably inaccurate and may also impact the comparison of the absolute treatment effects on the efficacy and safety outcomes.

These limitations and deficiencies, especially regarding the representativeness of the study participants and the statistical power of analysis, could be potentially addressed through observational studies based on large real-world populations, as further discussed in section 6.4.1 below.

6.3.2 Challenges in the development of the prediction model

As discussed in Chapters 4 and 5, the development of the StatinMD model encountered some challenges and difficulties in both the model derivation and validation, due to which some compromises had to be made in the study design and the statistical analyses and potential bias or uncertainty may have been introduced during the process. Some key issues are emphasised in this chapter. These should be taken into account when considering the use of the StatinMD model in clinical practice.

6.3.2.1 Compromises in the study design

- **Eligibility of study participants**

The target population for the use of the StatinMD model is the individuals eligible for statin treatment for prevention of CVD. According to the clinical guidelines, this eligibility should be determined by an individual's predicted CVD risk.[5, 6] However, patients' predicted CVD risk is poorly recorded in the primary care data used for developing the model in this thesis. As such, the eligibility of the study participants in both the model derivation and validation cohorts was determined by a cut-off age as a proxy of the predicted CVD risk based on the QRISK2 model.[59] This may have missed some eligible patients younger than the defined cut-off age and diminished the representativeness of the study participants for the target population.

- **Definition of outcome and predictors**

The desired outcome for prediction was serious muscle disorders that resulted in hospitalisation or death, which in this study was defined by a hospital admission or death record with a diagnosis of serious muscle disorders. However, the hospitalisation or death may not be necessarily caused by the diagnosed muscle disorders and it is difficult to

confirm the direct cause based on the healthcare records.[311] Consequently, the predicted risk from the StatinMD model may be higher than the risk of serious muscle disorders that could actually cause hospitalisation or death.

As predictors for the outcome, the use of the medications was determined by a recent prescription within a 12-month baseline exposure period before the start of follow-up. This design treated the individuals with previous use of the medications before the baseline exposure period as the same as those who never used the medications, whose risks of serious muscle disorders may be different, particularly if the previous users were on the medications for a long time and just recently stopped.[312] However, this design was considered better than a design without the baseline exposure period, where the use of the medications may be determined by a prescription years ago that has little predictive ability for the outcome.

6.3.2.2 Methodology issues in the model derivation

- **Categorisation of variables**

As one of the predictors, body mass index (BMI) was included in the model as a categorical variable instead of a continuous variable, in order to use the data on categorised BMI levels recorded by the medical codes to supplement the records of BMI values that contained considerable missing data. Due to the low incidence of the outcome, several categorical variables (ethnicity, alcohol consumption level, statin drug type) were re-categorised by combining some of the originally defined groups, in order to have a sufficient number of outcome events in each category to facilitate model convergence. Such categorisation and grouping may lead to a loss of information in the data and reduce the statistical power of model estimation.[313]

- **Model diagnosis and specification**

Due to the intensive computation for the competing risk model and the large sample size of the model derivation cohort, the model assumptions were checked based on a subsample of the cohort with complete-case data. Although the subsample was believed to be sufficient to detect any violations of the model assumptions, there may be potential differences between the complete-case subsample and the whole cohort and it would have been more reassuring to check the model assumptions with all the data.

Similarly, due to the model complexity and the intensive computation, there is currently no efficient algorithm in R to perform the selection of factional polynomial (FP) function for a competing risk model. In this thesis, the optimal FP function to transform age had to be selected using the available R package 'mfp' that implements the Cox model for a time-to-event outcome. Although the Cox model and the competing risk model share similar model assumptions and statistical fundamentals,[156] the selected FP function based on the Cox model may not still be optimal for the competing risk model.

- **Multiple imputation of missing data**

Another compromise to the intensive computation and the large sample size was that only 10 imputations were able to be performed to impute the missing data. Although this was believed to be a reasonable number, more imputations may lead to more robust imputed values, given the considerable proportion of missing data in this study.[314] For the practical purpose of imputing the four variables at once, the imputation models did not particularly include auxiliary variables for each imputed variable. Although the constructed imputation models performed well, the missingness of the imputed variables may not be fully explained by the explanatory variables in the imputation models.[315]

6.3.2.3 Pitfalls in the model validation

- **Sample similarity and non-independent validation**

The key purpose of external validation is to evaluate the generalisability of a model, in order to support its wider use beyond the population from where the model has been derived. However, the population characteristics in the external validation cohort for the StatinMD model turned out similar to the model derivation cohort, which may be more useful for demonstrating the reproducibility, rather than the generalisability, of the model.[145] Although both the model derivation and validation cohorts were representative of the primary care population in the UK, the generalisability of the StatinMD model may be only limited to this setting but not to others, such as in secondary care or other countries.

Independent model validation by investigators who are not involved in the model derivation is encouraged to reduce potentially inflated findings and 'spin'.[143, 144] However, in this thesis, the StatinMD model was derived and validated by the same author, which may lead to better validation results than by independent validation.[356]

- **Measurement of discrimination**

This study used the *c*-index as a measurement of model discrimination, which has been widely used in the validation of prognostic models.[356] The *c*-index reflects the average discrimination of the model across all possible risk thresholds in the theoretical range from 0% to 100%.[339] However, in this study, the predicted risk for most individuals in the validation cohort was within a small range of very low risk (around 0% - 5%), for which the model discrimination may be different from the average discrimination across the full theoretical range.[366] Partial *c*-index has been proposed to measure discrimination in a probability range of interest, but it is less interpretable than the full *c*-index and also has limitations.[366, 367]

- **Calibration at the individual level**

The model calibration at the individual level used the pseudo observations that were generated based on a restrictive assumption about the complete independence of the censoring to the outcome.[347] This assumption is difficult to confirm and some approaches to addressing potential dependent censoring have been proposed, such as computing pseudo observations separately in the subsets stratified by categorical covariates that are suspected to be related to the censoring mechanism.[347] However, these approaches would add much complexity to the calculation of pseudo values, especially when competing risk exists.[369] In fact, without using these approaches, the generation of the pseudo observations in this study has already encountered computation issues due to the huge sample size of the validation cohort. The whole sample was therefore divided into fifty subsamples to compute pseudo values, which may also introduce potential bias.

The calibration curves fitted to individuals' predicted risks and pseudo observations showed a large overestimation by the model in individuals with a high predicted risk. However, only a very small number of individuals had such a high predicted risk and the fitted curves may not be accurate due to the insufficient data. Overestimation in these individuals probably did exist, but the degree of overestimation remained uncertain.

- **No re-calibration attempted**

The StatinMD model showed some miscalibration in the external validation, which may be due to the sparsely distributed data of outcome that is likely to lead to the range of predicted risk across the population being too wide.[363] Some approaches could be used to adjust such miscalibration by shrinking the predictor coefficients to narrow down the range of predicted risk.[364] This study did not make such an adjustment because the miscalibration by the model would only affect a very small proportion of the population

and the calibration slope did not indicate a need for re-calibration either. Nevertheless, although the potential improvement may be limited, an attempt to re-calibrate the model may reduce the miscalibration to certain degree.

6.3.2.4 Restrictions by the use of routine healthcare records

The development of the StatinMD model in this thesis utilised the data from electronic healthcare records (EHR) that were collected in routine clinical practices and administration activities. This benefits from several features of EHR data, such as a large sample size, representative study population, long-term follow-up, and rich information.[277] Despite these advantages, the use of EHR data for clinical prediction modelling has presented some restrictions and challenges, due to the fact that these data are not collected for specific research purposes and the observational and retrospective nature of the records,[311, 380] including those discussed below.

- **Missing data and the imputation**

Four predictor variables in the StatinMD model had a considerable proportion of missing data in both the model derivation and validation cohorts. These missing data were imputed using multiple imputation, which was performed with only 10 imputations and did not include auxiliary variables in the imputation model due to the computation capacity, as discussed above. A more fundamental pitfall is that the multiple imputation was based on an untestable assumption of missing at random that may not be true in some circumstances.[162] For example, ethnicity may be more likely to be recorded if an individual is from a minority group, which may result in missing not at random in the data. Although the missing-at-random assumption has been widely accepted in epidemiological studies using healthcare records, it should be acknowledged that the

fulfilment of this assumption is not confirmed and there may be potential bias introduced to the model development based on the imputed data.[316]

- **Misclassification of the presence of the outcome and predictors**

Another restriction in using EHR data is that the presence of a condition can only be defined by a record of the condition and it is impossible to distinguish between a missing record and the actual absence of the condition.[162] For the StatinMD model, this may have caused misclassification of the outcome occurrence and the baseline characteristics regarding comorbidities and medications.[317]

Given the potential missed-out outcome events, the overall outcome incidence in the study cohorts may be lower than the actual incidence in the population and the individuals' risk predicted from the model may be underestimated. Furthermore, this definition approach may be subjected to recording bias.[381] Because muscle disorders may be more likely to be recorded in statin users since clinicians are aware of this potential adverse event of statins, or in younger patients as it is less common in this population. This may manually modify the predictive ability of these baseline characteristics, leading to biased coefficients of the predictor variables in the model.[381]

Similarly, the presence of the comorbidities at baseline may be also under-recorded in the study cohorts, especially those non-serious and common conditions in the population, such as degenerative joint disorders and vitamin D deficiency.[311] The potential incompleteness of the medical code lists used to identify these conditions may result in further under-capture of the comorbidities. In contrast, the recording of drug prescriptions is less likely to be missed and the codes for the medications are more definite. However, misclassification bias may still exist the other way around, because a prescription of a

medication does not necessarily mean the actual use of the medication but it is difficult to confirm using the EHR data.

6.4 Considerations for future research

6.4.1 Observational study for causal inference

Given the limitations of the evidence from clinical trials for causal inference, observational studies could be considered to provide supplementary evidence for a better understanding of the associations between statins and adverse events.[279] In particular, with large-scale longitudinal EHR data, such as the data from the Clinical Practice Research Datalink (CPRD) used in this thesis, observational studies could be designed with sufficient sample size and length of follow-up to examine serious but rare or late-onset adverse events of statins, including muscle disorders, liver injuries, renal failure, and surgery-required cataract, which have scarce data or are barely reported in clinical trials.[277] Compared to the highly selective participants in clinical trials, participants in observational studies could include a wide range of individuals, including some vulnerable groups with risk factors for adverse events, who are more representative of the target treatment population in clinical practice.[382] The adverse events identified by clinical diagnosis in observational studies may be better defined than most of the reported adverse events without consistent definition or pre-specification in clinical trials.[278]

In addition, the participants in observational studies are usually more diverse, which enables further understanding of the associations between statins and adverse events in different patient groups.[382] This may be useful for considering stratified statin treatment and preventing potential harm in certain patient groups who are more vulnerable to adverse events. Given the very little data for the comparison between statin drug types

and dose levels in clinical trials, observational studies could provide the opportunity to further explore the variations of the potential adverse effects by statin regimens, using data from a large number of statin prescriptions of different drugs and doses in clinical use.[277, 279]

A considerable number of observational studies examining the unintended outcomes of statin treatment have been previously conducted.[25] However, these studies did not lead to convincing conclusions on most adverse outcomes due to the conflicting results and the very high heterogeneity among them, as evaluated in a recent umbrella review involving 117 meta-analyses of observational studies.[179] The heterogeneity was most evident in terms of the outcome definitions and the populations in these studies. Most studies included a wide range of conditions with inconsistent definitions, particularly for muscle-related outcomes, and very few studies examined serious adverse events. Many studies focused on certain patient groups with specific comorbidities that were not representative of the general population eligible for statin treatment.[179] Few studies focused on the primary prevention population or compared the different statin regimens, which is of great clinical interest as pointed out in this thesis. Therefore, there remains an apparent evidence gap that could be filled by further observational studies.

Despite the advantages mentioned above, observational studies also have limitations for making causal inferences. The major drawback is the potential bias and confounding due to the absence of randomisation that could lead to differences between the participants in the treatment and the control groups.[279] The imbalance between the two groups may distort the association between the treatment and the outcome or modify the magnitude of the treatment effect observed in the study.[383] Appropriate design and analytical methods may help reduce systematic bias and control confounding, such as matched cohort design,[384] covariate adjustment,[148] propensity score methods,[385] and instrumental

variable approach.[386] However, these methods may not be able to completely eliminate bias and confounding. The fundamental concern is about the unrecognised or unmeasured confounding that is difficult to be addressed by a post hoc adjustment.[279] Furthermore, observational studies using EHR data may also suffer from the restrictions of such data sources as discussed above.

Previous reviews have compared the results from observational studies and RCTs that examined the same treatments and outcomes, showing that the agreement between the two types of studies varied by the particular treatment and outcome examined.[278, 279, 383] Regardless of the study type, it is more important to consider the specific designs and analysis methods used in an individual study, which determine the quality of the study and the reliability of its results.[383] Since observational studies and RCTs both have inherent limitations but provide valuable information, it is encouraged that evidence from both study types should be taken into account to accelerate the understanding of treatment effects.[387]

6.4.2 Development of a clinical decision tool

The purpose of the StatinMD prediction model in this thesis is to assist clinical decision-making on statin treatment and eventually improve the prevention of CVD. However, risk prediction itself does not automatically lead to appropriate decisions. Risk prediction is a fundamental step that provides the information required for decision-making, but how the information could be used to achieve appropriate decisions needs to be further studied, in order to develop a practical and effective tool supporting clinical decision-making on statin treatment.[138] This may involve considerations for risk communication and the theories of decision-making behaviour.[321, 388]

For a risk prediction model, a common approach to communicate the predicted risk is through a risk calculator that could allow personal data input and present the calculated risk to the users.[328] A possible web-based risk calculator for the StatinMD model is shown in Chapter 4. However, clinicians' and patients' preferences need to be further taken into account to design a user-friendly risk calculator.[327, 328] The manual efforts for data input should be minimised to avoid increasing too much workload for clinicians when making clinical decisions.[389] Ideally the calculator could be added to the clinical management system and take relevant information from the existing medical records. Careful consideration about the presentation of the predicted risk is also important, as it may influence how clinicians and patients perceive the risk.[390] It has also been suggested that a risk expressed in a natural frequency, instead of a probability, is easier for clinicians and patients to understand, particularly a very low risk like serious muscle disorders.[321] The use of pictographs could be also helpful to illustrate the predicted risk.[391]

As such, further work is warranted to understand in what way clinicians prefer to use the StatinMD model and to collect feedback on trial versions of the risk calculator with different designs of data input interface, perhaps by a survey delivered to some general practitioners (GP) involved in the CPRD network. Focus group interviews could be conducted with clinicians and patients to present the predicted risk from the StatinMD model in different ways and understand how they perceive the risk to help design the output panel of the risk calculator to facilitate appropriate perception of the risk.[392, 393] Consultation with the development teams of the EHR systems could also be initiated to discuss the feasibility of implementing the StatinMD risk calculator in the systems.

Although an effective risk communication tool could help establish an appropriate understanding of the risks of treatment, there are other critical factors contributing to patients' decisions on the treatment, including personal beliefs and opinions.[394] For

example, in the case of statin treatment, patients may have different views about the importance of serious muscle disorders compared to CVD and therefore attach unequal weights to the same predicted risk of serious muscle disorders in the trade-off between benefits and harms. They may also have different priorities in the treatment decision-making, such as to achieve the goal of treatment, to minimise the risk of adverse effects, or to have the most cost-effective treatment. These different preferences are usually formed in a complex context that involves personal experience, education level, socio-economic status, and the social and cultural environment.[395] A survey or interviews could be conducted among patients eligible for statin treatment, including those who have decided to take or not to take statins, to understand the personal considerations in the decision-making on stain treatment. This may help develop a decision aid that incorporates multiple risk prediction tools for different outcomes of treatment, such as the StatinMD model and the QRISK2 model, and takes into account personal preferences to guide users towards an appropriate treatment decision.[396]

As a complex intervention, such a decision aid needs to be further evaluated for its effectiveness at individual and system levels.[397, 398] This could be achieved by some extension designs to the standard RCT, such as cluster-randomised studies or N-of-1 trials.[399, 400] The outcomes to measure the effectiveness could be determined by working with relevant stakeholders and patient groups.[398] For example, the effectiveness of a decision aid for statin treatment may be measured by the increase of statin uptake in eligible patients, the reduction of treatment discontinuation in statin users, or the rate of restarting treatment in previous users. These approaches could also be used to evaluate the clinical consequences of using the StatinMD model or the risk calculator based on the model, also called the 'impact study' mentioned in Chapter 5.[376] Such evaluation studies usually require considerable time and resources and are rarely conducted for published

prediction models. However, this highlights the importance that the fundamental step – the development of the risk prediction model should be performed properly and therefore it is worth conducting further work to achieve the ultimate purpose of improving clinical decisions on the treatment.

6.5 Overall reflections on this thesis

6.5.1 The gap between research objectives and accomplishments

6.5.1.1 Unanswered questions in the systematic review

The systematic review and meta-analyses in this thesis aimed to assess the associations between statins and adverse events and explore the rationales for tailoring statin regimens to address the concerns about safety in primary prevention of CVD.[36, 51] However, most of the analyses turned out to be underpowered due to the insufficient data from clinical trials, leaving uncertainty in the estimates and making it difficult to draw conclusions about the associations, either found significant or not. Especially for the rare muscle disorders, which was probably the adverse event of the greatest interest,[20] the non-significant estimate with low precision did not eliminate the possibility of the association. Although this uncertainty itself and the observed low incidences of the adverse events may to some degree help alleviate the concerns about statin safety, which was one key reason to initiate this study, the meta-analysis results did not really answer the questions about the existence and the magnitude of the potential adverse effects of statins.

This was even more evident in the comparison between statin drug types and the examination of the dose-response relationships, owing to the more scarce data on specific drug types or doses. Even though it could be justified that the value of a systematic review is not only to draw an evidence-based conclusion but also to reveal the current evidence

gap for future research,[94, 96] performing the complex analyses with such sparse data, especially the dose-response meta-analysis, may be discouraged, if a preliminary inspection of existing data was conducted to understand the feasibility of using the data to achieve the objectives of these analyses.

6.5.1.2 Difficulties in the potential application of the StatinMD model

The main purpose of developing the StatinMD model was to provide a prediction tool to assist decision-making on statin treatment in clinical practice. When designing the modelling study, it was assumed that the model algorithm could be merged into the clinical system as a module, like the QRISK2 model for CVD risk implemented in the UK,[5, 59] which could automatically capture the data in patients' records and report the risk information through a few mouse clicks. With this expectation, the predictors in the StatinMD model were selected with consideration of the data availability and quality in primary care records in the UK and through an approach that maximised the comprehensiveness of the inclusion of predictors that were clinically important as judged by GPs. However, applying a prediction model to clinical practice is not straightforward, even though the model performs well. Designing a module in the clinical system to use the model algorithm in the UK needs collaboration with software developers and requires regulatory permission. This makes it difficult to implement the StatinMD model to assist decision-making in clinical practice and indeed very few published prediction models have been actually used in clinical practice.[138, 401]

An alternative application could be a web-based risk calculator using the model algorithm, as created and shown in this thesis, which requires manual data input. The StatinMD model included 22 predictors, which may need considerable effort for data input when using the web-based calculator. Especially the predictor frailty index requires

additional calculation using another model.[293] Therefore it is probably inefficient and impractical to use the risk calculator in clinical practice. If a web-based risk calculator was the expected application of the model, a different approach may be taken to select a smaller set of predictors and help make the application of the model easier. But even so, such risk calculators are still considered medical devices and require proper testing and regulatory permission for their clinical use.[402] This is also the reason that the algorithm of the StatinMD model is not fully presented in this thesis and instead accessible from the Oxford University Innovation platform, which protects the algorithm from being converted into an application tool used in the public without proper testing.

6.5.1.3 Limited clinical utility of potential stratified statin treatment strategy

Another purpose for developing the StatinMD model is to implement a potential stratified statin treatment strategy based on the stratification of the risk of serious muscle disorders in the population.[63] However, only a very small proportion of the population eligible for statin treatment would be classified as high-risk individuals. For example, with a potential risk threshold of 10% for the 10-year risk, only 2% of the population was classified as high-risk. In addition, the classification accuracy indicated that the model was likely to miss some cases of serious muscle disorders, which means that the number of individuals who would actually benefit from being classified as high-risk for stratified treatment may be even smaller. As estimated in this thesis, using the model for risk stratification may have a net benefit of identifying only 5 high-risk individuals for interventions in 10,000 persons, compared to simply ignoring the risk and doing nothing in the population.

As such, the clinical utility of the potential stratified statin treatment strategy would be very limited and provide little help to improve the use of statins in the whole population.

In the decision-making flowchart designed in this thesis, only the first two steps may be needed in most cases and the StatinMD model may be mainly used for supporting well-informed patient-centred decision-making for low-risk individuals, rather than implementing a stratified treatment strategy.[65]

6.5.2 The trade-off between methodology complexity and result improvement

A key strength in the development of the StatinMD model was that it took into account the competing risk to reduce the overestimation of the outcome risk.[151] This was believed to be crucial because the competing risk, which was the risk of death from causes other than muscle disorders, was high in the study population (10-year incidence of 33% in the model derivation cohort). More importantly, overestimating the risk of serious muscle disorders is particularly unwanted for making a decision on statin treatment, given the widespread concerns about statin safety and the common poor adherence to statin treatment.[34, 35] However, model predictions in the derivation and validation cohorts showed that the majority of the population had a very low risk of serious muscle disorders (the 10-year risk of 90% of the individuals was below 2%). Therefore the potential overestimation of the risk avoided by adjusting the competing risk may be very small and is unlikely to affect statin treatment decision-making in most people.

On the other hand, the Fine-Gray model used in this thesis to address the competing risk is more complex than the basic Cox model. Due to the model complexity and the huge sample size in this thesis, the model derivation and validation both required intensive, time-consuming computation. As such, some compromises had to be made when performing the statistical modelling in R, such as the model diagnosis and specification discussed in the preceding sections.

It remains unknown to what degree the competing risk model has reduced the overestimation, unless a Cox model is also derived and validated through the same process, which requires substantial work. Given the potentially small improvement and the actual computation difficulties, using a more complex model may not be worth the additional effort. If a simpler model had been used, some compromises in the process of the model development may not have been required; it might have been possible to include the continuous and multi-categorical variables in their original forms (to avoid loss of information from categorisation or grouping); and interaction terms, such the use of statins and other medications that interact with statins, may be also considered to improve the prediction ability of the model. All of these may have resulted in a model that performs as well as the competing risk model.

Similarly, examining the calibration of the StatinMD model at the individual level was more complicated than examining the calibration by risk groups, but the results did not actually change the conclusion about the overall validity of the model for clinical use. The most important information that the individual-level calibration added to the group-level calibration was the potential overestimation of very high predicted risks from the model. As previously discussed, the individuals with such high risk accounted for a very small proportion of the population and it did not change the fact that the model is validated for use in most people as seen in the calibration by risk groups.

Knowing the individual-level calibration across the full range of possible predicted risks may be more helpful when determining a threshold for risk stratification that may be located beyond the range of the group-level risks. However, as discussed above, it may be unnecessary to make a risk stratification for serious muscle disorders due to the limited clinical utility and the StatinMD model may be mainly used to better inform low-risk individuals who are well represented by the ten risk groups. For this purpose, examining

the calibration by risk groups may be sufficient to justify the clinical validity of the model. In fact, some widely used clinical prediction models, such as the QRISK2 model,[59] were only examined for their calibration by risk group.

Reflection on these suggests that theoretical superiority in methodology does not necessarily lead to pragmatic improvement in results. Although the use of all the complex analysis methods in this thesis was justified and did provide additional useful outputs, it may not be appropriate to claim that they are better than alternative methods. If simpler methods were chosen, the time and effort saved may be used to improve other aspects of this thesis. Perhaps it is possible to conduct additional studies, such as a cohort study of the associations between statins and serious adverse events, or the development of risk prediction models for other adverse events, which were proposed at the beginning of this project but not initiated due to all the time-consuming analyses included in this thesis. More efforts on patient and public involvement could also be made to add more clinical value to this thesis.

6.6 Conclusions

Based on current evidence from randomised controlled trials in patients without previous cardiovascular events, the increase of potential adverse events with statin treatment was small and did not outweigh the reduction of major cardiovascular events by statins. This should alleviate some concerns about statin safety in the public and reassure that the balance between potential benefits and harms of statins is favourable for treatment, supporting the clinical recommendations on the use of statins for primary prevention of CVD. Differences in the potential adverse effects between statin drug types and dose-response relationships in these adverse effects were inconclusive, suggesting that tailoring statin regimens to address safety concerns is currently not evidence-based.

The StatinMD model was designed to predict the risk of serious muscle disorders in 1, 5, and 10 years in the population potentially eligible for statin treatment, based on personal characteristics that are routinely recorded in clinical practice. The model was derived and externally validated in millions of patients in primary care in the UK. The model showed a good ability to distinguish between individuals who will and will not develop serious muscle disorders and was able to predict the personalised risk accurately for the majority of the population. The model could be used to identify high-risk individuals for proper clinical actions to address the concern about serious muscle disorders during statin treatment and also help avoid unnecessary interventions in low-risk individuals. The personalised risk information provided by the model could help better consider the trade-off between potential benefits and harms of statin treatment and assist the discussion between patients and doctors in clinical practice. This could lead to well-informed, patient-centred, and shared clinical decision-making on statin treatment and may eventually help improve the current low uptake of statins and poor adherence to statin treatment in the eligible population.

Bibliography

1. World Health Organisation (WHO). Cardiovascular Diseases Fact Sheet 2021 [Accessed 01 Nov 2022]. Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

2. Nelson RH. Hyperlipidemia as a Risk Factor for Cardiovascular Disease. Primary Care. 2013;40(1):195-+.

3. Rifkind BM. The Lipid Research Clinics Coronary Primary Prevention Trial Results .2. The Relationship of Reduction in Incidence of Coronary Heart-Disease to Cholesterol Lowering. Jama-J Am Med Assoc. 1984;251(3):365-74.

4. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J. 2017;38(32):2459-72.

5. National Institute for Health and Care Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification. NICE; 2014.

6. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2019;73(24):e285-e350.

7. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). Eur Heart J. 2020;41(1):111-88.

8. Endo A. The Discovery and Development of Hmg-Coa Reductase Inhibitors. J Lipid Res. 1992;33(11):1569-82.

9. Endo A. A historical perspective on the discovery of statins. P Jpn Acad B-Phys. 2010;86(5):484-93.

10. Law MR, Wald NJ, Rudnicka AR. Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: systematic review and meta-analysis. Brit Med J. 2003;326(7404):1423-7.

11. 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. 2005;366(9493):1267-78.

12. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet. 2016;388(10059):2532-61.

13. Blais JE, Wei Y, Yap KKW, Alwafi H, Ma TT, Brauer R, et al. Trends in lipid-modifying agent use in 83 countries. Atherosclerosis. 2021;328:44-51.

14. O'Keeffe AG, Nazareth I, Petersen I. Time trends in the prescription of statins for the primary prevention of cardiovascular disease in the United Kingdom: a cohort study using The Health Improvement Network primary care data. Clin Epidemiol. 2016;8:123-32.

15. Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol. 2006;97(8A):52C-60C.

16. Armitage J. The safety of statins in clinical practice. Lancet. 2007;370(9601):1781-90.

17. Staffa JA, Chang J, Green L. Cerivastatin and reports of fatal rhabdomyolysis. New Engl J Med. 2002;346(7):539-40.

18. Furberg CD, Pitt B. Withdrawal of cerivastatin from the world market. Curr Contr Trials C. 2001;2(5):205-7.

19. Grundy SM. The Issue of Statin Safety. Circulation. 2005;111(23):3016-9.

20. Rosenson RS, Baker SK, Jacobson TA, Kopecky SL, Parker BA. An assessment by the Statin Muscle Safety Task Force: 2014 update. J Clin Lipidol. 2014;8(3):S58-S71.

21. Bhardwaj SS, Chalasani N. Lipid-lowering agents that cause drug-induced hepatotoxicity. Clin Liver Dis. 2007;11(3):597-613, vii.

22. Bays H, Cohen DE, Chalasani N, Harrison SA. An assessment by the Statin Liver Safety Task Force: 2014 update. J Clin Lipidol. 2014;8(3):S47-S57.

23. Marcum ZA, Griend JPV, Linnebur SA. FDA Drug Safety Communications: A Narrative Review and Clinical Considerations for Older Adults. Am J Geriatr Pharmac. 2012;10(4):264-71.

24. Maki KC, Ridker PM, Brown WV, Grundy SM, Sattar N. An assessment by the Statin Diabetes Safety Task Force: 2014 update. J Clin Lipidol. 2014;8(3):S17-S29.

25. Macedo AF, Taylor FC, Casas JP, Adler A, Prieto-Merino D, Ebrahim S. Unintended effects of statins from observational studies in the general population: systematic review and meta-analysis. Bmc Med. 2014;12.

26. Newman CB, Preiss D, Tobert JA, Jacobson TA, Page RL, Goldstein LB, et al. Statin Safety and Associated Adverse Events: A Scientific Statement From the American Heart Association. Arterioscl Throm Vas. 2019;39(2):E38-E81.

27. Kocas C, Abaci O, Kocas BB, Cetinkal G, Arslan S, Yildiz A, et al. The role of media on statin adherence. Int J Cardiol. 2015;201:139-.

28. Schaffer AL, Buckley NA, Dobbins TA, Banks E, Pearson SA. The crux of the matter: did the ABC's Catalyst program change statin use in Australia? Med J Australia. 2015;202(11):591-5.

29. Saib A, Sabbah L, Perdrix L, Blanchard D, Danchin N, Puymirat E. Evaluation of the impact of the recent controversy over statins in France: The EVANS study. Arch Cardiovasc Dis. 2013;106(10):511-6.

30. Nielsen SF, Nordestgaard BG. Negative statin-related news stories decrease statin persistence and increase myocardial infarction and cardiovascular mortality: a nationwide prospective cohort study. Eur Heart J. 2016;37(11):908-16.

31. Lassenius MI, Toppila I, Bergius S, Perttila J, Airaksinen KJ, Pietila M. Cardiovascular event rates increase after each recurrence and associate with poor statin adherence. European Journal of Preventive Cardiology. 2020:2047487320904334.

32. Abramson JD, Rosenberg HG, Jewell N, Wright JM. Should people at low risk of cardiovascular disease take a statin? Bmj-Brit Med J. 2013;347.

33. Malhotra A. Saturated fat is not the major issue. Bmj. 2013;347:f6340.

34. Matthews A, Herrett E, Gasparrini A, Van Staa T, Goldacre B, Smeeth L, et al. Impact of statin related media coverage on use of statins: interrupted time series analysis with UK primary care data. Bmj-Brit Med J. 2016;353.

35. Marcus ME, Manne-Goehler J, Theilmann M, Farzadfar F, Moghaddam SS, Keykhaei M, et al. Use of statins for the prevention of cardiovascular disease in 41 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data. Lancet Glob Health. 2022;10(3):E369-E79.

36. Redberg RF, Katz MH. Statins for Primary Prevention: The Debate Is Intense, but the Data Are Weak. JAMA. 2016;316(19):1979-81.

37. Byrne P, Cullinan J, Smith SM. Statins for primary prevention of cardiovascular disease. BMJ. 2019;367:l5674.

38. Stone NJ, Robinson JG, Lichtenstein AH, Merz CNB, Blum CB, Eckel RH, et al. 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Journal of the American College of Cardiology. 2014;63(25):2889-934.

39. National Institute for Health and Care Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification (draft). NICE; 2023.

40. Yebyo HG, Zappacosta S, Aschmann HE, Haile SR, Puhan MA. Global variation of risk thresholds for initiating statins for primary prevention of cardiovascular disease: a benefit-harm balance modelling study. BMC Cardiovascular Disorders. 2020;20(1):418.

41. Newman DH, Redberg RF. Will new statin guidelines lead to overtreatment? Clin Lipidol. 2014;9(2):125-8.

42. Demasi M. Statin wars: have we been misled about the evidence? A narrative review. Brit J Sport Med. 2018;52(14):905-9.

43. 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. 2012;380(9841):581-90.

44. Taylor F, Huffman MD, Macedo AF, Moore THM, Burke M, Smith GD, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Db Syst Rev. 2013(1).

45. Therapeutics Initiative. Do statins have a role in primary prevention? An update. . Therapeutics Letter. 2010(77):1-2.

46. Ebrahim S, Casas JP. Statins for all by the age of 50 years? Lancet. 2012;380(9841):545-7.

47. National Institute for Health and Care Excellence (NICE). NICE recommends wider use of statins for prevention of CVD 2014 [Accessed 01 Nov 2022]. Available from: https://www.nice.org.uk/news/article/nice-recommends-wider-use-of-statins-for-prevention-of-cvd#:~:text=In%20updated%20guidance%20on%20lipid,a%2010%20per%20cent%20risk.

48. Schachter M. Chemical, pharmacokinetic and pharmacodynamic properties of statins: an update. Fund Clin Pharmacol. 2005;19(1):117-25.

49. Dimmitt SB, Stampfer HG, Warren JB. The pharmacodynamic and clinical trial evidence for statin dose. Br J Clin Pharmacol. 2018;84(6):1128-35.

50. Chong PH, Seeger JD, Franklin C. Clinically relevant differences between the statins: Implications for therapeutic selection. Am J Med. 2001;111(5):390-400.

51. Borges RD, Degobi NAH, Bertoluci MC. Choosing statins: a review to guide clinical practice. Arch Endocrin Metab. 2020;64(6):639-53.

52. Alonso R, Cuevas A, Cafferata A. Diagnosis and Management of Statin Intolerance. J Atheroscler Thromb. 2019;26(3):207-15.

53. Bellosta S, Paoletti R, Corsini A. Safety of statins - Focus on clinical pharmacokinetics and drug interactions. Circulation. 2004;109(23):50-7.

54. Warren JB. Translating the dose response into risk and benefit. Br J Clin Pharmacol. 2019;85(10):2187-93.

55. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients - The need for risk stratification. Jama-J Am Med Assoc. 2007;298(10):1209-12.

56. Fraenkel L, Fried TR. Individualized medical decision making: necessary, achievable, but not yet attainable. Arch Intern Med. 2010;170(6):566-9.

57. Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. Int J Epidemiol. 2016;45(6):2075-88.

58. de Vries TI, Stam-Slob MC, Peters RJG, van der Graaf Y, Westerink J, Visseren FLJ. Impact of a Patient's Baseline Risk on the Relative Benefit and Harm of a Preventive Treatment Strategy: Applying Trial Results in Clinical Decision Making. J Am Heart Assoc. 2022;11(1).

59. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. Bmj-Brit Med J. 2008;336(7659):1475-+.

60. Hageman S, Pennells L, Ojeda F, Kaptoge S, Kuulasmaa K, de Vries T, et al. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J. 2021;42(25):2439-54.

61. Eichler K, Puhan MA, Steurer J, Bachmann LM. Prediction of first coronary events with the Framingham score: A systematic review. Am Heart J. 2007;153(5):722-U14.

62. Collins GS, Altman DG. Predicting the adverse risk of statin treatment: an independent and external validation of Qstatin risk scores in the UK. Heart. 2012;98(14):1091-7.

63. Bell J. Stratified medicines: towards better treatment for disease. Lancet. 2014;383:3-5.

64. Sowden G, Hill JC, Konstantinou K, Khanna M, Main CJ, Salmon P, et al. Targeted treatment in primary care for low back pain: the treatment system and clinical training programmes used in the IMPaCT Back study (ISRCTN 55174281). Fam Pract. 2012;29(1):50-62.

65. Hingorani AD, van der Windt DA, Riley RD, Abrams K, Moons KGM, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: Stratified medicine research. Bmj-Brit Med J. 2013;346.

66. Wei MY, Ito MK, Cohen JD, Brinton EA, Jacobson TA. Predictors of statin adherence, switching, and discontinuation in the USAGE survey: understanding the use of statins in America and gaps in patient education. J Clin Lipidol. 2013;7(5):472-83.

67. 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. 2015;36(17):1012-22b.

68. Harper CR, Jacobson TA. The broad spectrum of statin myopathy: from myalgia to rhabdomyolysis. Curr Opin Lipidol. 2007;18(4):401-8.

69. Hansen KE, Hildebrand JP, Ferguson EE, Stein JH. Outcomes in 45 patients with statin-associated myopathy. Arch Intern Med. 2005;165(22):2671-6.

70. Abd TT, Jacobson TA. Statin-induced myopathy: a review and update. Expert Opin Drug Saf. 2011;10(3):373-87.

71. Mansi I, Frei CR, Pugh MJ, Makris U, Mortensen EM. Statins and Musculoskeletal Conditions, Arthropathies, and Injuries. Jama Intern Med. 2013;173(14):1318-26.

72. Tobert JA, Newman CB. The nocebo effect in the context of statin intolerance. J Clin Lipidol. 2016;10(4):739-47.

73. David S, Hamilton JP. Drug-induced Liver Injury. US Gastroenterol Hepatol Rev. 2010;6:73-80.

74. Karahalil B, Hare E, Koc G, Uslu I, Senturk K, Ozkan Y. Hepatotoxicity associated with statins. Arh Hig Rada Toksiko. 2017;68(4):254-60.

75. Cohen DE, Anania FA, Chalasani N. An assessment of statin safety by hepatologists. Am J Cardiol. 2006;97(8a):77c-81c.

76. Bader T. Liver tests are irrelevant when prescribing statins. The Lancet. 2010;376(9756):1882-3.

77. Reuben A, Koch DG, Lee WM. Drug-induced acute liver failure: results of a U.S. multicenter, prospective study. Hepatology. 2010;52(6):2065-76.

78. Bjornsson E, Jacobsen EI, Kalaitzakis E. Hepatotoxicity associated with statins: Reports of idiosyncratic liver injury post-marketing. J Hepatol. 2012;56(2):374-80.

79. Ridker PM, Danielson E, Fonseca FAH, Genest J, Gotto Jr AM, Kastelein JJP, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. New Engl J Med. 2008;359(21):2195-207.

80. 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. 2009;32(10):1924-9.

81. Sattar N, Preiss D, Murray HM, Welsh P, Buckley BM, de Craen AJ, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. Lancet. 2010;375(9716):735-42.

82. Wolfe SM. Dangers of rosuvastatin identified before and after FDA approval. The Lancet. 2004;363(9427):2189-90.

83. Davidson MH. Rosuvastatin safety: lessons from the FDA review and post-approval surveillance. Expert Opin Drug Saf. 2004;3(6):547-57.

84. Zipes DP, Zvaifler NJ, Glassock RJ, Gilman S, Muñoz A, Gogolak V, et al. Rosuvastatin: an independent analysis of risks and benefits. MedGenMed. 2006;8(2):73.

85. Honore PM, Jacobs R, Hendrickx I, De Waele E, Van Gorp V, De Regt J, et al. Statins and the Kidney: Friend or Foe? Blood Purification. 2017;43(1-3):91-6.

86. Verdoodt A, Honore PM, Jacobs R, De Waele E, Van Gorp V, De Regt J, et al. Do Statins Induce or Protect from Acute Kidney Injury and Chronic Kidney Disease: An Update Review in 2018. J Transl Int Med. 2018;6(1):21-5.

87. Hartman HA, Myers LA, Evans M, Robison RL, Engstrom RG, Tse FL. The safety evaluation of fluvastatin, an HMG-CoA reductase inhibitor, in beagle dogs and rhesus monkeys. Fundam Appl Toxicol. 1996;29(1):48-62.

88. Laties AM, Shear CL, Lippa EA, Gould AL, Taylor HR, Hurley DP, et al. Expanded clinical evaluation of lovastatin (EXCEL) study results. II. Assessment of the human lens after 48 weeks of treatment with lovastatin. Am J Cardiol. 1991;67(6):447-53.

89. Harris ML, Bron AJ, Brown NA, Keech AC, Wallendszus KR, Armitage JM, et al. Absence of effect of simvastatin on the progression of lens opacities in a randomised placebo controlled study. Oxford Cholesterol Study Group. Br J Ophthalmol. 1995;79(11):996-1002.

90. Yu S, Chu Y, Li G, Ren L, Zhang Q, Wu L. Statin Use and the Risk of Cataracts: A Systematic Review and Meta-Analysis. J Am Heart Assoc. 2017;6(3).

91. Yusuf S, Bosch J, Dagenais G, Zhu J, Xavier D, Liu L, et al. Cholesterol lowering in intermediate-risk persons without cardiovascular disease. New Engl J Med. 2016;374(21):2021-31.

92. Ma L, Wang Y, Du J, Wang M, Zhang R, Fu Y. The association between statin use and risk of age-related macular degeneration. Scientific Reports. 2015;5(1):18280.

93. Roizenblatt M, Naranjit N, Maia M, Gehlbach PL. The Question of a Role for Statins in Age-Related Macular Degeneration. Int J Mol Sci. 2018;19(11).

94. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: Synthesis of best evidence for clinical decisions. Ann Intern Med. 1997;126(5):376-80.

95. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. J Clin Nurs. 2003;12(1):77-84.

96. Kassirer JP. Clinical trials and meta-analysis. What do they do for us? N Engl J Med. 1992;327(4):273-4.

97. Mills EJ, Wu P, Chong G, Ghement I, Singh S, Akl EA, et al. Efficacy and safety of statin treatment for cardiovascular disease: a network meta-analysis of 170,255 patients from 76 randomized trials. Qjm. 2011;104(2):109-24.

98. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions (version 6.3). Cochrane 2022. Available from: www.training.cochrane.org/handbook.

99. Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. Lancet. 2010;376(9734):20-1.

100. Glass GV. Primary, Secondary, and Meta-Analysis of Research. Educational Researcher. 1976;5(10):3-8.

101. Haidich AB. Meta-analysis in medical research. Hippokratia. 2010;14(1):29-37.

102. Fagard RH, Staessen JA, Thijs L. Advantages and disadvantages of the meta-analysis approach. J Hypertens. 1996;14:S9-S12.

103. Sutton AJ, Higgins JP. Recent developments in meta-analysis. Stat Med. 2008;27(5):625-50.

104. Akobeng AK. Understanding systematic reviews and meta-analysis. Archives of Disease in Childhood. 2005;90(8):845-8.

105. Fletcher J. What is heterogeneity and is it important? Bmj. 2007;334(7584):94-6.

106. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003;327(7414):557-60.

107. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods. 2010;1(2):97-111.

108. Lumley T. Network meta-analysis for indirect treatment comparisons. Statistics in Medicine. 2002;21(16):2313-24.

109. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Statistics in Medicine. 2004;23(20):3105-24.

110. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. Bmj-Brit Med J. 2005;331(7521):897-900.

111. Li TJ, Puhan MA, Vedula SS, Singh S, Dickersin K, Methods AHNM-A. Network meta-analysis-highly attractive but more methodological research is needed. Bmc Med. 2011;9.

112. Dias S, Caldwell DM. Network meta-analysis explained. BMJ Publishing Group; 2019.

113. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods. 2012;3(2):98-110.

114. Lu GB, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. J Am Stat Assoc. 2006;101(474):447-59.

115. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Statistics in Medicine. 2010;29(7-8):932-44.

116. Ruberg SJ. Dose response studies I. Some design considerations. Journal of Biopharmaceutical Statistics. 1995;5(1):1-14.

117. Orsini N, Larsson SC, Salanti G. Dose–Response Meta-Analysis. Systematic Reviews in Health Research: Meta-Analysis in Context. 2022:258-69.

118. Upreti VV, Venkatakrishnan K. Model-Based Meta-Analysis: Optimizing Research, Development, and Utilization of Therapeutics Using the Totality of Evidence. Clin Pharmacol Ther. 2019;106(5):981-92.

119. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol. 1992;135(11):1301-9.

120. Macdougall J. Analysis of Dose–Response Studies—Emax Model. In: Ting N, editor. Dose Finding in Drug Development. New York, NY: Springer New York; 2006. p. 127-45.

121. Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. American journal of epidemiology. 2012;175(1):66-73.

122. Langford O, Aronson JK, van Valkenhoef G, Stevens RJ. Methods for meta-analysis of pharmacodynamic dose-response data with application to multi-arm studies of alogliptin. Stat Methods Med Res. 2018;27(2):564-78.

123. Mawdsley D, Bennetts M, Dias S, Boucher M, Welton NJ. Model-Based Network Meta-Analysis: A Framework for Evidence Synthesis of Clinical Trial Data. Cpt-Pharmacomet Syst. 2016;5(8):393-401.

124. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? Bmj-Brit Med J. 2009;338.

125. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. Bmj-Brit Med J. 2013;346.

126. Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. Plos Med. 2013;10(2).

127. van Diepen M, Ramspek CL, Jager KJ, Zoccali C, Dekker FW. Prediction versus aetiology: common pitfalls and how to avoid them. Nephrology Dialysis Transplantation. 2017;32(suppl_2):ii1-ii5.

128. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. Bmj-Brit Med J. 2009;338.

129. Haybittle J, Blamey R, Elston C, Johnson J, Doyle P, Campbell F, et al. A prognostic index in primary breast cancer. British journal of cancer. 1982;45(3):361-6.

130. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. Plos Med. 2013;10(2).

131. Le-Rademacher J, Wang XF. Time-To-Event Data: An Overview and Analysis Considerations. J Thorac Oncol. 2021;16(7):1067-74.

132. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine. 1996;15(4):361-87.

133. Hendriksen JMT, Geersing GJ, Moons KGM, de Groot JAH. Diagnostic and prognostic prediction models. J Thromb Haemost. 2013;11:129-41.

134. Sun G-W, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. Journal of clinical epidemiology. 1996;49(8):907-16.

135. Steyerberg EW, Eijkemans MJ, Harrell Jr FE, Habbema JDF. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Statistics in medicine. 2000;19(8):1059-79.

136. Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, et al. State of the art in selection of variables and functional forms in multivariable analysis-outstanding issues. Diagn Progn Res. 2020;4:3.

137. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453-73.

138. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. Brit Med J. 2009;338.

139. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925-+.

140. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. Bmj-Brit Med J. 2009;338.

141. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology. 2001;54(8):774-81.

142. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130(6):515-24.

143. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med. 2015;162(1):55-U103.

144. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013;13:33.

145. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol. 2015;68(3):279-89.

146. Cox DR. Regression Models and Life-Tables. J R Stat Soc B. 1972;34(2):187-+.

147. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: Basic concepts and first analyses. British Journal of Cancer. 2003;89(2):232-8.

148. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part II: Multivariate data analysis - an introduction to concepts and methods. British Journal of Cancer. 2003;89(3):431-6.

149. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annu Rev Public Health. 1997;18:83-104.

150. Satagopan J, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach A. A note on competing risks in survival data analysis. British journal of cancer. 2004;91(7):1229-35.

151. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol. 2012;41(3):861-70.

152. Allignol A, Schumacher M, Wanner C, Drechsler C, Beyersmann J. Understanding competing risks: a simulation point of view. BMC Med Res Methodol. 2011;11:86.

153. Noordzij M, Leffondre K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? Nephrology Dialysis Transplantation. 2013;28(11):2670-7.

154. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94(446):496-509.

155. Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic Models With Competing Risks Methods and Application to Coronary Risk Prediction. Epidemiology. 2009;20(4):555-61.

156. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. Statistics in Medicine. 2007;26(11):2389-430.

157. Grambauer N, Schumacher M, Beyersmann J. Proportional subdistribution hazards modeling offers a summary analysis, even if misspecified. Statistics in Medicine. 2010;29(7-8):875-84.

158. Sapir-Pichhadze R, Pintilie M, Tinckam KJ, Laupacis A, Logan AG, Beyene J, et al. Survival Analysis in the Presence of Competing Risks: The Example of Waitlisted Kidney Transplant Candidates. Am J Transplant. 2016;16(7):1958-66.

159. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016;133(6):601-9.

160. Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. Stat Methods Med Res. 2002;11(2):203-15.

161. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. American Journal of Epidemiology. 1995;142(12):1255-64.

162. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I. Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidem Dr S. 2010;19(6):618-26.

163. Petersen I, Welch CA, Nazareth I, Walters K, Marston L, Morris RW, et al. Health indicator recording in UK primary care electronic health records: key implications for handling missing data. Clin Epidemiol. 2019;11:157-67.

164. Groenwold RH, White IR, Donders ART, Carpenter JR, Altman DG, Moons KG. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. Cmaj. 2012;184(11):1265-9.

165. Donders ART, Van Der Heijden GJ, Stijnen T, Moons KG. A gentle introduction to imputation of missing values. Journal of clinical epidemiology. 2006;59(10):1087-91.

166. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. Yale J Biol Med. 2013;86(3):343-58.

167. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

168. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, Kristensen NR, Pham TM, Pedersen L, et al. Missing data and multiple imputation in clinical epidemiological research. Clin Epidemiol. 2017;9:157-65.

169. Austin PC, White IR, Lee DGS, van Buuren S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. Can J Cardiol. 2021;37(9):1322-31.

170. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. Bmc Medical Research Methodology. 2009;9.

171. White IR, Royston P. Imputing missing covariate values for the Cox model. Statistics in Medicine. 2009;28(15):1982-98.

172. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. Bmc Medical Research Methodology. 2017;17.

173. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011;45(3):1-67.

174. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine. 2011;30(4):377-99.

175. Kropko J, Goodrich B, Gelman A, Hill J. Multiple imputation for continuous and categorical data: comparing joint multivariate normal and conditional approaches. Political Analysis. 2014;22(4).

176. Seaman SR, Hughes RA. Relative efficiency of joint-model and full-conditional-specification multiple imputation when conditional models are compatible: The general location model. Stat Methods Med Res. 2018;27(6):1603-14.

177. Bittencourt MS, Cesena FHY. Statin dose in primary prevention: aim for the target! Heart. 2019;105(13):969-71.

178. Force UPST. Statin Use for the Primary Prevention of Cardiovascular Disease in Adults: US Preventive Services Task Force Recommendation Statement. JAMA. 2022;328(8):746-53.

179. He Y, Li X, Gasevic D, Brunt E, McLachlan F, Millenson M, et al. Statins and Multiple Noncardiovascular Outcomes: Umbrella Review of Meta-analyses of Observational Studies and Randomized Controlled Trials. Ann Intern Med. 2018;169(8):543-53.

180. Naci H, Brugts JJ, Fleurence R, Tsoi B, Toor H, Ades A. Comparative benefits of statins in the primary and secondary prevention of major coronary events and all-cause mortality: a network meta-analysis of placebo-controlled and active-comparator trials. European Journal of Preventive Cardiology. 2013;20(4):641-57.

181. Chou R, Dana T, Blazina I, Daeges M, Jeanne TL. Statins for Prevention of Cardiovascular Disease in Adults: Evidence Report and Systematic Review for the US Preventive Services Task Force. Jama. 2016;316(19):2008-24.

182. Yebyo HG, Aschmann HE, Kaufmann M, Puhan MA. Comparative effectiveness and safety of statins as a class and of specific statins for primary prevention of cardiovascular disease: A systematic review, meta-analysis, and network meta-analysis of randomized trials with 94,283 participants. Am Heart J. 2019;210:18-28.

183. Zhou Z, Albarqouni L, Curtis AJ, Breslin M, Nelson M. The Safety and Tolerability of Statin Therapy in Primary Prevention in Older Adults: A Systematic Review and Meta-analysis. Drugs Aging. 2020;37(3):175-85.

184. Naci H, Brugts J, Ades T. Comparative Tolerability and Harms of Individual Statins: A Study-Level Network Meta-Analysis of 246 955 Participants From 135 Randomized, Controlled Trials. Circ-Cardiovasc Qual. 2013;6(4):390-9.

185. Preiss D, Seshasai SRK, Welsh P, Murphy SA, Ho JE, Waters DD, et al. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. Jama. 2011;305(24):2556-64.

186. Villani R, Navarese EP, Cavallone F, Kubica J, Bellanti F, Facciorusso A, et al. Risk of Statin-Induced Hypertransaminasemia: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Mayo Clin Proc Innov Qual Outcomes. 2019;3(2):131-40.

187. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Plos Med. 2009;6(7):e1000097.

188. Alberton M, Wu P, Druyts E, Briel M, Mills EJ. Adverse events associated with individual statin treatments for cardiovascular disease: an indirect comparison meta-analysis. Qjm. 2012;105(2):145-57.

189. Armitage J, Baigent C, Barnes E, Betteridge DJ, Blackwell L, Blazing M, et al. Efficacy and safety of statin therapy in older people: a meta-analysis of individual participant data from 28 randomised controlled trials. The Lancet. 2019;393(10170):407-15.

190. Esmeijer K, Dekkers OM, de Fijter JW, Dekker FW, Hoogeveen EK. Effect of different types of statins on kidney function decline and proteinuria: a network meta-analysis. Sci Rep. 2019;9(1):16632.

191. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

192. Schünemann H, Brozek J, Guyatt G, Oxman A, editors. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach (Updated October 2013). The GRADE Working Group; 2013.

193. Puhan MA, Schunemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ. 2014;349:g5630.

194. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evidence Based Mental Health. 2019;22(4):153-60.

195. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. Statistics in Medicine. 2007;26(1):53-77.

196. Haviland MG. Yates's correction for continuity and the analysis of 2× 2 contingency tables. Statistics in medicine. 1990;9(4):363-7.

197. Petitti DB. Approaches to heterogeneity in meta-analysis. Stat Med. 2001;20(23):3625-33.

198. Rice K, Higgins JPT, Lumley T. A Re-Evaluation of Fixed Effect(s) Meta-Analysis. Journal of the Royal Statistical Society Series A: Statistics in Society. 2017;181(1):205-27.

199. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine. 2002;21(11):1539-58.

200. Rothman KJ. Epidemiology - An Introduction. Oxford: Oxford University Press; 2002.

201. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med. 2006;25(20):3443-57.

202. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010;1(2):112-25.

203. Rucker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods. 2012;3(4):312-24.

204. Krahn U, Binder H, König J. A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol. 2013;13:35.

205. Kirby S, Brain P, Jones B. Fitting E(max) models to clinical trial dose-response data. Pharm Stat. 2011;10(2):143-9.

206. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. Journal of computational and graphical statistics. 1998;7(4):434-55.

207. Bradford RH, Shear CL, Chremos AN, Dujovne C, Downton M, Franklin FA, et al. Expanded Clinical Evaluation of Lovastatin (EXCEL) Study results. I. Efficacy in modifying plasma lipoproteins and adverse event profile in 8245 patients with moderate hypercholesterolemia. Arch Intern Med. 1991;151(1):43-9.

208. Behounek BD, McGovern ME, Kassler-Taub KB, Markowitz JS, Bergman M, Bedard J, et al. A multinational study of the effects of low-dose pravastatin in patients with non-insulin-dependent diabetes mellitus and hypercholesterolemia. Clinical Cardiology. 1994;17(10):558-62.

209. Furberg CD, Adams Jr HP, Applegate WB, Byington RP, Espeland MA, Hartwell T, et al. Effect of lovastatin on early carotid atherosclerosis and cardiovascular events. Circulation. 1994;90(4 I):1679-87.

210. Jacobson TA, Chin MM, Curry CL, Miller V, Papademetriou V, Schlant RC, et al. Efficacy and safety of pravastatin in African Americans with primary hypercholesterolemia. Arch Intern Med. 1995;155(17):1900-6.

211. Salonen R, Nyyssonen K, Porkkala E, Rummukainen J, Belder R, Park JS, et al. Kuopio Atherosclerosis Prevention Study (KAPS): A population-based primary preventive trial of the effect of LDL lowering on atherosclerotic progression in carotid and femoral arteries. Circulation. 1995;92(7):1758-64.

212. Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. New Engl J Med. 1995;333(20):1301-7.

213. Mercuri M, Bond MG, Sirtori CR, Veglia F, Crepaldi G, Feruglio FS, et al. Pravastatin reduces carotid intima-media thickness progression in an asymptomatic hypercholesterolemic Mediterranean population: The Carotid Atherosclerosis Italian Ultrasound Study. Am J Med. 1996;101(6):627-34.

214. Bertolini S, Bon GB, Campbell LM, Farnier M, Langan J, Mahla G, et al. Efficacy and safety of atorvastatin compared to pravastatin in patients with hypercholesterolemia. Atherosclerosis. 1997;130(1-2):191-7.

215. Bak AAA, Huizer J, Leijten PA, Rila H, Grobbee DE. Diet and pravastatin in moderate hypercholesterolaemia: A randomized trial in 215 middle-aged men free from cardiovascular disease. Journal of Internal Medicine. 1998;244(5):371-8.

216. Downs JR, Clearfield M, Weis S, Whitney E, Shapiro DR, Beere PA, et al. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: Results of AFCAPS/TexCAPS. Journal of the American Medical Association. 1998;279(20):1615-22.

217. Gentile S, Turco S, Guarino G, Sasso CF, Amodio M, Magliano P, et al. Comparative efficacy study of atorvastatin vs. simvastatin, pravastatin, lovastatin and placebo in type 2 diabetic patients with hypercholesterolaemia. Diabetes, Obesity and Metabolism. 2000;2(6):355-62.

218. Furberg CD, Wright Jr JT, Davis BR, Cutler JA, Alderman M, Black H, et al. Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to pravastatin vs usual care: The antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT-LLT). Journal of the American Medical Association. 2002;288(23):2998-3007.

219. Holdaas H, Fellstrom B, Jardine AG, Holme I, Nyberg G, Fauchald P, et al. Effect of fluvastatin on cardiac outcomes in renal transplant recipients: A multicentre, randomised, placebo-controlled trial. Lancet. 2003;361(9374):2024-31.

220. Jones PH, Davidson MH, Stein EA, Bays HE, McKenney JM, Miller E, et al. Comparison of the efficacy and safety of rosuvastatin versus atorvastatin, simvastatin, and pravastatin across doses (STELLAR* trial). American Journal of Cardiology. 2003;92(2):152-60.

221. Kerzner B, Corbelli J, Sharp S, Lipka LJ, Melani L, LeBeaut A, et al. Efficacy and safety of ezetimibe coadministered with lovastatin in primary hypercholesterolemia. American Journal of Cardiology. 2003;91(4):418-24.

222. Melani L, Mills R, Hassman D, Lipetz R, Lipka L, LeBeaut A, et al. Efficacy and safety of ezetimibe coadministered with pravastatin in patients with primary hypercholesterolemia: A prospective, randomized, double-blind trial. Eur Heart J. 2003;24(8):717-28.

223. Mohler IER, Hiatt WR, Creager MA. Cholesterol reduction with atorvastatin improves walking distance in patients with peripheral arterial disease. Circulation. 2003;108(12):1481-6.

224. Sever PS, Dahlof B, Poulter NR, Wedel H, Beevers G, Caulfield M, et al. Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial - Lipid Lowering Arm (ASCOT-LLA): A multicentre randomised controlled trial. Lancet. 2003;361(9364):1149-58.

225. Asselbergs FW, Diercks GFH, Hillege HL, Van Boven AJ, Janssen WMT, Voors AA, et al. Effects of fosinopril and pravastatin on cardiovascular events in subjects with microalbuminuria. Circulation. 2004;110(18):2809-16.

226. Bays HE, Ose L, Fraser N, Tribble DL, Quinto K, Reyes R, et al. A multicenter, randomized, double-blind, placebo-controlled, factorial design study to evaluate the lipid-altering efficacy and safety profile of the ezetimibe/simvastatin tablet compared with ezetimibe and simvastatin monotherapy in patients with primary hypercholesterolemia. Clinical Therapeutics. 2004;26(11):1758-73.

227. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HAW, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): Multicentre randomised placebo-controlled trial. Lancet. 2004;364(9435):685-96.

228. 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. 2004;117(11):823-9.

229. Strandberg TE, Feely J, Sigurdsson EL. Twelve-week, multicenter, randomized, open-label comparison of the effects of rosuvastatin 10 mg/d and atorvastatin 10 mg/d in high-risk adults: A DISCOVERY study. Clinical Therapeutics. 2004;26(11):1821-33.

230. Zanchetti A, Crepaldi G, Bond MG, Gallus G, Veglia F, Mancia G, et al. Different effects of antihypertensive regimens based on fosinopril or hydrochlorothiazide with or without lipid lowering by pravastatin on progression of asymptomatic carotid atherosclerosis: Principal results of PHYLLIS - A randomized double-blind trial. Stroke. 2004;35(12):2807-12.

231. Anderssen SA, Hjelstuen AK, Hjermann I, Bjerkan K, Holme I. Fluvastatin and lifestyle modification for reduction of carotid intima-media thickness and left ventricular mass progression in drug-treated hypertensives. Atherosclerosis. 2005;178(2):387-97.

232. Berne C, Siewert-Delle A, Ekesbo R, Froberg L, Henningsen N, Ilestam G, et al. Comparison of rosuvastatin and atorvastatin for lipid lowering in patients with type 2 diabetes mellitus: Results from the URANUS study. Cardiovascular Diabetology. 2005;4 (no pagination).

233. Raggi P, Davidson M, Callister TQ, Welty FK, Bachmann GA, Hecht H, et al. Aggressive versus moderate lipid-lowering therapy in hypercholesterolemic postmenopausal women: Beyond endorsed lipid lowering with EBT scanning (BELLES). Circulation. 2005;112(4):563-71.

234. Stalenhoef AFH, Ballantyne CM, Sarti C, Murin J, Tonstad S, Rose H, et al. A COmparative study with rosuvastatin in subjects with METabolic Syndrome: Results of the COMETS study. Eur Heart J. 2005;26(24):2664-72.

235. Wolffenbuttel BHR, Franken AAM, Vincent HH. Cholesterol-lowering effects of rosuvastatin compared with atorvastatin in patients with type 2 diabetes - CORALL study. Journal of Internal Medicine. 2005;257(6):531-9.

236. Ferdinand KC, Clark LT, Watson KE, Neal RC, Brown CD, Kong BW, et al. Comparison of efficacy and safety of Rosuvastatin versus Atorvastatin in African-American patients in a six-week trial. American Journal of Cardiology. 2006;97(2):229-35.

237. Knopp RH, D'Emden M, Smilde JG, Pocock SJ. Efficacy and safety of atorvastatin in the prevention of cardiovascular end points in subjects with type 2 diabetes: The Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus (ASPEN). Diabetes Care. 2006;29(7):1478-85.

238. Milionis HJ, Rizos E, Kostapanos M, Filippatos TD, Gazi IF, Ganotakis ES, et al. Treating to target patients with primary hyperlipidaemia: Comparison of the effects of ATOrvastatin and ROSuvastatin (the ATOROS study). Current Medical Research and Opinion. 2006;22(6):1123-31.

239. Nakamura H, Arakawa K, Itakura H, Kitabatake A, Goto Y, Toyota T, et al. Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a prospective randomised controlled trial. Lancet. 2006;368(9542):1155-63.

240. Schmermund A, Achenbach S, Budde T, Buziashvili Y, Forster A, Friedrich G, et al. Effect of intensive versus standard lipid-lowering treatment with atorvastatin on the progression of calcified coronary atherosclerosis over 12 months: A multicenter, randomized, double-blind trial. Circulation. 2006;113(3):427-37.

241. Betteridge DJ, Gibson JM. Effects of rosuvastatin on lipids, lipoproteins and apolipoproteins in the dyslipidaemia of diabetes. Diabetic Medicine. 2007;24(5):541-9.

242. Bone HG, Kiel DP, Lindsay RS, Lewiecki EM, Bolognese MA, Leary ET, et al. Effects of atorvastatin on bone in postmenopausal women with dyslipidemia: A double-blind, placebo-controlled, dose-ranging trial. Journal of Clinical Endocrinology and Metabolism. 2007;92(12):4671-7.

243. Crouse IJR, Raichlen JS, Riley WA, Evans GW, Palmer MK, O'Leary DH, et al. Effect of rosuvastatin on progression of carotid intima-media thickness in low-risk individuals with subclinical atherosclerosis: The METEOR Trial. Journal of the American Medical Association. 2007;297(12):1344-53.

244. Lewis JH, Mortensen ME, Zweig S, Fusco MJ, Medoff JR, Belder R, et al. Efficacy and safety of high-dose pravastatin in hypercholesterolemic patients with well-compensated chronic liver disease: Results of a prospective, randomized, double-blind, placebo-controlled, multicenter trial. Hepatology. 2007;46(5):1453-63.

245. Mok VCT, Lam WWM, Chen XY, Wong A, Ng PW, Tsoi TH, et al. Statins for asymptomatic middle cerebral artery stenosis: The regression of cerebral artery stenosis study. Cerebrovascular Diseases. 2009;28(1):18-25.

246. Chan KL, Teo K, Dumesnil JG, Ni A, Tam J. Effect of lipid lowering with rosuvastatin on progression of aortic stenosis: Results of the aortic stenosis progression observation: Measuring effects of rosuvastatin (Astronomer) trial. Circulation. 2010;121(2):306-14.

247. Eriksson M, Budinski D, Hounslow N. Comparative efficacy of pitavastatin and simvastatin in high-risk patients: A randomized controlled trial. Advances in Therapy. 2011;28(9):811-23.

248. Saku K, Zhang B, Noda K. Randomized head-to-head comparison of pitavastatin, atorvastatin, and rosuvastatin for safety and efficacy (quantity and quality of LDL) - The PATROL trial. Circulation Journal. 2011;75(6):1493-505.

249. Ghia CJ, Panda SK, Khobragade LR, Jha RK. Impact of different starting doses of atorvastatin on reaching target low density lipoprotein cholesterol levels and health related quality of life in dyslipidemic patients. Indian Journal of Physiology and Pharmacology. 2013;57(3):293-307.

250. Parker BA, Capizzi JA, Grimaldi AS, Clarkson PM, Cole SM, Keadle J, et al. Effect of statins on skeletal muscle function. Circulation. 2013;127(1):96-103.

251. Stender S, Budinski D, Gosho M, Hounslow N. Pitavastatin shows greater lipid-lowering efficacy over 12 weeks than pravastatin in elderly patients with primary hypercholesterolaemia or combined (mixed) dyslipidaemia. Eur J Prev Cardiol. 2013;20(1):40-53.

252. Odawara M, Yamazaki T, Kishimoto J, Ito C, Noda M, Terauchi Y, et al. Pitavastatin reduced C-reactive protein without increasing the incidence of diabetes in Japanese individuals with impaired glucose tolerance (J-PREDICT study). Eur Heart J. 2014;35:392-.

253. Ogawa H, Matsui K, Saito Y, Sugiyama S, Jinnouchi H, Sugawara M, et al. Differences between Rosuvastatin and Atorvastatin in lipid-lowering action and effect on glucose metabolism in Japanese hypercholesterolemic patients with concurrent diabetes - Lipid-lowering with highly potent statins in hyperlipidemia with type 2 diabetes patients (LISTEN) study. Circulation journal. 2014;78(10):2512-5.

254. Sponseller CA, Morgan RE, Kryzhanovski VA, Campbell SE, Davidson MH. Comparison of the lipid-lowering effects of pitavastatin 4 mg versus pravastatin 40 mg in adults with primary hyperlipidemia or mixed (combined) dyslipidemia: a Phase IV, prospective, US, multicenter, randomized, double-blind, superiority trial. Clin Ther. 2014;36(8):1211-22.

255. Nakagomi A, Shibui T, Kohashi K, Kosugi M, Kusama Y, Atarashi H, et al. Differential Effects of Atorvastatin and Pitavastatin on Inflammation, Insulin Resistance, and the Carotid Intima-Media Thickness in Patients with Dyslipidemia. J Atheroscler Thromb. 2015;22(11):1158-71.

256. Patil CY, Baig MS, Doifode SM. Assessing the efficacy and safety of pitavastatin compared to atorvastatin in dyslipidemic patients: a double blind randomized controlled trial. International Journal of Basic & Clinical Pharmacology; Vol 5, No 3 (2016): May-June 2016DO - 1018203/2319-2003ijbcp20161530. 2016.

257. Chen X, Wang S, Lin L, Li Y, Zhang H. Drug effect of atorvastatin on middle cerebral atherosclerotic stenosis and high resolution NMR diagnosis. Pak J Pharm Sci. 2018;31(3(Special)):1169-73.

258. Itoh H, Komuro I, Takeuchi M, Akasaka T, Daida H, Egashira Y, et al. Intensive Treat-to-Target Statin Therapy in High-Risk Japanese Patients With Hypercholesterolemia and Diabetic Retinopathy: report of a Randomized Study. Diabetes care. 2018;41(6):1275-84.

259. Liu M, Liu G, Wang M. Comparison and analysis of the therapeutic effect of different statins in the treatment of atherosclerosis. Pak J Pharm Sci. 2018;31(4(Special)):1737-42.

260. Toribio M, Fitch KV, Stone L, Zanni MV, Lo J, de Filippi C, et al. Assessing statin effects on cardiovascular pathways in HIV using a novel proteomics approach: Analysis of data from INTREPID, a randomized controlled trial. EBioMedicine. 2018;35:58-66.

261. Bae JC, Min KW, Kim YH, Kim KA, Hong EG, Park CY, et al. Efficacy and safety of fixed-dose combination therapy with gemigliptin (50 mg) and rosuvastatin compared with monotherapy in patients with type 2 diabetes and dyslipidaemia (BALANCE): A multicentre, randomized, double-blind, controlled, phase 3 trial. Diabetes Obes Metab. 2019;21(1):103-11.

262. Kitas GD, Nightingale P, Armitage J, Sattar N, Belch JJF, Symmons DPM. A Multicenter, Randomized, Placebo-Controlled Trial of Atorvastatin for the Primary Prevention of Cardiovascular Events in Patients With Rheumatoid Arthritis. Arthritis Rheumatol. 2019;71(9):1437-49.

263. Peng M, Dong H, Jiang X, Che W, Zou Y, Zhang Y, et al. A randomized unblinded trial to compare effects of intensive versus conventional lipid-lowering therapy in patients undergoing renal artery stenting. J Cardiol. 2019;74(5):443-50.

264. Wang Y, Zhao B, Karlson BW, Wei D, Bots ML, Chen Y. Abstract 11154: Rosuvastatin Slows Progression of Carotid Intima-media Thickness in Low-risk Chinese Individuals With Subclinical Atherosclerosis (the Meteor-china Study). Circulation. 2019;140(Suppl_1):A11154-A.

265. Moroi M, Nagayama D, Hara F, Saiki A, Shimizu K, Takahashi M, et al. Outcome of pitavastatin versus atorvastatin therapy in patients with hypercholesterolemia at high risk for atherosclerotic cardiovascular disease. International Journal of Cardiology. 2020.

266. Thongtang N, Piyapromdee J, Tangkittikasem N, Samaithongcharoen K, Srikanchanawat N, Sriussadaporn S. Efficacy and Safety of Switching from Low-Dose Statin to High-Intensity Statin for Primary Prevention in Type 2 Diabetes: A Randomized Controlled Trial. Diabetes Metab Syndr Obes. 2020;13:423-31.

267. Reith C, Baigent C, Blackwell L, Emberson J, Spata E, Davies K, et al. Effect of statin therapy on muscle symptoms: an individual participant data meta-analysis of large-scale, randomised, double-blind trials. The Lancet. 2022;400(10355):832-45.

268. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376(9753):1670-81.

269. Smeeth L, Douglas I, Hall AJ, Hubbard R, Evans S. Effect of statins on a wide range of health outcomes: a cohort study validated by comparison with randomized trials. British Journal of Clinical Pharmacology. 2009;67(1):99-109.

270. Hippisley-Cox J, Coupland C. Unintended effects of statins in men and women in England and Wales: population based cohort study using the QResearch database. Bmj. 2010;340:c2197.

271. Yebyo HG, Aschmann HE, Puhan MA. Finding the Balance Between Benefits and Harms When Using Statins for Primary Prevention of Cardiovascular Disease. Ann Intern Med. 2018;170(1):1-10.

272. Ward NC, Watts GF, Eckel RH. Statin Toxicity. Circ Res. 2019;124(2):328-50.

273. von Hippel PT. The heterogeneity statistic I2 can be biased in small meta-analyses. BMC Medical Research Methodology. 2015;15(1):35.

274. J. Sweeting M, J. Sutton A, C. Lambert P. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Statistics in Medicine. 2004;23(9):1351-75.

275. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological). 1995;57(1):289-300.

276. Prinz H. Hill coefficients, dose-response curves and allosteric mechanisms. J Chem Biol. 2010;3(1):37-44.

277. Chute CG. Invited commentary: observational research in the age of the electronic health record. American journal of epidemiology. 2014;179(6):759-61.

278. Papanikolaou PN, Christidi GD, Ioannidis JPA. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. Canadian Medical Association Journal. 2006;174(5):635-41.

279. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. New Engl J Med. 2000;342(25):1878-86.

280. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clin Trials. 2005;2(3):209-17.

281. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. Bmj. 2010;340:c221.

282. Wei WQ, Feng Q, Jiang L, Waitara MS, Iwuchukwu OF, Roden DM, et al. Characterization of statin dose response in electronic medical records. Clin Pharmacol Ther. 2014;95(3):331-8.

283. Ford I, Norrie J. Pragmatic trials. New Engl J Med. 2016;375(5):454-63.

284. Thompson PD, Clarkson P, Karas RH. Statin-Associated Myopathy. JAMA. 2003;289(13):1681-90.

285. Hippisley-Cox J, Coupland C. Individualising the risks of statins in men and women in England and Wales: population-based cohort study. Heart. 2010;96(12):939-47.

286. Hopewell JC, Offer A, Haynes R, Bowman L, Li J, Chen F, et al. Independent risk factors for simvastatin-related myopathy and relevance to different types of muscle symptom. Eur Heart J. 2020;41(35):3336-42.

287. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.

288. Valiyil R, Christopher-Stine L. Drug-related myopathies of which the clinician should be aware. Current rheumatology reports. 2010;12:213-20.

289. Kellick KA, Bottorff M, Toth PP. A clinician's guide to statin drug-drug interactions. J Clin Lipidol. 2014;8(3):S30-S46.

290. Office for National Statistics (ONS). UK Census 2001 [Accessed 10 Jan 2021]. Available from: https://www.ons.gov.uk/census/2001censusandearlier/aboutcensus2001.

291. National Health Service (NHS). Alcohol misuse [Accessed 10 Jan 2021]. Available from: https://www.nhs.uk/conditions/alcohol-misuse/.

292. National Health Service (NHS). What is the body mass index (BMI)? [Accessed 10 Jan 2021]. Available from: https://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/.

293. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and ageing. 2016;45(3):353-60.

294. Li Q, Li X, Wang J, Liu H, Kwong JS-W, Chen H, et al. Diagnosis and treatment for hyperuricemia and gout: a systematic review of clinical practice guidelines and consensus statements. BMJ Open. 2019;9(8):e026677.

295. Association for Clinical Biochemistry (ACB), British Thyroid Association (BTA), British Thyroid Foundation (BTF). UK Guidelines for the Use of Thyroid Function Tests. BTA; 2006.

296. Scientific Advisory Committee on Nutrition (SACN). Vitamin D and health. SACN; 2016.

297. Devalia V, Hamilton MS, Molloy AM. Guidelines for the diagnosis and treatment of cobalamin and folate disorders. Br J Haematol. 2014;166(4):496-513.

298. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr., Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2019;38(7):1276-96.

299. Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Stat Methods Med Res. 2017;26(2):796-808.

300. Aalen O, Johansen S. An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on

Censored Observations. Scandinavian Journal of Statistics. 1978;5(3):141-50.

301. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res. 2010;1(4):274-8.

302. Kumar D, Klefsjo B. Proportional Hazards Model - a Review. Reliab Eng Syst Safe. 1994;44(2):177-88.

303. Schoenfeld D. Partial Residuals for the Proportional Hazards Regression-Model. Biometrika. 1982;69(1):239-41.

304. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part III: Multivariate data analysis - choosing a model and assessing its adequacy and fit. British Journal of Cancer. 2003;89(4):605-11.

305. Kawaguchi ES, Shen JI, Li G, Suchard MA. A Fast and Scalable Implementation Method for Competing Risks Data with the R Package fastcmprsk. R J. 2020;12(2):43-60.

306. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol. 1999;28(5):964-74.

307. Dupont WD. Review of Multivariable Model-building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modeling Continuous Variables, by Royston and Sauerbrei. Stata J. 2010;10(2):297-302.

308. Long JA, Bamba MI, Ling B, Shea JA. Missing race/ethnicity data in veterans health administration based disparities research: A systematic review. J Health Care Poor U. 2006;17(1):128-40.

309. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011;45(3):1 - 67.

310. 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 Drug Ther. 2005;19(6):403-14.

311. Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. Bmj. 2018;361.

312. Huddy K, Dhesi P, Thompson PD. Do the frequencies of adverse events increase, decrease, or stay the same with long-term use of statins? Current atherosclerosis reports. 2013;15:1-5.

313. Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. BMC medical research methodology. 2012;12:1-5.

314. Graham JW, Olchowski AE, Gilreath TD. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. Prevention Science. 2007;8(3):206-13.

315. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods. 2001;6(4):330-51.

316. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj-Brit Med J. 2009;339.

317. Beesley LJ, Mukherjee B. Statistical inference for association studies using electronic health records: handling both selection bias and outcome misclassification. Biometrics. 2022;78(1):214-26.

318. Office for National Statistics (ONS). UK Census 2021 [Accessed 10 Dec 2022]. Available from: https://census.gov.uk/census-2021-results.

319. NHS Digital. Quality and Outcomes Framework (QOF) 2022 [Accessed 10 Mar 2023]. Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/general-practice-data-hub/quality-outcomes-framework-qof#what-you-can-find-out.

320. Weiner SJ, Schwartz A, Sharma G, Binns-Calvey A, Ashley N, Kelly B, et al. Patient-centered decision making and health care outcomes: an observational study. Ann Intern Med. 2013;158(8):573-9.

321. Ahmed H, Naik G, Willoughby H, Edwards AGK. Communicating risk. BMJ : British Medical Journal. 2012;344:e3996.

322. Cardiology ESo. Governing policies and procedures for the writing of ESC clinical practice guidelines. European Society of Cardiology. 2017.

323. Garbi M. National Institute for Health and Care Excellence clinical guidelines development principles and processes. Heart. 2021;107(12):949-53.

324. The QResearch. QRISK-3 Risk Calculator 2018 [Accessed 10 Mar 2023]. Available from: https://qrisk.org/.

325. Winton Centre for Risk and Evidence Communication. The PREDICT for Breast Cancer 2010 [Accessed 10 Mar 2023]. Available from: https://breast.predict.nhs.uk/tool.

326. Harle CA, Downs JS, Padman R. Effectiveness of personalized and interactive health risk calculators: a randomized trial. Medical Decision Making. 2012;32(4):594-605.

327. Nolan T, Dack C, Pal K, Ross J, Stevenson FA, Peacock R, et al. Patient reactions to a web-based cardiovascular risk calculator in type 2 diabetes: a qualitative study in primary care. British Journal of General Practice. 2015;65(632):e152-e60.

328. Bonner C, Fajardo MA, Hui S, Stubbs R, Trevena L. Clinical validity, understandability, and actionability of online cardiovascular disease risk calculators: systematic review. Journal of medical Internet research. 2018;20(2):e29.

329. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Statistics in Medicine. 2004;23(5):723-48.

330. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis

Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med. 2015;162(1):W1-W73.

331. Pencina MJ, D'Agostino RB, Sr. Evaluating Discrimination of Risk Prediction Models: The C Statistic. JAMA. 2015;314(10):1063-4.

332. Vickers AJ, Cronin AM, Kattan MW, Gonen M, Scardino PT, Milowsky MI, et al. Clinical benefits of a multivariate prediction model for bladder cancer: a decision analytic approach. Cancer. 2009;115(23):5460-9.

333. Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. Medical Decision Making. 2006;26(6):565-74.

334. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. Brit Med J. 2012;344.

335. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol. 2019;48(6):1740-+.

336. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35(2):214-26.

337. Chen Y-C. A tutorial on kernel density estimation and recent advances. Biostatistics & Epidemiology. 2017;1(1):161-87.

338. Rahman MS, Ambler G, Choodari-Oskooei B, Omar RZ. Review and evaluation of performance measures for survival prediction models in external validation settings. Bmc Medical Research Methodology. 2017;17.

339. Hanley JA, Mcneil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (Roc) Curve. Radiology. 1982;143(1):29-36.

340. Heagerty PJ, Zheng YY. Survival model predictive accuracy and ROC curves. Biometrics. 2005;61(1):92-105.

341. Zhang ZH, Cortese G, Combescure C, Marshall R, Lee M, Lim HJ, et al. Overview of model validation for survival regression model with competing risks using melanoma study data. Ann Transl Med. 2018;6(16).

342. Bedogni G, Tsybakov A, Berlin S. Clinical prediction models—a practical approach to development, validation and updating. development. 2009;18(500):53-99.

343. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JDF. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Statistics in medicine. 2004;23(16):2567-86.

344. Stevens RJ, Poppe KK. Validation of clinical prediction models: what does the "calibration slope" really measure? J Clin Epidemiol. 2020;118:93-9.

345. Fan J, Gasser T, Gijbels I, Brockmann M, Engel J. Local polynomial regression: Optimal kernels and asymptotic minimax efficiency. Annals of the Institute of Statistical Mathematics. 1997;49:79-99.

346. Klein JP, Gerster M, Andersen PK, Tarima S, Perme MP. SAS and R functions to compute pseudo-values for censored data regression. Comput Methods Programs Biomed. 2008;89(3):289-300.

347. Andersen PK, Perme MP. Pseudo-observations in survival analysis. Stat Methods Med Res. 2010;19(1):71-99.

348. Efron B. The jackknife, the bootstrap and other resampling plans: SIAM; 1982.

349. Colin Cameron A, Windmeijer FAG. An R-squared measure of goodness of fit for some common nonlinear regression models. Journal of Econometrics. 1997;77(2):329-42.

350. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. Bmc Med Inform Decis. 2008;8.

351. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res. 2019;3:18.

352. Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpPIn" and "SnNOut": a note of caution. BMJ. 2004;329(7459):209-13.

353. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet. 2005;365(9469):1500-5.

354. McGee S. Simplifying likelihood ratios. J Gen Intern Med. 2002;17(8):646-9.

355. Royston P. Explained Variation for Survival Models. The Stata Journal. 2006;6(1):83-96.

356. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. Bmc Medical Research Methodology. 2014;14.

357. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J. 2003;24(11):987-1003.

358. Archer L, Koshiaris C, Lay-Flurrie S, Snell KIE, Riley RD, Stevens R, et al. Development and external validation of a risk prediction model for falls in patients with an indication for antihypertensive treatment: retrospective cohort study. Bmj-Brit Med J. 2022;379.

359. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. Stat Med. 2000;19(24):3401-15.

360. Koshiaris C, Archer L, Lay-Flurrie S, Snell KI, Riley RD, Stevens R, et al. Predicting the risk of acute kidney injury in primary care: derivation and validation of STRATIFY-AKI. Br J Gen Pract. 2023.

361. Perme MP, Andersen PK. Visualising survival data regression models using pseudo-observations. Iti. 2008:377-+.

362. Mortensen LM, Hansen CP, Overvad K, Lundbye-Christensen S, Parner ET. The Pseudo-Observation Analysis of Time-To-Event Data. Example from the Danish Diet, Cancer and Health Cohort Illustrating Assumptions, Model Validation and Interpretation of Results. Epidemiologic Methods. 2018;7(1).

363. Pavlou M, Ambler G, Seaman SR, Guttmann O, Elliott P, King M, et al. How to develop a more accurate risk prediction model when there are few events. BMJ : British Medical Journal. 2015;351:h3868.

364. Ambler G, Seaman S, Omar R. An evaluation of penalised survival methods for developing prognostic models with rare events. Statistics in medicine. 2012;31(11-12):1150-61.

365. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ. 2016;353:i3140.

366. Ma H, Bandos AI, Rockette HE, Gur D. On use of partial area under the ROC curve for evaluation of diagnostic performance. Stat Med. 2013;32(20):3449-58.

367. Walter SD. The partial area under the summary ROC curve. Stat Med. 2005;24(13):2025-40.

368. Parodi S, Verda D, Bagnasco F, Muselli M. The clinical meaning of the area under a receiver operating characteristic curve for the evaluation of the performance of disease markers. Epidemiol Health. 2022;44:e2022088.

369. Graw F, Gerds TA, Schumacher M. On pseudo-values for regression analysis in competing risks models. Lifetime Data Analysis. 2009;15(2):241-55.

370. Therneau T. A Package for Survival Analysis in R. R package version 3.5-5; 2023.

371. Tzoulaki I, Liberopoulos G, Ioannidis JPA. Assessment of Claims of Improved Prediction Beyond the Framingham Risk Score. Jama-J Am Med Assoc. 2009;302(21):2345-52.

372. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics. 2008;26:733-44.

373. Pandya A, Sy S, Cho S, Weinstein MC, Gaziano TA. Cost-effectiveness of 10-year risk thresholds for initiation of statin therapy for primary prevention of cardiovascular disease. Jama. 2015;314(2):142-50.

374. Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. The Lancet Oncology. 2018;19(5):e267-e74.

375. Ziebland S, Herxheimer A. How patients' experiences contribute to decision making: illustrations from DIPEx (personal experiences of health and illness). Journal of nursing management. 2008;16(4):433-9.

376. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res. 2018;2:11.

377. Hill AB. The environment and disease: association or causation? : Sage Publications; 1965.

378. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357:j2099.

379. Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. Bmj. 2014;349.

380. Callahan A, Shah NH, Chen JH. Research and reporting considerations for observational studies using electronic health record data. Ann Intern Med. 2020;172(11_Supplement):S79-S84.

381. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. J Med Internet Res. 2018;20(5):e185.

382. Hannan EL. Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. JACC: Cardiovascular Interventions. 2008;1(3):211-7.

383. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. Jama. 2001;286(7):821-30.

384. Cummings P, McKnight B. Analysis of matched cohort data. The Stata Journal. 2004;4(3):274-81.

385. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivar Behav Res. 2011;46(3):399-424.

386. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Statistics in medicine. 2014;33(13):2297-340.

387. Faraoni D, Schaefer ST. Randomized controlled trials vs. observational studies: why not just live together? BMC Anesthesiology. 2016;16(1):102.

388. Banning M. A review of clinical decision making: models and current research. J Clin Nurs. 2008;17(2):187-95.

389. Harle CA, Padman R, Downs JS, editors. Design, implementation, and preliminary evaluation of a web-based health risk calculator. 2009 42nd Hawaii International Conference on System Sciences; 2009: IEEE.

390. Harle C, Padman R, Downs J, editors. The impact of web-based diabetes risk calculators on information processing and risk perceptions. AMIA Annual Symposium Proceedings; 2008: American Medical Informatics Association.

391. Hess R, Visschers VH, Siegrist M. Risk communication with pictographs: The role of numeracy and graph processing. Judgment and Decision Making. 2011;6(3):263-74.

392. McLafferty I. Focus group interviews as a data collecting strategy. Journal of advanced nursing. 2004;48(2):187-94.

393. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. Journal of the National Cancer Institute. 2011;103(19):1436-43.

394. Smith M, Higgs J, Ellis E. Factors influencing clinical decision making. Clinical reasoning in the health professions. 2008;3:89-100.

395. Alaszewski A. Risk communication: identifying the importance of social context. Health, Risk & Society. 2005;7(2):101-5.

396. Roy B. Decision-aid and decision-making. European Journal of Operational Research. 1990;45(2-3):324-31.

397. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. Bmj. 2000;321(7262):694-6.

398. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. 2021;374:n2061.

399. Tudor K, Brooks J, Howick J, Fox R, Aveyard P. Tackling statin intolerance with n-of-1 trials (TaSINI) in primary care: protocol for a feasibility randomised trial to increase statin adherence. BMJ Open. 2020;10(2):e033070.

400. Puffer S, Torgerson DJ, Watson J. Cluster randomized controlled trials. J Eval Clin Pract. 2005;11(5):479-83.

401. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144(3):201-9.

402. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagnostic and Prognostic Research. 2018;2(1):11.

Publications from this thesis

- **Journal Article**

Cai T, Abel L, Langford O, Monaghan G, Aronson JK, Stevens RJ, et al. Associations between statins and adverse events in primary prevention of cardiovascular disease: systematic review with pairwise, network, and dose-response meta-analyses. *BMJ* 2021;374:n1537. doi:10.1136/bmj.n1537.

- **Conference Abstracts**

Cai T, Koshiaris C, Hirst JA, McManus RJ, Hobbs R, Sheppard JP. Predicting individual risk of muscle disorders in patients eligible for statin treatment: STRATIFY-StatinMD model derivation using data from electronic health records. *Atherosclerosis* 2022;355:216. doi: 10.1016/j.atherosclerosis.2022.06.853.

Cai T, Koshiaris C, Hirst JA, McManus RJ, Hobbs R, Sheppard JP. Predicting individual risk of muscle disorders in patients eligible for statin treatment: derivation and validation of STRATIFY-StatinMD prediction model. *J Hum Hypertens* 2022;36 (Suppl 1):1-22. doi: 10.1038/s41371-022-00734-5.

Appendixes

Contents

Appendix 1 Search strategies for previous systematic reviews of statins

Database: MEDLINE/PubMed

Search Queries:

1 statin[Title] OR statins[Title] OR atorvastatin[Title] OR fluvastatin[Title] OR lovastatin[Title] OR pitavastatin[Title] OR pravastatin[Title] OR rosuvastatin[Title] OR simvastatin[Title]

2 lipid lowering[Title] OR cholesterol lowering[Title] OR LDL lowering[Title]

3 statin[Title/Abstract] OR statins[Title/Abstract] OR atorvastatin[Title/Abstract] OR fluvastatin[Title/Abstract] OR lovastatin[Title/Abstract] OR pitavastatin[Title/Abstract] OR pravastatin[Title/Abstract] OR rosuvastatin[Title/Abstract] OR simvastatin[Title/Abstract]

4 2 AND 3

5 1 OR 4

6 systematic review[Title] OR meta analysis[Title]

7 systematic review[Publication Type] OR meta analysis[Publication Type]

8 6 OR 7

9 5 AND 8

10 adverse event*[Title/Abstract] OR adverse effect*[Title/Abstract] OR adverse outcome*[Title/Abstract] OR side effect*[Title/Abstract] OR unwanted effect*[Title/Abstract] OR unwanted outcome*[Title/Abstract] OR undesirable effect*[Title/Abstract] OR undesirable outcome*[Title/Abstract] OR unexpected effect*[Title/Abstract] OR unexpected outcome*[Title/Abstract] OR harm*[Title/Abstract] OR safety[Title/Abstract] OR toxicity[Title/Abstract] OR tolerance[Title/Abstract]

11 statin induced[Title/Abstract] OR statin associated[Title/Abstract] OR statin related[Title/Abstract]

12 myopathy[Title/Abstract] OR myalgia[Title/Abstract] OR myositis[Title/Abstract] OR rhabdomyolysis[Title/Abstract] OR muscle[Title/Abstract] OR muscular[Title/Abstract]

13 diabetes[Title/Abstract]

14 intracerebral haemorrhage[Title/Abstract] OR intracerebral hemorrhage[Title/Abstract] OR intracranial haemorrhage[Title/Abstract] OR intracranial hemorrhage[Title/Abstract] OR haemorrhagic stroke[Title/Abstract] OR hemorrhagic stroke[Title/Abstract]

15 cataract[Title/Abstract] OR glaucoma[Title/Abstract] OR macular degeneration[Title/Abstract]

16 liver[Title/Abstract] OR hepatic[Title/Abstract]

17 kidney[Title/Abstract] OR renal[Title/Abstract]

18 dementia[Title/Abstract] OR memory[Title/Abstract] OR cognition[Title/Abstract] OR cognitive[Title/Abstract]

19 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18

20 9 AND 19

21 ezetimibe[Title/Abstract] OR fibrates[Title/Abstract] OR alirocumab[Title/Abstract] OR evolocumab[Title/Abstract] OR PCSK9[Title/Abstract]

22 20 NOT 21

23 letter[Publication Type] OR comment[Publication Type] OR editorial[Publication Type] OR congress[Publication Type] OR protocol[Title]

24 21 NOT 23

25 English[Language]

26 "2010/01/01"[Date - Publication] : "2019/12/31"[Date - Publication]

27 24 AND 25 AND 26

Appendix 2 Search strategies for clinical trials of statins

Database: MEDLINE/PubMed

Search Queries:

1 "hydroxymethylglutaryl-coa reductase inhibitors"[MeSH Terms]

2 statin[Title/Abstract] OR statins[Title/Abstract] OR atorvastatin[Title/Abstract] OR fluvastatin[Title/Abstract] OR lovastatin[Title/Abstract] OR pitavastatin[Title/Abstract] OR pravastatin[Title/Abstract] OR rosuvastatin[Title/Abstract] OR simvastatin[Title/Abstract]

3 lipitor[Title/Abstract] OR lescol[Title/Abstract] OR mevacor[Title/Abstract] OR livalo[Title/Abstract] OR pravachol[Title/Abstract] OR crestor[Title/Abstract] OR zocor[Title/Abstract]

4 1 OR 2 OR 3

5 "randomized controlled trial"[Publication Type] OR "controlled clinical trial"[Publication Type]

6 "randomized controlled trials as topic"[MeSH Terms] OR "controlled clinical trials as topic"[MeSH Terms]

7 trial[Title/Abstract] AND (random*[Title/Abstract] OR control*[Title/Abstract] OR placebo[Title/Abstract])

8 5 OR 6 OR 7

9 animals[MeSH Terms] NOT humans[MeSH Terms]

10 8 NOT 9

11 4 AND 10

12 niacin[Title/Abstract] OR niaspan[Title/Abstract] OR bile acid[Title/Abstract] OR cholestyramine[Title/Abstract] OR colesevelam[Title/Abstract] OR colestipol[Title/Abstract] OR ezetimibe[Title/Abstract] OR fibrate[Title/Abstract] OR fibrates[Title/Abstract]OR fenofibrate[Title/Abstract] OR gemfibrozil[Title/Abstract] OR PCSK9[Title/Abstract] OR alirocumab[Title/Abstract] OR evolocumab[Title/Abstract]

13 11 NOT 12

14 comment[Publication Type] OR congress[Publication Type] OR duplicate publication[Publication Type] OR editorial[Publication Type] OR letter[Publication Type] OR meta analysis[Publication Type] OR news[Publication Type] OR published erratum[Publication Type] OR review[Publication Type] OR systematic review[Publication Type]

15 protocol[Title]

16 14 OR 15

17 13 NOT 16

18 "2013/01/01"[Date - Publication] : "present"[Date - Publication] (updated on 1st August 2020)
19 17 AND 18

Database: Embase (via Ovid)

Search Queries:

1 hydroxymethylglutaryl coenzyme a reductase inhibitor.sh.
2 (statin? or atorvastatin or fluvastatin or lovastatin or pitavastatin or pravastatin or rosuvastatin or simvastatin).ti,ab.
3 (lipitor or lescol or mevacor or livalo or pravachol or crestor or zocor).ti,ab.
4 1 or 2 or 3
5 limit 4 to (randomized controlled trial or controlled clinical trial)
6 exp animals/ not exp humans/
7 5 not 6
8 (niacin or niaspan or bile acid or cholestyramine or colesevelam or colestipol or ezetimibe or fibrate? or fenofibrate or gemfibrozil or PCSK9 or alirocumab or evolocumab).ti,ab.
9 7 not 8
10 (abstract or conference abstract or "conference review" or editorial or erratum or letter or note or "review").pt.
11 (comment* or protocol).ti.
12 10 or 11
13 9 not 12
14 limit 13 to yr="2013 - Current" (updated on 1st August 2020)

Database: CENTRAL (Cochrane Library)

Search Queries:

1 (statin OR statins OR atorvastatin OR fluvastatin OR lovastatin OR pitavastatin OR pravastatin OR rosuvastatin OR simvastatin) in Record Title
2 (niacin or niaspan or bile acid or cholestyramine or colesevelam or colestipol or ezetimibe or fibrate or fibrates or fenofibrate or gemfibrozil or PCSK9 or alirocumab or evolocumab) in Title Abstract Keyword
3 1 NOT 2

With search limits:

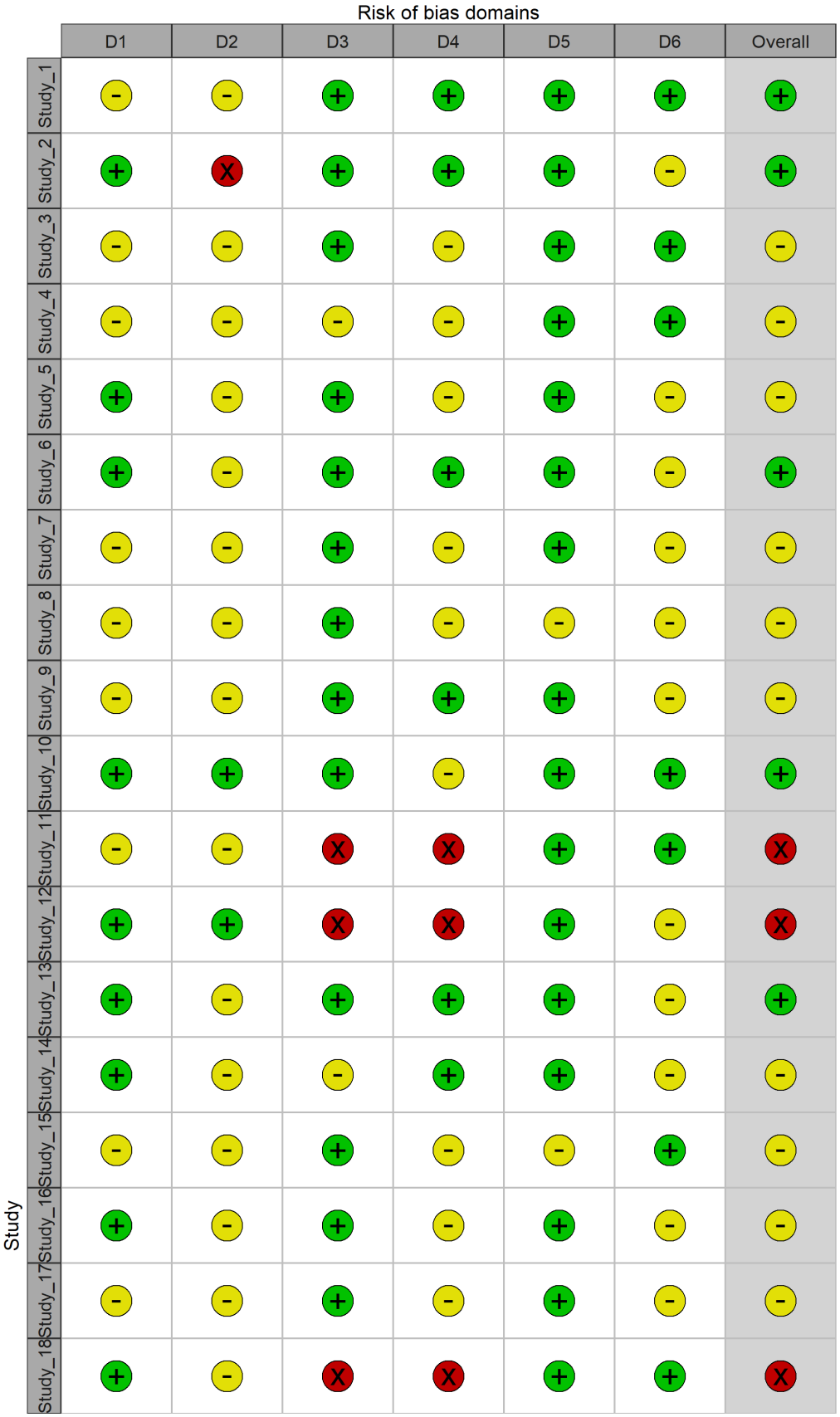
- Content type - Trials
- Trials original publication year - from 2013 to 2020 (updated on 1st August 2020)
- Search word variations - Yes

Appendix 3 Study eligibility criteria for the systematic review

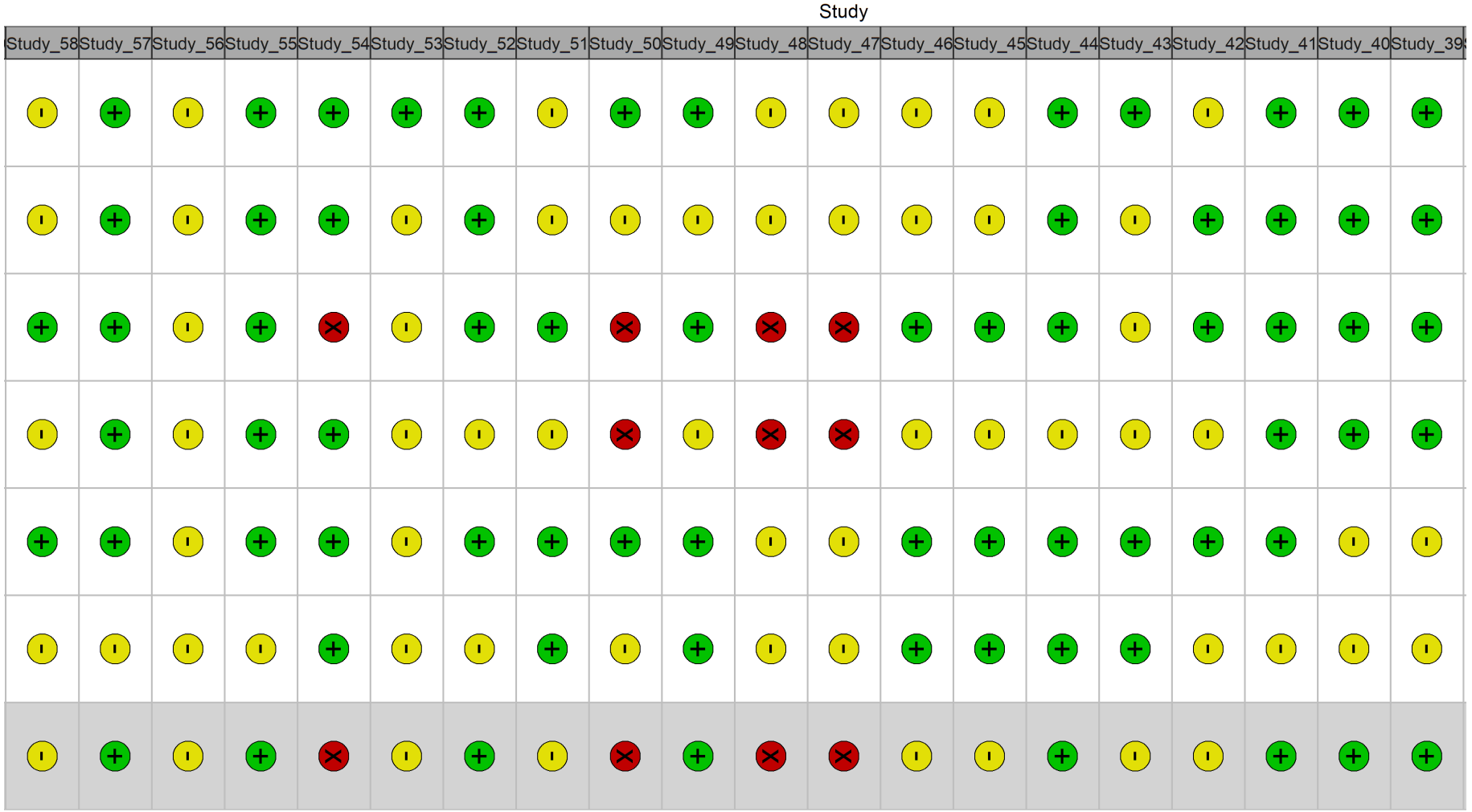
	Inclusion Criteria	Exclusion Criteria
Study design/ Settings	Randomised controlled clinical trial	• Study duration shorter than four weeks • Cross-over or self-controlled study designs • Re-analyses/subgroup analyses of the same trial data, or post-trial analyses of observational data from extended follow-up which did not adhere to the initial randomisation • Examination of the pleiotropy of statins for non-CVD conditions (e.g. cancer, chronic obstructive pulmonary disease, contrast-induced nephropathy) • Uncompleted or withdrawn studies
Participants	• Adults (over 18 years old) without existing diagnoses of CVD, previous cardiovascular events, or history of cardiac surgery • If a small proportion of CVD patients was involved, the proportion should be no more than 30%	Number of participants less than 100
Interventions	• Any of the seven types of statins in clinical use, including Atorvastatin, Fluvastatin, Lovastatin, Pitavastatin, Pravastatin, Rosuvastatin, and Simvastatin • Statins could be used in any doses • Statins could be applied as monotherapy or add-on treatment to other interventions	Statins were used in combination with other lipid-lowering medications
Comparators	• Non-statin controls: placebo, usual care, non-pharmaceutical treatment (e.g. dietary management, physical exercise), or no treatment	Comparison of different formulations (e.g. oral vs. topical) of one statin type at the same dose

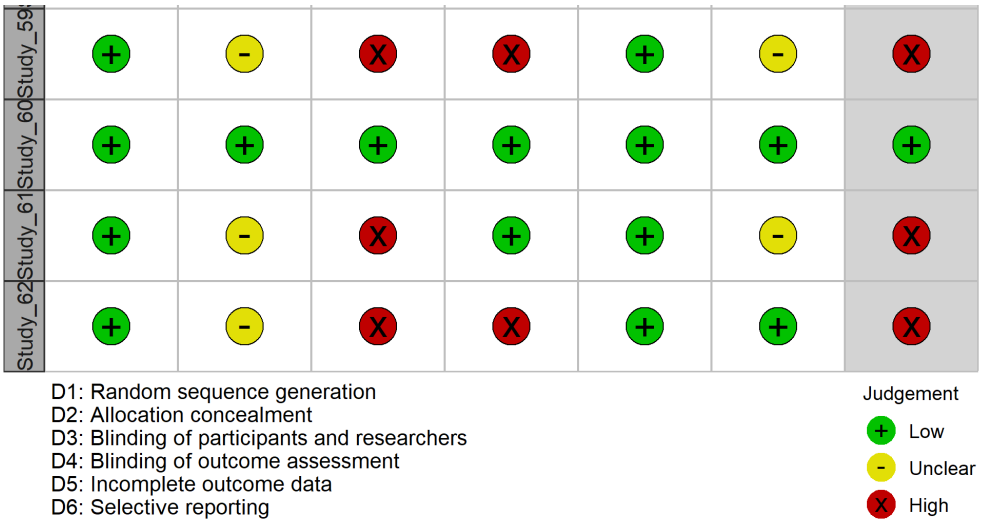
	• Statin comparators: different types or doses of statins	
Outcomes	Primary outcomes: • Muscle symptoms • Muscle disorders • Live dysfunction • Renal insufficiency • Type 2 Diabetes • Eye conditions Secondary outcomes: • Myocardial infarction • Stroke • Death from CVD	No desired outcome data available, either effect estimate with standard error for compared groups or number/rate of events in each group

Appendix 4 Risk of bias in each of the included studies



Study_19	+	+	+	+	+	-	+
Study_20	-	-	X	X	-	+	X
Study_21	-	-	+	-	+	+	-
Study_22	+	+	+	-	+	-	+
Study_23	+	-	+	-	-	+	-
Study_24	+	-	+	+	-	+	+
Study_25	-	-	+	-	X	+	X
Study_26	-	-	+	-	+	-	-
Study_27	-	-	X	X	-	+	X
Study_28	-	-	+	-	-	-	-
Study_29	-	-	+	-	+	+	-
Study_30	-	-	X	X	+	-	X
Study_31	-	-	+	+	-	-	-
Study_32	-	-	X	X	-	+	X
Study_33	+	-	X	+	+	-	X
Study_34	-	-	+	-	+	-	-
Study_35	+	-	+	-	+	+	+
Study_36	+	-	+	-	+	+	+
Study_37	-	-	+	+	X	+	X
Study_38	+	+	+	-	+	-	+


Study
Study_58
Study_57
Study_56
Study_55
Study_54
Study_53
Study_52
Study_51
Study_50
Study_49
Study_48
Study_47
Study_46
Study_45
Study_44
Study_43
Study_42
Study_41
Study_40
Study_39



*The study numbers are corresponding to their ‘study ID’ in Table 3.1 in the main text.

Appendix 5 GRADE assessment of the significant comparisons in the network meta-analyses

Outcome	Treatment	Comparator	Direct Evidence		Indirect Evidence		Combined Evidence	
			OR (95% CI)	Quality	OR (95% CI)	Quality	OR (95% CI)	Quality
Muscle Symptoms	Rosuvastatin	Control	1.08 (1.00, 1.16)	Moderate [b]	1.53 (0.99, 2.35)	Very Low [a b e]	1.09 (1.01, 1.16)	Moderate [*]
Liver Dysfunction	Atorvastatin	Control	1.30 (0.98, 1.72)	Moderate [b]	2.98 (1.30, 6.83)	Very Low [a b e]	1.41 (1.08, 1.85)	Moderate [*]
Liver Dysfunction	Lovastatin	Control	1.81 (1.24, 2.67)	High	\	\	1.81 (1.23, 2.66)	High
Liver Dysfunction	Lovastatin	Fluvastatin	\	\	2.57 (1.11, 5.93)	Low [b e]	2.57 (1.11, 5.93)	Low
Liver Dysfunction	Lovastatin	Pravastatin	1.01 (0.02, 51.64)	Very Low [a d e]	1.83 (1.09, 3.08)	Low [b e]	1.81 (1.08, 3.03)	Low [*]
Renal Insufficiency	Rosuvastatin	Control	1.13 (1.00, 1.28)	Moderate [c]	\	\	1.13 (1.00, 1.28)	Moderate
Type 2 Diabetes	Rosuvastatin	Control	1.14 (1.00, 1.30)	High	\	\	1.14 (1.00, 1.30)	High
Type 2 Diabetes	Atorvastatin	Pitavastatin	1.26 (0.34, 4.75)	Low [a e]	1.50 (1.08, 2.09)	Low [a b]	1.49 (1.08, 2.05)	Low
Type 2 Diabetes	Rosuvastatin	Pitavastatin	\	\	1.50 (1.16, 1.94)	Low [a b]	1.50 (1.16, 1.94)	Low
Eye Conditions	Rosuvastatin	Control	1.26 (1.04, 1.52)	High	\	\	1.26 (1.04, 1.52)	High

OR: odds ratio, CI: confidence interval

a. Overall risk of bias due to more than 3 studies or more than half of the included studies presenting high risk of bias.

b. Inconsistency due to different patient groups of specific conditions included in individual studies.

c. Inconsistency due to different specific conditions reported in individual studies for the outcome.

d. Indirectness due to included participants not representative of the targeted study population.

e. Imprecision due to small sample size for direct comparison or for certain treatment comparisons that were involved in the indirect evidence.

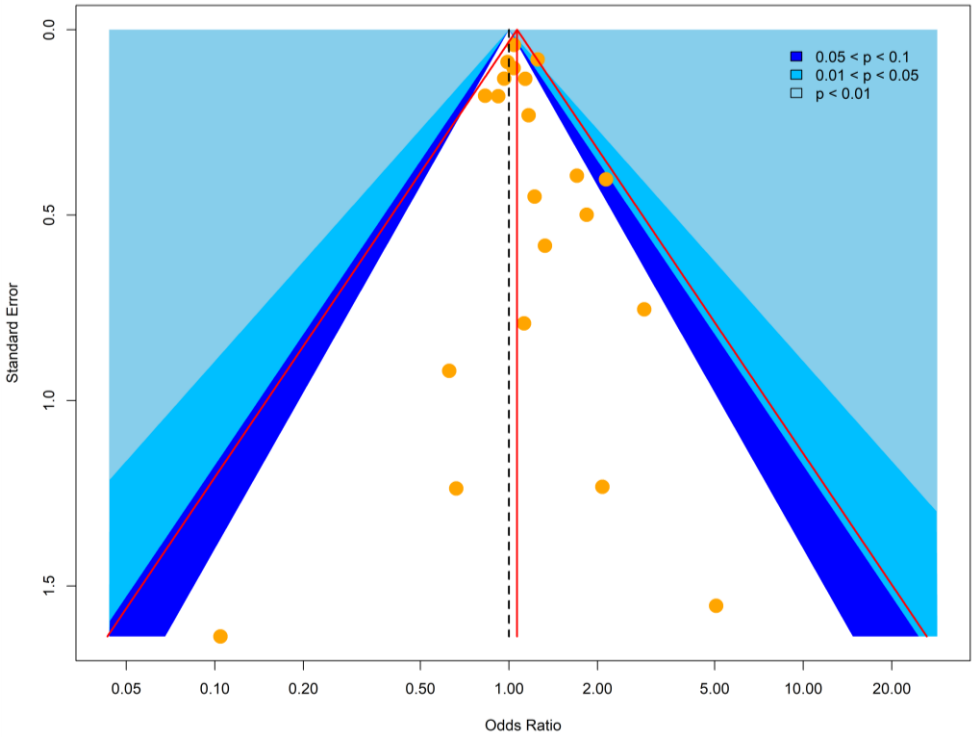
* Combined evidence quality is determined by the higher level when the quality of direct and indirect evidence is not equal.

Appendix 6 Funnel plots of publication bias in pair-wise meta-analyses

Primary (Safety) Outcomes:

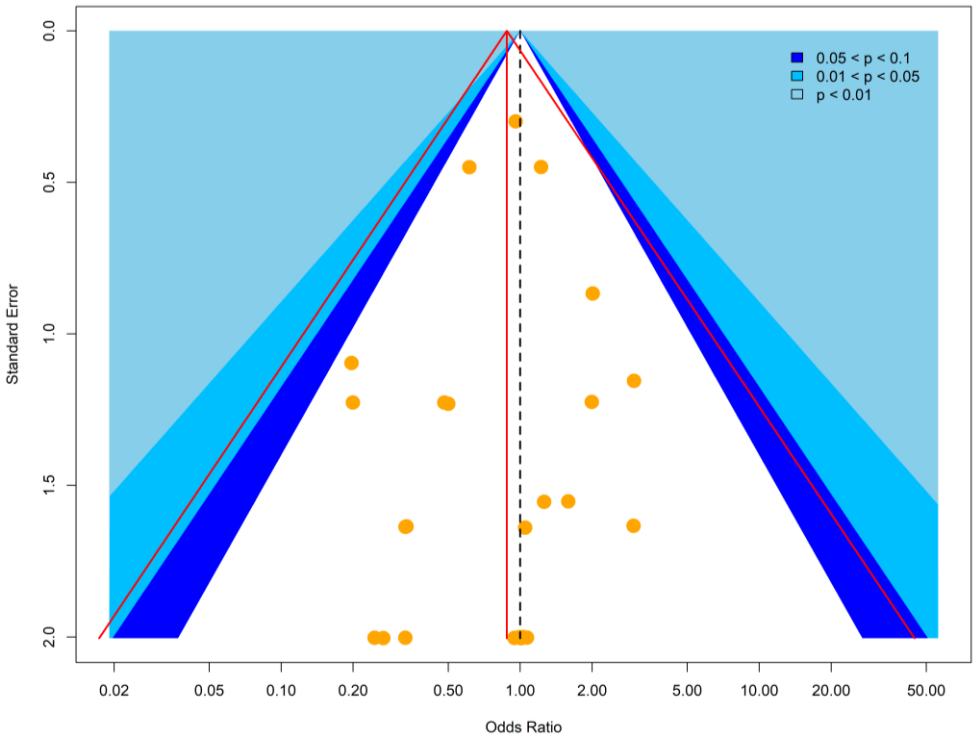
- **Muscle Symptoms**

Test of funnel plot asymmetry: P = 0.3440



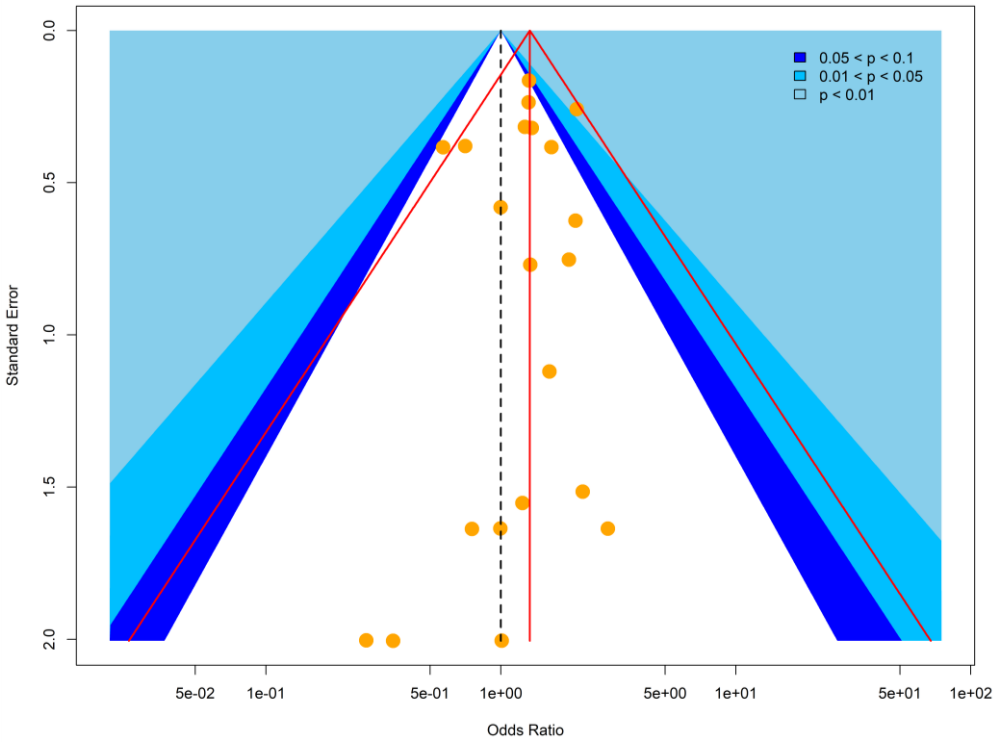
- **Muscle Disorders**

Test of funnel plot asymmetry: P = 0.9921



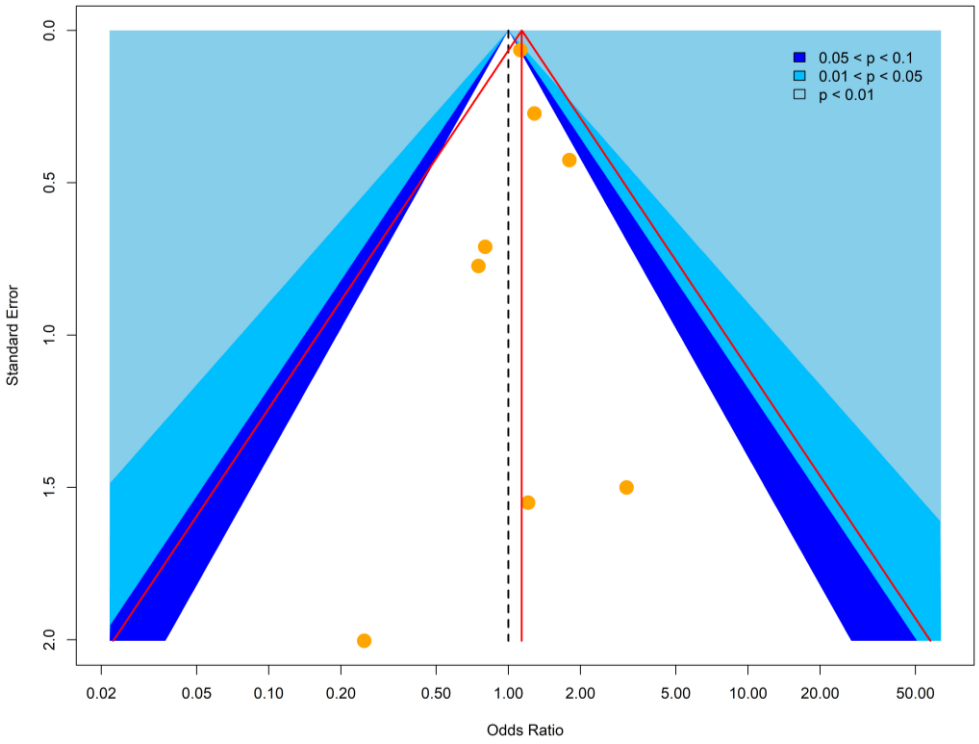
- **Liver Dysfunction**

Test of funnel plot asymmetry: P = 0.5029



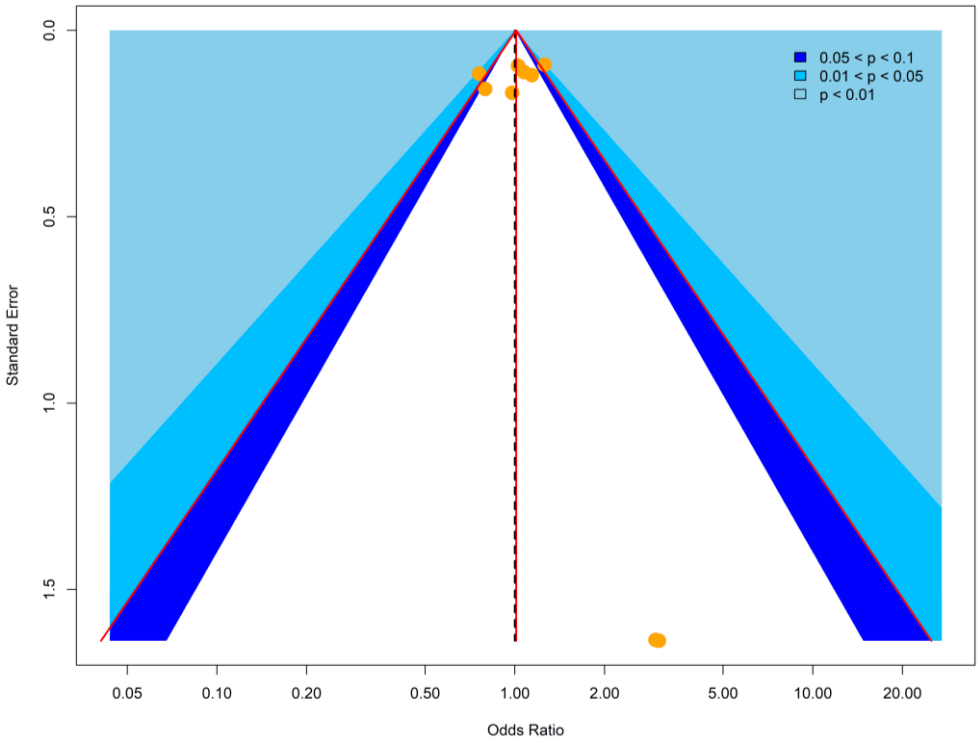
- **Renal Insufficiency**

Test of funnel plot asymmetry: the number of studies <10, no test was performed.



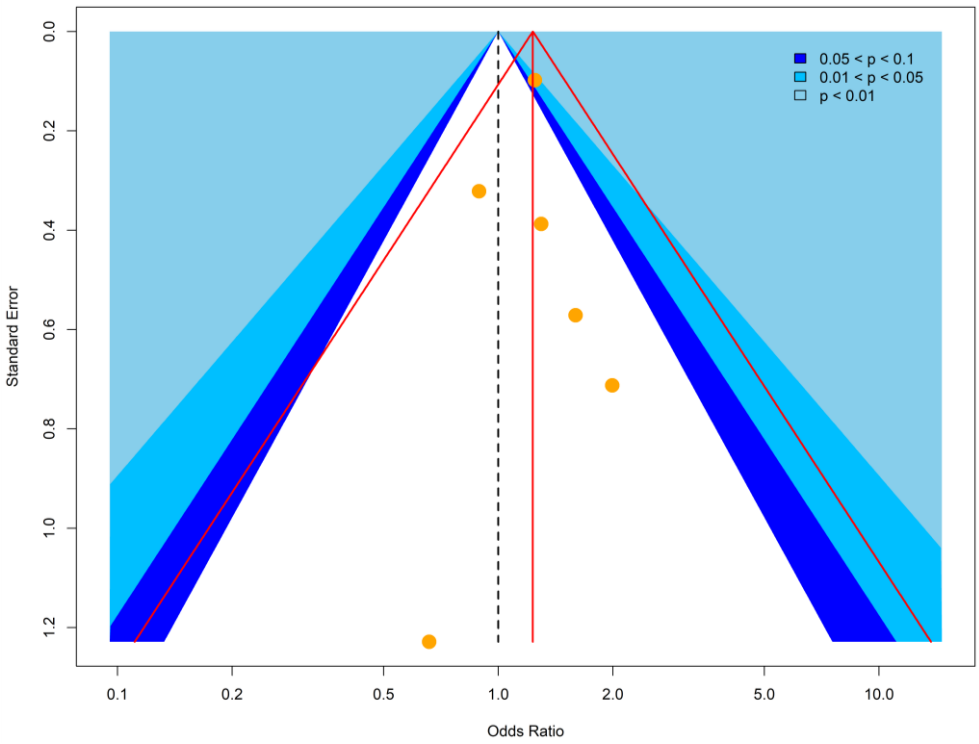
- **Type 2 Diabetes**

Test of funnel plot asymmetry: the number of studies <10, no test was performed.



- **Eye Conditions**

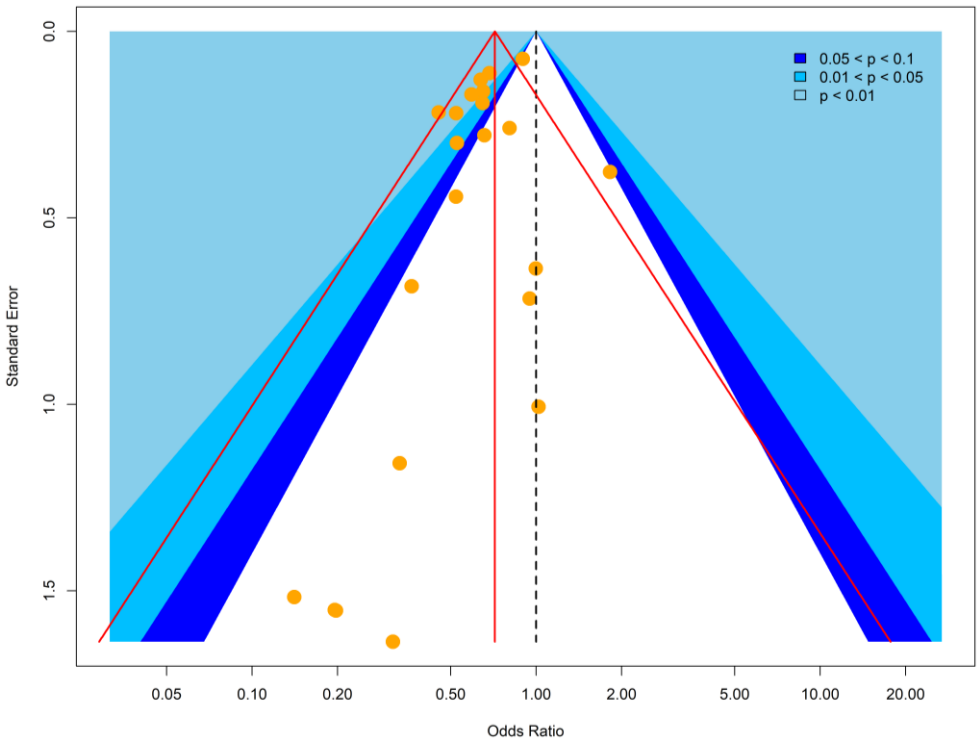
Test of funnel plot asymmetry: the number of studies <10, no test was performed.



Secondary (Efficacy) Outcomes:

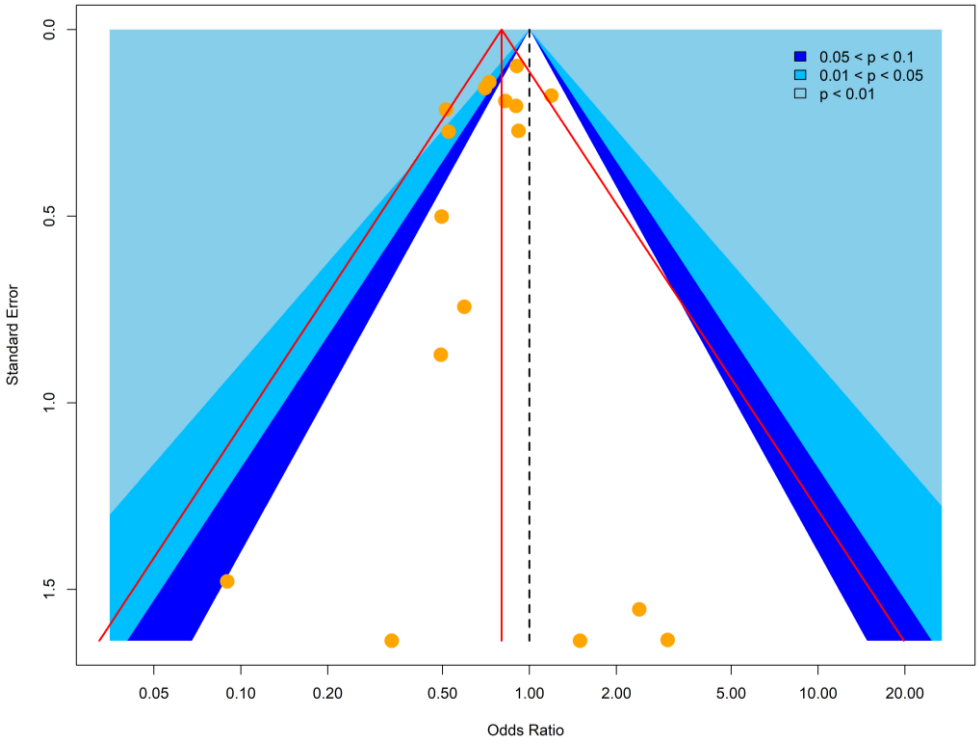
- **Myocardial Infarction**

Test of funnel plot asymmetry: P = 0.0360



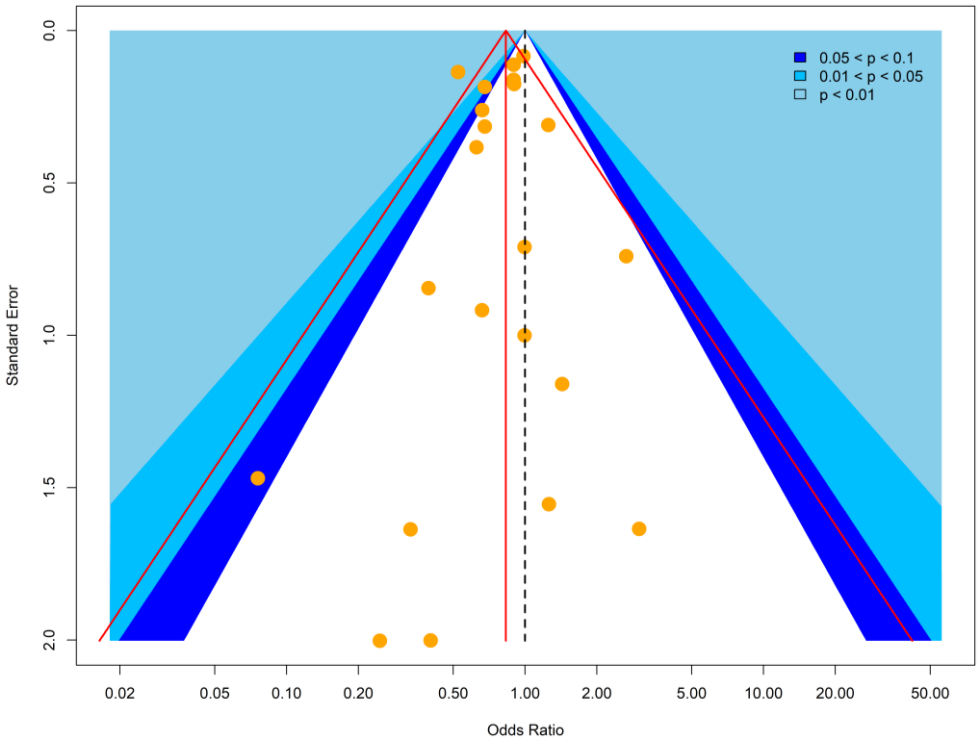
- **Stroke**

Test of funnel plot asymmetry: P = 0.4983



- **Death from CVD**

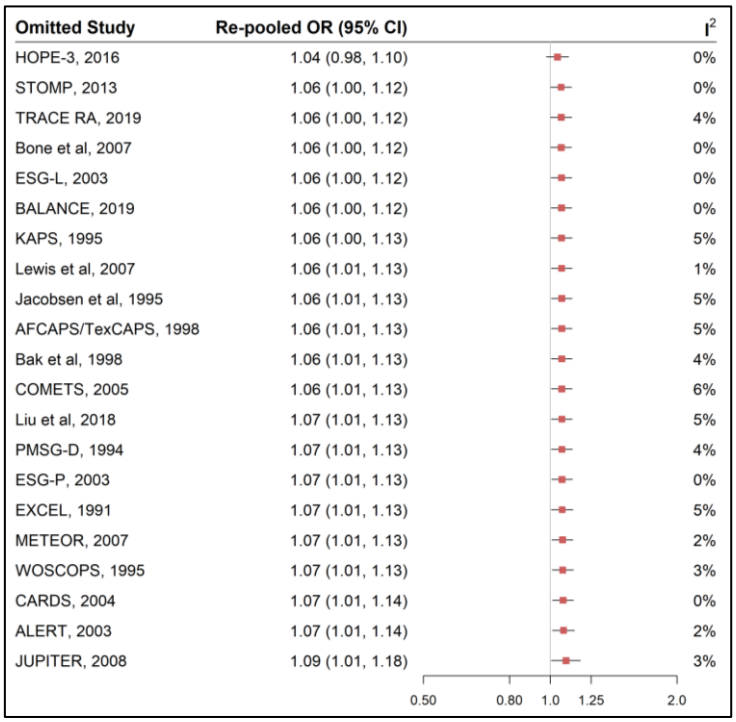
Test of funnel plot asymmetry: P = 0.6858



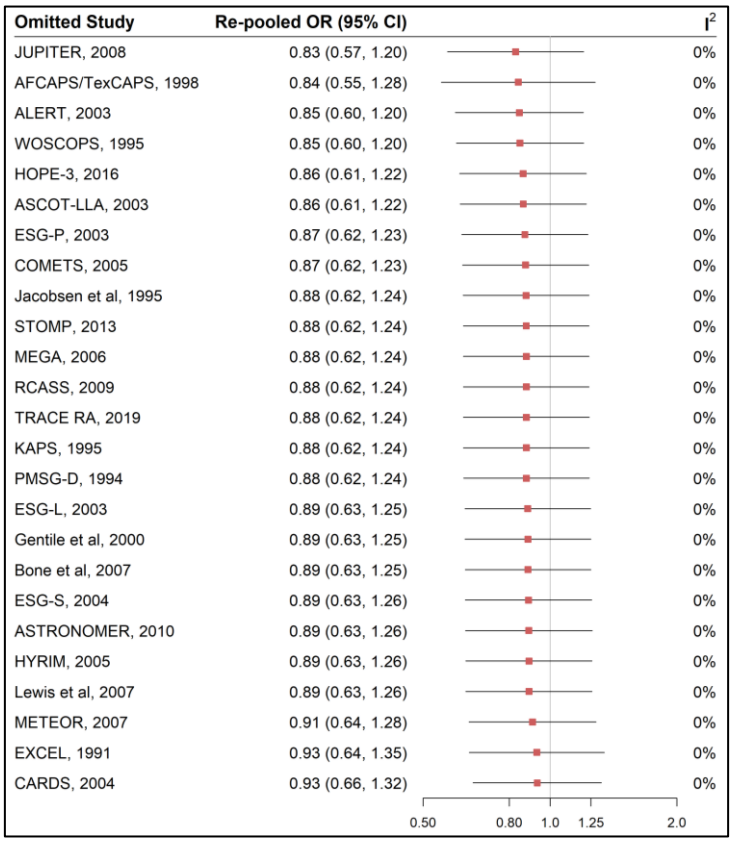
Appendix 7 Leave-one-out influence analyses for the pair-wise meta-analyses

Primary (Safety) Outcomes:

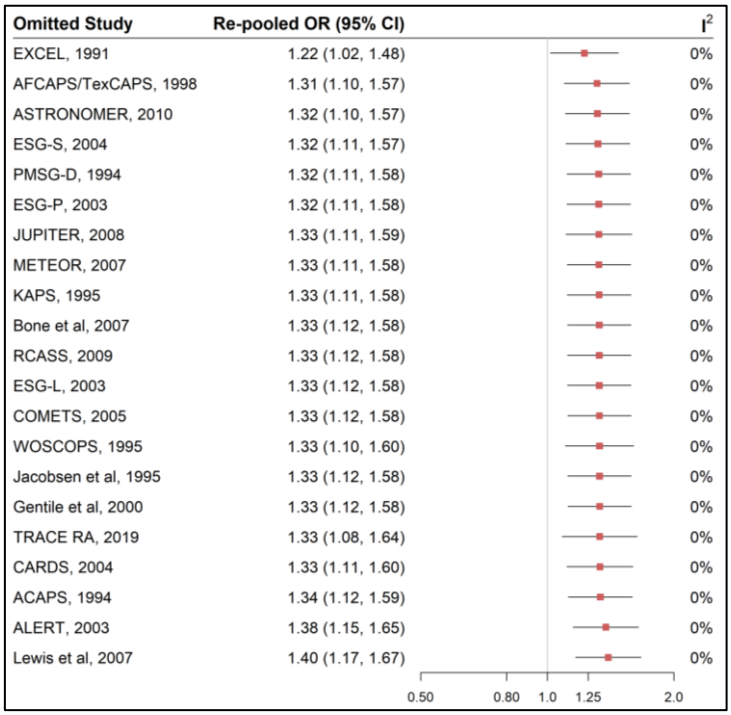
- **Muscle Symptoms**



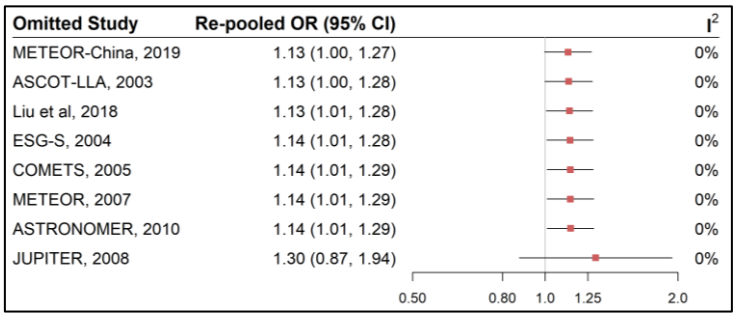
- **Muscle Disorders**



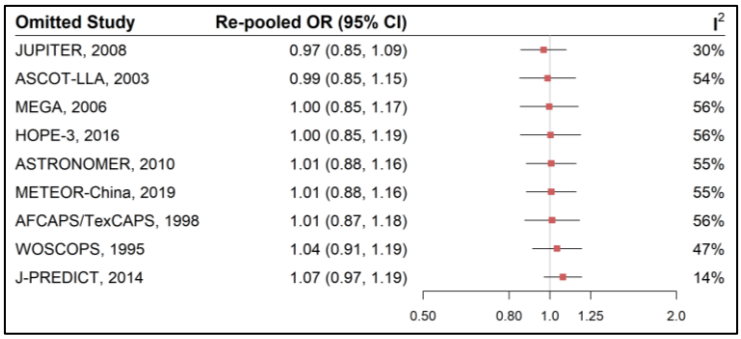
- **Liver Dysfunction**



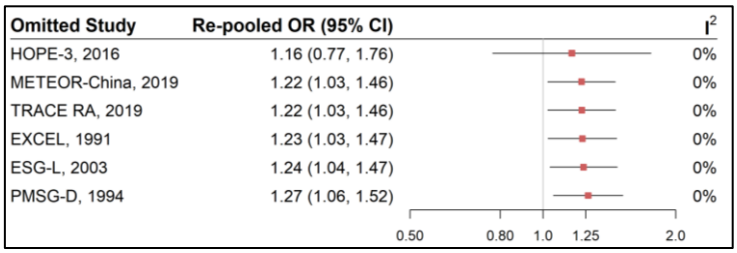
- **Renal Insufficiency**



- **Type 2 Diabetes**

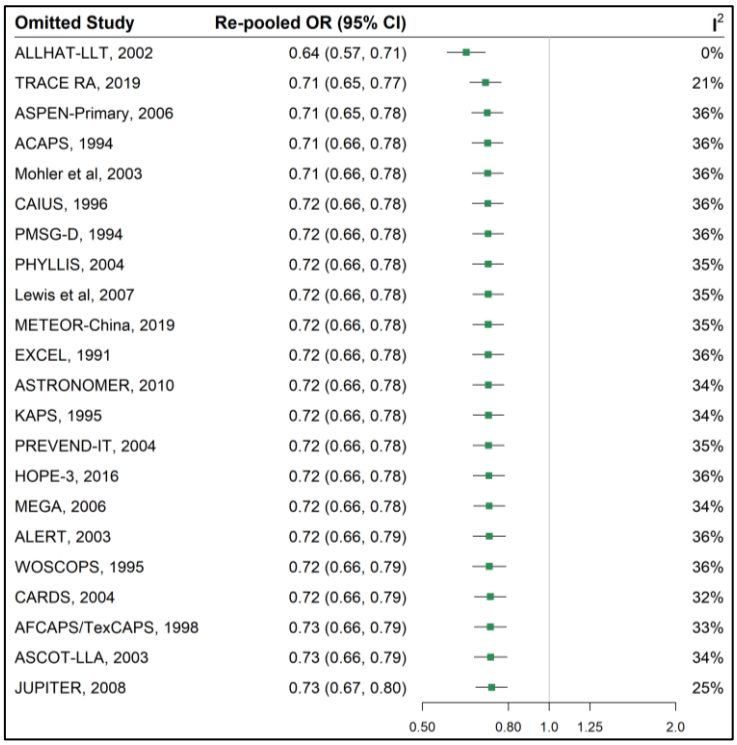


- **Eye Conditions**

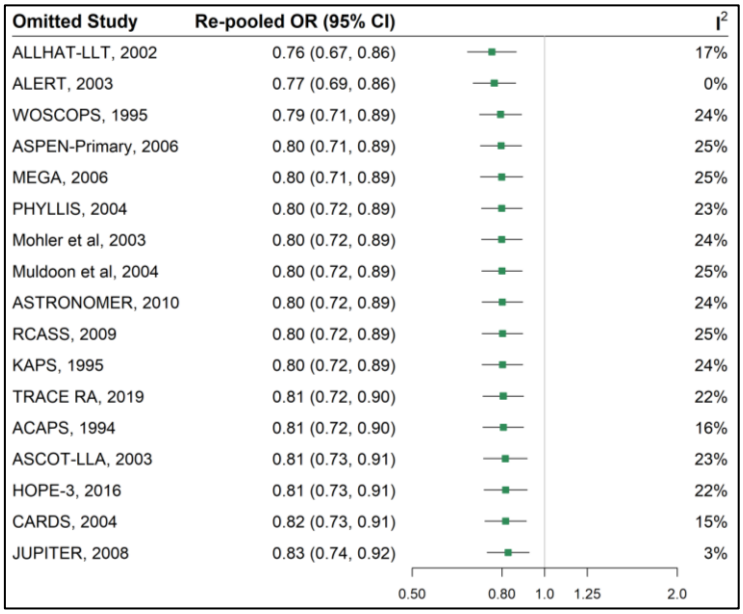


Secondary (Efficacy) Outcomes:

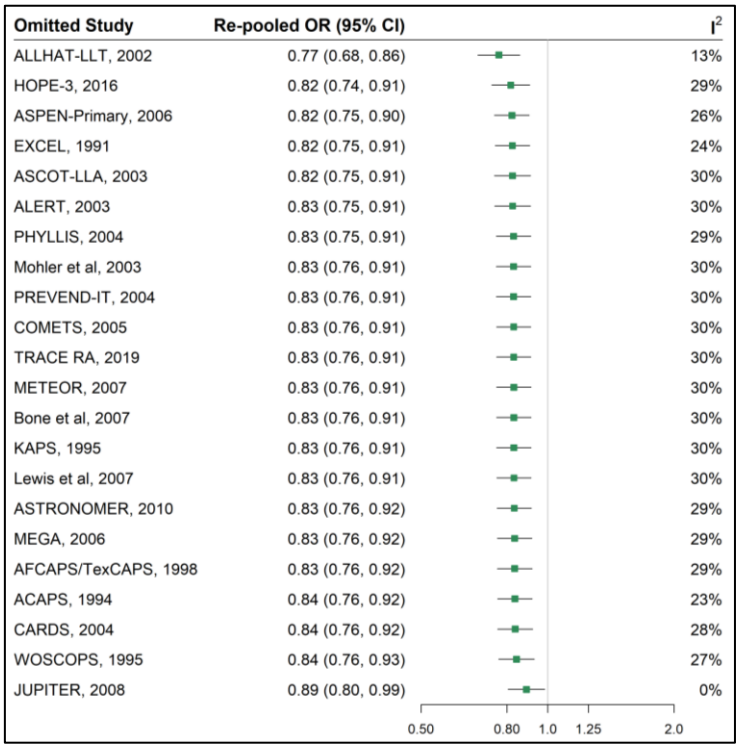
- **Myocardial Infarction**



- **Stroke**



- **Death from CVD**

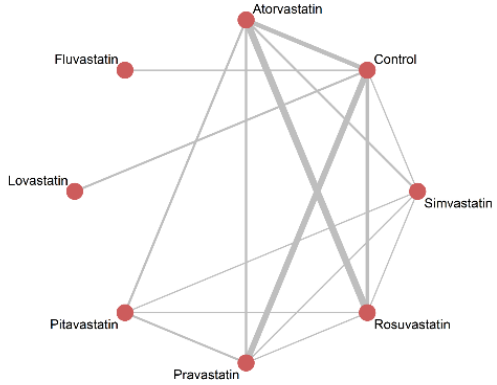


Appendix 8 Sensitivity analyses for the pair-wise meta-analyses of the secondary outcomes

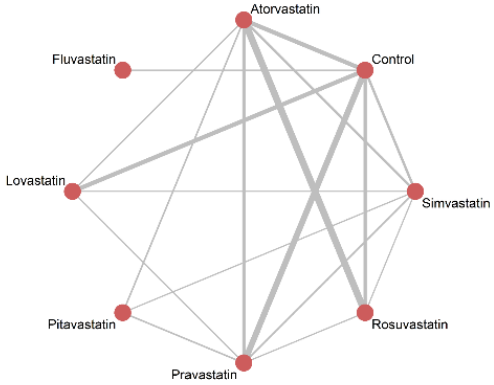
Outcome	Alternative Model of Meta-analysis	Excluding Studies with CVD Patients	Excluding Small Studies #
Myocardial Infarction	0.67 (0.59, 0.76)	0.60 (0.50, 0.71)	0.72 (0.66, 0.78)
Stroke	0.79 (0.69, 0.90)	0.65 (0.54, 0.78)	\
Death from CVD	0.81 (0.70, 0.93)	0.70 (0.60, 0.81)	\

Sensitivity analysis by excluding small studies was conducted only for myocardial infarction, of which publication bias was detected in the main analysis; 5 studies with less than 200 participants in any arm were excluded.

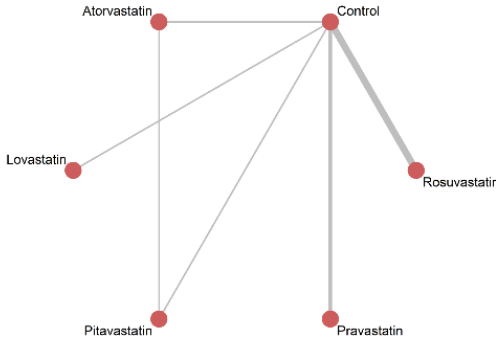
Appendix 9 Networks of treatment comparisons in the network meta-analyses



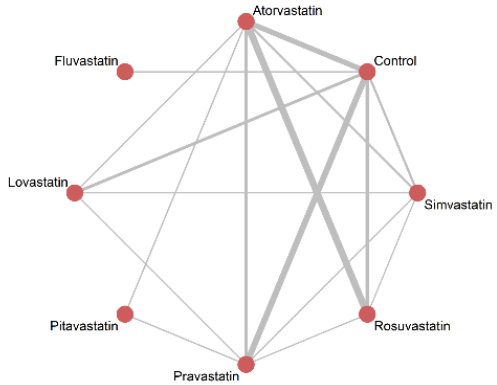
Muscle Symptoms



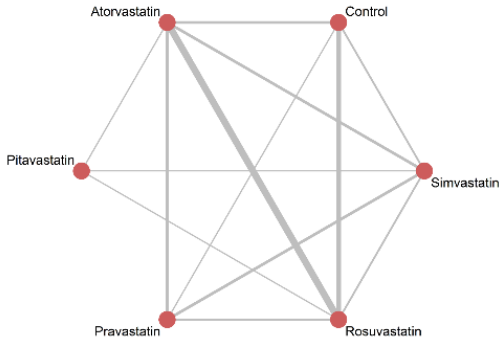
Liver Dysfunction



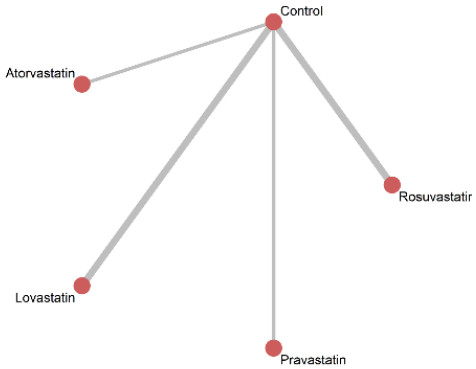
Diabetes



Muscle Disorders



Renal Insufficiency



Eye Conditions

Appendix 10 Q tests of global heterogeneity and inconsistency in network meta-analyses

Outcome	Total heterogeneity /inconsistency			Within-design heterogeneity			Between-design inconsistency (fixed-effect assumption)			Between-design inconsistency (random-effect assumption)			
	Q	df	P	Q	df	P	Q	df	P	Q	df	P	Tau^2
Muscle Symptoms	35.34	36	0.500	28.10	22	0.173	7.24	14	0.925	6.99	14	0.935	0.014
Muscle Disorders	13.27	39	0.999	9.11	26	0.999	4.16	13	0.989	4.16	13	0.989	0
Liver Dysfunction	22.22	35	0.954	10.34	21	0.974	11.88	14	0.616	11.88	14	0.616	0
Renal Insufficiency	7.47	18	0.986	4.68	8	0.791	2.79	10	0.986	2.79	10	0.986	0
Type 2 Diabetes	5.52	5	0.355	5.51	4	0.239	0.02	1	0.898	0.02	1	0.901	0.009
Eye Conditions*	0.70	2	0.705	0.70	2	0.705	/	/	/	/	/	/	/

* For eye conditions, there was no indirect comparison and only one type of study design (direct comparison) was involved in each treatment comparison.

Appendix 11 Results of the random-effects network meta-analyses

Muscle Symptoms							
Control	0.94 (0.79, 1.12)	1.01 (0.85, 1.20)	0.94 (0.77, 1.14)	1.43 (0.89, 2.28)	1.01 (0.82, 1.25)	0.92 (0.86, 0.99)	1.56 (0.76, 3.22)
1.06 (0.89, 1.26)	**Atorvastatin**	1.07 (0.84, 1.37)	1.00 (0.77, 1.29)	1.51 (0.94, 2.45)	1.07 (0.82, 1.40)	0.98 (0.82, 1.17)	1.66 (0.81, 3.39)
0.99 (0.83, 1.17)	0.93 (0.73, 1.19)	**Fluvastatin**	0.93 (0.72, 1.20)	1.41 (0.85, 2.33)	1.00 (0.76, 1.31)	0.91 (0.76, 1.10)	1.55 (0.74, 3.25)
1.07 (0.88, 1.30)	1.00 (0.77, 1.30)	1.08 (0.83, 1.40)	**Lovastatin**	1.52 (0.91, 2.53)	1.08 (0.81, 1.44)	0.98 (0.80, 1.21)	1.67 (0.79, 3.52)
0.70 (0.44, 1.12)	0.66 (0.41, 1.07)	0.71 (0.43, 1.17)	0.66 (0.40, 1.10)	**Pitavastatin**	0.71 (0.45, 1.12)	0.65 (0.40, 1.04)	1.10 (0.51, 2.34)
0.99 (0.80, 1.22)	0.93 (0.72, 1.21)	1.00 (0.76, 1.31)	0.93 (0.70, 1.23)	1.41 (0.89, 2.23)	**Pravastatin**	0.91 (0.73, 1.13)	1.54 (0.74, 3.22)
1.08 (1.01, 1.16)	1.02 (0.85, 1.23)	1.10 (0.91, 1.32)	1.02 (0.83, 1.25)	1.55 (0.96, 2.49)	1.10 (0.88, 1.37)	**Rosuvastatin**	1.70 (0.82, 3.50)
0.64 (0.31, 1.32)	0.60 (0.30, 1.23)	0.65 (0.31, 1.36)	0.60 (0.28, 1.27)	0.91 (0.43, 1.95)	0.65 (0.31, 1.35)	0.59 (0.29, 1.22)	**Simvastatin**

Muscle Disorders							
Control	1.26 (0.55, 2.90)	0.74 (0.16, 3.31)	1.21 (0.75, 1.95)	0.93 (0.13, 6.76)	1.00 (0.38, 2.63)	1.06 (0.56, 2.03)	1.10 (0.33, 3.63)
0.79 (0.34, 1.83)	**Atorvastatin**	0.58 (0.11, 3.26)	0.96 (0.37, 2.49)	0.74 (0.11, 5.03)	0.79 (0.26, 2.42)	0.84 (0.36, 1.98)	0.87 (0.25, 3.12)
1.36 (0.30, 6.09)	1.71 (0.31, 9.52)	**Fluvastatin**	1.64 (0.34, 7.94)	1.27 (0.11, 15.19)	1.36 (0.23, 8.09)	1.44 (0.28, 7.40)	1.50 (0.22, 10.17)
0.83 (0.51, 1.33)	1.04 (0.40, 2.70)	0.61 (0.13, 2.94)	**Lovastatin**	0.77 (0.10, 5.89)	0.83 (0.28, 2.41)	0.88 (0.39, 1.96)	0.91 (0.25, 3.26)
1.07 (0.15, 7.76)	1.35 (0.20, 9.18)	0.79 (0.07, 9.48)	1.30 (0.17, 9.91)	**Pitavastatin**	1.07 (0.15, 7.54)	1.14 (0.15, 8.54)	1.18 (0.13, 10.82)
1.00 (0.38, 2.63)	1.26 (0.41, 3.85)	0.74 (0.12, 4.40)	1.21 (0.41, 3.54)	0.93 (0.13, 6.58)	**Pravastatin**	1.06 (0.35, 3.19)	1.10 (0.27, 4.54)
0.94 (0.49, 1.80)	1.19 (0.51, 2.78)	0.69 (0.14, 3.56)	1.14 (0.51, 2.54)	0.88 (0.12, 6.58)	0.94 (0.31, 2.82)	**Rosuvastatin**	1.04 (0.30, 3.55)
0.91 (0.28, 2.99)	1.14 (0.32, 4.08)	0.67 (0.10, 4.55)	1.10 (0.31, 3.93)	0.85 (0.09, 7.75)	0.91 (0.22, 3.73)	0.96 (0.28, 3.30)	**Simvastatin**

Liver Dysfunction							
Control	0.72 (0.55, 0.94)	1.42 (0.67, 2.98)	0.56 (0.38, 0.83)	0.94 (0.46, 1.95)	0.98 (0.69, 1.39)	0.72 (0.44, 1.18)	0.78 (0.31, 1.95)
1.39 (1.06, 1.82)	**Atorvastatin**	1.97 (0.89, 4.35)	0.79 (0.49, 1.26)	1.31 (0.64, 2.69)	1.36 (0.89, 2.09)	1 (0.58, 1.73)	1.09 (0.43, 2.73)
0.71 (0.34, 1.48)	0.51 (0.23, 1.12)	**Fluvastatin**	0.40 (0.17, 0.92)	0.67 (0.24, 1.88)	0.69 (0.30, 1.57)	0.51 (0.21, 1.24)	0.55 (0.17, 1.79)
1.77 (1.20, 2.61)	1.27 (0.80, 2.04)	**2.51 (1.09, 5.81)**	**Lovastatin**	1.67 (0.73, 3.80)	**1.73 (1.03, 2.93)**	1.27 (0.68, 2.39)	1.39 (0.52, 3.72)
1.06 (0.51, 2.19)	0.76 (0.37, 1.57)	1.50 (0.53, 4.25)	0.60 (0.26, 1.36)	**Pitavastatin**	1.04 (0.50, 2.15)	0.76 (0.32, 1.82)	0.83 (0.27, 2.55)
1.02 (0.72, 1.45)	0.73 (0.48, 1.12)	1.45 (0.64, 3.30)	0.58 (0.34, 0.97)	0.96 (0.47, 1.99)	**Pravastatin**	0.73 (0.40, 1.35)	0.80 (0.30, 2.10)
1.39 (0.85, 2.28)	1.00 (0.58, 1.72)	1.97 (0.81, 4.82)	0.78 (0.42, 1.47)	1.31 (0.55, 3.13)	1.36 (0.74, 2.49)	**Rosuvastatin**	1.09 (0.39, 3.03)
1.28 (0.51, 3.18)	0.92 (0.37, 2.30)	1.81 (0.56, 5.88)	0.72 (0.27, 1.94)	1.21 (0.39, 3.71)	1.25 (0.48, 3.28)	0.92 (0.33, 2.56)	**Simvastatin**

Renal Insufficiency					
Control	0.83 (0.51, 1.34)	1.86 (0.15, 22.49)	0.48 (0.13, 1.72)	0.88 (0.78, 1.00)	0.85 (0.24, 3.00)
1.21 (0.75, 1.97)	**Atorvastatin**	2.26 (0.18, 27.72)	0.58 (0.16, 2.09)	1.07 (0.65, 1.76)	1.03 (0.29, 3.69)
0.54 (0.04, 6.47)	0.44 (0.04, 5.43)	**Pitavastatin**	0.26 (0.02, 3.77)	0.47 (0.04, 5.71)	0.46 (0.04, 5.87)
2.09 (0.58, 7.53)	1.73 (0.48, 6.23)	3.90 (0.27, 57.31)	**Pravastatin**	1.85 (0.51, 6.66)	1.78 (0.48, 6.63)
1.13 (1.00, 1.28)	0.93 (0.57, 1.54)	2.11 (0.18, 25.44)	0.54 (0.15, 1.95)	**Rosuvastatin**	0.96 (0.27, 3.40)
1.18 (0.33, 4.15)	0.97 (0.27, 3.48)	2.19 (0.17, 28.23)	0.56 (0.15, 2.10)	1.04 (0.29, 3.67)	**Simvastatin**

Diabetes					
Control	0.88 (0.68, 1.14)	1.02 (0.72, 1.44)	1.31 (1.02, 1.68)	1.04 (0.85, 1.26)	0.88 (0.75, 1.02)
1.13 (0.88, 1.47)	**Atorvastatin**	1.16 (0.75, 1.78)	**1.49 (1.04, 2.11)**	1.18 (0.85, 1.63)	0.99 (0.74, 1.34)
0.98 (0.69, 1.39)	0.86 (0.56, 1.33)	**Lovastatin**	1.28 (0.84, 1.97)	1.02 (0.68, 1.52)	0.86 (0.59, 1.25)
0.76 (0.59, 0.98)	0.67 (0.47, 0.96)	0.78 (0.51, 1.20)	**Pitavastatin**	0.79 (0.58, 1.09)	0.67 (0.50, 0.90)
0.96 (0.79, 1.17)	0.85 (0.61, 1.17)	0.98 (0.66, 1.47)	1.26 (0.92, 1.74)	**Pravastatin**	0.84 (0.66, 1.08)
1.14 (0.98, 1.33)	1.01 (0.75, 1.36)	1.17 (0.80, 1.70)	**1.50 (1.12, 2.01)**	1.19 (0.92, 1.52)	**Rosuvastatin**

Eye Conditions				
Control	0.63 (0.20, 1.92)	0.82 (0.40, 1.69)	1.12 (0.60, 2.11)	0.79 (0.66, 0.96)
1.60 (0.52, 4.89)	**Atorvastatin**	1.31 (0.35, 4.97)	1.79 (0.50, 6.48)	1.27 (0.41, 3.95)
1.22 (0.59, 2.52)	0.76 (0.20, 2.90)	**Lovastatin**	1.37 (0.52, 3.58)	0.97 (0.46, 2.05)
0.89 (0.47, 1.67)	0.56 (0.15, 2.02)	0.73 (0.28, 1.91)	**Pravastatin**	0.71 (0.37, 1.37)
1.26 (1.04, 1.52)	0.79 (0.25, 2.46)	1.03 (0.49, 2.18)	1.41 (0.73, 2.73)	**Rosuvastatin**

Appendix 12 Node-splitting analyses of inconsistency between direct and indirect evidence in the network meta-analyses

A. Muscle Symptoms

Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	1.11 (0.91, 1.34)	0.89 (0.61, 1.30)	1.06 (0.90, 1.26)	1.243	1.003	0.316
Fluvastatin	Control	0.99 (0.83, 1.17)	\	0.99 (0.83, 1.17)	\	\	\
Lovastatin	Control	1.07 (0.88, 1.31)	\	1.07 (0.88, 1.31)	\	\	\
Pitavastatin	Control	\	0.70 (0.44, 1.12)	0.70 (0.44, 1.12)	\	\	\
Pravastatin	Control	1.02 (0.82, 1.26)	0.67 (0.31, 1.47)	0.99 (0.81, 1.22)	1.513	-1.005	0.315
Rosuvastatin	Control	1.08 (1.00, 1.16)	1.53 (0.99, 2.35)	1.09 (1.01, 1.16)	0.705	1.565	0.118
Simvastatin	Control	0.33 (0.01, 8.22)	0.67 (0.32, 1.40)	0.64 (0.31, 1.32)	0.489	0.423	0.672
Fluvastatin	Atorvastatin	\	0.93 (0.73, 1.19)	0.93 (0.73, 1.19)	\	\	\
Lovastatin	Atorvastatin	\	1.01 (0.78, 1.31)	1.01 (0.78, 1.31)	\	\	\
Pitavastatin	Atorvastatin	0.54 (0.23, 1.25)	0.74 (0.41, 1.31)	0.66 (0.41, 1.06)	0.730	0.604	0.546
Pravastatin	Atorvastatin	1.11 (0.30, 4.17)	0.93 (0.71, 1.21)	0.93 (0.72, 1.21)	1.199	-0.263	0.792
Rosuvastatin	Atorvastatin	1.37 (0.93, 2.02)	0.95 (0.78, 1.16)	1.02 (0.86, 1.22)	1.438	-1.627	0.104
Simvastatin	Atorvastatin	0.51 (0.21, 1.21)	1.00 (0.28, 3.58)	0.60 (0.30, 1.23)	0.506	0.866	0.386
Lovastatin	Fluvastatin	\	1.09 (0.84, 1.41)	1.09 (0.84, 1.41)	\	\	\
Pitavastatin	Fluvastatin	\	0.71 (0.43, 1.16)	0.71 (0.43, 1.16)	\	\	\
Pravastatin	Fluvastatin	\	1.00 (0.77, 1.31)	1.00 (0.77, 1.31)	\	\	\
Rosuvastatin	Fluvastatin	\	1.10 (0.91, 1.32)	1.10 (0.91, 1.32)	\	\	\
Simvastatin	Fluvastatin	\	0.65 (0.31, 1.36)	0.65 (0.31, 1.36)	\	\	\
Pitavastatin	Lovastatin	\	0.65 (0.39, 1.08)	0.65 (0.39, 1.08)	\	\	\

Pravastatin	Lovastatin	\	0.92 (0.69, 1.23)	0.92 (0.69, 1.23)	\	\	\
Rosuvastatin	Lovastatin	\	1.01 (0.82, 1.25)	1.01 (0.82, 1.25)	\	\	\
Simvastatin	Lovastatin	\	0.60 (0.28, 1.26)	0.60 (0.28, 1.26)	\	\	\
Pravastatin	Pitavastatin	1.15 (0.66, 2.02)	2.08 (0.97, 4.47)	1.41 (0.90, 2.22)	0.554	1.219	0.223
Rosuvastatin	Pitavastatin	1.01 (0.17, 5.96)	1.64 (1.01, 2.67)	1.55 (0.97, 2.48)	0.616	0.516	0.606
Simvastatin	Pitavastatin	1.42 (0.44, 4.56)	0.67 (0.25, 1.80)	0.91 (0.43, 1.95)	2.123	-0.962	0.336
Rosuvastatin	Pravastatin	1.43 (0.27, 7.67)	1.10 (0.88, 1.36)	1.10 (0.88, 1.36)	1.307	-0.311	0.756
Simvastatin	Pravastatin	0.24 (0.03, 2.20)	0.74 (0.34, 1.62)	0.65 (0.31, 1.35)	0.327	0.935	0.350
Simvastatin	Rosuvastatin	0.13 (0.01, 2.66)	0.64 (0.30, 1.35)	0.59 (0.29, 1.21)	0.208	0.998	0.318

B. Muscle Disorders

Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	0.62 (0.21, 1.82)	0.99 (0.30, 3.20)	0.77 (0.35, 1.69)	0.625	-0.577	0.564
Fluvastatin	Control	1.29 (0.32, 5.24)	\	1.29 (0.32, 5.24)	\	\	\
Lovastatin	Control	0.84 (0.51, 1.36)	\	0.84 (0.52, 1.36)	\	\	\
Pitavastatin	Control	\	1.09 (0.16, 7.54)	1.09 (0.16, 7.54)	\	\	\
Pravastatin	Control	1.25 (0.42, 3.69)	0.55 (0.09, 3.31)	0.99 (0.40, 2.49)	2.275	-0.768	0.443
Rosuvastatin	Control	1.00 (0.50, 2.03)	0.74 (0.18, 3.07)	0.93 (0.50, 1.75)	1.362	-0.38	0.704
Simvastatin	Control	0.70 (0.11, 4.44)	1.14 (0.23, 5.79)	0.91 (0.28, 2.99)	0.609	0.394	0.693
Fluvastatin	Atorvastatin	\	1.68 (0.34, 8.38)	1.68 (0.34, 8.38)	\	\	\
Lovastatin	Atorvastatin	1.05 (0.02, 53.53)	1.10 (0.43, 2.82)	1.09 (0.43, 2.72)	0.958	0.021	0.984
Pitavastatin	Atorvastatin	2.00 (0.18, 22.15)	0.83 (0.04, 16.57)	1.42 (0.22, 9.23)	2.398	-0.447	0.655
Pravastatin	Atorvastatin	0.86 (0.15, 5.03)	1.80 (0.47, 6.87)	1.29 (0.45, 3.74)	0.476	0.656	0.512
Rosuvastatin	Atorvastatin	1.22 (0.40, 3.70)	1.36 (0.40, 4.67)	1.21 (0.53, 2.77)	0.892	0.135	0.892
Simvastatin	Atorvastatin	0.99 (0.22, 4.37)	0.81 (0.09, 7.71)	1.18 (0.33, 4.18)	1.217	-0.143	0.886
Lovastatin	Fluvastatin	\	0.65 (0.15, 2.84)	0.65 (0.15, 2.84)	\	\	\
Pitavastatin	Fluvastatin	\	0.84 (0.08, 9.17)	0.84 (0.08, 9.17)	\	\	\
Pravastatin	Fluvastatin	\	0.77 (0.14, 4.09)	0.77 (0.14, 4.09)	\	\	\
Rosuvastatin	Fluvastatin	\	0.72 (0.16, 3.34)	0.72 (0.16, 3.34)	\	\	\
Simvastatin	Fluvastatin	\	0.70 (0.11, 4.42)	0.70 (0.11, 4.42)	\	\	\
Pitavastatin	Lovastatin	\	1.30 (0.18, 9.54)	1.30 (0.18, 9.54)	\	\	\
Pravastatin	Lovastatin	0.99 (0.02, 50.38)	1.20 (0.41, 3.51)	1.19 (0.42, 3.33)	0.824	0.093	0.926
Rosuvastatin	Lovastatin	\	1.11 (0.51, 2.46)	1.11 (0.51, 2.46)	\	\	\

Simvastatin	Lovastatin	1.03 (0.02, 52.32)	1.08 (0.27, 4.25)	1.09 (0.30, 3.90)	0.950	0.024	0.981
Pravastatin	Pitavastatin	1.44 (0.09, 23.13)	0.60 (0.04, 8.50)	0.91 (0.13, 6.20)	2.398	-0.447	0.655
Rosuvastatin	Pitavastatin	\	0.86 (0.12, 6.16)	0.86 (0.12, 6.16)	\	\	\
Simvastatin	Pitavastatin	\	0.83 (0.09, 7.42)	0.83 (0.09, 7.42)	\	\	\
Rosuvastatin	Pravastatin	2.30 (0.09, 56.57)	0.86 (0.28, 2.64)	0.94 (0.33, 2.70)	2.666	-0.567	0.571
Simvastatin	Pravastatin	2.51 (0.25, 24.95)	0.34 (0.05, 2.47)	0.92 (0.23, 3.70)	7.357	-1.289	0.197
Simvastatin	Rosuvastatin	1.64 (0.22, 12.44)	0.51 (0.10, 2.60)	0.98 (0.28, 3.34)	3.217	-0.880	0.379

C. Liver Dysfunction

Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	1.30 (0.98, 1.72)	2.98 (1.30, 6.83)	1.41 (1.08, 1.85)	0.435	-1.860	0.063
Fluvastatin	Control	0.71 (0.34, 1.48)	\	0.71 (0.34, 1.48)	\	\	\
Lovastatin	Control	1.81 (1.24, 2.67)	\	1.81 (1.23, 2.66)	\	\	\
Pitavastatin	Control	\	1.04 (0.51, 2.12)	1.04 (0.51, 2.12)	\	\	\
Pravastatin	Control	1.08 (0.75, 1.55)	0.51 (0.17, 1.49)	1.00 (0.71, 1.41)	2.133	-1.302	0.193
Rosuvastatin	Control	1.49 (0.88, 2.54)	0.76 (0.20, 2.96)	1.39 (0.85, 2.28)	1.956	-0.902	0.367
Simvastatin	Control	1.74 (0.48, 6.28)	0.94 (0.25, 3.50)	1.30 (0.52, 3.21)	1.851	-0.656	0.512
Fluvastatin	Atorvastatin	\	0.50 (0.23, 1.10)	0.50 (0.23, 1.10)	\	\	\
Lovastatin	Atorvastatin	1.05 (0.02, 53.53)	1.29 (0.80, 2.06)	1.28 (0.80, 2.04)	0.815	0.101	0.919
Pitavastatin	Atorvastatin	0.70 (0.27, 1.82)	0.78 (0.27, 2.23)	0.74 (0.36, 1.49)	0.907	0.136	0.892
Pravastatin	Atorvastatin	0.20 (0.05, 0.88)	0.78 (0.51, 1.21)	0.71 (0.47, 1.08)	0.259	1.730	0.084
Rosuvastatin	Atorvastatin	0.62 (0.20, 1.89)	1.08 (0.59, 1.98)	0.98 (0.57, 1.69)	0.570	0.864	0.388
Simvastatin	Atorvastatin	0.40 (0.12, 1.40)	1.35 (0.36, 5.05)	0.92 (0.37, 2.29)	0.298	1.308	0.191
Lovastatin	Fluvastatin	\	2.57 (1.11, 5.93)	2.57 (1.11, 5.93)	\	\	\
Pitavastatin	Fluvastatin	\	1.48 (0.53, 4.13)	1.48 (0.53, 4.13)	\	\	\
Pravastatin	Fluvastatin	\	1.42 (0.62, 3.22)	1.42 (0.62, 3.22)	\	\	\
Rosuvastatin	Fluvastatin	\	1.97 (0.81, 4.81)	1.97 (0.81, 4.81)	\	\	\
Simvastatin	Fluvastatin	\	1.84 (0.57, 5.94)	1.84 (0.57, 5.94)	\	\	\
Pitavastatin	Lovastatin	\	0.57 (0.26, 1.29)	0.57 (0.26, 1.29)	\	\	\
Pravastatin	Lovastatin	0.99 (0.02, 50.38)	0.55 (0.33, 0.92)	0.55 (0.33, 0.92)	1.807	-0.292	0.770
Rosuvastatin	Lovastatin	\	0.77 (0.41, 1.43)	0.77 (0.41, 1.43)	\	\	\

Simvastatin	Lovastatin	1.03 (0.02, 52.32)	0.71 (0.25, 1.97)	0.72 (0.27, 1.91)	1.447	-0.178	0.859
Pravastatin	Pitavastatin	0.89 (0.33, 2.40)	1.04 (0.38, 2.89)	0.96 (0.47, 1.96)	0.854	0.218	0.827
Rosuvastatin	Pitavastatin	\	1.34 (0.57, 3.15)	1.34 (0.57, 3.15)	\	\	\
Simvastatin	Pitavastatin	1.95 (0.04, 99.09)	1.20 (0.37, 3.82)	1.24 (0.41, 3.79)	1.633	-0.235	0.814
Rosuvastatin	Pravastatin	0.77 (0.02, 38.64)	1.44 (0.79, 2.64)	1.39 (0.76, 2.53)	0.531	0.313	0.755
Simvastatin	Pravastatin	2.51 (0.25, 24.95)	1.20 (0.35, 4.12)	1.30 (0.50, 3.38)	2.082	-0.551	0.581
Simvastatin	Rosuvastatin	4.92 (0.24, 102.75)	0.84 (0.25, 2.84)	0.93 (0.34, 2.58)	5.839	-1.057	0.290

D. Renal Insufficiency

Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	1.30 (0.77, 2.19)	0.93 (0.24, 3.58)	1.23 (0.76, 1.99)	1.403	0.458	0.647
Pitavastatin	Control	\	0.54 (0.04, 6.44)	0.54 (0.04, 6.44)	\	\	\
Pravastatin	Control	5.23 (0.24, 112.06)	1.75 (0.39, 7.92)	2.03 (0.62, 6.66)	2.982	-0.627	0.531
Rosuvastatin	Control	1.13 (1.00, 1.28)	1.67 (0.46, 6.08)	1.13 (1.00, 1.28)	0.674	0.597	0.551
Simvastatin	Control	1.91 (0.21, 17.74)	1.00 (0.20, 5.07)	1.17 (0.35, 3.90)	1.909	-0.459	0.646
Pitavastatin	Atorvastatin	1.00 (0.02, 50.89)	0.48 (0.01, 29.90)	0.44 (0.04, 5.34)	2.067	-0.250	0.803
Pravastatin	Atorvastatin	1.65 (0.50, 5.50)	\	1.65 (0.50, 5.50)	\	\	\
Rosuvastatin	Atorvastatin	1.26 (0.39, 4.09)	0.86 (0.50, 1.48)	0.92 (0.56, 1.51)	1.460	-0.573	0.567
Simvastatin	Atorvastatin	1.41 (0.27, 7.25)	0.89 (0.06, 12.44)	0.95 (0.28, 3.25)	1.583	-0.290	0.772
Pravastatin	Pitavastatin	\	3.79 (0.26, 54.47)	3.79 (0.26, 54.47)	\	\	\
Rosuvastatin	Pitavastatin	3.06 (0.12, 76.02)	1.85 (0.03, 113.22)	2.11 (0.18, 25.28)	1.658	-0.190	0.850
Simvastatin	Pitavastatin	1.95 (0.04, 99.09)	2.37 (0.08, 66.88)	2.18 (0.17, 27.82)	0.826	0.073	0.942
Rosuvastatin	Pravastatin	0.87 (0.17, 4.50)	0.43 (0.07, 2.70)	0.56 (0.17, 1.82)	2.012	-0.556	0.578
Simvastatin	Pravastatin	0.58 (0.16, 2.14)	\	0.58 (0.16, 2.14)	\	\	\
Simvastatin	Rosuvastatin	0.66 (0.11, 3.96)	1.25 (0.23, 6.69)	1.04 (0.31, 3.44)	0.526	0.513	0.608

E. Type 2 Diabetes

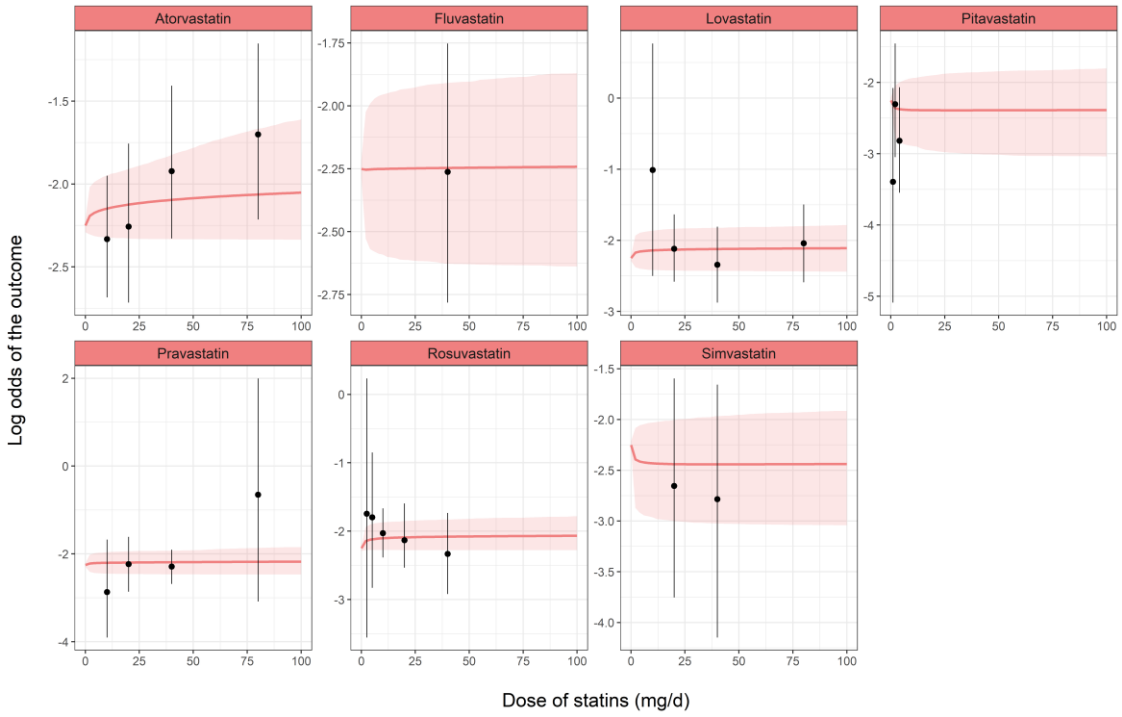
Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	1.14 (0.90, 1.44)	0.96 (0.25, 3.67)	1.14 (0.9, 1.43)	1.191	0.251	0.802
Lovastatin	Control	0.98 (0.71, 1.36)	\	0.98 (0.71, 1.36)	\	\	\
Pitavastatin	Control	0.76 (0.60, 0.95)	0.90 (0.24, 3.47)	0.76 (0.61, 0.96)	0.840	0.251	0.802
Pravastatin	Control	0.97 (0.81, 1.16)	\	0.97 (0.81, 1.16)	\	\	\
Rosuvastatin	Control	1.14 (1.00, 1.30)	\	1.14 (1.00, 1.30)	\	\	\
Lovastatin	Atorvastatin	\	0.86 (0.58, 1.29)	0.86 (0.58, 1.29)	\	\	\
Pitavastatin	Atorvastatin	0.79 (0.21, 2.98)	0.67 (0.48, 0.92)	0.67 (0.49, 0.92)	1.191	-0.251	0.802
Pravastatin	Atorvastatin	\	0.85 (0.64, 1.15)	0.85 (0.64, 1.15)	\	\	\
Rosuvastatin	Atorvastatin	\	1.01 (0.77, 1.31)	1.01 (0.77, 1.31)	\	\	\
Pitavastatin	Lovastatin	\	0.78 (0.52, 1.16)	0.78 (0.52, 1.16)	\	\	\
Pravastatin	Lovastatin	\	0.99 (0.68, 1.44)	0.99 (0.68, 1.44)	\	\	\
Rosuvastatin	Lovastatin	\	1.17 (0.82, 1.66)	1.17 (0.82, 1.66)	\	\	\
Pravastatin	Pitavastatin	\	1.27 (0.95, 1.70)	1.27 (0.95, 1.70)	\	\	\
Rosuvastatin	Pitavastatin	\	1.50 (1.16, 1.94)	1.50 (1.16, 1.94)	\	\	\
Rosuvastatin	Pravastatin	\	1.18 (0.95, 1.47)	1.18 (0.95, 1.47)	\	\	\

F. Eye Conditions

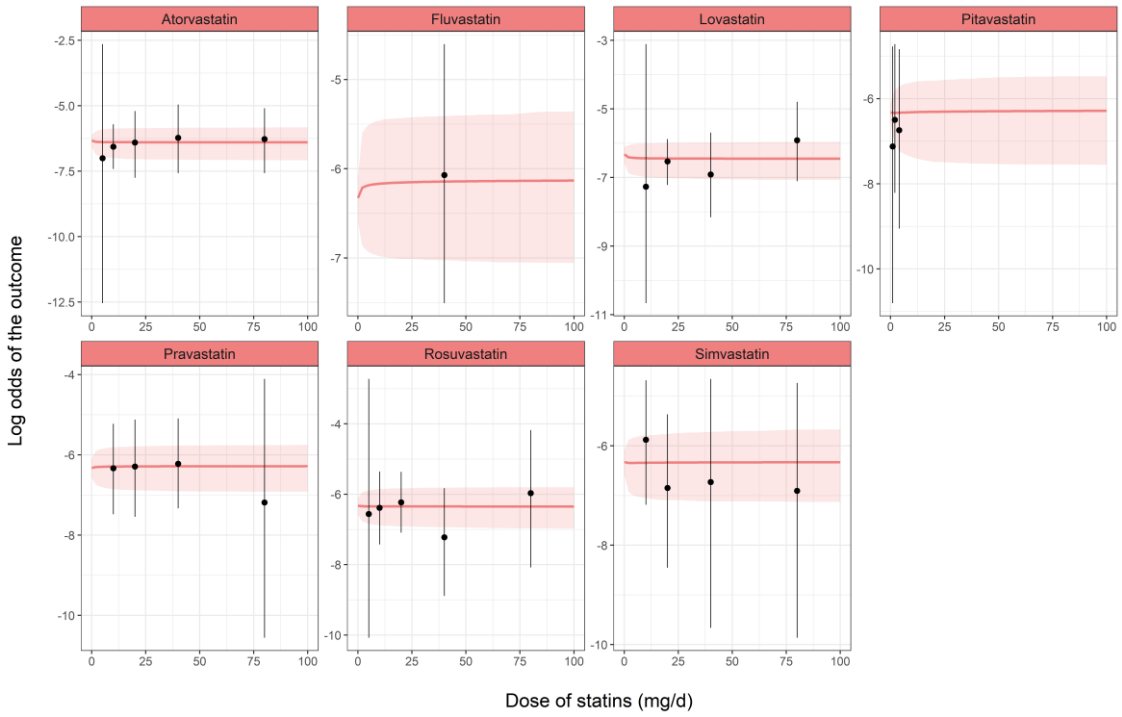
Treatment	Comparator	Direct	Indirect	NMA	Direct/Indirect	Z	P
Atorvastatin	Control	1.60 (0.52, 4.89)	\	1.60 (0.52, 4.89)	\	\	\
Lovastatin	Control	1.23 (0.60, 2.53)	\	1.23 (0.60, 2.53)	\	\	\
Pravastatin	Control	0.89 (0.47, 1.67)	\	0.89 (0.47, 1.67)	\	\	\
Rosuvastatin	Control	1.26 (1.04, 1.52)	\	1.26 (1.04, 1.52)	\	\	\
Lovastatin	Atorvastatin	\	0.77 (0.20, 2.92)	0.77 (0.20, 2.92)	\	\	\
Pravastatin	Atorvastatin	\	0.56 (0.15, 2.02)	0.56 (0.15, 2.02)	\	\	\
Rosuvastatin	Atorvastatin	\	0.79 (0.25, 2.46)	0.79 (0.25, 2.46)	\	\	\
Pravastatin	Lovastatin	\	0.72 (0.28, 1.89)	0.72 (0.28, 1.89)	\	\	\
Rosuvastatin	Lovastatin	\	1.02 (0.49, 2.16)	1.02 (0.49, 2.16)	\	\	\
Rosuvastatin	Pravastatin	\	1.41 (0.73, 2.73)	1.41 (0.73, 2.73)	\	\	\

Appendix 13 Dose-response curves of the adverse effects of individual statins

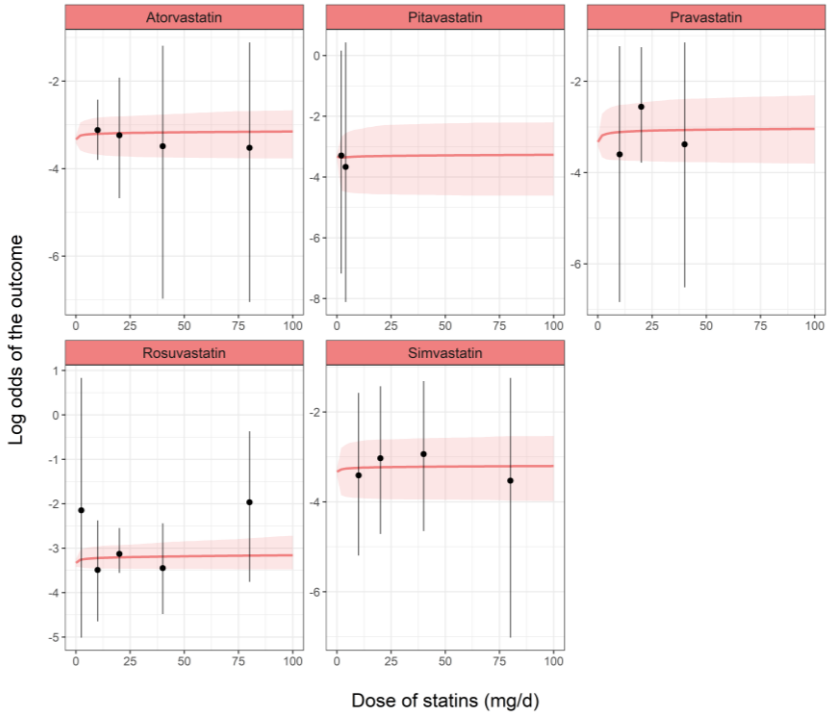
Muscle Symptoms



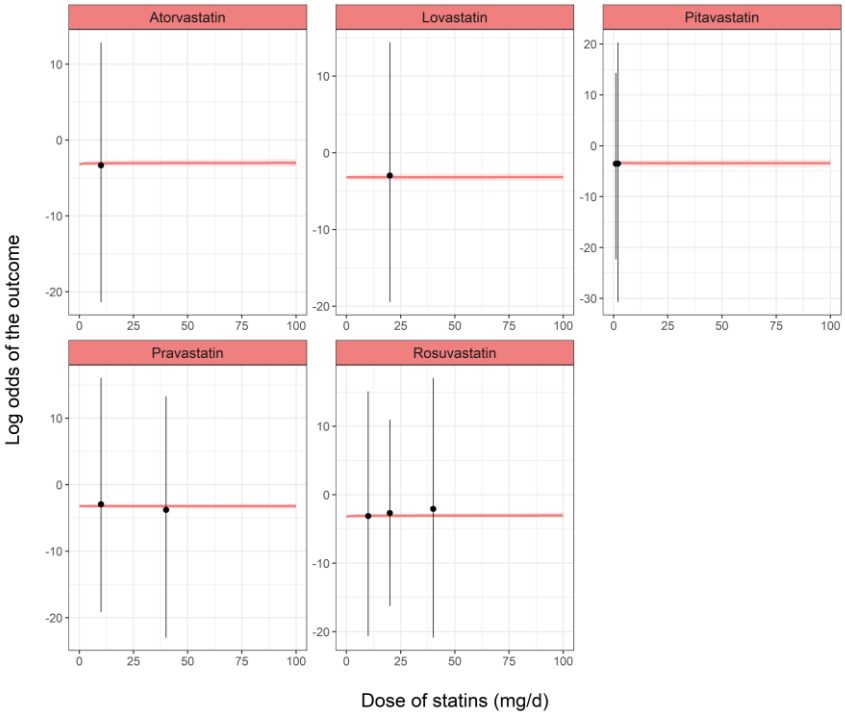
Muscle Disorders



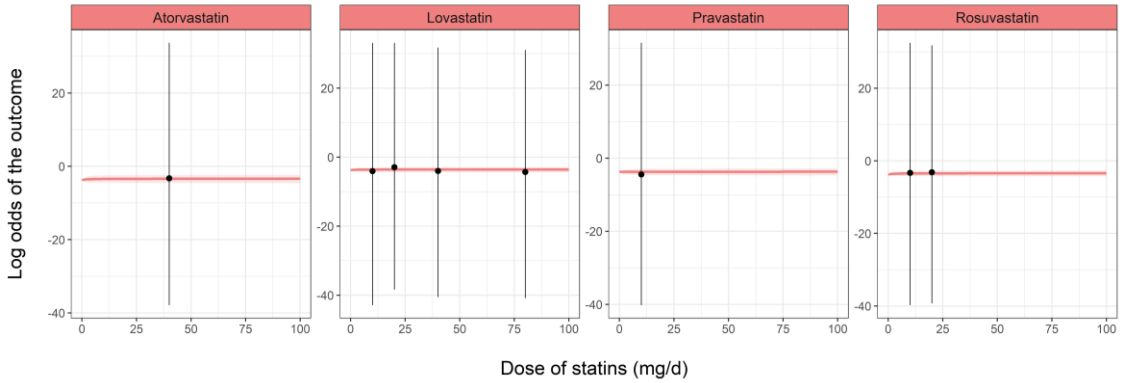
Renal Insufficiency



Type 2 Diabetes



Eye Conditions



Appendix 14 List of the ICD codes for the outcome (serious muscle disorders)

Code	Description
359.4	Toxic myopathy
359.7	Inflammatory and immune myopathies, not elsewhere classified
359.8	Other myopathies
359.9	Myopathy, unspecified
G72	Other myopathies
G72.0	Drug-induced myopathy
G72.2	Myopathy due to other toxic agents
G72.4	Inflammatory and immune myopathies, not elsewhere classified
G72.8	Other specified myopathies
G72.9	Myopathy, unspecified
M60	Myositis
M60.1	Interstitial myositis
M60.8	Other myositis
M60.9	Myositis, unspecified
M62	Other disorders of muscle
M62.0	Separation of muscle (nontraumatic)
M62.1	Other rupture of muscle (nontraumatic)
M62.4	Contracture of muscle
M62.5	Muscle wasting and atrophy, not elsewhere classified
M62.8	Other specified disorders of muscle (including muscle weakness, rhabdomyolysis, muscle spasm)
M62.9	Disorder of muscle, unspecified

Appendix 15 Search strategy for the literature on risk factors of muscle disorders

Database: MEDLINE/PubMed

Search Queries:

1 "myalgia*"[Title] OR "myopath*"[Title] OR "rhabdomyolysis"[Title] OR "muscle pain*"[Title] OR "muscle weakness*"[Title] OR "muscle stiffness*"[Title] OR "muscle symptom*"[Title] OR "muscle dysfunction*"[Title] OR "muscle disorder*"[Title] OR "muscle disease*"[Title] OR "muscle problem*"[Title]

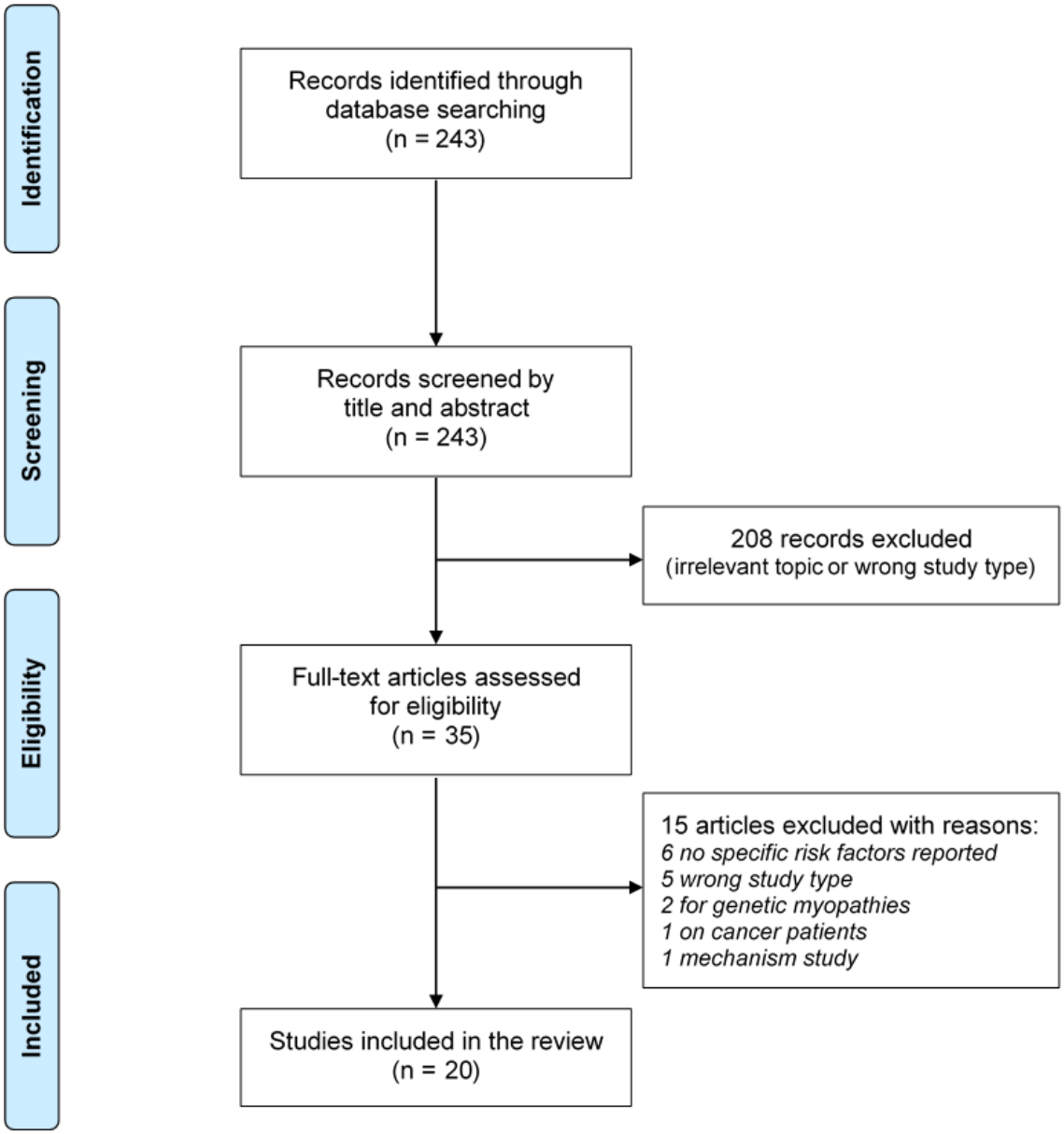
2 "risk factor*"[Title/Abstract] OR "predictor*"[Title/Abstract] OR "predictive factor*"[Title/Abstract] OR "association*"[Title/Abstract]

3 "review"[Title/Abstract] OR "overview"[Title/Abstract] OR "guideline"[Title/Abstract]

4 "0000/00/00"[Date - Publication] : "2020/12/31"[Date - Publication] (from database initiation)

5 1 AND 2 AND 3 AND 4

Appendix 16 Selection of the literature on risk factors of muscle disorders



Appendix 17 Identified risk factors from the included literature

No.	Reference	Reported Risk Factors
1	Guis S, Mattéi JP, Lioté F. Drug-induced and toxic myopathies. Best Pract Res Clin Rheumatol 2003;17(6):877-907.	age, renal diseases, hypothyroidism, diabetes, liver diseases, grapefruit juice, venoms, bacterial toxins, organophosphates, denatured/rerefined rapeseed oil, L-tryptophane contaminant product, aluminium hydroxide-containing vaccines, herbs, nutritional supplements, myotoxic drugs, statin-interactive drugs
2	Saleh FG, Seidman RJ. Drug-induced myopathy and neuropathy. J Clin Neuromuscul Dis 2003;5(2):81-92.	myotoxic drugs
3	Kuncl RW. Agents and mechanisms of toxic myopathy. Curr Opin Neurol 2009;22(5):506-15.	myotoxic drugs
4	Chatzizisis YS, Koskinas KC, Misirli G, et al. Risk factors and drug interactions predisposing to statin-induced myopathy: implications for risk assessment, prevention and treatment. Drug Saf 2010;33(3):171-87.	age, gender, ethnicity, alcohol consumption, exercise, genetic predispositions, frailty, body frame, history of major trauma, history of major surgery, renal diseases, hypothyroidism, diabetes, inherited muscle diseases, liver diseases, serious infections, acquired muscle diseases, grapefruit juice, red rice fungus, statin-interactive drugs
5	Valiyil R, Christopher-Stine L. Drug-related myopathies of which the clinician should be aware. Curr Rheumatol Rep 2010;12(3):213-20.	red rice fungus, myotoxic drugs
6	Abd TT, Jacobson TA. Statin-induced myopathy: a review and update. Expert Opin Drug Saf 2011;10(3):373-87.	age, gender, alcohol consumption, exercise, genetic predispositions, BMI, frailty, body frame, family history of statin intolerance, previous myopathy, history of major surgery, renal diseases, hypothyroidism, inherited muscle diseases, grapefruit juice, statin-interactive drugs
7	Feng Q, Wilke RA, Baye TM. Individualized risk for statin-induced myopathy: current knowledge, emerging	age, gender, ethnicity, alcohol consumption, exercise, family history of statin intolerance, previous myopathy, history of major trauma, previous CK elevation,

	challenges and potential solutions. Pharmacogenomics 2012;13(5):579-94.	previous unexplained cramp, renal diseases, hypothyroidism, diabetes, inherited muscle diseases, serious infections, malignant hyperthermia, hyperuricaemia, statin-interactive drugs
8	Mohassel P, Mammen AL. The spectrum of statin myopathy. Curr Opin Rheumatol 2013;25(6):747-52.	gender, alcohol consumption, genetic predispositions, BMI, family history of statin intolerance, previous myopathy, renal diseases, hypothyroidism, liver diseases, statin-interactive drugs
9	Ungprasert P, Bethina NK, Jones CH. Malignancy and idiopathic inflammatory myopathies. N Am J Med Sci 2013;5(10):569-72.	cancers
10	Keen HI, Krishnarajah J, Bates TR, et al. Statin myopathy: the fly in the ointment for the prevention of cardiovascular disease in the 21st century? Expert Opin Drug Saf 2014;13(9):1227-39.	age, gender, ethnicity, alcohol consumption, genetic predispositions, BMI, frailty, family history of statin intolerance, renal diseases, hypothyroidism, diabetes, inherited muscle diseases, liver diseases, vitamin D deficiency, statin-interactive drugs
11	Degens H, Gayan-Ramirez G, van Hees HW. Smoking-induced skeletal muscle dysfunction: from evidence to mechanisms. Am J Respir Crit Care Med 2015;191(6):620-5.	smoking
12	Chavez LO, Leon M, Einav S, et al. Beyond muscle destruction: a systematic review of rhabdomyolysis for clinical practice. Crit Care 2016;20(1):135.	alcohol consumption, exercise, history of major trauma, hypothyroidism, diabetes, inherited muscle diseases, serious infections, malignant hyperthermia, electrolyte imbalance, hypertension, sickle-cell disease, venoms, myotoxic drugs
13	Chiu N, Chiu L, Chow R, et al. Taxane-induced arthralgia and myalgia: A literature review. J Oncol Pharm Pract 2017;23(1):56-67.	myotoxic drugs

14	Colmenares EW, Pappas AL. Proton Pump Inhibitors: Risk for Myopathy? Ann Pharmacother 2017;51(1):66-71.	myotoxic drugs
15	Adler BL, Christopher-Stine L. Triggers of inflammatory myopathy: insights into pathogenesis. Discov Med 2018;25(136):75-83.	serious infections, cancers
16	Nguyen KA, Li L, Lu D, et al. A comprehensive review and meta-analysis of risk factors for statin-induced myopathy. Eur J Clin Pharmacol 2018;74(9):1099-109.	age, gender, genetic predispositions, previous CK elevation, renal diseases, diabetes, vitamin D deficiency, CVD, hypertension
17	Fonseca J, Nellessen AG, Pitta F. Muscle Dysfunction in Smokers and Patients With Mild COPD: A SYSTEMATIC REVIEW. J Cardiopulm Rehabil Prev 2019;39(4):241-52.	smoking, COPD
18	Lowe K, Kubra KT, He ZY, et al. Vitamin D Supplementation to Treat Statin-Associated Muscle Symptoms: A Review. Sr Care Pharm 2019;34(4):253-57.	vitamin D deficiency
19	Gupta R, Alcantara R, Popli T, et al. Myopathy Associated With Statins and SGLT2 - A Review of Literature. Curr Probl Cardiol 2020;46(4):100765.	statin-interactive drugs
20	Nikolic D, Banach M, Chianetta R, et al. An overview of statin-induced myopathy and perspectives for the future. Expert Opin Drug Saf 2020;19(5):601-15.	age, gender, alcohol consumption, exercise, genetic predispositions, BMI, renal diseases, hypothyroidism, diabetes, inherited muscle diseases, liver diseases, vitamin D deficiency, malignant hyperthermia, statin-interactive drugs

Appendix 18 Criteria for evidence assessment of identified risk factors

- **Adequacy of the evidence**

For each risk factor, the supporting evidence was regarded as adequate or inadequate according to the number of articles reporting this risk factor:

① Adequate evidence: there were 3 or more selected articles reporting the risk factor;

② Inadequate evidence: less than 3 articles reporting the risk factor.

- **Likelihood of the association**

The likelihood of the association between a risk factor and muscle disorders in each of the articles reporting this risk factor was assessed to be high, moderate, or low according to the following criteria:

① High likelihood: the association was reported with a significant RR or OR >1.5 or a well-established biological mechanism;

② Moderate likelihood: the association was reported with a significant RR or OR <1.5 or a possible biological mechanism;

③ Low likelihood: the association was reported without any statistical measurement or supporting mechanism (with empirical explanation only).

The overall likelihood of each risk factor was then determined by the highest level of its likelihood assessed in the individual relevant articles, to avoid omitting any potential predictors of muscle disorders.

- **Overall strength of evidence**

The overall strength of evidence supporting each risk factor was graded as strong, moderate, or weak, based on the evidence adequacy and association likelihood:

① Strong evidence: an association of high likelihood with adequate evidence;

② Moderate evidence: an association of high likelihood with inadequate evidence, or an association of moderate likelihood with adequate evidence;

③ Weak evidence: an association of moderate likelihood with inadequate evidence, or an association of low likelihood.

Appendix 19 Evidence assessment of identified risk factors

Category	Risk Factor	No. of References	Evidence Adequacy	No. of High-likelihood Association	No. of Moderate-likelihood Association	No. of Low-likelihood Association	Overall Likelihood of Association	Overall Strength of Evidence
Demographics	Age	7	Adequate	2	1	4	High	Strong
	Gender	7	Adequate	2	0	5	High	Strong
	Ethnicity	3	Adequate	2	0	1	High	Strong
Health Behaviour/ Status	Alcohol consumption	7	Adequate	0	2	5	Moderate	Moderate
	Smoking	2	Inadequate	0	1	1	Moderate	Weak
	Exercise	5	Adequate	0	0	5	Low	Weak
	BMI	4	Adequate	0	0	4	Low	Weak
	Body frame	2	Inadequate	0	0	2	Low	Weak
	Frailty	3	Adequate	0	0	3	Low	Weak
	Genetic predispositions	6	Adequate	3	0	3	High	Strong
Comorbidities/ Medical History	Family history of statin intolerance	4	Adequate	2	0	2	High	Strong
	History of major trauma	3	Adequate	0	1	2	Moderate	Moderate
	History of major surgery	2	Inadequate	0	0	2	Low	Weak
	Previous muscle problems	5	Adequate	3	0	2	High	Strong
	Previous CK elevation	2	Inadequate	2	0	0	High	Moderate

	Inherited muscle diseases	6	Adequate	2	0	4	High	Strong
	CVD	2	Inadequate	1	0	1	High	Moderate
	Hypertension	1	Inadequate	1	0	0	High	Moderate
	Diabetes	7	Adequate	1	1	5	High	Strong
	Liver diseases	5	Adequate	0	1	4	Moderate	Moderate
	Renal diseases	8	Adequate	2	1	5	High	Strong
	COPD	1	Inadequate	0	0	1	Low	Weak
	Hypothyroidism	8	Adequate	2	0	6	High	Strong
	Hyperuricaemia	1	Inadequate	0	1	0	Moderate	Weak
	Vitamin D deficiency	4	Adequate	0	2	2	Moderate	Moderate
	Serious infections	4	Adequate	0	1	3	Moderate	Moderate
	Electrolyte imbalance	1	Inadequate	0	0	1	Low	Weak
	Malignant hyperthermia	3	Adequate	0	0	3	Low	Weak
	Sickle-cell disease	1	Inadequate	0	0	1	Low	Weak
	Cancers	2	Inadequate	0	0	2	Low	Weak
Medications	Statin-interactive drugs	8	Adequate	6	2	0	High	Strong
	Myotoxic drugs	7	Adequate	4	3	0	High	Strong
Food/ Chemicals/ Environment	Grapefruit juice	3	Adequate	0	1	2	Moderate	Moderate
	Red rice fungus	2	Inadequate	0	1	1	Moderate	Weak
	Denatured/refined rapeseed oil	1	Inadequate	0	0	1	Low	Weak
	Nutritional supplements	1	Inadequate	0	0	1	Low	Weak

	Herbs	1	Inadequate	0	0	1	Low	Weak
	Organophosphates	1	Inadequate	0	0	1	Low	Weak
	L-tryptophane contaminant product	1	Inadequate	0	0	1	Low	Weak
	Aluminium hydroxide-containing vaccines	1	Inadequate	0	0	1	Low	Weak
	Bacterial toxins	1	Inadequate	0	0	1	Low	Weak
	Venoms	2	Inadequate	0	0	2	Low	Weak

Appendix 20 Criteria for clinical evaluation of identified risk factors and specific medication

- **For identified risk factors**

1. Clinical importance

The clinical importance of each identified risk factor was evaluated by two GPs as high, moderate, or low. Discrepancies between the two GPs were resolved through discussion among the GPs and the author to reach a consensus. Additional risk factors suggested by the GPs and considered with high clinical importance were added to the list of potential risk factors.

2. Population/Clinical prevalence of comorbidities

For each of the identified and suggested comorbidities, their prevalence in the population was sought from published studies, national surveys, and disease registries. A disease with a population prevalence below 0.05% was regarded as a rare condition. Where the population prevalence data were unavailable, the clinical prevalence of the comorbidities in primary care was judged by the GPs to determine if they were uncommon conditions.

- **For specific medications**

1. Severity and likelihood of interaction with statins

A list of specific statin-interactive drugs were identified from the British National Formulary (BNF) clinical guidance, including the particular types of statins that each drug interfere with and the severity of each interaction. The overall likelihood of interaction with statins of each drug was determined by the number of statin types that the drug has potential severe interaction with:

① High likelihood: potential severe interaction with 3 or more types of statins;

② Moderate likelihood: potential severe interaction with 1 or 2 types of statins;

③ Low likelihood: no severe interaction with any statin types.

2. Likelihood of association with muscle disorders

Specific drugs with myotoxicity were identified from the selected articles in the literature review. The likelihood of the association of each drug with muscle disorders was assessed to be high, moderate, or low by a pharmacist and the author, based on the evidence from the literature and clinical opinions from the pharmacist. Discrepancies were resolved through discussion between the pharmacist and the author.

3. Clinical usage in the UK

For the drugs with high likelihood of interaction with statins and those with high likelihood of association with muscle disorders, their number of annual prescriptions was sought by searching the OpenPrescribing platform. Drugs with annual prescriptions less than 10000 items were regarded as uncommonly used in the UK. Where the prescription data were unavailable, the clinical usage of the drugs in primary care was judged by the pharmacist to determine if they were not used or rarely used in primary care in the UK.

Appendix 21 Clinical evaluation of identified risk factor and specific medications

A. Clinical evaluation of identified risk factors

Category	Risk Factor	Clinical Importance (GP1)	Clinical Importance (GP2)	Clinical Importance (Consensus)	Population Prevalence Data	Population Prevalence Level	Clinical Prevalence Level
Demographics	Age	High	High	High	NA	NA	NA
	Gender	High	Low	Moderate	NA	NA	NA
	Ethnicity	Low	Low	Low	NA	NA	NA
Health Behaviour/ Status	Alcohol consumption	Moderate	Moderate	Moderate	NA	NA	NA
	Smoking	High	Moderate	High	NA	NA	NA
	Exercise	Low	Moderate	Moderate	NA	NA	NA
	BMI	Moderate	Low	Moderate	NA	NA	NA
	Body frame	Moderate	Low	Moderate	NA	NA	NA
	Frailty	Moderate	High	High	NA	NA	NA
	Genetic predispositions	Moderate	Moderate	Moderate	NA	NA	NA
Comorbidities/ Medical History	Family history of statin intolerance	Moderate	High	High	Unavailable	Unknown	Uncommon
	History of major trauma	Low	High	Moderate	Unavailable	Unknown	Uncommon

	History of major surgery	Low	Moderate	Moderate	Unavailable	Unknown	Uncommon
	Previous muscle problems	High	High	High	Unavailable	Unknown	Moderate
	Previous CK elevation	Moderate	High	High	Unavailable	Unknown	Uncommon
	Inherited muscle diseases	Moderate	High	High	<0.05% [1]	Rare	\
	CVD	Low	Moderate	Moderate	3% [2]	Moderate	\
	Hypertension	Low	Low	Low	14% [2]	Common	\
	Diabetes	Moderate	Low	Moderate	7% [2]	Moderate	\
	Liver diseases	Moderate	Moderate	Moderate	20% [3]	Common	\
	Renal diseases	Moderate	Moderate	Moderate	4% [2]	Moderate	\
	COPD	Low	High	Moderate	2% [2]	Moderate	\
	Hypothyroidism	Moderate	Low	Moderate	5% [4]	Moderate	\
	Hyperuricaemia	Low	High	Moderate	2.5% (gout) [5]	Moderate	\
	Vitamin D deficiency	Moderate	High	High	4%-29% (varying by season) [6]	Common	\
	Serious infections	Low	Low	Low	Unavailable	Unknown	Uncommon
	Electrolyte imbalance	Low	Low	Low	Unavailable	Unknown	Uncommon
	Malignant hyperthermia	Low	Moderate	Moderate	<0.05% [1]	Rare	\

	Sickle-cell disease	Low	High	Moderate	<0.05% [1]	Rare	\
	Cancers	Low	Moderate	Moderate	Varying by specific type	Unknown	Uncommon
Medications	Statin-interactive drugs	High	High	High	NA	NA	NA
	Myotoxic drugs	High	High	High	NA	NA	NA
Food/ Chemicals/ Environment	Grapefruit juice	Low	Low	Low	NA	NA	NA
	Red rice fungus	Low	Low	Low	NA	NA	NA
	Denatured/refined rapeseed oil	Low	Low	Low	NA	NA	NA
	Nutritional supplements	Moderate	Low	Moderate	NA	NA	NA
	Herbs	Moderate	Low	Moderate	NA	NA	NA
	Organophosphates	Low	Low	Low	NA	NA	NA
	L-tryptophane contaminant products	Low	Low	Low	NA	NA	NA
	Aluminium hydroxide-containing vaccines	Low	Low	Low	NA	NA	NA
	Bacterial toxins	Low	Moderate	Moderate	NA	NA	NA
	Venoms	Low	Low	Low	NA	NA	NA

Additional Suggested Risk Factors	Rheumatic arthritis (RA)	High	High	High	Unavailable	Unknown	Moderate
	Degenerative joint disorders	High	High	High	Unavailable	Unknown	Common
	Vitamin B12 deficiency	High	Moderate	Moderate	Unavailable	Unknown	Moderate
	Deprivation	Moderate	Moderate	Moderate	NA	NA	NA

1 National Organization for Rare Disorders database (https://rarediseases.org/for-patients-and-families/information-resources/rare-disease-information)

2 Quality and Outcomes Framework 2019-2020 (https://app.powerbi.com/view?r=eyJrIjoiMDZiMmI2MzEtMWVjZC00YTVlLWI5NjEtMTNkODM3M2M0NDk3IiwidCI6IjUwZjYwNzFmLWJiZmUtNDAxYS04ODAzLTY3Mzc0OGU2MjllMiIsImMiOjh9)

3 British Liver Trust (https://britishlivertrust.org.uk/information-and-support/living-with-a-liver-condition/liver-conditions/non-alcohol-related-fatty-liver-disease/)

4 National Institute for Health and Care Excellence (NICE) Guideline [NG145] (https://www.nice.org.uk/guidance/ng145/chapter/Context)

5 Kuo C, Grainge MJ, Mallen C, et alRising burden of gout in the UK but continuing suboptimal management: a nationwide population studyAnnals of the Rheumatic Diseases 2015;74:661-667.

6 National Diet and Nutrition Survey (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/772434/NDNS_UK_Y1-9_report.pdf)

B. Clinical evaluation of specific statin-interactive drugs

Interactive Drug	Atorva-statin	Fluva-statin	Prava-statin	Rosuva-statin	Simva-statin	No. of Severe Interactions	Interaction Likelihood	No. of Annual Prescriptions	Clinical Usage (Prescription Data)	Clinical Usage (Pharmacist Evaluation)
Amiodarone	Moderate	Severe	\	\	Severe	2	Moderate	\	\	\
Aprepitant	Severe	\	\	\	\	1	Moderate	\	\	\
Amlodipine	\	\	\	\	Mild	0	Low	\	\	\
Antacids	\	\	\	Moderate	\	0	Low	\	\	\
Apalutamide	\	\	\	Mild	Moderate	0	Low	\	\	\
Aprepitant	\	\	\	\	Severe	1	Moderate	\	\	\
Atazanavir	Severe	\	\	Severe	Severe	3	High	Unavailable	Unknown	Rare use in primary care
Bempedoic acid	\	\	\	\	Moderate	0	Low	\	\	\
Bezafibrate	Severe	Severe	Severe	Severe	Severe	5	High	446102	Common	\
Bosentan	Mild	\	\	\	Moderate	0	Low	\	\	\
Carbamazepine	Moderate	\	\	\	Severe	1	Moderate	\	\	\
Ceftobiprole	Moderate	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Ciclosporin	Severe	Severe	Severe	Severe	Severe	5	High	78252	Common	\
Ciprofibrate	Severe	Severe	Severe	Severe	Severe	5	High	29116	Common	\
Clarithromycin	Severe	\	Severe	\	Severe	3	High	2152224	Common	\

Clopidogrel	\	\	\	Moderate	\	0	Low	\	\	\
Cobicistat	Severe	\	\	\	Severe	2	Moderate	\	\	\
Colchicine	Severe	Severe	Severe	Severe	Severe	5	High	473565	Common	\
Crizotinib	Severe	\	Moderate	\	Severe	2	Moderate	\	\	\
Dabrafenib	Mild	\	\	\	Moderate	0	Low	\	\	\
Danazol	Severe	\	\	\	Severe	2	Moderate	\	\	\
Daptomycin	Severe	Severe	Severe	Severe	Severe	5	High	34	Uncommon	\
Darolutamide	Severe	Severe	Moderate	Severe	Moderate	3	High	Unavailable	Unknown	Rare use in primary care
Darunavir	Severe	\	\	Severe	Severe	3	High	52	Uncommon	\
Dasabuvir	\	\	\	Moderate	\	0	Low	\	\	\
Dasatinib	\	\	\	\	Moderate	0	Low	\	\	\
Diltiazem	Severe	\	\	\	Severe	2	Moderate	\	\	\
Dronedarone	Severe	\	\	Severe	Severe	3	High	23377	Common	\
Efavirenz	Mild	\	\	\	Moderate	0	Low	\	\	\
Elbasvir	Moderate	Moderate	\	Moderate	Moderate	0	Low	\	\	\
Elexacaftor	Moderate	\	Moderate	Moderate	Moderate	0	Low	\	\	\
Eltrombopag	Moderate	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Enzalutamide	\	\	\	\	Severe	1	Moderate	\	\	\
Erythromycin	Severe	\	Severe	\	Severe	3	High	557568	Common	\

Eslicarbazepine	Moderate	\	\	\	Moderate	0	Low	\	\	\
Ezetimibe	Severe	Severe	Severe	Severe	Severe	5	High	1788039	Common	\
Fenofibrate	Severe	Severe	Severe	Severe	Severe	5	High	557325	Common	\
Filgotinib	Moderate	\	Moderate	Moderate	Moderate	0	Low	\	\	\
Fluconazole	Severe	Severe	\	\	Severe	3	High	552789	Common	\
Fosamprenavir	Severe	\	\	Severe	Severe	3	High	Unavailable	Unknown	Rare use in primary care
Fosphenytoin	Moderate	\	\	\	Moderate	0	Low	\	\	\
Fostamatinib	\	\	\	Moderate	Moderate	0	Low	\	\	\
Fusidate	Severe	Severe	Severe	Severe	Severe	5	High	45855	Common	\
Gemfibrozil	Severe	Severe	Severe	Severe	Severe	5	High	10303	Common	\
Glecaprevir + Pibrentasvir	Severe	Moderate	Moderate	Moderate	Moderate	1	Moderate	\	\	\
Grapefruit	Mild	\	\	\	Severe	1	Moderate	\	\	\
Grazoprevir	Moderate	Moderate	\	Moderate	Moderate	0	Low	\	\	\
Idelalisib	Severe	\	\	\	Severe	2	Moderate	\	\	\
Imatinib	Severe	\	\	\	Severe	2	Moderate	\	\	\
Isavuconazole	Moderate	Moderate	\	Moderate	Severe	1	Moderate	\	\	\
Itraconazole	Moderate	\	\	\	Severe	1	Moderate	\	\	\
Ketoconazole	Moderate	\	\	\	Severe	1	Moderate	\	\	\

Ledipasvir	Mild	Moderate	Moderate	Severe	Moderate	1	Moderate	\	\	\
Leflunomide	Moderate	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Letermovir	Severe	Moderate	Moderate	Severe	Severe	3	High	Unavailable	Unknown	Rare use in primary care
Lomitapide	Mild	\	\	\	Moderate	0	Low	\	\	\
Lopinavir-Ritonavir	Severe	\	\	Severe	Severe	3	High	7	Uncommon	\
Miconazole	Severe	Severe	\	\	Severe	3	High	605936	Common	\
Mitotane	\	\	\	\	Severe	1	Moderate	\	\	\
Netupitant	Severe	\	\	\	Severe	2	Moderate	\	\	\
Nevirapine	Mild	\	\	\	Moderate	0	Low	\	\	\
Nicotinic acid	Severe	Severe	Severe	Severe	Severe	5	High	16	Uncommon	\
Nilotinib	Severe	\	\	\	Severe	2	Moderate	\	\	\
Osimertinib	\	\	\	Moderate	\	0	Low	\	\	\
Paritaprevir	Severe	\	\	\	Severe	2	Moderate	\	\	\
Paritaprevir + Ritonavir + Ombitasvir	\	Moderate	Moderate	Moderate	\	0	Low	\	\	\
Pazopanib	Moderate	\	Moderate	Moderate	Moderate	0	Low	\	\	\
Phenytoin	Moderate	\	\	\	Moderate	0	Low	\	\	\
Phenindione	\	\	\	Severe	\	1	Moderate	\	\	\

Pibrentasvir + Glecaprevir	\	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Posaconazole	Severe	\	\	\	Severe	2	Moderate	\	\	\
Ranolazine	Moderate	\	\	\	Moderate	0	Low	\	\	\
Regorafenib	Moderate	Moderate	\	Moderate	\	0	Low	\	\	\
Ribociclib	\	\	Moderate	Moderate	Moderate	0	Low	\	\	\
Rifampicin	Moderate	Moderate	\	\	Moderate	0	Low	\	\	\
Ritonavir	Severe	\	\	Severe	Severe	3	High	55	Uncommon	\
Sacubitril-Valsartan	Severe	Severe	Severe	Severe	Severe	5	High	140746	Common	\
Saquinavir	Severe	\	\	Severe	Severe	3	High	6	Uncommon	\
Sarilumab	Moderate	\	\	\	Moderate	0	Low	\	\	\
St John's wort	Mild	\	\	\	Moderate	0	Low	\	\	\
Tedizolid	Moderate	Moderate	\	Moderate	\	0	Low	\	\	\
Teriflunomide	Moderate	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Ticagrelor	\	\	\	\	Moderate	0	Low	\	\	\
Tipranavir	Severe	\	\	Severe	Severe	3	High	Unavailable	Unknown	Rare use in primary care
Tivozanib	\	\	\	Moderate	\	0	Low	\	\	\
Tocilizumab	Moderate	\	\	\	Moderate	0	Low	\	\	\
Velpatasvir	\	\	\	Severe	Severe	2	Moderate	\	\	\

Venetoclax	Moderate	Moderate	Moderate	Moderate	Moderate	0	Low	\	\	\
Verapamil	Severe	\	\	\	Severe	2	Moderate	\	\	\
Voriconazole	Severe	\	\	\	Severe	2	Moderate	\	\	\
Voxilaprevir + Sofosbuvir + Velpatasvir	Moderate	Moderate	Moderate	Severe	Moderate	1	Moderate	\	\	\

C. Clinical evaluation of specific myotoxic drugs

Myotoxic Drugs	Likelihood of Association (Assessor 1)	Likelihood of Association (Assessor 2)	Likelihood of Association (Consensus)	No. of Annual Prescriptions	Clinical Usage (Prescription Data)	Clinical Usage (Pharmacist Evaluation)
Gemfibrozil	Moderate	Moderate	Moderate	10303	Common	\
Bezafibrate	Moderate	Moderate	Moderate	446102	Common	\
Fenofibrate	Moderate	Moderate	Moderate	557325	Common	\
Ciprofibrate	Moderate	Moderate	Moderate	29116	Common	\
Clofibrate	Moderate	Moderate	Moderate	Unavailable	Unknown	Not marketed in the UK
Niacin	Low	Low	Low	\	\	\
Ezetimibe	Low	Low	Low	\	\	\
Metoprolol	Low	Low	Low	\	\	\
Minoxidil	Low	Low	Low	\	\	\
Enalapril	Low	Low	Low	\	\	\
Perhexiline	Low	High	Moderate	Unavailable	Unknown	Not marketed in the UK
Penicillin	Low	Low	Low	\	\	\
Nifuroxazide	Low	Low	Low	\	\	\
Sulfonamides (no specified drugs)	Low	Low	Low	\	\	\
Fluoroquinolones	Low	Low	Low	\	\	\

Voriconazole	Low	Moderate	Moderate	622	Uncommon	\
Chloroquine	High	Low	Moderate	1908	Uncommon	\
Hydroxychloroquine	Moderate	Low	Moderate	1160249	Common	\
Zidovudine	High	High	High	6	Uncommon	\
Clevudine	Low	High	Moderate	Unavailable	Unknown	Not marketed in the UK
Interferons (no specified drugs)	Low	Moderate	Moderate	Unavailable	Unknown	No/Rare use in primary care
Ciclosporin	High	Low	Low	\	\	\
Tacrolimus	Moderate	Low	Low	\	\	\
Adalimumab	Low	Low	Low	\	\	\
Infliximab	Low	Low	Low	\	\	\
Leflunomide	Moderate	Moderate	Moderate	179744	Common	\
Sulfasalazine	Low	Moderate	Low	\	\	\
Azathioprine	Low	Moderate	Low	\	\	\
Sodium Aurothiomalate	Low	High	Moderate	2975	Uncommon	\
Mycophenolate Mofetil	Low	Low	Low	\	\	\
Immune Checkpoint Inhibitors (no specified drugs)	Low	Moderate	Moderate	Unavailable	Unknown	No/Rare use in primary care

Colchicine	High	High	High	473565	Common	\
Diclofenac	Low	Moderate	Low	\	\	\
Phenylbutazone	Low	Low	Low	\	\	\
Ibuprofen	Low	Moderate	Low	\	\	\
Triamcinolone	High	High	High	282870	Common	\
Betamethasone	High	High	High	3875186	Common	\
Dexametasone	High	High	High	2144697	Common	\
Niflumic Acid	Low	Low	Low	\	\	\
Diphenhydramine	Low	Low	Low	\	\	\
Omeprazole	Moderate	Moderate	Moderate	31732881	Common	\
Esomeprazole	Moderate	Moderate	Moderate	1932194	Common	\
Lansoprazole	Moderate	Moderate	Moderate	26016659	Common	\
Rabeprazole	Moderate	Moderate	Moderate	363374	Common	\
Phenelzine	Low	Low	Low	\	\	\
Diazepam	Low	Moderate	Moderate	4942895	Common	\
Neuroleptics (no specified drugs)	Low	Moderate	Low	\	\	\
Chlorpromazine	Low	Low	Low	\	\	\
Carbimazole	Low	Low	Low	\	\	\
Propylthiouracil	Low	Low	Low	\	\	\

Penicillamine	Moderate	Moderate	Moderate	9279	Uncommon	\
Tiopronin	Low	Moderate	Moderate	52	Uncommon	\
Pamidronate	Low	High	Low	\	\	\
Zoledronate	Low	High	Low	\	\	\
Paracetamol	Low	Low	Low	\	\	\
Topical Menthol	Low	Low	Low	\	\	\
Opioids (Morphine)	Low	Moderate	Low	\	\	\
Anesthetics (no specified drugs)	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Diuretics (no specified drugs)	Low	Moderate	Low	\	\	\
Vasopressin	Low	Low	Low	\	\	\
Emetine	Moderate	High	High	Unavailable	Unknown	Not marketed in the UK
Etretinate	Low	Moderate	Low	\	\	\
Vincristine	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Methotrexate	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Cyclophosphamide	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Cytarabine	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Docetaxel	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care
Paclitaxel	Low	High	Moderate	Unavailable	Unknown	No/Rare use in primary care

Cocaine	Low	Moderate	Low	\	\	\
Lysergic Acid Diethylamide (LSD)	Low	Low	Low	\	\	\
Ecstasy	Low	Moderate	Low	\	\	\
Heroin	Low	Moderate	Low	\	\	\

Appendix 22 CPRD data quality assessment for identified and suggested risk factors

Category	Risk Factor	Data Quality	Reason
Demographics	Age	Acceptable	
	Gender	Acceptable	
	Ethnicity	Acceptable	
Health Behaviour/Status	Alcohol consumption	Acceptable	
	Smoking	Acceptable	
	Exercise	Unacceptable	Not well recorded in primary care
	BMI	Acceptable	
	Body frame	Unacceptable	Unavailable in CPRD data
	Frailty	Acceptable	
	Genetic predispositions	Unacceptable	Unavailable in CPRD data
Comorbidities/Medical History	Family history of statin intolerance	Unacceptable	Not well recorded in primary care
	History of major trauma	Unacceptable	Not well recorded in primary care
	History of major surgery	Unacceptable	Not well recorded in primary care
	Previous muscle problems	Acceptable	
	Previous CK elevation	Unacceptable	Not well recorded in primary care
	Inherited muscle diseases	Unacceptable	Not well recorded in primary care
	CVD	Acceptable	

	Hypertension	Acceptable	
	Diabetes	Acceptable	
	Liver diseases	Acceptable	
	Renal diseases	Acceptable	
	Hypothyroidism	Acceptable	
	Hyperuricaemia	Acceptable	
	Vitamin D deficiency	Acceptable	
	COPD	Acceptable	
	Serious infections	Acceptable	
	Electrolyte imbalance	Unacceptable	Not well recorded in primary care
	Malignant hyperthemia	Unacceptable	Not well recorded in primary care
	Sickle-cell disease	Unacceptable	Not well recorded in primary care
	Cancers	Acceptable	
Medications	Statin-interactive drugs	Acceptable	
	Myotoxic drugs	Acceptable	
Food/Chemicals/Environment	Grapefruit juice	Unacceptable	Unavailable in CPRD data
	Red rice fungus	Unacceptable	Unavailable in CPRD data
	Denatured/refined rapeseed oil	Unacceptable	Unavailable in CPRD data
	Nutritional supplements	Unacceptable	Unavailable in CPRD data

	Herbs	Unacceptable	Unavailable in CPRD data
	Organophosphates	Unacceptable	Unavailable in CPRD data
	L-tryptophane contaminant products	Unacceptable	Unavailable in CPRD data
	Aluminium hydroxide-containing vaccines	Unacceptable	Unavailable in CPRD data
	Bacterial toxins	Unacceptable	Unavailable in CPRD data
	Venoms	Unacceptable	Unavailable in CPRD data
Additional Suggested Risk Factors	RA	Acceptable	
	Degenerative joint disorders	Acceptable	
	Vitamin B12 deficiency	Acceptable	
	Deprivation	Acceptable	

Appendix 23 Criteria for final selection of candidate risk factors and specific medications

- **For candidate risk factors, the final selection followed the process below:**

1. From all the risk factors identified in the literature review or suggested by the GPs, those with unacceptable data quality in CPRD were first excluded;
2. Comorbidities that are rare or uncommon conditions in primary care were further excluded;
3. Among the remaining factors, those with strong or moderate strength of evidence on their associations with muscle disorders or those with high or moderate clinical importance were finally selected.

- **For specific medications, the final selection used the criteria below:**

1. Statin-interactive drugs were selected if they had high likelihood of interactions and common clinical usage in the UK;
2. Myotoxic drugs were selected if they had strong or moderate strength of evidence on their associations with muscle disorders and common clinical usage in the UK.

Appendix 24 Final selection of candidate risk factors and specific medications

A. Final selection of candidate risk factors

Category	Risk Factor	Strength of Evidence	Clinical Importance	Population/Clinical Prevalence	Data Quality	Selection
Demographics	Age	Strong	High	NA	Acceptable	Selected
	Gender	Strong	Moderate	NA	Acceptable	Selected
	Ethnicity	Strong	Low	NA	Acceptable	Selected
Health Behaviour/Status	Alcohol consumption	Moderate	Moderate	NA	Acceptable	Selected
	Smoking	Weak	High	NA	Acceptable	Selected
	Exercise	Weak	Moderate	NA	Unacceptable	
	BMI	Weak	Moderate	NA	Acceptable	Selected
	Body frame	Weak	Moderate	NA	Unacceptable	
	Frailty	Weak	High	NA	Acceptable	Selected
	Genetic predispositions	Strong	Moderate	NA	Unacceptable	
Comorbidities/ Medical History	Family history of statin intolerance	Strong	High	Uncommon	Unacceptable	
	History of major trauma	Moderate	Moderate	Uncommon	Unacceptable	
	History of major surgery	Weak	Moderate	Uncommon	Unacceptable	
	Previous muscle problems	Strong	High	Moderate	Acceptable	Selected
	Previous CK elevation	Moderate	High	Uncommon	Unacceptable	

	Inherited muscle diseases	Strong	High	Rare	Unacceptable	
	CVD	Moderate	Moderate	Moderate	Acceptable	Selected
	Hypertension	Moderate	Low	Common	Acceptable	Selected
	Diabetes	Strong	Moderate	Moderate	Acceptable	Selected
	Liver diseases	Moderate	Moderate	Common	Acceptable	Selected
	Renal diseases	Strong	Moderate	Moderate	Acceptable	Selected
	COPD	Weak	Moderate	Moderate	Acceptable	Selected
	Hypothyroidism	Strong	Moderate	Moderate	Acceptable	Selected
	Hyperuricaemia	Weak	Moderate	Moderate	Acceptable	Selected
	VitaminD deficiency	Moderate	High	Common	Acceptable	Selected
	Serious infections	Moderate	Low	Uncommon	Acceptable	
	Electrolyte imbalance	Weak	Low	Uncommon	Unacceptable	
	Malignant hyperthemia	Weak	Moderate	Rare	Unacceptable	
	Sickle-cell disease	Weak	Moderate	Rare	Unacceptable	
	Cancers	Weak	Moderate	Uncommon	Acceptable	
Medications	Statin-interactive drugs	Strong	High	NA	Acceptable	Selected
	Myotoxic drugs	Strong	High	NA	Acceptable	Selected
Food/Chemicals/ Environment	Grapefruit juice	Moderate	Low	NA	Unacceptable	
	Red rice fungus	Weak	Low	NA	Unacceptable	

	Denatured/refined rapeseed oil	Weak	Low	NA	Unacceptable	
	Nutritional supplements	Weak	Moderate	NA	Unacceptable	
	Herbs	Weak	Moderate	NA	Unacceptable	
	Organophosphates	Weak	Low	NA	Unacceptable	
	L-tryptophane contaminant product	Weak	Low	NA	Unacceptable	
	Aluminium hydroxide-containing vaccines	Weak	Low	NA	Unacceptable	
	Bacterial toxins	Weak	Moderate	NA	Unacceptable	
	Venoms	Weak	Low	NA	Unacceptable	
Additional Suggested Risk Factors	RA	NA	High	Moderate	Acceptable	Selected
	Degenerative joint disorders	NA	High	Common	Acceptable	Selected
	Vitamin B12 deficiency	NA	Moderate	Moderate	Acceptable	Selected
	Deprivation	NA	Moderate	NA	Acceptable	Selected

B. Final selection of statin-interactive drugs

Interactive Drug	Likelihood of Interaction	Clinical Usage	Selection
Amiodarone	Moderate	\	
Aprepitant	Moderate	\	
Amlodipine	Low	\	
Antacids	Low	\	
Apalutamide	Low	\	
Aprepitant	Moderate	\	
Atazanavir	High	Rare use in primary care	
Bempedoic acid	Low	\	
Bezafibrate	High	Common	Selected
Bosentan	Low	\	
Carbamazepine	Moderate	\	
Ceftobiprole	Low	\	
Ciclosporin	High	Common	Selected
Ciprofibrate	High	Common	Selected
Clarithromycin	High	Common	Selected
Clopidogrel	Low	\	
Cobicistat	Moderate	\	

Colchicine	High	Common	Selected
Crizotinib	Moderate	\	
Dabrafenib	Low	\	
Danazol	Moderate	\	
Daptomycin	High	Uncommon	
Darolutamide	High	Rare use in primary care	
Darunavir	High	Uncommon	
Dasabuvir	Low	\	
Dasatinib	Low	\	
Diltiazem	Moderate	\	
Dronedarone	High	Common	Selected
Efavirenz	Low	\	
Elbasvir	Low	\	
Elexacaftor	Low	\	
Eltrombopag	Low	\	
Enzalutamide	Moderate	\	
Erythromycin	High	Common	Selected
Eslicarbazepine	Low	\	
Ezetimibe	High	Common	Selected

Fenofibrate	High	Common	Selected
Filgotinib	Low	\	
Fluconazole	High	Common	Selected
Fosamprenavir	High	Rare use in primary care	
Fosphenytoin	Low	\	
Fostamatinib	Low	\	
Fusidate	High	Common	Selected
Gemfibrozil	High	Common	Selected
Glecaprevir+Pibrentasvir	Moderate	\	
Grapefruit	Moderate	\	
Grazoprevir	Low	\	
Idelalisib	Moderate	\	
Imatinib	Moderate	\	
Isavuconazole	Moderate	\	
Itraconazole	Moderate	\	
Ketoconazole	Moderate	\	
Ledipasvir	Moderate	\	
Leflunomide	Low	\	
Letermovir	High	Rare use in primary care	

Lomitapide	Low	\	
Lopinavir-Ritonavir	High	Uncommon	
Miconazole	High	Common	Selected
Mitotane	Moderate	\	
Netupitant	Moderate	\	
Nevirapine	Low	\	
Nicotinic acid	High	Uncommon	
Nilotinib	Moderate	\	
Osimertinib	Low	\	
Paritaprevir	Moderate	\	
Paritaprevir + Ritonavir + Ombitasvir	Low	\	
Pazopanib	Low	\	
Phenytoin	Low	\	
Phenindione	Moderate	\	
Pibrentasvir + Glecaprevir	Low	\	
Posaconazole	Moderate	\	
Ranolazine	Low	\	
Regorafenib	Low	\	
Ribociclib	Low	\	

Rifampicin	Low	\	
Ritonavir	High	Uncommon	
Sacubitril-Valsartan	High	Common	Selected
Saquinavir	High	Uncommon	
Sarilumab	Low	\	
St John's wort	Low	\	
Tedizolid	Low	\	
Teriflunomide	Low	\	
Ticagrelor	Low	\	
Tipranavir	High	Rare use in primary care	
Tivozanib	Low	\	
Tocilizumab	Low	\	
Velpatasvir	Moderate	\	
Venetoclax	Low	\	
Verapamil	Moderate	\	
Voriconazole	Moderate	\	
Voxilaprevir + Sofosbuvir + Velpatasvir	Moderate	\	

C. Final selection of myotoxic drugs

Myotoxic Drug	Likelihood of Association	Clinical Usage	Selection
Gemfibrozil	Moderate	Common	Selected
Bezafibrate	Moderate	Common	Selected
Fenofibrate	Moderate	Common	Selected
Ciprofibrate	Moderate	Common	Selected
Clofibrate	Moderate	Not marketed in the UK	
Niacin	Low	\	
Ezetimibe	Low	\	
Metoprolol	Low	\	
Minoxidil	Low	\	
Enalapril	Low	\	
Perhexiline	Moderate	Not marketed in the UK	
Penicillin	Low	\	
Nifuroxazide	Low	\	
Sulfonamides (no specified drugs)	Low	\	
Fluoroquinolones	Low	\	
Voriconazole	Moderate	Uncommon	
Chloroquine	Moderate	Uncommon	

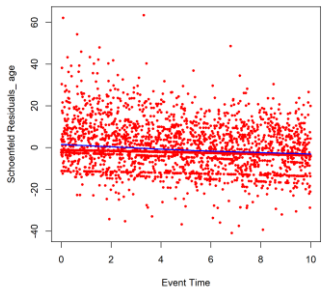
Hydroxychloroquine	Moderate	Common	Selected
Zidovudine	High	Uncommon	
Clevudine	Moderate	Not marketed in the UK	
Interferons (no specified drugs)	Moderate	Rare use in primary care	
Ciclosporin	Low	\	
Tacrolimus	Low	\	
Adalimumab	Low	\	
Infliximab	Low	\	
Leflunomide	Moderate	Common	Selected
Sulfasalazine	Low	\	
Azathioprine	Low	\	
Sodium Aurothiomalate	Moderate	Uncommon	
Mycophenolate Mofetil	Low	\	
Immune Checkpoint Inhibitors (no specified drugs)	Moderate	Rare use in primary care	
Colchicine	High	Common	Selected
Diclofenac	Low	\	
Phenylbutazone	Low	\	
Ibuprofen	Low	\	
Triamcinolone	High	Common	Selected

Betamethasone	High	Common	Selected
Dexametasone	High	Common	Selected
Niflumic Acid	Low	\	
Diphenhydramine	Low	\	
Omeprazole	Moderate	Common	Selected
Esomeprazole	Moderate	Common	Selected
Lansoprazole	Moderate	Common	Selected
Rabeprazole	Moderate	Common	Selected
Phenelzine	Low	\	
Diazepam	Moderate	Common	Selected
Neuroleptics (no specified drugs)	Low	\	
Chlorpromazine	Low	\	
Carbimazole	Low	\	
Propylthiouracil	Low	\	
Penicillamine	Moderate	Uncommon	
Tiopronin	Moderate	Uncommon	
Pamidronate	Low	\	
Zoledronate	Low	\	
Paracetamol	Low	\	

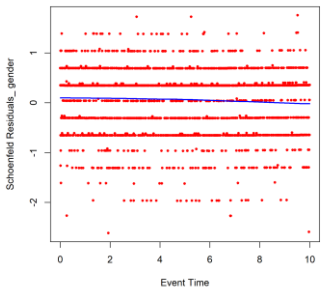
Topical Menthol	Low	\	
Opioids (Morphine)	Low	\	
Anesthetics (no specified drugs)	Moderate	Rare use in primary care	
Diuretics (no specified drugs)	Low	\	
Vasopressin	Low	\	
Emetine	High	Not marketed in the UK	
Etretinate	Low	\	
Vincristine	Moderate	Rare use in primary care	
Methotrexate	Moderate	Rare use in primary care	
Cyclophosphamide	Moderate	Rare use in primary care	
Cytarabine	Moderate	Rare use in primary care	
Docetaxel	Moderate	Rare use in primary care	
Paclitaxel	Moderate	Rare use in primary care	
Cocaine	Low	\	
Lysergic Acid Diethylamide (LSD)	Low	\	
Ecstasy	Low	\	
Heroin	Low	\	

Appendix 25 Schoenfeld residuals* plots for the 24 candidate predictor variables

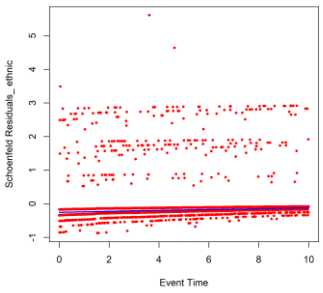
Age



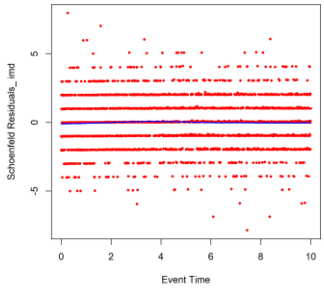
Gender



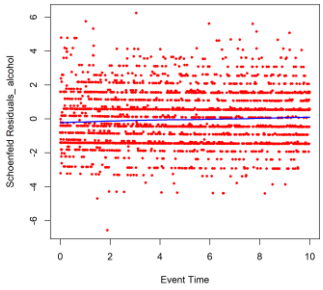
Ethnicity



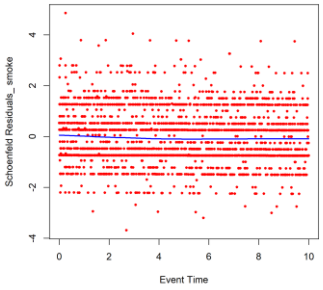
Deprivation



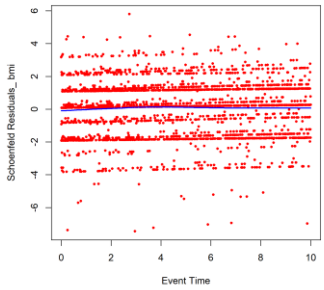
Alcohol consumption



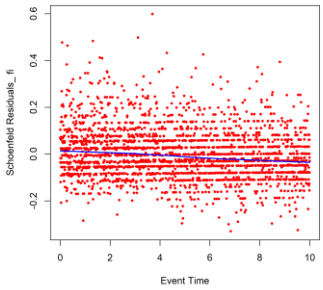
Smoking status



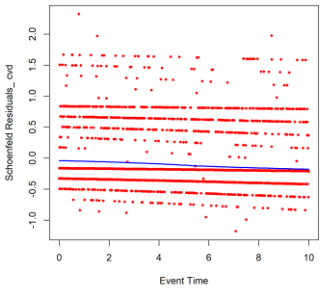
BMI level



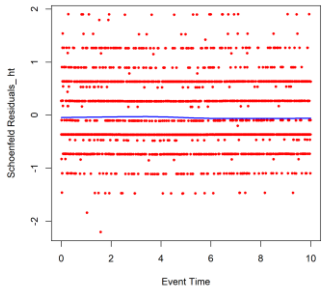
Frailty index



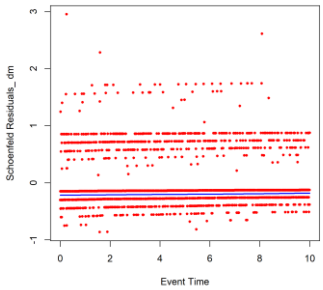
CVD



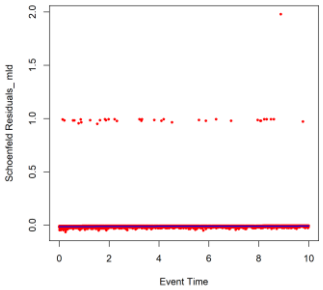
Hypertension



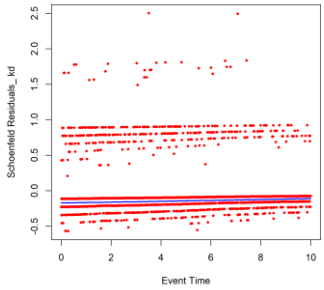
Diabetes



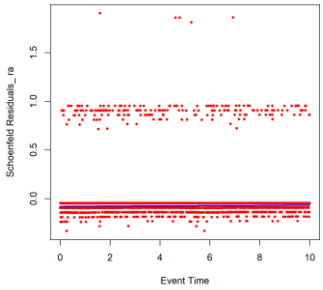
Mild liver disease



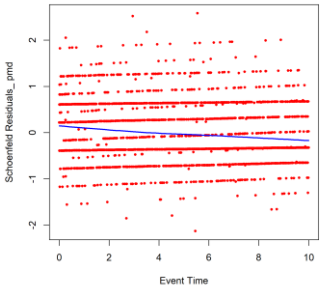
Kidney disease



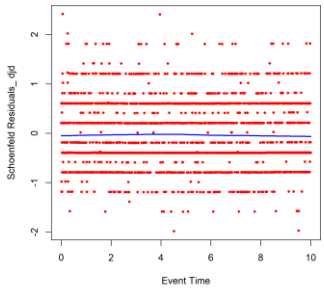
Rheumatic arthritis



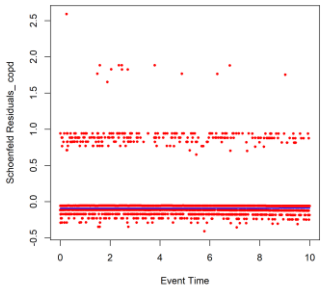
Previous muscle problems



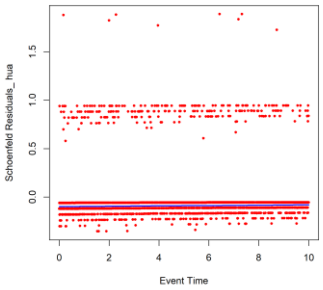
Degenerative joint disorders



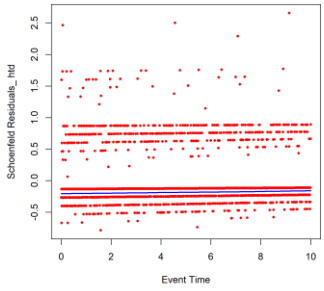
COPD



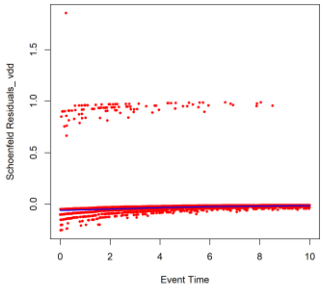
Hyperuricaemia



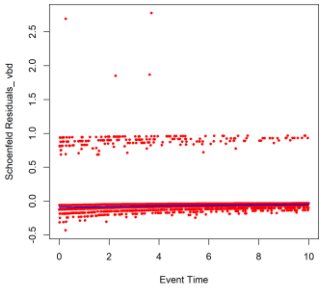
Hypothyroidism



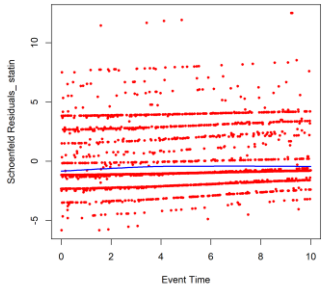
Vitamin D deficiency



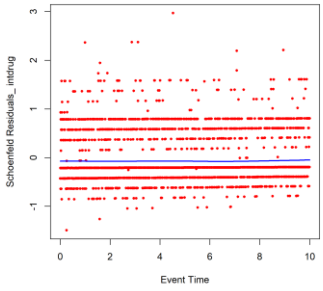
Vitamin B12 deficiency



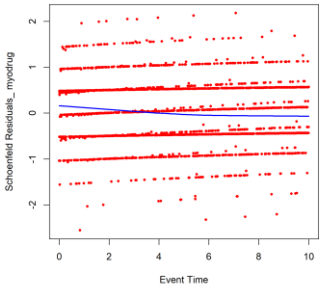
Statins



Statin-interactive drugs



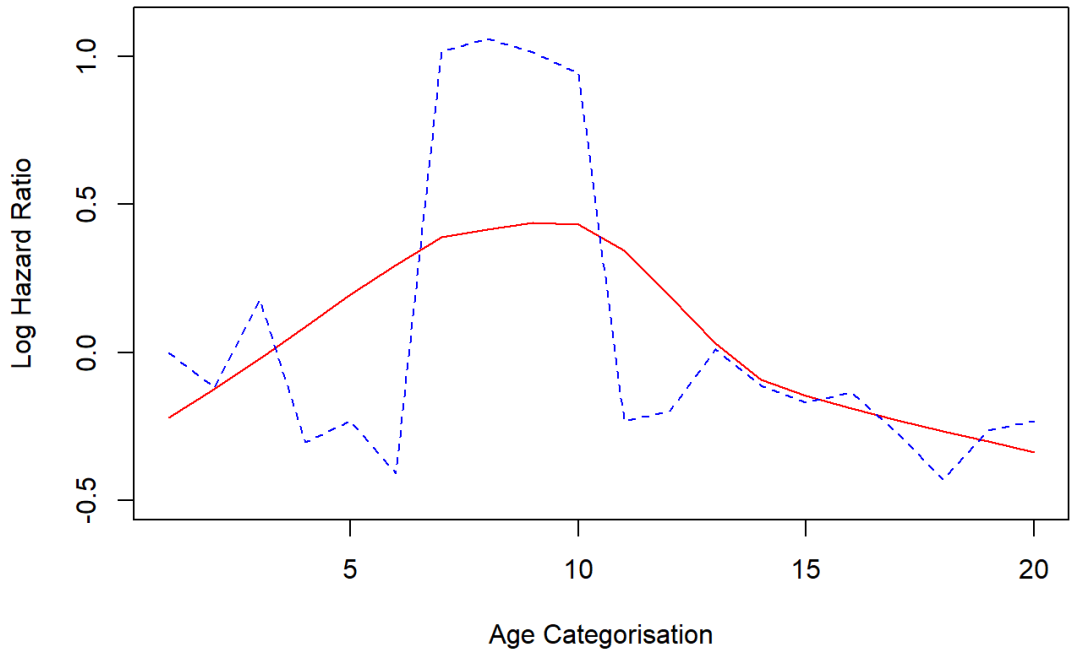
Myotoxic drugs



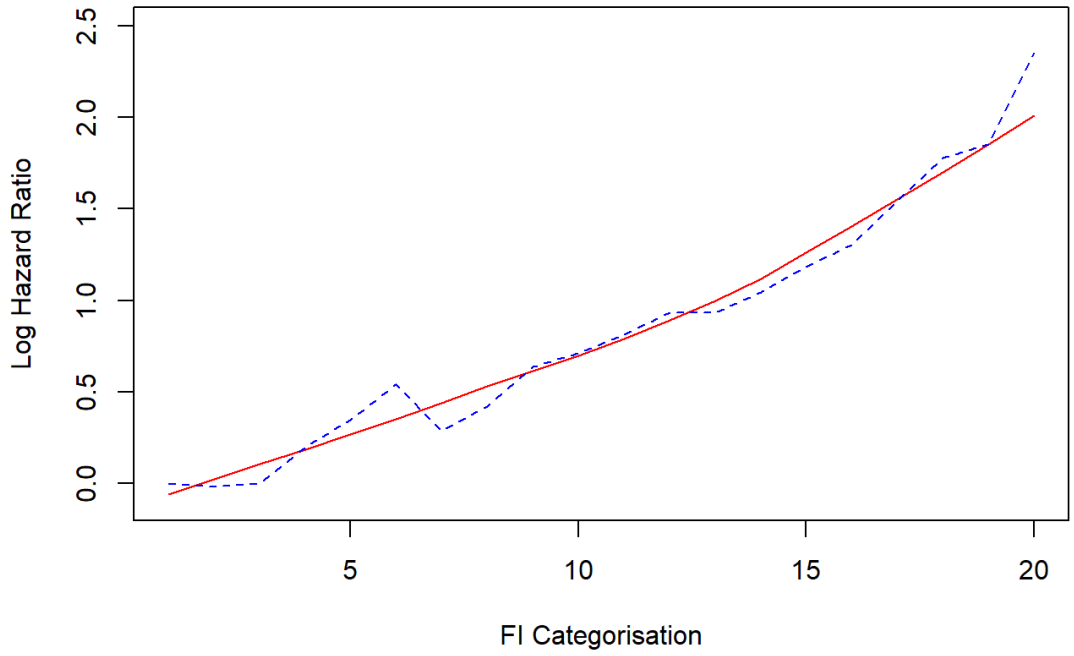
*The red dots represent the Schoenfeld residuals for each candidate predictor variable from the initial competing risk model, based on the complete-case data that included 4597 outcome events (serious muscle disorders) at 2601 time points throughout the 10 years. The blue lines were the LOESS smoothed curves showing the associations between the residuals and time.

Appendix 26 Examination of linearity of the relationship between the continuous variables and the outcome

Variable: Age



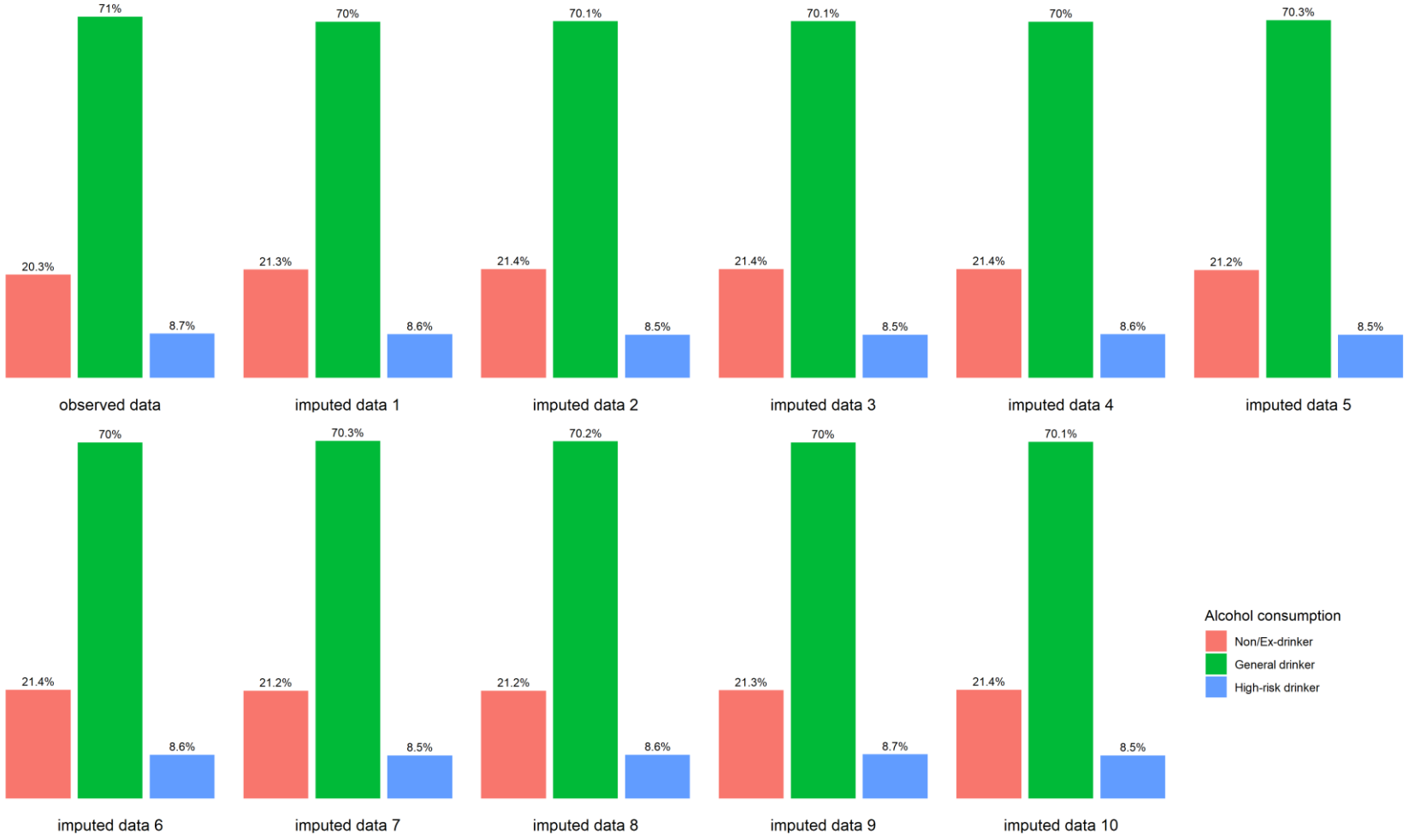
Variable: Frailty index (FI)



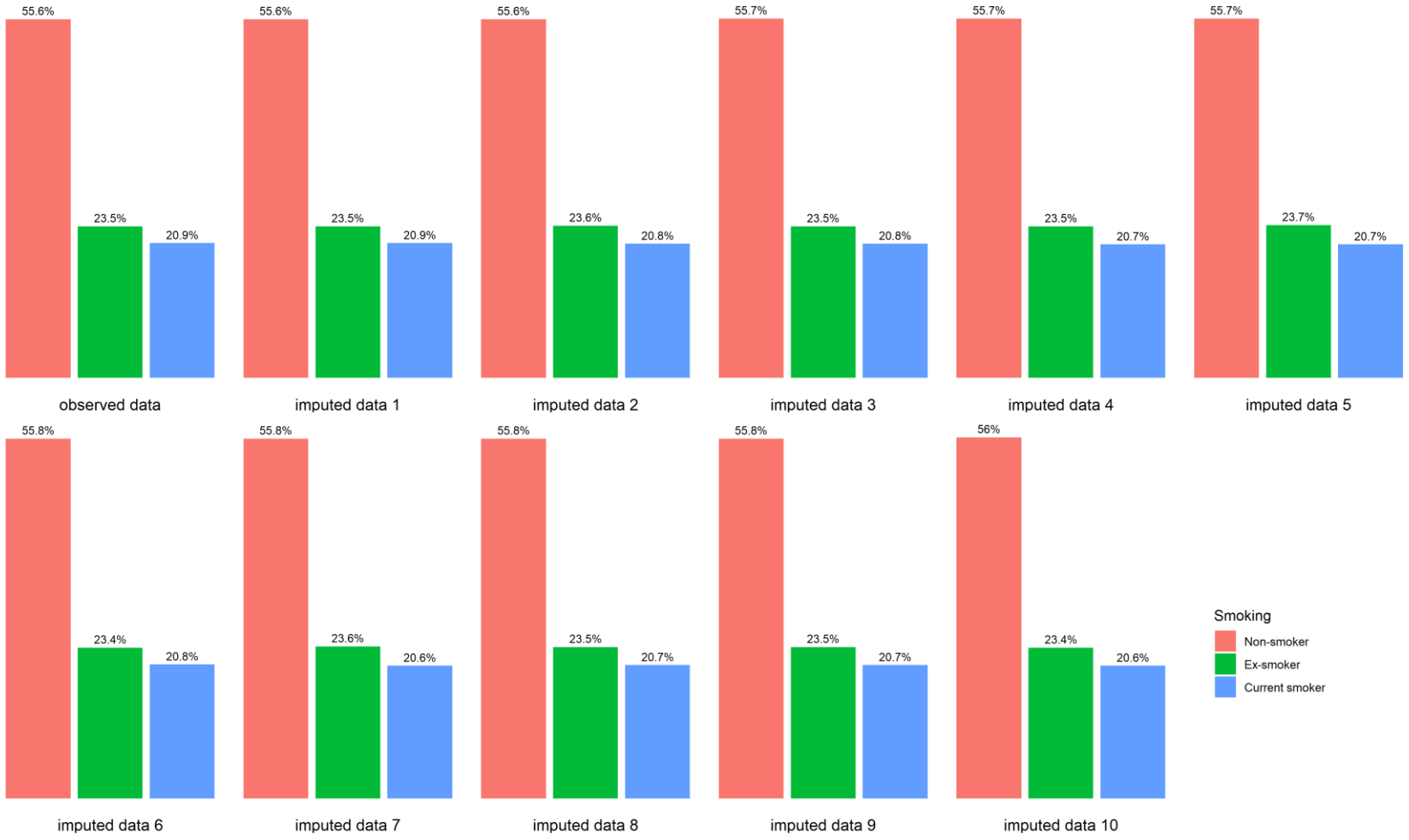
Appendix 27 Comparison of the imputed data and the observed data

Ethnicity

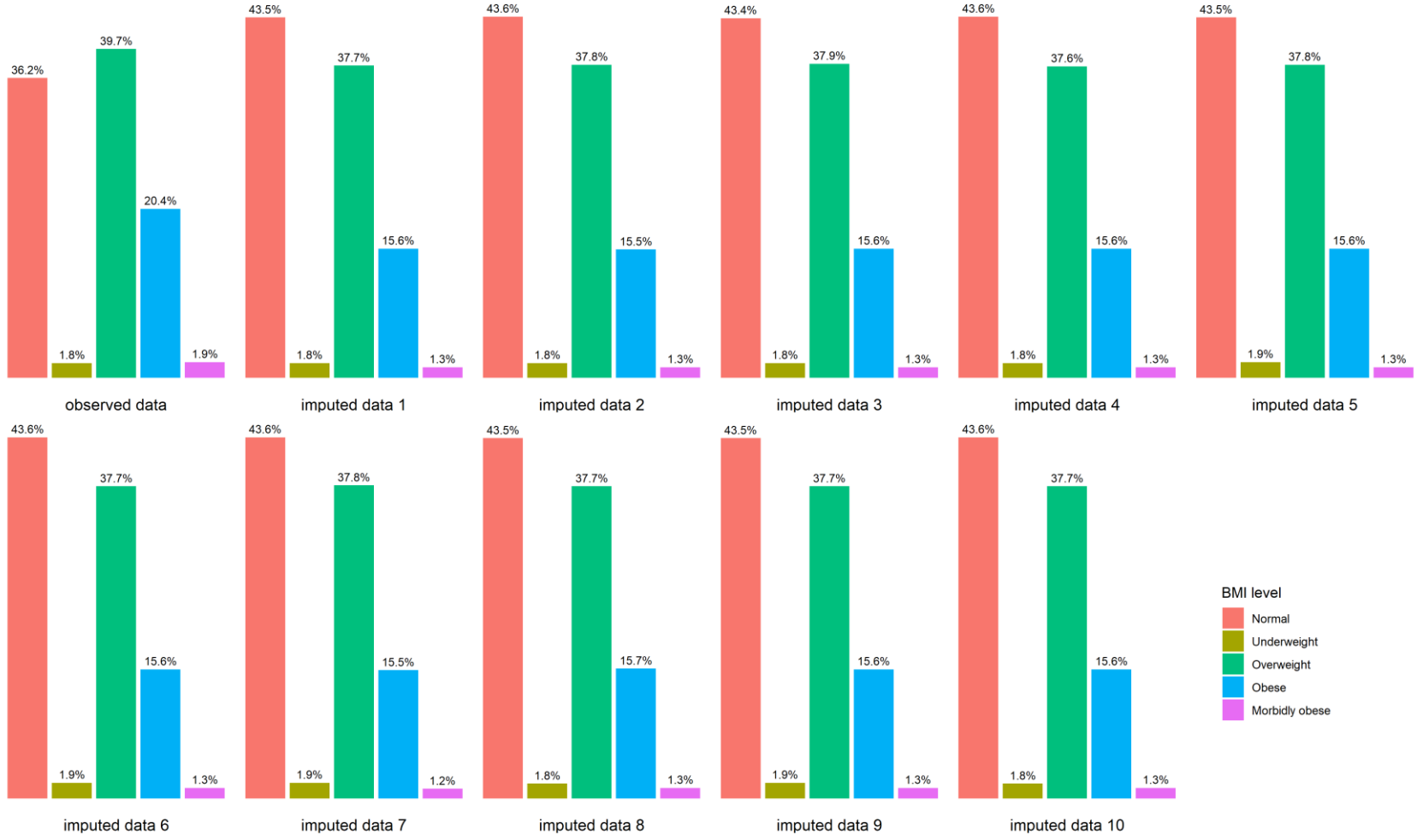
Alcohol consumption



Smoking status



BMI level


36.2%
1.8%
39.7%
20.4%
1.9%
observed data
43.5%
1.8%
37.7%
15.6%
1.3%
imputed data 1
43.6%
1.8%
37.8%
15.5%
1.3%
imputed data 2
43.4%
1.8%
37.9%
15.6%
1.3%
imputed data 3
43.6%
1.8%
37.6%
15.6%
1.3%
imputed data 4
43.5%
1.9%
37.8%
15.6%
1.3%
imputed data 5
43.6%
1.9%
37.7%
15.6%
1.3%
imputed data 6
43.6%
1.9%
37.8%
15.5%
1.2%
imputed data 7
43.5%
1.8%
37.7%
15.7%
1.3%
imputed data 8
43.5%
1.9%
37.7%
15.6%
1.3%
imputed data 9
43.6%
1.8%
37.7%
15.6%
1.3%
imputed data 10
BMI level
Normal
Underweight
Overweight
Obese
Morbidly obese

Appendix 28 Predictor coefficients in the initial model

Candidate Predictor	Coefficient (SE)	SHR (95% CI)
Demographics		
Age		
Age $_{FP1}$	3.809 (0.170)	45.12 (32.31 - 63.00)
Age $_{FP2}$	3.275 (0.175)	26.44 (18.76 - 37.25)
Gender		
Male		
Female	0.590 (0.033)	1.80 (1.69 - 1.93)
Ethnicity		
White		
Black	0.030 (0.160)	1.03 (0.75 - 1.41)
South Asian	-0.043 (0.124)	0.96 (0.75 - 1.22)
Mixed and other	0.150 (0.109)	1.16 (0.94 - 1.44)
Deprivation		
Level 1 (least deprived)		
Level 2	0.103 (0.043)	1.11 (1.02 - 1.21)
Level 3	0.138 (0.042)	1.15 (1.06 - 1.25)
Level 4	0.179 (0.041)	1.20 (1.10 - 1.29)
Level 5 (most deprived)	0.163 (0.046)	1.18 (1.08 - 1.29)
Health Status Indicators		
Alcohol consumption		
Non-drinker/Ex-drinker		
General drinker	-0.173 (0.032)	0.84 (0.79 - 0.90)
High-risk drinker	-0.430 (0.075)	0.65 (0.56 - 0.75)
Smoking status		
Non-smoker		
Ex-smoker	0.289 (0.033)	1.34 (1.25 - 1.42)
Current smoker	0.205 (0.041)	1.23 (1.13 - 1.33)
BMI level		
Normal		
Underweight	0.099 (0.104)	1.10 (0.90 - 1.35)
Overweight	0.137 (0.033)	1.15 (1.07 - 1.22)

Obese	0.393 (0.043)	1.48 (1.36 - 1.61)
Morbidly obese	0.486 (0.067)	1.63 (1.42 - 1.85)
Frailty index	0.368 (0.026)	1.44 (1.37 - 1.52)
Comorbidities		
CVD	-0.384 (0.043)	0.68 (0.63 - 0.74)
Hypertension	-0.128 (0.033)	0.88 (0.83 - 0.94)
Diabetes	-0.218 (0.047)	0.80 (0.73 - 0.88)
Mild liver diseases	0.685 (0.182)	1.98 (1.39 - 2.83)
Kidney diseases	0.296 (0.051)	1.34 (1.22 - 1.48)
Rheumatic arthritis	0.454 (0.068)	1.58 (1.38 - 1.80)
Previous muscle problems	1.830 (0.032)	6.24 (5.86 - 6.63)
Degenerative joint disorders	0.312 (0.032)	1.37 (1.28 - 1.46)
COPD	-0.111 (0.064)	0.90 (0.79 - 1.02)
Hyperuricaemia	0.070 (0.055)	1.07 (0.96 - 1.20)
Hypothyroidism	0.254 (0.045)	1.29 (1.18 - 1.41)
Vitamin D deficiency	0.976 (0.112)	2.66 (2.13 - 3.31)
Vitamin B12 deficiency	0.506 (0.072)	1.66 (1.44 - 1.91)
Concomitant Medications		
Statin-interactive drugs	0.219 (0.037)	1.24 (1.16 - 1.34)
Myotoxic drugs	0.714 (0.030)	2.04 (1.93 - 2.16)
Statin Treatment		
No statins		
Atorvastatin	0.567 (0.056)	1.76 (1.58 - 1.97)
Rosuvastatin	0.711 (0.129)	2.04 (1.58 - 2.62)
Simvastatin	0.455 (0.042)	1.58 (1.45 - 1.71)
Fluvastatin/Pravastatin	0.324 (0.099)	1.38 (1.14 - 1.68)

Appendix 29 Predictor coefficient estimates from individual imputed datasets

Predictor	Subdistribution Hazard Ratio (SHR)									
	Estimate1	**Estimate2**	**Estimate3**	**Estimate4**	**Estimate5**	**Estimate6**	**Estimate7**	**Estimate8**	**Estimate9**	**Estimate10**
Demographics										
Age										
Age_FP1	43.69	44.16	44.12	44.34	44.55	44.72	44.45	44.33	44.62	43.39
Age_FP2	25.57	25.89	25.84	25.96	26.00	26.12	25.97	25.92	26.05	25.41
Gender										
Male										
Female	1.80	1.81	1.80	1.80	1.80	1.80	1.80	1.79	1.80	1.80
Ethnicity										
White										
Black	1.00	1.05	1.08	0.95	1.12	1.00	1.02	1.01	1.08	1.05
South Asian	0.91	1.01	0.97	1.04	0.94	0.94	0.91	0.95	0.94	1.02
Mixed and other	1.16	1.24	1.15	1.18	1.18	1.19	1.09	1.13	1.22	1.13
Deprivation										
Level 1 (least deprived)										
Level 2	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11
Level 3	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15

Level 4	1.19	1.19	1.19	1.20	1.19	1.20	1.19	1.19	1.19	1.19
Level 5 (most deprived)	1.17	1.17	1.17	1.18	1.17	1.18	1.17	1.17	1.17	1.17
Health Indicators										
Alcohol consumption										
Non/Ex-drinker										
General drinker	0.84	0.84	0.85	0.85	0.84	0.84	0.84	0.84	0.84	0.85
High-risk drinker	0.65	0.68	0.66	0.67	0.65	0.65	0.67	0.63	0.64	0.64
Smoking										
Non-smoker										
Ex-smoker	1.34	1.33	1.32	1.32	1.33	1.32	1.33	1.32	1.33	1.33
Current smoker	1.24	1.21	1.22	1.21	1.21	1.20	1.21	1.20	1.22	1.24
BMI level										
Normal										
Underweight	1.11	1.07	1.09	1.07	1.09	1.11	1.09	1.16	1.10	1.05
Overweight	1.15	1.14	1.14	1.16	1.15	1.14	1.16	1.16	1.13	1.17
Obese	1.49	1.50	1.51	1.48	1.47	1.46	1.48	1.52	1.45	1.49
Morbidly obese	1.64	1.61	1.62	1.64	1.58	1.66	1.62	1.65	1.62	1.68
Frailty Index	1.43	1.44	1.44	1.44	1.44	1.43	1.43	1.43	1.43	1.44
Comorbidities										
Cardiovascular diseases	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68

Hypertension	0.89	0.88	0.88	0.88	0.89	0.89	0.88	0.88	0.89	0.88
Diabetes	0.81	0.80	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.80
Mild liver diseases	1.99	1.98	1.99	1.99	1.99	1.99	1.99	1.99	1.98	1.99
Kidney diseases	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35
Rheumatic arthritis	1.58	1.58	1.58	1.58	1.57	1.57	1.57	1.57	1.57	1.58
Previous muscle problems	6.25	6.24	6.24	6.24	6.24	6.25	6.25	6.24	6.25	6.24
Degenerative joint disorders	1.37	1.37	1.37	1.37	1.37	1.37	1.37	1.37	1.37	1.37
Hypothyroidism	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.30	1.29
Vitamin D deficiency	2.70	2.61	2.64	2.61	2.65	2.65	2.70	2.66	2.65	2.63
Vitamin B12 deficiency	1.66	1.65	1.66	1.65	1.66	1.65	1.66	1.66	1.66	1.66
Concomitant Medications										
Statin-interactive drugs	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24
Myotoxic drugs	2.04	2.04	2.04	2.04	2.04	2.04	2.04	2.04	2.04	2.04
Statin Treatment										
No statins										
Atorvastatin	1.76	1.76	1.77	1.77	1.77	1.77	1.77	1.76	1.77	1.76
Rosuvastatin	2.03	2.04	2.04	2.04	2.04	2.04	2.04	2.04	2.04	2.04
Simvastatin	1.57	1.57	1.57	1.57	1.58	1.58	1.58	1.57	1.58	1.57
Fluvastatin/Paravastatin	1.39	1.38	1.38	1.39	1.39	1.38	1.39	1.39	1.39	1.38

Appendix 30 Measurements of the prediction performance of the StatinMD model in individual imputed datasets

A. Internal Validation

Measures	Estimate1	Estimate2	Estimate3	Estimate4	Estimate5	Estimate6	Estimate7	Estimate8	Estimate9	Estimate10
***c*-index**										
1-Year	0.792	0.791	0.789	0.790	0.790	0.789	0.790	0.790	0.789	0.790
5-Year	0.788	0.789	0.788	0.788	0.788	0.788	0.788	0.789	0.788	0.789
10-Year	0.775	0.775	0.775	0.775	0.775	0.774	0.775	0.775	0.775	0.775
D statistic (SE)										
1-Year	2.246 (0.078)	2.243 (0.078)	2.232 (0.080)	2.235 (0.079)	2.231 (0.079)	2.231 (0.079)	2.234 (0.080)	2.237 (0.079)	2.233 (0.079)	2.242 (0.079)
5-Year	2.202 (0.036)	2.205 (0.035)	2.197 (0.035)	2.197 (0.035)	2.200 (0.035)	2.197 (0.035)	2.200 (0.035)	2.201 (0.035)	2.199 (0.035)	2.206 (0.035)
10-Year	2.016 (0.023)	2.014 (0.023)	2.013 (0.023)	2.012 (0.023)	2.013 (0.023)	2.011 (0.023)	2.013 (0.023)	2.016 (0.023)	2.012 (0.023)	2.014 (0.023)
E/O ratio										
1-Year	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984
5-Year	0.927	0.927	0.927	0.927	0.927	0.927	0.927	0.927	0.928	0.927
10-Year	0.862	0.862	0.862	0.862	0.862	0.862	0.862	0.862	0.862	0.862

Calibration slope (SE)										
1-Year	1.071 (0.022)	1.071 (0.022)	1.069 (0.023)	1.07 (0.022)	1.069 (0.022)	1.068 (0.022)	1.069 (0.022)	1.070 (0.022)	1.069 (0.022)	1.072 (0.022)
5-Year	1.067 (0.011)	1.068 (0.011)	1.066 (0.011)	1.066 (0.011)	1.067 (0.011)	1.066 (0.011)	1.067 (0.011)	1.067 (0.011)	1.066 (0.011)	1.068 (0.011)
10-Year	1.009 (0.008)	1.009 (0.008)	1.008 (0.008)	1.008 (0.008)	1.008 (0.008)	1.008 (0.008)	1.008 (0.008)	1.009 (0.008)	1.008 (0.008)	1.009 (0.008)
R_D^2										
1-Year	0.546	0.546	0.543	0.544	0.543	0.543	0.544	0.544	0.543	0.546
5-Year	0.536	0.537	0.535	0.535	0.536	0.535	0.536	0.536	0.536	0.537
10-Year	0.492	0.492	0.492	0.491	0.492	0.491	0.492	0.492	0.491	0.492

B. External Validation

Measures	Estimate1	Estimate2	Estimate3	Estimate4	Estimate5	Estimate6	Estimate7	Estimate8	Estimate9	Estimate10
***c*-index**										
1-Year	0.840	0.840	0.839	0.838	0.840	0.839	0.840	0.840	0.839	0.840
5-Year	0.806	0.806	0.806	0.806	0.807	0.806	0.806	0.807	0.806	0.805
10-Year	0.782	0.782	0.782	0.782	0.782	0.782	0.782	0.783	0.782	0.782
D statistic (SE)										
1-Year	2.659 (0.041)	2.661 (0.041)	2.657 (0.042)	2.654 (0.041)	2.662 (0.041)	2.659 (0.042)	2.661 (0.041)	2.662 (0.041)	2.656 (0.042)	2.661 (0.041)
5-Year	2.419 (0.022)	2.420 (0.022)	2.420 (0.023)	2.418 (0.023)	2.422 (0.023)	2.419 (0.023)	2.417 (0.023)	2.422 (0.023)	2.418 (0.023)	2.415 (0.023)
10-Year	2.177 (0.018)	2.174 (0.019)	2.176 (0.019)	2.174 (0.019)	2.178 (0.019)	2.177 (0.019)	2.174 (0.019)	2.180 (0.019)	2.174 (0.019)	2.172 (0.019)
E/O ratio										
1-Year	1.677	1.677	1.677	1.677	1.678	1.677	1.677	1.677	1.677	1.678
5-Year	1.172	1.171	1.172	1.172	1.172	1.172	1.172	1.172	1.171	1.172
10-Year	0.869	0.869	0.869	0.869	0.869	0.869	0.869	0.869	0.869	0.869

Calibration slope (SE)										
1-Year	1.146 (0.010)	1.146 (0.010)	1.145 (0.010)	1.145 (0.010)	1.147 (0.010)	1.146 (0.010)	1.146 (0.010)	1.146 (0.010)	1.145 (0.010)	1.146 (0.010)
5-Year	1.113 (0.006)	1.113 (0.006)	1.113 (0.006)	1.113 (0.006)	1.114 (0.006)	1.113 (0.006)	1.113 (0.006)	1.114 (0.006)	1.113 (0.006)	1.113 (0.006)
10-Year	1.063 (0.005)	1.062 (0.005)	1.063 (0.005)	1.063 (0.005)	1.063 (0.005)	1.063 (0.005)	1.062 (0.005)	1.063 (0.005)	1.062 (0.005)	1.062 (0.005)
$\boldsymbol{R_D^2}$										
1-Year	0.628	0.628	0.628	0.627	0.629	0.628	0.628	0.628	0.627	0.628
5-Year	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.582
10-Year	0.531	0.530	0.531	0.530	0.531	0.531	0.530	0.531	0.530	0.530

Appendix 31 Predicted and observed risks in the ten risk groups in the calibration plots

A. Internal Validation

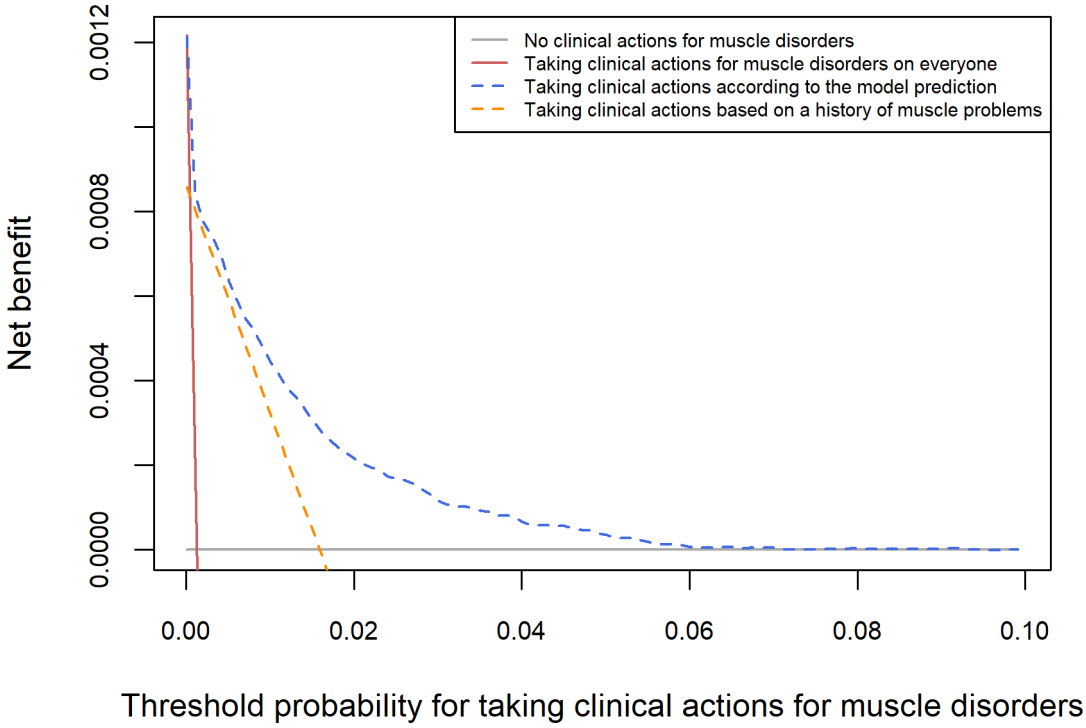
Risk Group (low to high)	1-Year Risk		5-Year Risk		10-Year Risk	
	Predicted Risk	Observed Risk (95% CI)	Predicted Risk	Observed Risk (95% CI)	Predicted Risk	Observed Risk (95% CI)
1st	0.01	0.02 (0.01-0.02)	0.03	0.07 (0.06-0.08)	0.07	0.12 (0.10-0.13)
2nd	0.01	0.02 (0.01-0.02)	0.05	0.07 (0.06-0.08)	0.13	0.16 (0.14-0.18)
3rd	0.01	0.01 (0.01-0.02)	0.07	0.06 (0.05-0.07)	0.17	0.16 (0.14-0.18)
4th	0.02	0.02 (0.01-0.02)	0.08	0.06 (0.05-0.07)	0.20	0.18 (0.16-0.20)
5th	0.02	0.01 (0.01-0.02)	0.09	0.06 (0.05-0.08)	0.23	0.21 (0.18-0.23)
6th	0.02	0.01 (0.01-0.02)	0.11	0.09 (0.07-0.10)	0.27	0.25 (0.22-0.28)
7th	0.03	0.02 (0.01-0.02)	0.13	0.12 (0.10-0.14)	0.33	0.35 (0.32-0.39)
8th	0.03	0.03 (0.02-0.04)	0.17	0.15 (0.13-0.17)	0.44	0.41 (0.37-0.45)
9th	0.05	0.04 (0.03-0.05)	0.27	0.24 (0.22-0.27)	0.67	0.62 (0.57-0.67)
10th	0.21	0.23 (0.21-0.25)	1.12	1.17 (1.11-1.22)	2.73	2.66 (2.56-2.77)

B. External Validation

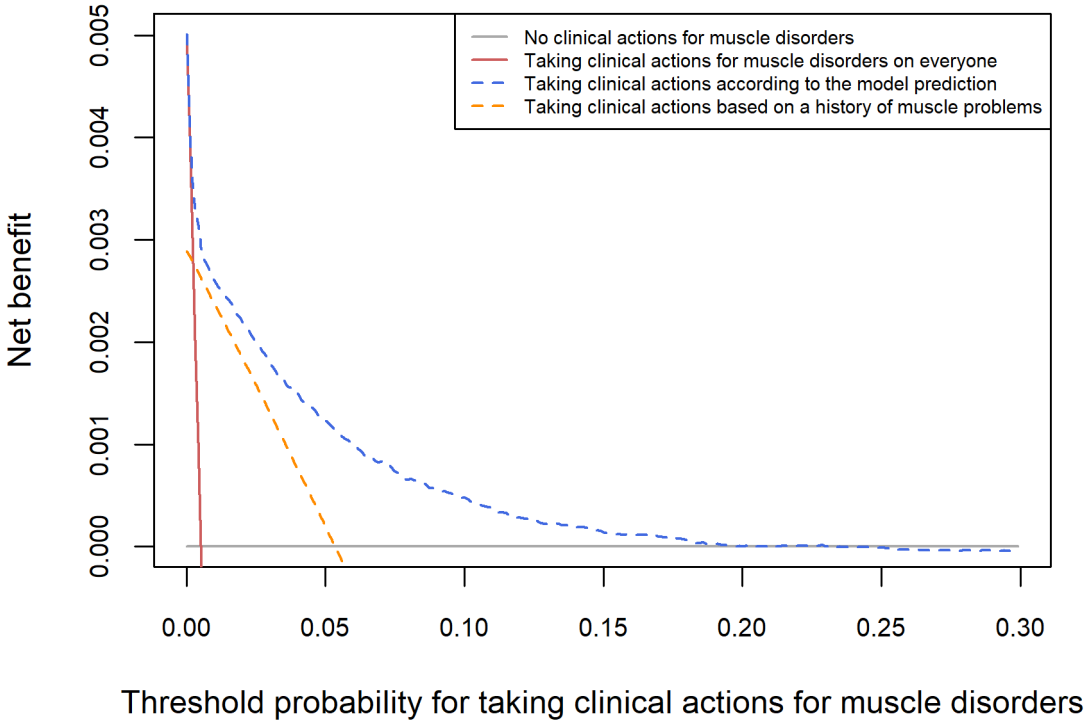
Risk Group (low to high)	1-Year Risk		5-Year Risk		10-Year Risk	
	Predicted Risk	Observed Risk (95% CI)	Predicted Risk	Observed Risk (95% CI)	Predicted Risk	Observed Risk (95% CI)
1st	0.01	0.01 (0.01-0.02)	0.03	0.06 (0.05-0.07)	0.08	0.12 (0.11-0.13)
2nd	0.01	0.01 (0.01-0.02)	0.06	0.06 (0.06-0.07)	0.14	0.13 (0.12-0.14)
3rd	0.01	0.01 (0.01-0.01)	0.07	0.06 (0.05-0.07)	0.18	0.14 (0.13-0.15)
4th	0.02	0.01 (0.01-0.02)	0.08	0.07 (0.06-0.07)	0.21	0.15 (0.14-0.16)
5th	0.02	0.01 (0.01-0.02)	0.10	0.08 (0.07-0.09)	0.24	0.19 (0.17-0.21)
6th	0.02	0.02 (0.02-0.03)	0.11	0.12 (0.11-0.13)	0.27	0.26 (0.24-0.28)
7th	0.02	0.05 (0.04-0.05)	0.13	0.17 (0.16-0.18)	0.33	0.35 (0.32-0.37)
8th	0.03	0.06 (0.05-0.06)	0.17	0.22 (0.20-0.24)	0.42	0.44 (0.42-0.47)
9th	0.05	0.07 (0.06-0.07)	0.25	0.29 (0.27-0.31)	0.62	0.60 (0.57-0.63)
10th	0.21	0.42 (0.40-0.44)	1.09	1.46 (1.42-1.50)	2.62	2.61 (2.55-2.68)

Appendix 32 Decision curves of the StatinMD model among the statins users in the external validation cohort

A. 1-year risk classification



B. 5-year risk classification



C. 10-year risk classification