A Precision Medicine Approach to Lipid Monitoring in the Secondary Prevention of Cardiovascular Disease

Rosemary Elisabeth Brown

MSc, BSc (Hons)

Submitted in fulfilment of the requirements for the

Degree of Doctor of Philosophy

University of Glasgow

Institute of Cardiovascular and Medical Sciences

College of Medical, Veterinary & Life Sciences

University of Glasgow

May 2021

Abstract

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality globally. Individuals should have their risk factors monitored and clinically managed to lower this risk, and many countries adopt similar prevention strategies targeting those at highest risk. One such high-risk group are those who have previously experienced a cardiovascular event, and consequently these patients are universally identified and targeted in risk management strategies. This thesis initially compared recommendations for the management of lipids in secondary prevention populations in a systematic review of national clinical guidelines. This found that statins were consistently recommended, but there were substantial differences in the use of lipid targets and the frequency of lipid monitoring after statin initiation, although annual lipid tests were the most frequently recommended. Aside from expert opinion, however, there was little robust evidence to support the recommendations for targets and monitoring frequency.

Therefore, the frequency of lipid testing under the current guidelines is unlikely to be optimal for the management of many patients. Increased use of electronic health records could allow the development of algorithms to result in a personalised approach to lipid testing. Specifically, if most patients have lipid tests annually, the possibility that there are a group of patients that need less frequent monitoring can be assessed. The remainder of this thesis, therefore, aimed to explore this within a subgroup of the secondary prevention population, survivors of myocardial infarctions (MIs), within Greater Glasgow and Clyde (GGC), where current and previous guidelines recommend an annual lipid test.

To achieve this, the cohort was first described to facilitate comparisons with external literature and to understand the ongoing real-life clinical management of these patients. Associations between adherence and the achievement of guideline-recommended lipid targets were then investigated with further hospitalisations for MIs and mortality. Finally, factors associated with non-adherence and non-target lipids were identified and used to predict patients' subsequent adherence and cholesterol levels, i.e. those who could receive reduced lipid monitoring.

Data was obtained from NHS GGC's Safe Haven for 11,110 patients who experienced a non-fatal MI between 2009 and 2014, with follow up available until July 2017. Demographics were consistent with similar observational cohorts from other countries in the literature, including a greater proportion of males to females, an average age of 67 years, and approximately a fifth diagnosed with diabetes before their baseline MI. Estimated statin adherence, obtained through encashed prescribing records, found that two thirds achieved an average adherence during follow up ≥80% and 85% ≥50%. Three quarters of those with at least one lipid test achieved LDL ≤1.8mmol/l during follow up. Statin adherence did not fully account for LDL target achievement, but those with higher adherence were significantly more likely to achieve it.

High adherence and lipid target attainment were common suggesting that there was a subset of patients for whom an annual lipid test could be considered unnecessary, as a further lipid test was unlikely to change any clinical decisions. Non-adherence and elevated lipids were separately significantly associated with increased mortality within this cohort (<80% adherence, HR (95% CI): 1.4 (1.3-1.5); LDL>1.8mmol/l: 1.3 (1.2-1.4)), and with CVD mortality specifically (<80% adherence: 1.3 (1.1-1.5); LDL>1.8mmol/l: 1.3 (1.1-1.5)). Therefore, careful and accurate identification of low risk patients is needed to avoid increased mortality.

Latent class analysis, a type of mixture model for categorical variables, was implemented to explore clustering and patterns within the data associated with lipid target achievement and adherence, into latent classes. For LDL ≤1.8mmol/l, sensitivity of these classes was 83% and positive predictive value was 100%, meaning all those predicted to achieve the target did so. This positive predictive value was also observed when ≥50% adherence was considered, and sensitivity was 99%. The class share of those predicted to have ≥50% adherence was substantially larger, with 85% predicted to do so, compared to 42% for LDL ≤1.8mmol/l. Associations between predicted classes and mortality showed that those predicted to have ≥50% adherence experienced lower rates of mortality, than those predicted not to. This was not the case for the LDL ≤1.8mmol/l, although the predictions performed no worse than those observed.

In conclusion, given lipid tests as part of an annual review can be expensive in terms of the time needed for repeat appointments and biochemistry, the purpose of these for secondary prevention CVD patients needs to be considered and clarified by guideline committees. Once a patient meets a lipid target and adherence continues to be high, clinical decisions are unlikely to change with further blood tests. The results in this thesis have shown that using previous adherence and lipid results, and demographic information, patients' adherence can be accurately predicted and therefore could be used as a practical marker of lipid test's necessity within a review. However, before the implementation of this approach should be considered, further validation of these results within other external observational cohorts is required, and a non-inferiority randomised controlled trial should also be implemented.

Table of Contents

List of Tables

List of Figures

List of Equations

Posters and Abstracts

Brown, R. *et al*. (2019) 'OP16 Associations of cholesterol targets and mortality with statin adherence in NHS Greater Glasgow & Clyde's post myocardial infarction population', in *J Epidemiol Community Health*. BMJ, p. A7. doi: 10.1136/jech-2019-ssmabstracts.16. *Oral Presentation at the Joint European Congress of Epidemiology and Society for Social Medicine and Population Health Annual Scientific Meeting, 4-6th September 2019, Cork, Ireland.*

Brown, R. *et al*. (2019) 'Glasgow's Post Myocardial Infarction Population: Demographics, Statin Adherence and Cholesterol Targets' *Poster Presentation at the Association of Physicians of Great Britain & Ireland Annual Meeting 28-29th March 2019, Glasgow, UK.*

Brown, R. *et al*. (2018) 'Personalising Blood Testing Schedules in Chronic Disease Management' *ePoster Presentation at the Farr Institute Symposium and Innovation Workshop, 21-22nd May 2018, Manchester, UK.*

Publications

Brown, R. E., Welsh, P. and Logue, J. (2020) 'Systematic review of clinical guidelines for lipid lowering in the secondary prevention of cardiovascular disease events', *Open Heart*, 7(2), p. e001396. doi: 10.1136/openhrt-2020-001396.

Additional Papers

Ferguson, L. D. *et al*. (2019) 'Association of central adiposity with psoriasis, psoriatic arthritis and rheumatoid arthritis: A cross-sectional study of the UK Biobank', *Rheumatology (Oxford, England)*, 58(12), pp. 2137–2142. doi: 10.1093/rheumatology/kez192.

Muilwijk, M. *et al*. (2019) 'Contribution of type 2 diabetes to all-cause mortality, cardiovascular disease incidence and cancer incidence in white Europeans and South Asians: Findings from the UK Biobank population-based cohort study', *BMJ Open Diabetes Research and Care*, 7(1), p. e000765. doi: 10.1136/bmjdrc-2019-000765.

Petermann-Rocha, F., Brown, R. E., *et al*. (2019) 'Association of leisure time and occupational physical activity with obesity and cardiovascular risk factors in Chile', *Journal of Sports Sciences*, 37(22), pp. 2549–2559. doi: 10.1080/02640414.2019.1647738.

Petermann-Rocha, F., Sillars, A., *et al*. (2019) 'Sociodemographic patterns of urine sodium excretion and its association with hypertension in Chile: A cross-sectional analysis', *Public Health Nutrition*, 22(11), pp. 2012–2021. doi: 10.1017/S1368980018003889.

Petermann-Rocha, F. *et al*. (2020) 'Physical capability markers used to define sarcopenia and their association with cardiovascular and respiratory outcomes and all-cause mortality: A prospective study from UK Biobank', *Maturitas*, 138, pp. 69–75. doi: 10.1016/j.maturitas.2020.04.017.

Troncoso, C. *et al*. (2018) ‘Patterns of healthy lifestyle behaviours in older adults: Findings from the Chilean National Health Survey 2009–2010’, *Experimental Gerontology*, 113, pp. 180–185. doi: 10.1016/j.exger.2018.09.026.

Welsh, C. *et al*. (2019) ‘Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease: Data from UK Biobank’, *Circulation*. doi: 10.1161/CIRCULATIONAHA.119.041149.

Welsh, C. *et al*. (2020) ‘Glycated Hemoglobin, Prediabetes, and the Links to Cardiovascular Disease: Data From UK Biobank’, *Diabetes Care*, 43(2), p. dc191683. doi: 10.2337/dc19-1683.

Welsh, C. E. *et al*. (2020) ‘Grip Strength and Walking Pace and Cardiovascular Disease Risk Prediction in 406,834 UK Biobank Participants’, *Mayo Clinic Proceedings*, 95(5), pp. 879–888. doi: 10.1016/j.mayocp.2019.12.032.

Welsh, P. *et al*. (2020) ‘Lipoprotein(a) and cardiovascular disease: prediction, attributable risk fraction, and estimating benefits from novel interventions’, *European Journal of Preventive Cardiology*, p. zwaa063. doi: 10.1093/eurjpc/zwaa063.

Acknowledgements

Firstly, I would like to thank those at the University of Glasgow, University of Edinburgh, and the Medical Research Council for accepting me onto the Precision Medicine Doctoral Training Programme and funding this project (Project No: 1952050; Grant Ref: MR/N013166/1) as otherwise this thesis would not exist.

Four people also need a huge thank you for supervising me throughout the last three and a half years: Dr Paul Welsh, Dr Jennifer Logue, Professor Jim Lewsey, and Professor Sarah Wild. Sarah, thank you for your help with the diabetes data and stepping in to present my results at such short notice. Jim, thank you for guiding the statistical aspects of this research when we were at a loss for next steps, and helping with interpreting increasingly confusing R output. Jen, thank you for your clinical insights, endless support and guidance, your pushing to make me a better scientist and getting me out of my comfort zones, and your constant stream of cat photos – all of which were appreciated! Paul, thank you for all you have contributed to this work – for checking code, helping with my writing, balancing the statistical and clinical aspects carefully, and your encouragement. Thank you for being there and listening, making time for our meetings, being my friend when I needed one, making me laugh when I did not think it was possible, and seeing solutions when I could only see problems.

Thank you to those at the Safe Haven who performed the extraction and dealt with my many queries regarding it. There would be no project without the data. Thank you to all at the Robertson Centre for hosting the data and your patience with me constantly needing to install R packages and slowing down the servers.

Thank you to all in the metabolic medicine team for always being happy to help. Thank you to Anne Sillars, Alan Cameron, Evan Campbell, Stuart Gray, and Naomi Rankin for your company. Thank you to Carlos Celis for looking after me in my first few weeks and for your unwavering belief in me. Thank you to Fernanda Gabler Trisotti, Ayan Merchant, Solange Parra Sotto, and Jirapitcha Boonpor for all the office (remember those?!) madness, support, and laughter.

Thank you to all in Lilybank Gardens for letting me become an honorary member and allowing the frequent invasion of your coffee mornings and offices. Thank

you to Zoe Lyall for your cuddles on bad days. Thank you to Rona Strawbridge for looking out for me. Thank you to Heather Waddell for your constant pictures of Clyde, understanding the challenges, and celebrating the good.

Thank you to John McClure for the opportunity to assist with teaching statistics throughout this. This has shone a light on a potential career for me that I would never have known I could do. Thank you also for your humour and friendship.

Thank you to Rachael McDonald, Andy Cornwell, and Kamilla Michniewicz. Thank you for looking after me despite having no clue as to what I do. Thank you to Abi Teekaram, Katie Lombardo, Uuriintuya Munkhbayar, and Melina Anetaki, for being great flatmates, tolerating my grumpiness, untidiness, and stress-baking.

Thank you to Rebecca and Matt Thomas for the superb task of making your wedding cakes in the middle of it (and worrying Paul and Jen that I would leave academia to go and run a bakery!).

Thank you to Keira Johnston and Stephen Wilkie. We have got through this madness together. Thank you for all the support and laughter. You have made dark days light and I will never be able to thank you enough. Thank you to Lyn Ferguson for hearing many rants and your frequent statistical calls (even if I said otherwise at the time). Thank you to Fanny Petermann-Rocha. I am going to miss your constant singing, dancing, and our random conversations. Thank you for your optimism (even if I did not enjoy the salsa dancing when I had not slept!). Thank you for always being there. I could not have done it without you all.

Thank you to numerous other friends and my family members for putting up with years of constant PhD talk, supporting me, and asking all the things that PhD students hate. Last, and by no means least, an enormous thank you to my wonderful mum who has borne the brunt of my struggles. Thank you for never doubting me, even when I did. Thank you for listening and allowing me to persevere when the easiest option would have been to walk away. Thank you for putting up with the moods, tears, and long hours. I got there in the end.

Oh, and thank you to Monte, the tabby menace, who, thanks to a global pandemic, helped with far more of this thesis than I ever thought he would.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work. The contents of this thesis have not been submitted for any other degree at the University of Glasgow or any other institution.

Rosemary Elisabeth Brown

February 2021

Abbreviations

ACCORD	Action to Control Cardiovascular Risk in Diabetes
ACE	angiotensin-converting enzyme
ACR	albumin-to-creatinine ratio
ACS	acute coronary syndromes
ADVANCE	Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation
AGREE	Appraisal of Guidelines for Research and Evaluation
AIC	Akaike's information criteria
ARB	angiotensin II receptor blocker
ASCVD	atherosclerotic cardiovascular disease
Ave PP	average posterior probability
BIC	bayesian information criteria
BMI	body mass index
CAD	coronary artery disease
CETP	cholesteryl ester transfer protein
CHD	coronary heart disease
CHI	community health index
CI	confidence interval

CINAHL	Cumulative Index to Nursing and Allied Health Literature
CKD	chronic kidney disease
CPRD	Clinical Practice Research Datalink
CTT	Cholesterol Treatment Trialists
CVD	cardiovascular disease
eGFR	estimated glomerular filtration rate
ESC	European Society of Cardiology
EU	European Union
FH	familial hypercholesterolaemia
FU	follow up
GGC	Greater Glasgow and Clyde
GP	general practitioner
HbA1c	glycated haemoglobin
HDL	high density lipoprotein
HSCIC	Health & Social Care Information Centre
ICD	International Classification of Diseases
IQR	interquartile range
IT	information technology

LCA	latent class analysis
LCGA	latent class growth analysis
LDL	low density lipoprotein
LLM	lipid-lowering medication
MI	myocardial infarction
MONICA	Monitoring Trends and Determinants in Cardiovascular Disease
MPR	medication possession ratio
MRFIT	Multiple Risk Factor Intervention Trial for the Prevention of Coronary Heart Disease
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NRS	National Records of Scotland
NSTEMI	non-ST-elevation myocardial infarction
OECD	Organisation for Economic Co-operation and Development
PAD	peripheral arterial disease
PCSK9	Proprotein convertase subtilisin/kexin type 9
PDC	proportion of days covered
PIS	Prescribing Information System

PREDIMED	prevención con dieta mediterránea
PROSPERO	International prospective register of systematic reviews
QOF	Quality Outcomes Framework
RA	rheumatoid arthritis
RCT	randomised controlled trial
SCI	Scottish Care Information
SCORE	Systematic Coronary Risk Evaluation
SD	standard deviation
SIGN	Scottish Intercollegiate Guidelines Network
SIMD	Scottish Index of Multiple Deprivation
SMR	Scottish Morbidity Record
STEMI	ST-elevation myocardial infarction
TIA	transient ischaemic attack
TRIP	Turning Research into Practice
UK	United Kingdom
USA	United States of America
VADT	Veterans Affairs Diabetes Trial
WOSCOPS	West of Scotland Coronary Prevention Study

Chapter 1 Introduction

1.1 Cardiovascular Disease

The cardiovascular system allows the circulation of blood around the body, which facilitates the quick and efficient delivery of nutrients, including oxygen, water, and amino acids, as well as the removal of the waste products of metabolism, such as carbon dioxide and urea, from the body's tissues. This system also plays a key role in the transportation of hormones and components of the body's immune system, as well as ensuring temperature regulation through the dispersion of heat from the body's core (Evans, Horton-Szar and Newby, 2015).

Conditions which affect the circulation of blood are often referred to under the term of cardiovascular disease (CVD), with the most common examples being Coronary Heart Disease and Stroke (Brown *et al.*, 2015). These typically develop as either a result of thrombosis (the formation of blood clots) or atherosclerosis (SIGN, 2017). The latter is the deposition of lipids in the inner walls of blood vessels leading to their subsequent hardening and narrowing. Whilst this is thought to be present to some extent in all adults, this process can subsequently lead to a decreased capacity for blood flow due to the narrowing in the arteries (stenosis) or, in the worst case, a complete obstruction of a blood vessel. In both cases, a thrombus develops as a consequence of plaque rupture or plaque erosion, with the former the most likely to result in a complete occlusion (Evans, Horton-Szar and Newby, 2015).

1.1.1 Role of Cholesterol in Atherosclerosis

Atherosclerosis is the development of lesions in the intima (inner layer) of arteries causing it to thicken and consequently lead to the narrowing of the arteries (Hansson, 2005). The first step of their development occurs due to changes in the endothelium, often triggered by damage caused by known risk factors such as smoking, diabetes, and hypertension (Mundi *et al.*, 2018). This typically occurs at sites in the arteries where blood flow is low or oscillating, such as near branch points (Bentzon *et al.*, 2014; Mundi *et al.*, 2018). The damage to the endothelium allows low-density lipoprotein (LDL) cholesterol to

permeate into, and accumulate in, the intima of the artery. This forms a fatty streak in the initial stages of lesion development, although these are present in many individuals and do not always progress to lesions (Hansson, 2005; Mundi *et al.*, 2018).

Free radicals oxidise the LDL in the intima, triggering an immune and inflammatory response (Hansson, 2005; Bentzon *et al.*, 2014). Specifically, endothelial cells and smooth muscle cells attract further monocytes to the region so that they can differentiate as macrophages and infiltrate the streak (Bentzon *et al.*, 2014). As the macrophages become lipid-laden, they progress into foam cells which, combined with the abundance of smooth muscle cells, leads to the development of an atherosclerotic lesion. This lesion (or plaque) has smooth muscle cells and a collagen matrix forming a cap to it (Hansson, 2005). Furthermore, as the process continues, the endothelium becomes damaged further, allowing more LDL to be retained within the lesion (Bentzon *et al.*, 2014).

The thinnest point of the cap of the lesion is where a rupture is most likely to occur, and this is typically the region where macrophages have infiltrated to the greatest degree. This is as a result of macrophage-derived proteinases degrading collagen in the cap of the plaque and preventing its further formation (Hansson, 2005; Bentzon *et al.*, 2014). The increasing complexity of the lesion can lead to the development of artery tertiary lymphoid organs, which increases mechanical instability (Yin *et al.*, 2016). When rupture does occur, this triggers the formation of a thrombus (blood clot), which can result in the blockage of the artery, causing ischaemia and infarction as a consequence (Hansson, 2005).

1.1.2 Types of Cardiovascular Disease

CVD is often used as an umbrella term to refer to many conditions that are illustrated in Figure 1.1. One such condition is coronary heart disease (CHD) which is another umbrella term which includes stable angina and acute coronary syndromes (ACS). Stable angina, characterised by central chest pain which subsides at rest, occurs as a result of atherosclerotic stenosis in one or more coronary arteries, with symptoms usually only experienced when the area of the affected vessel is reduced by approximately 70%. ACS, on the other hand, refers

to unstable angina, non-ST-elevation myocardial infarctions (NSTEMIs), and ST-elevation myocardial infarctions (STEMIs), which are given in increasing order of severity. These conditions all occur when the cap on the lesions in the arteries becomes unstable leading to either its erosion or rupture. This then leads to the formation of a blood clot either on the surface of, or inside, the deposit causing further expansion of the lesion, with both locations leading to, at least partial, blockage of the blood vessel (Evans, Horton-Szar and Newby, 2015).

Myocardial infarctions (MIs), which are often more commonly referred to as 'heart attacks', refers to both NSTEMIs and STEMIs. These are triggered by a lack of supply to the myocardium (an area of heart muscle) resulting in cell death, with irreversible damage caused once the lack of blood supply has persisted for between twenty and forty minutes. The consequences of MIs include sudden death, arrhythmias (irregular, fast, or slow heart rates), and heart failure (Evans, Horton-Szar and Newby, 2015).

Peripheral arterial disease (PAD) develops when atherosclerosis develops in the arteries supplying the lower limbs (and rarely the upper limbs). It is likely that most individuals with CHD will have PAD to some extent (Evans, Horton-Szar and Newby, 2015).

Ischaemic strokes, which account for approximately 85% of all strokes, occur when blood flow is reduced due to a blockage of one of the arteries supplying the brain, causing cells in the affected area to die. This blockage is usually as a result of atherosclerosis, specifically, the blood clot formed as a result of a ruptured lipid deposit in a blood vessel. If this blockage is only temporary, this can result in a transient ischaemic attack (TIA). These often precede ischaemic strokes, with the risk being at its greatest in the week after and for those with longer lasting TIAs. Haemorrhagic strokes, which make up the remaining 15% of all strokes, occur as a result of bleeds into the brain's functional tissue, with the primary mechanism thought to be hypertension (Albers *et al.*, 2010).

Finally, as atherosclerosis causes the walls of arteries to weaken, it is also the primary cause of aneurysms, though, as with haemorrhagic strokes, high blood pressure forms the main driver of their development. These occur when the three layers of the wall of an artery allow it to become permanently abnormally

dilated. Whilst this can occur in any artery, they are most common in the aorta (the largest artery in the body), specifically in the abdomen, with complications, such as a rupture and subsequent blood loss, more likely to occur as the size of aneurysm increases (Evans, Horton-Szar and Newby, 2015).

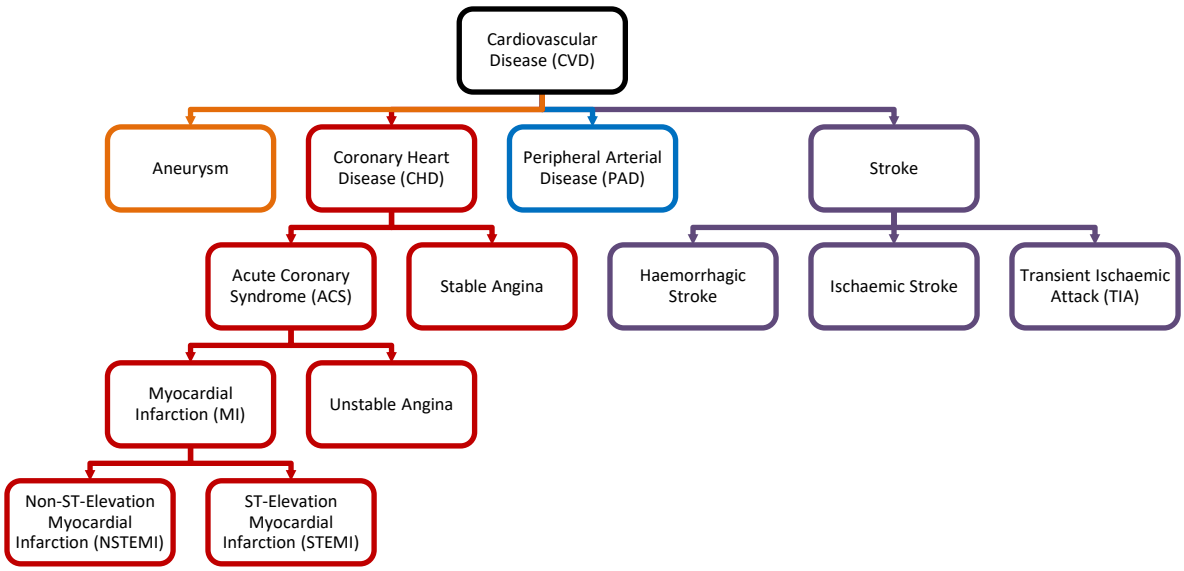


Figure 1.1: Types of CVD

1.1.3 Mortality and Morbidity

The World Health Organization reported that CVD was responsible for 17.9 million deaths globally in 2016 (World Health Organization, 2017b, 2020), with this anticipated to increase to more than 23 million by 2030 (Mathers and Loncar, 2006). In 2016, this made CVD the most common cause of death, and accounted for 44% of all deaths due to noncommunicable diseases. Furthermore, whilst the probability of death under the age of 70 due to CVD has decreased rapidly since 2000 worldwide and resulted in cancer becoming the leading cause of premature death in some high-income countries, this is not the case for lower-income countries. Indeed, in these instances, CVD is not only responsible for the most premature deaths, but also progress in reducing this mortality has been the slowest. This is possibly due to increases in the prevalence of CVD risk factors, such as obesity, but also the challenges faced by healthcare systems with many countries facing a dual burden of both communicable and non-communicable diseases (World Health Organization, 2020).

In Europe, the fifth edition of the European Cardiovascular Disease Statistics found that diseases of the heart and circulatory system were responsible for 45% of all deaths (equivalent to nearly four million deaths), making them the leading cause of death in 2017. When individual causes were considered, CHD and stroke were the two largest. When this was restricted to just EU countries, the percentage of deaths as a result of CVD decreased to 37%. However, CVD remained the leading cause of death, ahead of cancer, and CHD and stroke remained the two single leading causes (Wilkins *et al.*, 2017).

In the UK, the number of deaths from CVD has decreased considerably since 1961 when it was responsible for over half of all deaths. In 2016, according to figures published by the British Heart Foundation, 26% (equating to approximately 150,000 deaths) of all deaths in the UK were as a result of CVD. Consistent with the European figures, CHD was one of the leading causes of mortality in the UK and was responsible for 66,000 deaths in 2016 (British Heart Foundation, 2018). When looking at Scotland in isolation, which has the highest CVD incidence and mortality rates within the UK, it was estimated that just over 15,000 people died as a result of CVD, with 6,700 of these due to CHD, in 2016 (Information Services Division, 2017; British Heart Foundation, 2018).

Premature deaths, defined as deaths occurring prior to 75 years of age, are often considered to be preventable. In such cases, CVD was responsible for 1.3 million deaths in Europe in 2017, with CHD being the single largest contributing cause. Within only EU countries, CVD was the second biggest cause of premature death, after cancer (Wilkins *et al.*, 2017). In 2016, 42,000 premature deaths were due to CVD, of which approximately 5,000 were in Scotland. However, at a regional level, there were higher premature death rates reported in Northern England, Central Scotland, and South Wales, with the greatest death rates observed in the Glasgow and Greater Manchester regions (British Heart Foundation, 2018).

The majority of the estimated 85 million people living with CVD in Europe in 2015 had PAD or CHD, with the number likely to be higher when undiagnosed cases are additionally included. These two conditions were also thought to be the most prevalent when only EU countries were considered, where an estimated 49 million people were thought to be living with CVD (Wilkins *et al.*,

2017). In the UK in 2016, there were an estimated seven million people living with CVD, with 685,000 in Scotland. Furthermore, the British Heart Foundation predicted that this number was likely to increase over the coming years given the increasingly ageing and growing population, together with improving survival rates from major cardiovascular events (British Heart Foundation, 2018).

When major types of CVD were considered individually in the UK in 2016, CHD was the most prevalent with an estimated 2.3 million people affected, with over 900,000 heart attack survivors. There were also an estimated 1.3 million people living with atrial fibrillation, which occurs when the heart rhythm becomes abnormal and can result in a stroke. There were 1.2 million stroke or TIA survivors (British Heart Foundation, 2018).

1.2 Risk Classification for Cardiovascular Disease

Many countries, including Scotland, adopt high-risk prevention strategies when endeavouring to reduce the incidence of CVD. These strategies can facilitate more easily achievable benefits than population-wide strategies, as those at greater risk can reduce their risk in absolute terms more substantially than those at a lower risk. Furthermore, in the case of the management of those with CVD, those at high risk are usually recommended medication which may have undesirable side effects. Therefore, if medication was prescribed to the whole population, those at very low risk of a cardiovascular event would have a lower benefit-risk ratio of such medications. Additionally, whilst high-risk prevention strategies do not prevent all cases of CVD, they are practical and cost-effective in a population and can complement population-wide strategies. As population-wide strategies are more challenging to implement, greater attention is usually placed on identifying those at high risk to maximise a strategy's effectiveness in reducing CVD (Wilson *et al.*, 2017).

In Scotland, where the data used for much of this project is based (as outlined subsequently in Chapter 1.8), patients are classified by healthcare professionals as being at high or low risk of a cardiovascular event in the next ten years. According to guidelines, risk assessments are recommended in all individuals over 40 years or those with a first degree relative with premature CVD or familial hypercholesterolaemia. The definition of high risk, which is illustrated in

Figure 1.2, classifies many groups of patients, including those with established disease, as high risk regardless of the presence of other known risk factors. For patients who do not fall within any of these categories, the Scottish national guidelines recommend the Scotland-specific CVD risk assessment tool ASSIGN, where a score of greater than 20% also represents the patient as being at high risk (ASSIGN, 2014; SIGN, 2017).

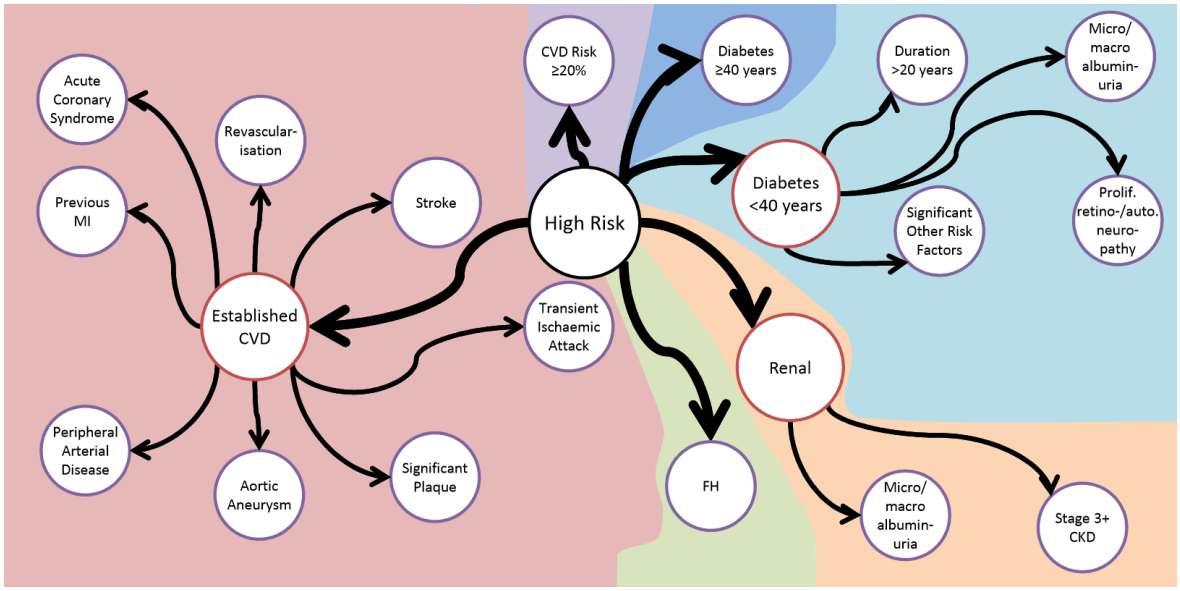


Figure 1.2: Definition of High Risk According to SIGN's 2017 Guidelines for the Prevention and Management of CVD (adapted from SIGN, 2017)

There are many modifiable and unmodifiable risk factors associated with the development of CVD which are well established both within the literature and the current CVD risk management guidelines, with many included in the ASSIGN assessment tool (ASSIGN, 2014). Therefore, within this section, risk factors are divided by those included within the ASSIGN tool, populations automatically identified as high risk by the current Scottish guidelines, and other risk factors discussed by them. However, many of these risk factors and high-risk populations are universal across guidelines for the risk management and prevention of cardiovascular disease with minimal variation between them.

1.2.1 Risk Factors used in Cardiovascular Risk Scores

1.2.1.1 Age

The risk of experiencing a cardiovascular event increases with age to an extent greater than that identified by any modifiable risk factors. This is likely due to the resulting increased exposure to known risk factors that occurs as a result of

a longer lifespan, thereby increasing the likelihood of critical atherosclerosis to develop (Sniderman and Furberg, 2008). Furthermore, age also has an impact on known risk factors such as plasma total:high density lipoprotein (HDL) cholesterol ratios and blood pressure, where both are known to increase with age. Indeed, Jousilahti et al (1999), found in a Finnish follow up study with over 14,000 individuals that increases in total cholesterol and blood pressure, together with relative weight and diabetes prevalence (also more likely to increase with age) were associated with cardiovascular events. These associations were particularly strong in women, suggesting the impact of age on cardiovascular risk is modified by a patient's sex (Jousilahti *et al.*, 1999).

1.2.1.2 Sex

Despite CVD being one of the leading causes of mortality in both men and women, the rates of the disease are considerably higher in men than in women. This disparity does decrease as age increases, with women only experiencing CVD at a higher rate than men in populations over the age of 75 years. In this age category though, this may be explained by survival bias arising from the greater life expectancy of women, as women are likely to represent a greater proportion of the population over the age of 75 years, and men who are susceptible to CVD have already died, making a higher rate of CVD more likely (Mosca, Barrett-Connor and Kass Wenger, 2011).

The difference in cardiovascular risk between the two sexes is likely due to a combination of the differences in the prevalence of risk factors, differences in the impact of risk factors in each sex, and the role of the predominant sex hormones in men and women, specifically their influence on risk factors and an individual's biology (Vitale, Mendelsohn and Rosano, 2009; Mosca, Barrett-Connor and Kass Wenger, 2011; Appelman *et al.*, 2014). In the case of the latter, it is widely believed that endogenous oestrogen reduces women's cardiovascular risk, whilst endogenous androgen increases men's. However, both sexes have each hormone present and their effects on cardiovascular risk are likely the same in both sexes, meaning that the difference in risk observed is likely due to the ratio of the hormones and their effects on vascular processes (Vitale, Mendelsohn and Rosano, 2009). Some risk factors for CVD are thought to have similar effects in men and women, such as hypertension and raised cholesterol

(Appelman *et al.*, 2014). However, in women, prolonged smoking and diabetes increase cardiovascular risk to a greater extent than observed in men (Appelman *et al.*, 2014), with the estimated effect of diabetes approximately negating the reduced cardiovascular risk experienced by women (Vitale, Mendelsohn and Rosano, 2009).

1.2.1.3 Diabetes Mellitus

Although the mortality and incidence of cardiovascular events in patients with Type 1 and Type 2 diabetes mellitus has decreased considerably over recent years, the risks of both remain between two and four times greater than those observed in the general population without diabetes (Rawshani *et al.*, 2017), including in Scotland (Read *et al.*, 2019). Whilst this decrease is likely to be as a result of a corresponding decrease in the prevalence, and increase in the control, of major cardiovascular risk factors both in patients with and without diabetes (Rawshani *et al.*, 2017; Read *et al.*, 2019), this difference in risk of cardiovascular events cannot be completely explained by the associations and coexistences of diabetes with other cardiovascular risk factors (Low Wang *et al.*, 2016). Consequently, various processes, with some likely interactions between them, have therefore been additionally implicated in the development of atherosclerosis in patients with diabetes including hyperglycaemia, insulin resistance, and endothelial dysfunction (Low Wang *et al.*, 2016).

However, the impact of hyperglycaemia in the increased risk of cardiovascular events is disputed. In many epidemiological studies, where confounding factors are likely present, its role is often the most evident. For example, one analysis of over 250,000 patients with Type 2 diabetes on the Swedish National Diabetes Register matched at a ratio of 1:5 with controls sourced from the population register between 1998 and 2012, found that a patient's glycated haemoglobin level was a significant predictor of all cardiovascular outcomes, including acute MI. Furthermore, lower levels of glycated haemoglobin than those recommended as a treatment target were associated with even lower risks of a patient experiencing acute MI and stroke (Aidin Rawshani *et al.*, 2018).

In contrast, in three randomised controlled trials (RCTs) (ACCORD, ADVANCE and VADT) intensive glycaemic control did not show a significant reduction across a

variety of cardiovascular outcomes, despite significant decreases in microvascular disease. Furthermore, in ACCORD, an increased mortality rate was observed in those randomised to the intensive arm, resulting in early discontinuation of the trial, with exploratory analysis uncovering no explanation (Skyler *et al.*, 2009). Moreover, in a recent analysis of the UK Biobank cohort, the association between HbA1c levels and cardiovascular risk has been shown to be substantially attenuated following adjustment for conventional risk factors, in populations with and without a diabetes diagnosis (Welsh *et al.*, 2020). However, both trial populations and those in the UK Biobank have been shown to be a generally healthier population than those seen routinely in clinical practice (Welsh *et al.*, 2020), possibly leading to discrepancies with other observational cohorts. Nonetheless, whilst other processes may have a role in the increased cardiovascular risk, it is likely that the elevated risk experienced by those with diabetes is largely due to the presence of other known and coexisting risk factors for CVD.

1.2.1.4 Family History of Coronary Heart Disease and Stroke

Information regarding a patient's family history of cardiovascular events, especially those events which are considered to be premature, can be used to aid in the prediction of a patient's cardiovascular risk. The impact of family history is likely to be mediated by the patient's risk factors, including those occurring as a result of genetics and their environment. Despite this, its effect has still been shown to be a significant predictor of future events when adjustment for such factors has been performed (Kinra *et al.*, 2003; Yarnell *et al.*, 2003; Lloyd-Jones *et al.*, 2004; Woodward, Brindle and Tunstall-Pedoe, 2007).

For example, in a study of around 10,000 men aged 50-59 years who completed self-administered questionnaires and were subsequently followed up for five years, total cholesterol, blood pressure, and age differed substantially between those who reported having a family history of CVD and those who did not. However, following adjustment for these risk factors and several others, the odds of a coronary event were 93% higher in those with a family history (Yarnell *et al.*, 2003). This conclusion was also supported by Kinra et al (2003) who also found, over 43 years of follow up in approximately 8,500 males aged 16-30 at

baseline, that parental history of CVD was associated with fatal CHD following adjustment for other risk factors. Additionally, whilst a stronger association was found with paternal, rather than maternal, history, there was no evidence to suggest that this difference was significant, and therefore the cardiovascular history of both parents is relevant to risk prediction (Kinra *et al.*, 2003).

The risk of recall bias when acquiring family history from a patient has raised concerns. However, an analysis of 2,000 individuals in the Framingham Offspring cohort over eight years, where parental cardiovascular history could be validated, reported adjusted odds ratios for cardiovascular events of 2.0 and 1.7 for men and women respectively, in individuals whose parents had experienced premature CHD (Lloyd-Jones *et al.*, 2004). This estimate would therefore suggest that the effects of recall bias are likely to be small and the use of family history of CVD is likely to be effective in clinical practice. Indeed, family history was initially included in the ASSIGN equations with a view to it acting as an approximate marker for ethnicity as the sample size used to create the equations was insufficient to allow the detection of any ethnic differences in cardiovascular risk. Furthermore, the creators argued that an enquiry of the patient's family history is likely to be less intimidating in a clinical scenario (Woodward, Brindle and Tunstall-Pedoe, 2007).

1.2.1.5 Blood Pressure (Systolic)

High blood pressure, or specifically hypertension, has a strong association with a patient's cardiovascular risk. The Global Burden of Disease Study in 2010 found that, of 67 risk factors examined, high blood pressure had the largest attributable risk for CVD and one of the largest impacts on a patient's disability-adjusted life years (Lim *et al.*, 2012).

The relationship between a patient's blood pressure and their cardiovascular risk appears to be linear, with values below those conventionally recommended as the intervention threshold still considered to be at an elevated risk when compared to patients with systolic blood pressure nearing 120mmHg (Vasan *et al.*, 2001; Bundy *et al.*, 2017). This has been made most evident in a meta-analysis of over 140,000 individuals, where the lowest risk of CVD and mortality was found in patients with systolic blood pressure between 120 and 124mmHg

(Bundy *et al.*, 2017). However, in some patients the consequences of lowering blood pressure to this extent (such as syncope) may make this target impractical, meaning that the risks and benefits of such a threshold should be considered carefully. Furthermore, another meta-analysis of 300,000 individuals showed that offering treatments to primary prevention patients whose baseline blood pressure is below the current threshold for treatment offered little benefit in terms of cardiovascular events (Brunström and Carlberg, 2018). This was also shown in a further meta-analysis by the Blood Pressure Lowering Treatment Trialists' Collaboration, who found that whilst relative risk reductions in cardiovascular risk were similar, the greatest absolute risk reductions were observed in those with a higher risk of an event (The Blood Pressure Lowering Treatment Trialists' Collaboration, 2014). As a consequence, many current cardiovascular guidelines provide a threshold for anti-hypertensive treatments (Piepoli *et al.*, 2016; SIGN, 2017; Whelton *et al.*, 2018; NICE, 2019a).

As with many risk factors included, high blood pressure has also been associated with, and tends to cluster with, other known risk factors including diabetes, obesity, and elevated cholesterol, as well as naturally increasing with age. However, the associated risk of high blood pressure cannot be fully explained by these other risk factors, making it an important risk factor for CVD (Kannel, 1996).

1.2.1.6 Smoking

Despite the decreases in the prevalence of cigarette smoking in some countries (including Scotland) over the preceding decades, it remains one of the leading causes of acute MI. Indeed, the INTERHEART study of 27,000 patients found smoking to be the most strongly associated risk factor, with a population attributable risk of 35.7%, the highest of all the lifestyle factors considered (Yusuf *et al.*, 2004).

Smoking has also been shown to be associated with other known cardiovascular risk factors. For example, in the five year follow up of the Edinburgh Artery Study, lower dietary consumption of antioxidants, higher alcohol intake, lower HDL cholesterol, and elevated triglycerides were all associated with cigarette smoking. However, whilst the risk of PAD and coronary artery disease (CAD) was

attenuated following adjustment for these factors, hazard ratios of 2.72 and 1.61 were observed for PAD and CAD in heavy smokers, though only the former retained statistical significance. A dose-response relationship was also observed, with those with a higher number of pack-years more likely to experience an event, particularly PAD (Price *et al.*, 1999).

The attenuation of effects observed by Price et al (1999) may suggest that the effects of cigarette smoking could be reduced by simply improving the detection and prevention of the coexisting risk factors. However, an analysis on the Framingham Offspring cohort noted that the effect of smoking had remained consistent across three decades of follow up when such prevention efforts had been occurring, thus suggesting that this may not be the case (Burke *et al.*, 2017). Furthermore, the full impact of the decrease in the prevalence of cigarette smoking, especially in the young, may not be observed for many years. Additionally, whilst former smokers have a lower risk than those who currently smoke and continue to, their risk of CVD does not fully reduce for a significant period after cessation (Yusuf *et al.*, 2004). Nonetheless, some benefits can be observed early. For example, in Scotland, the implementation of smoke-free legislation in 2006 significantly reduced the number of ACS admissions and fatalities in a ten month period. This reduction occurred in both non-smokers and smokers, although the greatest reduction was in non-smokers and was likely due to a decrease in second-hand smoke exposure (Pell *et al.*, 2008).

1.2.1.7 Plasma Lipid Concentrations (Total and HDL Cholesterol)

Details regarding the role of cholesterol, specifically LDL cholesterol, in the development of atherosclerosis are given in Chapter 1.1.1. Both total and HDL cholesterol values have been shown to have significant associations with the cardiovascular risk, with higher total cholesterol conferring a greater risk (Neaton *et al.*, 1992; Smith *et al.*, 1992), and higher HDL cholesterol associated with lower risk (Cooney *et al.*, 2009; Bartlett *et al.*, 2016).

A patient's total cholesterol level has consistently demonstrated a strong, positive, and graded relationship with cardiovascular mortality. In the MRFIT study of 350,000 men over a period of 12 years, the relative risk of cardiovascular mortality was 17% higher, and 56% higher for those whose total

cholesterol was between 4.1 and 5.2mmol/l and greater than 5.2 mmol/l respectively, when compared to those with total cholesterol less than 4.1mmol/l (Neaton *et al.*, 1992). A similar trend was also observed in the 18 years of follow up of the Whitehall Study involving 17,718 men. In this instance, a decrease of 1.21mmol/l in total cholesterol was associated with a 17% lower hazard of death as a result of CVD (Smith *et al.*, 1992).

In contrast, HDL cholesterol has been shown to have an inverse relationship with CVD. In the SCORE dataset of 96,000 individuals, significant hazard ratios for cardiovascular mortality of 0.60 and 0.76 in women and men respectively were reported per 0.5mmol/l increase in HDL cholesterol. These associations remained significant in subsequent analyses of smokers, non-smokers, and those with and without hypertensive treatments, as well as being shown to be independent of triglyceride levels and family history (Cooney *et al.*, 2009). Despite these associations, whether HDL cholesterol is causally protective is controversial, with several trials of HDL-raising cholesteryl ester transfer protein (CETP) inhibitors failing to reduce CVD risk (Barter *et al.*, 2007; Schwartz *et al.*, 2012), and another CETP inhibitor perhaps lowering CVD risk through LDL cholesterol reduction rather than HDL raising (Bowman *et al.*, 2017).

The predictive value of HDL could be as a result of its association with other components of the lipid profile. For example, in the Framingham Offspring cohort, low HDL in isolation was found to be considerably less predictive in the presence of elevated triglycerides or LDL cholesterol, suggesting a role of the ratios of components in risk prediction (Bartlett *et al.*, 2016). Nonetheless, an analysis of over 300,000 individuals reported that the strength of the relationships with CVD were similar for non-HDL cholesterol, HDL cholesterol, and direct LDL cholesterol. Therefore, the authors of this analysis suggested that total cholesterol and one of the above should be used when estimating cardiovascular risk, with the addition of further components offering minimal additional predictive value (The Emerging Risk Factors Collaboration, 2009). This was further supported by our analysis of 500,000 participants in the UK Biobank, which found that there was no meaningful improvement in the prediction of events when apolipoproteins or direct or calculated LDL were included or

substituted in existing models utilising Total and HDL cholesterol (Welsh *et al.*, 2019).

1.2.1.8 Rheumatoid Arthritis

The presence of rheumatoid arthritis (RA) is associated with an increased risk of CVD. For example, one meta-analysis of observational studies found a 50% increase of cardiovascular mortality in those with RA, compared to those without (Aviña-Zubieta *et al.*, 2008). A further meta-analysis also found a similar increase in the risk of experiencing cardiovascular events, with a 48% increased risk of CVD and a higher risk of 68% for experiencing an MI (Aviña-Zubieta *et al.*, 2012). However, both analyses also reported significant heterogeneity (Aviña-Zubieta *et al.*, 2008, 2012).

Despite this, the estimated effect of RA is comparable to that which was observed in the QRESEARCH data, the data used in the development and validation of the QRISK2 prediction equation recommended under the National Institute for Health and Care Excellence (NICE) 2014 guidelines (Hippisley-Cox *et al.*, 2008; NICE, 2014). In this analysis, following adjustment for other CVD risk factors, the effect of RA remained a significant predictor of cardiovascular events, with hazard ratios of 1.50 and 1.38, for women and men respectively (Hippisley-Cox *et al.*, 2008). When this equation was subsequently updated in 2017, the effect of RA was reduced to 1.24 and 1.23 in men and women respectively, but remained significant (Hippisley-Cox, Coupland and Brindle, 2017). Consequently, rheumatoid arthritis is included in the QRISK2, QRISK3, and ASSIGN equations for risk prediction, where in the case of the latter, its effect is estimated to be equivalent to light smoking (ten cigarettes a day) (Hippisley-Cox *et al.*, 2008; ASSIGN, 2014).

1.2.1.9 Deprivation

The risk of experiencing a cardiovascular event is higher in areas of greater socioeconomic deprivation, with one analysis finding that the increase in risk between the least and most deprived 20% is broadly comparable to a ten year increase in age or a diagnosis of diabetes (Tunstall-Pedoe and Woodward, 2006). Similarly, the MONICA register, which involved more than 5,000 patients in

Glasgow, reported that the rate of coronary events in the most deprived quarter was 1.7 and 2.4 times higher for men and women respectively than those in the least deprived. Furthermore, those in the most deprived 25% were less likely to be treated at the hospital and had a higher case fatality than those in less deprived areas (Morrison *et al.*, 1997).

Therefore, with socioeconomic deprivation having an important effect on cardiovascular event rates, primary and secondary prevention strategies have a greater potential to reduce the effect of deprivation than acute hospital-based interventions (Morrison *et al.*, 1997). Attendance at screening appointments to assess cardiovascular risk may be one factor contributing to socio-economic inequalities. A study of 8,000 patients at practices across the UK reported that whilst the most deprived patients were more likely to be at high risk of an event, they were also considerably less likely to attend screening appointments (Lang *et al.*, 2016).

In the Scottish guidelines, the inclusion of deprivation in the ASSIGN tool is an attempt to adjust for the increased risk experienced by the most deprived (Woodward, Brindle and Tunstall-Pedoe, 2007), as well as to facilitate access to prevention among deprived populations (Tunstall-Pedoe and Woodward, 2006). This aspect of CVD risk would be ignored if the Framingham score was implemented in the Scottish population. Whilst the expected risk, calculated by the Framingham score, in Scottish patients overall is higher than the observed risk, there was scope for potential overtreatment of the least deprived populations and undertreatment of the most deprived (Tunstall-Pedoe and Woodward, 2006; Woodward, Brindle and Tunstall-Pedoe, 2007). Thus, there is a need for the Scottish-specific risk score, ASSIGN, which includes deprivation (ASSIGN, 2014). This did not yield substantial improvements in discriminating between those who will and will not experience an event but instead aims to improve equity in the prevention of cardiovascular events (Woodward, Brindle and Tunstall-Pedoe, 2007).

1.2.2 Default High Risk Populations

1.2.2.1 Established Cardiovascular Disease

Patients with established disease are widely considered to be at elevated risk of further cardiovascular events and consequently are excluded from most risk prediction algorithms (NICE, 2014; Stone *et al.*, 2014; Piepoli *et al.*, 2016; SIGN, 2017; Mach *et al.*, 2020). The rationale behind this is self-evident. If atherosclerotic plaques have developed to such an extent as to cause an event in one location, it is therefore likely that similarly advanced lesions are present elsewhere within the circulatory system and may lead to further cardiovascular events. Indeed, the rate per year of further cardiovascular events in patients not treated with statins has been estimated at 5.6% in those with established disease, compared to 1.8% in the primary prevention population (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010). Survival rates are also lower in the secondary prevention population, with the World Health Organization reporting a death rate six times higher than those without CVD (World Health Organization, 2017a).

Therefore, risk management strategies are needed in this population, regardless of other risk factors. However, their exclusion from risk prediction equations does not mean that the co-existence of known risk factors in patients should not also require management. For example, additional medical therapies are recommended for those with established disease to effectively manage the risks presented by elevated lipids and blood pressure as well as the risk of clotting, as detailed in Chapters 1.3.1 and 1.3.2. There is also a wide consensus among major guidelines that other lifestyle factors, such as smoking and physical activity, should also be addressed to further reduce a patient's cardiovascular risk (NICE, 2014; SIGN, 2017; Mach *et al.*, 2020).

1.2.2.2 Renal Disease

Patients with stage 3 or higher chronic kidney disease (CKD) (defined as an estimated glomerular filtration rate (eGFR) <60ml/min/1.73m^2) or micro- or macro-albuminuria (an albumin-to-creatinine ratio (ACR) >30 and 300 mg/g respectively) are also automatically considered to be at high risk of CVD in many major guidelines, and therefore should not have their cardiovascular risk

assessed (NICE, 2014; SIGN, 2017; Grundy *et al.*, 2019; Mach *et al.*, 2020). In the case of this patient group, studies have consistently reported that cardiovascular events occur more frequently in patients with CKD than in the general population (Chronic Kidney Disease Prognosis Consortium, 2010; Gansevoort *et al.*, 2013). There is also evidence to suggest that such events occur at a greater severity, may go unrecognised, and be undertreated due to the additional presence of CKD. Furthermore, the association between CKD and CVD has been shown regardless of age and sex, and across many geographical regions (Gansevoort *et al.*, 2013).

CKD is a common sequela of hypertension and diabetes (Gansevoort *et al.*, 2013), both of which have already been discussed as significant risk factors for CVD (Chapters 1.2.1.3 and 1.2.1.5). However, two large meta-analyses of published and unpublished cohort data have shown that the association with CVD remains significant following adjustment for traditional risk factors. This association was shown using both eGFR and ACR measures, and the two measures additionally appeared to be independent of each other in their association with CVD. Moreover, thresholds used in the SIGN guidelines for the definition of the high-risk populations coincide with a significant increase in risk (Chronic Kidney Disease Prognosis Consortium, 2010, 2011). Nonetheless, whilst the association between CKD and CVD may be evident after adjustment for risk factors, the complications as a result of renal disease can make the management of the co-existing risk factors more challenging. Consequently, this population's risk of a cardiovascular event may be further elevated (Gansevoort *et al.*, 2013).

1.2.2.3 Duration of Diabetes and Diabetes Complications

Diabetes as a risk factor is discussed in Chapter 1.2.1.3. For many patients with a diagnosis of diabetes, the associated elevated risk (through the combination of their age and type of diabetes) will result in them automatically being considered as high risk under many current guidelines. However, there is some discrepancy between the guidelines in terms of the criteria used to classify patients with diabetes as high risk. For example, some take into account the duration of the disease and the presence of diabetes complications such micro- or macro-albuminuria, proliferative retinopathy or autonomic neuropathy, or

other significant risk factors (NICE, 2014; SIGN, 2017; Mach *et al.*, 2020) whilst others consider a diagnosis of diabetes sufficient (Grundy *et al.*, 2019).

Duration of disease is considered as the lifetime risk increases with longer exposure to diabetes and with age (Aidin Rawshani *et al.*, 2018). Therefore, those who develop diabetes at a younger age, who are typically patients with Type 1 diabetes, have substantially higher rates of mortality and morbidity from CVD than the general population, and consequently their risk should be managed (Rawshani *et al.*, 2017). Duration of Type 2 diabetes has also been shown to be associated with the likelihood of a range of cardiovascular outcomes, with its greatest predictive value in the prediction of stroke (Aidin Rawshani *et al.*, 2018).

Typically, younger patients with diabetes would not frequently have been exposed for a sufficiently long period to start developing comorbidities associated with diabetes. However, even in younger patients, the development of micro-or macro-albuminuria (as discussed in Chapter 1.2.2.2) confers a significant cardiovascular risk regardless of a patient's diabetes status, although diabetes often precedes its development (Gansevoort *et al.*, 2013). Proliferative retinopathy or autonomic neuropathy is also likely an indication of uncontrolled diabetes and therefore suggestive of the presence of other significant cardiovascular risk factors that need to be addressed (Skyler *et al.*, 2009).

1.2.2.4 Familial Hypercholesterolaemia

Familial Hypercholesterolaemia (FH) is a common genetic disorder that primarily manifests as elevated LDL cholesterol levels, ultimately resulting in an increased risk of premature CVD (Nordestgaard *et al.*, 2013; Lee *et al.*, 2019). There are two forms: heterozygous and homozygous. Heterozygous FH is the more common form with estimates of prevalence approximately 1 in 250 to 500 individuals (Nordestgaard *et al.*, 2013; Akioyamen *et al.*, 2017; Lee *et al.*, 2019), and patients are likely to develop CHD before the ages of 55-60 if left untreated (Nordestgaard *et al.*, 2013). Homozygous FH patients typically have more marked elevations of LDL cholesterol levels and are likely to develop CHD by their early teens, dying before the age of 20 if left untreated (Nordestgaard *et al.*, 2013). Estimates of homozygous FH's prevalence would classify it as an

orphan disease with it effecting between 1 in 250,000 to 1,000,000 (Nordestgaard *et al.*, 2013; Lee *et al.*, 2019). However, in both cases, it is widely believed that FH is likely underdiagnosed in the general population (Nordestgaard *et al.*, 2013; Akioyamen *et al.*, 2017).

Treatment with lipid-lowering medications and management of concurrent risk factors can attenuate some of the increased cardiovascular risk, especially amongst those with heterozygous FH (henceforth referred to as FH), however, the individuals would remain at high risk (Nordestgaard *et al.*, 2013). Those with FH have considerably higher lifetime cardiovascular risk (Nordestgaard *et al.*, 2013; Perak *et al.*, 2016; Akioyamen *et al.*, 2017; Lee *et al.*, 2019). For example, one meta-analysis found that the risk of MI ranged from 4.4-6.8 times that of the general population, depending on the particular variant of the LDL receptor gene (Lee *et al.*, 2019). Meanwhile, a different meta-analysis found that the 30-year risk of CHD could be as high as 5.0 times that of the general population, with a similar estimated risk (max 4.1 times) for the risk of atherosclerotic CVD (ASCVD), with the hazard ratios decreasing as age increased. Furthermore, this analysis reported that the presence of FH accelerated a patient's CHD risk 10-20 years in men and 20-30 years in women (Perak *et al.*, 2016).

1.2.3 Other Cardiovascular Risk Factors

1.2.3.1 Diet

Various components of an individual's diet can have important roles in their cardiovascular risk, as well as influencing, and being influenced by, other risk factors. Therefore, whilst dietary intake is not usually included within an individual's risk assessment nor is its influence alone considered sufficient to warrant the identification of an automatically high-risk population, its impact is still discussed within many current guidelines for CVD risk management with some recommendations made (NICE, 2014; SIGN, 2017; Grundy *et al.*, 2019; Mach *et al.*, 2020).

One recommendation, which encompasses many of the guideline's recommendations for components of an individual's diet, is that those deemed

to be at high risk of CVD should follow a Mediterranean diet pattern, specifically one that is supplemented with either extra virgin olive oil or unsalted nuts. Although a universal definition of a Mediterranean diet is yet to be widely accepted, leaving some ambiguity, this diet is often identified through its high vegetable and fruit, moderate fat, and low red meat intakes (NICE, 2014; SIGN, 2017; Grundy *et al.*, 2019; Mach *et al.*, 2020). The association between the recommended diet and a reduction in cardiovascular risk has been discussed in both RCTs and Cochrane reviews (Rees *et al.*, 2013, 2019; Estruch *et al.*, 2018). For example, in the PREDIMED trial of nearly 7,500 high-risk patients followed up for an average of five years, those who received a Mediterranean diet with extra virgin olive oil or unsalted nuts had a 30% lower risk of a major cardiovascular event than those who were randomised to a control diet (Estruch *et al.*, 2013, 2018).

In contrast, a recent Cochrane review of 12,500 patients from 30 RCTs suggested that in primary prevention a Mediterranean diet showed little to no effect on cardiovascular, and overall, mortality, albeit that it considered the quality of evidence to be low. In secondary prevention, a statistically significant reduction in cardiovascular, and overall, mortality was observed, although the number of patients included was small and the quality of evidence was also assessed to be low by Cochrane (Rees *et al.*, 2019). However, there was some evidence of associated significant reductions in other known cardiovascular risk factors, such as components of the lipid panel and in blood pressure, in both primary and secondary prevention (Rees *et al.*, 2019), with similar findings also reported in the previous larger (n>50,000) primary prevention Cochrane review (Rees *et al.*, 2013). Therefore, this may suggest that any effects of this diet on mortality may not only require research with a longer exposure or follow up but may also occur as a result of the effects in other established cardiovascular risk factors.

1.2.3.2 Alcohol Intake

The consumption of excess alcohol has been shown to have a causal relationship with several hundred diseases and injuries, resulting in it being implicated in an estimated 5% of all global deaths (World Health Organization, 2018). One of these relationships is with CVD where the harmful use of alcohol is a well-established risk factor (Costanzo *et al.*, 2010; Ronksley *et al.*, 2011).

Much of the evidence supporting the association between alcohol consumption and risk of CVD arises from observational studies. These typically report a J-shaped relationship i.e. that those with a low alcohol intake have lower cardiovascular morbidity and mortality than abstainers (Costanzo *et al.*, 2010; Ronksley *et al.*, 2011). For example, one meta-analysis of 84 prospective cohort studies (n>1m) found, in a dose-response analysis, that the equivalent of less one drink a day (2.5-14.9g/day) was associated with a significant reduction in CVD mortality, CHD incidence and mortality, and stroke incidence and mortality by between 14-25% when compared to those who consumed no alcohol. However, the risks for all outcomes did also increase as alcohol consumption increased, particularly for stroke (Ronksley *et al.*, 2011). A similar reduction in cardiovascular mortality was observed in a further dose-dependent meta-analysis, where the lowest risk was observed in patients whose daily consumption was 5-10g/day in both primary and secondary prevention populations (Costanzo *et al.*, 2010).

However, the shape of this relationship may be explained due to methodological issues arising from the use of observational data, particularly the risk of reverse causality. People who do not drink may not do so due to the presence of pre-existing ill-health which may also elevate their cardiovascular risk, resulting in a higher observed rate of cardiovascular events in abstainers (Shaper, Wannamethee and Walker, 1988). Furthermore, a Mendelian randomisation found that some genetic variations associated with reduced alcohol intake were also associated with a more favourable combination of cardiovascular risk factors (Holmes *et al.*, 2014). As a result, current Scottish guidelines recommend that all patients irrespective of their current alcohol consumption should be advised to reduce theirs to decrease their CVD risk (SIGN, 2017), whilst other guidelines suggest that consumption should not exceed recommended levels (NICE, 2014; Mach *et al.*, 2020).

1.2.3.3 Psychological Factors

The perception that psychological stress can contribute to the development of CVD is prevalent amongst both cardiac patients and the general public. However, reasonable uncertainty exists as to the exact nature of its role in the development of CVD, which is likely exacerbated by a lack of universal definition

and measurement standards in individuals. This makes the impact of stress on a patient's cardiovascular risk difficult to quantify (Bunker *et al.*, 2003). Despite this, common consequences of prolonged stress are depression and anxiety (Bunker *et al.*, 2003), and both have been shown to be associated with an increased CVD risk (Bunker *et al.*, 2003; Batelaan *et al.*, 2016).

In depression, a review of systematic reviews found a dose-response relationship between the severity of a patient's depression and the increased risk of developing CHD, independent of many conventional risk factors. Furthermore, social isolation, which is commonly associated with depression, also demonstrated a similar dose-response relationship with CHD mortality (Bunker *et al.*, 2003). However, treating a patient's depression has not been shown to reduce their CVD risk, and as such, it is advised that greater emphasis should be placed on treating the conventional risk factors for CVD that are likely to be co-existing in these patients (Carney *et al.*, 2004).

Anxiety has also been shown to be an independent risk factor for CVD in a meta-analysis of over 1.5 million individuals, with a hazard ratio of 1.52. This hazard ratio remained statistically significant and broadly similar following adjustment for suspected high publication bias, comorbid depression, and multiple CVD risk factors (hazard ratios: 1.41, 1.57, and 1.50 respectively). However, given that psychotropic medication use can also elevate an individual's cardiovascular risk and few of the included studies adjusted for this, it is possible that the true effect of anxiety on a patient's CVD risk may have been inflated in this analysis (Batelaan *et al.*, 2016).

Nonetheless, whilst there may not be sufficient cause to include psychological wellbeing indicators as part of risk score calculators, some guidelines recommend that a patient's depression, anxiety, and social isolation should be considered in combination with this score when assessing their cardiovascular risk (Piepoli *et al.*, 2016; SIGN, 2017; Mach *et al.*, 2020).

1.3 Current Management of Established Cardiovascular Disease

As much of this thesis will involve the analysis of a cohort derived from the NHS Greater Glasgow and Clyde Safe Haven (as outlined later in Chapter 1.8), all sections from Chapter 1.3.3 onwards will take a closer look at the current Scottish Guidelines. These are produced by the Scottish Intercollegiate Guidelines Network (SIGN), and the principal guideline of interest is the guideline for the risk estimation and prevention of cardiovascular disease (SIGN 149). However, the initial subsections will focus on the medications used within the management of cardiovascular disease more generally.

Regardless of their overall risk, all patients should be encouraged to make lifestyle changes to lower their risk of CVD further. For example, unhealthy diet, physical inactivity with or without increased body mass index (BMI), smoking, and high alcohol consumption (all discussed in Chapter 1.2) are examples of established lifestyle risk factors. Consequently, advice that may be given to the patient would therefore include a Mediterranean diet and an increase in physical activity, the cessation of cigarette smoking, and a reduction in alcohol consumption, respectively.

1.3.1 Use of Lipid-Lowering Medication in Cardiovascular Disease

Given the evidence of the association between LDL cholesterol and cardiovascular risk, lipid-lowering therapies are often prescribed to individuals to reduce that risk, with statins forming the principal choice in both primary and secondary prevention (World Health Organization, 2017b). Statins work by inhibiting the 3-hydroxy-3-methylglutaryl-coenzyme A reductase activity, which forms an important part in the synthesis of cholesterol in the liver. This results in an increased LDL receptor expression in the liver, leading to a greater amount of LDL cholesterol being removed from the patient's blood (Ward, Watts and Eckel, 2019).

There is a strong evidence base for the use of statins to reduce the risk of cardiovascular events. One meta-analysis of 26 double-blinded RCTs, which included over 170,000 patients showed a significant reduction in many clinical

outcomes in those randomised to receive statins. This reduction was present regardless of baseline risk (including primary and secondary prevention), and a greater, but proportional, reduction was observed as the intensity of the statin was increased. When compared to those treated with placebo, those treated with statins were 21% (95% CI: 0.77-0.81, p<0.0001) less likely to experience a major vascular event per 1mmol/l reduction in LDL cholesterol. This reduction of cholesterol is attainable for many patients, as the average reduction in LDL cholesterol in the meta-analysis in the first year of statin treatment was 1.07mmol/l (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010). However, as risk reductions are relative, those at a higher risk of a cardiovascular event have the greatest potential benefit from treatment. In addition, no lower limit has been identified where cholesterol lowering is not beneficial to CVD risk (Cholesterol Treatment Trialists' (CTT) Collaborators, 2012). For this reason, current and previous guidelines recommend statin therapy to those considered at high risk (regardless of cholesterol), with more intensive doses recommended for those with established disease and greatest overall risk (SIGN, 2007, 2017).

One of the studies included in the above meta-analysis was the West of Scotland Coronary Prevention Study (WOSCOPS), the first large endpoint trial of statins in primary prevention. This recruited 6,500 men aged 45-64 years who were randomised to receive either placebo or pravastatin 40mg over an average follow up period of five years. All patients were considered to be in the primary prevention population, but 5% had been diagnosed with angina and therefore had established CVD (The WOSCOPS Study Group, 1995). LDL cholesterol was lowered by just over 1mmol/l in the pravastatin treatment group, compared to no change in the placebo group, putting the result in line with the other trials included in the Cholesterol Treatment Trialists' meta-analysis (Shepherd *et al.*, 1995; Cholesterol Treatment Trialists' (CTT) Collaboration, 2010). Furthermore, a 31% reduction in the risk of the composite endpoint, of non-fatal MI and death from coronary heart disease, was observed at five years in the treatment group. A significant reduction was also observed when each of the events was considered separately (Shepherd *et al.*, 1995). Following the completion of the trial, patients were transferred back to primary care, where those patients who had been in the treatment group continued to have a significantly lower risk of

mortality at both ten and twenty years after the end of the trial (Ford *et al.*, 2007, 2016), even though less than two fifths were on statins in the five years after the trial completion (Ford *et al.*, 2007).

There is strong evidence suggesting that statins are largely considered to be safe, with minor muscle discomfort being the most commonly reported adverse effect of their use, which is resolved with the discontinuation of treatment (Collins *et al.*, 2016). Other adverse effects such as myopathy and haemorrhagic strokes are considered to be very rare. It is estimated that in 2,000 people treated for five years, there would be one new case of myopathy and up to two haemorrhagic strokes (Collins *et al.*, 2016). One meta-analysis of RCTs also found that statins were associated with a small increase in absolute risk of developing diabetes (Sattar *et al.*, 2010), with a second finding that this risk is associated with the dose intensity of the statin (Preiss *et al.*, 2011). However, recently published guidelines conclude that the cardiovascular benefits of statins exceed this risk and therefore should continue to be prescribed as the principal lipid-lowering therapy (SIGN, 2017; Grundy *et al.*, 2019; Mach *et al.*, 2020).

Despite this, some patients will be unable to tolerate their side effects, statins will be contraindicated, or an insufficient reduction in LDL cholesterol will be observed on the patient's maximum tolerated dose. Other lipid-lowering therapies are also recommended by many guidelines either to be used as an alternative to, or in combination with, statins. These include ezetimibe, and, more recently, PCSK9 inhibitors. Fibrates and nicotinic acids are no longer commonly routinely recommended as alternative lipid-lowering medications (SIGN, 2017; Grundy *et al.*, 2019; Mach *et al.*, 2020).

1.3.2 Other Medications Prescribed in Cardiovascular Disease

Further classes of medication have also been effective at reducing the incidence of CVD. This is due to the decrease in the impact of known risk factors, specifically high blood pressure and the formation of blood clots. Whilst hypertension can be treated in both the primary and secondary prevention of CVD (Piepoli *et al.*, 2016; SIGN, 2017; Whelton *et al.*, 2018; NICE, 2019a), antithrombotic treatment is typically only recommended for those with established disease (Piepoli *et al.*, 2016; SIGN, 2017; Arnett *et al.*, 2019).

However, unlike cholesterol where no lower limit has been identified (Cholesterol Treatment Trialists' (CTT) Collaborators, 2012), it is possible to lower blood pressure, and thin blood, too far which can result in unwanted and potentially fatal adverse events.

For hypertension, thresholds for intervention with anti-hypertensive therapies and treatment targets are provided by current guidelines (Piepoli *et al.*, 2016; SIGN, 2017; Whelton *et al.*, 2018; NICE, 2019a). In the UK, medications should be prescribed according to the British Hypertension Society algorithm (as shown in Figure 1.3), which provides the order that the medications should be considered in patients with hypertension (SIGN, 2017; NICE, 2019a). However, if a patient has experienced an acute MI, then they should be offered an Angiotensin-Converting Enzyme (ACE) inhibitor or an Angiotensin II Receptor Blocker (ARB) regardless of their blood pressure, in addition to a beta-blocker (NICE, 2013; World Health Organization, 2017b). Furthermore, in Scotland, patients who have had a stroke or TIA should also be offered anti-hypertensives even if they are not hypertensive (SIGN, 2017).

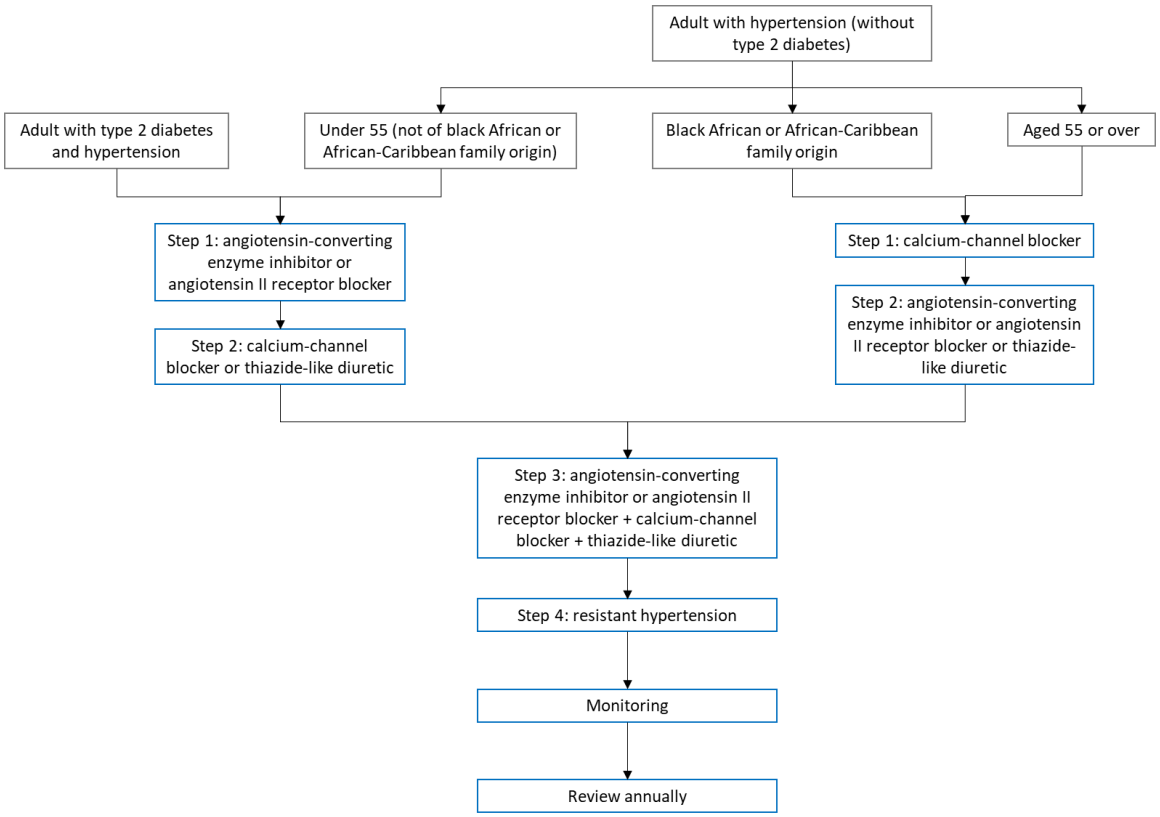


Figure 1.3: British Hypertension Society Algorithm for the use of antihypertensives (NICE, 2019b)

Antithrombotic therapy in established CVD consists of antiplatelets and anticoagulants which are not generally co-prescribed due to elevated bleeding risks. For most patients in this population, aspirin should be prescribed with clopidogrel (or other $P2Y_{12}$ inhibitor) considered in patients where this is not tolerated or contraindicated. However, there is some variation between guideline bodies as to which patients with established disease should receive solely aspirin and which should receive dual antiplatelet therapy. There is a wider consensus that in patients with stroke or TIA, dual antiplatelet (such as clopidogrel or aspirin in combination with dipyridamole) therapy should be offered (Levine *et al.*, 2016; Piepoli *et al.*, 2016; SIGN, 2017). For those patients who also have atrial fibrillation, anticoagulants should be prescribed, although care should be taken to avoid interactions with other medications (including antiplatelet medications) that are co-prescribed to minimise bleeding risk (SIGN, 2013; Levine *et al.*, 2016). Finally, due to the associated risk of serious bleeding, antiplatelet therapy is not recommended in the primary prevention population (Piepoli *et al.*, 2016; SIGN, 2017; Arnett *et al.*, 2019).

1.3.3 Management of High-Risk Individuals under Current Scottish Guidelines

Regardless of their baseline plasma lipid levels, patients classified as high risk are typically offered drug treatment, conventionally statins, to lower this further, with the expectation that this will reduce non-HDL cholesterol by 1mmol/l or 40%. However, this should not be used as a treatment target, but rather a measure to check a patient's adherence following the initiation of treatment. In cases where the patient is also experiencing hypertension, blood pressure lowering therapy should also be offered (SIGN, 2017).

Assessing and identifying those at high risk of an event is not the only aim of the guideline, however. Once identified, a patient should be monitored to see if interventions have been successful in reducing risk factors, with the aim of preventing such events from occurring. Patients considered at low risk should be reviewed every five years, with patients at high risk reviewed annually as part of good practice (SIGN, 2017). These annual reviews should be considered to "*discuss lifestyle modification, medicines adherence and address CVD risk factors*" (SIGN, 2017, p. 15).

Given plasma lipid levels are a known risk factor for disease, it is therefore likely that a clinician will recommend blood tests to ascertain the effectiveness of the treatment, and potentially use the results as a starting point when discussing the patient's adherence to their current treatment regimen, or the potential introduction of further treatment. Furthermore, the guidelines subsequently provide an expected reduction: a 40% reduction in non-HDL cholesterol or a 1 mmol/l reduction in non-HDL, in cases where the former is unattainable (SIGN, 2017). As these suggested aims are provided, this could further facilitate the use of blood test results as part of a patient's annual review.

1.3.4 Comparison with Previous SIGN Guidelines

Prior to the current (2017) guidelines, the previous SIGN guidelines were published in 2007. The definition of the high-risk population stayed broadly similar between the two. However, in 2007, patients with stage 3 or higher CKD, micro- or macro-albuminuria, or diabetes and under the age of 40 (with additional risk factors) were not considered to be automatically at high risk. Annual monitoring in this population was still encouraged in the prior guideline, as was monitoring every five years in those considered to be at low risk. Indeed, all recommendations for the low-risk population remained identical between the two guidelines, with only levels of supporting evidence altering (SIGN, 2007, 2017).

Recommended medications have altered slightly. With respect to statins, high-risk patients should have been offered a statin since the 2007 guidelines. In 2007, simvastatin 40mg was the recommended drug and dose for those at high risk, with more intensive therapy considered for those with established disease (though the specific medication is not detailed). In 2017, atorvastatin is the recommended drug, with 80mg for secondary prevention and 20mg for primary prevention patients.

In 2007, aspirin 75mg should have been considered for all at high risk but offered to those with diabetes and over 50 years of age or with additional risk factors. In 2017, aspirin is only recommended for those with established disease. The same medications and doses of anticoagulants are recommended for stroke patients in both guidelines (SIGN, 2007, 2017).

There are also some differences in blood pressure recommendations, but the target for the high-risk population remained at 140/90mmHg. In patients with diabetes, those with a blood pressure higher than 130/80mmHg should have been considered for treatment, though in the current guidelines this is revised. In the current guidelines, if the systolic blood pressure is above 140mmHg then treatment should be offered; if it is less than 140mmHg then treatment should be considered, with a treatment target of 135/85mmHg. The same target is provided for patients with CKD or micro- or macro-albuminuria, and in patients with a history of stroke, antihypertensives should be prescribed even if their blood pressure is normal (SIGN, 2007, 2017).

1.3.5 Differences with NICE Guidelines

These SIGN recommendations are broadly in line with the recommendations published by the National Institute for Health and Care Excellence (NICE). In the case of low-risk patients, these recommendations are similar, with both recommending five yearly reviews (NICE, 2014; SIGN, 2017). The identification of those subgroups automatically considered high risk is also similar, with only real differences occurring regarding individuals with diabetes mellitus. In patients with Type 1 diabetes, both guidelines identify people over the age of 40 years as high risk, but under the age of 40, there are differences regarding the duration of disease which constitutes high risk, 10 years in the case of NICE, compared to 20 years in SIGN. For those patients with Type 2 diabetes, SIGN has the same recommendations as Type 1, but for the NICE 2014 recommendations, the patient's risk should be assessed using the most recent QRISK risk calculator (NICE, 2014; SIGN, 2017). Perhaps the largest differences between the two guidelines are the recommended risk calculators and the thresholds used to define high risk. In Scotland, a patient's risk is calculated using the ASSIGN tool and they are considered high risk if their 10 year risk of an event is greater than 20%, whereas NICE defines a patient as being at high risk if their risk exceeds 10% using the most recent QRISK tool (NICE, 2014; SIGN, 2017).

When it comes to high-risk populations, lipid-lowering therapy, preferably statins, is also recommended by NICE, together with an annual review for each patient (NICE, 2014). NICE specify only that such a review should include the topic of the patient's adherence to medication but go onto state that a fasting

non-HDL cholesterol blood result could be used to assist with this aspect. Both guidelines have the same anticipated reduction in cholesterol to assess the effectiveness and patient's adherence to medication: a 40% reduction in non-HDL cholesterol from a non-medicated baseline, although SIGN additionally provides a supplementary guide reduction of 1mmol/l, and both values are only expectations and not targets (NICE, 2014; SIGN, 2017).

1.4 Adherence to Recommended Medications

1.4.1 Approaches to Assessing Adherence

Medication adherence has been defined by Osterberg and Blaschke (2005) as *"the extent to which patients take medications as prescribed by their healthcare providers"*, with methods used to assess adherence divided into two categories: direct and indirect (Osterberg and Blaschke, 2005). This differs from medication persistence, which is the duration that a medication is taken by an adherent patient without a break (Ho, Bryson and Rumsfeld, 2009; Ryan *et al.*, 2017).

Direct methods of measuring adherence include the act of physically observing the patients taking the medication, as well as the measuring of the levels of medicine, its metabolites, or an associated biological marker in the patient's blood or urine at regular intervals. In contrast, indirect methods often rely on patient self-reporting of their adherence, pill counting, or the use of electronic prescription data (Osterberg and Blaschke, 2005). None of these methods are without their disadvantages, though, and as a result, there is no gold standard method of assessing adherence (Osterberg and Blaschke, 2005; Ho, Bryson and Rumsfeld, 2009; Brown *et al.*, 2016; Stephenson *et al.*, 2018).

Whilst direct methods are generally thought to be more robust than indirect methods, they cannot be implemented easily within routine clinical practice (Osterberg and Blaschke, 2005). This is largely due to the costs of the more intensive monitoring and testing, as well as the higher patient burdens (Ho, Bryson and Rumsfeld, 2009). Additionally, for methods involving blood samples, the timing of doses, especially for medications with shorter half-lives, can make their detection within the blood challenging (Ryan *et al.*, 2017; Sutherland *et*

al., 2018). Furthermore, if patients are informed that such a test is going to occur this has been shown to result in higher adherence within the samples collected (Sutherland *et al.*, 2018). This effect has also been documented within the literature as "white coat adherence", whereby patients restart medications just before their review appointments (Brown *et al.*, 2016; Ryan *et al.*, 2017). One approach to reducing such a bias would be to introduce random blood tests within a population, but this would be difficult to implement in clinical practice and as a result has largely been confined to drugs of abuse (Ryan *et al.*, 2017).

Observer effects, however, are not confined to direct methods of estimation and are also likely when using indirect methods, specifically those involving patient-reported adherence (Osterberg and Blaschke, 2005; Ho, Bryson and Rumsfeld, 2009; Brown *et al.*, 2016). For example, when pill counting methods are utilised, patients may discard pills to appear more adherent, and when medication diaries are used, both inaccurate recall and social desirability bias are likely to lead to overestimation (Ho, Bryson and Rumsfeld, 2009). Nonetheless, one study looking at the adherence of anticoagulants in non-valvular atrial fibrillation found that patient-reported adherence (captured through four validated questionnaires over a year) was lower than the adherence derived from electronic health records. One explanation for this discrepancy is that the questionnaires reflected a patient's adherence in the recent run-up to its completion, whereas the electronic health record data reflected their adherence throughout the year. Furthermore, the associated reduction in healthcare costs with the adherence estimated from the health records was larger than with patient-reported adherence (Stephenson *et al.*, 2018).

Electronic prescription data, which can be divided into two distinct types (prescribing and dispensing), is less susceptible to the biases discussed and often more easily applicable to routine clinical practice or larger trial populations (Brown *et al.*, 2016). Prescribing data, which is available in the Clinical Practice Research Datalink (CPRD), includes all prescriptions that have been written by healthcare professionals (Herrett *et al.*, 2015). This allows the identification of patients who should be in possession of medication but does not indicate whether the medication was ever dispensed by the pharmacy and collected by the patient. Therefore, this data allows the analysis of healthcare professionals'

prescribing patterns but can also result in adherence overestimation. In contrast, dispensing data, which is available in the Scottish National Prescribing Information System (PIS), includes all prescriptions that have been dispensed (and encashed, in the case of PIS) (Information Services Division, 2010). As a result, it can be inferred that the patient is in possession of the medication. However, like prescribing methods, adherence could be overestimated in patients who receive their medication but do not take it and is particularly likely where medication is automatically dispensed at weekly or monthly time intervals. Furthermore, the size of populations which may be identified through prescription records, such as the primary prevention population for CVD, could be underestimated (Sutherland *et al.*, 2018).

Using either type of prescription refill data, the medication possession ratio (MPR) and the proportion of days covered (PDC) are two of the most widely applied methods to estimate a patient's adherence, with no gold standard method identified, and often differences in the specific calculations of the estimates reported (Krueger *et al.*, 2018). The MPR is principally the total number of doses prescribed in a time window divided by the number of doses that should have been administered in a time window and can therefore give values greater than one if early refilling has occurred between prescriptions. In contrast, the PDC, whilst similar, ignores any overlaps that may occur between prescriptions and as a result is truncated at the value of one (Ho, Bryson and Rumsfeld, 2009; Krueger *et al.*, 2018). However, these measures can only be applied in closed pharmacy systems (Osterberg and Blaschke, 2005; Ho, Bryson and Rumsfeld, 2009). Furthermore, underestimation can arise in situations where electronic health records are not complete. For example, one study from the US found that in 1,000 patients randomly sampled from routine clinical care, the proportion who had matching electronic health records with the medications detected within their blood was less than half (Sutherland *et al.*, 2018). Nonetheless, these methods only allow clinicians to gauge if a patient is in possession of the medication and not whether they are taking it correctly or, indeed, at all (Ho, Bryson and Rumsfeld, 2009).

1.4.2 Adherence to Cardiovascular Medications

Patients with chronic diseases are less likely to be adherent to their medication, especially if the disease of interest is asymptomatic, when compared to those receiving short-term prescriptions for symptomatic conditions. Therefore, as lipid-lowering medications should be prescribed to patients indefinitely and, in primary prevention, the patient has experienced no symptoms of high cholesterol, non-adherence to their medication is increasingly likely as the time from first administration increases (Ho, Bryson and Rumsfeld, 2009). In secondary prevention, adherence to all cardiovascular medications has been found to be significantly higher than in primary prevention (Naderi, Bestwick and Wald, 2012). Furthermore, two meta-analyses of prospective studies identified no substantial differences in adherence between the classes of medication prescribed to secondary prevention patients (Naderi, Bestwick and Wald, 2012; Chowdhury *et al.*, 2013). This may suggest that, especially in clinical scenarios where polypharmacy is recommended, non-adherence may not be as a consequence of a specific medication, rather as a result of patient or condition factors.

Estimates of the proportion of patients adhering to statin medication vary between 54% and 71%, with adherence rates decreasing as the length of follow up increases (Naderi, Bestwick and Wald, 2012; Chowdhury *et al.*, 2013; Chen *et al.*, 2019). This rate of decrease was examined more closely in over 150,000 patients with established CVD in the Taiwan National Health Insurance database between 2006 and 2012. It was found that although adherence did decrease over the seven years, the largest decrease was in the first year after statin initiation, with the rate nearly plateauing between years three and six of follow up (Chen *et al.*, 2019). This pattern has also been observed in secondary and primary prevention populations in the UK and Finland (Nordstrom *et al.*, 2015; Lavikainen *et al.*, 2016).

Characterising patients who are likely to not adhere to their medication is not straightforward. However, non-adherence to cardiovascular medication has been consistently associated with being female (Vupputuri *et al.*, 2016; Rodriguez *et al.*, 2017; Chen *et al.*, 2019; Hope *et al.*, 2019) and not white (Vupputuri *et al.*, 2016; Rodriguez *et al.*, 2017; Hope *et al.*, 2019). The relationship with age is

also consistently observed in the literature but not a linear association. Younger and older patients are more likely to be non-adherent (Vupputuri *et al.*, 2016; Rodriguez *et al.*, 2017; Chen *et al.*, 2019; Hope *et al.*, 2019), with adherence peaking at approximately 65 years of age (Hope *et al.*, 2019). Additionally, it is not just the patient's characteristics that have been shown to influence a patient's adherence. Statin prescriptions issued by a cardiologist and further appointments with a cardiologist have also been associated with increased adherence to medication (Vupputuri *et al.*, 2016; Chen *et al.*, 2019).

The patient's number of comorbidities or other prescriptions and their association with adherence is not clear (Hope *et al.*, 2019). Patients with depression (Chen *et al.*, 2019; Hope *et al.*, 2019), renal impairment, liver damage, and COPD (Chen *et al.*, 2019) are less likely to be adherent to statin medication, whilst patients with hypertension and diabetes are more likely (Hope *et al.*, 2019). However, the use of insulin has been associated with non-adherence (Chen *et al.*, 2019), suggesting the relationship between adherence and diabetes is more complex, especially as patients with diabetes have been shown to be more likely to achieve the guideline-recommended cholesterol targets (Danese *et al.*, 2017).

1.4.3 Impact of Adherence on Guideline Targets and Patient Outcomes

Cardiovascular medication adherence and the achieving of LDL cholesterol targets have been examined in CPRD, a primary care database from selected GP practices across the UK, by Danese et al (2017). Over 24,000 patients were identified following a hospitalisation with a cardiovascular event, with prescriptions for lipid-lowering medications (LLMs) issued in the preceding six months. At twelve months following the event, the percentage of patients achieving an LDL cholesterol result <1.8mmol/l increased from pre-event levels (by 5-9%), with patients with diabetes more likely to reach the target than those without. Adherence, defined as >80% MPR, was high, with over 70% adherent in the first year of follow up, although persistence (taking the medication without any breaks) with LLMs was considerably lower with only 50% achieving this (Danese *et al.*, 2017).

Some of this increase in the achievement of targets is likely to be as a result of the increase of the intensity of the treatment following a cardiovascular event. Indeed, it was observed that between 12% and 16% of patients had the intensity of their statin therapy increased (which consisted of either a higher dose or a change to a more potent statin) following their event, with many others augmenting therapy with ezetimibe or fibrates (Danese *et al.*, 2017). Nevertheless, an increase in the intensity of therapy alone is unlikely to increase the likelihood of achieving LDL targets in patients with poor adherence to medication. For example, in over 1,000 secondary prevention patients in Georgia, USA, those with adherence <50%, estimated with PDC, were at 88% greater risk of not achieving a reduction of 30% in LDL cholesterol (Vupputuri *et al.*, 2016).

However, an interaction between the intensity of therapy and adherence to therapy could seem plausible, particularly if the intensity of therapy is viewed as a surrogate measure for severity of illness. This has been found in both the primary and secondary prevention of CVD, with patients with higher numbers of comorbidities, or more severe comorbidities, likely to be adherent to LLM (Vupputuri *et al.*, 2016; Danese *et al.*, 2017; Hope *et al.*, 2019). In contrast, in a primary prevention systematic review, whilst increased comorbidities were associated with an increased likelihood of being adherent, the intensity of statin was found to have a negative association, possibly as a result of the risk of adverse events (Hope *et al.*, 2019). In secondary prevention populations though, the overall benefit of LLM may negate this. For example, Khunti et al (2018) used CPRD to examine 16,000 patients with established CVD and found that those prescribed higher intensity therapies were more likely to be adherent (Khunti *et al.*, 2018).

Furthermore, higher adherence to LLM has been associated with a decreased likelihood of cardiovascular events. In patients with established CVD, a 10% increase in adherence was found to be associated with a 5% reduction in the risk of the patient experiencing further events (Khunti *et al.*, 2018). This was also shown in the reduction of the risk of stroke in a meta-analysis of primary and secondary prevention populations (n>700,000 patients), where a 20% increase in

adherence was associated with an 8% lower risk of any stroke (ischaemic or haemorrhagic) (Xu *et al.*, 2017).

Overall, there is some evidence to suggest that many patients who are adherent in the first year remain so beyond this. In one study of >60,000 established CVD patients in CPRD, 71.3% were adherent (MPR >80%) two years after diagnosis, down from 74.8% at the end of the first year (Nordstrom *et al.*, 2015). This pattern has also been observed in a primary prevention population in Finland (n>40,000), where 63.8% of those who were adherent (PDC >80%) in the first year of treatment, remained so for three years. This pattern was stronger in non-adherers, where 73.7% of those who were non-adherent in the first year, were non-adherent for the three years of follow up (Lavikainen *et al.*, 2016). This would suggest that few patients alternate between being adherent and non-adherent to LLMs, with higher rates of discontinuation most likely in the first year of follow up. Therefore, non-adherence may pose a greater problem in secondary prevention populations where there is a greater risk of further cardiovascular events in the first year after an event (Khunti *et al.*, 2018).

1.5 Developments within the IT Structure in the NHS

1.5.1 Linkage of Routine Healthcare Data

Extensive quantities of healthcare and demographic data are collected routinely within various parts of the NHS. When these sources are linked together they can prove hugely beneficial for research, with various projects demonstrating the potential impact that routine linked data could have for both the patients and the health service (Barker *et al.*, 2018; The Health Foundation, 2018; Lloyd *et al.*, 2019). For example, when a survey of the confidence in patients to manage their chronic conditions was linked with primary and secondary care data, it revealed that those more confident had a lower usage of health services, including Accident and Emergency admissions (Barker *et al.*, 2018). In a separate project, the identification of patients living in care homes from GP records facilitated the introduction of enhanced care packages which subsequently reduced the number of recipients' emergency admissions to hospital (Lloyd *et al.*, 2019). Routine healthcare data linkage can also facilitate longer and accurate follow up periods for studies relatively inexpensively, and with minimal

impact on the participants. Furthermore, when multiple sources of routine healthcare data are linked to create a patient's electronic health record, their engagement with various healthcare services can also be easily followed (NHS Research Scotland, 2015a; Wellcome Trust, 2015).

With these benefits identified, there is a growing aspiration to implement linkage between health record datasets routinely, such as through the Integrated Care Programmes in England. This not only aims to improve care decisions through the consistent sharing of health data, but also aims to enable the development of system interface standards, resulting in the standardisation of electronic health records and, consequently, improving the consistency of recording between partnerships (NHS Digital, 2018).

1.5.2 Information Contained within NHS Scotland Data

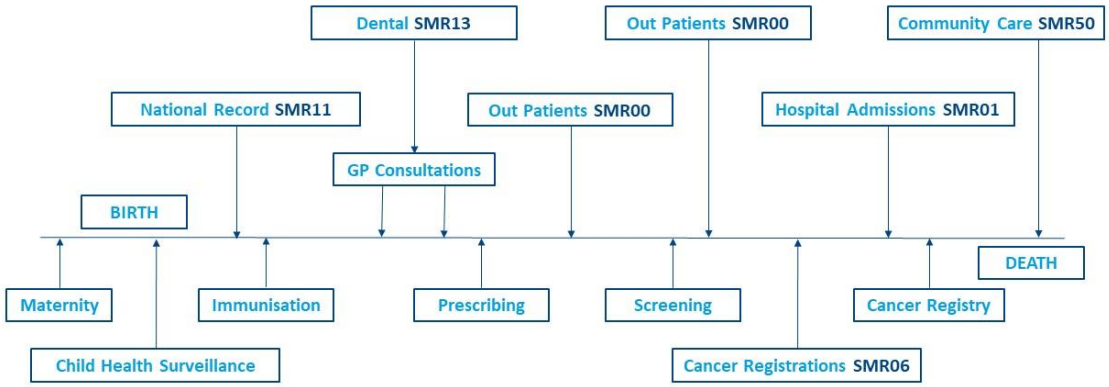


Figure 1.4: Health Data Captured Across the Lifespan in Scotland and the associated datasets (NHS Research Scotland, 2015a)

In Scotland, data linkage between a patient's health records from different services is made possible using a patient's Community Health Index (CHI) number. This is entered in the capture of all health-related activities for the individual including prescriptions and test results, as well as all appointments and surgeries. The exact repositories where this information from across an individual's life span is stored is shown in Figure 1.4. As a result of this linkage capability, rigorous anonymisation, and complete population coverage, health and biomedical research on this data is common (NHS Research Scotland, 2015a). Access to such data for research purposes is usually provided through data safe havens, which are available for regions of Scotland and nationally.

These provide secure, confidential, and non-identifiable access to such datasets to trained researchers and allow linkages between health and non-health data for analysis (NHS Research Scotland, 2015b).

For clinical practice, NHS Scotland produced an eHealth Strategy for 2014-2017, and developed Scotland's Digital Health and Care Strategy in 2018 (The Scottish Government, 2015, 2018). This original strategy provided an update on the current progress and modified vision towards the streamlining of the various components of a patient's electronic health record, with the goal of combining the information and making it available in one portal, with a parallel system available for patients to view their data. It is hoped that this would ultimately include all interactions between a patient and the health and social care services, and in the required level of detail for the professional accessing it. It was anticipated that this would be achieved by 2020 (The Scottish Government, 2015).

1.6 Plasma Lipid Testing

1.6.1 Incidence and Costs of Testing

In 2010, a review for the Department of Health estimated that tests in patients for chronic diseases, including CVD, accounted for 50% of all pathology work in the UK (Department of Health, 2010), with the number of tests requested increasing on an annual basis (Smellie, 2012; O'Sullivan *et al.*, 2018). In an analysis of more than 250 million tests in 11 million patients from the CPRD database between April 2000 and March 2016, the age and sex-adjusted rate of laboratory testing increased by an average of 8.7% annually, with 44,847 tests per 10,000 person-years in 2015/16 (O'Sullivan *et al.*, 2018). This is broadly comparable to an estimate from a 2004 survey of UK laboratories which reported an annual rise of 10% in workload over the preceding three years, with the majority of this rise accounted for by primary care workload (Smellie, 2012).

The increasing number of tests cannot be completely accounted for by an increase in patient numbers. Indeed, in the analysis of testing patterns in CPRD, an increasingly greater proportion of tests were ordered for fewer patients, with 32.7% of tests ordered for patients with more than 10 tests annually in 2015/16

compared to just 9.5% in 2000/01. The Quality Outcomes Framework (QOF), which ran between 2004/5 and 2015/6 in Scotland, but continues in other parts of the UK, and incentivised the monitoring of those with chronic diseases may partially explain this. Tests included in the QOF had a higher average annual percentage increase than those not included in the QOF (9.8% vs 7.4%). However, in terms of CVD, despite lipid tests being included in QOF, their annual percentage testing increase was lower than other QOF tests (6.6%). Furthermore, the annual rate of testing over the sixteen years showed an increase in testing rate which then decreased resulting in an inverted U-shaped temporal trend in the number of tests performed each year (O'Sullivan *et al.*, 2018).

Whilst an individual test has been conservatively estimated to cost approximately £6 including some direct staff and processing costs (National Institute for Health and Care Excellence, 2015), the volume of tests being conducted represents a significant expenditure within the NHS, with estimates of £1.8 billion for all laboratory testing in 2015/16, not including administrative and reviewing costs (O'Sullivan *et al.*, 2018).

The burden of an increase in testing is not solely a financial one and will also have a substantial impact on primary care workload, both in terms of consultation time and the associated administration. For example, one study estimated that if each test result took two to three minutes to be reviewed, this would translate to up to two hours a day (O'Sullivan *et al.*, 2018), without considering the time needed for any follow up consultations with patients. This is broadly in line with Thompson and Walter (2016) who, using data from the Eighth National GP Worklife Survey in 2015 (Gibson *et al.*, 2015), estimated that work outside of consultation hours was responsible for an additional eight or more hours for an average GP a week (Thompson and Walter, 2016). Furthermore, the number and duration of consultations also increased significantly between 2007 and 2014 (by 10.5% and 6.7% respectively), leading to a workload increase equivalent to 16% over the seven years for GPs (Hobbs *et al.*, 2016). This has led to widespread and justified beliefs that workload in primary care is increasing unsustainably, resulting in difficulties in the recruitment and retention of staff, burnout, and concerns that the problem will

continue to worsen as the populations grows and ages (Hobbs *et al.*, 2016; Thompson and Walter, 2016).

1.6.2 Inappropriate Lipid Testing (Over and Under)

As numbers of laboratory tests conducted are consistently shown to be increasing annually and carry with them a significant associated burden on the health service (O'Sullivan *et al.*, 2018), the question as to the necessity and appropriateness of the testing is raised. However, the definition of an inappropriate lipid test is not universal within the literature, and many authors argue that the absence of the clinical context for a test in databases used for such analysis may have resulted in its misclassification. As a result, estimates of the proportion of inappropriate testing in lipids ranges from 10-42%, with definitions of an inappropriate test focussing on the frequency and intervals between repeated tests (Doll *et al.*, 2011; Morgen and Naugler, 2015; Chami *et al.*, 2017; Hajati *et al.*, 2018).

Two analyses from Canada both defined a lipid test as inappropriate if it took place within three months of another test, in line with the recommendations in the Canadian guidelines, and reported similar results (Morgen and Naugler, 2015; Chami *et al.*, 2017). Chami et al (2017) looked at 4.6 million lipid tests in 3.5 million patients from the Ontario Health Insurance plan between 2006 and 2010 and found that 10.2% of tests appeared to be inappropriate, with 9.2% of patients having at least one inappropriate test in the period (Chami *et al.*, 2017). Meanwhile, an analysis of 100,000 randomly selected patients in the laboratory services data in Calgary in 2010 who were followed up for a year reported that 10.5% of tests were inappropriately repeated within 12 weeks of an earlier test, and after 6.5 months a quarter of the population had been retested (Morgen and Naugler, 2015).

However, in the UK, guidelines recommend monitoring cholesterol levels in high-risk patients annually (NICE, 2014; SIGN, 2017), rather than quarterly as was assumed by Morgen and Naugler (2015) and Chami et al (2017). An analysis of 1.6 million tests from a subset of the Medicare Benefits Schedule in Australia from 2008 to 2014 defined more than one HDL cholesterol test per year to be unnecessary. This revealed a slight but steady decrease in the proportion of

over-testing during the seven years (20.8% to 16.9%) but estimated that the economic burden of this testing was approximately $A4.3 million each year. Furthermore, whilst more tests may be necessary following the commencement of adjustment of lipid-lowering medication, Hajati et al (2018) only included patients who had reported a stable prescription, and therefore tests conducted for these purposes are unlikely to have been included in these estimates. Despite this, in a sensitivity analysis where the number of tests considered necessary was increased to two, 6% of tests remained unnecessary (Hajati *et al.*, 2018).

Nonetheless, all of these estimates are considerably lower than estimates from a 20-year period in Oxfordshire, which defined a test as unnecessary if more than three tests were reported within three years (i.e. more than annual), and found the rate to be increasing across the two decades with the highest rate of 42% observed in the final period between 2005 and 2007 (Doll *et al.*, 2011). However, regardless as to which estimate or definition is used, it is clear that over-utilisation could be contributing to the burden placed in the health system as the demand for testing continues to grow in primary care.

Furthermore, whilst reducing the number of tests conducted more than those recommended, the guidelines may reduce the impact of such testing. One suggestion to further mitigate this could be to scale back the number of tests recommended in guidelines. Indeed, there is some evidence to suggest that this may already be happening. In Australia, it was found that in any given year between 2008 and 2014, nearly half of patients on stable lipid-lowering medication did not have their HDL cholesterol checked, and consequently, an estimated $A11 million was saved in laboratory expenditure (Hajati *et al.*, 2018). However, the reasons for these being missed were unknown (Hajati *et al.*, 2018), and the long-term consequences of this could ultimately have resulted in a greater burden for both the patients and the health service if opportunities were missed to reduce the potential number of cardiovascular events (Perera *et al.*, 2015).

As a result, a careful balancing act must be performed when deciding when a test is necessary. Testing patients too often can lead to an increased likelihood of misleading results due to biological variation in cholesterol levels within an

individual. This could result in potentially over- or under-treating patients (Perera *et al.*, 2015), in addition to the resource implications discussed in Chapter 1.6.1. In contrast, testing too infrequently could increase the risk of cardiovascular events with a significant impact on families and the health service (Perera *et al.*, 2015). In balancing these scenarios, a Health Technology Assessment funded by the National Institute for Health Research (NIHR) found that, for patients considered to be at high risk of CVD, annual reviews of lipid levels represented the optimal benefit-risk ratio (Perera *et al.*, 2015), in line with the recommendation in current national guidelines (NICE, 2014; SIGN, 2017).

1.6.3 Implementation of Personalising Testing Schedules

Whilst annual reviews may be optimal for the high-risk population as a whole, it is likely that there is a group of patients at each annual review for whom their lipid test result will not result in a change in the patient's management of their cardiovascular risk. Therefore, the successful identification of these patients would allow the overall number of tests conducted to be reduced without increasing the number of cardiovascular events experienced within the population. Routinely collected data, such as a patient's prescription collection records, their demographics, and their previous test results, which all form part of a patient's electronic health record (Chapter 1.5.2), could all potentially facilitate the accurate identification of this sub-population at any given time.

The idea of personalising monitoring schedules is not novel, but research has, thus far, largely been confined to screening programmes including diabetic retinopathy, and cancers such as cervical and breast. In the case of cervical cancer screening programmes, such approaches are in the early stages of development: one analysis sought to develop risk scores in patients with abnormalities detected (Kyrgiou *et al.*, 2016), and another has attempted to devise a risk-stratification for all patients (Baltzer *et al.*, 2017). Both, however, recognise that further research is needed before either could be implemented safely, be clinically useful, or cost-effective (Kyrgiou *et al.*, 2016; Baltzer *et al.*, 2017). Within breast cancer, models have shown that personalised risk-management is cost-effective and results in more favourable harm-benefit ratios through the reduced likelihood of false-positives and overestimation of cancer-

incidence (Vilaprinyo *et al.*, 2014; Pashayan *et al.*, 2018). However, the accurate determination of an individual's risk category, and their corresponding risks of overdiagnosis, remain key areas for further research before any widespread implementation is likely to be successful (Vilaprinyo *et al.*, 2014; Román *et al.*, 2017; Pashayan *et al.*, 2018).

In contrast, the literature seeking to personalise screening schedules for diabetic retinopathy is better established and has analysed large datasets in various countries, including Denmark (Aspelund *et al.*, 2011), Scotland (Looker *et al.*, 2013), Wales (Thomas *et al.*, 2012), and England (Eleuteri *et al.*, 2017). The largest of these studies involved data from 150,000 patients over six years from the Scottish Diabetic Retinopathy Screening Programme, which found that, for patients with Type 2 diabetes, two consecutive clear scans were sufficient to recommend biennial screening with minimal additional risk (Looker *et al.*, 2013). This was also supported by data from nearly 50,000 patients in Wales, where a clear scan in patients with Type 2 diabetes was found to be adequate evidence to extend the screening interval beyond the standard 12 months. However, it also reported that patients on insulin and with a duration of disease longer than ten years should continue with annual scans, irrespective of their previous scan result (Thomas *et al.*, 2012).

The accurate identification of those who could benefit from reduced monitoring is not the sole step in this process. Consideration also needs to be given towards the duration of the extended intervals that could be implemented safely without increasing the patients' risks of negative outcomes. For example, in cervical screening, the introduction of the human papillomavirus vaccines and testing has led to the development of models derived from national screening programmes. These, in turn, have suggested that those with negative results could safely benefit from screening at reduced intervals from current guidelines, with comparable risks (Dijkstra *et al.*, 2016; Bains *et al.*, 2019).

Meanwhile, in diabetic retinopathy, some models have looked at the effects of implementing a range of intervals depending on the calculated risks. Both the analysis of the Danish Diabetes database (Aspelund *et al.*, 2011) and the Liverpool Diabetic Eye Screening Programme (Eleuteri *et al.*, 2017) calculated a patient's risk score and offered a range of screening intervals. Aspelund et al's

(2011) risk score, developed using data from 5,000 Danish patients with 20 years of follow up, included several clinical predictors of retinopathy: type and duration of diabetes, HbA1c, blood pressure, and presence and grade of retinopathy at the previous scan. The corresponding risk score was then used to output a screening interval for the patient ranging between 6 months and 5 years, with the average recommendation for the population between 2 and 2.5 years (Aspelund *et al.*, 2011). Similarly, Eleuteri et al (2017) recommended screening intervals of 6, 12, or 24 months based on their risk score which was derived from a 5 year follow up of approximately 12,000 patients in Liverpool. This score was calculated based on the duration of disease, HbA1c, systolic blood pressure, age, and total cholesterol (Eleuteri *et al.*, 2017).

The cost-effectiveness of any reduced monitoring schedule should also be considered, and within diabetic retinopathy, the predicted impact of such methods on resourcing varies between studies. For example, the Danish risk score corresponded to the greatest reduction in visits, at 59% (Aspelund *et al.*, 2011), whilst a simulation of UK data, where a clear scan was sufficient to extend the screening interval to two years, yielded a 25% reduction in costs and went on to suggest that such a transition was also safe and cost-effective (Chalk *et al.*, 2012).

However, a systematic review published in 2016, exercised caution concerning extending the screening interval for diabetic retinopathy, concluding that whilst there was a minimal difference in clinical outcomes for annual and biennial screening for low-risk patients, the poor quality of the evidence due to the experimental designs used limited the ability to draw reliable conclusions (Taylor-Phillips *et al.*, 2016). Indeed, whilst Apselund et al's (2011) algorithm has now been validated in many diabetes registers, concern has also been raised regarding missed opportunities (Van Der Heijden *et al.*, 2014; Soto-Pedre, Pinies and Hernaez-Ortega, 2015; Lund *et al.*, 2016). This is especially true for patients for whom longer screening intervals were recommended, where a patient's risk profile could change substantially (McGhee, Harding and Wong, 2012). Therefore, when Lund et al (2016) validated this algorithm, a maximum of two years between screenings was used and resulted in a similar risk profile to annual screenings and a reduction in costs of 40%. This represented a reasonable

and cautious compromise towards reduced screening, with the authors reasoning that the risks could be further minimised if linked to real-time clinical records to detect changes in patients' risk profiles (Lund *et al.*, 2016).

Nonetheless, further steps are also needed to validate the evidence for the safe implementation of reduced testing with the systematic review of the evidence strongly advocating the need for clinical trials in this area (Taylor-Phillips *et al.*, 2016), and one such trial has since been designed for the Eleuteri et al's (2017) Liverpool risk score (Broadbent *et al.*, 2019).

1.7 Overall Summary

Risk stratification of the population, using established risk factors, forms a crucial component of the clinical guidelines for the management and prevention of CVD. For those at high risk, lipid-lowering therapies, principally statins, are recommended, with other classes of medication including antihypertensives, anticoagulants, and antiplatelets also recommended for subpopulations of those at high risk where clinically indicated and the benefit-risk ratio is favourable. This current approach is broadly similar to the previous guidelines issued ten years prior. Despite a difference in threshold defining high risk (20% vs 10%), the guidelines are also comparable to those adopted by the rest of the UK.

Medication represents a key component of the management of those at high risk, but, due to the chronic nature of CVD, long-term adherence can be difficult to achieve. Perhaps unsurprisingly, those with established disease are more likely to be adherent than those without, although, as in primary prevention, many patients discontinue their medication within the first few years of commencing it, thereby reducing the potential benefits. However, adherence to LLM has reassuringly been linked with an increased likelihood of achieving cholesterol targets, as well as reducing the likelihood of cardiovascular events and mortality. This confirms the clinical benefits of such medication and reinforces the need to encourage compliance among high-risk patients.

Annual reviews with the patient, which are likely to include a lipid test, can be used as a mechanism to facilitate this, as are recommended by current guidelines. However, the number of tests is increasing year on year, as is the

workload for GPs. Whilst some tests are likely occurring in excess of the guideline's recommendations, it is plausible that for some patients the annual review would not improve their adherence or alter clinical decisions for their risk management. As a result, being able to identify such patients accurately could reduce the burden of such reviews in primary healthcare, and data captured within a patient's electronic health record could be utilised to enable this.

1.7.1 Current Gaps in the Literature

Whilst many of the risk factors for CVD and the use of LLMs in managing patients' risks are well-established, there are some areas within their prevention and risk management where the evidence is less conclusive. For example, the use of statins to lower a patient's cholesterol is likely to be universal worldwide, but there has been a lack of consensus over the use of lipid targets. Nonetheless, these targets may be one reason for a patient to be reviewed frequently, although it could also be used as an opportunity to assess a patient's adherence to their medication. Increasing evidence from electronic prescribing records has shown that adherence to LLM decreases with treatment duration. However, the evidence base that such reviews translate into improved adherence, lipid levels, and consequently patient outcomes when considered together, is also minimal with often relatively short durations of follow up.

With the exception of one simulation study (Perera *et al.*, 2015), there is also little evidence for optimal testing and review schedules of such patients, and as a result any guideline recommendations are likely to be derived from expert opinion. As evidence suggesting that the burden associated with these reviews is growing, there is a clear need to establish a scientific basis for them and their frequency, as well as explore options that could safely reduce their number.

One such option is the personalisation of a patient's monitoring, which would necessitate the identification of those likely to require continued frequent reviews. To date, research looking to describe patients who are likely to become non-adherent or not achieve lipid targets have successfully identified some possible characteristics. Despite this, for many other potential factors, the associations are less clear and often contradictory, with patients' adherence

behaviours likely multifactorial and possible variables confounding with each other. Consequently, there is still a need to pursue the identification of such individuals and investigate the plausibility and effectiveness of any personalised schedules.

1.8 Thesis Overview

1.8.1 Aims and Objectives

This thesis aims to:

1. Assess the current guidance, within international guidelines, for the secondary prevention population for cardiovascular disease regarding lipid-lowering treatment, cholesterol targets (and if used), and frequency of monitoring as part of long-term follow-up.

2. Understand the demographics of the post myocardial infarction population in Greater Glasgow and Clyde, together with their adherence to statins and achievements of cholesterol-lowering targets.

3. Explore the associations between statin adherence, achievements of cholesterol targets, further myocardial infarctions, and mortality.

4. Examine the associations between demographic factors, and previous statin adherence and lipid results, with the likelihood of statin non-adherence or non-target lipids.

5. Identify patients who would not require any intervention in response to a lipid test obtained during the annual review process and could, therefore, benefit from biennial testing.

With these aims in mind, specific objectives were also established, with each numbered aim corresponding to the equivalently numbered objective. Therefore, the objectives of this thesis are:

1. Conduct a systematic review of international guidelines relating to the management of those with established cardiovascular disease to compare

recommendations for lipid-lowering treatments, the use of and any specific cholesterol targets, and long-term monitoring.

2. Utilise linked data from the NHS Greater Glasgow and Clyde Safe Haven to establish a post myocardial infarction cohort, and use descriptive statistics to summarise their demographics, adherence to statins, and achievements of cholesterol targets.

3. Utilise data from this post myocardial infarction cohort to describe and quantify the strength and direction of the associations between statin adherence, the achievement of cholesterol targets, further myocardial infarctions, and mortality.

4. Utilise the post myocardial infarction cohort data to identify, describe, and quantify the strength and direction of the associations between demographic factors, previous statin adherence, and previous lipid results, with statin non-adherence and failure to achieve cholesterol targets.

5. Develop an algorithm, using factors identified through the earlier objectives, which identifies patients for whom the results of an annual lipid test would not require any intervention, and therefore could safely benefit from biennial testing.

The aims and objectives outlined here will each be addressed within the subsequent chapters, as outlined in further detail in Chapter 1.8.3 below.

1.8.2 Hypotheses

With the aims and objectives of this thesis now outlined, the following hypotheses were generated:

1. In secondary prevention populations, the use of statins as a lipid-lowering therapy will be universal in current guidelines due to the strong evidence base for their use (Chapter 1.3.1). However, there will be some variation in the use, and nature, of cholesterol targets, and in the frequency of monitoring these patients.

2. The demographics, statin adherence, and achievement of cholesterol targets of the post myocardial infarction cohort derived from data within the NHS Greater Glasgow and Clyde Safe Haven will be broadly similar to other post-myocardial infarction and secondary prevention cohorts.

3. There will be significant positive associations between statin adherence and the achievement of cholesterol targets, and both, in turn, will be positively significantly associated with further myocardial infarctions and mortality.

4. A patient's previous statin adherence and previous lipid results will be strongly positively associated with their current adherence and lipid results. Some demographic factors will also be significantly associated but these associations will be smaller than those observed with previous data.

5. There will be a group of patients identified by an algorithm, utilising associations identified from earlier objectives, who could be safely considered for biennial lipid testing.

1.8.3 Structure of the Thesis

The rest of this thesis will systematically cover each of the aims, objectives, and hypotheses listed above. Chapter 2 addresses the first of these and contains a systematic review of secondary prevention guidelines published during the past decade and qualitatively compares the recommendations surrounding the management of lipids within this population. The work outlined in Chapter 3 details the methods and the decisions implemented as part of the cleaning of the data from the NHS Greater Glasgow and Clyde (GGC) Safe Haven and the derivation of the post MI cohort. Whilst this does not directly address any of the specific aims listed above, this work was necessary to facilitate the remainder of this thesis with it all depending upon the successful derivation of this cohort. Using the cohort derived in the previous chapter, Chapter 4 seeks to address the second aim using a series of descriptive statistics to characterise the post MI population and includes their demographics, lipid testing frequency, lipid test results, and estimated statin adherence. The third aim is investigated in Chapter 5, where the nature of any associations between statin adherence and lipid testing results with further MIs and mortality are elucidated. Both final two aims

are covered within Chapter 6, which first identifies significant predictors of non-adherence or non-target plasma lipids (and the direction of these associations). These results are then used to develop an algorithm to identify patients who are likely to meet targets with reasonable adherence to LLM, and as such could safely benefit from a reduced testing schedule. Finally, Chapter 7 provides a general discussion and summary of the results generated as part of this thesis as well as outlining their key conclusions.

Chapter 2 Systematic Review of Clinical Guidelines for Lipid-Lowering in the Secondary Prevention of Cardiovascular Disease Events

This chapter is reproduced from a manuscript published in Open Heart with a CC-BY 4.0 license in 2020 (Brown, Welsh and Logue, 2020). However, the supplementary materials are included within this chapter and not as a separate appendix, as published, and the text has been updated to reflect this. Therefore, the example search strategy has been added into Chapter 2.2.1.1, and the supplementary table of recommendations for lipid lowering drugs other than statins is included as Table 2.4 in Chapter 2.3.3.1. The full reference for the manuscript is:

Brown, R. E., Welsh, P. and Logue, J. (2020) 'Systematic review of clinical guidelines for lipid lowering in the secondary prevention of cardiovascular disease events', *Open Heart*, 7(2), p. e001396. doi: 10.1136/openhrt-2020-001396.

2.1 Introduction

In 2016, cardiovascular disease (CVD) was the leading cause of death worldwide and was responsible for an estimated 17.9 million deaths, with heart attacks and strokes accounting for 85% of these (World Health Organization, 2017b). This is broadly similar to the Global Burden of Disease study's estimate for 2015 of 17.92 million deaths, which additionally estimated that the number of cases that year was 422.7 million (Roth *et al.*, 2017). With the cardiovascular death rate falling between 1990 and 2015 in most high-income countries (Roth *et al.*, 2017), there is an increasing focus on the management and risk prevention of CVD in the secondary prevention setting.

The need for risk management in secondary prevention, which encompasses coronary heart disease, stroke, and peripheral artery disease, is clear. The rate of further cardiovascular events per annum in unmedicated patients with previous events has been estimated to be around 5.6% and 3.7% depending on whether the previous event was coronary heart disease related, compared to 1.8% in those without (Cholesterol Treatment Trialists' (CTT) Collaboration,

2010). The mortality rate is also six times higher in this population (World Health Organization, 2017a). As a result of this elevated risk, the World Health Organization states that individuals with established cardiovascular disease should be treated with lipid-lowering therapy, aspirin, beta-blockers and ACE inhibitors, as well as engage in smoking cessation to reduce the risk of further events by up to 75% (World Health Organization, 2017b). This has led to the use of statins as lipid-lowering therapy being considered a cornerstone of clinical practice in order to manage secondary CVD risk throughout the world, due to their relative safety, cost, efficacy in lowering cholesterol and consequently CVD prevention. Additionally, there is no threshold beyond which cholesterol-lowering is considered dangerous (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010).

However, the World Health Organization has offered no specific guidance regarding target plasma lipid levels or the monitoring of these since the publication of their prevention of cardiovascular guidelines in 2007 (World Health Organization, 2007). Indeed, different national or international guidelines have contrasting recommendations. For example, the use of on-treatment cholesterol targets has proved controversial in recent years (Leibowitz *et al.*, 2017; Pallazola *et al.*, 2018). Prescriptions of higher doses of lipid-lowering therapy are more likely in the pursuit of increasingly lower lipid targets with the aim of reducing a patient's risk of further cardiovascular events (Allahyari *et al.*, 2020). However, higher doses of medication also lead to an increased likelihood of side effects (Cannon *et al.*, 2006; Preiss *et al.*, 2011), which could result in further costs or even non-adherence or discontinuation in patients, reducing the potential reduction in risk. The long-term follow-up of lipids in high-risk populations poses a significant burden of time to patients (Kim *et al.*, 2014), and costs to healthcare. Specifically, increased biochemistry costs from expanding clinical demand have been flagged as a major financial burden (Smellie, 2012). Therefore, effectively balancing the costs of follow-up with the reduction in cardiovascular risk within a given population represents a significant challenge, with countries and regions likely to have differing approaches.

This systematic review aimed to investigate similarities and differences in clinical guidelines surrounding the recommendations for the therapeutic

treatment, targets, and monitoring of lipid risk factors in adults who have established cardiovascular disease. This will help to highlight variation in the guidelines, thereby providing guidance for future research priorities.

2.2 Methods

A protocol which documented the pre-specified analysis and the inclusion criteria for this systematic review was first registered on PROSPERO (https://www.crd.york.ac.uk/prospero/) on 19th June 2018 [Ref: CRD42018098582] (Brown, Welsh and Logue, 2018).

2.2.1 Literature Search

A search of Medline, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Turning Research into Practice (TRIP) databases was conducted for all guidelines published in the ten years prior to 31 December 2019. In addition, several guideline specific databases were searched: National Guideline Clearinghouse (USA), the National Library for Health Guidelines Finder (UK), the Canadian Medical Association Clinical Practice Guidelines Infobase and Guidelines International Network International Guideline Library. Finally, an additional hand search was performed to identify the most recent versions of the guidelines identified through the systematic search. A copy of the search strategy used for MEDLINE is included in Chapter 2.2.1.1 with comparable searching strategies employed for other searched databases.

2.2.1.1 Example MEDLINE Search Strategy

((Cardiovascular Diseases/) OR (exp Aortic Aneurysm/) OR (exp Myocardial Ischemia/) OR (exp ARTERIOSCLEROSIS/) OR (exp Cerebrovascular Disorders/) OR (Peripheral Vascular Diseases/) OR (exp Heart Failure/) OR ((cardiovascular adj3 disease*).mp.) OR ((coronary adj3 disease*).mp.) OR (heart disease*.mp.) OR ((stroke* or cerebrovasc* or cva*).mp.) OR ((aort* adj5 aneurysm).mp.) OR ((abdominal adj5 aneurysm).mp.) OR ((thoracoabdominal adj5 aneurysm).mp.) OR ((arteri* adj3 occlusi*).mp.) OR ((arteri* adj3 stenosis).mp.) OR ((peripher* adj5 occlusi*).mp.) OR ((peripher* adj5 arteri*).mp.) OR ((peripher* adj5 vascular).mp.) OR (heart failure.mp.) OR (atherosclerosis.mp.) OR (arteriosclerosis.mp.) OR (HYPERTENSION/) OR (Hyperlipidemias/) OR (Diabetes

Mellitus/) OR (hypertension.mp.) OR (hyperlipid?emia.mp.) OR (dyslipid?emia.mp.) OR (cholesterol.mp.) OR (diabetes.mp.) OR (metabolic syndrome.mp.))

AND

((Cardiovascular Diseases/pc [Prevention & Control]) OR (Secondary Prevention/) OR (Risk Assessment/) OR ((established adj3 disease).mp.) OR ((subsequent adj3 event*).mp.) OR ((recurrent adj3 event*).mp.) OR ((multiple adj3 event*).mp.) OR ((secondary adj3 prevention).mp.) OR ((subsequent adj3 episode*).mp.) OR ((recurrent adj3 episode*).mp.) OR ((prior adj3 cardiovascular).mp.) OR ((previous adj3 cardiovascular).mp.) OR ((multiple adj3 episode*).mp.))

AND

((Critical Pathways/) OR (exp Clinical Protocols/) OR (exp CONSENSUS/) OR (exp Consensus Development Conference/) OR (exp Consensus Development Conferences as Topic/) OR (exp GUIDELINE/) OR (Guidelines as Topic/) OR (exp Practice Guideline/) OR (Practice Guidelines as Topic/) OR (Health Planning Guidelines/) OR ((guideline or practice guideline or consensus development conference or consensus development conference, NIH).pt.) OR ((position statement* or policy statement* or practice parameter* or best practice*).ti,ab,kf,kw. OR (standards or guideline or guidelines).ti,kf,kw.) OR (((practice or treatment* or clinical) adj guideline*).ab.) OR ((CPG or CPGs).ti.) OR (consensus*.ti,kf,kw.) OR (consensus*.ab. /freq=2) OR (((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti,ab,kf,kw.) OR (recommendat*.ti,kf,kw.) OR ((care adj2 (standard or path or paths or pathway or pathways or map or maps or plan or plans)).ti,ab,kf,kw.) OR ((algorithm* adj2 (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti,ab,kf,kw.) OR ((algorithm* adj2 (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti,ab,kf,kw.))

NOT ((comment.pt.) OR (letter.pt.) OR (editorial.pt.))

Limit results to English language

Limit results to yr="2010 -Current"

2.2.2 Selection Process

Papers were retained if they met the Institute of Medicine's 2011 definition of a clinical guideline, "Clinical Practice Guidelines are statement that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options" (Institute of Medicine (US), 2011). As the focus of this systematic review was the management of patients with established cardiovascular disease, guidelines were only retained within the review if their specific management was detailed, regardless of whether they covered established cardiovascular disease as a whole, or for the management of patients after a specific event, such as myocardial infarction or stroke. Only the most recent version of the guidelines was retained, with any previous versions removed. Finally, included guidelines had to apply to OECD countries, produced by a professional organisation, and have the full version of the guidelines available in English.

Two reviewers (RB and JL) independently reviewed the titles and abstracts of the results against the eligibility criteria. The same two reviewers also performed the full text review, where the reason for exclusion of the guidelines was also documented. In both instances, any discrepancies of opinion were resolved through discussion.

2.2.3 Data Extraction

Data extraction was performed initially by one reviewer (RB) with accuracy checked by a second reviewer (JL). Data extracted included the target population, the publishing society, the country or region the guideline applied to and the year it was published. Recommendations specifically for the secondary prevention population surrounding the frequency that plasma lipid monitoring should be performed, therapies that should be used, and any lipid target values were also extracted if given within the guideline. The strength and the level of evidence of each recommendation were also extracted. Once extracted, the recommendations were compared by all authors.

2.2.4 Quality Assessment

The quality of the development processes of each of these guidelines was then assessed using the 2009 Appraisal of Guidelines for Research and Evaluation (AGREE) II tool by two reviewers (RB and JL). The AGREE II tool consists of 23 questions covering six domains (Scope and Purpose, Stakeholder Involvement, Rigour of Development, Clarity of Presentation, Applicability, and Editorial Independence) and an overall assessment of the quality of the guideline. Each of the items, including the overall quality assessment, is scored on a 7 point scale (1, Strongly Disagree; 7, Strongly Agree), with scaled domain scores then calculated (AGREE Next Steps Consortium, 2010). As the AGREE II tool does not facilitate an aggregated score across the domains nor a specific cut point for high or low quality, scores for all domains are presented.

2.3 Results

2.3.1 Results of Literature Search

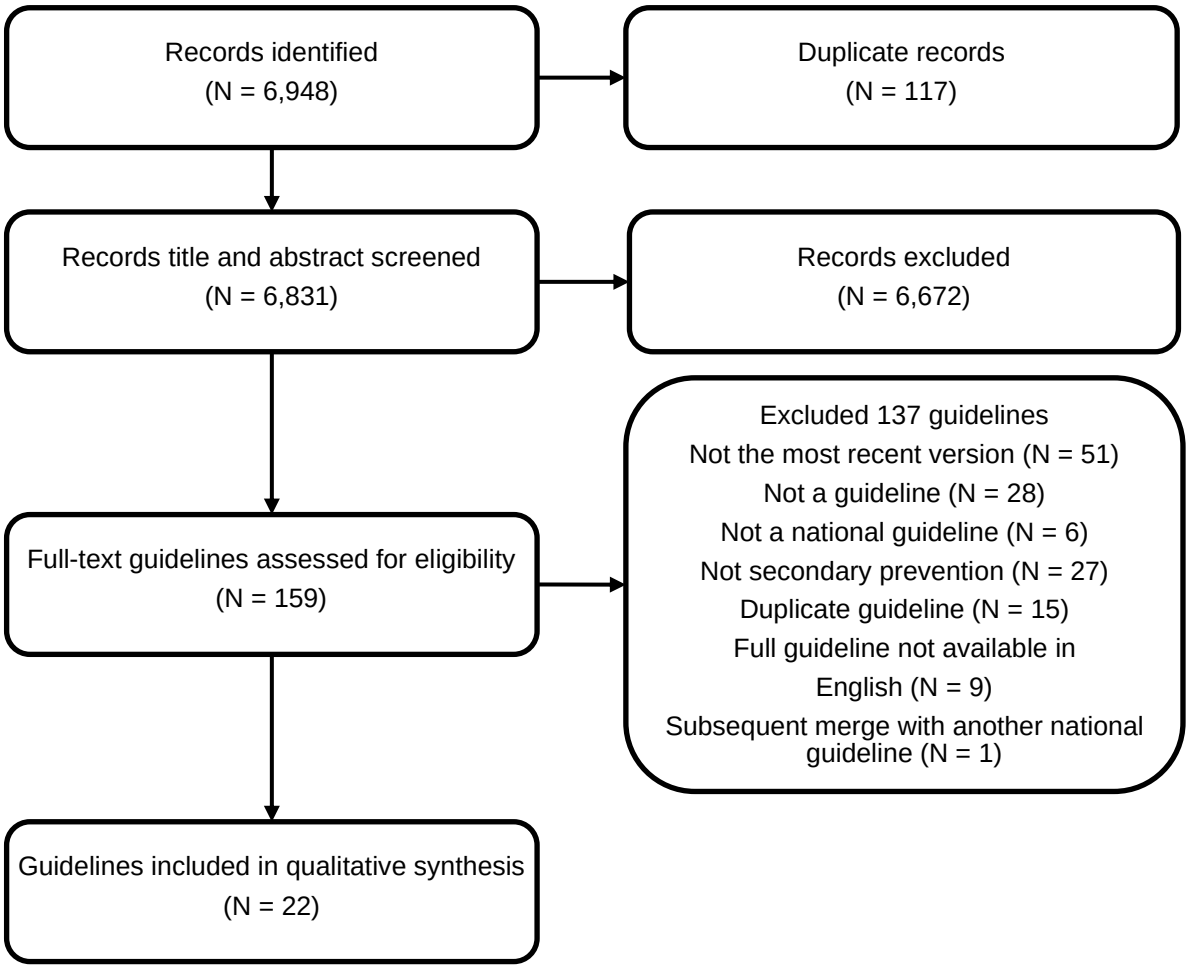


Figure 2.1: Selection Process of Relevant Guidelines

The literature search found 6,948 results (Figure 2.1), of which 117 were identified as duplicates. Of the 6,831 unique results, 6,672 were excluded following title and abstract screening. Following a full text review of the remaining 159 records, a further 137 were excluded. Common reasons for this exclusion were that a more recent version of the guideline existed (n=51), the record was not a guideline (n=28), the guideline did not apply to the secondary prevention population (n=27), or that it was a duplicate copy of another guideline published in a different journal (n=15). The remaining 22 guidelines (Bryer *et al.*, 2010; JBS3 Board, 2014; NICE, 2014; Expert Dyslipidemia Panel of the International Atherosclerosis Society Panel Members, 2014; Herdy *et al.*, 2014; International Diabetes Federation Guideline Development Group, 2014; Jacobson *et al.*, 2015; Anderson *et al.*, 2016; Toplak *et al.*, 2016; Chew *et al.*, 2016; (Australian) Stroke Foundation, 2017; SIGN, 2017; Tai *et al.*, 2017; Wang, Liu and Pu, 2017; Cheung *et al.*, 2017; Kinoshita *et al.*, 2018; Klug *et al.*, 2018; Li *et al.*, 2018; Naylor *et al.*, 2018; New Zealand Ministry of Health, 2018; Grundy *et al.*, 2019; Mach *et al.*, 2020) were assessed for their quality using AGREE II and included in the qualitative comparisons.

2.3.2 Characteristics and Quality of Guidelines

The guidelines included are summarised in Table 2.1 and were for 16 different regions. Two of the guidelines were global, with two guidelines each for the US and Europe. There were also two guidelines each for the UK, South Africa, Australia, and New Zealand including one which was applicable to both Australia and New Zealand. Finally, there was one guideline for the following regions: Austria, Canada, China, Hong Kong, Japan, South America, Scotland, Singapore, and Taiwan. Most of the guidelines were published in 2014 (n=7), 2016 (n=4), 2017 (n=5), and 2018 (n=4), with one guideline each published in 2010 and 2019.

Table 2.1: Summary of Included Guidelines

Abbrev.	Development Group	Title	Population	Region	Year
ACD	American College of Cardiology, American Heart Association, American Association of Cardiovascular and Pulmonary Rehabilitation, American Association Academy of Physician Assistants, Association of Black Cardiologists, American College of Preventive Medicine, American Diabetes Association, American Geriatrics Society, American Pharmacists Association, American Society for Preventive Cardiology, National Lipid Association, and Preventive Cardiovascular Nurses Association	2018 AHA/ACC/AACVPR/AAPA/ ABC/ACPM/ADA/AGS/APhA/ASPC/ NLA/PCNA Guideline on the Management of Blood Cholesterol (Grundy *et al.*, 2019)	All	USA	2018
AUSS	Stroke Foundation	Australian Clinical Guidelines for Stroke Management 2017 ((Australian) Stroke Foundation, 2017)	Stroke	Australia	2017
AUST	Austrian Obesity Association, Austrian Atherosclerosis Society, Austrian Diabetes Association, Austrian Society of Hypertension, Austrian Society for Internal Angiology, Austrian Society of Nephrology, Austrian Society of Cardiology, Austrian Stroke Society	Austrian Lipid Consensus on the management of metabolic lipid disorders to prevent vascular complications: A joint position statement issued by eight medical societies. 2016 update (Toplak *et al.*, 2016)	All	Austria	2016
CCSG	Canadian Cardiovascular Society	2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult (Anderson *et al.*, 2016)	All	Canada	2016

Abbrev.	Development Group	Title	Population	Region	Year
CSN	Chinese Society of Neurology, Cerebrovascular Disease Group	2014 Chinese guidelines for secondary prevention of ischemic stroke and transient ischemic attack (Wang, Liu and Pu, 2017)	Stroke	China	2014
ESCEAS	European Society of Cardiology, European Atherosclerosis Society	2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk (Mach *et al.*, 2020)	All	Europe	2019
ESVS	European Society for Vascular Surgery	Management of Atherosclerotic Carotid and Vertebral Artery Disease: 2017 Clinical Practice Guidelines of the European Society for Vascular Surgery (ESVS) (Naylor *et al.*, 2018)	Coronary Artery Disease	Europe	2017
HKCTF	Hong Kong Cardiovascular Task Force	2016 Consensus statement on prevention of atherosclerotic cardiovascular disease in the Hong Kong population (Cheung *et al.*, 2017)	All	Hong Kong	2016
IAS	International Atherosclerosis Society	An International Atherosclerosis Society Position Paper: global recommendations for the management of dyslipidemia (Expert Dyslipidemia Panel of the International Atherosclerosis Society Panel Members, 2014)	All	Global	2014
IDF	International Diabetes Federation	Global guideline for type 2 diabetes (International Diabetes Federation Guideline Development Group, 2014)	Diabetes	Global	2014
JAS	Japan Atherosclerosis Society	Japan Atherosclerosis Society Guidelines for Prevention of Atherosclerotic Cardiovascular Diseases 2017 (Kinoshita *et al.*, 2018)	All	Japan	2017
JBS3	Joint British Societies	Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3) (JBS3 Board, 2014)	All	UK	2014

Abbrev.	Development Group	Title	Population	Region	Year
NHF	National Heart Foundation of Australia, Cardiac Society of Australia and New Zealand	National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand: Australian clinical guidelines for the management of acute coronary syndromes 2016 (Chew *et al.*, 2016)	Acute Coronary Syndromes	Australia, New Zealand	2016
NLA	National Lipid Association	National Lipid Association recommendations for patient-centered management of dyslipidemia (Jacobson *et al.*, 2015)	All	USA	2014
NICE	National Institute for Health and Clinical Excellence	NICE Cardiovascular disease: risk assessment and reduction, including lipid modification (NICE, 2014)	All	UK	2014
NZ	New Zealand Ministry of Health	New Zealand Cardiovascular Disease Risk Assessment and Management for Primary Care (New Zealand Ministry of Health, 2018)	All	New Zealand	2018
SAF	South African Stroke Society	South African guideline for management of ischaemic stroke and transient ischaemic attack 2010 (Bryer *et al.*, 2010)	Stroke	South Africa	2010
SAHA	South African Heart Association, Lipid and Atherosclerosis Society of Southern Africa	South African Dyslipidaemia Guideline Consensus Statement: 2018 Update (Klug *et al.*, 2018)	All	South Africa	2018
SAM	Sociedade Brasileira de Cardiologia	South American guidelines for cardiovascular disease prevention and rehabilitation (Herdy *et al.*, 2014)	All	South America	2014
SIGN	Scottish Intercollegiate Guidelines Network	SIGN 149: Risk estimation and the prevention of cardiovascular disease (SIGN, 2017)	All	Scotland	2017
SMH	Singapore Ministry of Health	Singapore Ministry of Health Clinical Practice Guidelines: Lipids (Tai *et al.*, 2017)	All	Singapore	2017

Abbrev.	Development Group	Title	Population	Region	Year
TSC	Taiwan Society of Cardiology, Taiwan Society of Emergency Medicine, Taiwan Society of Cardiovascular Interventions	2018 Guidelines of the Taiwan Society of Cardiology, Taiwan Society of Emergency Medicine and Taiwan Society of Cardiovascular Interventions for the management of non ST-segment elevation acute coronary syndrome (Li *et al.*, 2018)	Acute Coronary Syndromes (Not ST)	Taiwan	2018

Abbrev, Abbreviation;

Table 2.2: AGREE Assessment Summary Scores

Guideline	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	Domain 6	Overall Quality (/7)
ACD	89%	50%	65%	100%	63%	100%	7
AUSS	100%	83%	90%	89%	50%	100%	6
AUST	78%	33%	27%	61%	0%	33%	4
CCSG	78%	39%	65%	94%	38%	100%	6
CSN	56%	44%	48%	78%	17%	83%	4
ESCEAS	56%	61%	63%	94%	54%	100%	7
ESVS	78%	56%	75%	72%	13%	33%	5
HKCTF	50%	33%	46%	78%	25%	58%	4
IAS	44%	39%	42%	44%	25%	42%	4
IDF	56%	33%	48%	83%	100%	83%	6
JAS	61%	44%	71%	67%	4%	25%	4
JBS3	56%	56%	29%	94%	25%	50%	5
NHF	89%	50%	73%	94%	79%	67%	7
NICE	100%	78%	94%	100%	96%	100%	7
NLA	72%	61%	54%	94%	17%	92%	6
NZ	83%	39%	27%	78%	17%	17%	4
SAF	39%	56%	67%	78%	29%	92%	6
SAHA	61%	56%	31%	72%	29%	83%	5
SAM	33%	28%	29%	61%	33%	33%	4
SIGN	89%	100%	96%	89%	88%	100%	7
SMH	67%	67%	54%	94%	17%	0%	4
TSC	67%	50%	52%	94%	8%	50%	5

Domain 1, Scope and Purpose; Domain 2, Stakeholder Involvement; Domain 3, Rigour of Development; Domain 4, Clarity of Presentation; Domain 5, Applicability; Domain 6, Editorial Independence.

ACD, American College of Cardiology/American Heart Association/American Association of Cardiovascular and Pulmonary Rehabilitation/American Association Academy of Physician Assistants/Association of Black Cardiologists/American College of Preventive Medicine/American Diabetes Association/American Geriatrics Society/American Pharmacists Association/American Society for Preventive Cardiology/National Lipid Association/Preventive Cardiovascular Nurses Association; AUSS, Australia Stroke Society; AUST, Austrian Obesity Association/Austrian Atherosclerosis Society/Austrian Diabetes Association/Austrian Society of Hypertension/Austrian Society for Internal Angiology/Austrian Society of Nephrology/Austrian Society of Cardiology/Austrian Stroke Society; CCSG, Canadian Cardiovascular Society; CSN, Chinese Society of Neurology and Cerebrovascular Disease Group; ESCEAS, European Society of Cardiology and European Atherosclerosis Society; ESVS, European Society for Vascular Surgery; HKCTF, Hong Kong Cardiovascular Task Force; IAS, International Atherosclerosis Society; IDF, International Diabetes Federation; JAS, Japan Atherosclerosis Society; JBS3, Joint British Societies; NHF, National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand; NICE, National Institute for Health and Clinical Excellence; NLA, National Lipid Association; NZ, New Zealand Ministry of Health; SAF, South African Stroke Society; SAHA, South African Heart Association; SAM, Sociedade Brasileira de Cardiologia; SIGN, Scottish Intercollegiate Guidelines Network; SMH, Singapore Ministry of Health; TSC, Taiwan Society of Cardiology/Taiwan Society of Emergency Medicine/Taiwan Society of Cardiovascular Interventions.

Table 2.2 contains the AGREE II scores for each of the guidelines. Reflecting generally high quality, eight guidelines were ranked as 4, four ranked as 5, five

as 6 and five as 7. Guidelines scored highest in Domain 4 (Clarity of Presentation) on average, and lowest in Domain 5 (Applicability) with many guidelines scoring below 50%. Scores of 0% were only attained in Domains 5 and 6 (Editorial Independence), with the latter occurring when no funding information or conflicts of interest were documented within the text. In terms of maximum values, 100% was attained in all domains by at least one guideline except for Domain 3 (Rigour of Development), where the highest score was 96%. This was the largest domain of the six and focussed on the development process of the guideline. Many did not document this fully or provide necessary references to additional material, and few detailed their update procedures.

2.3.3 Summary of Recommendations

Recommendations for the use of statin medication, cholesterol targets, and the frequency of monitoring are presented in Table 2.3.

2.3.3.1 Treatment Recommendations

All of the guidelines presented recommendations for the treatment of the secondary prevention population, with all recommending statins as the primary therapy. Only UK/Scottish guidelines suggested the specific drug and dose, namely, atorvastatin 80mg, with many instead recommending the maximally tolerated high-intensity doses in general, with lower doses considered when contraindications were present, or they were poorly tolerated by the patients. Few caveats were stated regarding the prescription of statins. For example, all stroke guidelines recommended statins only when the cause of the stroke was likely to be atherosclerotic, with the South African Stroke Society (SAF) additionally only recommending them in the case of total cholesterol >3.5mmol/l, and the Australia Stroke Society (AUSS) only considering them appropriate when the patient's life expectancy was considered reasonable. For the secondary prevention population as a whole, the European Society of Cardiology and European Atherosclerosis Society (ESCEAS) tailored their recommendations for patients whose baseline low-density lipoprotein (LDL) was <1.4mmol/l at baseline respectively, with therapy considered rather than offered to these patients. Finally, the 2018 American Consensus (ACD) guidelines

stated that the benefit-risk ratio should be considered when offering medication to patients over the age of 75 years.

For the guidelines which reported corresponding levels of evidence with their treatment recommendations (n=12), all considered that the level of evidence for statins was high, resulting in strong recommendations for their administration to the secondary prevention population. For situations where lower doses of therapy may be needed, such as in cases of contraindications or lack of tolerance, if specified at all, guidelines often considered the level of evidence supporting these changes to be lower than for the main treatment recommendation. Specifically, the level of evidence was typically assessed to be moderate (rather than high) or such alterations to medications were considered to be only good practice.

Besides statins, other lipid-lowering medications were also discussed within the guidelines (Table 2.4). The most commonly recommended of these was ezetimibe (n=17), both as an additional medication (n=15), and as a monotherapy (n=10) predominantly for patients with statin intolerance (n=8). Fibrates, niacin derivatives, and omega-3 supplements were also commonly recommended (n=15, n=10 and n=8, respectively) though under two different circumstances: elevated triglyceride levels and LDL cholesterol-lowering. For the former, fourteen recommended fibrates, whilst five guidelines each recommended considering niacin derivatives and Omega-3 supplements. Three guidelines suggested Omega-3 as lipid-lowering therapy, although the roles of fibrates and niacin derivatives were more disputed. Fibrates and niacin derivatives were recommended routinely in five and eight guidelines, respectively. However, three guidelines each did not recommend the use of fibrates and niacin derivatives. Bile acid sequestrant use was debated in thirteen guidelines, with only the South African Heart Association (SAHA) discouraging their use. PCSK9 inhibitors were only included in six guidelines, all of which were published from 2016 onwards, and all recommending them as an additional therapy or in cases of statin intolerance. Four guidelines did not give any recommendations for lipid-lowering medications other than statins. Evidence supporting these recommendations (if stated) was generally assessed by the guidelines to be of lower quality than for statins, and consequently the

associated strength of recommendations was typically lower. Ezetimibe and PCSK9 inhibitors tended to have higher levels of supporting evidence behind them, although the strength of recommendations for PCSK9 inhibitors was lower due to their recent approval and limited long term follow-up of cardiovascular events.

Table 2.3: Summary of Recommendations

	Statin Medication			Cholesterol Targets			Frequency of Monitoring		
Guideline	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**
ACD	≤75yr: High intensity statins >75yr: Initiate moderate or high intensity statins if benefit-risk ratio favourable	A B	I IIa	--	--	--	Fasting lipids 4-12 weeks after initiation, then every 3-12 months	A	I
AUSS	Atherosclerotic: High intensity statins if reasonable life expectancy	High	Strong	--	--	--	--	--	--
AUST	Statins	--	--	LDL-C <1.8mmol/L Or LDL-C >50% reduction Non-HDL-C <2.6mmol/L	-- -- --	-- -- --	--	--	--
CCSG	Moderate/high intensity statins	High	Strong	LDL-C <2.0mmol/L Or LDL-C >50% reduction Alternatively, apoB <0.8g/L Or non-HDL-C <2.6mmol/L ACS: LDL-C <1.8mmol/L Or LDL-C >50% reduction	Mod Mod Mod Mod -- --	Strong Strong Strong Strong V&Ps V&Ps	Until stable	--	PO

	Statin Medication			Cholesterol Targets			Frequency of Monitoring		
Guideline	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**
CSN	Ischaemic: High intensity statins	A	I	LDL-C <1.8mmol/L Or LDL-C ≥50% reduction	B B	II II	--	--	--
ESCEAS	Baseline LDL-C >1.4mmol/L: offer med <1.4mmol/L: consider med	 A A	 I IIa	LDL-C <1.4mmol/L Or LDL-C >50% reduction Further event <2yrs: LDL-C <1.0mmol/L	A A B	I I IIb	Starting/adjusting: 8 (±4) weeks Once achieved: Annually	-- --	-- --
ESVS	Statins prior to endarterectomy or stenting	A	I	--	--	--	--	--	--
HKCTF	High intensity statins	--	--	LDL-C <1.8mmol/L Or LDL-C >50% reduction if baseline 1.8-3.5 mmol/L	-- --	-- --	--	--	--
IAS	Maximal statins	--	--	LDL-C <1.8mmol/L. Non-HDL-C <2.6mmol/L	-- --	-- --	--	--	--
IDF	Statins	--	RC	Triglyceride <2.3mmol/L HDL-C >1.0mmol/L. Non-HDL-C <2.5mmol/L LDL-C <1.8mmol/L	-- -- -- --	RC RC RC RC	At least annually	--	RC

	Statin Medication			Cholesterol Targets			Frequency of Monitoring		
Guideline	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**
JAS	Statins	1+	A	LDL-C <2.6mmol/L Or LDL-C >50% reduction if target cannot be met <u>If additional conditions:</u> LDL-C <1.8mmol/L <u>If triglyceride high:</u> Non-HDL-C <3.4mmol/L Non-HDL-C <2.6mmol/L if additional conditions	3 3 -- -- -- --	A A -- -- -- --	Regular blood testing Every 3-6 months	-- --	B --
JBS3	Atorvastatin, up to 80mg in ACS	--	--	LDL-C <2.0mmol/L Non-HDL-C <2.5mmol/L	-- --	-- --	Annual non-fasting TC and HDL-C once stable	--	--
NHF	Highest tolerated dose of statins	1A	Strong	LDL-C ≤1.8mmol/L	--	--	TC and LDL-C approx. 3 months after starting	--	--
NICE	Atorvastatin 80mg, lower dose if not tolerated.	--	Strong	Non-HDL-C >40% reduction	--	Strong	3 months after treatment start. Annual non-fasting lipids	-- --	Strong Weak
NLA	Moderate/high intensity statins	High	A	Non-HDL-C <2.6mmol/L LDL-C <1.8mmol/L	High High	A A	4-12 months once achieved	Low	E
NZ	Statins	--	--	LDL-C 1.6-1.8mmol/L	--	--	Non-fasting 6-12 months until target achieved. Annually.	-- --	-- --

	Statin Medication			Cholesterol Targets			Frequency of Monitoring		
Guideline	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**	**Recommendation**	**LoE**	**SoR**
SAF	Atherosclerotic and TC >3.5mmol/L: Statins	I	A	--	--	--	--	--	--
	Trial strength e.g. 40mg simvastatin	--	--						
SAHA	High-intensity statins	--	--	LDL-C <1.8mmol/L	--	--	Starting/adjusting:	--	--
				Or LDL-C >50% reduction if baseline 1.8-3.5mmol/L	--	--	8(±4) weeks	--	--
							Once achieved: 6 months	--	--
SAM	Statins	--	--	LDL-C <2.6mmol/L	--	--	--	--	--
SIGN	Atorvastatin 80mg	--	Strong	--	--	--	Annual Review	--	GP
	Lower if not tolerated	--	GP						
SMH	Statins	1++	A	LDL-C <2.1mmol/L	1++	A	Annually	--	GP
TSC	Statins	A	I	LDL-C <1.8mmol/L	B	I	--	--	--
				Diabetes: LDL-C <1.4mmol/L	B	IIa			

ACD, American College of Cardiology/American Heart Association/American Association of Cardiovascular and Pulmonary Rehabilitation/American Association Academy of Physician Assistants/Association of Black Cardiologists/American College of Preventive Medicine/American Diabetes Association/American Geriatrics Society/American Pharmacists Association/American Society for Preventive Cardiology/National Lipid Association/Preventive Cardiovascular Nurses Association; ACS, acute coronary syndromes; AUSS, Australia Stroke Society; AUST, Austrian Obesity Association/Austrian Atherosclerosis Society/Austrian Diabetes Association/Austrian Society of Hypertension/Austrian Society for Internal Angiology/Austrian Society of Nephrology/Austrian Society of Cardiology/Austrian Stroke Society; CCSG, Canadian Cardiovascular Society; CSN, Chinese Society of Neurology and Cerebrovascular Disease Group; ESCEAS, European Society of Cardiology and European Atherosclerosis Society; ESVS, European Society for Vascular Surgery; GP, good practice; ; (non) HDL-C, (non) HDL-cholesterol; HKCTF, Hong Kong Cardiovascular Task Force ; IAS, International Atherosclerosis Society; IDF, International Diabetes Federation; JAS, Japan Atherosclerosis Society; JBS3, Joint British Societies ; LDL-C, LDL cholesterol; LoE, level of evidence; Med, medication; Mod, moderate; NHF, National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand; NICE, National Institute for Health and Clinical Excellence; NLA, National Lipid Association; NZ, New Zealand Ministry of Health; PO, panel opinion; RC, recommended care; SAF, South African Stroke Society; SAHA, South African Heart Association; SAM, Sociedade Brasileira de Cardiologia; SIGN, Scottish Intercollegiate Guidelines Network; SMH, Singapore Ministry of Health; SoR, strength of recommendation; TC, total cholesterol; TSC, Taiwan Society of Cardiology/Taiwan Society of Emergency Medicine/Taiwan Society of Cardiovascular Interventions; V&Ps, values and preferences.

Table 2.4: Summary of Non-Statin Medication Recommendations

	Ezetimibe			Fibrates			PCSK9 Inhibitors			Others		
Guideline	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**
ACD	If max statins and LDL ≥1.8 mmol/l, +	B-R	IIa	If trigs ≥5.7 mmol/l	B - NR	IIa	If max LLM, non-HDL ≥2.6 or LDL ≥1.8 mmol/l, +	A	IIa	--	--	--
AUSS	--	--	--	Not routine	Mod	Weak	--	--	--	--	--	--
AUST	First-line option	--	--	If trigs ≥5.6 mmol/l For lowering LDL, 2nd choice	-- --	-- --	--	--	--	For lowering LDL, BAS Niacin 2nd choice If trigs ≥5.6 mmol/l Niacin Omega-3 3-4g	-- -- -- --	-- -- -- --
CCSG	If max statins, +	High	Strong	For lowering LDL once met, + not recommended If high trigs, low HDL, may benefit	High --	Strong V&Ps	If max statins (/+ezetimibe), consider +	Mod	Cond	For lowering LDL once met, + niacin not recommended If max statins (/+ezetimibe), consider + BAS	High Low	Strong Cond
CSN	--	--	--	--	--	--	--	--	--	--	--	--

	Ezetimibe			Fibrates			PCSK9 Inhibitors			Others		
Guideline	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**
ESCEAS	If max statins and LDL above target, + If statin intolerant, consider	B C	I IIa	If trigs >2.3 mmol/l and LDL at target, consider + feno- or bezafibrate	C	IIa	If max statin + ezetimibe, + If statin intolerant, consider + ezetimibe	A C	I IIb	If max statins, consider + BAS If trigs 1.5-5.6 mmol/l, consider + Omega-3	C B	IIb IIa
ESVS	--	--	--	--	--	--	--	--	--	--	--	--
HKCTF	If max statins, + If statin intolerant	-- --	-- --	If very high trigs	--	--	--	--	--	If low HDL or high trigs, + nicotinic acid If statin intolerant: BAS Nicotinic acid	-- -- --	-- -- --
IAS	If max statins, LDL ≥ 1.8mmol/l, consider +	--	--	If high trigs, low HDL, target LDL, consider +	--	--	--	--	--	If max statins, LDL ≥1.8 mmol/l, consider + BAS If high trigs, low HDL, target LDL, consider +: Niacin High dose Omega-3	-- -- --	-- -- --

	Ezetimibe			Fibrates			PCSK9 Inhibitors			Others		
Guideline	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**
IDF	If max statins,	--	RC	If trigs > 2.3mmol/l,	--	RC	--	--	--	<u>If max statins, consider +:</u>		
	consider +	--	RC	low HDL,						BAS	--	RC
	If statin intolerant			consider fenofibrate						Nicotinic Acid	--	RC
										Omega-3	--	RC
										<u>If statin intolerant:</u>	--	RC
										BAS	--	RC
										Nicotinic Acid	--	RC
										Omega-3		
JAS	+ with statins	1+	B	Monotherapy	--	--	+ with statins	1+	B	BAS	--	--
	Monotherapy	--	--	If high trigs,	--	--				Omega-3	--	--
				consider +						Nicotinic Acid	--	--
JBS3	If max statins,	--	--	If trigs >10mmol/l	--	--	--	--	--	If max statins, consider BAS	--	--
	consider + 10mg	--	--									
	If statin intolerant, consider 10mg											
NHF	If max statins,	--	--	--	--	--	--	--	--	--	--	--
	consider + 10mg	--	--									
	If statin intolerant, consider 10mg											

	Ezetimibe			Fibrates			PCSK9 Inhibitors			Others		
Guideline	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**
NICE	--	--	--	--	--	--	--	--	--	--	--	--
NLA	If max statins, consider + If statin intolerant	Mod High	A A	If max statins, consider + If statin intolerant	Mod High	A A	--	--	--	<u>If max statins, consider +:</u> BAS Omega-3 Nicotinic Acid <u>If statin intolerance</u> BAS Omega-3 Nicotinic Acid	 Mod Mod Mod High High High	 A A A A A A
NZ	If statin intolerant, consider	--	--	If high trigs If statin intolerant, consider	-- --	-- --	--	--	--	If statin intolerant, consider niacin	--	--
SAF	--	--	--	--	--	--	--	--	--	--	--	--
SAHA	If max statins, consider +	--	--	If trigs >2.3 mmol/l, fenofibrate	--	--	If max statins, consider +	--	--	<u>LDL lowering:</u> BAS discouraged Nicotinic acid discouraged <u>If trigs >2.3 mmol/l:</u> + Omega-3 (2-4g) + Nicotinic Acid	 -- -- -- --	 -- -- -- --

	Ezetimibe			Fibrates			PCSK9 Inhibitors			Others		
Guideline	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**	**Recom.**	**LoE**	**SoR**
SAM	If max statins, +	--	--	If max statins, + If trigs >5.6 mmol/l, +	-- C	-- I	--	--	--	If low HDL, + niacin <u>If max statins:</u> + BAS + Niacin	-- -- -- --	-- -- -- --
SIGN	If max statins, consider +	--	Strong	For lipid lowering, not routine If trigs high, low HDL, consider	-- --	Strong GP	If max LLM, consider	--	Strong	If max statins, consider + BAS Nicotinic acid not recommended	-- --	Strong Strong
SMH	If max statins, + If statin intolerant	1++ 1++	A A	If low HDL, trigs 2.3-4.5 mmol/l, consider + Prefer fenofibrate If trigs >10 mmol/l	2+ 3 1+	C D A	--	--	--	If max statins, + BAS If max statins, consider + niacin If trigs ≥4.5 mmol/l, consider niacin If trigs >10 mmol/l, + Omega-3	4 1+ 1+ 1+	D A A A
TSC	+ with statin	A	I	--	--	--	--	--	--	--	--	--

ACD, American College of Cardiology/American Heart Association/American Association of Cardiovascular and Pulmonary Rehabilitation/American Association Academy of Physician Assistants/Association of Black Cardiologists/American College of Preventive Medicine/American Diabetes Association/American Geriatrics Society/American Pharmacists Association/American Society for Preventive Cardiology/National Lipid Association/Preventive Cardiovascular Nurses Association; AUSS, Australia Stroke Society; AUST, Austrian Obesity Association/Austrian Atherosclerosis Society/Austrian Diabetes Association/Austrian Society of Hypertension/Austrian Society for Internal Angiology/Austrian Society of Nephrology/Austrian Society of Cardiology/Austrian Stroke Society; BAS, bile acid sequestrant; CCSG, Canadian Cardiovascular Society; Cond, conditional; CSN, Chinese

Society of Neurology and Cerebrovascular Disease Group; ESCEAS, European Society of Cardiology and European Atherosclerosis Society; ESVS, European Society for Vascular Surgery; GP, good practice; (non) HDL-C, (non) HDL-cholesterol; HKCTF, Hong Kong Cardiovascular Task Force ; IAS, International Atherosclerosis Society; IDF, International Diabetes Federation; JAS, Japan Atherosclerosis Society; JBS3, Joint British Societies ; LDL-C, LDL cholesterol; LLM, lipid-lowering medication; LoE, level of evidence; Med, medication; Mod, moderate; NHF, National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand; NICE, National Institute for Health and Clinical Excellence; NLA, National Lipid Association; NZ, New Zealand Ministry of Health; PCSK9, Proprotein Convertase Subtilisin/Kexin type 9; RC, recommended care; Recom, recommendations; SAF, South African Stroke Society; SAHA, South African Heart Association; SAM, Sociedade Brasileira de Cardiologia; SIGN, Scottish Intercollegiate Guidelines Network; SMH, Singapore Ministry of Health; SoR, strength of recommendation; TC, total cholesterol; Trigs, triglycerides; TSC, Taiwan Society of Cardiology/Taiwan Society of Emergency Medicine/Taiwan Society of Cardiovascular Interventions; V&Ps, values and preferences; +, adding.

2.3.3.2 Plasma Lipid Recommendations

Seventeen of the guidelines provided at least one target, with all except one of these providing an LDL cholesterol goal. Target values ranged between 1.0 and 2.6mmol/l, although the most frequently recommended was <1.8mmol/l (n=12). Many guidelines additionally suggested that a 50% reduction in LDL could be used as an alternative where this target may be unattainable or for patients whose baseline values were already <3.5mmol/l (n=7). A non-high-density lipoprotein (HDL) cholesterol target was also common (n=8), with target values ranging from 2.5-3.4mmol/l, with 2.6mmol/l the most frequent (n=5). These targets were usually given in combination with a target for LDL cholesterol, though in the case of the Japan Atherosclerosis Society (JAS), the non-HDL target was considered only relevant when a patient's triglycerides were elevated. Meanwhile, the National Institute for Health and Clinical Excellence (NICE) did not provide a numerical target for non-HDL, recommending a 40% reduction from the patient's baseline only. Only two guidelines referred to other lipid parameters in their recommendations. The Canadian Cardiovascular Society (CCSG) provided an apolipoprotein B target as an alternative for LDL, and the International Diabetes Federation (IDF) provided additional targets for triglycerides and HDL cholesterol. There were no apparent differences in recommendations for stroke or diabetes specific guidelines, although some guidelines for all secondary prevention populations provided different targets for those patients with additional comorbidities (n=2), or specific cardiovascular events (n=2).

For the majority of the guidelines that provided targets, the recommendations provided either no supporting evidence or graded it as low. Consequently, the associated strength of the recommendations was often either not given or was given as preferences and opinions of those involved in the guideline's construction. There were few exceptions. Singapore's (SMH) recommendation of LDL<2.1mmol/l considered the level of evidence to be high, resulting in a strong recommendation. The target of LDL<1.8mmol/l, when stated in some guidelines, was also strongly recommended, with the National Lipid Association (NLA) considering the supporting evidence to be high, whilst others graded it as only moderate (n=2). CCSG's target for patients who have not experienced an acute coronary syndrome of LDL<2.0mmol/l was also given as a strong

recommendation, with the guideline viewing the supporting evidence for it as moderate. Its recommendation of LDL<1.8mmol/l for patients who have experienced an acute coronary syndrome cited no supporting evidence, and consequently was given as a preference amongst the guideline's creators.

2.3.3.3 Frequency of Monitoring

Thirteen guidelines detailed recommendations regarding the ongoing monitoring of this population, with the specifics falling into three categories: monitoring following the initiation of treatment (n=5), monitoring prior to stable lipids (n=2), and long term follow-up (n=11), with some providing recommendations in more than one of these categories (n=5). Of those who detailed monitoring following statin initiation, all recommended a review of the patient's lipids within three months. Within these, two guidelines recommended also measuring the patient's alanine aminotransferase at this review (NICE, 2014; Mach et al., 2020), one suggested this should be conducted only if symptoms were present (Grundy et al., 2019), and another two did not refer to this safety blood indicator (Chew et al., 2016; Klug et al., 2018). Furthermore, the measurement of creatine kinase was only considered where side effects were reported at this initial review, and was recommended in five of these guidelines (NICE, 2014; Klug et al., 2018; Grundy et al., 2019; Mach et al., 2020), and implied in the remaining guideline (Chew et al., 2016). Meanwhile, for the guidelines which recommended monitoring prior to stability, the criteria for this was not clearly defined. Both CCSG and New Zealand Ministry of Health (NZ) did not detail a specific purpose for these reviews. In terms of long term follow-up, many recommended reviewing the patient annually (n=9), although there was some variation in recommendations between 3 and 12 months.

The majority of guidelines considered the evidence behind their frequency of monitoring recommendations to either be low (n=1) or gave no evidence to support them (n=11), sometimes referring to them as good practice points or clinician's opinions or preferences (n=5). However, the ACD guidelines were the exception to this, which graded the evidence behind their monitoring recommendations of every 3-12 months as strong.

2.4 Discussion

This systematic review illustrates the variation in recommendations surrounding optimal on-statin lipid monitoring within secondary prevention. Specifically, there were considerable differences in the recommendations for cholesterol targets (including their use) and the ongoing monitoring of lipid levels over the longer term. These findings reflect the fact that no guideline identified a specifically designed randomised controlled trial to assess either treatment targets or monitoring of therapy. However, such trials are likely to be expensive, although in the future advances with electronic health records may facilitate the evidence base for this. Nonetheless, this systematic review illustrates that better evidence is needed to provide an optimal approach to lipid monitoring in order to balance safety, adherence, cost, and time burden to patients.

All guideline committees are likely to be searching a broadly similar evidence base, where the efficacy and safety of statins has been well established in the prevention of further cardiovascular events (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010; Cholesterol Treatment Trialists' (CTT) Collaborators, 2012). Furthermore, the World Health Organisation recommends the use of statins as part of their secondary prevention program (World Health Organization, 2017b), with the increased risk widely accepted within this population. It is therefore not surprising that all guidelines agree that statins should be commenced as the lipid-lowering therapy with a high level of evidence, commensurate with the availability of randomised controlled trials. Guidelines that specify drugs or doses generally recommend high dose therapy and titrating down as necessary to a tolerated dose, rather than titrating up. This is in line with large randomised controlled trials over the preceding decades that have shown that higher dose statin therapy improves outcomes (Cannon *et al.*, 2006). Nonetheless, despite this widespread recommendation, there is some evidence to suggest that statins are not consistently prescribed within the secondary prevention population depending on the cardiovascular events experienced (Tibuakuu *et al.*, 2020). Meanwhile, recommendations for the use of other lipid-lowering therapies illustrate that such guidelines are likely a reflection of the evidence available when the guidelines were created e.g. PCSK9 inhibitors were only discussed in guidelines published after 2016.

Therefore, in clinical practice, consideration may need to be given to the timing of the publication where new evidence has emerged.

The evidence that "lower is better" for LDL and non-HDL cholesterol in terms of secondary CVD prevention is well supported by evidence, and the recent advent of Proprotein Convertase Subtilisin/Kexin type 9 (PCSK9) inhibitors further supports this notion (Sabatine et al., 2017; Schwartz et al., 2018). The issue at hand is how to use this information clinically to support a testing regimen. In this regard the use of cholesterol targets for therapy is contentious, in part because the evidence for their use is less strong. Whilst research has shown that achieving targets is associated with better outcomes (Boekholdt et al., 2014; Sabatine et al., 2017; Schwartz et al., 2018), no specific randomised controlled trial has shown that randomising patients to a target improves adherence or event rates Indeed, it might be argued that the maximally tolerated statin should be initiated as the default, and therefore a hard target may be moot. Clearly, other lipid-lowering medications could be added to therapy. Despite the lack of strong evidence, many guidelines recommended specific lipid targets, with many choosing similar values, suggesting that guideline committees are likely to be examining the same evidence. However, one of the most recently published guidelines, the 2019 ESC/EAS Guidelines for the Management of Dyslipidaemias (ESCEAS) advised the lowest cholesterol targets of all the included guidelines (LDL<1.4mmol/l), with further lower levels for those with multiple recent cardiovascular events (LDL<1.0mmol/l). This was rated as being supported by strong and moderate levels of evidence, although the guidelines acknowledged that both targets are based on the LDL levels achieved in the trials for PCSK9 inhibitors (Sabatine et al., 2017; Schwartz et al., 2018; Mach et al., 2020).

There was virtually no evidence to support any recommendations regarding the frequency of ongoing monitoring once lipid-lowering therapies had been commenced. Guidelines that recommended retesting following statin initiation tended to additionally recommend liver function tests were performed as a safety indicator, especially if hepatic symptoms were present, including those produced by the ACC/AHA and the ESC/EAS (Grundy *et al.*, 2019; Mach *et al.*, 2020). Guidelines frequently conflate the issue of short term safety bloods with

longer-term lipid monitoring when reporting the strength of evidence. Despite this only one guideline (ACD) cited evidence that they considered to be strong in their recommendations. The evidence referenced would suggest that monitoring patients regularly is associated with improved adherence to medication, and, consequently, patient outcomes (Benner *et al.*, 2004). However, this study was open to confounding due to its observational nature, and as far as we are aware, this has never been tested within a randomised controlled trial. Furthermore, the purpose of such follow-up testing in guidelines is seldom stated. Where evidence was cited, though, this would suggest that the purpose of such reviews is to promote adherence, but this may not be the rationale for all of the guidelines included, which could also include the monitoring of lipids to check if targets are still being achieved or of safety concerns. Regardless, the majority of guidelines recommended that secondary prevention patients were reviewed annually based purely on clinician's opinion. However, in the UK, simulations have suggested that this is likely to be optimal economically as well as reducing the impact of any natural variation in an individual's cholesterol levels (Perera *et al.*, 2015). Nonetheless, some patients will not be optimally managed under these recommendations, and by integrating algorithms into electronic health records to aid clinical decision making, there is the potential to personalise an individual patient's lipid management.

2.4.1 Strengths and Limitations

To our knowledge, this is the first review to compare guidelines surrounding the management of lipids in the secondary prevention population, as previous research has focussed on comparing guidelines for assessing risk and managing it through lifestyle interventions in the primary prevention population only (Khanji *et al.*, 2016, 2018). Furthermore, following a comprehensive search, this identified current guidelines from 22 different professional bodies, covering 16 different geographical regions. Nonetheless, guidelines were only included if their full guideline was available in English, which is likely to have resulted in a bias in the regions included and impacted the number of guidelines compared within this review. Another limitation is that the guidelines' methodological quality, as assessed by the AGREE II tool, was not used to restrict their inclusion in this review, and no comparisons were made between either the recommendations given, their considered level of evidence or their strength,

and the AGREE scores. However, given that all included guidelines were assessed to be of generally high quality, such stratification is unlikely to have yielded meaningful differences.

2.4.2 Conclusion

The safe and optimal treatment of plasma lipids within the secondary prevention population is key to reducing the increasing burden of cardiovascular disease in society. However, given the paucity of evidence for the frequency of ongoing monitoring, there is a clear need for further research in these two key areas of its management. This will improve patient care while optimising costs in an evidence-based manner.

Chapter 3 Deriving the Cohort

3.1 Data Source

The source data for this project consisted of an extract of all individuals in NHS Greater Glasgow and Clyde (NHS GGC) who had any lipid profile result or a prescription for a statin, ezetimibe, or PCSK9 inhibitor before 29th December 2017 (when the data extraction occurred). Further details surrounding the definitions of a lipid profile result and the lipid-regulating medications of interest are shown in Figure 3.1 below. These medications were selected as statins, ezetimibe, and PCSK9 inhibitors are all recommended routinely in Scotland, whereas fibrate use is not routine and niacin is no longer recommended (SIGN, 2017). Initial data extraction from all available NHS GGC data was performed by Safe Haven staff and contained data for 652,441 individuals (Project Reference: GSH/17/CA/012 PMCVD).

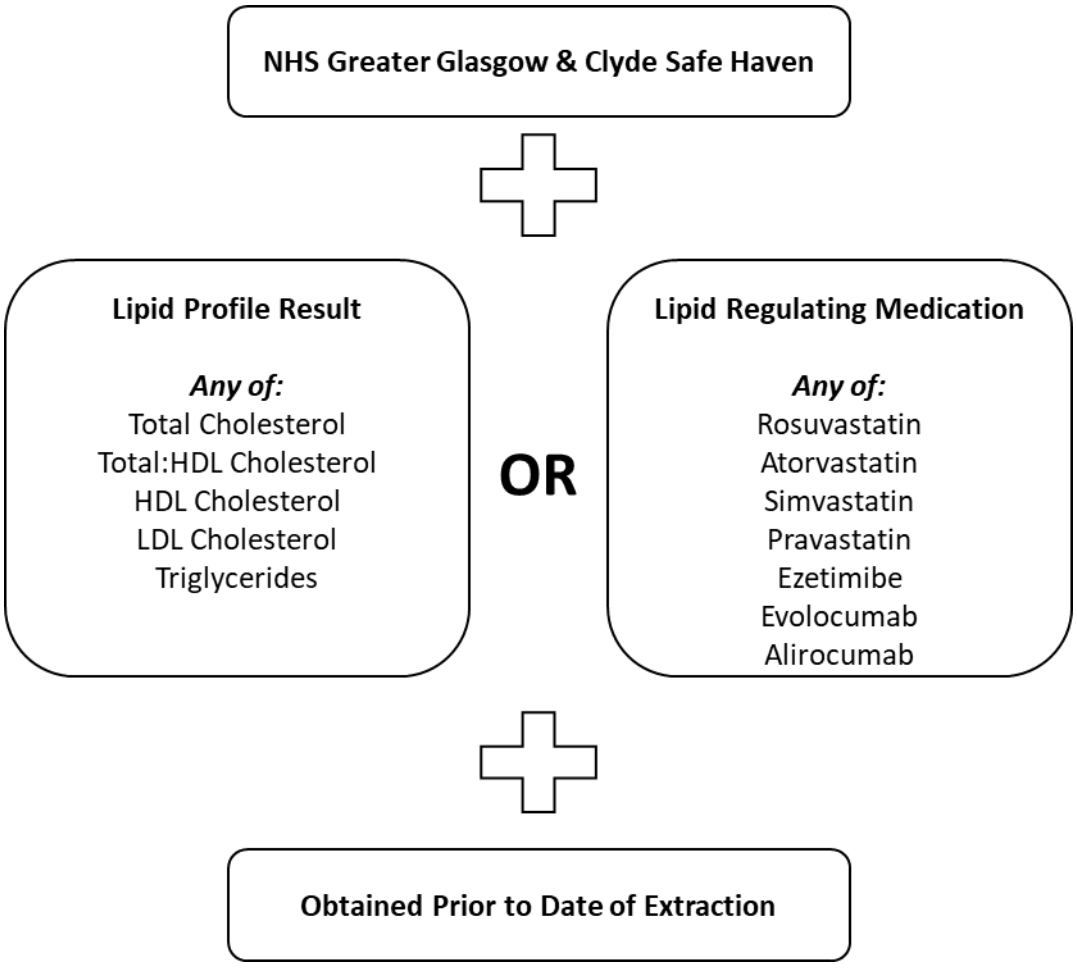


Figure 3.1: Initial Inclusion Criteria for Data Extract

The extracted data for these individuals included their demographic data (such as their date of birth, sex, postcode sector, and deprivation quintiles), their laboratory results (from SCI Store), their dispensed prescriptions, their hospital admissions (from SMR01 records), and death records captured by the National Records of Scotland (NRS). If applicable, information captured regarding an individual's diabetes diagnosis was also extracted from the diabetes register (SCI Diabetes). The number of records contained within each of these datasets before data cleaning are shown in Table 3.1.

Table 3.1: Size of Datasets from Initial Extraction Performed by Safe Haven Staff.

Dataset	# Records
Demographics	652,441
Laboratory Tests (SCI Store)	256,361,353
*Lipid Tests**	14,626,093
Prescribing (Pharmacy)	40,899,219
Lipid-Lowering Medications†	9,812,303
Hospital Admissions (SMR01)	775,172
Deaths (NRS)	67,930
Diabetes Register (SCI Diabetes)	82,579,801
Date and Type‡	1,791,215

***Results for Total Cholesterol, LDL Cholesterol, HDL Cholesterol, Triglycerides, and Total:HDL Cholesterol. †Prescriptions with a BNF Section Description equal to "Lipid-Regulating Drugs". ‡Data items with a description of Date of Diagnosis or Diabetes Mellitus Type.**

Records between the datasets for an individual were linked by Safe Haven staff using the individual's CHI number, which was then removed, and an anonymous patient ID assigned. Dates of birth were also set to the fifteenth of the month and year applicable, and postcodes were limited to the postcode sector of the individual. Access to this anonymised data was granted via a secure platform hosted by the Robertson Centre for Biostatistics. All results included in this thesis were extracted in tabular and graphical form and were quality checked by Safe Haven staff before extraction. All data cleaning and analysis for this project was performed using R v3.5.0 (R Core Team, 2018) within the secure platform. Packages for specific analyses are referenced at the corresponding results.

After extraction, data cleaning occurred in two stages: first, to remove those with clinically implausible values, and records captured outside the period of interest, and second, to identify those individuals with a valid MI (determined using hospital admission records with ICD10 codes I21 and I22) to be included in the analysis. An overview of these processes is shown in Figure 3.2 below, and further details of the decisions and processes implemented are described in the remainder of this chapter.

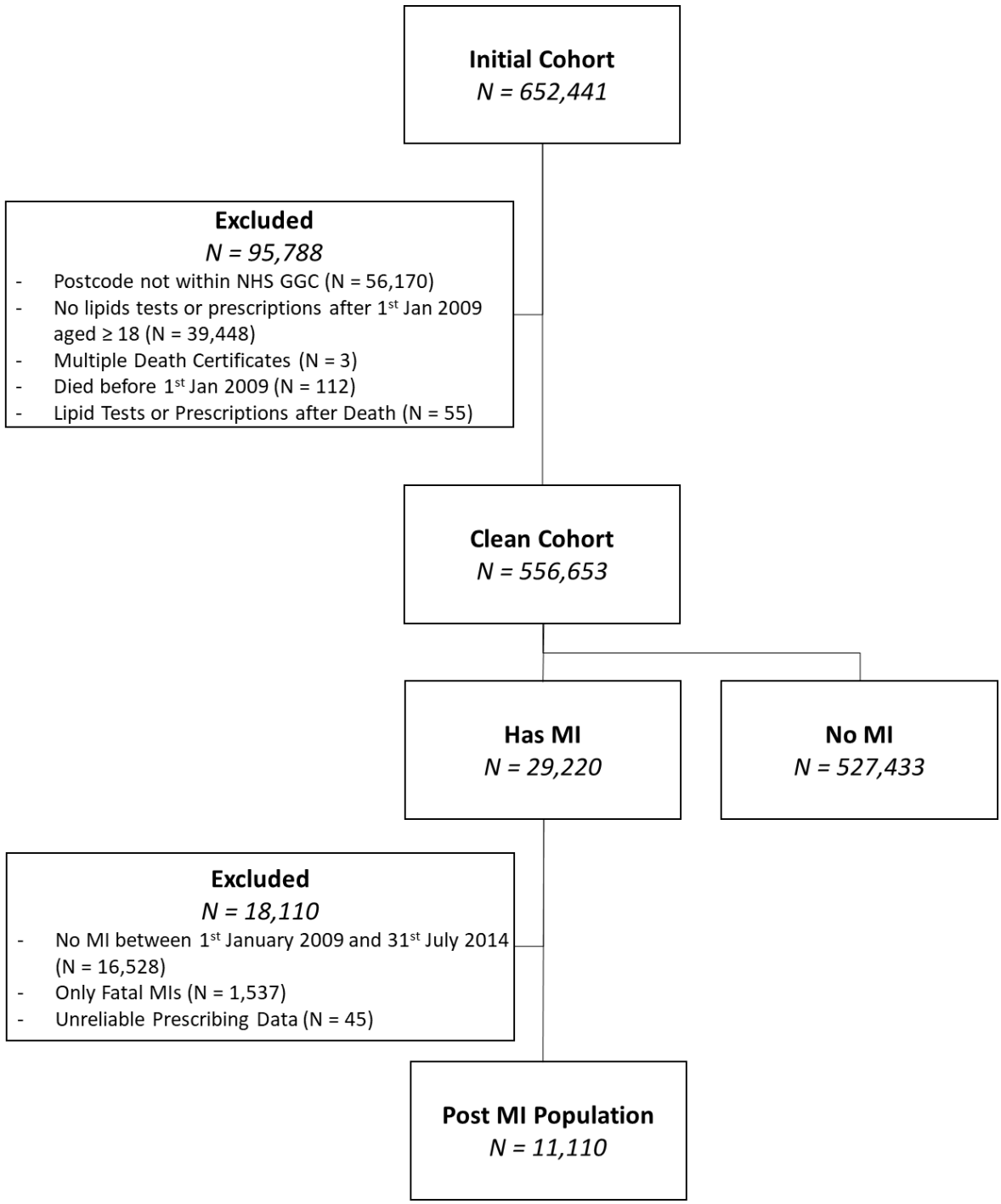


Figure 3.2: Derivation of the Post Myocardial Infarction Population; MI = Myocardial Infarction.

3.2 Initial Data Cleaning to Establish Basic Cohort

3.2.1 Start and End Date

In the initial data extraction, no start date was formally defined to maximise the data available for use. However, the earliest dates in the laboratory data did not suggest reliability or plausibility for the earliest records (e.g. earliest dates were in 1931). Therefore, a cumulative distribution plot was used to identify the first day that such laboratory values were likely to be collected (Figure 3.3). This identified two possible dates where a sharp increase in the number of records was reported, 25th June 2006 and 1st March 2008.

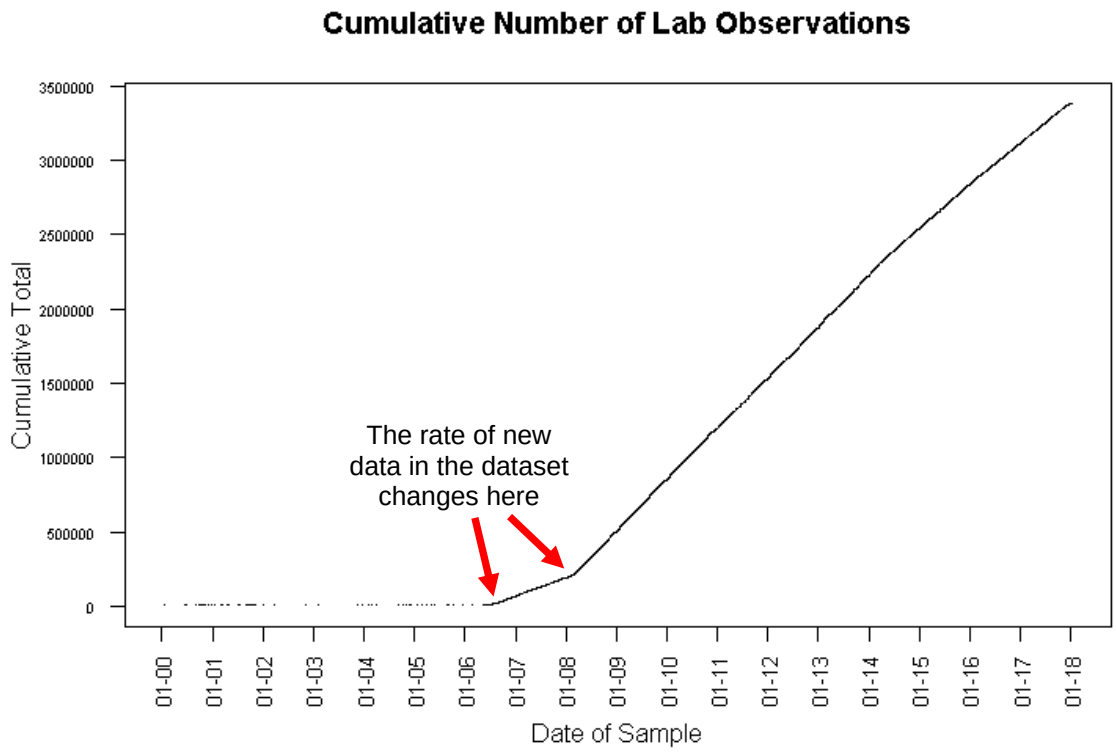


Figure 3.3: Cumulative Observation Plot for Laboratory Samples within SCI Store. The total number of lipid profile observations contained within the data by each date between January 2000 and January 2018 were plotted to reveal two dates where a sharp increase was observed.

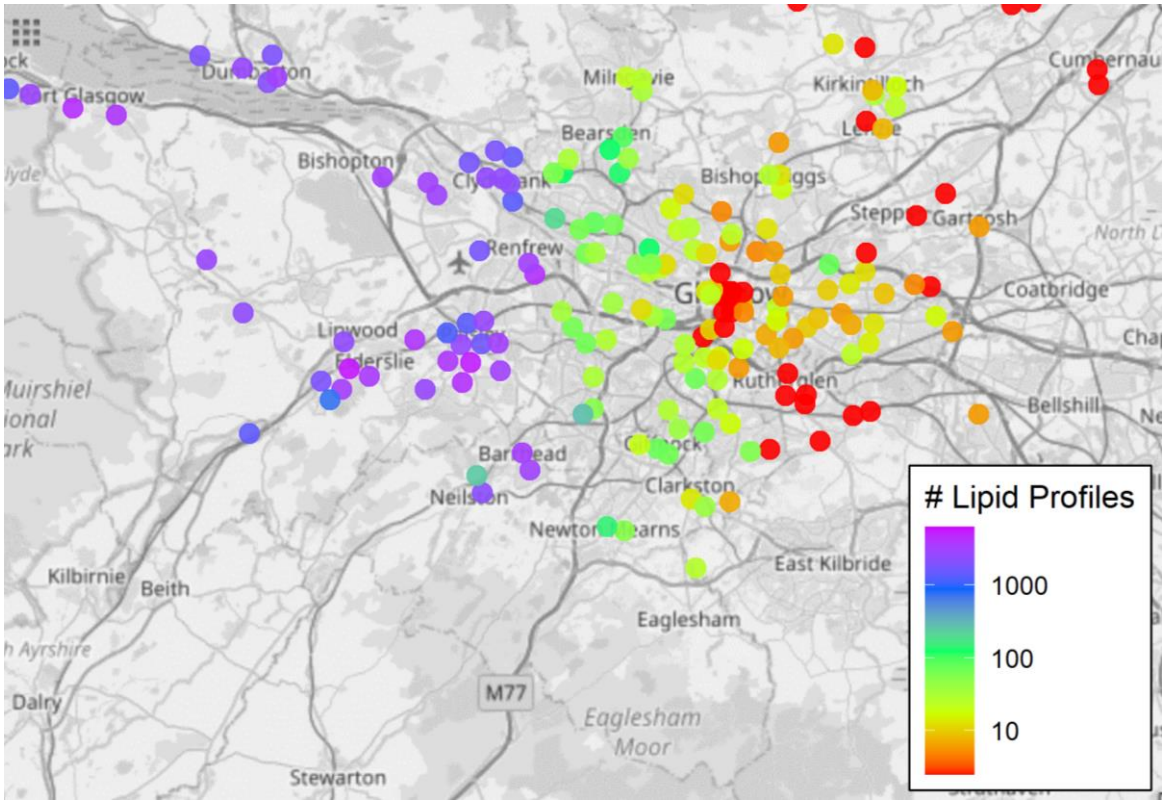


Figure 3.4: Location Plot of Lipid Profile Tests. The number of tests in each postcode sector collected between 25th June 2006 and 1st March 2008 were plotted geographically, with purple and red areas indicating high and low numbers of tests respectively. Background map was sourced from OpenStreetMap (OpenStreetMap, 2020b) **with extraction coordinates: latitude (55.67, 55.97), longitude (-4.75, -3.95).**

To determine which of the dates to select, the number of tests per postcode sector between these two dates were cross-tabulated and subsequently plotted in R to aid the identification of any geographical patterns (Figure 3.4). These were plotted using coordinates obtained from the 2019 Scottish Postcode Directory (National Records of Scotland, 2019), which were subsequently transformed using the Mercator projection (OpenStreetMap, 2020a), and plotted on a map sourced from OpenStreetMap (OpenStreetMap, 2020b). This revealed that the majority of the tests collected between the two dates were from individuals residing in the west of the NHS GGC region. This is likely due to not all laboratories in the region starting to upload results into SCI store on the same date, resulting in the second notable increase in the rate of observations on 1st March 2008. Therefore, 1st March 2008 was considered the start of reliable data captured within SCI store.

However, whilst laboratory data was found to be accurate from 1st March 2008, concerns were raised about the completeness of the prescription records before 2009. The cumulative incidence graph showed a relatively small increase in the

rate of prescribing at the start of 2009 (Figure 3.5). As a result of this, and the prescribing documentation recommendation regarding the CHI linkage of prescriptions before 2009 at the national level (Information Services Division, 2010, p. 2), the start date for this analysis was selected as 1st January 2009.

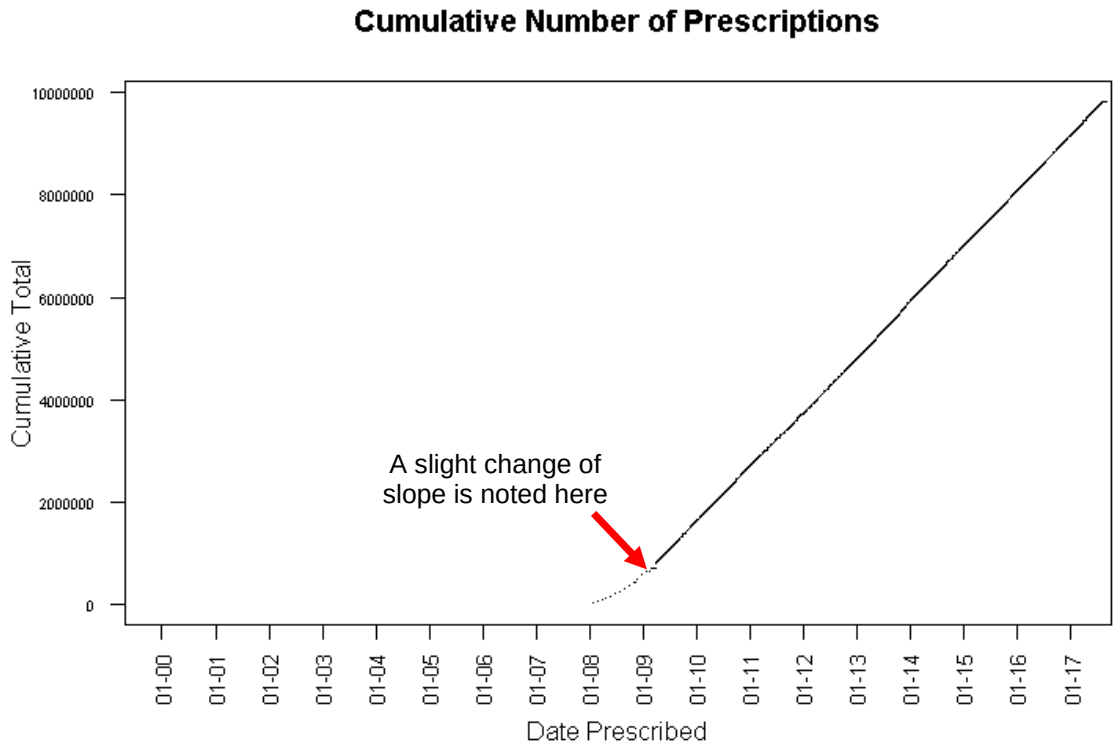


Figure 3.5: Cumulative Observation Plot for Lipid Regulating Prescriptions. The total number of prescriptions contained within the data by each date between January 2008 and January 2018 were plotted. This revealed one date where a slight increase was observed, suggesting that records were becoming more complete at the start of 2009.

Whilst the laboratory records were generally considered to be complete until the date of extraction (29th December 2017), concern was raised for complete record ascertainment in the prescription records and the hospital admission data (SMR01 file). The number of prescriptions and MI admissions appeared to taper off throughout 2017 (Figure 3.6), suggesting that not all the records had been completed at the time of extraction. When examined on a monthly basis for 2017, it became apparent that whilst the SMR01 records appeared complete until the end of August 2017, the prescriptions records were only complete until the end of July 2017. For this reason, the end of follow up for this analysis was selected as 31st July 2017.

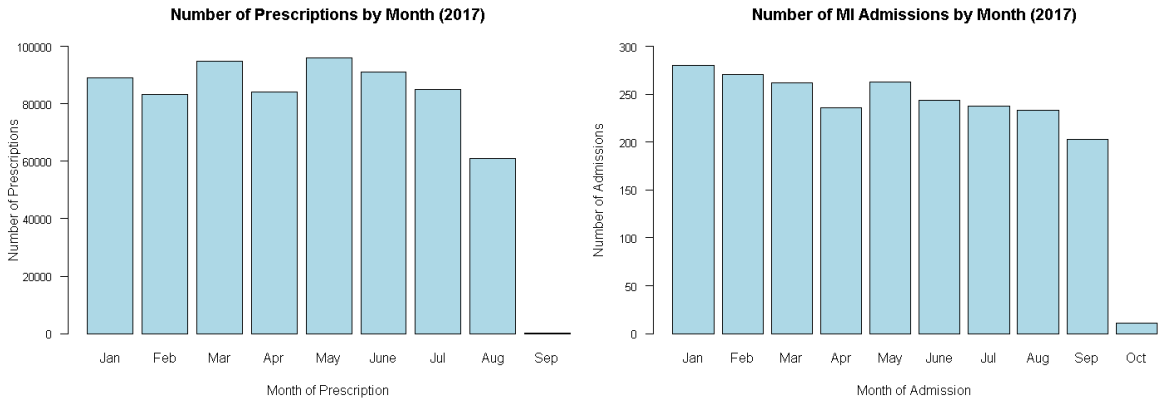


Figure 3.6: Bar Plots Illustrating Prescriptions and Number of MI Admissions by Month in 2017. The decrease in the number of prescriptions appears to start in August 2017, whilst the decrease in the number of MI admissions is not evident until September 2017.

3.2.2 Events Occurring Around Time of Death

Individuals included in the analysis were followed up until their date of death or 31st July 2017, whichever occurred first. Data cleaning efforts were therefore undertaken to ensure the accuracy and reliability of death records.

The date of death was available in the NRS deaths file, together with details surrounding the cause and location from the individual's death certificate. Firstly, those whose date of death was reported before 1st January 2009 were removed (n=112), and individuals who had more than one death certificate contained within the NRS file were removed from the cohort (n=3) as the correct record could not be ascertained. The date of death was also recorded in the separately derived demographics file, with the date being derived through a hospital discharge notice where the discharge occurs as a result of 'DEATH'. A check was made to ensure that the dates in both sources matched, and there were approximately 100 individuals where this was not the case. As the NRS date was likely to be the more reliable, this date was used in such instances. Where the death certificate details were unavailable, the demographic file date was assumed. If an individual had no date of death recorded in either file before 31st July 2017, then they were assumed to still be alive at the end of follow up.

To further assess the plausibility of such dates of death, checks for lipid profile results recorded after the date of death were conducted. Individuals with apparently posthumous blood results were also removed from the population as it could not be determined whether the date of the test or the date of death

was correct. Similarly, checks for prescriptions of lipid-regulating medications prescribed after the date of death were performed. However, due to the possibility that some individuals could have their medications automatically dispensed (such as those residing in care homes), a six-month time window (182 days) was allowed. Any individuals whose prescriptions were issued more than six months after death were removed from the cohort, as this would suggest that one of the dates was inaccurate and determining where the error had occurred was not possible. These two checks combined removed a further 55 individuals from the population.

3.2.3 Missing Causes of Death

The number of deaths reported in the two sources also highlighted that there were considerably more deaths captured in the demographics file than in the NRS data. This meant that for many deaths, the cause of death was unknown. A temporal check was performed and was found not to be the cause of this discrepancy. A geographical check, which consisted of plotting the number of deaths and the number of missing deaths per postcode sector (Figure 3.7), was implemented. This highlighted that the areas with the highest rates of missing information were in postcode sectors outside the NHS GGC region. These individuals are likely to have travelled into the region for some aspects of their care, and therefore a complete representation of their health status was unlikely to have been obtained in this data extract. As a result, all individuals with postcodes completely outside the health board (NHS Greater Glasgow and Clyde, 2015a, p. 7), as shown in Figure 3.8, were excluded (n=56,170).

Despite this, a moderate degree of missingness was still present, as around 25% of those with a recorded death did not have a cause of death available due to a missing death certificate. This is most likely to have occurred for patients who died in an NHS GGC hospital, but whose death certificate was registered outwith the region. This highlighted a possible bias within the data, as a patient's location was only ascertained by their postcode sector at the time of their first lipid test or lipid-regulating prescription. Consequently, those appearing to not be tested or dispensed medications may have moved out of the area during the period of interest. As a result, the cause of death may not be missing at random if relocations are more common in patients who have more severe levels of

frailty (Scheibl *et al.*, 2019), which may result in deaths from certain causes more likely.

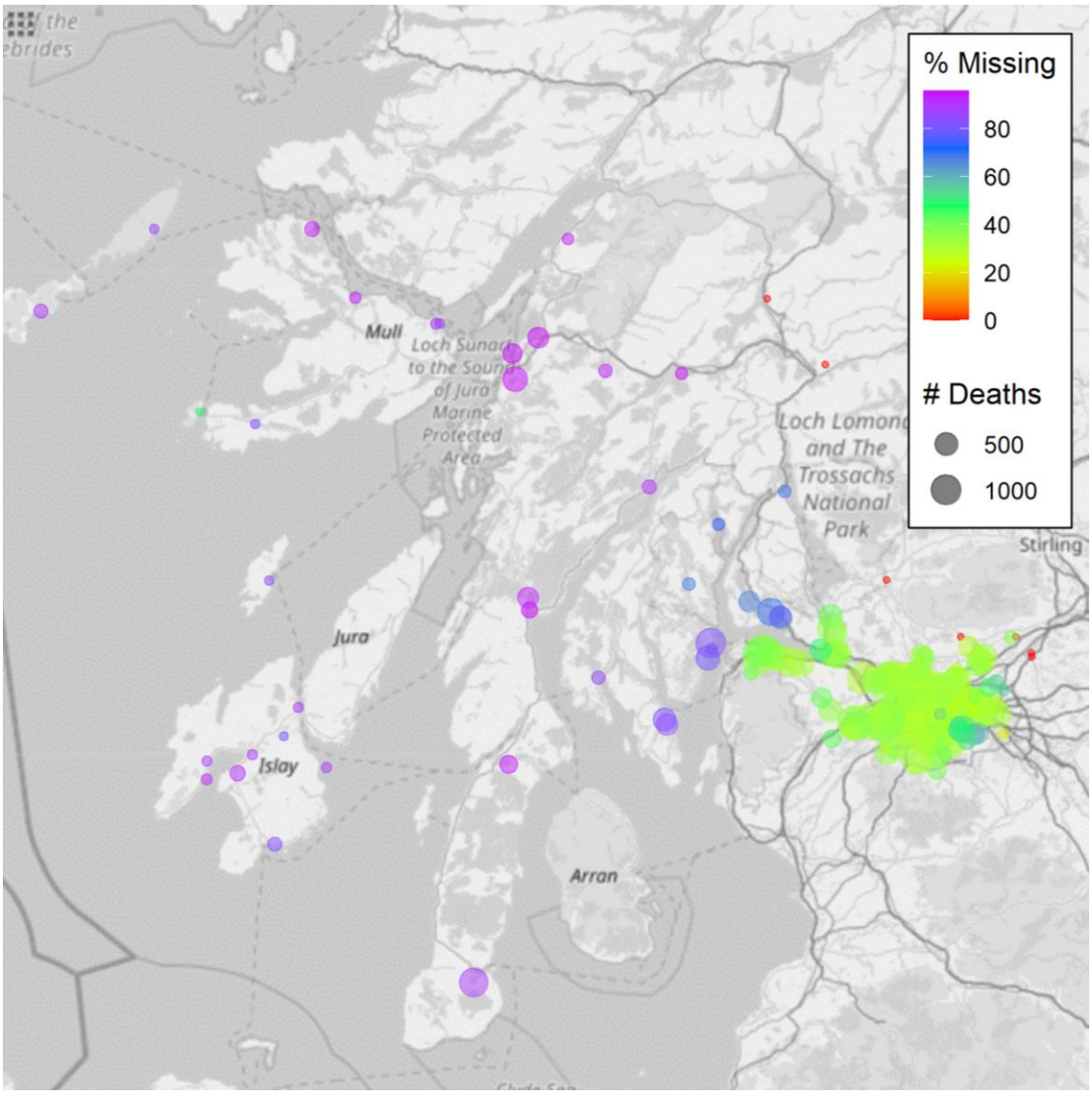


Figure 3.7: Geographical Plotting of Missing Death Certificates by Postcode Sector. Larger circles represent regions with higher numbers of deaths. Purple circles indicate a higher proportion of death certificates are missing, whilst red indicates that a lower proportion is missing. Background map was sourced from OpenStreetMap (OpenStreetMap, 2020b) **with extraction coordinates: latitude (55.2, 57.0), longitude (-7.0, -3.9).**

Figure 3.8: Map of Postcode Sectors within NHS GGC Region 2015 (History & Maps, 2015)

3.2.4 Age Restrictions

Lipid-lowering prescriptions and lipid profile blood tests were removed from the records if they occurred when the individual was less than 18 years old at the time. The individual's age was calculated using the following equation:

$$(date\ of\ event - date\ of\ birth)/365.25$$

Equation 3.1: Calculation of Age in Years

As dates of birth were provided as the fifteenth of the relevant month and year for the individual, this age will be an approximation. Individuals were subsequently removed completely if there were no prescriptions for lipid-lowering drugs or lipid profile blood test results reported for them in the period of interest when they were over the age of 18 years (n=39,448).

3.3 Defining the Post Myocardial Infarction Population

3.3.1 Justification of Choice of Population

The post MI population was selected as it is widely accepted that patients who have previously experienced an MI are, and remain, at high risk of further MIs without the need to consider other risk factors. Indeed, the majority of cardiovascular risk assessments exclude those with established CVD from their use with guidelines enforcing the notion that they should be treated and monitored regardless (NICE, 2014; Stone *et al.*, 2014; Piepoli *et al.*, 2016; SIGN, 2017), and the World Health Organisation has issued clear guidance regarding the treatment of such patients (World Health Organization, 2017a). For the purposes of this study, a post MI population would be expected to have a high proportion of patients with a statin prescription given the clear clinical guidelines, and therefore this group represents a cleaner population than primary prevention patients. Although the terms 'secondary prevention' and 'established cardiovascular disease' do not exclusively refer to MIs alone, they form one of the key subsections of these populations (SIGN, 2017). Patients could also be defined with relative ease from the data available (all MIs would result in a hospital admission), and benefit from an easily definable start date for follow up (the date of their MI).

3.3.2 Definition of New Hospital Stay

During an inpatient stay at a hospital, multiple records are often generated in the SMR01 database with each transfer to a new ward (or transfer between hospitals) counting as a new episode and consequently generating a new record. Without a method of identifying which records were part of the same stay, this would result in the overestimation of the number of stays in a hospital made by a patient (and result in a single event such as an MI being counted multiple times). The method implemented to identify new stays was similar to that taken by The Health & Social Care Information Centre (HSCIC) methodology used in the national datasets (Health and Social Care Information Centre, 2014).

An episode was considered to be part of the same stay if for the same individual:

- The admission date was on or before the previous discharge date

  OR

- The admission date was less than or equal to two days after the previous discharge date

  AND

  - The previous discharge type was ‘Transfer within the same Health Board/ Health Care Provider’ (Code 12)

    OR

  - The previous discharge type was ‘Transfer to another Health Board/ Health Care Provider’ (Code 13)

    OR

  - The admission type was ‘Planned Transfers’ (Code 18)

This is further illustrated as an example in Figure 3.9. Under these conditions, a new stay is the first episode which is not considered to be part of the previous stay for the same patient. Once identified, each new stay was numbered consecutively to create a hospital stay ID.

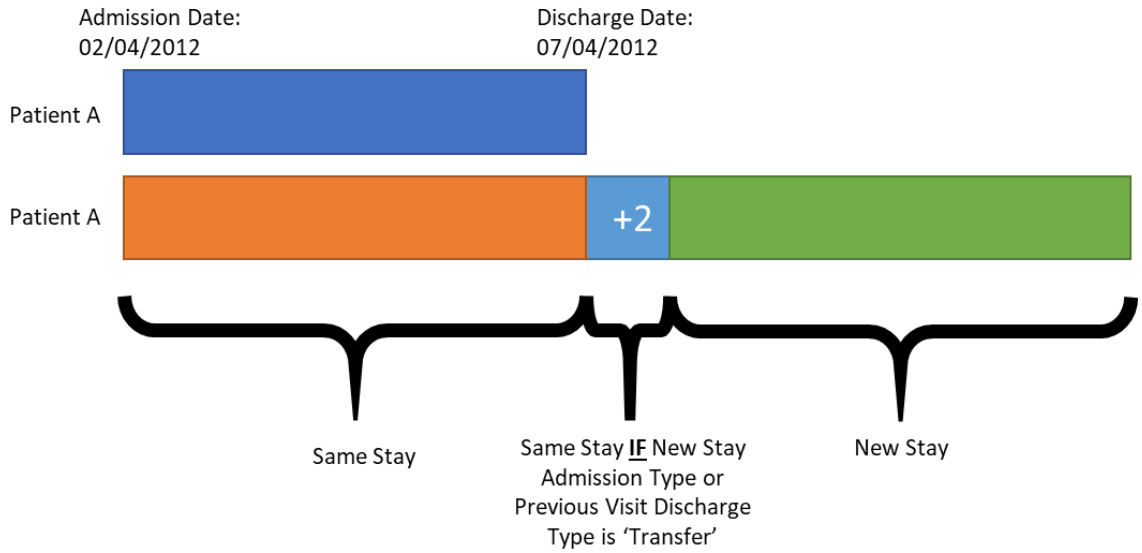


Figure 3.9: Methodology to Determine New Hospital Stays (Schematic Example)

3.3.3 Definition of Myocardial Infarction of Interest for Analysis

Figure 3.10 details the number of records retained at each stage of the derivation process to extract the post MI population. The hospital admission (SMR01) file was first restricted to records where any of the six diagnosis codes captured contained the ICD10 codes I21 (acute myocardial infarction) or I22 (subsequent myocardial infarction). Duplicate records were then removed (which occurred as a result of other episodes captured as part of the same stay) so that the number of unique MIs could be identified.

The dates of admission were then used to select events where there were potentially three years or more of follow up data (death may occur before that, however), as the longer-term management of this population was the focus of this thesis. As the date for the end of follow up was 31st July 2017, this made the time window of interest for MI admissions 1st January 2009 and 31st July 2014 (inclusive). However, it is important to note that in the absence of a reported death, patients are assumed to be still alive and within the NHS GGC region until 31st July 2017, with no formal indicator to confirm that this is the case. Consequently, by restricting the cohort to those with the possibility of three years follow up, this extended time window may mean that the population has a higher proportion of patients who have moved out of the region than in a population of post MI patients without this constraint. Patient relocation could be due to systematic differences in their characteristics, such as frailty or

severity of illness, and could result in some biases in the estimation of the frequency of patient outcomes and behaviour.

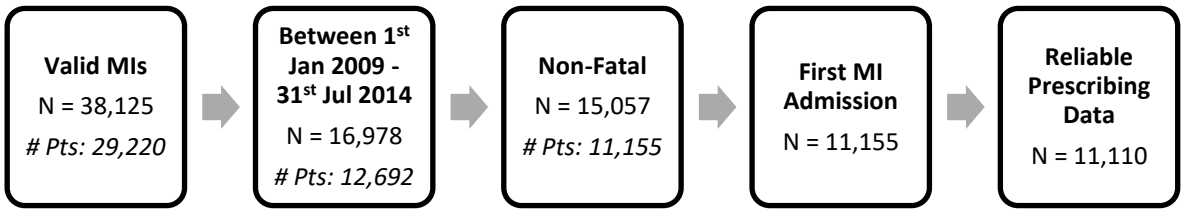


Figure 3.10: Identification of MIs for the Post MI population. A valid MI was defined as ICD10 codes I21 or I22. An MI was non-fatal if no death occurred during admission or within 30 days of discharge. Reliable prescribing data refers to patients where estimated adherence did not exceed 200% for any given year (further details given in Chapter 3.4.3).

3.3.4 Creation of Subsequent Time Windows

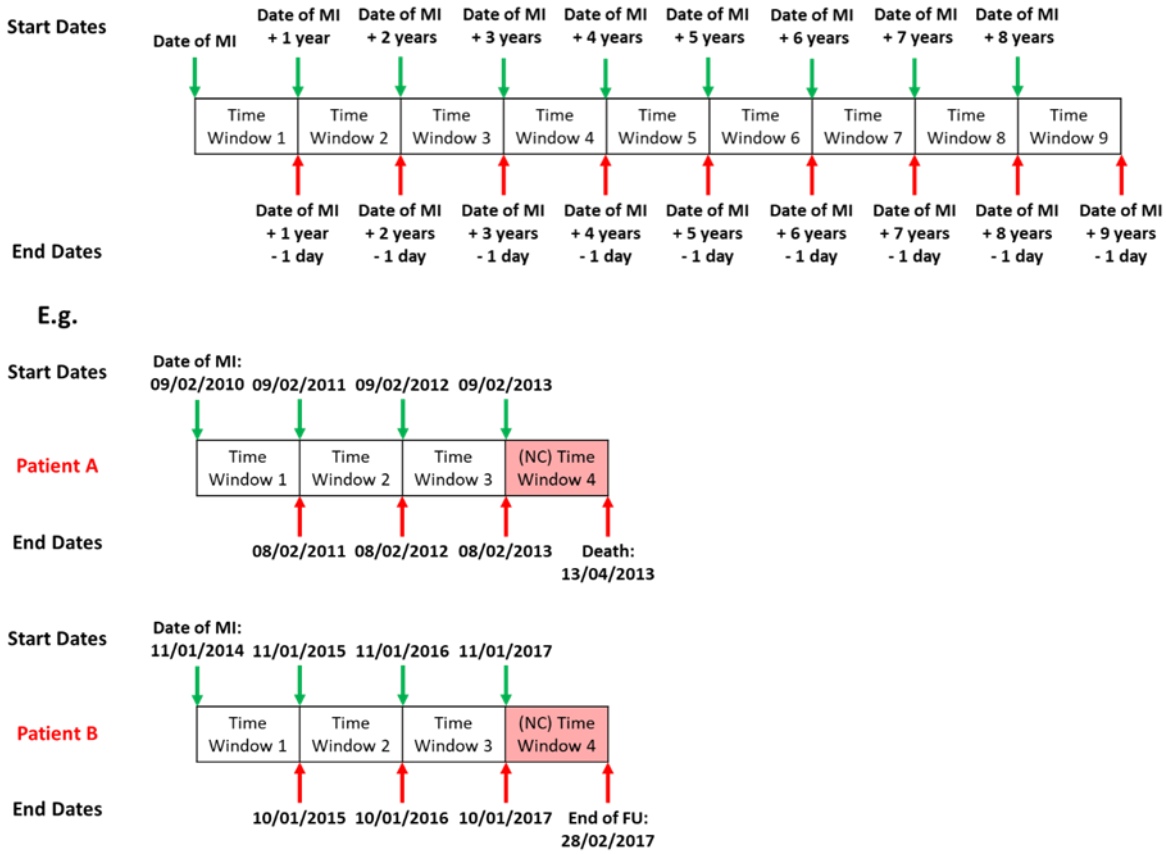


Figure 3.11: Annual Time Windows Constructed for each Patient (Schematic Examples). NC = Not Complete (year). Green and red arrows represent the start and end dates for each time window respectively.

The list of relevant MIs was then restricted further to those which were non-fatal. MIs were considered to be fatal if the patient died during the admission or within the first 30 days after hospital discharge, with non-fatal MIs being the remainder. For patients with more than one MI captured during the period, only

the first MI was used. The admission date for the MI was used as the start date of follow up for each patient. The list of patient IDs was then used to extract the corresponding prescribing, laboratory, and demographic data.

For each patient in the post MI population, a set of time windows was constructed to capture information (such as adherence) available in each year of follow up since the date of admission (as illustrated in Figure 3.11). These windows were adjusted for leap years to ensure that each time window started on the same date of each year. Given the earliest date that an MI could be counted was 1st January 2009, the maximum number of windows available for a patient was nine. Any time windows starting after the end of follow up (31st July 2017) or after the patient's death were removed. For time windows where the end of follow up or death occurs, the end date for the window was adjusted accordingly and an indicator variable was created to highlight that this time window did not represent a complete year of follow up.

3.4 Statin Prescription Data

Statin prescription records were used to approximate a patient's statin adherence through the calculation of the Medication Possession Ratio (MPR) for each year post MI. Prescriptions for Atorvastatin, Fluvastatin, Pravastatin, Simvastatin, Rosuvastatin, Simvastatin/Ezetimibe, and Simvastatin/Fenofibrate, where the latter two represented combination drugs, were included. Dispensed dates were used preferentially over prescribed dates as this was the first date that the patient could be in possession of the medication. If the dispensed date had been left blank on entry then this is automatically set as the paid date by the Information Services Division which is the last day of the month (Information Services Division, 2010).

3.4.1 Adjusting for Concomitant Statin Use

A prescription's end date was calculated as day the drug was dispensed plus the day-coverage of statins dispensed. This model assumes that one statin tablet was taken per day (NHS, 2018), that the patient took their first dose on the day it was dispensed, and that the end date is the first day that the patient would have no medication. Prescriptions with an estimated end date after a patient's

death had the end date revised to the date of death, and the quantity dispensed was updated to reflect the number of doses that could have been taken by the patient. Prescriptions dispensed after death were removed completely.

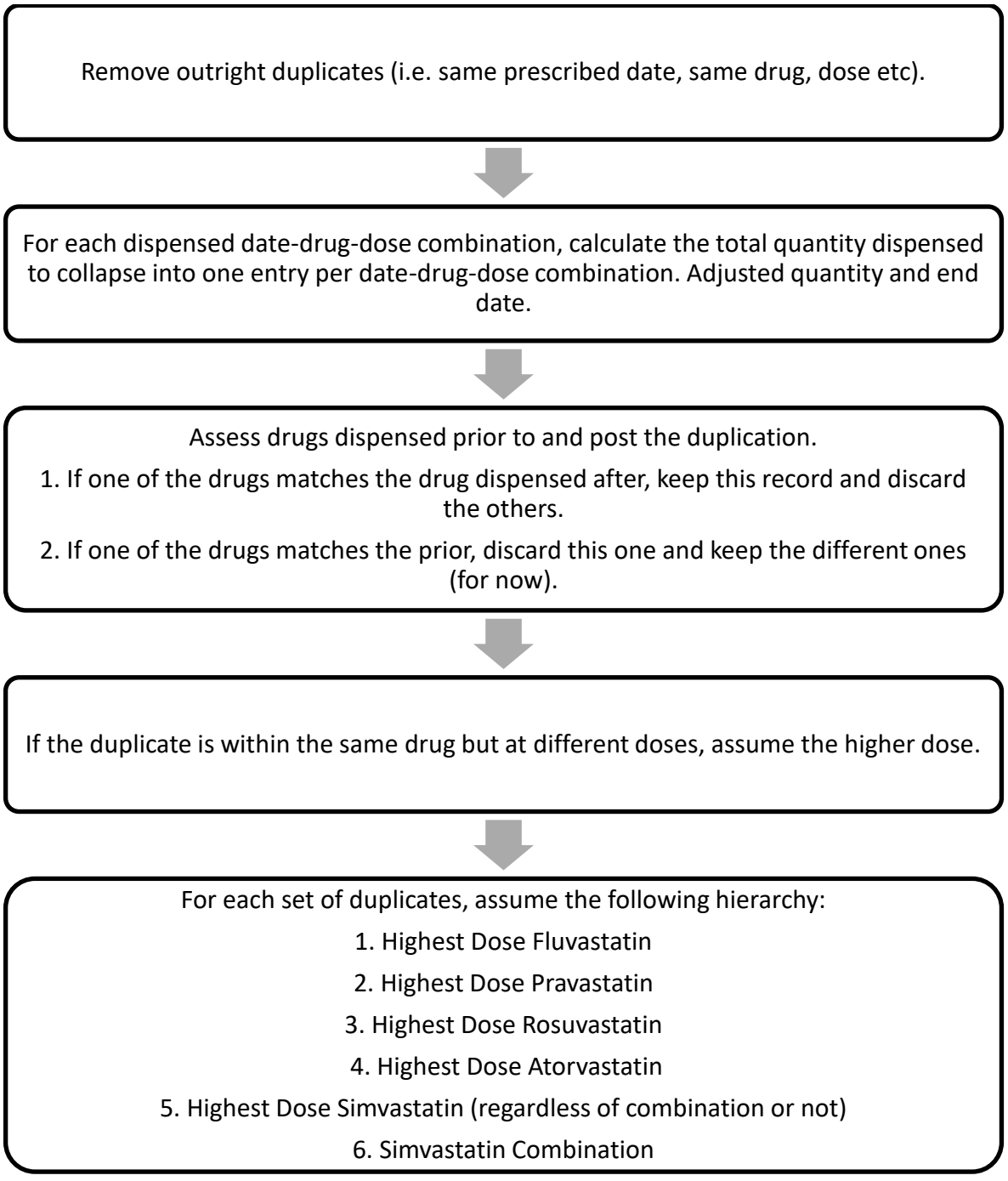


Figure 3.12: Algorithm and Hierarchy for Multiple Prescriptions. This method was only implemented on cases where an individual was dispensed multiple prescriptions on the same date (regardless of drug or dose). For the final part, the first drug that occurs in each duplicate set should be kept at its highest dose e.g. if Fluvastatin 40mg, Fluvastatin 20mg and Rosuvastatin 5mg were all dispensed on the same day, Fluvastatin 40mg should be retained only.

Before further adjustment, an algorithm was implemented to adjust for cases where a patient was dispensed more than one prescription for a statin on the

same day. This was necessary as a patient should not be assumed to be taking more than one statin at a time, and such duplication is likely the result of a delay in updating the patient's current prescription. Therefore, inclusion of the multiple prescriptions in the calculation would result in an inflated and inaccurate estimation of the patient's adherence. The algorithm used to handle such instances is detailed in Figure 3.12.

Once this was complete, if any overlap occurred between the dates for two different statins (and doses) in a patient, an adjustment was made such that the end date for the previous prescription was the dispensed date of the new one (as the end date would be the first day with no medication), and quantities were also adjusted. Prescriptions for the combination drugs, Simvastatin/Ezetimibe and Simvastatin/Fenofibrate, were then merged with the Ezetimibe and Fibrate prescriptions respectively where a similar process was conducted to prevent any further overlap occurring. The adjusted dates and quantities from this process were then recombined with the statin prescriptions. The complete details on the stages of this process can be found in Appendix A.

3.4.2 Estimating Adherence with Medication Possession Ratio

The MPR was used to estimate a patient's adherence to statin medication for each of the full year time windows constructed. Whilst concerns regarding overestimation of adherence using this method have been documented (Raebel *et al.*, 2014), this likelihood was reduced as dispensed quantities and end dates of treatment had been adjusted to ensure that a patient was only dispensed one medication in the class at a time. However, early refilling issues within the same statin and dose would have remained (Raebel *et al.*, 2014). Despite this, the fixed length of the time windows (at 365 or 366 days) used for this analysis does make MPR preferable in this scenario (Sperber, Samarasinghe and Lomax, 2017). Nonetheless, as with all methods used for estimating adherence using secondary data, MPR assumes that possession of the medication corresponds to a patient's compliance, which may not be the case (Lam and Fresco, 2015). However, whilst possession of the medication isn't a sufficient condition to ensure a patient's adherence with the medication, it is a necessary one, and as a result, such methods have been extensively used in the literature (Donnan, MacDonald and

Morris, 2002; Donnelly *et al.*, 2008; Wei, Fahey and MacDonald, 2008; Shetty *et al.*, 2016).

For the post MI population, the statins prescription file was first restricted to those only issued with an end date after the MI admission date of interest and with a start date before the end of follow up. Prescriptions were then matched to the time windows. Prescriptions which fell over two time windows were matched to both, and their quantities adjusted in each time window to reflect the number of doses that should have been available to be taken by the patient if they were fully compliant with the medication schedule. The MPR for each time window was then calculated using the formula (Raebel *et al.*, 2014; Lam and Fresco, 2015; Sperber, Samarasinghe and Lomax, 2017):

$$\left({}^{total\ number\ of\ doses\ dispensed\ in\ time\ window}/_{length\ of\ time\ window}\right) x\ 100$$

Equation 3.2: Medication Possession Ratio (MPR)

This calculation was only made for time windows that were a full year (365 or 366 days) in duration, and not those which were shorter due to end of follow up or death. The full details of this method can be found in Appendix B.

3.4.3 Handling of Excessively High Adherers

After this calculation, however, it was noted that some patients achieved an MPR higher than 100% (i.e. taking the statin as prescribed). Small increases above 100% would be expected if a patient was dispensed further medication before the previous prescription was completely used, which is a known limitation of MPR (Raebel *et al.*, 2014). However, where the calculated MPR was considerably higher than 100%, this raised concerns regarding the accuracy and reliability of the dispensing records. To fully assess the scale of this, an individual line graph was plotted to examine trends of MPR for patients who had an MPR greater than 120% at any point during follow up (Figure 3.13). Patients with an MPR greater than 200% (n=45, 0.4% Post MI) were subsequently removed. A look at the prescription records for these patients found they were due to, sometimes persistent, human-error in the quantities dispensed variable.

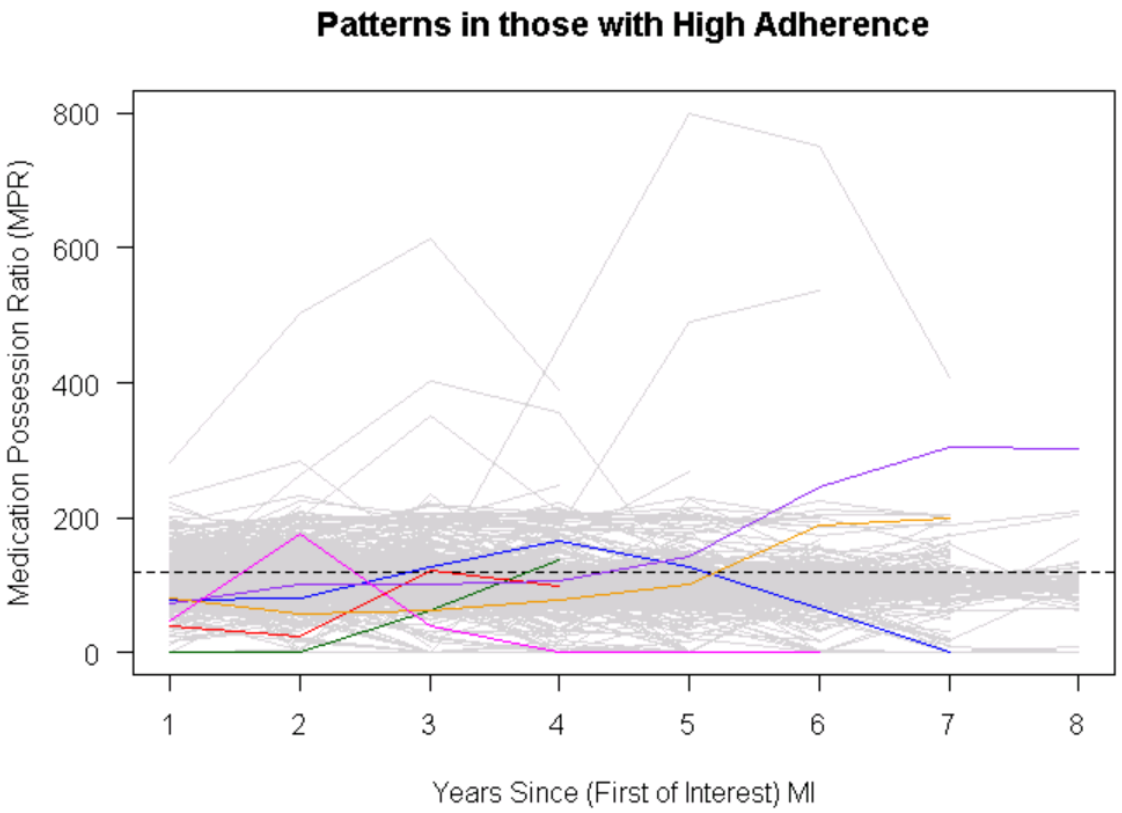


Figure 3.13: MPR Patterns for Patients with High Adherence. Patients were included if their estimated MPR exceeded 120% for any year during follow up. Coloured lines show some of the patterns in MPR for a selection of individual patients.

3.5 Lipid Profile Results

A lipid profile result consisted of five components ("a set"): Total Cholesterol, Total Cholesterol:HDL Cholesterol, HDL Cholesterol, LDL Cholesterol, and Triglycerides. However, not all of these components were available for each sample, and in some cases, multiple results for the same component were reported.

3.5.1 Handling Duplicate Results

Multiple results for one component from the same sample became problematic when the five components of the set were subsequently transposed from long to wide format and matched by the patient, sample date, and time, as shown in Figure 3.14. In such instances, these records needed to be simplified such that only one test result was reported at each time point.

First, complete duplicates were removed. Total Cholesterol:HDL Cholesterol and LDL Cholesterol values were then manually calculated, using the formulae detailed in the SIGN guidelines (SIGN, 2017, p. 39):

$$Total : HDL\ Cholesterol = \frac{Total\ Cholesterol}{HDL\ Cholesterol}$$

Equation 3.3: Total Cholesterol:HDL Cholesterol

$$LDL\ Cholesterol = Total\ Cholesterol - HDL\ Cholesterol - \frac{Triglycerides}{2.2},$$
$$where\ all\ values\ are\ measured\ in\ mmol/l, and\ triglycerides\ are\ \leq 4.5\ mmol/l.$$

Equation 3.4: LDL Cholesterol

ID	Date	Test	Result
1	03/09/2012	TC	5.4
1	03/09/2012	TC	5.0
1	03/09/2012	HDL	3.2
1	03/09/2012	HDL	3.4
1	03/09/2012	LDL	0.7
1	03/09/2012	LDL	0.6
1	03/09/2012	Trigs	3.4
1	03/09/2012	Trigs	2.2
1	03/09/2012	T:HDL	1.7
1	03/09/2012	T:HDL	1.5

ID	Date	TC	HDL	LDL	Trigs	T:HDL
1	03/09/2012	5.4	3.2	0.7	3.4	1.7
1	03/09/2012	5.4	3.2	0.7	3.4	1.5
1	03/09/2012	5.4	3.2	0.7	2.2	1.7
1	03/09/2012	5.4	3.2	0.7	2.2	1.5
1	03/09/2012	5.4	3.2	0.6	3.4	1.7
1	03/09/2012	5.4	3.2	0.6	3.4	1.5
1	03/09/2012	5.4	3.2	0.6	2.2	1.7
1	03/09/2012	5.4	3.2	0.6	2.2	1.5
1	03/09/2012	5.4	3.4	0.7	3.4	1.7
1	03/09/2012	5.4	3.4	0.7	3.4	1.5
1	03/09/2012	5.4	3.4	0.7	2.2	1.7
1	03/09/2012	5.4	3.4	0.7	2.2	1.5
1	03/09/2012	5.4	3.4	0.6	3.4	1.7
1	03/09/2012	5.4	3.4	0.6	3.4	1.5
1	03/09/2012	5.4	3.4	0.6	2.2	1.7
1	03/09/2012	5.4	3.4	0.6	2.2	1.5
1	03/09/2012	5.0	3.2	0.7	3.4	1.7
1	03/09/2012	5.0	3.2	0.7	3.4	1.5
1	03/09/2012	5.0	3.2	0.7	2.2	1.7
1	03/09/2012	5.0	3.2	0.7	2.2	1.5
1	03/09/2012	5.0	3.2	0.6	3.4	1.7
1	03/09/2012	5.0	3.2	0.6	3.4	1.5
1	03/09/2012	5.0	3.2	0.6	2.2	1.7
1	03/09/2012	5.0	3.2	0.6	2.2	1.5
1	03/09/2012	5.0	3.4	0.7	3.4	1.7
1	03/09/2012	5.0	3.4	0.7	3.4	1.5
1	03/09/2012	5.0	3.4	0.7	2.2	1.7
1	03/09/2012	5.0	3.4	0.7	2.2	1.5
1	03/09/2012	5.0	3.4	0.6	3.4	1.7
1	03/09/2012	5.0	3.4	0.6	3.4	1.5
1	03/09/2012	5.0	3.4	0.6	2.2	1.7
1	03/09/2012	5.0	3.4	0.6	2.2	1.5

Figure 3.14: Example of Transposition-Induced Duplication of Lipid Profiles. This example is included for illustration purposes only. In this case, each of the five components is duplicated just once, but when these are transposed 32 records are produced.

Records where both manually calculated values matched the reported values (allowing for rounding error) (Goldberg, 1991) were then retained, along with the most complete sets. Those with triglycerides too high for an LDL result (>4.5mmol/l) to match were then retained if they matched on Total

Cholesterol:HDL Cholesterol alone. The mean values for each of the components for any remaining duplicates for a sample were then calculated. The full details of the method undertaken to handle duplicated lipid results can be found in Appendix C.

3.5.2 Matching Lipid Profile Results to Time Windows

The lipid profile results were then restricted to the post MI population and subsequently combined with the time window dates. Tests were matched to a time window if the sample was taken within the dates of the time window. In cases where multiple tests were obtained within a time window, the mean value of each of the components from all of the tests in that window was derived. These values could then be used to generate indicator variables to ascertain whether a patient met guideline specific lipid reduction criteria in any given time window (on average) e.g. LDL Cholesterol ≤1.8mmol/l from the 2016 European Society of Cardiology (ESC) Guidelines (Piepoli *et al.*, 2016).

3.5.3 Calculating Percentage Change in Non-HDL Cholesterol

One target of interest was the percentage reduction of non-HDL cholesterol; the primary focus of the current NICE guidelines (NICE, 2014). For this, a baseline sample where the patient was not on statin medication was required, and the number of the patients included at each stage of its derivation is shown in Figure 3.15. Firstly, the lipid profile data was restricted to tests conducted before the MI at the start of a patient's follow up. These tests were then combined with the statin prescription dates. A test was considered a suitable baseline if it was more than six months after the estimated end date (182 days) of a statin prescription or if the patient had no prescriptions for statins prior to the date of the MI. If multiple tests were considered to be suitable for a patient's baseline result, the test closest to the date of MI was selected. If no tests were considered to be suitable, then a percentage reduction could not be calculated for any of the patient's time windows, and the result was set to missing.

Once a baseline test was selected and merged with the time window information, non-HDL Cholesterol was calculated as (SIGN, 2017, p. 39):

$$nonHDL\ Cholesterol = Total\ Cholesterol - HDL\ Cholesterol$$

Equation 3.5: Non-HDL Cholesterol

This was only calculated for time windows and baseline tests where both of these values were present. Finally, the percentage change was calculated as:

$$\left({}^{(time\ window\ nonHDL\ -\ baseline\ nonHDL)}/_{baseline\ nonHDL} \right) x\ 100$$

Equation 3.6: Percentage Change in Non-HDL Cholesterol

Full details of the method used can be found in Appendix D.

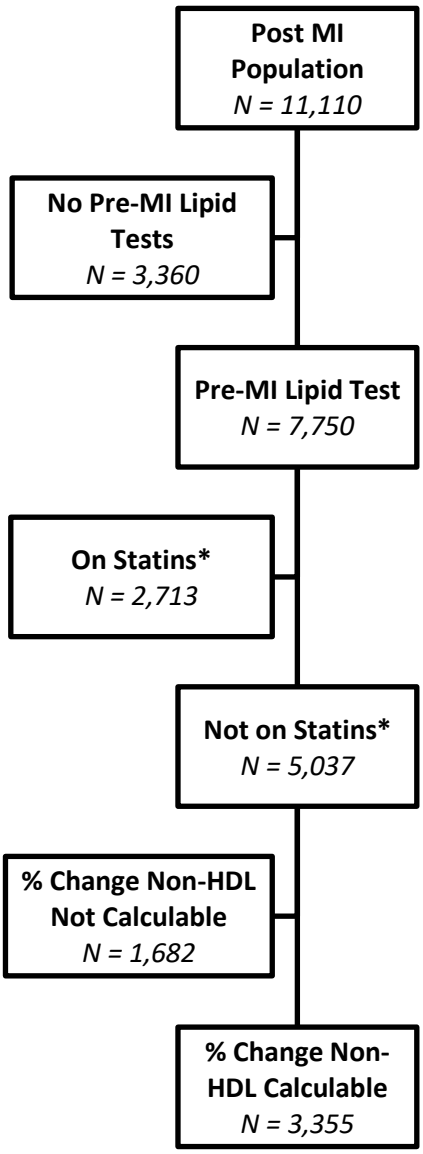


Figure 3.15: Selection Process for Deriving the Percentage Change in Non-HDL Cholesterol. *On statins defined as either the result was obtained during a prescription or within 182 days of its end date.

3.5.4 Lipid Testing Rates

The total number of tests during a patient's follow up was also calculated, along with the number of tests within each time window. Within each time window, patients were then categorised as having no tests, recommended testing (one test), or over-testing (more than one test). Over the entire duration of follow up, a patient's testing rate was calculated as:

$$\left({}^{total\ number\ of\ tests\ in\ follow\ up}\Big/_{duration\ of\ follow\ up\ (days)}\right) x\ 365.25$$

Equation 3.7: Rate of Lipid Testing

3.6 Demographics and Covariates

3.6.1 Age at Myocardial Infarction

The age of the patient was calculated on the day they were first admitted with an MI. This was calculated as:

$$(date\ of\ admission - date\ of\ birth)\ /_{365.25}$$

Equation 3.8: Age (in years) at MI

As stated earlier (Chapters 3.1 and 3.2.4), the calculated age of the patient will be an approximation as the dates of birth were given as the fifteenth of the relevant month and year in the demographics file.

3.6.2 NHS GGC Specific Scottish Index of Multiple Deprivation

The demographics file provided various measures of the patient's deprivation status which were derived from their postcode. This included the Scottish Index of Multiple Deprivation (SIMD) 2012 rank and its associated quintile. These were derived by ranking data zones (which represent small areas) across Scotland in 2012 from the most deprived to the least deprived. Areas in the top 20% of the ranks would, therefore, constitute the most deprived quintile (Scottish Government, 2018).

However, approximately 40% of the population of the NHS GGC area are estimated to live in areas in the most deprived quintile (NHS Greater Glasgow and Clyde, 2015b, chap. 1), with 30% of the 15% most deprived zones nationally found in Glasgow City in 2013 (Scottish Government, 2013). This meant that the NHS GGC population was unlikely to be representative of the overall Scottish population, or its post MI population. Therefore, NHS GGC specific SIMD quintiles were derived, so that the role of deprivation could be effectively compared within the post MI population of NHS GGC. For the majority of summaries, only these NHS GGC specific quintiles are presented.

To generate these quintiles, a list of the data zones in the NHS GGC area was sourced, which showed that there were 1,473 data zones in the area in 2012 (Scottish Government, 2015). These were ordered by rank and subsequently divided into quintiles, based on the number of data zones included. The cut points for these quintiles were then equated back to the Scottish SIMD ranks for 2012 and are shown in Table 3.2.

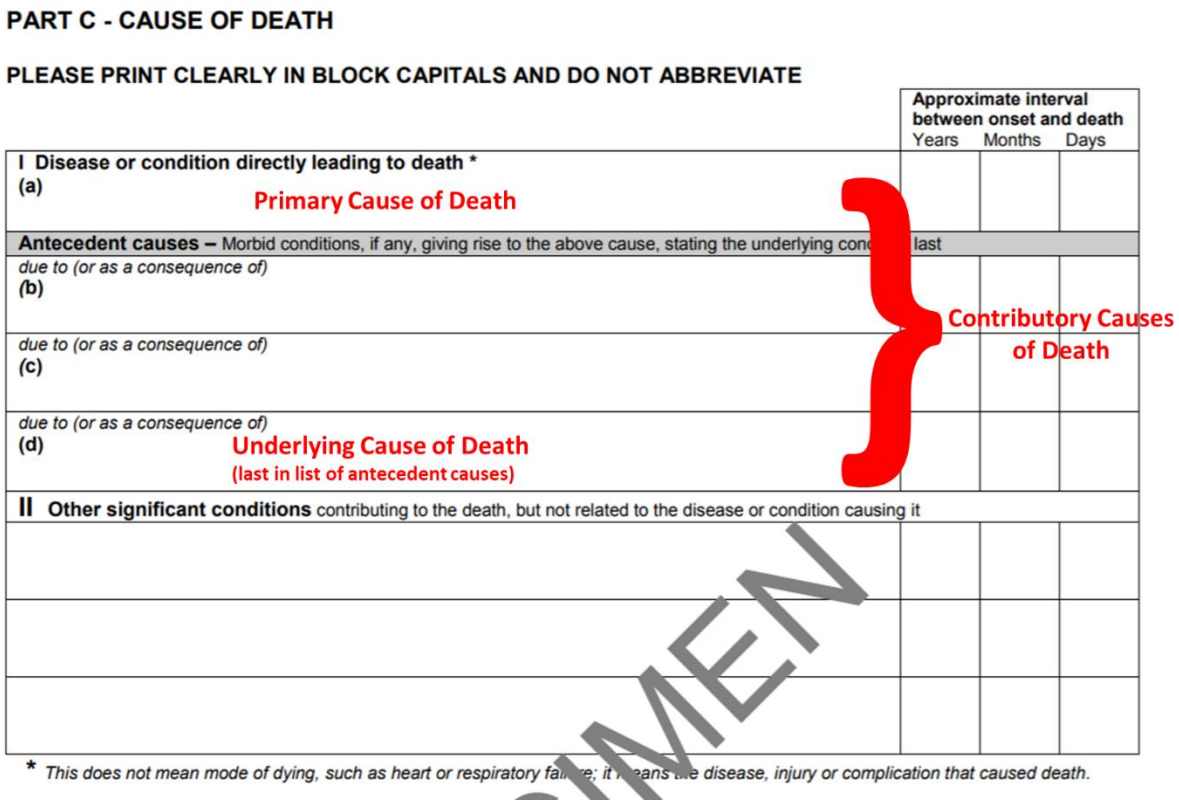
Table 3.2: NHS GGC Specific SIMD 2012 Quintiles. All cut points are inclusive, and ranks refer to the ISD (National) Ranks

Rank	Range of Ranks
1	1-542
2	543-1,527
3	1,528-3,104
4	3,105-5,101
5	5,102-6,505

3.6.3 Causes of Death

Causes of death were available for all individuals with a death certificate recorded. These were first formatted to the shortened ICD10 codes (retaining only the letter and the first two numbers). A series of consecutively numbered cause of death variables were included, with the primary cause of death coded with zero, and the underlying cause of death given in a separate non-numerically coded variable. Primary and antecedent causes of death were then extracted and separated from other significant conditions. The primary cause and antecedent causes were also referred to as contributory causes when considered together. The process of extraction was derived to reflect how such information

would have been captured and structured on the paper version of a death certificate, as shown in Figure 3.16 (National Records Scotland, 2014a).



PART C - CAUSE OF DEATH

PLEASE PRINT CLEARLY IN BLOCK CAPITALS AND DO NOT ABBREVIATE

	Approximate interval between onset and death		
	Years	Months	Days
I **Disease or condition directly leading to death** * (a) **Primary Cause of Death**			
Antecedent causes – Morbid conditions, if any, giving rise to the above cause, stating the underlying cond last			
due to (or as a consequence of) (b)			
due to (or as a consequence of) (c)			
due to (or as a consequence of) (d) **Underlying Cause of Death** (last in list of antecedent causes)			
II **Other significant conditions** contributing to the death, but not related to the disease or condition causing it			

Contributory Causes of Death

* *This does not mean mode of dying, such as heart or respiratory fa re; it means the disease, injury or complication that caused death.*

Figure 3.16: Annotated Death Certificate. Annotations indicate assumed source of the causes of death definitions. Example death certificate was taken from National Records Scotland, 2014.

3.6.4 Diabetes Type and Duration

Diabetes type and the date of diagnosis were extracted from the SCI Diabetes data, which contains health and treatment data for all diabetes patients in Scotland (Scottish Care Information Diabetes Collaboration, 2015). The date of diagnosis is captured in the ‘date of entry’ column for the ‘Type of Diabetes’ observations within the data, and therefore all of these observations were extracted. Multiple records were available for each patient by visit. Dates of entry before the patient’s date of birth were considered to be invalid and were excluded. The first valid date for each patient was then assumed to be the date of diagnosis of diabetes. For patients with no valid dates of diagnosis, the date of diagnosis and consequently the duration of disease was set to missing. However, diabetes type could still be ascertained. Some discrepancy was present for some patients with the type of diabetes often listed differently at each visit, or simply listed as unknown. As a result, an algorithm was

implemented to determine the type of diabetes for each patient (as shown in Figure 3.17). Full details of the method used can be found in Appendix E.

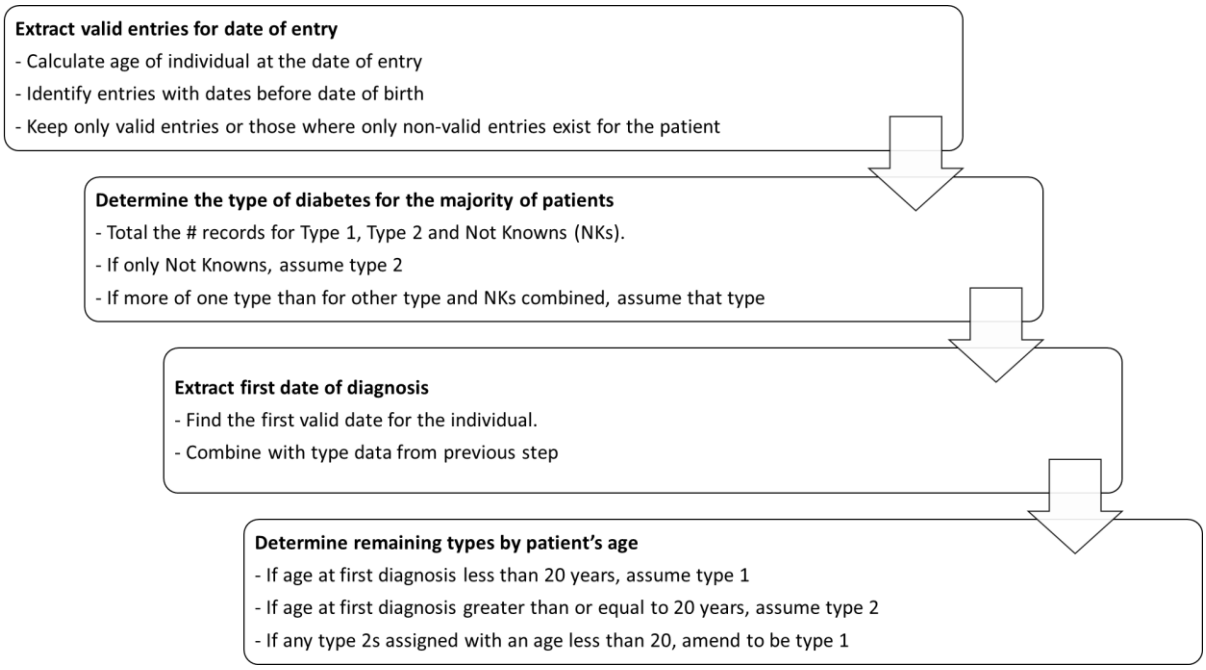


Figure 3.17: Type and Date of Diagnosis of Diabetes Algorithm. This algorithm was used for patients included in the SCI Diabetes database using the information captured in all observations with the description 'Type of Diabetes'.

Once a date of diagnosis and type of diabetes was determined, this was restricted to only members of the post MI population, and those diagnosed before the end of follow up. This was combined with the demographic and time window data. The duration of disease (in years) at the time of MI was calculated using the formula below. The minus one was included to identify those whose diabetes status was confirmed before their admission with an MI (as it could have been confirmed at the visit). This also prevented any discrepancy when identifying patients diagnosed during follow up, as the date of MI forms the start date of the first window.

$$(date\ of\ MI - 1 - date\ of\ diagnosis)\,/\,365.25$$

Equation 3.9: Diabetes Duration Prior to MI

In the time window data, a similar calculation was used to determine the duration of disease at the end of the time window. A minus one adjustment was not necessary for this instance as the start and end dates of the time windows were inclusive.

$$(end\ date\ of\ time\ window - date\ of\ diagnosis)/365.25$$

Equation 3.10: Diabetes Duration in Time Windows

An indicator variable was also derived to highlight time windows where the diagnosis of diabetes was new.

3.6.5 Duration of Follow Up

Each patient was followed up from the day of the admission with their first MI (on or after 1st January 2009) until death (or 31st July 2017), whichever occurred first. Therefore, a patient's time at risk, or duration of follow up, was calculated using the equation below. The plus one adjustment is included to allow the inclusion of both the start and end date in the time at risk. Though the equation below would yield the duration of follow up in days, it can be converted into years by dividing by 365.25.

$$Patient's\ Time\ at\ Risk = Patient's\ End\ Date - Patient's\ Start\ Date + 1$$

Equation 3.11: Patient's Duration of Follow Up (days). A patient's end date was defined as 31st July 2017 if death did not occur earlier. A patient's start date was the first day of their admission with an MI on or after 1st January 2009.

3.6.6 Duration of Stay

The duration of the hospital stay was defined similarly (Equation 3.12). The plus one adjustment is included so that if a patient was admitted and discharged on the same day, the duration of stay would be equal to one day. Therefore, the duration of stay does not reflect the number of nights admitted, but the number of partial days in the hospital.

$$Duration\ of\ Stay = Visit\ Discharge\ Date - Visit\ Admission\ Date + 1$$

Equation 3.12: Duration of Hospital Stay (days)

3.6.7 Time Since MI (Prior)

MIs documented in the ten years prior to the first non-fatal MI of interest were also captured. If an individual had multiple previous MIs, then all were retained, and the number of them captured in a separate variable. The time elapsed

between the prior MI and the MI of interest was calculated for all patients with a prior MI using a similar method to the other duration variables previously discussed (Equation 3.13). Values were converted into years by dividing by 365.25.

$$Time\ Since\ Prior\ MI = Date\ of\ Admission - Prior\ Date\ of\ Admission + 1$$

Equation 3.13: Time Since Prior MI (days). Date of Admission refers to the date the patient was first admitted with the non-fatal MI of interest on or after 1st January 2009 (the patient's start of follow up date). Prior Date of Admission is the visit start date for the most recent MI in the ten years before this start date.

Chapter 4 Describing the Post Myocardial Infarction Population

4.1 Introduction

NHS GGC is formed from six local authorities: Glasgow City, Renfrewshire, East Dunbartonshire, East Renfrewshire, West Dunbartonshire, and Inverclyde. In 2013, its population was estimated to represent 21.4% of the Scottish population with over 1.1 million individuals, with a slightly higher percentage of females (51.4%) (NHS Greater Glasgow and Clyde, 2015b). This gender divide was similar to the overall Scottish population (National Records Scotland, 2014b). The age distribution is also comparable, with similar numbers of those under the age of 20 (NHS GGC: 21.6%, Scotland: 21.9%), although there was a slightly lower percentage of those over the age of 65 in NHS GGC (16.2% vs 17.8%). There was also some regional variation in the age distribution within NHS GGC, with 13.9% in the Glasgow City local authority over the age of 65, compared to 20.8% in East Dunbartonshire (NHS Greater Glasgow and Clyde, 2015b).

The population of interest in the remainder of this thesis, namely those who have experienced a non-fatal MI, however, is not the entire population of NHS GGC region as the presence of the risk factors identified in Chapter 1.2 will impact on an individual's likelihood of experiencing a cardiovascular event. Consequently, the prevalence of, as well as the interactions between, these will all influence the prevalence of CVD within the NHS GGC population, and therefore the size of the population of interest for this thesis. For example, a higher proportion than expected live in areas within the most deprived quintile (35.9%) (NHS Greater Glasgow and Clyde, 2015b), and the Scottish Health Survey in 2015 found higher rates of smoking and exceeding recommended alcohol limits than the national average (NHS Greater Glasgow and Clyde, 2015c).

Furthermore, the prevalence of diabetes mellitus was estimated to be 5.5% in the 2016 Scottish Diabetes Survey, which was broadly in line with the national average (5.4%), and 4.7% in the Quality and Outcomes Framework (QOF) registers for 2015/16, which was lower than the national average (5.0%) (ScotPHO, 2016). For several comorbidities and types of established CVD, the QOF estimates for prevalence were: 12.7% for hypertension, 2.8% for CKD, 4.0%

for CHD, and 2.1% for stroke. Except for hypertension, where this rate was slightly lower than the national average (13.9%), all of these estimates were approximately similar to national averages (Information Services Division, 2016).

Nonetheless, the British Heart Foundation statistics for 2020 report that the premature death rate from CVD in Glasgow is the highest in the UK, with rates in Scotland also higher than the other nations. This is likely the consequence of combinations of many risk factors, with deprivation often demonstrating large variations in the rates of CVD, hypertension, and diabetes within Scotland, with all approximately twice as prevalent in areas that are more deprived than least deprived (British Heart Foundation, 2020a).

However, research has also shown that many of the risk factors associated with CVD are also associated with a patient's adherence to medication and subsequently their achievement of target lipid levels, as discussed in Chapters 1.4.2 and 1.4.3. For example, those with a diagnosis of diabetes are more likely to be adherent (Hope *et al.*, 2019) and meet guideline-recommended lipid targets (Danese *et al.*, 2017).

This chapter, therefore, aims to understand the post MI population in NHS GGC. To achieve this, descriptive statistics of the cohort derived in Chapter 3 will summarise their demographics, their adherence to LLM, their achievements of lipid-lowering targets from clinical guidelines, and the frequency and patterns of the testing of their lipid levels. It is hypothesised, that this population will be broadly similar to other post MI and secondary prevention cohorts within the literature.

4.2 Population Summary

4.2.1 Baseline Demographics

Baseline demographics for the post MI population in NHS GGC are shown in Table 4.1. There were 11,110 patients who experienced at least one non-fatal MI (ICD10: I21, I22) between 1st January 2009 and 31st July 2014 (inclusive), without an annual MPR greater than 200%.

Table 4.1: Baseline Demographics of NHS GGC's Post MI Population

	Post MI N = 11,110
Gender	
Male	6,732 (60.6%)
Female	4,378 (39.4%)
Age at MI (years)	
Mean (SD)	66.9 (13.9)
Median	67.4
IQR	56.1 – 77.9
Range	19.3 – 102.6
SIMD 2012 Quintile (Scotland)	
1 (Most)	4,760 (43.2%)
2	2,023 (18.3%)
3	1,393 (12.6%)
4	1,305 (11.8%)
5 (Least)	1,550 (14.1%)
Missing	79
SIMD 2012 Quintile (NHS GGC)	
1 (Most)	2,650 (24.0%)
2	2,528 (22.9%)
3	2,294 (20.8%)
4	1,877 (17.0%)
5 (Least)	1,682 (15.2%)
Missing	79
Diabetes at MI	2,099 (18.9%)
Type 1	120 (1.1%)
Type 2	1,979 (17.8%)
Prior MI	714 (6.4%)
1	558 (5.0%)
>1	156 (1.4%)

Numbers are N (%) unless otherwise specified. SD = Standard Deviation; IQR = Interquartile Range. Population summarised is the Post MI cohort derived for this thesis; specifically, patients who experienced a non-fatal MI in NHS GGC between 2009 and 2014, as outlined in Chapter 3.

This population was predominantly male (60.6%) and had an average age at their baseline MI admission of 67 years. Examination of the Scottish SIMD quintiles would suggest that a larger percentage of the population reside in the most deprived quintile of Scotland, with non-fatal MIs less prevalent in the least deprived areas (most vs least deprived quintile: 43.2% vs 14.1%). However, this does not take account of the high proportion of the most deprived regions of Scotland being within the NHS GGC area. When the NHS GGC specific SIMD

deprivation quintiles were used, this trend, whilst still present, is attenuated (24.0% vs 15.2%). Finally, approximately one fifth (18.9%) of the population was diagnosed with diabetes before their MI, and 6.4% had experienced a prior MI in the ten years before the start of follow up.

4.2.2 Duration of Follow Up

Details surrounding the duration of follow up and the number of years available per patient are shown in Table 4.2. The average duration of follow up per patient was approximately four years and six months, with the maximum follow up of eight years and seven months (as expected given the dates used for extraction).

Table 4.2: Summary of Duration of Follow Up

	Summary Statistic	
Duration of FU	**(days)**	**(years)**
Mean (SD)	1,646.0 (807.7)	4.5 (2.2)
Median	1,653.0	4.5
IQR	1,155.0 – 2,278.0	3.2 – 6.2
Range	35.0 – 3,134.0	0.1 – 8.6
Complete Years Available		
0 years		1,028 (9.3%)
1 years		767 (6.9%)
2 years		643 (5.8%)
3 years		2,115 (19.1%)
4 years		1,802 (16.2%)
5 years		1,619 (14.6%)
6 years		1,452 (13.1%)
7 years		1,117 (10.1%)
8 years		567 (5.1%)

Numbers are N (%) unless otherwise specified. FU = Follow Up; SD = Standard Deviation; IQR = Interquartile Range. Post MI population was followed up from the date of their first non-fatal MI between 1st January 2009 and 31st July 2014, until 31st July 2017 or death, whichever occurred first.

However, for much of the analysis, a patient's follow up is divided into year-long time windows to allow the capture of time-varying variables such as their estimated adherence and lipid profile results. Therefore, the latter half of Table 4.2 categorises patients by the number of complete (full) years of follow up that they had before death or the end of follow up occurred. Within these numbers, patients with less than three complete years of follow up, were all patients who died; for example, this means that just over 1,000 patients (9.3%) died in the

year following their MI. This is because the cohort was set up to allow at least three years of follow up before censoring occurred. Therefore, those with greater than or equal to three full years of follow up represent those alive at end of follow up (31st July 2017) as well as individuals who had died, thus explaining the marked increase in the number of patients observed from the three complete years of follow up onwards.

4.2.3 Myocardial Infarctions Prior to Period of Interest

There were 714 (6.4%) patients who had a documented MI in the ten years before their start of follow up, with 156 (1.4%) experiencing more than one. Using just the patient's most recent prior MI, the average time since their previous MI to the start of follow up was 5 years (Table 4.3). However, this ranged from 6 days to 10 years (maximum length of the lookback period).

Table 4.3: Time Since Prior MI in Post MI Population with a Prior MI

	Time Since MI (days)	Time Since MI (years)
Mean (SD)	1,851.0 (1,003.8)	5.1 (2.7)
Median	1,835.0	5.0
IQR	1,057.0 – 2,674.0	2.9 – 7.3
Range	6.0 – 3,653.0	0.0 – 10.0

SD = Standard Deviation; IQR = Interquartile Range. Only the time since the most recent prior MI is included, and only calculated for patients with a prior MI. Prior MIs were only included if the admission occurred in the ten years before the patient's start of follow up.

4.3 Baseline Admission for Myocardial Infarction

Details surrounding the MIs which defined the start of a patient's follow up are presented in Table 4.4 and include the year of admission, the first hospital of admission, and the duration of their stay.

The number of MIs in each year was calculated for each calendar year. The numbers recorded increased from around 1,700 in 2009 to approximately 2,000 per year by 2011 which then remained reasonably consistent for the remainder of the period. The decrease observed in 2014 occurs as a result of the fact that a full year's worth of data was not included (until 31st July only).

The hospital admitted was the first hospital that the patient was admitted to with the MI. This does not, therefore, allow for any transfers that may have

occurred subsequently, or prior, to the diagnosis of the MI. The Golden Jubilee National Hospital had the most admissions, and given that it is the location of the national and regional heart and lung services (Golden Jubilee National Hospital, 2019), this was in line with expectations.

There was some apparent skewing in the duration of stay data, with several long stays likely to lead to an elevated mean average stay, which is further reflected in the wide range of values. The maximum length of stay was 1,043 days, which is equivalent to 2.9 years. Many of those with longer durations of stays were for patients who went on to be transferred to long-stay geriatric medicine units or rehabilitation wards, with these stays considered to be part of the same visit under the continuous visit methodology implemented (Chapter 3.3.2).

Table 4.4: Details of Baseline MIs for the Post MI Population

	Summary Statistic
Year of MI	
2009	1,733 (15.6%)
2010	1,994 (17.9%)
2011	2,080 (18.7%)
2012	2,114 (19.0%)
2013	2,029 (18.3%)
2014	1,160 (10.4%)
Hospital Admitted	
Golden Jubilee National Hospital	3,764 (33.9%)
Glasgow Royal Infirmary	1,657 (14.9%)
Royal Alexandra Hospital	1,334 (12.0%)
New Victoria Hospital	1,197 (10.8%)
West Glasgow	1,179 (10.6%)
Queen Elizabeth University Hospital	706 (6.4%)
Inverclyde Royal Hospital	437 (3.9%)
Stobhill Hospital	317 (2.9%)
Hairmyres Hospital	243 (2.2%)
Vale of Leven General Hospital	147 (1.3%)
Other	129 (1.2%)
Duration of Stay (days)	
Mean (SD)	9.5 (17.8)
Median	5.0
IQR	4.0 – 9.0
Range	1.0 – 1,043.0

Year totals are for each calendar year. Statistics are N (%) unless otherwise stated. SD = Standard Deviation; IQR = Interquartile Range. Information is given for baseline admissions only, and not subsequent events. Hospital admitted is the first hospital patient was admitted to as part of the hospital admission, so subsequent transfers may have occurred. Numbers are lower for 2014 as baseline MIs were only included up until 31st July 2014.

4.4 Plasma Lipid Testing

4.4.1 Frequency of Lipid Testing

There were 63,940 plasma lipid concentration tests recorded on or after the day of admission with the first MI of interest. These were obtained in 10,428 individuals (93.9%). The number of tests per patient and testing rates is shown below in Table 4.5. The average number of tests per individual was between five and six, with some individuals receiving considerably more than this. On a crude level, given the mean follow up in years was four and a half years, this would suggest that individuals in this population may be being tested more frequently than the recommended annual testing. This is confirmed in the categorisation of testing rates, where nearly half of the patients had a testing rate higher than would be expected if a patient was being tested annually. Nonetheless, only a third of patients had a testing rate in line with guideline recommendations, with around one fifth having a lipid sample taken less frequently.

Table 4.5: Number of Lipid Plasma Tests (per patient) in the Post MI Population

	# Tests per Patient	Tests/Year Rate (r)
Mean (SD)	5.8 (4.2)	1.4 (1.7)
Median	5.0	1.2
IQR	3.0 – 8.0	0.8 – 1.6
Range	0.0 – 56.0	0.0 – 51.1
r=0	--	682 (6.1%)
0<r≤0.75	--	1,710 (15.4%)
0.75<r≤1.25	--	3,642 (32.8%)
r>1.25	--	5,076 (45.7%)

SD = Standard Deviation; IQR = Interquartile Range. r = rate of tests per year, calculated by the total number of tests during follow up divided by the length of follow up.

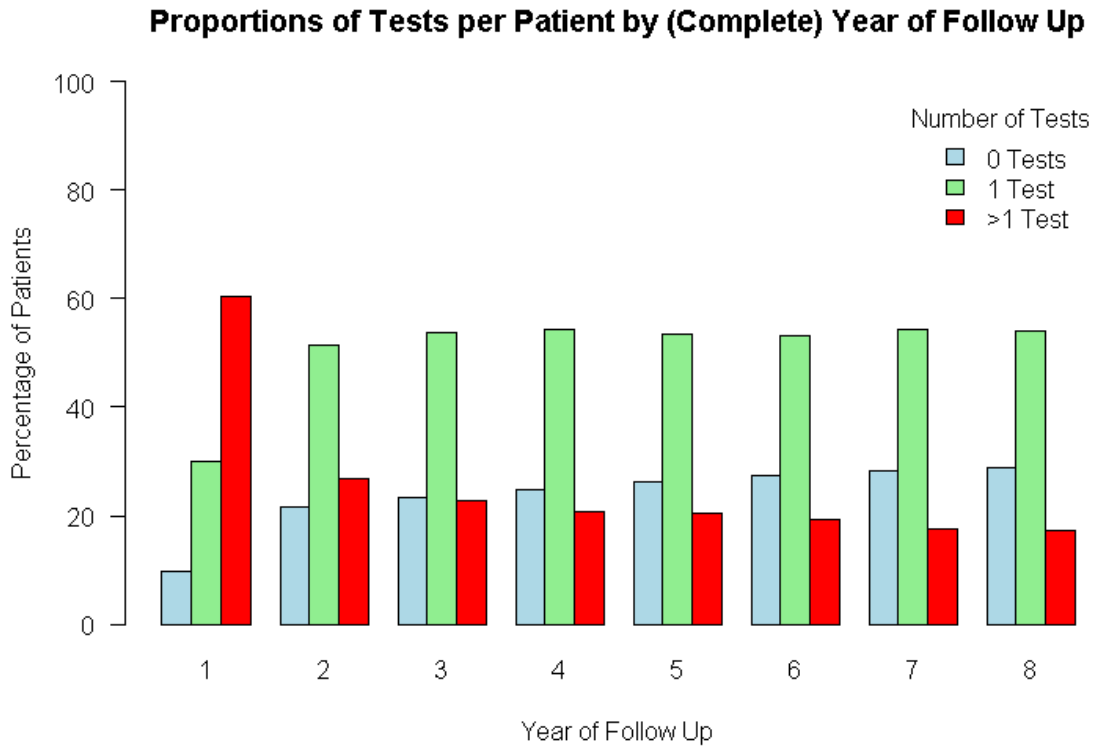


Figure 4.1: Lipid Testing per Patient by Year of Follow Up in the Post MI Population. Only complete years of follow up are included. Partial years of follow up were excluded as tests could have subsequently occurred/or been scheduled to occur before the end of the full year but after the end of follow up (31st July 2017).

When considering the complete years of follow up separately (Figure 4.1, Table 4.6), patients were more likely to have more than one test in the first year after an MI, with 60% of patients having more than one test in year 1. This could be expected as LLMs are likely to be initiated or altered after an MI and clinicians may need to review the patient more frequently to assess if sufficient reductions in cholesterol occur. However, 10% of patients did not have their lipids checked at all in the first year after their MI. The number of patients who were not tested increased to 20% in the second year of follow up, and gradually increased to nearly 30% by the eighth year. Conversely, the percentage of patients having more than one test per year decreased sharply to 27% in the second year of follow up and tapered to 17% by the last year. Thirty percent received one lipid test in their first year, in line with guideline recommendations (SIGN, 2017), which subsequently increased to over 50% in the second year and remained reasonably stable for the duration of follow up. This trend is also reflected in the summary statistics for each year. The median number of tests in the first year was two, and one test for all subsequent years. However, this analysis does

not show patterns within each patient, and patients may switch between each of these categories throughout their follow up (Figure 4.2).

Figure 4.2 attempts to elucidate the testing patterns within patients during follow up, both as a whole and within those who died during follow up. This confirmed that over-testing was more common within the first year of follow up (even when incomplete years were included), though to a lesser extent in patients who died. This difference may be partly explained by the fact that a significant number of deaths occurred in the first year, and therefore there was reduced available time to have multiple tests. Indeed, of those who died in the first year, most had no lipid tests post MI. This pattern was also observed throughout follow up where many patients had no tests in the year that they died. In both graphs, of those who did not have lipid tests in their first year, a significant proportion did not receive lipid tests throughout their follow up. This was not the case for those who had more than one test in their first year, with many having only one test in their second year and throughout the remainder of their follow up. Furthermore, most changes in the frequency of lipid testing occurred in the first couple of years, with many of these changes representing a decrease in the frequency of testing.

Table 4.6: Number of Tests per Patient by Year of Follow Up in the Post MI Population

	Year 1 10,082	Year 2 9,315	Year 3 8,672	Year 4 6,557	Year 5 4,755	Year 6 3,136	Year 7 1,684	Year 8 567
Mean (SD)	2.1 (1.5)	1.2 (1.0)	1.1 (0.9)	1.0 (0.9)	1.0 (0.9)	1.0 (0.9)	1.0 (0.9)	1.0 (0.9)
Median	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
IQR	1.0 – 3.0	1.0 – 2.0	1.0 – 1.0	1.0 – 1.0	0.0 – 1.0	0.0 – 1.0	0.0 – 1.0	0.0 – 1.0
Range	0.0 – 15.0	0.0 – 13.0	0.0 – 19.0	0.0 – 8.0	0.0 – 10.0	0.0 – 13.0	0.0 – 10.0	0.0 – 9.0
0 Tests	982 (9.7%)	2,013 (21.6%)	2,035 (23.5%)	1,624 (24.8%)	1,244 (26.2%)	862 (27.5%)	478 (28.4%)	163 (28.7%)
1 Test	3,015 (29.9%)	4,800 (51.5%)	4,672 (53.9%)	3,568 (54.4%)	2,540 (53.4%)	1,666 (53.1%)	913 (54.2%)	307 (54.1%)
>1 Test	6,085 (60.4%)	2,502 (26.9%)	1,965 (22.7%)	1,365 (20.8%)	971 (20.4%)	608 (19.4%)	293 (17.4%)	97 (17.1%)

Only complete years of follow up are included in summary statistics, as incomplete years may lead to underestimation of testing patterns (e.g. follow up may have ended before further tests being conducted within that year). Numbers included are for the number of patients with a complete year of follow up for that year post MI e.g. for year 1, 10,082 patients had a complete year of follow up available after their baseline MI.

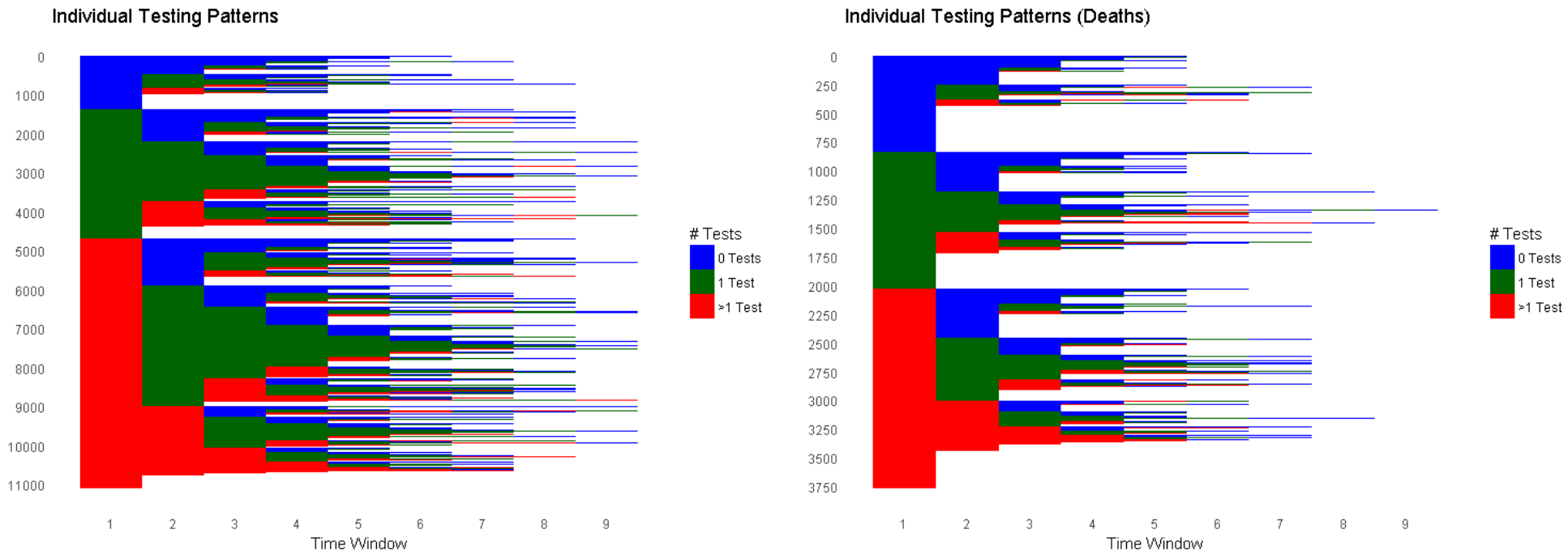


Figure 4.2: Individual Line Plots of Individual Testing Patterns During Follow Up (Overall and Patients who Died) including incomplete time windows. Patient numbers differ between the two graphs, i.e. patient 100 in the overall graph will not be patient 100 in the patients who died graph. Therefore, for both graphs, the patient number is arbitrary, and patients were sorted by their testing pattern before plotting.

4.4.2 Plasma Lipid Concentrations

Just under half of the patients had a baseline lipid measurement when they were not receiving statin therapy (5,037, 45.3%), with 3,360 (30.2%) patients having no pre-MI lipid measurement, and 2,713 (24.4%) with tests obtained only whilst already receiving statin medication. As only one baseline sample was required and extracted where it was available for a patient, the number of tests before the MI per individual was not investigated. Both the baseline (pre-MI statin-naïve) and post MI lipid plasma concentrations are summarised in Table 4.7.

There were no substantial differences identified in the post MI lipid plasma results between those with and without a baseline result (Table 4.7). However, the converse was not the case; some differences were present in the baseline lipid plasma results between those with and without post MI results. Whilst the numbers were small, those with a baseline result but no post MI result tended to have more favourable lipid concentrations (e.g. lower total and LDL cholesterol, and slightly higher HDL cholesterol). Comparisons of the baseline and post MI plasma lipid concentrations within the same patients suggested that the lipids in this population were subsequently reasonably well controlled following an MI, with the mean total cholesterol lower by over one mmol/l (5.4 vs 4.2mmol/l). As the HDL cholesterol stays broadly similar, changes in the Total:HDL cholesterol ratio are, therefore, likely the result of the reduction of the total cholesterol rather than an increase in HDL cholesterol. The mean values for LDL cholesterol and triglycerides also show some evidence of reduction post MI. However, given that there are often multiple tests per person in the post MI results compared to just one baseline result per person, such comparisons between the summary statistics should be interpreted with caution.

Table 4.7: Plasma Lipid Summary Statistics Comparing Pre and Post MI Concentrations

	Pre MI (w/ Post MI) Pts: 4,725 Tests: 4,725	Post MI (w/ Pre MI) Pts: 4,725 Tests: 28,887	Pre MI (no Post MI) Pts: 312 Tests: 312	Post MI (no Pre MI) Pts: 5,703 Tests: 35,053
Total Cholesterol				
Mean (SD)	5.4 (1.4)	4.2 (1.4)	4.9 (1.3)	4.1 (1.3)
Median	5.3	4.1	4.8	4.0
IQR	4.5 – 6.3	3.4 – 4.9	4.0 – 5.7	3.4 – 4.7
Range	0.0 – 14.6	0.0 – 16.0	0.0 – 9.3	0.0 – 18.6
Missing (%)	2 (<0.1%)	23 (0.1%)	0 (0.0%)	37 (0.1%)
Total:HDL Cholesterol				
Mean (SD)	4.8 (1.7)	3.8 (1.5)	4.1 (1.5)	3.8 (1.5)
Median	4.6	3.6	3.9	3.6
IQR	3.6 – 5.7	2.9 – 4.5	3.0 – 4.9	3.0 – 4.5
Range	0.0 – 15.7	0.0 – 25.0	0.0 – 11.6	0.0 – 30.8
Missing (%)	1,152 (24.4%)	9,314 (32.2%)	95 (30.4%)	11,305 (32.3%)
HDL Cholesterol				
Mean (SD)	1.2 (0.4)	1.1 (0.4)	1.3 (0.4)	1.1 (0.4)
Median	1.1	1.1	1.2	1.1
IQR	1.0 – 1.4	0.9 – 1.3	1.0 – 1.5	0.9 – 1.3
Range	0.0 – 3.9	0.0 – 4.4	0.0 – 2.9	0.0 – 4.4
Missing (%)	1,142 (24.2%)	9,246 (32.0%)	95 (30.4%)	11,246 (32.1%)
LDL Cholesterol				
Mean (SD)	1.7 (1.9)	1.5 (1.3)	1.3 (1.6)	1.4 (1.3)
Median	1.6	1.7	0.0	1.6
IQR	0.0 – 3.4	0.0 – 2.3	0.0 – 2.6	0.0 – 2.3
Range	0.0 – 9.3	0.0 – 10.2	0.0 – 5.9	0.0 – 8.1
Missing (%)	1,208 (25.6%)	9,414 (32.6%)	97 (31.1%)	11,398 (32.5%)
Triglycerides				
Mean (SD)	2.0 (1.5)	1.8 (1.5)	1.7 (1.1)	1.8 (1.5)
Median	1.7	1.5	1.4	1.5
IQR	1.2 – 2.4	1.1 – 2.2	1.0 – 2.2	1.1 – 2.2
Range	0.0 – 25.3	0.0 – 54.3	0.0 – 9.2	0.0 – 57.8
Missing (%)	199 (4.2%)	342 (1.2%)	16 (5.1%)	559 (1.6%)

All measurements are in mmol/l, except for the ratio where no units are given. Missing % calculated from the number of tests. Pts = Patients; SD = Standard Deviation; IQR = Interquartile Range. Pre MI (w/ Post MI) are pre MI tests obtained in patients with post MI test results available. Similarly, Post MI (with Pre MI) are post MI results in patients with a baseline lipid test. The remaining columns are pre MI test results where patients had no post MI tests and post MI results for patients with no pre MI baseline tests. Only one pre MI statin-naïve test per patient was included. Multiple post MI tests per patients were included.

4.4.3 Achieving Target Values Post Myocardial Infarction

Two lipid concentration targets were examined. The percentage reduction of non-HDL cholesterol, which is the focus of the NICE guidelines (NICE, 2014),

targets ≥40% reduction, whilst in the 2016 ESC guidelines, a target value of ≤1.8 mmol/l for LDL cholesterol is recommended (Piepoli *et al.*, 2016). For population comparisons, if a patient had achieved one of the targets in any of their time windows, then they were considered to have met that target, regardless of the consistency of this achievement.

There were 10,700 time windows (complete or partial years of follow up) where the percentage of non-HDL reduction could be calculated, occurring in 3,355 individuals. There were 3,835 (35.8%) time windows where the average reduction in non-HDL was ≥40% from their baseline, which occurred in 1,721 individuals. In 6,865 (64.2%) windows, this target was not met.

There were 30,319 windows where LDL values were reported in 9,504 individuals. Of these, 16,336 (53.9%) reported LDL to be ≤1.8mmol/l, with the remainder (13,983, 46.1%) not achieving this. The 16,336 windows where this target was achieved occurred in 7,215 individuals.

4.4.4 Demographics by Plasma Lipid Concentration Targets

The demographics summarising those who met and did not meet the lipid cholesterol targets are shown in Table 4.8. Males were more likely to achieve the percentage reduction in non-HDL and LDL targets, with a similar gender split observed in both groups (non-HDL: 64.6% vs 35.4%; LDL: 63.7% vs 36.3%). However, a slightly higher percentage of females did not achieve the percentage non-HDL reduction target compared to the LDL cholesterol target (43.9% vs 40.9%). Patients who achieved the target percentage non-HDL reduction were typically younger than those who did not (63.6 vs 67.1 years), but no age differences were identified when comparing patients who met and did not meet the LDL cholesterol recommendation. For the LDL target, a broadly U-shaped relationship appeared to exist between deprivation and those who met such targets with those in the most and least deprived areas more likely to meet it. This is not the case for the non-HDL target where no differences by deprivation quintile were evident. Whilst patients with a diagnosis of diabetes were more likely to achieve the LDL target (20.1% vs 13.8%), they were less likely to meet the percentage non-HDL reduction target (16.7% vs 22.8%). This pattern was also present amongst those patients with a prior MI.

Table 4.8: Demographics by Lipid Plasma Targets (from NICE and ESC Recommendations)

	Target: ≥40% Reduction non-HDL*		Target: LDL ≤1.8mmol/l	
	Met	Not Met	Met	Not Met
	1,721 (51.3%)	1,634 (48.7%)	7,215 (75.9%)	2,289 (24.1%)
Gender				
Male	1,111 (64.6%)	917 (56.1%)	4,593 (63.7%)	1,352 (59.1%)
Female	610 (35.4%)	717 (43.9%)	2,622 (36.3%)	937 (40.9%)
Age at MI (years)				
Mean (SD)	63.6 (12.5)	67.1 (13.6)	65.4 (13.2)	65.5 (14.0)
Median	62.7	67.8	65.4	65.2
IQR	53.9 – 73.8	56.4 – 77.5	55.2 – 75.9	54.8 – 76.6
Range	30.0 – 96.4	23.4 – 99.3	19.3 – 102.2	22.5 – 98.1
SIMD 2012 Quintile (NHS GGC)				
1 (Most)	398 (23.3%)	381 (23.5%)	1,824 (25.4%)	412 (18.2%)
2	389 (22.8%)	363 (22.4%)	1,620 (22.6%)	551 (24.3%)
3	349 (20.4%)	349 (21.5%)	1,438 (20.1%)	535 (23.6%)
4	284 (16.6%)	283 (17.5%)	1,140 (15.9%)	454 (20.0%)
5 (Least)	289 (16.9%)	244 (15.1%)	1,149 (16.0%)	317 (14.0%)
Missing	12	14	44	20
Diabetes at MI	288 (16.7%)	373 (22.8%)	1,447 (20.1%)	315 (13.8%)
Type 1	12 (0.7%)	29 (1.8%)	73 (1.0%)	32 (1.4%)
Type 2	276 (16.0%)	344 (21.1%)	1,374 (19.0%)	283 (12.4%)
Prior MI	34 (2.0%)	99 (6.1%)	473 (6.6%)	115 (5.0%)
1	28 (1.6%)	80 (4.9%)	361 (5.0%)	95 (4.2%)
>1	6 (0.3%)	19 (1.2%)	112 (1.6%)	20 (0.9%)

***refers to the percentage reduction in non-HDL cholesterol from the pre-MI baseline. Numbers are N (%) unless otherwise specified. Percentages are calculated within columns, except for the header where percentages are calculated from total with information available. SD = Standard Deviation; IQR = Interquartile Range. Patients only included if the necessary test results were available to determine whether the target was met.**

4.5 Estimated Statin Adherence

4.5.1 Achieving Targets for Adherence

Estimated adherence, using MPR, was not calculated for time windows where the death occurred before the end of the window, or the end of follow up (31st July 2017). In total, there were 55,866 time windows, with 11,098 (19.9%) representing an incomplete year (i.e. there were 44,768 complete years of follow up).

Two adherence targets were used, 50% and 80%. There were 7,127 (7127/44768, 15.9%) time windows where a patient's adherence was <50%, occurring in 2,798 patients. For 80%, there were 13,546 (13546/44768, 30.3%) time windows where the threshold was not achieved, which occurred in 5,532 patients.

For demographic comparisons, a patient's adherence was (mean-)averaged over their available time windows with the average value used to classify a patient as adherent or non-adherent depending on the threshold used (i.e. not everyone who ever had an adherence of lower than 50% will fall under the <50% category). Using these averages, the majority of patients had an average MPR greater than the conventional adherence threshold of 80% (68.9%), with a higher percentage achieving the lower threshold of 50% (84.5%).

4.5.2 Other Lipid Lowering Medications in Non-Adherent Patients

There were 434 (4.3%) patients who had an average estimated statin adherence equal to zero (i.e. were never dispensed any statins post MI). Of these, 37 patients were prescribed different LLMs at any point during their follow up. The most common of these was Ezetimibe, where 349 prescriptions were dispensed to 23 patients. A further 331 fibrate prescriptions, specifically for Bezafibrate and Fenofibrate, were dispensed to 14 patients, with two patients prescribed omega 3 ethyl esters. Very few patients were dispensed more than one LLM that was not a statin during follow up.

4.5.3 Demographics by Statin Adherence

For both adherence thresholds (MPR 50% and 80%), males were more likely to be classified as adherent (50%: 63.1% vs 53.6%; 80%: 63.4% vs 57.7%). In the case of the lower threshold, those with an average adherence <50% were several years older than those with average adherence ≥50% (68.1 vs 65.3 years). This was less evident when the threshold of 80% was used, with only a difference of a year (66.4 vs 65.4 years). Those in the more deprived areas were more likely to achieve the 50% threshold for adherence (24.3% vs 22.4%), but this was less evident for the 80% threshold, where the greatest difference was in the second most deprived quintile (23.6% vs 21.6%). Patients who were diagnosed with Type 2 diabetes before their MI were more likely to meet the adherence thresholds

(50%: 17.2% vs 15.1%; 80%: 17.2% vs 16.1%), particularly in the case of the 50% threshold. In contrast, patients with Type 1 diabetes were more likely to have lower adherence (50%: 0.9% vs 1.7%; 80%: 0.9% vs 1.4%). The proportion of patients with prior MIs was similar between those with average adherence above and below both thresholds (50%: 6.2% vs 5.5%; 80%: 6.1% vs 6.1%).

Table 4.9: Demographics by Average Statin Adherence Thresholds (50% and 80%)

	MPR ≥50% 8,522 (84.5%)	MPR <50% 1,560 (15.5%)	MPR ≥80% 6,942 (68.9%)	MPR <80% 3,140 (31.1%)
Gender				
Male	5,375 (63.1%)	836 (53.6%)	4,400 (63.4%)	1,811 (57.7%)
Female	3,147 (36.9%)	724 (46.4%)	2,542 (36.6%)	1,329 (42.3%)
Age at MI (years)				
Mean (SD)	65.3 (13.2)	68.1 (15.9)	65.4 (13.0)	66.4 (15.2)
Median	65.2	69.9	65.4	67.1
IQR	55.1 – 75.8	55.9 – 81.1	55.4 – 75.6	54.5 – 78.6
Range	21.9 – 101.0	19.3 – 102.2	21.9 – 101.0	19.3 – 102.2
SIMD 2012 Quintile (NHS GGC)				
1 (Most)	2,058 (24.3%)	347 (22.4%)	1,652 (24.0%)	753 (24.2%)
2	1,979 (23.4%)	318 (20.5%)	1,625 (23.6%)	672 (21.6%)
3	1,750 (20.7%)	338 (21.8%)	1,389 (20.1%)	699 (22.4%)
4	1,395 (16.5%)	289 (18.7%)	1,153 (16.7%)	531 (17.0%)
5 (Least)	1,279 (15.1%)	256 (16.5%)	1,075 (15.6%)	460 (14.8%)
Missing	61	12	48	25
Diabetes at MI	1,547 (18.2%)	261 (16.7%)	1,258 (18.1%)	550 (17.5%)
Type 1	80 (0.9%)	26 (1.7%)	63 (0.9%)	43 (1.4%)
Type 2	1,467 (17.2%)	235 (15.1%)	1,195 (17.2%)	507 (16.1%)
Prior MI	532 (6.2%)	86 (5.5%)	426 (6.1%)	192 (6.1%)
1	415 (4.9%)	69 (4.4%)	331 (4.7%)	153 (4.9%)
>1	117 (1.4%)	17 (1.1%)	95 (1.4%)	39 (1.2%)

Numbers are N (%) unless otherwise specified. Percentages are calculated within columns, except for the header where percentages are calculated from total with information available. SD = Standard Deviation; IQR = Interquartile Range. Patients were only included if at least one complete year of follow up was available and then classified by average statin adherence across the course of their complete years of follow up available.

However, some of these differences are likely due to confounding, for example, females were typically older, and therefore any associations with adherence are unlikely to be straightforward. For this reason, models adjusting for confounders, which are covered in later chapters, are perhaps better suited to understanding the relationships between a patient's demographics and their statin adherence.

4.5.4 Statin Adherence and Lipid Plasma Concentrations

Table 4.10 shows the proportion of patients meeting lipid plasma concentration targets recommended by guidelines by their average statin adherence during their follow up. As anticipated, given the mechanism of action of statin medication, those with higher average statin adherence were more likely to achieve the recommended cholesterol targets from the guidelines. This was most evident when looking at the percentage reduction of non-HDL cholesterol, where the proportion of those who achieved the target was more than double in those with an MPR ≥50% than those who did not (58.5% vs 22.9%). A similar pattern was observed when the 80% threshold for adherence was used (61.1% vs 34.2%).

Table 4.10: Average Statin Adherence and Lipid Plasma Concentrations (Adherence: 50% and 80%; Lipids: NICE and ESC Recommendations)

	MPR ≥50%	MPR <50%	MPR ≥80%	MPR <80%
≥40% non-HDL*	1,546 (58.5%)	131 (22.9%)	1,312 (61.1%)	365 (34.2%)
<40% non-HDL*	1,095 (41.5%)	442 (77.1%)	835 (38.9%)	702 (65.8%)
Total	2,641	573	2,147	1,067
LDL ≤1.8 mmol/l	6,097 (78.0%)	816 (64.2%)	5,053 (78.7%)	1,860 (69.7%)
LDL >1.8 mmol/l	1,719 (22.0%)	455 (35.8%)	1,367 (21.3%)	807 (30.3%)
Total	7,816	1,271	6,420	2,667

***refers to the percentage reduction in non-HDL cholesterol from the pre-MI baseline. Numbers are N (%). Patients were only included if at least one complete year of follow up was available and then classified by average statin adherence across the course of their complete years of follow up available. Patients were also only included if the necessary test results were available to determine whether the target was met. The number of patients in each group for each target and with adherence data are included on the total line of each set of information, e.g. 2,641 patients had calculable non-HDL cholesterol and MPR ≥50%.**

4.5.5 Statin Adherence and Lipid Testing

To examine the association between a patient's adherence and the frequency of lipid testing in a given year, the two were cross-tabulated (Table 4.11 and Table 4.12). Due to the different testing pattern in the first year of follow up (Chapter 4.4.1), only years two to eight of follow up were tabulated. This highlighted that there may be some association between the two, with those with higher adherence more likely to have the guideline-recommended one lipid test, and those with lower adherence more likely to have had no lipid tests that year. There were no differences between over-testing and a patient's adherence when the 80% threshold was used, though those with higher adherence were slightly more likely to have more than one test when the 50% cut point was used.

Nonetheless, despite some differences observed, a patient's adherence does not appear to be substantially explained by the frequency of their lipid testing.

Table 4.11: Adherence and Number of Lipid Tests by Year of Follow Up (50% MPR)

	MPR≥50%			MPR<50%		
Year of FU	N	row %	col %	N	row %	col %
Year 2						
0 Tests	1,515	75.3%	19.2%	498	24.7%	35.3%
1 Test	4,260	88.8%	53.9%	540	11.3%	38.3%
>1 Test	2,130	85.1%	26.9%	372	14.9%	26.4%
Year 3						
0 Tests	1,480	72.7%	20.5%	555	27.3%	38.2%
1 Test	4,057	86.8%	56.2%	615	13.2%	42.4%
>1 Test	1,684	85.7%	23.3%	281	14.3%	19.4%
Year 4						
0 Tests	1,158	71.3%	21.5%	466	28.7%	40.0%
1 Test	3,071	86.1%	57.0%	497	13.9%	42.7%
>1 Test	1,163	85.2%	21.6%	202	14.8%	17.3%
Year 5						
0 Tests	914	73.5%	23.2%	330	26.5%	40.4%
1 Test	2,199	86.6%	55.8%	341	13.4%	41.8%
>1 Test	826	85.1%	21.0%	145	14.9%	17.8%
Year 6						
0 Tests	633	73.4%	24.5%	229	26.6%	41.3%
1 Test	1,427	85.7%	55.3%	239	14.3%	43.1%
>1 Test	521	85.7%	20.2%	87	14.3%	15.7%
Year 7						
0 Tests	340	71.1%	24.4%	138	28.9%	47.9%
1 Test	791	86.6%	56.7%	122	13.4%	42.4%
>1 Test	265	90.4%	19.0%	28	9.6%	9.7%
Year 8						
0 Tests	115	70.6%	24.5%	48	29.4%	49.0%
1 Test	268	87.3%	57.1%	39	12.7%	39.8%
>1 Test	86	88.7%	18.3%	11	11.3%	11.2%

FU = Follow Up. Complete years of follow up included only. Patients classified by adherence in the year of follow up, so individual patients may move between adherence categories and testing frequencies between follow up years. Patients were defined as having a test if at least one component of the lipid panel was reported. Percentages are calculated by row and column separately but within each year of follow up.

Table 4.12: Adherence and Number of Lipid Tests by Year of Follow Up (80% MPR)

	MPR≥80%			MPR<80%		
Year of FU	N	row %	col %	N	row %	col %
Year 2						
0 Tests	1,260	62.6%	19.0%	753	37.4%	28.1%
1 Test	3,631	75.6%	54.7%	1,169	24.4%	43.7%
>1 Test	1,749	69.9%	26.3%	753	30.1%	28.1%
Year 3						
0 Tests	1,220	60.0%	19.8%	815	40.0%	32.6%
1 Test	3,564	76.3%	57.8%	1,108	23.7%	44.3%
>1 Test	1,386	70.5%	22.5%	579	29.5%	23.1%
Year 4						
0 Tests	972	59.9%	21.1%	652	40.1%	33.5%
1 Test	2,669	74.8%	57.9%	899	25.2%	46.1%
>1 Test	967	70.8%	21.0%	398	29.2%	20.4%
Year 5						
0 Tests	783	62.9%	23.3%	461	37.1%	33.2%
1 Test	1,881	74.1%	55.9%	659	25.9%	47.5%
>1 Test	703	72.4%	20.9%	268	27.6%	19.3%
Year 6						
0 Tests	529	61.4%	24.0%	333	38.6%	35.8%
1 Test	1,239	74.4%	56.2%	427	25.6%	45.9%
>1 Test	437	71.9%	19.8%	171	28.1%	18.4%
Year 7						
0 Tests	290	60.7%	24.4%	188	39.3%	38.0%
1 Test	681	74.6%	57.3%	232	25.4%	46.9%
>1 Test	218	74.4%	18.3%	75	25.6%	15.2%
Year 8						
0 Tests	98	60.1%	23.8%	65	39.9%	41.7%
1 Test	240	78.2%	58.4%	67	21.8%	42.9%
>1 Test	73	75.3%	17.8%	24	24.7%	15.4%

FU = Follow Up. Complete years of follow up included only. Patients classified by adherence in the year of follow up, so individual patients may move between adherence categories and testing frequencies between follow up years. Patients were defined as having a test if at least one component of the lipid panel was reported. Percentages are calculated by row and column separately but within each year of follow up.

4.6 Diagnosis of Diabetes

4.6.1 Diagnosis Prior to Myocardial Infarction

There were 2,099 (18.9%) patients diagnosed with diabetes before their baseline MI admission, the majority of whom had Type 2 Diabetes (1979/2099, 94.3%), as

shown in Table 4.13. The average duration of disease was substantially different between the two types of diabetes. This would be expected given the average age of patients at the time of diagnosis for Type 1 diabetes compared to Type 2, and the average age of patients when they experience an MI.

Table 4.13: Duration of Diabetes at Baseline MI Admission

	Type 1 N = 120	Type 2 N = 1,979	Total N = 2,099
Duration of Diabetes (years)			
Mean (SD)	30.1 (16.6)	10.5 (7.8)	11.6 (9.7)
Median	30.4	9.2	9.7
IQR	17.2 – 42.2	4.4 – 14.8	4.7 – 15.7
Range	0.3 – 75.1	0.0 – 45.7	0.0 – 75.1

SD = Standard Deviation; IQR = Interquartile Range. Duration of diabetes calculated using the date of diagnosis and date of admission with baseline MI.

4.6.2 Diagnosis Post Myocardial Infarction

Table 4.14 contains the number of patients subsequently diagnosed with diabetes in follow up (i.e. the day of admission with MI onwards). There were 733 further patients diagnosed with diabetes, the majority with Type 2 (727/733, 99.2%), and within the first year of follow up (236/733, 32.2%). However, some of the decreases in the number of those being newly diagnosed per year are as a result of the few patients remaining in follow up in each year. Two patients had an unknown date of diagnosis, so it could not be ascertained whether this occurred prior to or after the start of follow up and are therefore not included in these summaries.

Table 4.14: Diagnosis of Diabetes in Follow Up by Year of Follow Up

	Type 1 N = 6	Type 2 N = 727	Total N = 733
Year of FU			
1	3	233	236
2	0	133	133
3	0	122	122
4	0	97	97
5	2	66	68
6	1	42	43
7	0	24	24
8	0	10	10
9	0	0	0

Numbers are N. FU = Follow Up. Year of follow up where diagnosis occurred was obtained from dates of diagnosis and start and end dates of years of follow up.

4.6.3 Demographics by Diabetes Diagnosis at MI

Table 4.15: Demographics by Diabetes Diagnosis at Baseline MI Admission

	Type 1 120 (1.1%)	Type 2 1,979 (17.8%)	Total 2,099 (18.9%)	No Diabetes 9,011 (81.1%)
Gender				
Male	63 (52.5%)	1,138 (57.5%)	1,201 (57.2%)	5,531 (61.4%)
Female	57 (47.5%)	841 (42.5%)	898 (42.8%)	3,480 (38.6%)
Age at MI (years)				
Mean (SD)	59.7 (13.9)	70.0 (11.9)	69.4 (12.2)	66.3 (14.2)
Median	60.6	71.1	70.7	66.4
IQR	49.6 – 69.9	62.4 – 78.6	61.7 – 78.3	55.2 – 77.7
Range	28.1 – 90.8	30.0 – 101.0	28.1 – 101.0	19.3 – 102.6
SIMD 2012 Quintile (NHS GGC)				
1 (Most)	26 (21.7%)	497 (25.3%)	523 (25.1%)	2,127 (23.8%)
2	26 (21.7%)	450 (22.9%)	476 (22.9%)	2,052 (22.9%)
3	28 (23.3%)	441 (22.4%)	469 (22.5%)	1,825 (20.4%)
4	27 (22.5%)	336 (17.1%)	363 (17.4%)	1,514 (16.9%)
5 (Least)	13 (10.8%)	237 (12.1%)	250 (12.0%)	1,432 (16.0%)
Missing	0	18	18	61
Prior MI	14 (11.7%)	205 (10.4%)	219 (10.4%)	495 (5.5%)
1	12 (10.0%)	161 (8.1%)	173 (8.2%)	385 (4.3%)
>1	2 (1.7%)	44 (2.2%)	46 (2.2%)	110 (1.2%)

Numbers are N (%). Percentages are calculated within columns, except for the header where percentages are calculated from total with information available. SD = Standard Deviation; IQR = Interquartile Range. Patients diagnosed during follow up are included in the no diabetes group, as demographics are presented by diabetes status at baseline MI admission.

The demographics for those diagnosed with Type 1, Type 2, either and no diabetes before their baseline MI admission are summarised in Table 4.15. There was a higher proportion of females with a diagnosis of diabetes, especially Type 1 diabetes (47.5%), than in the population without diabetes (42.8% vs 38.6%, respectively). Patients diagnosed with Type 1 diabetes were also, on average, around five years younger at their baseline MI admission compared to those without (60.6 vs 66.4 years), whereas those with Type 2 diabetes were on average around four years older at their baseline MI (71.1 years). When looking at differences in deprivation, in the case of patients with Type 1 diabetes, the relationship did not appear to be straightforward. However, in the case of patients with Type 2 diabetes, there was a slightly higher proportion of patients diagnosed who resided in an area in the most deprived quintile compared with those with no diagnosis of diabetes before their baseline MI admission (25.3% vs

23.8%). Patients with diabetes were also more likely to have experienced an MI prior to the one occurring at the start of follow up (10.4% vs 5.5%).

4.7 Discussion

This chapter sought to describe the post MI population of the NHS GGC region between 2009 and 2017 and specifically summarises their demographics, lipid results, testing patterns, and statin adherence following an MI.

4.7.1 Overall Post Myocardial Infarction Population

The post MI population was not a representative sample of the NHS GGC population and, instead, was characterised by the presence of the known risk factors for CVD. Indeed, patients within this cohort were more likely to be male, have a diagnosis of diabetes, be over the age of 65, and live in more deprived areas than in the NHS GGC overall population.

However, this population is broadly comparable to other secondary prevention cohorts. These included one high-risk cohort from an insurance database in Georgia, USA who initiated statins in 2011 (n=1066) (Vupputuri *et al.*, 2016), and three cohorts derived from UK CPRD data: a secondary prevention cohort from 2006 and 2012 with at least two statin prescriptions before their event (n=24,093) (Danese *et al.*, 2017), a cohort with documented CVD and newly initiated statins or ezetimibe between 2010 and 2013 (n=16,701) (Khunti *et al.*, 2018), and a high-risk cohort formed of those newly diagnosed with documented disease or diabetes from 2008 to 2011 (n=131,603) (Nordstrom *et al.*, 2015).

A higher proportion of males is reported in established CVD populations (Danese *et al.*, 2017; Khunti *et al.*, 2018), although is less evident in broader high-risk populations (Nordstrom *et al.*, 2015; Vupputuri *et al.*, 2016). The age profiles are also similar with average ages consistently over 65 years (Nordstrom *et al.*, 2015; Danese *et al.*, 2017; Khunti *et al.*, 2018). There is some discrepancy in the prevalence of diabetes reported within secondary prevention cohorts, with percentages ranging from 6% to 36% (Nordstrom *et al.*, 2015; Danese *et al.*, 2017; Khunti *et al.*, 2018). This is possibly a consequence of the different inclusion criteria for each of the cohorts. In the NHS GGC population, the

proportion of patients diagnosed with diabetes (including those diagnosed during follow up) is closer to the upper bound of this range and is closest to the estimated prevalence in patients who have been hospitalised with an acute cardiovascular event (Danese *et al.*, 2017).

4.7.2 Frequency and Patterns of Lipid Testing

The frequency of testing within this population showed that, excluding the first year post MI, just over half of the cohort received one test in any given year, in line with current (and previous) Scottish guidelines (SIGN, 2007, 2017). However, the consistency of the annual test (an overall testing rate over follow up close to one) was only met by approximately a third of patients, with the highest proportion of patients reporting a much higher, rather than lower, rate. Nonetheless, when each year was considered individually, under-testing became more common than over-testing as follow up duration increased. In the first year of follow up, over-testing was the most common and therefore may explain the discrepancy between the two measures. Long-term trends in testing have also been shown in an analysis of 1.6 million patients at high risk of CVD on stable statin prescriptions, where 33% received annual testing (33%) between 2008 and 2014 (Hajati *et al.*, 2018). This was consistent with the overall rates in the post MI population but lower than the proportions observed in their individual years of follow up. Furthermore, in Hajati et al's (2018) analysis, under-testing occurred twice as often as over-testing; a pattern that was not observed in this post MI cohort, possibly due to more frequent monitoring in those with established CVD.

The higher rate of over-testing in the first year could be driven by two factors: medication initiation or alteration and further cardiovascular events. Following the MI at the start of their follow up, many patients are likely to either initiate LLM for the first time or increase its intensity (Danese *et al.*, 2017). As such, extra testing may occur to assess whether targets have been achieved, as suggested by some major guidelines (NICE, 2014; Stone *et al.*, 2014; Catapano *et al.*, 2016). Secondly, the greatest risk of further cardiovascular events is in the first year after an initial event (Khunti *et al.*, 2018). Such events occurred in this cohort and could have led to extra testing through hospital admissions or

subsequent interactions with primary care as medications are further adjusted or intensified (Danese *et al.*, 2017).

4.7.3 Achievement of Guideline-Recommended Lipid Targets

The average post MI plasma lipid results showed a more favourable lipid profile when contrasted with statin-naïve results before the start of follow up. Specifically, reductions were observed in total cholesterol, as a consequence of decreases in LDL cholesterol, which in turn result in an improved Total:HDL cholesterol ratio. This pattern is similar to UK derived secondary prevention populations, although average baseline LDL levels in this population were considerably lower than values reported in secondary prevention populations in CPRD data (Danese *et al.*, 2017; Khunti *et al.*, 2018).

When translated into the achievement of guideline-recommended targets, 76% achieved the 2016 ESC target of LDL cholesterol ≤1.8mmol/l (Piepoli *et al.*, 2016) during their follow up. This is higher than has been observed in other secondary prevention cohorts, with estimates in various high-risk populations ranging between 23% and 42% (Nordstrom *et al.*, 2015; Danese *et al.*, 2017). This difference may, in part, be accounted for by population differences. However, many analyses have focussed on the achievement of targets at a specific time point during follow up, e.g. one (Danese *et al.*, 2017), or two years (Nordstrom *et al.*, 2015). In contrast, patients in this population were classified as achieving the target if they achieved it in any year of follow up. Around 50% achieved the NICE recommended ≥40% reduction in non-HDL cholesterol (NICE, 2014). This is also higher than percentages achieving a 30% reduction in LDL cholesterol in a high-risk USA population (Vupputuri *et al.*, 2016).

Differences observed in the baseline demographics by LDL target achievement were not all reflected for the non-HDL cholesterol target. This may be due to the need for a suitable baseline to calculate the percentage reduction, causing a selection bias. Many patients with test results available before their MI are likely to have already been identified as being at high risk, either by default or using the ASSIGN risk prediction tool. Consequently, they should have been prescribed statins, making it less likely that a suitable (statin-naïve) baseline result was available. For example, those with prior MIs should have, and those with a

diagnosis of diabetes are likely to have, been on statins before their start of follow up (SIGN, 2007, 2017). Patients may also have no prior test results available because no CVD risk assessment had been performed before their MI, whilst others may have no post MI results due to a pre-existing concurrent illness, resulting in both low pre-MI lipids and an absence of lipid tests following their baseline MI (also preventing the calculation of a percentage reduction). As a result, patients included in the non-HDL cholesterol comparisons may not be representative of the overall post MI population. However, the LDL cholesterol target used in this analysis does not encounter this, as this is based on its absolute value (Piepoli *et al.*, 2016).

Despite this, some differences noted in baseline demographics by the achievement of the LDL target may have arisen due to confounding. For example, those residing in the least deprived areas were more likely to meet the non-HDL target, whereas those residing in the most and least deprived areas were more likely to meet the LDL target. This U-shaped relationship for LDL may have arisen due to higher rates of multimorbidity in more deprived areas, where these conditions may also be associated with the greater achievement of lipid levels and a decreased likelihood of statin naivety before their baseline MI (thus excluding them from non-HDL comparisons). For example, in this cohort, patients with a diagnosis of diabetes were more likely to reside in the most deprived areas (Chapter 4.7.5), as well as being more likely to achieve recommended LDL levels. The observations in this cohort that those with diabetes and those with prior MIs were more likely to achieve LDL ≤1.8mmol/l are also consistent with findings from another secondary prevention cohort within the UK (Danese *et al.*, 2017).

Furthermore, females were less likely to achieve the LDL cholesterol target; an observation that has been observed within the literature (Reiner *et al.*, 2016; Leskelä *et al.*, 2020), despite evidence to suggest that the absolute reduction in cardiovascular events from statins is equivalent in both sexes (Gutierrez *et al.*, 2012). Confounding may offer one potential explanation. For example, women who experience MIs are typically older than men (Smolina *et al.*, 2015), and adherence to statin medication has also been shown to decrease with age over 70 years (Hope *et al.*, 2019), which would influence a patient's LDL levels.

However, whilst adherence did decrease with age in this population, this did not translate into age differences between those who achieved and did not achieve the LDL target. Nonetheless, the management of females who experience MIs may explain the differences observed in both lipid target achievement and adherence. Several studies have reported that women are more likely to be undertreated following an MI (Smolina *et al.*, 2015; Eindhoven *et al.*, 2018), with the largest differences for statin use in NSTEMIs (Eindhoven *et al.*, 2018). Consequently, undertreated patients could be classified as non-adherers when adherence is captured through dispensing prescription data, as well as being less likely to achieve lipid targets.

4.7.4 Medication Adherence

A third of the NHS GGC post MI population had an average statin adherence <80% which is slightly higher than estimates from other similar cohorts which utilised MPR or PDC as estimation methods (Nordstrom *et al.*, 2015; Danese *et al.*, 2017; Khunti *et al.*, 2018). However, this may be due to length of follow up where calculations are often confined to the first few years of follow up, where adherence is likely to be the highest (Khunti *et al.*, 2018). Although there is some evidence to suggest that the greatest decrease occurs in the first year of follow up (Chen *et al.*, 2019), some decline is still present beyond this, and likely to affect a patient's long-term average adherence (Naderi, Bestwick and Wald, 2012; Chowdhury *et al.*, 2013; Khunti *et al.*, 2018; Chen *et al.*, 2019). In contrast, the proportion of patients with average adherence <50% (16%) was lower than observed in another high-risk cohort, where 36% reported a PDC <50% in the first year of statin initiation in Georgia, USA (Vupputuri *et al.*, 2016). This could be due to differences within the populations, where those who have experienced an event are more likely to remain adherent to their medication (Naderi, Bestwick and Wald, 2012) than the overall high-risk population. Prescription charges could also offer some explanation for this, with their costs associated with lower adherence to medications (Leslie, McCowan and Pell, 2018). Therefore, the lack of prescription charge in Scotland since 2011 (NHS Inform Scotland, 2020) could have resulted in fewer patients with adherence <50%.

Despite this, in line with adherence to cardiovascular medications research, females and older-aged patients were less likely to be adherent (Hope *et al.*, 2019), whilst those with a prior MI (Vupputuri *et al.*, 2016; Danese *et al.*, 2017) were more likely. This would suggest that the factors associated with adherence to statins in both primary and secondary prevention are universal. In this population, those with a diagnosis of diabetes were also more likely to be adherent. This association has been identified through a meta-analysis in primary prevention (Hope *et al.*, 2019), but within secondary prevention has either been shown to have no association with adherence (Danese *et al.*, 2017) or patients with diabetes are just as likely to have low adherence as they are high, but not intermediate (50-80%) (Vupputuri *et al.*, 2016).

Those with higher average adherence were also more likely to have achieved guideline-recommended lipid targets. This is consistent with the rationale of prescribing and mechanism of action LLMs, and findings in other high-risk cohorts (Nordstrom *et al.*, 2015; Vupputuri *et al.*, 2016). However, not all patients who had high adherence achieved the lipid targets, and therefore meeting these targets cannot be due to adherence alone. Indeed, treatment intensity may also be a factor. In one high-risk cohort, patients receiving a high-intensity statin were more likely to achieve a lower LDL cholesterol result (Nordstrom *et al.*, 2015). This concurs with findings in a different high-risk cohort, but which also highlighted that adherence decreases as intensity increases (Vupputuri *et al.*, 2016), suggesting that adherence and intensity should be balanced carefully. Regardless, in both analyses, not all patients receiving high-intensity LLM met cholesterol targets (Nordstrom *et al.*, 2015; Vupputuri *et al.*, 2016), and therefore, achieving these involves factors beyond medication.

The frequency of testing was also shown to have some influence and may impact the association described above. Patients with higher adherence were more likely to have an annual lipid test and those with lower adherence less likely. Therefore, for patients with lower adherence, it was less likely that it could be ascertained whether a lipid target had been met. Given LLMs lower cholesterol, this may mean that the number of patients with lower adherence who did not achieve the target is far higher, and the association is underestimated.

4.7.5 Diabetes Mellitus

Around one in four patients in this cohort were diagnosed with diabetes before their baseline MI or during follow up. This percentage diagnosed before their start of follow up is similar to findings for the whole of Scotland, which reported that 19% of patients who experienced an MI (including fatal MIs), angina, or revascularisation between 2006 and 2015 had a diagnosis of Type 2 diabetes at the time (Read *et al.*, 2019). For those diagnosed during follow up, this was most likely to occur within the first-year post MI and is likely to have been detected during or shortly after their admission with their baseline MI (and was therefore likely present but undocumented at the time of the event).

In line with a secondary prevention cohort, there was a higher proportion of females with a diagnosis than without, although this difference was more pronounced in this post MI population (Danese *et al.*, 2017). Surprisingly, and in contrast to Danese et al's (2017) established CVD population, which found minimal differences in age at the time of their index event, patients diagnosed with diabetes were also more likely to be older than those without at the time of their baseline MI. For those diagnosed with Type 2 diabetes (the majority of diagnoses in this cohort), this may in part be explained by the increased risk of diabetes with age (Read *et al.*, 2016), and hence patients would need to be older to have diabetes and an MI concurrently. Nonetheless, there is also a larger variation in the age range of those experiencing MIs without a diagnosis of diabetes, suggesting that there may be subpopulations within this group with different risk factor profiles. Consistent with analysis of the diabetes register (SCI Diabetes) for Scotland in 2016, patients with a diagnosis of Type 2 diabetes were also more likely to live in deprived areas (Whittaker *et al.*, 2020). Furthermore, a separate analysis of SCI Diabetes demonstrated an increased risk of acute MI as a result of a significant interaction between deprivation and diabetes (Read *et al.*, 2019).

The majority of patients with a diagnosis of diabetes in this population were diagnosed with Type 2, with approximately 5% diagnosed with Type 1, which is broadly similar to an established CVD cohort in CPRD (Danese *et al.*, 2017). As a result of this, many of the differences observed in the demographics are the result of differences between those with a Type 2 diagnosis and those without a

diagnosis of diabetes, and consequently, the Type 1 population could be overlooked. Those diagnosed with Type 1 diabetes were considerably younger at the time of their baseline MI: around ten years younger than those with a Type 2 diagnosis, and six years younger than those without. This, in combination with the greater duration of disease at the time of their MI and the higher rates of prior MIs reported, demonstrates the increased cardiovascular risk experienced by this population, which was also highlighted by findings from the Swedish National Diabetes Register (Araz Rawshani *et al.*, 2018). Furthermore, in this cohort, those with a diagnosis of Type 1 diabetes were also less likely to achieve average adherence thresholds of 50% and 80%, and less likely to achieve guideline-recommended lipid targets, suggesting that risk management of this population could be improved.

4.7.6 Strengths and Limitations

One of the strengths of this analysis is the size and coverage of this cohort. This population includes over 11,000 survivors of MIs and follows them up for at least three years (unless death occurred), with an average follow up of four and a half years. This cohort was derived from hospital admission records for all patients across NHS GGC who had ever had a lipid test or prescription for LLM, making it representative of the post MI population in this area. Furthermore, the length of follow up is also a strength. Many similarly derived cohorts from national datasets often focus on shorter average periods of follow up (Nordstrom *et al.*, 2015; Vupputuri *et al.*, 2016; Danese *et al.*, 2017), whereas this data allows longer-term trends in adherence, lipid levels, testing patterns, and further cardiovascular events to be examined in later chapters.

Nonetheless, this data is observational and thus open to confounding. This is the main limitation of these descriptive summaries as no adjustment for these has been performed. For example, those with a diagnosis of diabetes and who had experienced a prior MI were each more likely to be adherent, but those with a prior MI were also more likely to have a diagnosis of diabetes. Such interactions are likely to be common, leading to complex and multifactorial relationships with adherence or achievement of lipid targets. Likewise, the direction of an association may be difficult to determine, such as with the association between the frequency of tests and adherence. In a review of systematic reviews for

factors associated with adherence to CVD medications, lower adherence has been associated with a reduced likelihood of attending appointments (Leslie, McCowan and Pell, 2018), thereby reducing the potential for testing. However, as annual reviews should include discussions to address a patient's adherence (SIGN, 2007, 2017), those who do not attend appointments are also less likely to become adherent in the future (Leslie, McCowan and Pell, 2018). Finally, due to the observational nature of this cohort, no conclusions regarding causality can be made.

4.7.7 Conclusions

This chapter summarised the demographics, lipid tests, testing frequency, and statin adherence of over 11,000 patients within the NHS GGC post MI population and found them to be broadly in line with similar secondary prevention cohorts. However, due to the observational nature of the data and the consequential likelihood of confounding, further analysis is needed to fully understand the associations between them. This will be addressed in later chapters.

Chapter 5 Consequences of Non-Adherence and Non-Target Lipids

5.1 Introduction

For those with established CVD, the rationale for managing their risk has been outlined in Chapter 1.2.2.1. In one meta-analysis of RCTs, for patients not treated with statins, the rate of further events in those with a previous cardiovascular event has been estimated to be more than double the rate in patients with no previous events (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010). For all-cause mortality, the rate per year has been estimated to be six times greater in survivors of MIs than those without CVD (World Health Organization, 2017a). Furthermore, the associated healthcare costs of CVD to the UK economy are nearly £9bn per year (British Heart Foundation, 2018).

However, managing and monitoring risk factors within this population can be a highly effective strategy to reduce the burdens of CVD. One of the cornerstones of this risk-management is the prescribing of LLMs to these patients, with statins as the primary therapy (as described in Chapter 1.3.1). The evidence for this is strong, with the Cholesterol Treatment Triallists' Collaboration meta-analysis of 21 RCTs comprising 129,526 participants reporting a 21% reduction in the risk of further events per 1mmol/l reduction in LDL cholesterol (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010). Furthermore, no adverse safety limit for cholesterol lowering has been identified, and absolute risk reductions are greater for those at the highest risk (Cholesterol Treatment Trialists' (CTT) Collaborators, 2012).

With this strong evidence base for therapies to reduce a patient's cardiovascular risk, a patient's adherence to these medications and their corresponding lipid levels become key components of a patient review (SIGN, 2017). Higher patient adherence and the achievement of target lipid levels (set by clinical guidelines) have both been associated with a reduced likelihood of further events and mortality (Chapters 1.3 and 1.4.3) with evidence for latter the result of the supporting evidence for the use of statins. The association between adherence and the risk of cardiovascular events has been shown in both primary and

secondary prevention (Xu *et al.*, 2017; Khunti *et al.*, 2018). For example using CPRD data, in patients with CVD, a 10% increase in statin adherence was associated with a 5% reduction in further cardiovascular events (Khunti *et al.*, 2018). However, as electronic prescribing records in CPRD are prescribing data (Herrett *et al.*, 2015), the extent of this association may have been underestimated as not all included prescriptions will have been dispensed.

Adherence to medication in secondary prevention settings is generally higher than observed in primary prevention (Naderi, Bestwick and Wald, 2012), and several analyses of secondary prevention cohorts have found levels of adherence to be reasonable; with between 67.9% and 74.8% of participants with statin adherence ≥80% in the 12 months following an event (Nordstrom *et al.*, 2015; Danese *et al.*, 2017). However, this does not fully translate into achievements of target lipid levels, with less than half of patients achieving them in the year following an event (Danese *et al.*, 2017). Therefore, there superficially appears to be a disconnect between adherence to LLMs and achievement of lipid targets stated in guidelines, as discussed in Chapter 1.4.3. These issues require further study.

Consequently, this chapter aims to investigate the association between statin adherence, and achievement of lipid-lowering targets, with cardiovascular outcomes separately. The two cardiovascular events of interest are further hospital admissions for MIs and death (all-cause and circulatory), with the relationships examined in separate analyses.

5.2 Statistical Methods

As a patient may experience multiple further hospital admissions for MIs in follow up, but death occurs only once, different statistical methods were used for each of these outcomes. Descriptive analyses and logistic regression were used to understand the associations between average adherence and the achievement of lipid targets, and both separately with further MIs. For mortality outcomes, descriptive analyses, Kaplan-Meier curves and Cox regression modelling was conducted. For both the logistic and Cox regression, unadjusted and adjusted (for age at MI, sex, deprivation quintile, and year of MI) models were generated using available case analysis. Cohen's kappa statistic was also

calculated to determine the level of agreement between the two lipid targets to be considered: ≥40% reduction in non-HDL from a statin-naïve baseline (NICE, 2014), LDL ≤1.8mmol/l (Piepoli *et al.*, 2016).

5.2.1 Descriptive Statistics, Confounding, and Reverse Causality

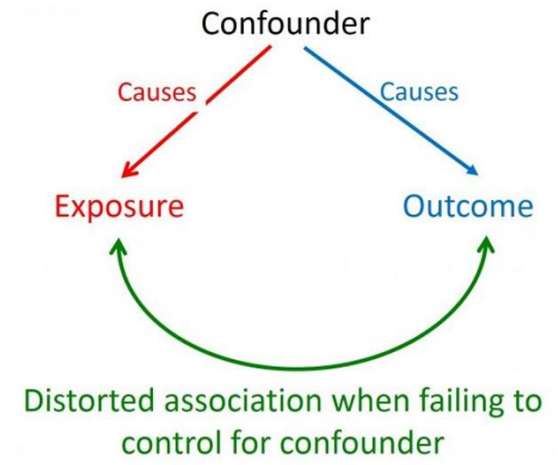
Initially, descriptive statistics are presented to simply illustrate potential associations between exposures (adherence and lipid targets) with outcomes (further hospital MIs and death). For continuous variables, such as age at the start of follow up, several such statistics are presented: mean, standard deviation, median, interquartile range (the 25th and 75th percentiles), minimum and maximum. For categorical variables, such as deprivation quintiles, the number of observations (N) and percentages are provided. For each descriptive summary, the Ns are the numbers of observations with the necessary data available, unless stated otherwise and labelled as missing. Likewise, the summaries of continuous variables are calculated using only those with the data available.

However, whilst descriptive statistics are useful for summarising the data, any associations highlighted by them will require further analytical investigation. This is particularly the case within observational cohorts, such as this one, where confounding variables are uncontrolled for, and therefore true associations may be masked or inflated, as illustrated in Figure 5.1. Adjusting for such confounders using other statistical methods may generate a clearer picture of the extent of associations (McNamee, 2003; Catalogue of Bias Collaboration *et al.*, 2018). In observational data, though, this is seldom sufficient to establish causality due to the possibility of unmeasured confounders which could explain the presence or absence of such a relationship (Catalogue of Bias Collaboration *et al.*, 2018). Consequently, causal relationships can be difficult to establish reliably using observational data (Catalogue of Bias Collaboration *et al.*, 2018; Coscia Requena, Muriel and Peñuelas, 2018; Sheetz and Nathan, 2020).

Nevertheless, some methods can facilitate causal inferences in observational settings, including propensity scoring and instrumental variable analysis, which are commonly utilised. Propensity scores allow the estimation of the probability of treatment assignment based on the patient's characteristics, and these scores

can then be matched, stratified, used as a covariable in models, or used to calculate the inverse probability treatment weighting (Coscia Requena, Muriel and Peñuelas, 2018). This facilitates a structure to the data similar to that of a clinical trial by balancing the exposure groups. However, matching can result in the exclusion of unmatched patients from the study, and the correct selection of the variables for inclusion in the study is crucial to minimise bias and impact on variance (Coscia Requena, Muriel and Peñuelas, 2018; Sheetz and Nathan, 2020). Although the identification of strong, suitable variables is its principal limitation, instrumental variable analysis, which uses variables that predict exposure but not the outcome, offers the advantage of removing the effects of both measured and unmeasured confounders. If such a variable can be identified, though, patients are then compared by their likelihood of exposure, similar to intention-to-treat analyses in RCTs (Sheetz and Nathan, 2020). However, neither of these methods were employed within the analysis in this chapter, and therefore no causal inferences are made regarding the associations observed.

Within the descriptive statistics and the subsequent statistical models performed, there is also the possibility of reverse causality. This is where one may assume that the exposure causes the disease, but the opposite is true. Whilst particularly likely in cross-sectional studies (where temporal relationships are hard to ascertain), reverse causality can also occur in longitudinal analyses. Underlying processes for the disease may have already commenced and caused the exposure to be more likely and lead to the outcome in question (Katz, 2006; Wanberg, 2012). Unlike confounding, there is no standard approach for addressing such situations. Methods used by researchers typically involve sensitivity analyses which may involve the exclusion of those unwell at baseline, stratification by potential markers of underlying processes (such as age), or the exclusion of those who experience the outcome early on in follow up (Sattar and Preiss, 2017).



Criteria:

1. Be a cause of the disease, or a surrogate measure of a cause, in unexposed people
2. Be correlated, positively or negatively, with exposure in the study population
3. *Not* be affected by the exposure

Figure 5.1: Criteria for Confounding Variable. The diagram is taken from the Catelogue of Bias Collaboration et al (2018), with the criteria for confounding taken from McNamee (2003).

5.2.2 Cohen's Kappa Statistic

Cohen's kappa statistic is used to determine the level of agreement between two categorical variables. It is calculated as a ratio, as illustrated in Equation 5.1. The numerator represents the proportion of disagreements between the two variables that were observed beyond those which would have been expected, whilst the denominator is the proportion of disagreements that would have been expected. The expected number of disagreements represents the number that would be anticipated if the two variables were allocated randomly by chance (Cohen, 1960).

$$\kappa = \frac{p_0 - p_e}{1 - p_e},$$

$$where\ p_0 = proportion\ of\ cases\ where\ variables\ agreed, and$$
$$p_e = proportion\ of\ cases\ where\ variables\ were\ expected\ to\ agree\ by\ chance.$$

Equation 5.1: Cohen's Kappa Statistic for Agreement (Cohen, 1960)

Using this value, the level of agreement could then be assessed using the arbitrary thresholds provided in Table 5.1, as first outlined by Landis and Koch (1977). In this chapter, for patients where it could be ascertained whether a ≥40% reduction in non-HDL from a pre-MI baseline and an LDL ≤1.8mmol/l at any point during follow up was achieved, a Cohen's kappa statistic for agreement was calculated.

Table 5.1: Agreement Thresholds for Cohen's Kappa Statistic from (Landis and Koch, 1977)

Kappa Statistic	Level of Agreement
<0.0	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

5.2.3 Logistic Regression

Logistic regression is a method used to investigate the associations between one or more exposure variables with the binary outcome variable, where p is the probability that the outcome occurs, and the odds of the outcome are $\frac{p}{1-p}$. This is (natural) log-transformed to form the logit function, and the general form of the model is shown in Equation 5.2 (Kirkwood and Sterne, 2003).

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots$$

Equation 5.2: General Form of a Logistic Regression Model (Kirkwood and Sterne, 2003)

Within the model, β_0, represents the baseline odds, whilst all subsequent parameters (β_i) indicate the odds associated with the corresponding exposure variables (x_i). Following exponentiation, these parameters become the odds ratio (OR) associated with that exposure (Kirkwood and Sterne, 2003).

In this chapter, logistic regression is performed for the outcome variables: further hospital admissions for MIs (yes, no), and the achievement of lipid targets (yes, no), with odds ratios and confidence intervals, plotted using the 'forestplot' package in R (Gordon and Lumley, 2020). For each, the exposure variable is average adherence (of a patient's complete year-long time windows), with cut-offs 50% and 80%, and as a continuous measure. Where adherence is categorised with the cut-offs, the reference group is those whose average adherence was above it, so the odds are for those who are less adherent. When continuous adherence is used, the odds are calculated per 10% decrease, which can be obtained by inverting the parameter and raising it to the power of ten.

This approach was also taken in a sensitivity analyses where a patient's average adherence was calculated including time windows which were shorter than one year. The achievement of lipid targets is also used as an exposure variable for further hospital admissions for MIs, where, similar to models with adherence as the exposure, the achievement of each target is the reference category.

5.2.4 Kaplan-Meier Curves

Survival analysis methods are used when the outcome of interest is the time to a defined event, as not all individuals have experienced the event and their length of follow up, or 'time at risk', is different for each individual. Using these methods, individuals are censored at the last time point that it is known that they had not experienced the event, which is usually the end of their follow up (Bull and Spiegelhalter, 1997; Kirkwood and Sterne, 2003; Harrell, 2015b).

$Let\ n_t = no.patients\ at\ risk, and\ \mathrm{d}_t = no.events\ occuring\ exactly, at\ time\ t.$
$Then\ the\ risk\ of\ an\ event\ at\ time\ t,$

$$r_t = \frac{\mathrm{d}_t}{n_t},$$

$So, the\ survival\ probability,$

$$s_t = 1 - r_t = \frac{n_t - \mathrm{d}_t}{n_t}$$

$Let\ t_i\ be\ the\ time\ where\ the\ ith\ event\ occurs, then,$

$$S(t) = \prod_{i:t_i \le t} s_{t_i} = \prod_{i:t_i \le t} \frac{n_{t_i} - \mathrm{d}_{t_i}}{n_{t_i}}$$

Equation 5.3: Cumulative Survival Probability at time t The number at risk is the number of individuals who have not experienced the event, or been censored, prior to time t, but includes those who experience the event or are censored at time t.(Kirkwood and Sterne, 2003; Harrell, 2015b)

Kaplan-Meier curves are a non-parametric method of survival analysis which does not assume the shape of the survival distribution and plots the estimated cumulative survival probability, *S(t)*, against time as a step function (Bull and Spiegelhalter, 1997; Kirkwood and Sterne, 2003; Harrell, 2015b). The calculation of *S(t)* is in two stages. First, for each time *t*, that an event occurs within the dataset, the survival probability at that time is calculated. These instantaneous survival probabilities are then used to generate the cumulative survival probability *S(t)*, as shown in Equation 5.3 (Kirkwood and Sterne, 2003; Harrell, 2015b).

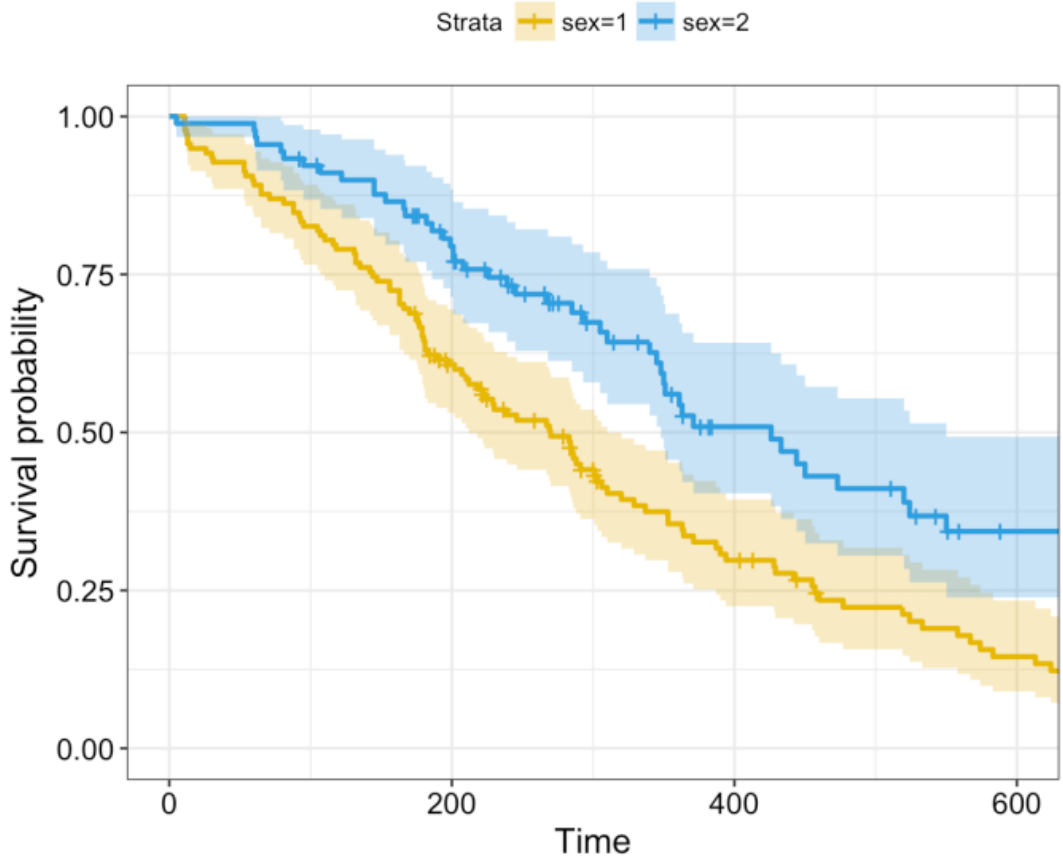


Figure 5.2: Example of Kaplan-Meier Curve. Graph taken from R-Bloggers (R-Bloggers, 2016b)**, generated using data available in the 'survival' R package** (Therneau, 2020)**, and plotted using the 'survminer' package** (Kassambara *et al.*, 2020)**.**

These estimates are plotted against time (Bull and Spiegelhalter, 1997; Kirkwood and Sterne, 2003; Harrell, 2015b), as shown in Figure 5.2. Typically, Kaplan-Meier curves are used to compare the survival rates between two or more discrete groups, with larger differences between the lines indicating higher survival in one group (Bull and Spiegelhalter, 1997). Quantiles can also be estimated from these graphs or from the table of calculated cumulative survival probabilities. Censored observations (usually denoted by a '+') and confidence intervals can be included on the graphs to illustrate where greater uncertainty may lie and indicate the strength of any differences between groups (Bull and Spiegelhalter, 1997; Harrell, 2015b). In the R package used for this analysis, these are constructed using the Normal approximation on the log scale and then exponentiated (Therneau, 2020).

For this analysis, Kaplan-Meier curves were produced for the mortality (all-cause and circulatory) outcome only and were plotted for overall survival, and by average adherence (with 50% and 80% cut-offs), and achievement of lipid targets

during follow up. In a sensitivity analysis, the 50% and 80% cut-offs were also applied to average adherence calculated including incomplete years of follow up. Due to the number of censored observations, these were not marked on the graphs but the number at risk and number of events at the end of each year of follow up are tabulated below each graph. They were generated using the 'survival' (Therneau, 2020) and 'survminer' (Kassambara *et al.*, 2020) packages within R v3.5.0 (R Core Team, 2018). Corresponding baseline hazards were plotted using the 'muhaz' package (Hess, Gentleman and Winsemius, 2019).

5.2.5 Cox Regression

The Cox Proportional Hazard Model consists of a baseline hazard and a collection of predictor variables with their estimated parameters, as shown in Equation 5.4. This model is semi-parametric as although the regression of the predictor variables is fully parameterised, there are no assumptions made regarding the form of the baseline hazard (Cox, 1972; Kleinbaum and Klein, 2012; Harrell, 2015a).

$$\log\big(h(t)\big) = \log\big(h_0(t)\big) + \beta_1 x_1 + \beta_2 x_2 + \cdots, \qquad (1)$$
$$where\ h(t)\ is\ the\ hazard\ function\ at\ time\ t, and\ h_0(t)\ is\ the\ baseline\ hazard.$$

$$By\ exponentiating\ both\ sides, (1)\ becomes:$$
$$h(t) = h_0(t) * \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots), \qquad (2)$$

Equation 5.4: Hazard Function of the Cox Proportional Hazards Model (Cox, 1972; Kirkwood and Sterne, 2003; Kleinbaum and Klein, 2012; Harrell, 2015a)

As with logistic regression, the β_i are the estimated parameters associated with the covariates or predictor variables, x_i, where the exponentiated parameter $(\exp(\beta_i))$ yields the hazard ratio for that covariate. This estimated effect of the covariate is independent of the time t and emphasises the proportional hazards assumption of the Cox model (Cox, 1972; Kirkwood and Sterne, 2003; Kleinbaum and Klein, 2012; Harrell, 2015a).

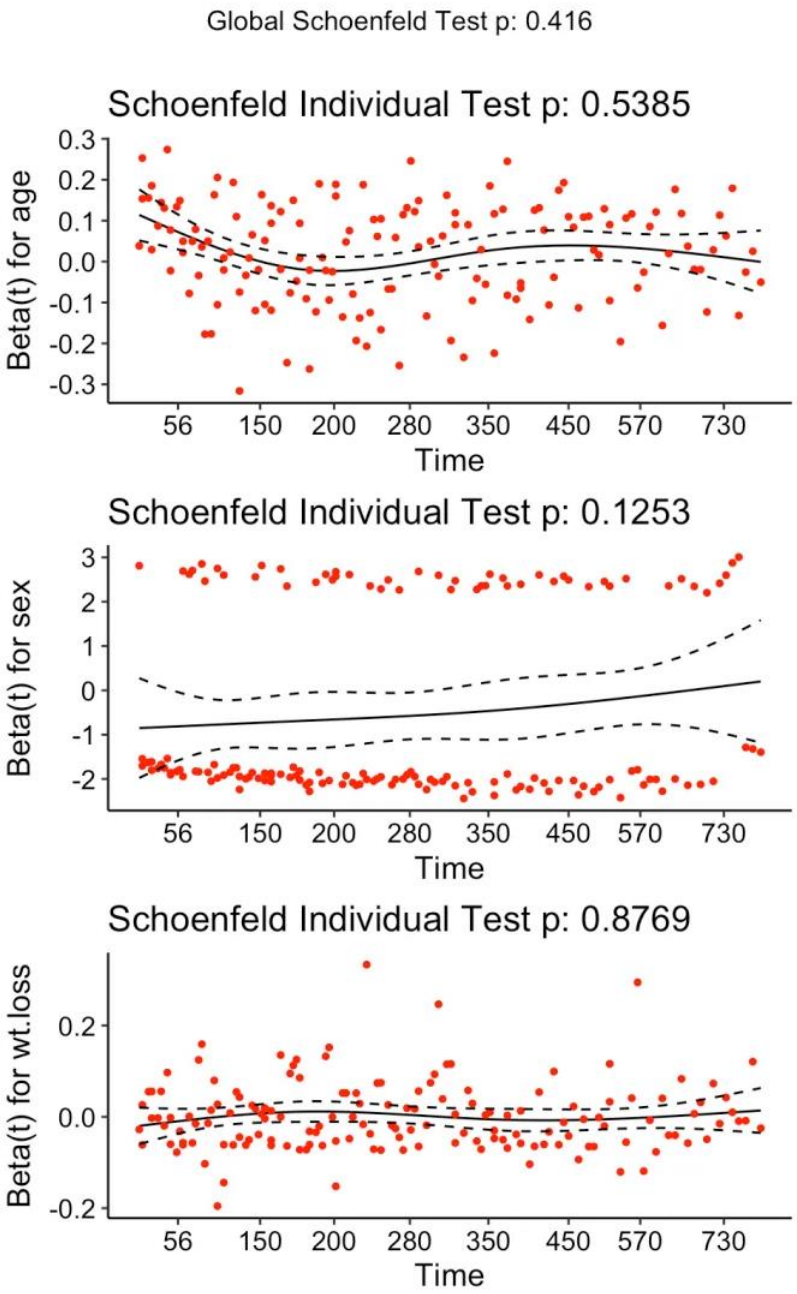


Figure 5.3: Schoenfeld Residual Plot to Check the Proportional Hazards Assumption (R-Bloggers, 2016a)

In this chapter, this assumption was checked using the partial Schoenfeld residuals for each covariate (Harrell, 2015a). These residuals are defined as the observed value of the covariates minus its expected (or mean) value given the number at risk at that time (Schoenfeld, 1982), which are then scaled (Harrell, 2015a; Therneau, 2020). If the proportional hazards assumption holds, then when plotted against time, there should be no pattern between the residuals and time (Harrell, 2015a). In the 'survival' package, an accompanying formal test can also be implemented, both for an individual covariate and the overall model (Harrell, 2015a; Therneau, 2020). This takes the form of a score test,

which calculates the correlation between the residuals and the order of the failure times (Grambsch and Therneau, 1994; Harrell, 2015a). When plotted using the 'survminer' package, the p-values corresponding to these tests are included as part of the graphs by default (Kassambara *et al.*, 2020), as shown in Figure 5.3.

Mortality (all-cause and circulatory) was the only outcome considered in the analysis, with its relationship with average adherence (continuous, 50% and 80% cut-offs), and achievement of lipid targets investigated. In line with the logistic models, hazard ratios reported for continuous adherence are per 10% decrease in average adherence, and in a sensitivity analysis, hazard ratios were also calculated for average adherence estimated using incomplete, as well as complete, year-long time windows. Hazard ratios and their 95% confidence intervals were plotted using the 'forestplot' package in R (Gordon and Lumley, 2020).

5.3 Recurrent Hospital Admissions for Myocardial Infarctions

All further MI admissions to hospital were included irrespective of the patient's mortality status following them. However, this number is likely to be an underestimate of the number of recurrent MIs experienced as not all fatal MIs will have resulted in hospital admission. Consequently, deaths due to circulatory conditions are examined separately within the subsequent sections focussing on mortality within this population (Chapter 5.4). Nonetheless, fatal recurrent MIs were still included in this analysis as the economic costs to the healthcare system of non-fatal and fatal MIs resulting in a hospital admission are likely to be similar.

There were a further 4,209 hospital admissions for MIs in the post MI population before 31st July 2017, which occurred in 3,038 patients (3038/11110, 27.3%). The number of further MIs resulting in a hospital admission that a patient had (i.e. excluding the first non-fatal one defining the start of follow up) is shown in Table 5.2. The majority of those with further hospital admissions for an MI had only one, but some patients had a substantially higher number of MI admissions recorded, with the maximum number being 16. The higher number of individuals

with MI admissions in follow up compared to those with MIs before their baseline date was not unexpected. To be included in this population, patients would have had to have survived not only the MI of interest but also the MIs that had occurred in the prior ten years and the years in between. This makes the those with prior MIs and those with subsequent MIs subtly distinct. However, the higher number could also be partially as a result of changes in MI definition over the lookback period and during follow up (Thygesen *et al.*, 2007, 2012) or incomplete early reporting in hospital records.

Table 5.2: Number of Further Hospital Admissions for MIs experienced in each patient who had at least one after the index admission.

No. Further MIs	N (%)
1	2,290 (75.4%)
2	502 (16.5%)
3	160 (5.3%)
4	49 (1.6%)
5	21 (0.7%)
>5	16 (0.5%)

Numbers exclude MIs which defined the patient's start date. Numbers are N (%) of those with Further Hospitalised MIs.

Figure 5.4 details the individual patterns of further hospital admissions for MIs throughout follow up. From this, and Table 5.3, it is clear that the greatest risk of experiencing a further MI admission was in the first year after an MI admission, with the risk decreasing as follow up duration increases. Indeed, over half of the subsequent hospital admissions for MIs occurred in the first year (59.3%), and in nearly three quarters (70.6%) of the patients who experienced further MI admissions. A decrease in the number of patients with data available at each year of follow up (due to follow up ending 31st July 2017) may account for some of the apparent decreasing risk, but in the first three years, the number of patients available only decreases due to patient deaths. Furthermore, as shown in Figure 5.4, this pattern was not confined to the patient's baseline MI, but with all subsequent MI admissions. Moreover, as Table 5.3 illustrates, some patients experienced multiple further MI admissions within a year of follow up, with only the ninth year reporting the same number of hospitalised MIs as patients being admitted with them.

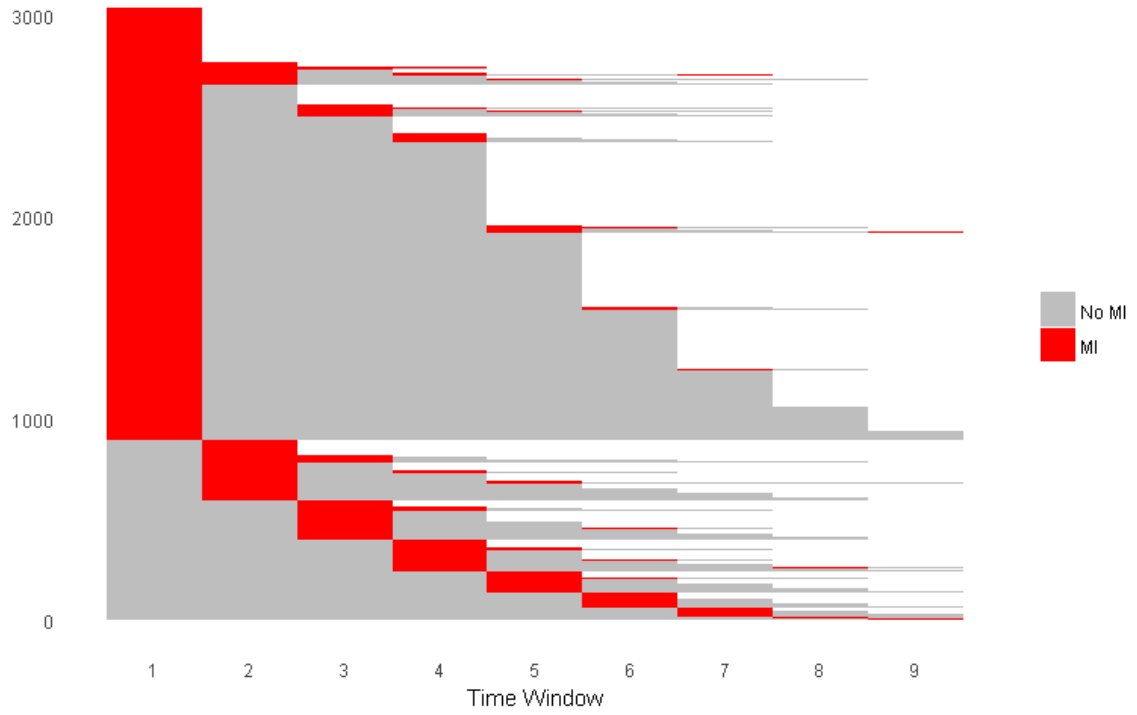


Figure 5.4: Individual Patterns of Further Hospitalised MIs During Follow Up. Each time window was categorised by whether further MI admissions occurred during it, with red indicating patients had an admission that year and grey indicating they did not. Patients were sorted by their further MI admission patterns before plotting, with lower patient numbers assigned to those with a longer period before their first recurrent MI admission. Only patients with a recurrent MI admission during their follow up were included (N = 3,038).

Table 5.3: Hospitalised MIs by Year of Follow Up

	Year of Follow Up								
	1	2	3	4	5	6	7	8	9
# MIs	2,496	525	359	308	231	159	88	35	8
# Pts	2,144	418	311	250	191	133	72	32	8
# At Risk	11,110	10,082	9,316	8,675	6,558	4,753	3,136	1,684	567

Patients may have more than one MI in any given year of follow up. # at risk is the number of patients still alive and not reached the end of follow up at the start of that year. # = Number; Pts = Patients.

When comparing the population of those with and without further hospital admissions for MIs (Table 5.4), the gender split was broadly similar between the two groups. There was also no substantial difference in age, with those who did not have a subsequent MI admission just half a year older than those who did (67.0 vs 66.5 years) on average. In terms of deprivation, patients who had a further MI admission were more likely to reside in more deprived areas (by quintile). However, the biggest differences between the two groups were observed in the presence of a diagnosis of diabetes, and the number of MIs prior to the start of follow up. In patients who had been diagnosed with diabetes approximately 6% more went on to have an MI admission than in patients without

diabetes (23.1% vs 17.3%). Higher proportions were observed in both Type 1 and Type 2 diabetes. For prior MIs, the percentage was nearly double in those with further MI admissions than those without (9.3% vs 5.4%).

Table 5.4: Demographics by Further Hospitalised MIs During Follow Up

	No Further MI 8,072 (72.7%)	Further MI 3,038 (27.3%)
Gender		
Male	4,897 (60.7%)	1,835 (60.4%)
Female	3,175 (39.3%)	1,203 (39.6%)
Age at MI (years)		
Mean (SD)	67.0 (14.0)	66.5 (13.8)
Median	67.4	67.2
IQR	56.2 – 77.9	55.8 – 77.5
Range	19.3 – 102.6	20.5 – 97.9
SIMD 2012 Quintile (NHS GGC)		
1 (Most)	1,883 (23.5%)	767 (25.4%)
2	1,829 (22.8%)	699 (23.1%)
3	1,659 (20.7%)	635 (21.0%)
4	1,404 (17.5%)	473 (15.7%)
5 (Least)	1,236 (15.4%)	446 (14.8%)
Missing	61	18
Diabetes at MI	1,397 (17.3%)	702 (23.1%)
Type 1	71 (0.9%)	49 (1.6%)
Type 2	1,326 (16.4%)	653 (21.5%)
Prior MI	432 (5.4%)	282 (9.3%)
1	337 (4.2%)	221 (7.3%)
>1	95 (1.2%)	61 (2.0%)

Numbers are N (%) unless otherwise specified. Percentages are calculated within columns, except for the header where percentages are calculated from the total with information available. SD = Standard Deviation; IQR = Interquartile Range.

5.4 Deaths

5.4.1 Underlying Causes of Death

For those who died before the end of follow up, the underlying causes of death are presented in Table 5.5. There were 3,768 (33.9%) deaths, of which 2,894 patients (76.8%) had the cause of death data available. Causes of death were grouped by the ICD10 chapters (World Health Organization, 2016), with only the underlying causes of death presented. Of those with the cause of death data available, the most common causes of death were diseases of the circulatory

system (I00–I99), with more than twice as many deaths than any other ICD10 chapter. The next two biggest underlying causes of death were neoplasms (C00-C97, D00-D48) and respiratory diseases (J00-J99), representing 19.0% and 16.4% of deaths respectively. All other ICD10 chapters were responsible for fewer than 150 deaths each, with many responsible for fewer than 100.

Table 5.5: Distribution of Underlying Causes of Death (by ICD10 Chapter)

Cause of Death	N (%)
Diseases of the Circulatory System	1,207 (41.7%)
Neoplasms	551 (19.0%)
Diseases of the Respiratory System	476 (16.4%)
Mental and Behavioural Disorders	142 (4.9%)
Diseases of the Digestive System	138 (4.8%)
Other	380 (13.1%)
Missing	874

Numbers are N (%) of those with the cause of death available.

Table 5.6: Distribution of Circulatory Underlying Causes of Death (by ICD10 Subchapter)

Circulatory Causes of Death	N (%)
Ischaemic Heart Disease (I20-I25)	836 (69.3%)
Cerebrovascular Diseases (I60-I69)	173 (14.3%)
Other Forms of Heart Disease (I30-I52)	128 (10.6%)
Other Circulatory Deaths*	70 (5.8%)
Total (I00-I99)	**1,207**

Numbers are N (%) of those with a circulatory cause of death (defined by ICD10 codes I00-I99). *including Chronic Rheumatic Heart Diseases (I05-I09); Hypertensive Diseases (I10-I15); Pulmonary Heart Disease and Diseases of Pulmonary Circulation (I26-I28); Diseases of Arteries, Arterioles and Capillaries (I70-I79); Diseases of Veins, Lymphatic Vessels and Lymph Nodes, Not Elsewhere Classified (I80-I89). No circulatory deaths were recorded for Acute Rheumatic Fever (I00-I02) or Other and Unspecified Disorders of the Circulatory System (I95-I99).

An examination of the ICD10 subchapters for those who died due to diseases of the circulatory system (Table 5.6) revealed that over half of the circulatory deaths were due to ischaemic heart disease (I20-I25), including MIs. Cerebrovascular diseases (I60-I69), principally strokes, and other forms of heart disease (I30-I52), including heart failure, were also responsible for over 100 deaths within this population. No other subchapters in this group were responsible for more than 50 deaths.

5.4.2 Time to Death

When looking at the number of deaths by year of follow up (Table 5.7), it is evident that more deaths occurred nearer the beginning of follow up (i.e. closer to the MI). This trend is also present in the numbers of deaths due to circulatory diseases, although the percentage of deaths due to circulatory disease does slightly increase between years two and six, before subsequently decreasing. It is also important to note that the number of circulatory deaths is likely to be underestimated, with patients missing a cause of death conservatively assigned to non-circulatory causes for much of the subsequent analyses unless otherwise specified.

Table 5.7: Deaths by Year of Follow Up

	Year of Follow Up								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
# Deaths	1,028	767	643	540	358	221	157	44	10
# Circ.	309	225	205	191	130	81	46	18	2
# At Risk	11,110	10,082	9,316	8,675	6,558	4,753	3,136	1,684	567

 # at risk is the number of patients still alive and not reached the end of follow up at the start of that year. Circulatory (Circ.) cause of death defined by ICD10 codes I00-I99 as the underlying cause of death.

Using the dates of death, the Kaplan-Meier plots (Figure 5.5) also show a slightly sharper decrease in the first couple of years of follow up before maintaining a shallower decline. This was less evident in the circulatory deaths plot, though this is likely largely due to the smaller number of events in this population. Furthermore, whilst there was no censoring during the first three years of follow up by the design of the cohort, there was a high number of censored observations after this time point (not marked in Figure 5.5). For both plots, the data was not mature enough to estimate the median survival time. However, for all-cause mortality only, there were enough events to estimate the lower quartile as 3.6 years (95% CI: 3.4–3.7 years).

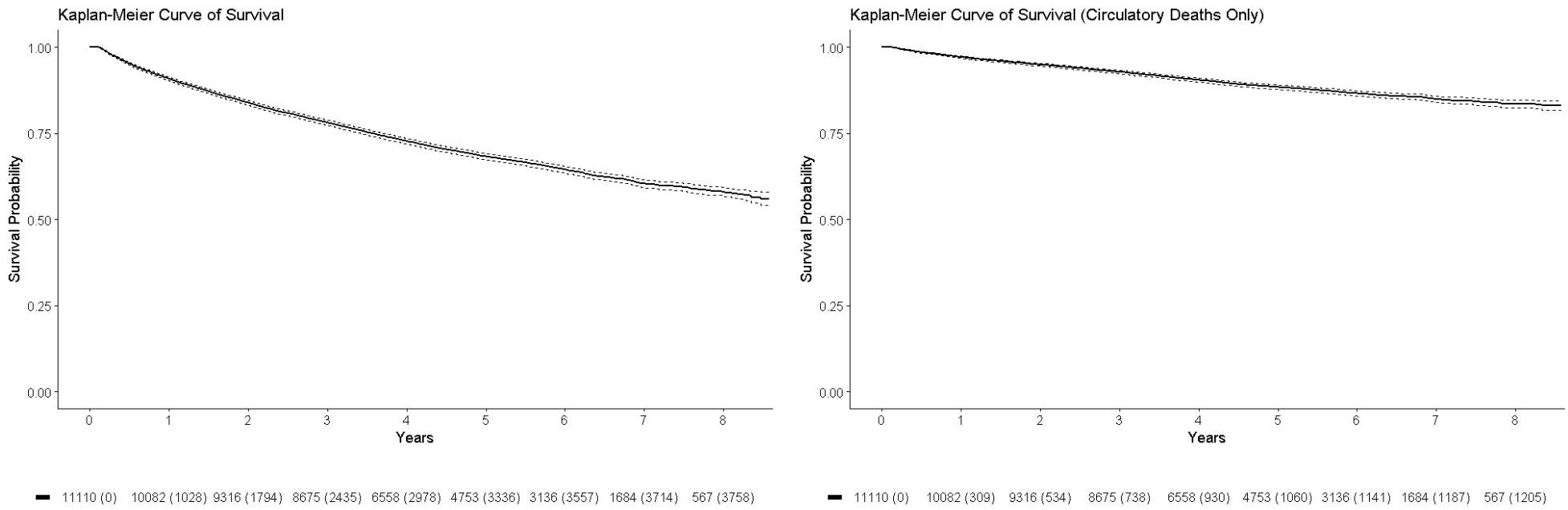


Figure 5.5: Kaplan-Meier Curves of Mortality (Overall and Circulatory Deaths) (with 95% CI). Numbers are Number at Risk (Number of Events). Circulatory cause of death defined by ICD10 I00-I99 as the underlying cause.

Table 5.8: Demographics by Mortality Status at End of Follow Up

	Alive 7,342 (66.1%)	Dead 3,768 (33.9%)	Circ. Death 1,207 (10.9%)	Non-Circ. Death 1,687 (15.2%)	Unknown Cause 874 (7.9%)
Gender					
Male	4,800 (65.4%)	1,932 (51.3%)	656 (54.3%)	860 (51.0%)	416 (47.6%)
Female	2,542 (34.6%)	1,836 (48.7%)	551 (45.7%)	827 (49.0%)	458 (52.4%)
Age at MI (years)					
Mean (SD)	62.0 (12.5)	76.5 (11.3)	76.3 (11.4)	76.2 (11.1)	77.2 (11.7)
Median	61.3	78.1	78.0	77.5	79.5
IQR	52.7 – 71.6	69.8 – 84.8	69.4 – 84.5	69.7 – 84.3	70.5 – 85.9
Range	19.3 – 98.1	26.5 – 102.6	36.2 – 102.6	30.6 – 102.4	26.5 – 100.7
SIMD 2012 Quintile (NHS GGC)					
1 (Most)	1,760 (24.2%)	890 (23.7%)	280 (23.3%)	399 (23.8%)	211 (24.2%)
2	1,638 (22.5%)	890 (23.7%)	288 (24.0%)	412 (24.6%)	190 (21.8%)
3	1,470 (20.2%)	824 (22.0%)	262 (21.8%)	365 (21.8%)	197 (22.6%)
4	1,219 (16.7%)	658 (17.6%)	215 (17.9%)	283 (16.9%)	160 (18.4%)
5 (Least)	1,196 (16.4%)	486 (13.0%)	155 (12.9%)	218 (13.0%)	113 (13.0%)
Missing	59	20	7	10	3
Diabetes at MI	1,045 (14.2%)	1,054 (28.0%)	364 (30.2%)	445 (26.4%)	245 (28.0%)
Type 1	61 (0.8%)	59 (1.6%)	16 (1.3%)	31 (1.8%)	12 (1.4%)
Type 2	984 (13.4%)	995 (26.4%)	348 (28.8%)	414 (24.5%)	233 (26.7%)
Prior MI	344 (4.7%)	370 (9.8%)	122 (10.1%)	139 (8.2%)	109 (12.5%)
1	266 (3.6%)	292 (7.7%)	94 (7.8%)	114 (6.8%)	84 (9.6%)
>1	78 (1.1%)	78 (2.1%)	28 (2.3%)	25 (1.5%)	25 (2.9%)

Numbers are N (%) unless otherwise specified. Percentages are calculated within columns, except for the header where they are calculated from total with information available. SD = Standard Deviation; IQR = Interquartile Range; Circulatory (Circ.) cause of death defined by ICD10 I00-I99 as underlying cause.

5.4.3 Demographics by Mortality

The demographics of those who had died and those who were alive at the end of follow up are compared in Table 5.8. A higher proportion of those who died were female (48.7% vs 34.6% in survivors). Additionally, compared to those who were alive at the end of follow up, those who died were typically around 15 years older (76.5 vs 62.0 years), were twice as likely to have a diagnosis of diabetes before the start of follow up (28.0% vs 14.2%), and the percentage who had experienced a prior MI was also more than double (9.8% vs 4.7%). There did not appear to be any substantial differences in the patient's deprivation quintile between the groups.

When differences between circulatory and non-circulatory causes of death were compared (Table 5.8), there were no differences in age or deprivation. Those who died due to a circulatory disease were slightly more likely to be male (54.3% vs 51.0%), as well as having had a prior MI (10.1% vs 8.2%). Those with a circulatory cause of death were also more likely to have had a diagnosis of diabetes before the start of follow up (30.2% vs 26.4%), although there were some differences by the type of diabetes. There was a higher prevalence of Type 2 diabetes (28.8% vs 24.5%) and a lower prevalence of Type 1 diabetes (1.3% vs 1.8%), in those with a circulatory cause of death, compared to those who died due to other causes.

5.5 Plasma Lipid Levels and Subsequent Outcomes

Two lipid targets were considered: a non-HDL target in line with NICE recommendations, where ≥40% reduction from a pre-MI non-medicated baseline was targeted (NICE, 2014), and an LDL target of ≤1.8mmol/l from the ESC guidelines (Piepoli *et al.*, 2016). Patients were classified as achieving the target if they achieved it in any of the years of follow up available for them, otherwise, they were defined as not achieving the target. It is important to note that the at-risk group for the non-HDL target is smaller as many patients did not have an available pre-MI statin-naïve baseline to facilitate the calculation of the percentage reduction. Further information surrounding the derivation of these targets within the data and descriptive statistics for those who met and did not meet these targets is given in Chapters 3.5 and 4.4 respectively.

Table 5.9: Events Stratified by Lipid Plasma Targets (NICE and ESC Recommendations)

	Target: ≥40% Reduction non-HDL*		Target: LDL ≤1.8mmol/l	
	Met	Not Met	Met	Not Met
	1,721 (51.3%)	1,634 (48.7%)	7,215 (75.9%)	2,289 (24.1%)
Further Hospitalised MI				
Yes	477 (27.7%)	499 (30.5%)	2,063 (28.6%)	635 (27.7%)
No	1,244 (72.2%)	1,135 (69.5%)	5,152 (71.4%)	1,654 (72.3%)
Death				
Yes	287 (16.7%)	518 (31.7%)	1,893 (26.2%)	663 (29.0%)
No	1,434 (83.3%)	1,116 (68.3%)	5,322 (73.8%)	1,626 (71.0%)
Cause of Death†	243	433	1,526	524
Circulatory‡	94 (38.7%)	177 (40.9%)	632 (41.4%)	224 (42.7%)
Neoplasms	64 (26.3%)	86 (19.9%)	312 (20.4%)	110 (21.0%)
Respiratory‡	33 (13.6%)	84 (19.4%)	247 (16.2%)	72 (13.7%)
Digestive‡	14 (5.8%)	17 (3.9%)	75 (4.9%)	26 (5.0%)
Other	38 (15.6%)	69 (15.9%)	260 (17.0%)	92 (17.6%)
Missing	44	85	367	139

***refers to the percentage reduction in non-HDL cholesterol from the pre-MI baseline. †where underlying cause of death information is available. ‡ refers to the ICD10 chapter with the title 'Diseases of the 'X' system'. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up. Numbers are N (%). Percentages are calculated within columns, except for the header where percentages are calculated from total with information available. Percentages for specific causes of death are calculated based on those with the information available.**

On examination of the raw frequencies of further MI admissions and deaths by the achievement of these lipid targets (Table 5.9), some differences were clear. Those who did not meet the non-HDL target were more likely to have had a further MI admission (30.5% vs 27.7%). The proportion of people in this group who died was nearly double that of those who did achieve the target (31.7% vs 16.7% respectively). There were also higher proportions of deaths from circulatory (40.9% vs 38.7%) and respiratory causes (19.4% vs 13.6%) amongst those who did not meet the non-HDL target than those who did meet the target. In contrast, there were no major differences in the proportion of patients who experienced a further hospital admission for MI between those who did and did not achieve the LDL target (28.6% vs 27.7%). However, as with the non-HDL target, a greater percentage of patients who did not meet the LDL target died (29.0% vs 26.2%), although this difference was less pronounced. The percentage of deaths due to circulatory diseases was also only slightly higher in those who did not meet the target (42.7% vs 41.4%). In contrast to what was observed with the non-HDL target, the proportion of deaths due to respiratory diseases was

higher in those that did meet the LDL target compared to those who did not (16.2% vs 13.7%).

These discrepancies in outcomes between the two targets may have been because there was only a slight agreement between the two targets, with a Cohen's kappa statistic for agreement of 0.16 (95% CI: 0.13, 0.20). In the group of 1,413 individuals that met both targets (compared to those who met only one or none of the targets), the differences observed were similar to those observed in the achievement of the non-HDL target; those who met both targets were less likely to have further MI admissions (28.3% vs 29.7%) or to die from any cause (16.6% vs 29.3%) or circulatory causes (5.7% vs 9.7%), than those who did not.

5.5.1 Association Between Targets and Recurrent Hospital Admissions for Myocardial Infarctions

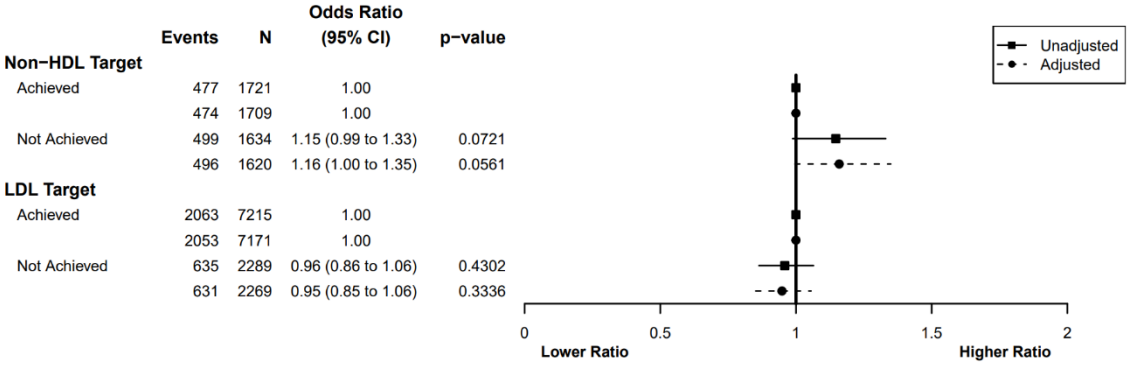


Figure 5.6: Odds Ratios of Further Hospitalised MIs by Achievement of Plasma Lipid Targets (unadjusted and adjusted for age, sex, year of MI, and SIMD quintile). Non-HDL target is ≥40% non-HDL reduction from a pre-MI baseline (NICE, 2014) **and the LDL target is ≤1.8mmol/l** (Piepoli *et al.*, 2016)**. Patients only included in analyses where all needed information was available. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up.**

In unadjusted and adjusted logistic regression models (Appendix G), there were no statistically significant associations, at the 5% level, between not achieving lipid targets and experiencing further MI admissions during follow up (Figure 5.6). Nonetheless, the non-HDL target did show trends towards associations; not achieving the non-HDL target during follow up was associated with a borderline significant 15% higher odds of a patient experiencing a further MI admission. In contrast, not achieving the LDL target was associated with a non-significant 4% lower odds. Furthermore, adjustment for age, sex, deprivation, and year of MI

resulted in minimal impact on the estimated odds ratios, with the odds increasing to 16% higher for the non-HDL target, and decreasing to 5% for LDL. In a sensitivity analysis (not presented), where the association with LDL was investigated amongst those with a non-HDL result, the odds ratio decreased to 15% lower odds of a further hospital admission for an MI, but this was not statistically significant in the adjusted model.

5.5.2 Association Between Targets and Mortality

Kaplan-Meiers for both targets and both overall mortality and circulatory causes (Figure 5.7 and Figure 5.8) show some separation in the survival probabilities between those who achieved and did not achieve the lipid targets, with the differences established early on in follow up and non-achievers faring worse. For the LDL target and all-cause mortality, this difference is established early and maintained throughout, with confidence intervals only overlapping at the maximal follow up, likely due to greater uncertainty. For circulatory causes, this difference is less pronounced, with the two lines similar throughout follow up and non-achievers faring marginally worse. For the non-HDL target, the difference is evident early on for both all-cause and circulatory causes, with the distance between the lines increasing as the length of time post MI increases.

As with the overall mortality Kaplan-Meier (Figure 5.5), only the lower quartile for each of the all-cause curves could be estimated. For both targets, the estimated lower quartile for not meeting the specific target was 4.2 years, although due to smaller numbers included for the non-HDL analysis, the corresponding confidence interval was slightly wider than the LDL target analysis (95% CI: 3.8–4.6 years vs 95% CI: 3.9–4.5 years). However, there was a bigger difference in the estimated lower quartiles for those who did achieve the targets. For those who achieved the LDL target, the lower quartile was estimated as 5.5 years (95% CI: 5.3–5.8 years), whereas the lower quartile for meeting the non-HDL target was 7.7 years, and the corresponding upper bound of the confidence interval inestimable (95% CI: 7.4–NA years). No other quartiles could be estimated from the graphs, except the median survival for those who did not meet the non-HDL target, which was 8.0 years (95% CI: 7.8–NA years), with the upper bound of the confidence interval also inestimable. No quartiles could be estimated for the mortality due to circulatory causes curves.

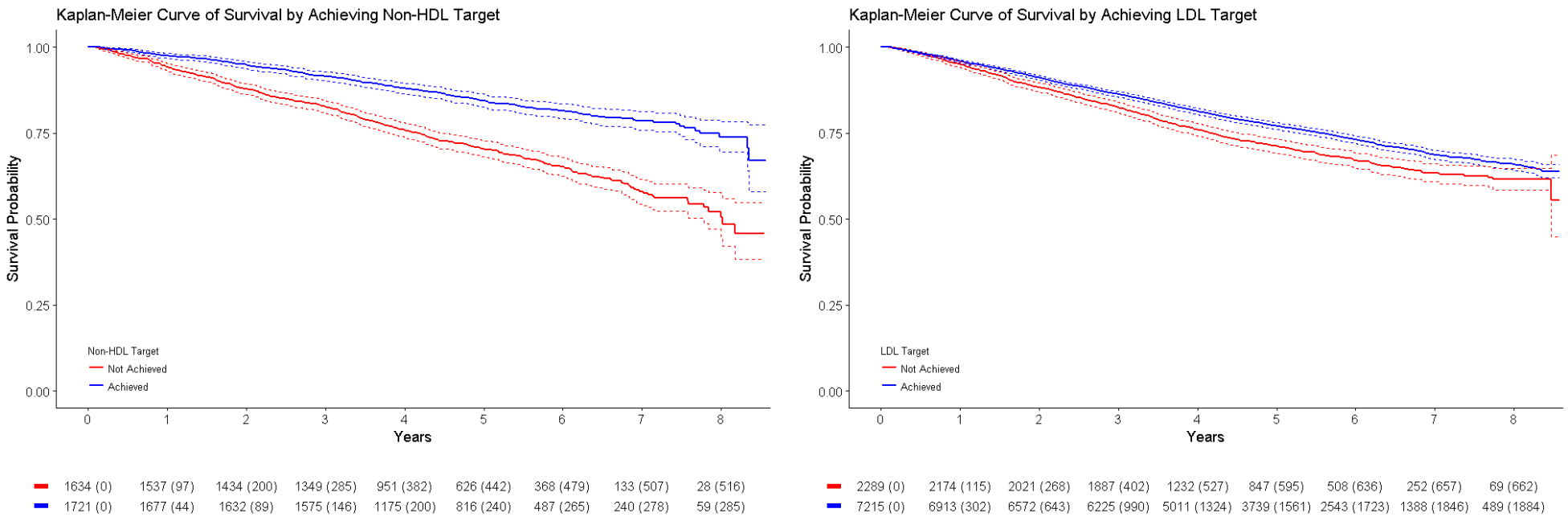


Figure 5.7: Kaplan-Meier Curves of Overall Mortality by Plasma Lipid Targets (with 95% Confidence Intervals). Numbers are Number at Risk (Number of Events). Non-HDL target is ≥40% non-HDL reduction from a pre-MI baseline (NICE, 2014) **and the LDL target is ≤1.8mmol/l** (Piepoli *et al.*, 2016)**. Patients only included in analyses where information was available. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up.**

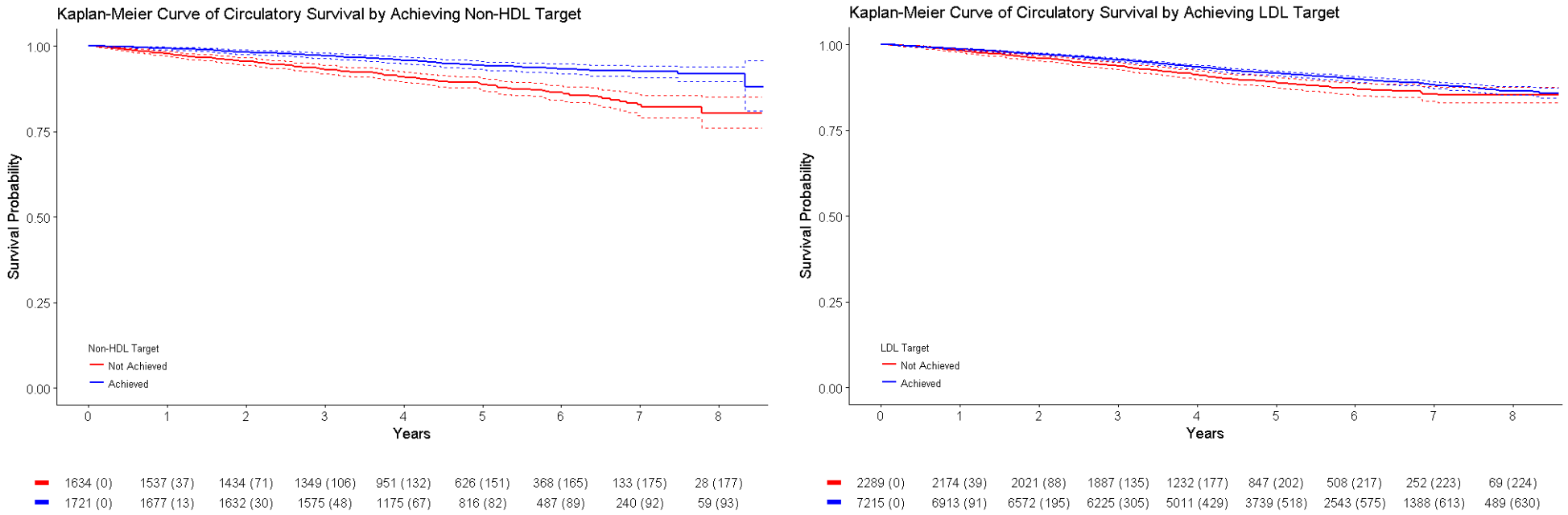
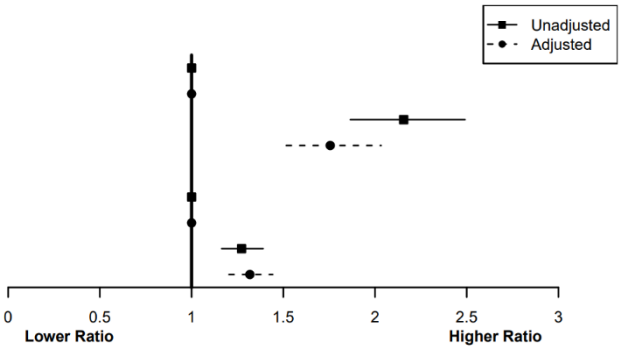


Figure 5.8: Kaplan-Meier Curves of Circulatory Mortality by Plasma Lipid Targets (with 95% Confidence Intervals). Numbers are Number at Risk (Number of Events). Circulatory cause of death defined by ICD10 I00-I99 as underlying cause. Non-HDL target is ≥40% non-HDL reduction from a pre-MI baseline (NICE, 2014) and the LDL target is ≤1.8mmol/l (Piepoli *et al.*, 2016). Patients only included in analyses where information was available. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up.

A - Overall Mortality

	Events	N	Hazard Ratio (95% CI)	p-value
Non-HDL Target				
Achieved	287	1721	1.00	
	286	1709	1.00	
Not Achieved	518	1634	2.16 (1.87 to 2.49)	<0.0001
	516	1620	1.76 (1.52 to 2.03)	<0.0001
LDL Target				
Achieved	1893	7215	1.00	
	1886	7171	1.00	
Not Achieved	663	2289	1.27 (1.16 to 1.39)	<0.0001
	658	2269	1.32 (1.20 to 1.44)	<0.0001



B - Circulatory Causes

	Events	N	Hazard Ratio (95% CI)	p-value
Non-HDL Target				
Achieved	94	1721	1.00	
	94	1709	1.00	
Not Achieved	177	1634	2.24 (1.74 to 2.87)	<0.0001
	175	1620	1.82 (1.41 to 2.34)	<0.0001
LDL Target				
Achieved	632	7215	1.00	
	629	7171	1.00	
Not Achieved	224	2289	1.30 (1.11 to 1.51)	0.0011
	222	2269	1.30 (1.11 to 1.52)	0.0012

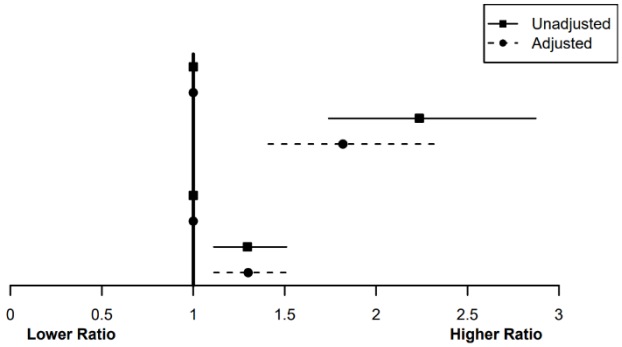


Figure 5.9: Hazard Ratios of Mortality by Achievement of Plasma Lipid Targets (unadjusted and adjusted for age, sex, year of MI, and SIMD quintile). A = all-cause mortality; B = circulatory cause of death (defined by ICD10 I00-I99 as the underlying cause). Non-HDL target is ≥40% non-HDL reduction from a pre-MI baseline (NICE, 2014) **and the LDL target is ≤1.8mmol/l** (Piepoli *et al.*, 2016)**. Patients only included in analyses where all needed information was available. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up.**

Cox regression showed that not achieving the non-HDL target was associated with 2.2 times greater hazard of all-cause mortality compared to those who did achieve the target and that not meeting the LDL target was associated with 1.3 times greater hazard, before any adjustment for other variables (Figure 5.9a). Nearly identical hazards were also observed when circulatory causes alone were considered (Figure 5.9b). When adjusted for age, sex, deprivation quintile, and year of MI, this did not attenuate the elevated hazard associated with not meeting the LDL target for either cause but did reduce the hazard associated with failure to achieve the non-HDL target to 1.8 times higher for both all-cause and circulatory causes. This result, along with all others, remained statistically significant at the 5% level. Full details of the models, including Schoenfeld residual plots, are included in Appendix I. As with earlier models, a sensitivity analysis of the LDL association in the non-HDL at-risk group (not presented), the

associations with all-cause and circulatory mortality remained statistically significant and the hazard ratios increased marginally.

5.6 Statin Adherence and Subsequent Outcomes

Patient statin adherence was assessed using the MPR and was calculated for each complete year of follow up that the individual patient had available. For each patient, their mean adherence was then derived. As this value needed at least one complete year of follow up, only patients who were alive at the end of the first year post MI were included in these analyses (9.3% died in the first year). Additionally, patients were then classified as adherent or non-adherent, using two different thresholds, 50% and 80% (Chapter 4.5.1). The derivation of a patient's statin adherence and further descriptive statistics can be found in Chapters 3.4 and 4.5 respectively.

Table 5.10: Events by Statin Adherence Thresholds (50% and 80%)

	≥50% MPR 8,522 (84.5%)	<50% MPR 1,560 (15.5%)	≥80% MPR 6,942 (68.9%)	<80% MPR 3,140 (31.1%)
Further Hospitalised MI				
Yes	2,366 (27.8%)	399 (25.6%)	1,875 (27.0%)	890 (28.3%)
No	6,156 (72.2%)	1,161 (74.4%)	5,067 (73.0%)	2,250 (71.7%)
Death				
Yes	2,175 (25.5%)	565 (36.2%)	1,666 (24.0%)	1,074 (34.2%)
No	6,347 (74.5%)	995 (63.8%)	5,276 (76.0%)	2,066 (65.8%)
Cause of Death*	1,808	469	1,396	881
Circulatory†	712 (39.4%)	186 (39.7%)	561 (40.2%)	337 (38.3%)
Neoplasms	375 (20.7%)	71 (15.1%)	282 (20.2%)	164 (18.6%)
Respiratory†	301 (16.6%)	102 (21.7%)	233 (16.7%)	170 (19.3%)
Men.&Beh.‡	95 (5.3%)	29 (6.2%)	64 (4.6%)	60 (6.8%)
Digestive†	82 (4.5%)	20 (4.3%)	67 (4.8%)	35 (4.0%)
Other	243 (13.4%)	61 (13.0%)	189 (13.5%)	115 (13.1%)
Missing	367	96	270	193

***where the underlying cause of death information is available. †refers to the ICD10 chapter with the title 'Diseases of the 'X' system'. ‡refers to the ICD10 Chapter 'Mental and Behavioural Disorders'. MPR = Medication Possession Ratio. NF/F-MI = Non-Fatal/Fatal Further Myocardial Infarction. Numbers are N (%). Percentages are calculated within columns, except for the header where percentages are calculated from total with information available. Percentages for specific causes of death are calculated based on those with information available.**

The relationship between adherence thresholds and further hospital admissions for MIs was not consistent (Table 5.10). For the 50% adherence threshold, those

with an average adherence above the cut-off were more likely to have further MI admissions than those with adherence below it (27.8% vs 25.6%). However, when the 80% cut-off is considered, there was little difference in the percentages of patients with further hospital admissions for MIs, with those whose adherence was lower slightly more likely to have further admissions (28.3% vs 27.0%).

For mortality, those with average adherence lower than both thresholds were more likely to die during follow up (50%: 36.2% vs 25.5%, 80%: 34.2% vs 24.0%). When the cause of death was considered, at the 50% cut-off there was no difference in the percentage who died due to circulatory disease (<50%: 39.7% vs ≥50%: 39.4%). Those whose average adherence was ≥80% were slightly more likely to have a circulatory cause of death than those whose adherence was below the threshold (40.2% vs 38.3%). Additionally, for both thresholds, the proportion of deaths due to respiratory diseases was greater in those whose adherence was lower than the threshold compared to those who exceeded it (50%: 21.7% vs 16.6%, 80%: 19.3% vs 16.7%).

5.6.1 Association Between Statin Adherence and Recurrent Hospital Admissions for Myocardial Infarctions

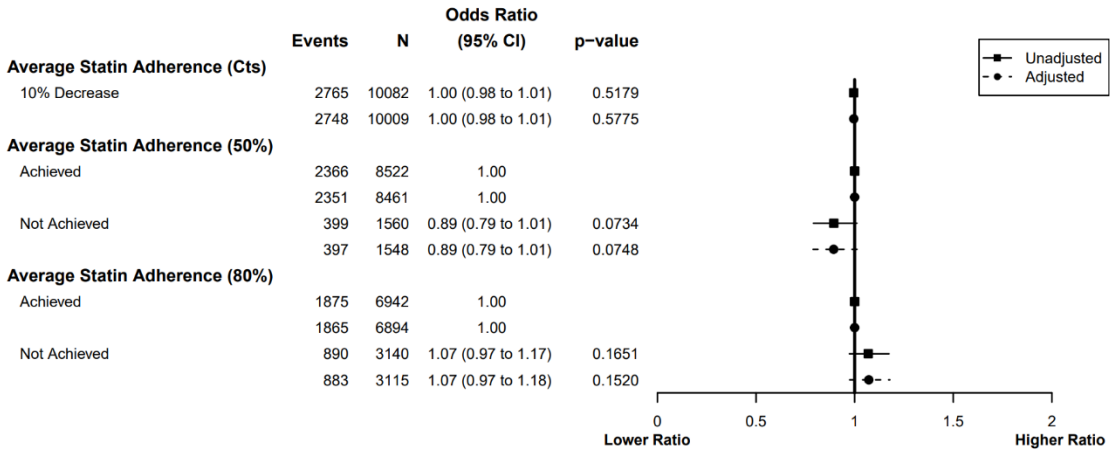


Figure 5.10: Odds of Further Hospitalised MIs by Average Statin Adherence During Follow Up (unadjusted and adjusted for age, sex, year of MI, and SIMD quintile). Achieving an adherence threshold defined as an average adherence, measured using MPR, with thresholds ≥ 50 or 80%. Average adherence calculated for patients with at least one full year of follow up. Patients only included in analyses where all information needed was available.

The associations between average statin adherence and further hospital admissions for MIs, when modelled using logistic regression, were all non-significant at the 5% level (Figure 5.10). Indeed, when modelled continuously, a

10% decrease in average statin adherence was associated with no difference in the odds of experiencing a further MI admission, with an estimated odds ratio of one. When the cut-offs of 50% and 80% were used, the odds ratios reflected the inconsistencies observed in Table 5.10, with adherence <50% associated with an 11% decrease in odds of a further MI admission, and adherence <80% associated with a 7% increase in odds. The differences in the directions with these cut-offs may be suggestive that the relationship between adherence and further MI admissions is not linear, but equally, as neither reached significance, this could have occurred randomly. Additionally, adjusting for age, sex, deprivation and year of MI did not alter the lack of associations observed between adherence and further MI admissions in follow up. All associations remained consistent when incomplete years of follow up were included in the average adherence calculation (results not presented).

5.6.2 Association Between Statin Adherence and Mortality

The Kaplan-Meier plots by average adherence showed that those whose average adherence was below 50% or 80% have lower overall and circulatory survival probabilities than those whose adherence was above these thresholds (Figure 5.11 and Figure 5.12). For both thresholds and causes, differences appear soon after the first year of follow up and are maintained throughout, although given there were fewer circulatory deaths, the differences are less pronounced in this instance. For all-cause mortality, the similar shape of the two graphs is also reflected in the estimates of the lower survival time quartiles. For patients whose average adherence was ≥50%, the lower quartile was 5.5 years (95% CI: 5.3–5.8 years), whilst for adherence ≥80%, the lower quartile was 5.8 years (95% CI: 5.6–6.1 years). The estimated lower quartiles were around two years earlier for those whose adherence was below the thresholds. The lower quartiles were 3.6 years (95% CI: 3.2–4.0 years) and 3.8 years (95% CI: 3.6–4.1 years), for <50% and <80% respectively. There was an insufficient number of events to estimate the median or upper quartile survival times, and no quartiles could be estimated for circulatory causes of mortality.

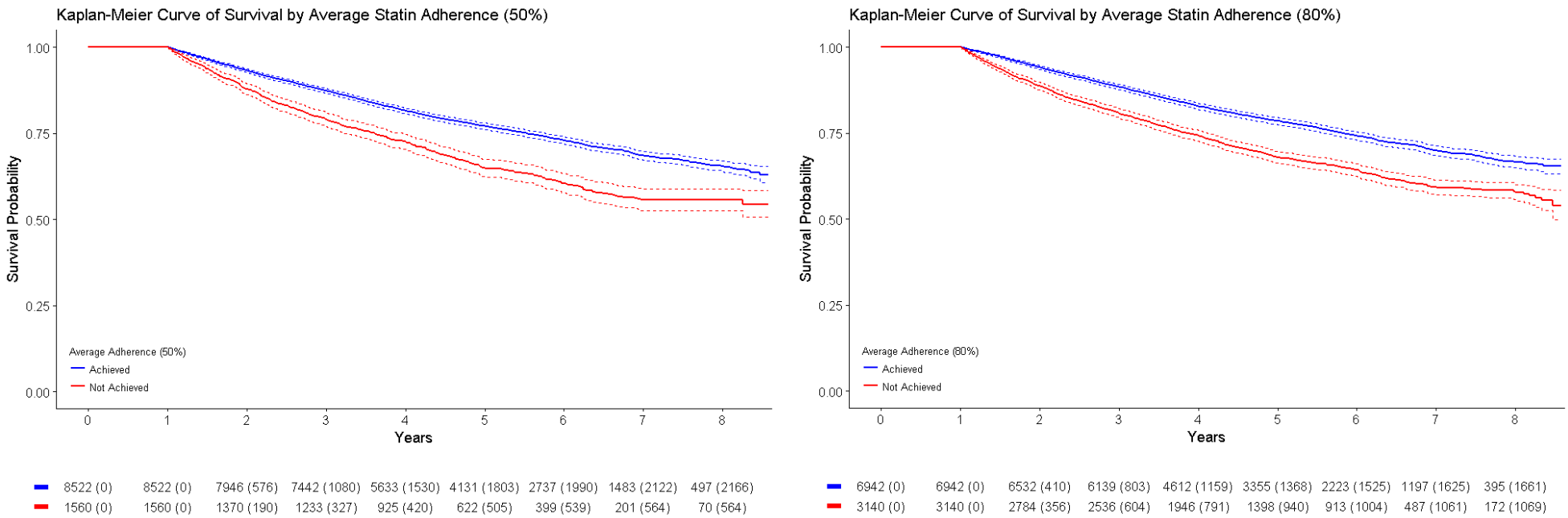
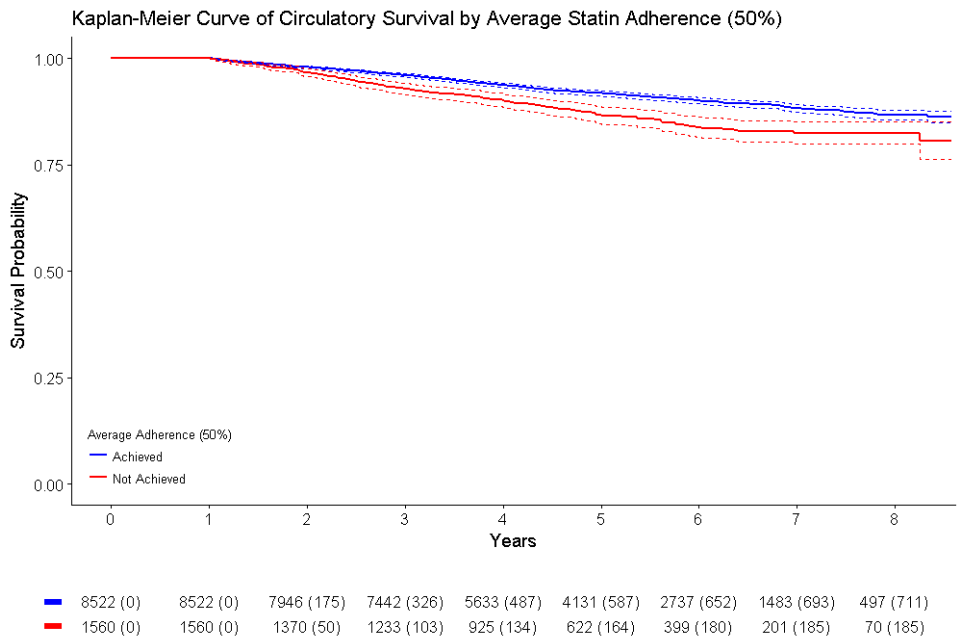


Figure 5.11: Kaplan-Meier Curves of Overall Mortality by Statin Adherence Thresholds (with 95% Confidence Intervals). Numbers are Number at Risk (Number of Events). Patients classified by average statin adherence during follow up, measured using MPR, with thresholds ≥ 50 or 80%. Average adherence calculated for patients with at least one full year of follow up. Patients only included in analyses where information was available.



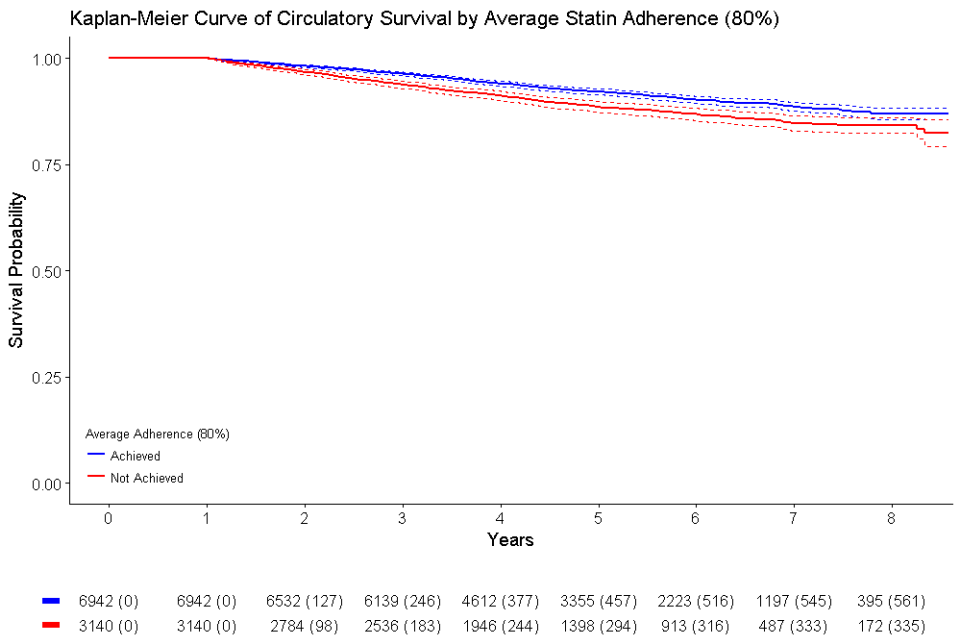


Figure 5.12: Kaplan-Meier Curves of Circulatory Mortality by Statin Adherence Thresholds (with 95% Confidence Intervals). Numbers are Number at Risk (Number of Events). Circulatory cause of death defined by ICD10 I00-I99 as underlying cause. Patients classified by average statin adherence during follow up, measured using MPR, with thresholds ≥ 50 or 80%. Average adherence calculated for patients with at least one full year of follow up. Patients only included in analyses where information was available.

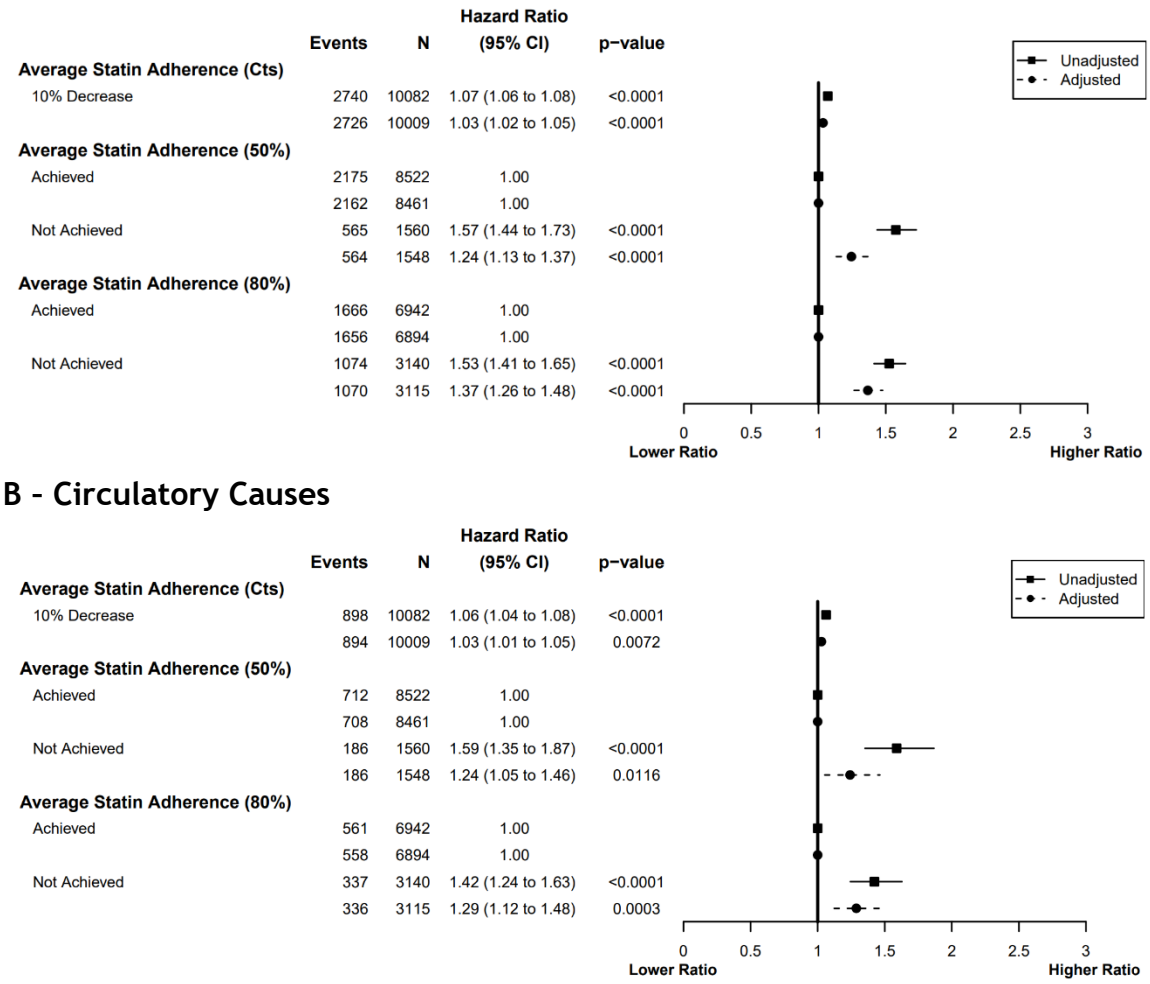


Figure 5.13: Hazard Ratios of Mortality by Average Statin Adherence (unadjusted and adjusted for age, sex, year of MI, and SIMD quintile). A = all-cause mortality; B = circulatory cause of death (defined by ICD10 I00-I99 as the underlying cause). Achieving an adherence threshold defined as an average adherence, measured using MPR, with thresholds ≥ 50 or 80%. Average adherence calculated for patients with at least one full year of follow up. Patients only included in analyses where all information needed was available.

These differences are also reflected in the estimated hazard ratios for all-cause and circulatory mortality from the Cox regression (Figure 5.13). Statistically significant increases were observed in the hazards of all-cause and circulatory mortality for those not meeting adherence thresholds, both in unadjusted analyses and when adjusted for age, sex, deprivation, and year of MI. In unadjusted analyses, average adherence <50% was associated with a hazard of all-cause mortality 1.6 times that of average adherence ≥50%, with estimates similar for circulatory causes. Patients with an average adherence <80% had a hazard of all-cause mortality 1.5 times that of those who exceeded this threshold, and 1.4 times for circulatory causes. Following adjustment, the hazard ratio slightly attenuated for the 80% threshold to a hazard ratio of 1.4 and 1.3 for all-cause and circulatory mortality respectively. Additionally,

adjustment reduced the hazard ratios to 1.2 for both causes when the 50% threshold was considered. On a continuous scale, a 10% decrease in a patient's average adherence was associated with a 7% and 6% increase in the risk of all-cause and circulatory mortality respectively, which reduced to 3% in the adjusted analyses. However, despite this comparatively small increase in risk, both results were statistically significant at the 5% level. Full details of these models and plots of the Schoenfeld residuals are available in Appendix L. Finally, the inclusion of incomplete years of follow up in the average adherence calculation resulted in marginally larger effect sizes in all associations with mortality and cardiovascular mortality, especially following adjustment (results not presented).

5.7 Association Between Adherence and Achievement of Lipid Targets

Finally, an analysis of the association between average adherence and achieving lipid targets was performed. An initial cross-tabulation in Chapter 4.5.4 (in Table 4.10) showed that those with higher average statin adherence were more likely to have achieved a lipid target during follow up than those who did not. This is confirmed with the odds ratios shown in Figure 5.14, which show that those who had lower adherence had greater associated odds of not achieving both lipid targets individually, with the greatest odds observed when the 50% threshold was used. However, the reported odds ratios were also statistically significant in the analyses when the 80% threshold, and when continuous adherence was considered. Full details of the logistic regression models are included in Appendix M (non-HDL) and Appendix N (LDL).

For the non-HDL target (Figure 5.14a), those with adherence <50% had 4.8 times greater odds of not achieving the recommended reduction than those with adherence ≥50%, which slightly reduced to 4.5 when adjusted for age, sex, deprivation, and year of MI. Meanwhile, when the 80% threshold was used, patients had 3 times higher odds of not achieving the target if their adherence was lower, with adjustment for other factors not impacting on this estimate. Adjusting for these potential confounders also had limited effect when adherence was measured continuously, with both models finding that a 10% decrease in statin adherence was associated with a 20% increase in odds of not

achieving a 40% reduction in non-HDL levels. These associated increased odds were slightly lower when incomplete years of follow up were included in the average adherence calculation, although the overall conclusions remained the same (results not presented).

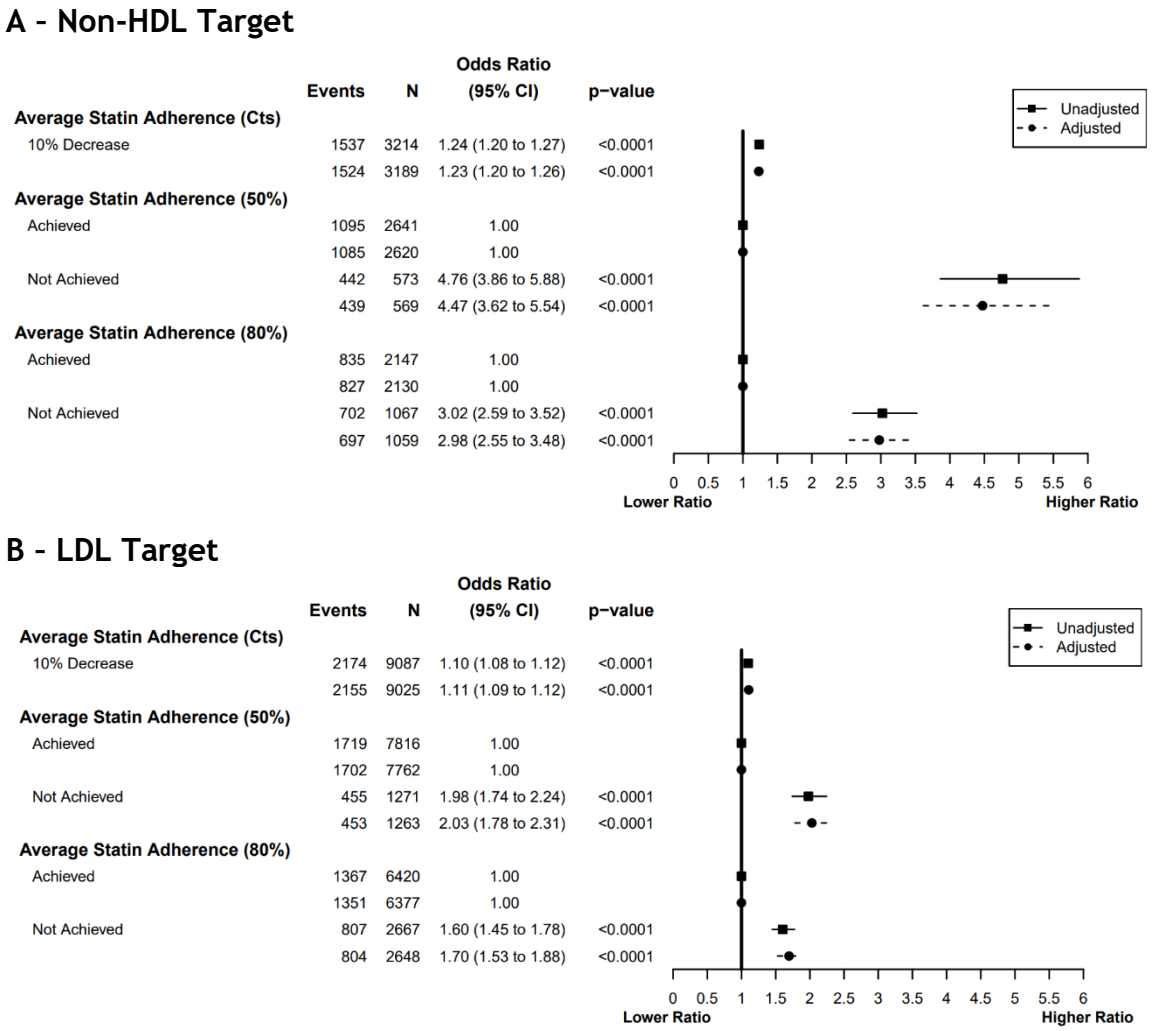


Figure 5.14: Odds Ratios of Not Achieving Lipid Targets During Follow Up (unadjusted and adjusted for age, sex, year of MI, and SIMD quintile). A = Odds of Not Achieving Non-HDL Target; B = Odds of Not Achieving LDL Target. Non-HDL target is ≥40% non-HDL reduction from a pre-MI baseline (NICE, 2014) **and the LDL target is ≤1.8mmol/l** (Piepoli *et al.*, 2016)**. Patients were classified as meeting a target if they achieved the target in any year of their follow up, and not meeting a target if they did not achieve it in every year of their follow up. Achieving an adherence threshold was defined as an average adherence, measured using MPR, ≥50% or ≥80%. Average adherence was calculated for patients with at least one full year of follow up. Patients only included in analyses where all information needed was available.**

The associated odds ratios were, overall, lower for the odds of not meeting the LDL target than they were observed for the non-HDL target (Figure 5.14b). As before, the 50% threshold yielded the greatest increase in odds, with adherence below this resulting in double the odds of not achieving LDL≤1.8mmol/l, compared to average adherence ≥50%. When 80% was used, the associated odds

of not achieving this target were 1.6 times that of those with adherence ≥80%. Adjustment for confounders did not alter the odds ratio at the 50% level but did result in a minor increase of the odds ratio to 1.7 when 80% was used. Considering age, sex, deprivation, and year of MI did not alter the model when adherence was used as a continuous variable, where a 10% decrease in average statin adherence was associated with 10% higher odds of not achieving the LDL target, in both the unadjusted and adjusted models. Utilising incomplete years of follow up in the adherence calculation, as seen with the non-HDL target, resulted in slightly lower associated odds ratios, but the overall conclusions remained unaltered (results not presented).

5.8 Discussion

5.8.1 Associations Between Lipid Targets, Average Adherence, Recurrent Myocardial Infarctions, and Mortality

In this Scottish secondary prevention population, 51% achieved ≥40% reduction in non-HDL from a pre-MI baseline, 76% achieved LDL ≤1.8mmol/l, and 85% and 69% had an average statin adherence above 50% and 80% respectively. Furthermore, one third of patients had died and just over a quarter experienced a further MI requiring hospitalisation during a median follow up of 4.5 years. Both lower average statin adherence, and failure to achieve guideline-recommended lipid targets, were strongly associated with a higher risk of all-cause mortality, as well as being significantly associated with each other. In particular, when adjusted for age, sex, deprivation, and year of MI, those not achieving 80% adherence were 40% more likely to die, and those not achieving the non-HDL target were 2-fold more likely to die, with similar patterns observed when circulatory causes were considered.

Therefore, in addition to clearly illustrating that lipid levels and statin adherence are important predictors of death, these results confirm and expand on previous data. The length of follow up, which commenced on the date of MI, builds on previous research which has often focussed on the initial year (or two years) following an event or statin initiation. Instead, this analysis allowed for a longer-term average statin adherence to be estimated, and a larger period for target lipid levels to have been achieved. Furthermore, the associations with

mortality and recurrent MI admissions were also investigated and compared between continuous adherence and two different adherence thresholds, and between two lipid targets from national and international guidelines. Additionally, they also suggest that there may be a cohort of patients where personalised testing regimens would be advantageous but highlight that the identification of these patients needs to be accurate.

As expected, this cohort had a higher proportion of patients dying due to CVD than in the general population in Scotland in 2017, where CVD accounted for 26% of deaths (vs 42% in this cohort). Among those dying of CVD, deaths due to ischaemic heart disease in this cohort also accounted for a higher proportion in this population than in the general population (69% vs 45% respectively) (National Records of Scotland, 2017). However, given the moderate degree of missingness for cause of death (as discussed in Chapter 3.2.3), these figures for death due to CVD may be an underestimate of the true percentage, as the cause was conservatively assumed not to be as a result of CVD for those where this was unknown. Additionally, as missingness is possibly due to systematic differences within the population, such as patient frailty following their baseline MI leading to their relocation, some subchapters within deaths due to CVD may also be under- or over-represented.

Nevertheless, for both further MI admissions and deaths, the highest numbers of these events were reported in the first year of follow up. Indeed, 75% of patients who experienced at least one further MI during follow up experienced one during their first year, and 27% of deaths that occurred during follow up were within one year. These proportions were fairly similar to the rate of cardiovascular events reported in those with established disease in a CPRD cohort, with the decreasing pattern throughout follow up also present (Khunti *et al.*, 2018). Nonetheless, these numbers were higher than those reported within clinical trials for statins (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010), although this is likely to be due to lower statin adherence outside of trial settings and differences between trial populations and these observational cohorts.

There were no significant associations between average statin adherence nor the achievement of lipid targets with experiencing further MI admissions. This

second observation is consistent with findings from a cohort of 25,000 patients in Sweden where post MI LDL demonstrated poor predictive performance for recurrent ASCVD events (Ohm *et al.*, 2019). Similar findings were also reported in a Finnish cohort (n=25,000), where despite LDL reaching statistical significance for its association with further cardiovascular events, its effect size was small following a multivariable adjustment, with an increased hazard ratio of 3%. However, within this study, statin non-adherence did demonstrate a stronger, and intensity-dependent, association with further events, contrary to findings in this analysis (Lassenius *et al.*, 2020). One explanation for this could be that the average statin adherence variables and those used for determining the achievement of lipid targets, captured information from before and after the further hospital admissions for MIs were experienced. With many experiencing these early on in their follow up, however, much of the data used to derive these variables will likely reflect a patient's behaviour following multiple MIs, rather than the intervening time between them. For example, patients may initially increase their statin adherence after they have experienced an event (Lassenius *et al.*, 2020) or have only achieved lipid targets after a further MI admission. This latter possibility has been shown in CPRD data by Danese et al (2017), although a comparatively small increase (5%) in the proportion of patients meeting the target was observed (Danese *et al.*, 2017). An increase in treatment intensity could also have occurred, which may increase the likelihood of patients achieving target values. In the same CPRD cohort, this occurred in 7% of patients after their second cardiovascular event (Danese *et al.*, 2017).

Confounding comorbidities could also mask any association. Adherence in secondary prevention populations is higher in those with a greater number of comorbidities (Vupputuri *et al.*, 2016; Danese *et al.*, 2017), which may also increase the risk of further events. Indeed, within this population, those experiencing further events were more likely to have a diagnosis of diabetes or have experienced a prior MI, both of which have been associated with higher adherence (Danese *et al.*, 2017; Hope *et al.*, 2019).

This latter confounding problem could also have occurred in the mortality analysis, although the timing of the achievement of lipid targets or changes in

adherence cannot be disputed. Additionally, in the adherence models, only those who survived at least one year of follow up were included which reduced the potential for reverse causality, although sensitivity analyses including incomplete years of follow up did not yield substantially different results. However, those who died also had a shorter duration of follow up and therefore had reduced opportunity to achieve lipid targets and to change their adherence behaviour.

Nonetheless, the significant and consistent associations reported with mortality are similar to findings observed in other cohorts (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010; Xu *et al.*, 2017; Khunti *et al.*, 2018). In an analysis of high-risk patients in CPRD, adherence less than 80% was associated with 50% greater hazard of all cardiovascular events following adjustment for demographic factors and co-morbidities (Khunti *et al.*, 2018). This is slightly higher than the observed hazard ratios of 1.4 in this population for all-cause and circulatory mortality, but this CPRD cohort considered all cardiovascular events and not mortality outcomes only. Furthermore, whilst the use of lipid targets is controversial (Leibowitz *et al.*, 2017; Pallazola *et al.*, 2018), and there is a lack of robust evidence to support their use in guidelines (as documented in Chapter 2 (Brown, Welsh and Logue, 2020)), lowering LDL has consistently been shown to significantly reduce mortality within those with established disease (Cholesterol Treatment Trialists' (CTT) Collaboration, 2010).

The significant associations observed between statin adherence and the achievement of lipid targets also expand on associations with mortality. This was expected given the mechanism of action of statins and the extensive literature supporting their use (as outlined in Chapter 1.3.1). These associations have also been observed in similar observational cohorts. For example, in 1,000 secondary prevention patients in Georgia, USA, those with adherence <50%, were at 88% greater risk of not achieving a 30% reduction in LDL (Vupputuri *et al.*, 2016). However, in Danese et al's (2017) analysis, in the first year following a cardiovascular event, the proportion of patients achieving at least 80% statin adherence was estimated to be approximately 70%, despite much higher numbers failing to meet the LDL target. This suggested that whilst adherence may be associated with cholesterol levels, it is not the only contributing factor

(Danese *et al.*, 2017). This pattern is also observed within this cohort, and there is little evidence of any strong confounding factors.

Additionally, the associations with the non-HDL target were much stronger than the associations with the LDL target. This is likely to have arisen due to the differing nature of the targets e.g. percentage change (non-HDL) vs absolute value (LDL), rather than the lipid profile components themselves. For example, associations between absolute values of non-HDL and LDL with a composite endpoint of further non-fatal and fatal cardiovascular endpoints have been shown to be approximately equivalent (Welsh *et al.*, 2019). Nonetheless, the use of percentage change targets within secondary prevention populations could prove problematic, with many patients unlikely to have a suitable baseline for the calculation to be performed. This was evident in this population where percentage change could only be calculated for 30% of the cohort.

5.8.2 Strengths and Limitations

This is a contemporary, large data set from a real-world population. Generalisability of these findings, particularly to the Scottish population, is therefore good. One key limitation is the lack of assessment or adjustment for the consistency of meeting lipid targets in the models, which is likely to be an important factor in their association with further events. Similarly, the use of average adherence does not reflect fluctuations in adherence throughout follow-up, although there is some evidence to suggest that this may be minimal after the first year (Nordstrom *et al.*, 2015). Nonetheless, when prediction models of adherence and lipid levels are developed in the subsequent chapter, these models will need to allow for time-varying covariates. The measure of adherence used is arguably stronger evidence of medication taking than data from CPRD, due to prescriptions being encashed.

Another limitation is the use of logistic regression to investigate associations with further hospital MI admissions, where statin adherence and lipid target variables involved components from before and after the admissions took place. To further investigate these associations, adherence and target variables using only data collected before the event could be derived. However, this time window is likely to be small for many patients as most further admissions

occurred during the first year of follow up, and therefore the potential for reverse causality is increased. This difficulty would also be faced when considering a time-to-first-event analysis, where competing risks, such as mortality, would additionally need to be accounted for. This could be conducted using either cause-specific, which estimates the instantaneous rate of a recurrent MI in event-free patients, or subdistribution hazards, which estimates the instantaneous rate of a recurrent MI in patients who are event-free or have died. In this instance, the former method would be preferable given the aetiological nature of the question being investigated (Austin, Lee and Fine, 2016). Finally, a recurrent analysis could be considered, although given 75% of patients with a further admission experienced only one event, this is unlikely to substantially alter any observed associations.

The models presented in this chapter showed both unadjusted and adjusted analyses, and in all cases, the adjustment did not result in substantial alterations to the estimated effect size or its significance. However, only a few demographic variables were accounted for, with no interactions between them considered. For example, significant interactions between sex and other covariates were likely to be present. Females typically experience MIs at a later age compared to males, and as cholesterol also increases with age, this further increases the risk of mortality (Jousilahti *et al.*, 1999; Mosca, Barrett-Connor and Kass Wenger, 2011). It has also been shown that females are less likely to be prescribed statins following an MI, reducing their adherence, and increasing their cardiovascular risk (Eindhoven *et al.*, 2018). Consequently, future analysis should consider stratifying models presented by sex. Furthermore, the variables considered are not comprehensive, and other confounders, such as other comorbidities, are likely to remain, as discussed earlier. As a result of this, the methodology employed in this chapter, and the observational nature of these data, no causal inferences regarding these associations are drawn.

Nevertheless, the strength of the associations between adherence and lipid targets, and them both separately with mortality in this post MI population are consistent with, and expand upon, current literature, and therefore demonstrate the external validity of this cohort.

5.8.3 Conclusions

This analysis has highlighted that a significant proportion of the population achieves lipid targets and reasonable average statin adherence and that both are associated with reduced mortality. No significant associations were observed with the likelihood of experiencing further hospitalised MIs, but this is likely due to a combination of confounding and the definitions of lipid target achievement and the use of average adherence. Despite this, the associations observed with mortality and between lipids and adherence are consistent and in line with the evidence base for statins. We also broadly validate the existing Scottish SIGN non-HDL guidance as clinically relevant, however low coverage for percentage change variables in secondary prevention populations may prevent its practical use in these settings.

These results also have implications for the potential use of personalised plasma lipid testing schedules within this population. Specifically, these associations highlight the importance of the accurate identification of patients for whom less frequent monitoring is considered. If reduced testing schedules are to be effective and safe, then this successful identification is crucial to avoid the risk of further MI admissions or deaths that could have been prevented through an annual test.

Chapter 6 Identifying Patients for Reduced Lipid Monitoring

6.1 Introduction

In an earlier chapter (Chapter 2), a systematic review of clinical guidelines for the established disease population found that many guidelines (n=17/22) specified a lipid target for these patients, but that evidence to support the use of these was limited (Brown, Welsh and Logue, 2020). Furthermore, when it came to monitoring these patients, half of the included guidelines gave recommendations for their long term follow up, although there was often no evidence to support these suggestions beyond expert opinion. In Scotland, current and previous guidelines have both recommended that annual reviews are conducted as part of good practice for those at high risk of CVD (SIGN, 2007, 2017).

However, the monitoring of chronic diseases is a significant and growing burden to primary care, as discussed in Chapter 1.6.1, especially as the population continues to grow and age (Thompson and Walter, 2016). The number of laboratory tests requested has increased each year since 2000 and cannot be solely attributed to increased patient numbers (O'Sullivan *et al.*, 2018). This has also led to a greater workload for primary care, with a survey of GPs reporting an increase in time taken for both consultations and the associated administration outside of it (Hobbs *et al.*, 2016; Thompson and Walter, 2016).

The purpose of the lipid test at a review is also unclear; both addressing patient's adherence and monitoring risk factors are listed as possible discussion points for the annual review in current Scottish guidelines (SIGN, 2017). For adherence, several analyses, including those in CPRD, have demonstrated reasonable, and sustained, statin adherence in both primary and secondary prevention populations (Nordstrom *et al.*, 2015; Lavikainen *et al.*, 2016; Danese *et al.*, 2017). Furthermore, a lower, but reasonable, proportion of patients with established disease has been shown to achieve lipid targets in the years immediately following an event (Nordstrom *et al.*, 2015; Danese *et al.*, 2017). Both of these findings have also been demonstrated in this post MI population of

NHS GGC in Chapter 4, with adherence rates and achievement of the LDL target of ≤1.8mmol/l slightly higher than those in the literature.

Therefore, with reasonable numbers of patients achieving lipid targets and acceptable adherence, it is plausible that there is a subgroup within this population whose lipid test as part of their annual review could be considered unnecessary. However, such an approach would need to accurately identify these patients to be considered safe, due to the elevated risks of mortality in those non-adherent or failing to meet lipid targets, as shown in Chapter 5. Such an approach has been attempted before in the context of other screening programmes (as described in Chapter 1.6.3). This has been most successful in diabetic retinopathy where one risk score has been tested in an RCT which reported no significantly increased risk as a consequence of reduced frequency of screening and a reduction of 40% in required appointments (Broadbent *et al.*, 2020).

However, for a similar approach to be considered in cardiovascular prevention, patients who could safely receive fewer lipid tests need to be accurately identified. Therefore, this chapter has two aims. Firstly, it seeks to use a post MI cohort to identify associations between demographic factors, previous adherence, lipid results, and testing frequency, with subsequent statin non-adherence and failure to achieve cholesterol targets. Second, it seeks to develop an algorithm that will identify patients within this cohort for whom the results of an annual lipid test are unlikely to change clinical advice or decisions, and therefore it could be omitted.

6.2 Methods

This chapter uses data previously derived and described in Chapter 4. In this chapter, identical methods were used for both the prediction of a patient's adherence and predicted achievement of cholesterol targets in the next year. However, only the LDL target of ≤1.8mmol/l from the 2016 ESC guidelines was considered in the case of the latter outcome (Piepoli *et al.*, 2016). Although the target of ≥40% reduction in non-HDL recommended by NICE (NICE, 2014) showed stronger associations with outcomes than the 2016 ESC target in Chapter 5, percentage change variables are likely to prove difficult to implement in

secondary prevention populations with many patients lacking a suitable baseline to facilitate their calculations. This was the case here, with just 30% of this cohort having at least one percentage change result available. Therefore, as the LDL target relies on its absolute value and has greater coverage, only this target was used throughout this chapter. Both 50% and 80% thresholds were used in the estimated statin adherence analyses.

Descriptive demographics are presented for all patients using the methods outlined in Chapter 5.2.1. Demographics were split by the consistency of meeting the adherence threshold or LDL target, although in the latter patients were only included if they had at least one LDL test result available post MI. For all subsequent analyses, patients were excluded if they were never prescribed a statin at any point in the duration of their follow up as it is likely these individuals are statin intolerant and will therefore have been monitored differently to the majority of the population.

In addition to descriptive analysis, Kaplan-Meier plots and Cox regression models were fitted to determine factors associated with time to (first) non-adherence or non-target LDL. In these models, some time-varying covariates were considered, thus building on the methods outlined in Chapters 5.2.4 and 5.2.5 and are described below. Using the factors identified, latent class analysis (LCA) was used to identify groups of patients and their likelihood of adherence thresholds or LDL targets being met in each year of follow up. The diagnostic accuracy of these was assessed using the predicted class assignments from the model, and these predictions were then bootstrapped with one time window per patient to assess the robustness of the analysis to a violation of independence (due to multiple time windows included per patient).

Cox regression models (as described in Chapter 5.2.5) were then used for models in years 2 and 3 of follow up separately to investigate any association between the predicted classes and mortality. These years were selected due to the higher number of patients with data available than for other periods of follow up. However, due to the need for the previous year's information, year 1 could not be considered for these models where the mortality rate was highest. Furthermore, a longer-term and time-varying model could not be implemented in this instance due to immortality bias (arising from the need to have survived

Year 1 to have a predicted class), as well as the possible impact of legacy effects for those changing between assigned classes.

Finally, latent class growth analysis (LCGA) was conducted using baseline covariates identified in the earlier Cox models to identify trajectories in proportions of patients who are adherent or on target in each year of follow up. Kaplan-Meier plots and Cox regression models were used to examine the associations between the predicted classes and mortality.

6.2.1 Cox Regression with Time-Varying Covariates

Cox regression models with time-independent covariates were first implemented in Chapter 5.2.5 and seek to compare the survival distributions of patients by certain risk factors (Xu, 2020). These models can be extended to incorporate time-dependent covariates, that is, the presence or the measurement of risk factors may alter during a patient's follow up (Fisher and Lin, 1999; Zhang *et al.*, 2018; Xu, 2020). Consequently, models with time-varying covariates instead compare the risks of an outcome by covariates at the time of each outcome (Xu, 2020).

The structure of such models is similar to that of the conventional Cox model but includes the addition of a time-interaction function within the main hazard equation, as shown in Equation 6.1. As a result, the hazard function at time t is only dependent on the covariate values at that particular time, and the corresponding hazard ratio is not constant and varies over time (Zhang *et al.*, 2018; Xu, 2020). The estimation of the model's coefficients is then the same as for a time-independent model using partial likelihood, with the only difference being that the set of covariates changes at each event time (Xu, 2020).

$$Consider\ the\ convention\ Cox\ regression\ hazard:$$
$$h(t) = h_0(t) * \exp(\beta X),$$
$$where\ X\ is\ a\ vector\ of\ time\ independent\ covariates\ and\ \beta\ is\ a\ vector\ of\ coefficients.$$

$$Now, allow\ time\ dependent\ covariates\ by\ altering\ h(t)\ to:$$
$$h(t) = h_0(t) * \exp(\beta' X(t)),$$
$$where\ \beta\ is\ a\ vector\ of\ coeffients, and\ X(t)\ is\ a\ vector\ of\ time\ varying\ covariates.$$

Equation 6.1: Cox Regression Model with Time-Varying Covariates (Zhang *et al.*, 2018; Xu, 2020)

The ‘survival’ package in R was used to produce time-varying Cox regression models in this chapter for time to (first) non-adherence (for both the 50% and 80% thresholds) and non-target LDL (Therneau, 2020), with patients assumed to be adherent or on target at time 0. To assess the impact of this assumption, sensitivity analyses were also performed starting at the time window where the patient was first adherent or had LDL on target. The time-varying covariates considered were whether the patient experienced a further MI in the previous year of follow up and whether the patient received a diagnosis of diabetes. Time-independent covariates considered were age at baseline MI, sex, deprivation quintile, whether the patient had experienced an MI before their baseline, and the year of MI. Further sensitivity analyses were considered whereby year of MI was removed as a covariate in the final model, as, whilst this variable may improve the fit, its use within a clinical setting is unclear unless significant changes in the clinical management of these patients had occurred at a particular time point. Age was categorised into <50, 50-<60, 60-<70, 70-<80, ≥80 years to allow for its inclusion in the subsequent latent class analysis (Chapter 6.2.2).

In the final models, the number of lipid tests conducted in the previous year was also added as a time-varying covariate, with the value set to the recommended annual 1 test for Year 1. However, changing this assumption to 0 tests and >1 test did not alter the results, as changes between time 0 and time 1 are cancelled out during the partial likelihood calculation (Xu, 2020). Before models were run, the ‘tmerge’ function was used to divide a patient’s follow up into multiple records with a new record created each time a covariate changed (Therneau, 2020; Therneau, Crowson and Atkinson, 2020). The optimal model for each outcome was determined through best subsets regression using Akaike’s Information Criteria (AIC) to determine the best fit. AIC was used instead of hypothesis testing for the selection criteria, as the aim was to build a prediction model and it is asymptomatically equivalent to leave-one-out cross-validation prediction error (Stone, 1977). The results of the final models were plotted using the ‘forestplot’ package in R (Gordon and Lumley, 2020).

6.2.2 Latent Class Analysis Adjusting for Covariates

LCA is a type of mixture model applied to data whose variables are categorical only, either dichotomous or polytomous (Agresti, 2012; Masyn, 2013; Oberski, 2015). The nature of these variables leads to a distribution that is also qualitative, and the model in analyses clustering and patterns within the data arising from the generation of multi-way tables (Agresti, 2012; Linzer and Lewis, 2016). These patterns are then considered to be indicators for an unobserved variable which would divide the data into specific subgroups. This variable, which is also categorical, is called a latent class (Agresti, 2012; Masyn, 2013; Oberski, 2015). Figure 6.1 illustrates a generic path diagram for such a model (Oberski, 2015).

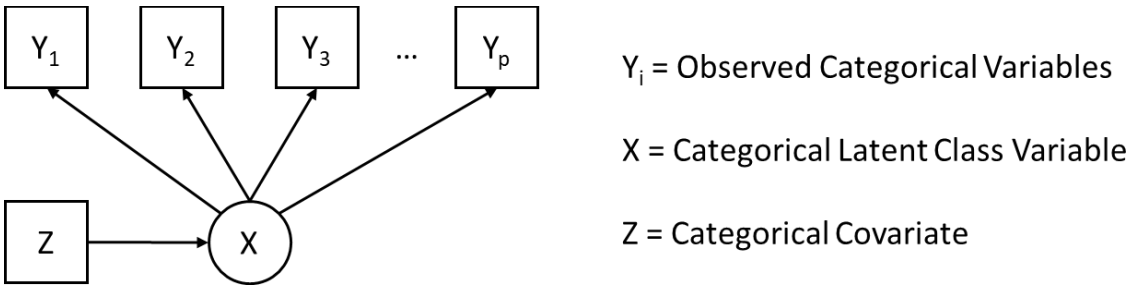


Figure 6.1: Path Diagram for an LCA Model with *p* observed categorical variables (Y_i), one latent class variable, and one categorical covariate (Oberski, 2015)**.**

For the main part of the LCA model, there are two key assumptions. The first of these is the mixture assumption, which specifies that a given combination of the observed variables is a joint distribution of the class-specific distributions, as shown in Equation 6.2 (Agresti, 2012; Oberski, 2015).

$$Let\ \mathbb{P}(y_1, y_2, \dots y_p) = \mathbb{P}(Y_1 = y_1, Y_2 = y_2, \dots Y_p = y_p),$$
$$\mathbb{P}(y_1, y_2, \dots y_p) = \sum_{k=1}^{K} \mathbb{P}(X = k) * \mathbb{P}(y_1, y_2, \dots y_p | X = k)$$

Equation 6.2: Assumed Joint Distribution Mixture of Class-Specific Distributions for an LCA Model with p observed parameters and a latent class variable with K classes (Agresti, 2012; Oberski, 2015)**.**

The second assumption is that of conditional or local independence of the observed variables (Y_i) included in the model within the classes, i.e. within a particular latent class (k), these observations are assumed to be independent, as shown in Equation 6.3 (Agresti, 2012; Oberski, 2015).

$For\ the\ kth\ class, i.e. X = k,$

$$\mathbb{P}(y_1, y_2, \dots y_p | X = k) = \prod_{i=1}^{p} \mathbb{P}(y_i | X = k)$$

Equation 6.3: Assumed Local Independence of Observed Variables in an LCA Model with p observed parameters for the kth class of K latent classes (Agresti, 2012; Oberski, 2015)**.**

$$\mathbb{P}(y_1, y_2, \dots y_p) = \sum_{k=1}^{K} \mathbb{P}(X = k) * \prod_{i=1}^{p} \mathbb{P}(y_i | X = k)$$

Equation 6.4: Latent Class Measurement Model Assuming Local Independence with p observed parameters and K latent classes (Agresti, 2012; Masyn, 2013)**.**

Using this assumption, the earlier joint distribution of the mixture model can be simplified by substituting Equation 6.3 into Equation 6.2, to give the equation given above (Equation 6.4) (Agresti, 2012; Masyn, 2013). This is also known as the latent class measurement model (Masyn, 2013).

Once the LCA model has been fitted, posterior class probabilities can be calculated to assign observations to latent classes. These can be calculated using the parameters from the LCA model, the above equations, and Bayes theorem. Equation 6.5 illustrates how this can be derived for class c (Oberski, 2015). Once these have been calculated for each class, the observation is assigned to the class with the highest probability, in a process known as modal assignment (Oberski, 2015; Linzer and Lewis, 2016).

$$\mathbb{P}(X = c | y_1, y_2, \dots y_p) = \frac{\mathbb{P}(X = c) * \mathbb{P}(y_1, y_2, \dots y_p | X = c)}{\mathbb{P}(y_1, y_2, \dots y_p)}$$

$$= \frac{\mathbb{P}(X = c) * \prod_{i=1}^{p} \mathbb{P}(y_i | X = c)}{\sum_{k=1}^{K} \mathbb{P}(X = k) * \prod_{i=1}^{p} \mathbb{P}(y_i | X = k)}$$

Equation 6.5: Posterior Class Probability of a Latent Class Model with p observed parameters, and K latent classes, for a class c where c≤K (Oberski, 2015)**.**

As suggested earlier in Figure 6.1, LCA models can be extended to adjust for further covariates to ascertain the presence of any association with the latent classes (Bolck, Croon and Hagenaars, 2004; Oberski, 2015; Linzer and Lewis, 2016). However, this should be conducted either simultaneously to the LCA model fitting or in a three-step process. Otherwise, any estimates from the multinomial model generated are highly likely to be biased and standard errors likely to be incorrect (Bolck, Croon and Hagenaars, 2004; Oberski, 2015).

In this chapter, LCA was selected as the method to identify classes of time windows which would be likely to meet adherence thresholds or the LDL target. For the eventual model to have sufficient predictive value, it was likely that higher-order interactions between the covariates would need to be included in a standard regression model, and therefore the use of a clustering-based method would be highly appropriate (Vermunt and Magidson, 2003). Such methods can be split into two types: distance-based and probability-based. Distance-based methods, such as k-means and hierarchical clustering, first create a measure of the differences between individuals and then endeavour to minimise the variability of this within subgroups and maximise the variability between them. Probability-based methods, such as LCA, identify distributions in the data and where individuals fall within them first, before seeking to maximise the variability explained in the simplest terms (Kent, Jensen and Kongsted, 2014). The advantages of this latter approach include greater classification accuracy and more easily interpretable output including the classification probabilities for individuals and parameters that can be used to classify new individuals (Vermunt and Magidson, 2003; Kent, Jensen and Kongsted, 2014). This would be particularly beneficial as the purpose of this work is to ultimately develop a clinical prediction tool for new patients to ascertain the necessity of a lipid test within an annual review. This, combined with the output of a categorical latent variable, which would aid this implementation further, made LCA the preferred method.

LCA models, which adjusted for either achievement of adherence thresholds or the LDL target, were produced using the 'polca' package in R, which performed the adjustment simultaneously (Linzer and Lewis, 2011, 2016). Models were also fitted using available case analysis, with each time window taken as an observation. Variables included in the LCA models were those identified in the time-varying cox models as well as binary variables indicating whether the adherence threshold and LDL target had been met in the previous year. Testing frequency in the previous year was not included. For each outcome, models were run with 2, 3, and 4 classes, replicated 10 times, allowing for 7,000 iterations, with a tolerance for convergence of $1e^{-7}$. Within the ten replications, the optimal model was selected using the Bayesian Information Criterion (BIC) (Oberski, 2015; Linzer and Lewis, 2016). To determine the optimal number of

classes, the diagnostic measures outlined in Chapter 6.2.3 were used, selecting the model with the highest positive predictive value.

The likelihood of a class meeting an adherence threshold or LDL target was determined using the regression coefficient from the resulting adjusted multinomial model. Throughout this chapter, classes have been renumbered to ensure consistency between the models, making class 1 always the most likely to meet the threshold/target.

Two measures were calculated to assess the overall fit of the final models: relative entropy and average classification rates. Relative entropy (Equation 6.6), which ranges from 0 to 1, provides an estimate of the precision of the classification of the time windows fitted using the model. A value close to 0 is suggestive that class assignment is no better than random chance, and values close to 1 indicate high certainty in the predicted classes (Masyn, 2013).

$$Let\ \hat{p}_{ik} = posterior\ probability\ of\ observation\ i\ in\ latent\ class\ k,$$
$$E_K = 1 - \frac{\sum_{i=1}^{n}\sum_{k=1}^{K}[-\hat{p}_{ik} * \log(\hat{p}_{ik})]}{n * \log(K)}$$

Equation 6.6: Relative Entropy for an LCA model for n observations with k classes (Masyn, 2013)

$$Let\ \hat{c}_{modal,i} = modally\ assigned\ class\ for\ observation\ i,$$
$$AvePP_k = mean\{\hat{p}_{ik}, \forall\ i: \hat{c}_{modal,i} = k\}$$

Equation 6.7: Average Posterior Class Probabilities for an LCA model with K classes (Masyn, 2013)

Average posterior class probabilities were also calculated using the formula given above (Equation 6.7), which returns the average probability for those modally assigned to that class. As this is an average of probabilities, these values are also bounded at 0 and 1. However, unlike relative entropy, these look at the certainty of the classification within a specific class rather than looking at the model's certainty in classification as a whole. Nonetheless, higher values suggest greater model fit, with an arbitrary threshold of 0.7 for all classes considered to indicate the accuracy of class assignment is adequate (Masyn, 2013).

6.2.3 Diagnostic Accuracy Measures

As part of the LCA modelling process, classes were categorised based on their likelihood of meeting either the adherence threshold (50% or 80%) or the LDL target depending on the model specified. This likelihood was determined by the regression coefficients as described above, but the accuracy of this was yet to be assessed.

For each time window included, posterior probabilities for the assignment to each class were generated and time windows were assigned to the class which had the highest posterior probability, in a process known as modal assignment (Masyn, 2013). These predicted classes were then cross-tabulated with whether the adherence threshold or LDL target had been achieved, as shown in Table 6.1. For optimal models which contained more than two classes, some rows of these tables were combined to form a 2x2 table so that the diagnostics accuracy measures introduced below could be calculated.

Table 6.1: Example Diagnostic Accuracy Table for LCA Class Assignment

		Adherence/LDL Target Achieved		
		Yes	**No**	**Total**
Class Likelihood of Meeting Adherence/LDL Target	**Most**	A	B	A+B
	Least	C	D	C+D
	Total	A+C	B+D	N

Using this cross-tabulation, the accuracy of using this class assignment to predict whether a patient's adherence would be above a threshold, or their LDL would be ≤1.8mmol/l in a given time window could then be determined. The first measure of this, sensitivity (Equation 6.8a), is the percentage of time windows where the threshold or target was met that were predicted to have achieved it. This differs from the positive predictive value (Equation 6.8c) which is the percentage of time windows where the threshold or target was met within the time windows where it was predicted to be. Both of these focus on the accurate identification of true positives, i.e. time windows who are predicted to meet the threshold or target. In contrast, the remaining measures, specificity and negative predictive value focus on the accurate identification of those who will not meet the threshold or target. Specificity (Equation 6.8b) is the percentage

of time windows where the threshold or target was not met that were predicted to have not achieved it, and negative predictive value (Equation 6.8d) is the percentage of time windows where the threshold or target was not achieved within the time windows where it was predicted not to be (Kirkwood and Sterne, 2003).

$$Sensitivity = \left(\frac{A}{A+C}\right) * 100 \quad \text{(a)}$$

$$Specificity = \left(\frac{D}{B+D}\right) * 100 \quad \text{(b)}$$

$$Positive\ Predictive\ Value = \left(\frac{A}{A+B}\right) * 100 \quad \text{(c)}$$

$$Negative\ Predictive\ Value = \left(\frac{D}{C+D}\right) * 100 \quad \text{(d)}$$

Equation 6.8: Measures of Diagnostic Accuracy (Kirkwood and Sterne, 2003)**. A, B, C, D are values from the appropriate cell in Table 6.1**

6.2.4 Bootstrapping

LCA modelling assumes that each time window included in the dataset is independent (Masyn, 2013). However, this is not the case in this analysis as multiple time windows were included per patient. Therefore, an internal bootstrap of the final model was used to gauge the impact of this violation.

Within a bootstrap, one time window per patient was randomly selected, giving a sample size of 8,626 and 6,274 for the adherence and LDL models respectively. The predicted classes for these selected observations were then cross-tabulated with whether the threshold or target was achieved, as in the main model. From this, the diagnostic accuracy measures, namely: sensitivity, specificity, positive predictive value, and negative predictive value were calculated. Once bootstrapped 5,000 times, the mean-average for each measure was obtained. Confidence intervals for these were calculated using the percentile method, with bounds obtained from the 2.5th and 97.5th percentiles (Helwig, 2017).

However, under this method, if diagnostic accuracy measures were 100% under the main LCA model, no different value will arise during this bootstrapping

method. This occurs because in the cross-tabulation stage the adjacent cell in the table for that particular measure will always be equal to 0, regardless of the sampling, making any percentage calculations 100%.

6.2.5 Latent Class Growth Analysis

As multiple time windows were available per patient, LCGA, a type of growth mixture modelling, was used to examine patient's trajectories in the achievement of adherence thresholds and LDL targets separately. Growth mixture models differ from multilevel random-effects models by allowing for patient variations, thereby considering that subgroups with different trajectories may exist within the single dataset (Jung and Wickrama, 2008; Cécile Proust-Lima, Philipps and Liquet, 2017). In LCGA, this classification of individual trajectories into identifiable subgroups is performed in a similar probability-based approach to that of LCA for the class-membership component of the model and consequently yields similar advantages to those outlined above (Section 6.2.2). Specifically, these models allow for easy interpretation and description of patient trajectories, which would prove beneficial within a clinical setting, and allow subsequent individuals to be classified using their baseline characteristics (Jung and Wickrama, 2008; Kent, Jensen and Kongsted, 2014).

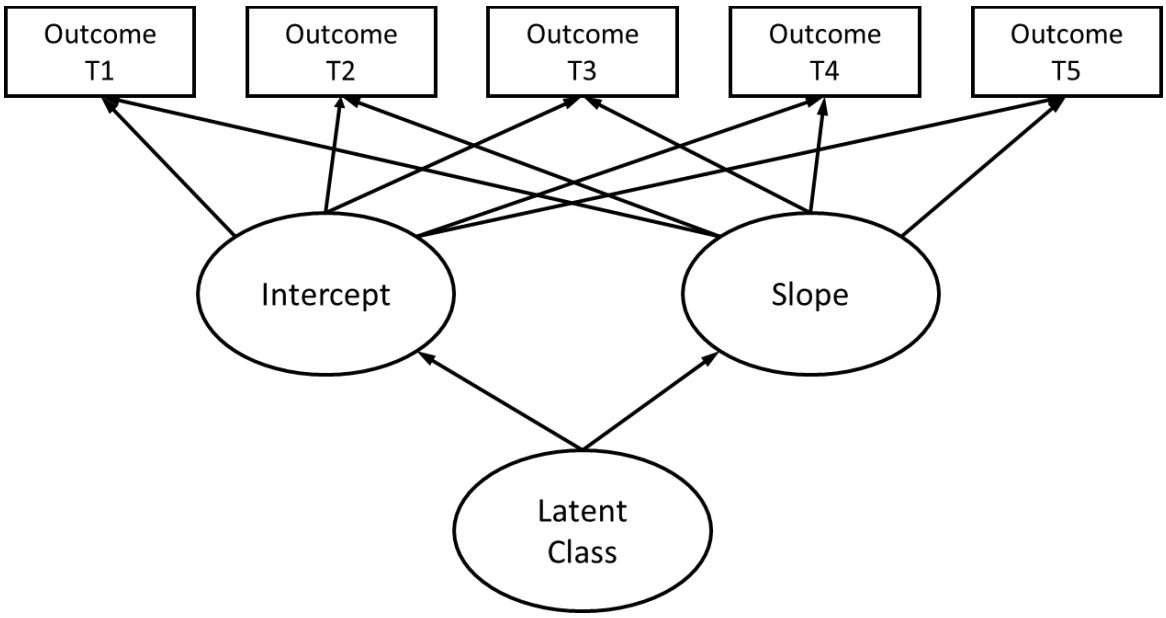


Figure 6.2: Path Diagram for a Latent Class Growth Model (adapted from (Berlin, Parra and Williams, 2014)**).**

LCGA (shown in Figure 6.2) is a special case of these models, where trajectories within such classes are assumed to be homogenous and that, as a result, all variation occurs between subgroups and not within them (Jung and Wickrama, 2008; Berlin, Parra and Williams, 2014; Cécile Proust-Lima, Philipps and Liquet, 2017). This is induced by the assumption that the variance of the slope and intercept coefficients within each of the latent classes is fixed at zero. This, in turn, results in the covariance of these parameters also assumed to be zero, and therefore variation can only occur across the latent classes, allowing groups of patients to be identified, rather than the prediction of individual trajectories. Practically, this can substantially reduce computation time due to the need to estimate fewer parameters and increased likelihood of model convergence (Jung and Wickrama, 2008; Berlin, Parra and Williams, 2014). However, whilst covariates can be included within the class-membership model, only those obtained at baseline that are time-independent should be included (Cécile Proust-Lima, Philipps and Liquet, 2017; Proust-Lima *et al.*, 2020). Additionally, as with all growth mixture models, data is assumed to be missing at random across repeated measures (Cécile Proust-Lima, Philipps and Liquet, 2017).

In this chapter, LCGA was performed using the 'lcmm' package in R using the 'gridsearch' function and the 'lcmm' function with the 'thresholds' link. Models were fitted using available case analysis for the class-membership component, and those with at least one estimated adherence or LDL result, as applicable (Cecile Proust-Lima, Philipps and Liquet, 2017; Proust-Lima *et al.*, 2020). As before, the covariates included within the class-membership model were those ascertained in the earlier time-varying cox models, but only those that were captured at their baseline MI. Models were generated with 2, 3, and 4 classes, replicated 10 times and allowing for 100 iterations with a tolerance for convergence of $1e^{-4}$. Within each set of replications, the optimal model was selected with the best log-likelihood, and the model with the optimal number of classes was determined using BIC (Cecile Proust-Lima, Philipps and Liquet, 2017; Proust-Lima *et al.*, 2020).

Similar to LCA, conditional class probabilities can be obtained, and patients are assigned to classes using modal assignment from the posterior probabilities. Both relative entropy and average posterior probabilities were also calculated and

compared as detailed above (Chapter 6.2.2). Using the predicted class assignments, the percentage of patients who were adherent or at target LDL at each year of follow up could also be calculated and were plotted to indicate class trajectories. Finally, predicted classes were used to compare survival probabilities between the classes in a Kaplan-Meier and an unadjusted Cox regression model.

6.3 Factors Associated with Statin Adherence

6.3.1 Patterns in Statin Adherence (within patients)

Figure 6.3 shows the statin adherence patterns of patients during their follow up, with green indicating that the adherence threshold (50% and 80%) was met in that year of follow up. These graphs suggest that for both adherence thresholds, the majority of patients remained adherent throughout their follow up, with those who were non-adherent, also likely to remain so. With the 50% threshold, after becoming non-adherent, comparatively few became adherent again, with these numbers slightly higher at the 80% threshold. At the 80% threshold, a small number of patients who were non-adherent in the first year became adherent in the second year, with many then remaining so for the remainder of their follow up. However, more than two crossovers between states (adherent/non-adherent) was relatively rare.

Using these adherence patterns, for each threshold, patients were then manually classified as being consistently adherent, consistently non-adherent, or inconsistently adherent, and their demographics are shown in Table 6.2. Using the 50% adherence threshold, 66% of the patients were consistently adherent, 23% were inconsistently adherent, and 10% were never adherent to statins during follow up. In comparison, only 38% were consistently adherently adherent at the 80% threshold, with 44% inconsistent, and 18% never adherent.

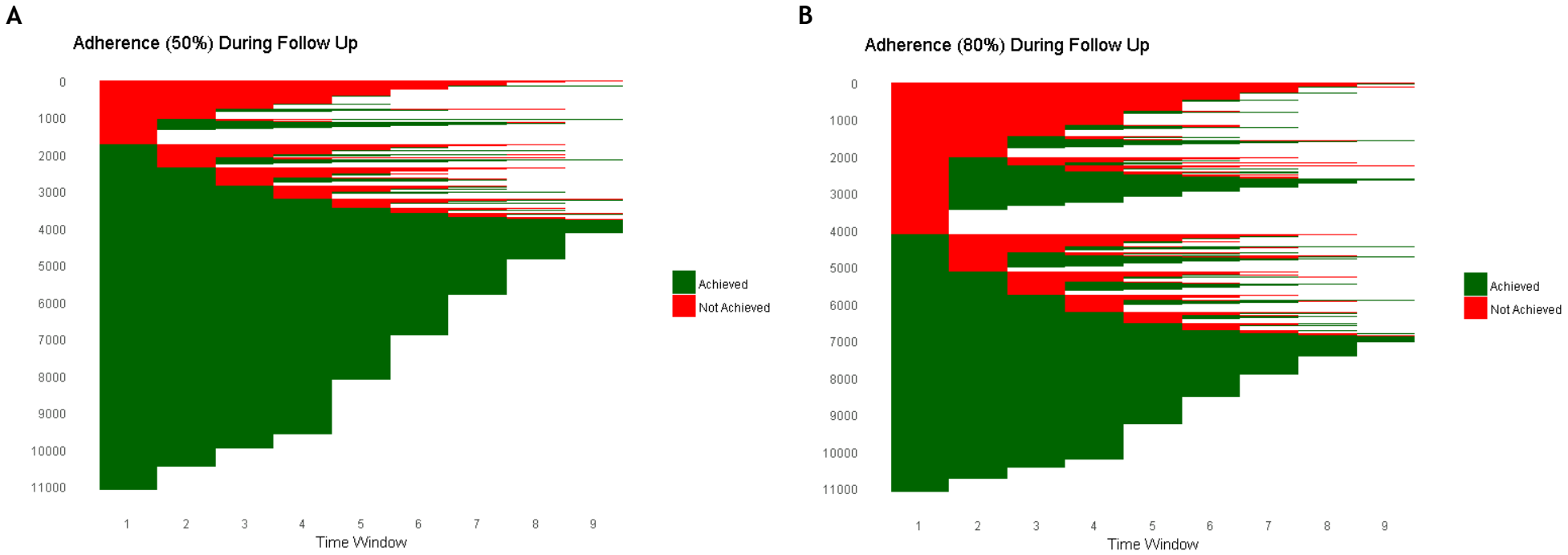


Figure 6.3 Individual Patterns of Statin Adherence During Follow Up (including incomplete years). Each time window was categorised by whether the patient achieved the adherence threshold (50% and 80%). Patient number on the y-axis is arbitrary, and patients were sorted by their adherence achievement patterns before plotting, with lower patient numbers assigned to those with a longer period of non-adherence.

Table 6.2: Demographics by Consistency of Achieving Adherence Thresholds (50% and 80%)

	50% Adherence			80% Adherence		
	All ≥50% 7,335 (66.0%)	**Inconsistent** 2,570 (23.1%)	**Never ≥50%** 1,205 (10.8%)	**All ≥80%** 4,229 (38.1%)	**Inconsistent** 4,871 (43.8%)	**Never ≥80%** 2,010 (18.1%)
Gender						
Male	4,641 (63.3%)	1,530 (59.5%)	561 (46.6%)	2,668 (63.1%)	3,025 (62.1%)	1,039 (51.7%)
Female	2,694 (36.7%)	1,040 (40.5%)	644 (53.4%)	1,561 (36.9%)	1,846 (37.9%)	971 (48.3%)
Age at MI (years)						
Mean (SD)	66.2 (13.3)	65.4 (14.1)	74.0 (15.2)	66.7 (13.0)	64.9 (13.6)	72.0 (15.2)
Median	66.4	65.4	77.1	67.1	64.5	74.8
IQR	55.9 – 76.8	54.3 – 76.8	65.2 – 85.3	56.9 – 76.9	54.4 – 75.9	61.8 – 83.9
Range	21.9 – 102.4	24.9 – 101.0	19.3 – 102.6	21.9 – 99.2	23.5 – 101.0	19.3 – 102.6
SIMD 2012 Quintile (NHS GGC)						
1 (Most)	1,769 (24.3%)	632 (24.8%)	249 (20.8%)	981 (23.4%)	1,213 (25.1%)	456 (22.9%)
2	1,727 (23.7%)	556 (21.8%)	245 (20.5%)	970 (23.1%)	1,121 (23.2%)	437 (21.9%)
3	1,463 (20.1%)	558 (21.9%)	273 (22.8%)	848 (20.2%)	999 (20.7%)	447 (22.4%)
4	1,232 (16.9%)	417 (16.3%)	228 (19.1%)	752 (17.9%)	774 (16.0%)	351 (17.6%)
5 (Least)	1,094 (15.0%)	388 (15.2%)	200 (16.7%)	648 (15.4%)	730 (15.1%)	304 (15.2%)
Missing	50	19	10	30	34	15
Diabetes at MI	1,386 (18.9%)	474 (18.4%)	239 (19.8%)	798 (18.9%)	873 (17.9%)	428 (21.3%)
Type 1	73 (1.0%)	27 (1.1%)	20 (1.7%)	44 (1.0%)	48 (1.0%)	28 (1.4%)
Type 2	1,313 (17.9%)	447 (17.4%)	219 (18.2%)	754 (17.8%)	825 (16.9%)	400 (19.9%)
Prior MI	473 (6.4%)	172 (6.7%)	69 (5.7%)	267 (6.3%)	304 (6.2%)	143 (7.1%)
1	370 (5.0%)	132 (5.1%)	56 (4.6%)	206 (4.9%)	237 (4.9%)	115 (5.7%)
>1	103 (1.4%)	40 (1.6%)	13 (1.1%)	61 (1.4%)	67 (1.4%)	28 (1.4%)

All patients in cohort included (n=11,110). Consistency of achievement of adherence thresholds defined using all full and partial years of follow up.

For both adherence thresholds, in univariable analysis, those who never met them were more likely to be female (50%: 53.4% vs 36.7%, 40.5%; 80%: 48.3% vs 36.9%, 37.9%). For 80%, approximately similar proportions were male between those who constantly adhered and those whose adherence was inconsistent (63.1% vs 62.1%), whereas for 50%, there was a slightly lower proportion of males who were inconsistent, compared to those who remained adherent (63.3% vs 59.5%). Those who were never adherent tended to be older, whereas those who were inconsistent were very slightly younger than those who were always adherent, for both thresholds. There were no clear patterns by deprivation quintile for the 80% threshold, though those never adherent at the 50% threshold were less likely to come from the most deprived quintile (20.8% vs 24.3%, 24.8%). Those who were inconsistent in their adherence were less likely to have been diagnosed with diabetes before their baseline MI, especially at the 80% cut-off. Those with a diagnosis of Type 1 diabetes were slightly more likely to never be adherent at both cut-offs. At the 80% cut-off, this pattern was also observed in those with a diagnosis of Type 2 diabetes. At the 50% threshold, those who were never adherent were less likely to have experienced an MI before their baseline MI, and more likely at the 80% threshold.

6.3.2 Time to Statin Non-Adherence

Before the optimal models for each threshold were found, Kaplan-Meier survival probabilities were plotted for the non-time-varying covariates considered for inclusion into the optimal models (not presented). These revealed the biggest differences were present for sex and age category, where females, and the oldest and youngest, were more likely to become non-adherent. There were minimal differences by deprivation quintile and whether the patient had experienced an MI before their baseline one.

Only patients who were prescribed statins at any point during their follow up were considered in these models (and subsequent analyses in the remainder of this chapter for adherence). The optimal model for time to adherence <50% contained the variables, age at MI, sex, and deprivation quintile only, all of which were non-time-varying. Consistent with the Kaplan-Meiers, a U-shaped association with age was observed: those over the age of 80 years were most likely to become non-adherent, whilst those aged 70-80 years, and less than 50

years had a similar hazard, and those aged 50-70 years were less likely. Additionally, males were also less likely to become non-adherent than females. Finally, hazard ratios for the association between becoming non-adherent and deprivation did not demonstrate a discernible trend. Similar hazard ratios were observed when patients were only followed up from the first time that they were adherent after their baseline MI (sensitivity results not presented).

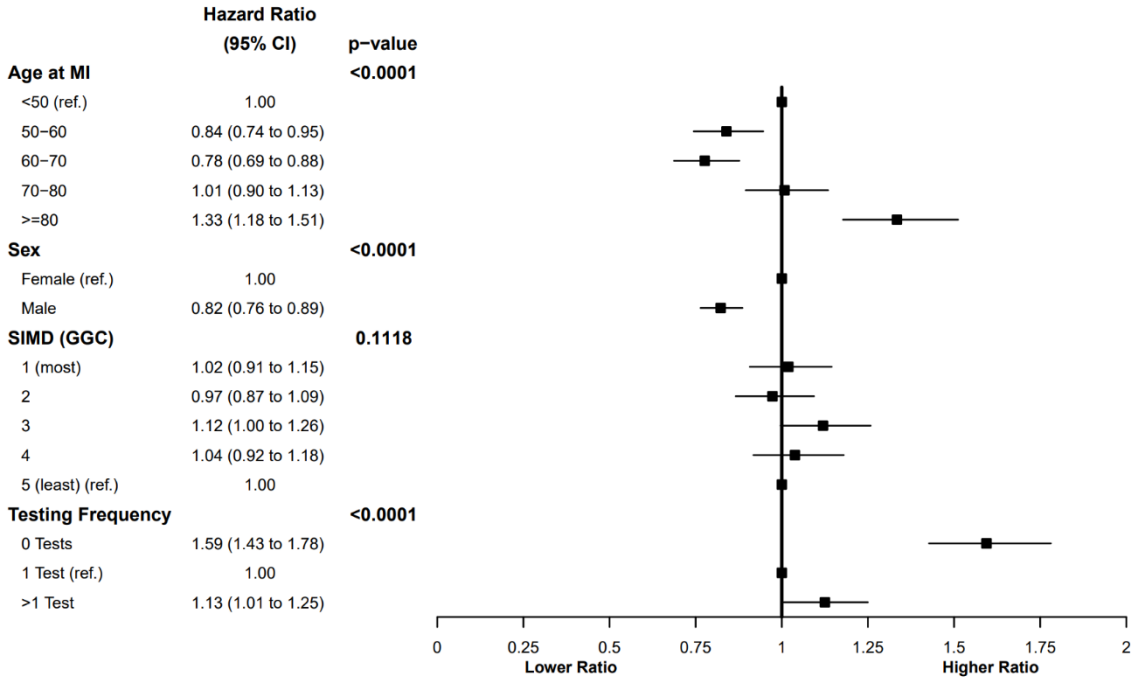


Figure 6.4: Time-Varying Cox Regression for Time to Non-Adherence (50% Threshold) with Lipid Testing Frequency Covariate. Testing frequency was the number of lipid tests conducted in the previous year of follow up, with the recommended one test per year as the reference level. Testing frequency was a time-varying covariate with one test assumed for the first year of follow up. AIC for this model was 55,302.72. Patients only included if ever prescribed statins during follow up.

These associations were not altered substantially following the inclusion of the number of lipid tests results in the previous year as a time-varying covariate, with this final model shown in Figure 6.4, and full details provided in Appendix O1. Furthermore, the inclusion of lipid testing frequency was statistically significant at the 5% level and slightly improved the fit of the model, decreasing the AIC from 55362.50 to 55302.72. Under-testing was associated with an increased likelihood of becoming non-adherent, with those with no lipid tests in the year before 60% more likely to become non-adherent at the 50% threshold. Meanwhile, those with more than one test per year were also more likely to become non-adherent, but to a lesser extent than those under-tested, with a hazard only 13% higher than those tested once in the previous year.

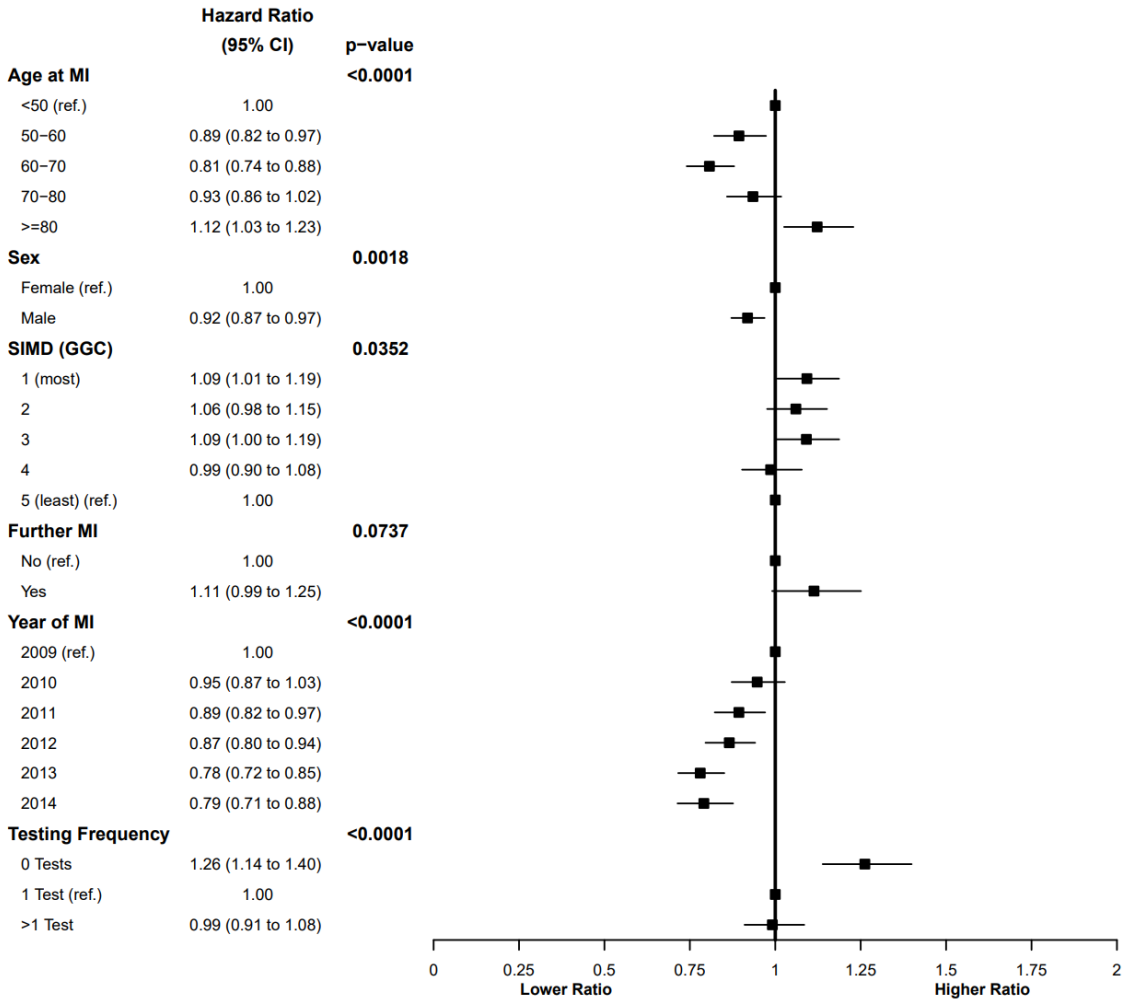


Figure 6.5: Time-Varying Cox Regression for Time to Non-Adherence (80% Threshold) with Lipid Testing Frequency Covariate. Testing frequency was the number of lipid tests conducted in the previous year of follow up, with the recommended one test per year as the reference level. Testing frequency was a time-varying covariate with one test assumed for the first year of follow up. AIC for this model was 108,472.8. Patients only included if ever prescribed statins during follow up.

Using the 80% adherence threshold, the optimal model for time to non-adherence contained the variables: age at MI, sex, deprivation quintile, further MI in the last year, and year of MI. In keeping with findings for the 50% threshold, there was a U-shaped association with age, males were less likely to become non-adherent than females, and there was no clear trend between deprivation and non-adherence despite its statistical significance at the 5% level. The only time-varying covariate in this model, a further MI in the previous year, was associated with an increased likelihood of becoming non-adherent, although this was not statistically significant despite its inclusion in the optimal model. Finally, the year of the baseline MI showed that those with more recent baseline MIs were less likely to become non-adherent. However, this could be a reflection of the reduced follow-up time available for those with later baseline MIs

compared to those whose follow up commenced in 2009. Nevertheless, whilst year of MI was significant in the final model, when it was removed in a sensitivity analysis the effect sizes observed for the remaining variables remained consistent, with only further MI in the previous year's hazard ratio increasing marginally (model not presented). Meanwhile, in a separate sensitivity analysis of the optimal model (not presented), similar associations with becoming non-adherent were also obtained when patients were considered only from the first year of follow up that they were adherent at the 80% threshold.

When the number of tests in the previous year was added as a time-varying covariate to this final model (Figure 6.5), these associations were minimally affected, and the model fit marginally improved (AIC decreased from 108490.1 to 108472.8). As noted in the model with the 50% threshold, no lipid tests in the previous year were associated with an increased likelihood of becoming non-adherent, although this increased hazard was lower at only 26%. However, unlike the 50% model, over-testing in the previous year was not associated with an increased likelihood of adherence <80%, with this hazard nearly equivalent to that of those tested once in the previous year. Further details of the model can be found in Appendix O2.

6.4 Latent Class Analysis for Adherence

6.4.1 Optimal Model for 50% Adherence

There were 26,565 time windows from 8,626 patients used in the generation of the adherence LCA models. The variables used to determine class-membership were whether the patient's adherence was ≥50% the previous year, whether the patient's LDL was ≤1.8mmol/l the previous year, and those variables identified in the earlier time-varying cox models (Chapter 6.3.2): age at MI, sex, and deprivation quintile.

The optimal LCA model for the achievement of the 50% adherence threshold converged after 795 iterations and contained four classes. Using modal assignment, the predicted class shares were 21.7%, 33.4%, 30.7%, and 14.2%, for

classes 1-4 respectively. Full model details, including conditional class probabilities, can be found in Appendix P1.

When the conditional class probabilities were compared, the largest differences between these classes were in the patient's achievement of the 50% adherence threshold in the previous year. Specifically, there were low numbers of those who were non-adherent in the first three classes, and then the majority of patients in class 4 (the least likely to achieve the threshold) had not achieved the target the previous year. This pattern was also observed for the achievement of the LDL target in the previous year in this model, with more meeting the target of LDL≤1.8mmol/l in classes 1, 2, and 3 in the previous year, and fewer in class 4.

The predicted class assignments for the time windows were cross-tabulated with the achievement of the 50% adherence threshold (Table 6.3). This revealed that all patients who were assigned to classes 1, 2, and 3 would have adherence ≥50% that year. Therefore, if these classes were all considered as predictions that the adherence would be achieved, the table could be collapsed to form a 2x2 table, with model diagnostics calculated, as shown in Table 6.4.

Table 6.3: Predicted Class Assignments for Optimal 50% Adherence LCA Model by Achievement of Adherence Threshold

	Meet	Not Meet	Total
Class 1 (Most)	5,762	0	5,762
Class 2	8,871	0	8,871
Class 3	8,156	0	8,156
Class 4 (Least)	313	3,463	3,776
Total	23,102	3,463	26,565

Time windows assigned to classes using predicted modal assignment. Only time windows with complete data available included in the LCA model. Likelihood of classes to meet the adherence threshold of 50% was ascertained through multinomial regression coefficients produced during a one-step model fitting process.

Therefore, if classes 1, 2, and 3, are combined, this yields a positive predictive value of 100%. Specifically, this means that of those predicted to have adherence ≥50% (by being assigned to class 1, 2, or 3), all of them will achieve this. This also resulted in a specificity of 100%, with all instances where adherence would be <50% being assigned to class 4 under this model. There were

only 313 time windows in class 4 where the adherence threshold would have been met. This meant that sensitivity was also high, with 98.6% of those who would have adherence ≥50% being assigned to these classes 1, 2, and 3. Additionally, this also translated into a fairly high negative predictive value of 91.7%; meaning that nearly 92% of those assigned to class 4 would not achieve the adherence target of 50%.

Table 6.4: Diagnostic Table for Optimal LCA Predictions for 50% Adherence

	1,2,3 vs 4	
	LCA	**Bootstrap (95% CI)**
Sensitivity	98.6%	98.7% (98.5%, 98.8%)
Specificity	100.0%	100.0% (100.0%, 100.0%)
Positive Predictive Value	100.0%	100.0% (100.0%, 100.0%)
Negative Predictive Value	91.7%	91.6% (90.5%, 92.8%)

LCA = Latent Class Analysis; CI = Confidence Interval. LCA diagnostics calculated by collapsing rows for classes 1-3 in Table 6.3. Diagnostics were bootstrapped 5,000 times with confidence intervals ascertained using the percentile method. Within each bootstrap, one time window per patient and its predicted class was selected, and cross-tabulations and diagnostic measures were calculated as on the full set of time windows above.

Following 5,000 bootstraps using one time window per patient, these diagnostic values were largely unaffected, with a minor increase in sensitivity observed, and a minor decrease in negative predictive value. Both specificity and positive predictive values, which were 100%, were not affected by this bootstrap, as outlined in Chapter 6.2.4.

Table 6.5: Average Posterior Probabilities for Optimal LCA for 50% Adherence

Class	Ave PP
1	0.7345
2	0.7086
3	0.7406
4	0.9609

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all time windows modally assigned to that class (i.e. only includes the time windows where the posterior probability was highest for that class).

Furthermore, the relative entropy for this model was also reasonable, at 0.6015, suggesting that there was some adequate separation between the classes. This was reinforced by the average posterior probabilities for each class (as shown in Table 6.5). These further suggested that the classes were distinct, and the LCA model fitted adequately with all values above 0.7. This was especially true for class 4, where the average probability was approaching one. This was

particularly beneficial, with class 4 being the only class to be treated as indicative of likely to not be adherent at the 50% threshold.

6.4.2 Optimal Model for 80% Adherence

Like the 50% adherence model, the optimal model for the 80% threshold also contained four classes and converged after 2,638 iterations. The variables used to determine class membership were similar, with the addition of whether the patient had experienced a further MI in the previous time window and the year of the baseline MI, and adherence in the previous year captured using the 80% threshold instead of 50%. All other variables considered in the model remained the same. Following modal assignment, the predicted class shares were 29.7%, 39.8%, 14.2%, and 16.4% for classes 1–4 respectively. Full model details, including conditional class probabilities, are provided in Appendix P2.

The conditional class probabilities revealed that the largest differences between these classes were in the patient's adherence last year, with low numbers of those with adherence <80% the previous year in the first three classes and nearly all of the patients in class 4 (the least likely) were non-adherent by this measure. This was also the case for achieving the LDL target in the previous year although the differences in proportions of those meeting and not meeting the target was less pronounced between the classes.

Table 6.6: Predicted Class Assignments for Optimal 80% Adherence LCA Model by Achievement of Adherence Threshold

	Meet	Not Meet	Total
Class 1 (Most)	7035	858	7893
Class 2	9416	1143	10559
Class 3	3305	462	3767
Class 4 (Least)	0	4346	4346
Total	19756	6809	26565

Time windows assigned to classes using predicted modal assignment. Only time windows with complete data available included in the LCA model. Likelihood of classes to meet the adherence threshold of 80% was ascertained through multinomial regression coefficients produced during a one-step model fitting process.

Cross-tabulating the class predictions with the achievement of the 80% adherence threshold (Table 6.6) yielded less conclusive results than had been seen at the 50% threshold. Nonetheless, the optimal way of grouping the classes

to form a 2x2 table remained the same, producing the diagnostics in Table 6.7, and patients assigned to classes 1, 2, and 3 more likely to have adherence ≥80%, and class 4 the least likely.

However, unlike in the 50% model, some patients assigned to classes 1, 2, and 3 did not meet the 80% target, resulting in a positive predictive value of 88.9%. Despite this, all patients who did meet the threshold were assigned to one of these three classes, meaning that all patients that were predicted to be in class 4 did not meet the threshold. Consequently, the sensitivity and negative predictive value of the model were both equal to 100%. Therefore, the diagnostics of this model mean that whilst all patients predicted to not have adherence ≥80% (class 4) would do so, only 88.9% of those predicted to meet the target (classes 1-3) would. As a result, implementing this model to identify those who could safely receive less frequent monitoring may be problematic, with around 1 in 10 patients incorrectly considered for reduced monitoring.

Table 6.7: Diagnostic Table for Optimal LCA Predictions for 80% Adherence

	1,2,3 vs 4	
	LCA	**Bootstrap**
Sensitivity	100.0%	100.0% (100.0%, 100.0%)
Specificity	63.8%	63.3% (62.2%, 64.4%)
Positive Predictive Value	88.9%	88.7% (88.1%, 89.2%)
Negative Predictive Value	100.0%	100.0% (100.0%, 100.0%)

LCA = Latent Class Analysis; CI = Confidence Interval. LCA diagnostics calculated by collapsing rows for classes 1-3 in Table 6.6. Diagnostics were bootstrapped 5,000 times with confidence intervals ascertained using the percentile method. Within each bootstrap, one time window per patient and its predicted class was selected, and cross-tabulations and diagnostic measures were calculated as on the full set of time windows above.

Bootstrapping these diagnostics 5,000 times and using one observation per patient (n=8,626) resulted in slight decreases in the two diagnostic measures not equal to 100% (for reasons detailed in Chapter 6.2.4). The largest decrease was in the model's specificity, the percentage of those who didn't meet the threshold who were assigned to class 4, which decreased by 0.5% to 63.3%. A minimal decrease was noted in the positive predictive value to 88.7%.

Nevertheless, there was evidence of some adequate separation between the classes, as the relative entropy of the LCA model was 0.6231. In Table 6.8, the average posterior probabilities for each class also support this, with all above

0.7. Furthermore, the highest average posterior probability was for class 4, the only class indicative of being not likely to achieve the adherence target of 80%, with a probability above 0.9, suggesting that this class was highly separated from the other three.

Table 6.8: Average Posterior Probabilities for Optimal LCA for 80% Adherence

Class	Ave PP
1	0.7359
2	0.7115
3	0.8220
4	0.9153

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all time windows modally assigned to that class (i.e. only includes the time windows where the posterior probability was highest for that class).

6.4.3 Mortality by Predicted Class

6.4.3.1 Mortality by Predicted Class for 50% Adherence

As combining classes 1-3 offered the optimal model diagnostics, these classes were also grouped when associations with mortality were examined (Table 6.9). In both years 2 and 3, there was a higher ratio of mortality amongst those predicted to not meet the adherence threshold of 50% (class 4) than those predicted to meet it. All results were statistically significant at the 5% level.

Table 6.9: Mortality by Predicted 50% Adherence Class (from Optimal LCA Model)

	Year 2		Year 3	
Classes	**1, 2, 3** (Most)	**4** (Least)	**1, 2, 3** (Most)	**4** (Least)
N	6,234	902	4,711	760
Events (%)	324 (5.2%)	96 (10.6%)	212 (4.5%)	79 (10.4%)
Cox PH				
Hazard Ratio	1.00 (ref.)	2.13	1.00 (ref.)	2.40
95% CI	--	(1.70, 2.67)	--	(1.86, 3.11)
		p-value<0.0001		***p-value <0.0001***

CI = Confidence Interval, PH = Proportional Hazards, ref = reference category. Patients only included if a predicted class assignment was available for that year of follow up from the 50% adherence LCA model. Models fitted are unadjusted for further covariates.

The percentage of patients who died in class 4 (the least likely to have adherence ≥50%) for both year 2 and 3, was more than double the percentage who died in classes 1, 2, and 3 (the most likely). This was also reflected in the

associated hazard ratios, of which both were statistically significant at the 5% level. Hazards of mortality in the second year of follow up were 2.13 times greater in class 4 than classes 1, 2, and 3 combined, and they were 2.40 times greater in the third year of follow up. Full details of all models are provided in Appendix Q.

6.4.3.2 Mortality by Predicted Class for 80% Adherence

The same approach was also taken using the predicted class assignments for the optimal 80% model, with classes 1, 2, and 3 combined, as shown in Table 6.10, with full details of the models fitted provided in Appendix R. In both years 2 and 3, those assigned to the least likely to achieve adherence ≥80% class (class 4) were at an increased risk of mortality compared to those assigned to the other three classes (classes 1-3), with hazard ratios statistically significant at the 5% level.

Table 6.10: Mortality by Predicted 80% Adherence Class (from Optimal LCA Model)

	Year 2		Year 3	
Classes	**1, 2, 3** (Most)	**4** (Least)	**1, 2, 3** (Most)	**4** (Least)
N	6,043	1,093	4,590	881
Events (%)	327 (5.4%)	93 (8.5%)	230 (5.0%)	61 (6.9%)
Cox PH				
Hazard Ratio	1.00 (ref.)	1.61	1.00 (ref.)	1.40
95% CI	--	(1.28, 2.03)	--	(1.06, 1.86)
		p-value: 0.0001		***p-value: 0.0239***

CI = Confidence Interval, PH = Proportional Hazards, ref = reference category. Patients only included if a predicted class assignment was available for that year of follow up from the 80% adherence LCA model. Models fitted are unadjusted for further covariates.

Overall, the difference in the risk of mortality was less evident in these 80% adherence threshold models than in the 50% adherence threshold models discussed above. However, the percentage of patients who died in the second year of follow up was higher amongst patients assigned to class 4, compared to the other three classes combined (8.5% vs 5.4% respectively). This was also the case in the third year of follow up, with the number of events slightly lower and a smaller difference in the percentages (6.9% vs 5.0%). The associated hazard ratios reflected this, with those assigned to class 4 with 61% greater hazard of

mortality compared to classes 1, 2, and 3 in the second year of follow up, with 40% higher risk in the third year.

6.5 Latent Class Growth Analysis for Adherence

6.5.1 Optimal Model for 50% Adherence

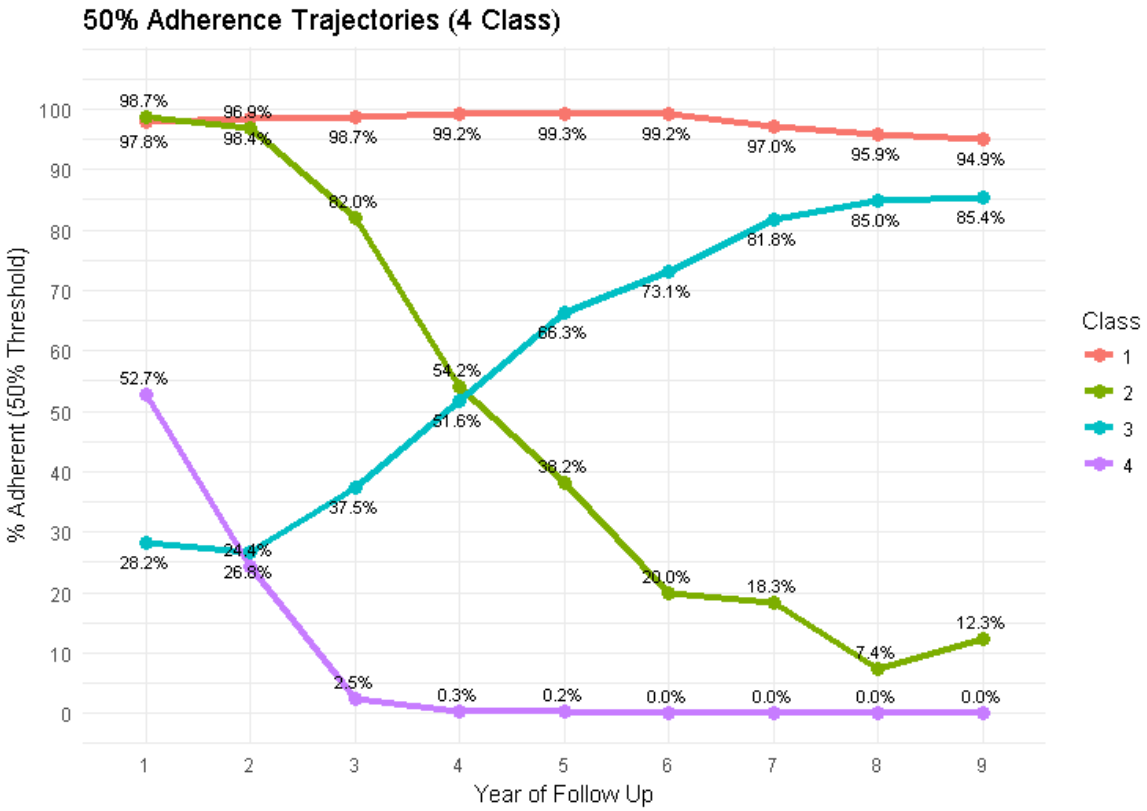


Figure 6.6: Trajectory Curves for % Adherent (50% Threshold) by Assigned Class from Optimal LCGA Model. Class was predicted for a patient using modal assignment. Percentages calculated using the number of patients with data available for that year of follow up.

For the 50% adherence LCGA, the variables considered were those identified in the earlier time-varying Cox models (Chapter 6.3.2) that were captured at baseline: age at MI, sex, and deprivation quintile. The optimal model for the 50% adherence threshold was run on data from 10,418 patients and contained four classes. Over three-quarters of these patients were assigned to class 1, and 8%, 5%, and 11% assigned to classes 2, 3, and 4 respectively. Full model details, including comparisons of the model diagnostics for those with 2, 3, and 4 classes are provided in Appendix S1.

As shown in Figure 6.6, within the largest class (class 1), in any given year, nearly 95% of the patients had adherence ≥50%. Unsurprisingly, as a result of its

large size, the characteristics of this group were similar to those of the overall population, with the highest proportion of those aged between 60 and 70 years, and male of all the classes (63%). The proportion in each SIMD quintile was also similar to that of the overall cohort.

In contrast, those in classes 2 and 4 show similar trajectories in adherence achievement, with rapid declines in the proportion achieving the threshold after the first couple of years, although class 2's adherence is initially comparable to those in class 1. These classes have the highest proportions of females (compared to classes 1 and 3), and higher proportions within the oldest age brackets. This is particularly true for class 4 (the worst-performing adherence trajectory), where just over half of the class were female, and 55% were over the age of 70. In terms of deprivation, there were some differences between the two; class 4 had the highest proportions of patients in the third and fourth quintiles of all classes, whilst class 2 had the highest proportion in the fifth (and least deprived) quintile.

Finally, class 3, which was the smallest of the classes, initially has a low proportion of patients who are adherent, with the percentage gradually increasing throughout follow up to around 85% in years 8 and 9. This group was similar to class 1 in its proportion of males yet had the highest proportion of patients under the age of 60 years and from the most deprived quintile of all the classes. This trajectory could be partly explained by survival bias, as there was a higher rate of death among those who were non-adherent within this class near the beginning of their follow up time with only those who are adherent remaining alive causing the proportion to increase. However, this impact is likely to be minimal, as a sensitivity analysis run only in patients who did not die during follow up assigned only 0.5% of patients to a different class.

Classification diagnostics suggested that the model fit was reasonable and the distinction between the class trajectories was adequate. Relative entropy was fairly high with a value of 0.7639 and all average posterior probabilities for each class (Table 6.11) above 0.7. This is especially true for the consistently adherent class (class 1), which had the highest average posterior probability of all the groups, with a probability of 0.9185. This suggests that whilst classification in

groups 2-4 is sufficiently high, there is very little uncertainty in the assignment of patients to the consistent group.

Table 6.11: Average Posterior Probabilities for Optimal LCGA for 50% Adherence

Class	Ave PP
1	0.9185
2	0.7602
3	0.7821
4	0.7869

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all patients modally assigned to that class (i.e. only includes the patients where the posterior probability was highest for that class).

6.5.2 Optimal Model for 80% Adherence

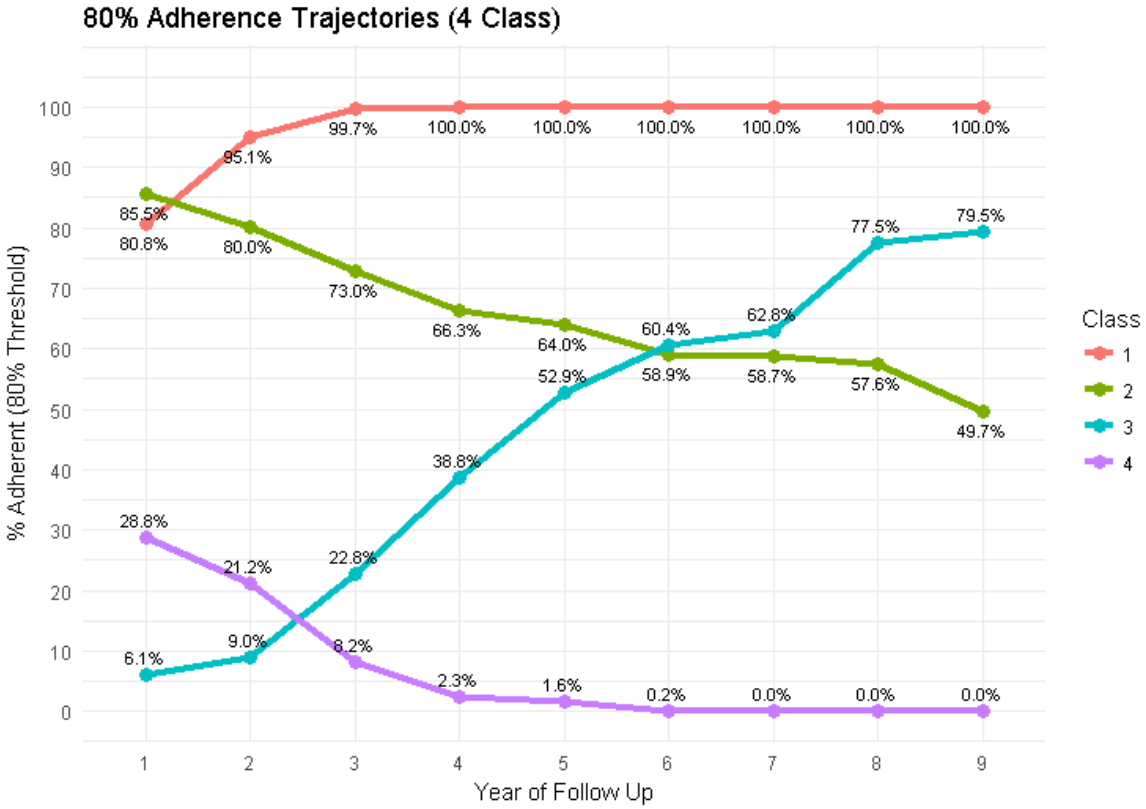


Figure 6.7: Trajectory Curves for % Adherent (80% Threshold) by Assigned Class from Optimal LCGA Model. Class was predicted for a patient using modal assignment. Percentages calculated using the number of patients with data available for that year of follow up.

As with the 50% adherence model, only variables captured at baseline that were included in the optimal time-varying cox models discussed earlier (Chapter 6.3.2) were included within the latent classes. Specifically, these were the age at the time of MI, sex, deprivation quintile, and the year of MI, with only the latter variable differing from the model fitted to the 50% adherence threshold.

The optimal model contained four classes and revealed similar patient trajectories to those observed with the 50% model. However, there were substantial differences in the estimated class shares for the 10,418 patients included. Class 1 remained the largest group, with an estimated class share of 48%, with the remaining three classes, representing 26%, 8%, and 19%, respectively. Comparisons of model diagnostics for the LCGA models with 2, 3, and 4 classes, and full model details of the optimal model are provided in Appendix S2.

The trajectories of these groups (Figure 6.7) are approximately similar to those seen in the 50% adherence model. However, class 1, whilst generally remaining consistently adherent, starts with only 81% of the patients achieving the 80% adherence threshold in the first year of follow up, with this increasing to 100% in year 4 and then remaining so. Furthermore, in class 2, the decline in the proportion with adherence ≥80% is less pronounced than in the 50% model, with this decreasing to just less than 50% adhering compared to less than 10% using the other adherence threshold. In contrast, class 3 started with lower levels of patients with adherence ≥80%, although the proportion did increase to just below 80% in year 9 of follow up. Finally, class 4 showed a similar trend in the two adherence models, although the percentage adherent in the first year of follow up for the 80% model was approximately half that of the 50% model.

The consistent class (class 1) is characterised by high proportions of those aged between 60 and 80 years at the time of their MI and a gender divide comparable to the overall population. There was also the highest proportion of patients in the least deprived quintile within class 1, with also relatively high proportions of those assigned to the third and fourth quintiles. Those assigned to class 1 were also more likely to have experienced their baseline MIs later in the period of interest, with a third of patients assigned to this group commencing their follow up in 2013 or 2014.

The gender divide in class 1 was similar in class 2, but those assigned to this class were marginally older with a high proportion aged over 70 years at the start of their follow up. There was also a trend towards residing in more deprived areas in patients assigned to this class. The year of the baseline MI was

also earlier with the highest proportion of MIs in 2011 of any classes and a relatively high proportion in 2010, resulting in a longer average duration of follow up for this class.

The characteristics of class 3 in this model were similar to those of class 3 in the model for the 50% adherence threshold. Specifically, patients assigned to this class were more likely to be male, be younger than 60 years at the time of their baseline MI and reside in areas in the two most deprived quintiles. This class were also likely to have a longer duration of follow up available, with 57% of those assigned experiencing their baseline MIs in 2009 and 2010.

Finally, within the class whose proportion of patients who were adherent rapidly approached 0% (class 4), there was the highest percentage of females of the four classes. This class was also most likely to reside in the third deprivation quintile and were the oldest, with more than a quarter of the patients over the age of 80 years. However, average follow up was likely to be fairly short for this group, with many patients experiencing their baseline MI in 2012 and 2013.

Table 6.12: Average Posterior Probabilities for Optimal LCGA for 80% Adherence

Class	Ave PP
1	0.7324
2	0.8054
3	0.6897
4	0.7109

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all patients modally assigned to that class (i.e. only includes the patients where the posterior probability was highest for that class).

Overall, the classification diagnostics for this model suggested that the fit of this model was likely to be reasonably adequate. However, these were less convincing than those that had been observed for the LCGA model for the 50% adherence threshold. Relative entropy for this 80% model was 0.5914, suggesting that there was some distinction between the classes assigned, this was considerably lower than the 0.7639 observed for the earlier adherence model. Furthermore, only three of the four classes had an average posterior probability above 0.7 (Table 6.12), and class 3 (which had poor initial adherence before improving) had the lowest average posterior probability, equal to 0.6897. Lower average posterior probabilities lead to greater uncertainty in the assignment of

patients to the class and therefore increase the possibility of incorrect assignments and reduced accuracy of the model.

This model was also more sensitive to mortality within the classes, possibly as a consequence of the inclusion of year of MI within the class-membership model. Indeed, following exclusion of patients who died before the end of follow up, whilst the four adherence trajectories were broadly similar, 22.8% of the remaining were assigned to a different adherence trajectory in this sensitivity analysis.

6.5.3 Mortality by Predicted LCGA Class

6.5.3.1 Mortality by Predicted LCGA Classes for 50% Adherence

The Kaplan-Meier in Figure 6.8 compares the overall survival probabilities within each of the groups identified by the LCGA for 50% adherence trajectories. Patients assigned to class 2, where the proportion adherent started high and then decreased rapidly experienced no deaths in the first two years of follow up where their adherence was highest. However, after two years, the number of events increased rapidly resulting in a decrease in survival probability with only those assigned to the least adherent class (class 4) faring worse by the end of follow up. This pattern in class 2 occurred despite no substantial differences between this group with any others in terms of class-membership variables (Appendix S1), and no information included in the model could have resulted in overfitting. Consequently, such a curve is likely the result of substantial clinical differences in the baseline health and the comparatively small group size (8%).

For patients in class 3, where the proportion with adherence ≥50% increased throughout follow up, there was an initial rapid decline in survival probability where adherence was worst followed by a plateauing in the curve when the proportion of those adherent within the class was increasing. For patients assigned to the consistently adherent group (class 1), the survival probability curve showed a steadily decreasing survival probability over follow up. Initially, these probabilities were lower than those observed within class 2, but due to the steep increase in events for class 2 after two years of follow up, class 1 had the optimal survival curve from year 5 onwards.

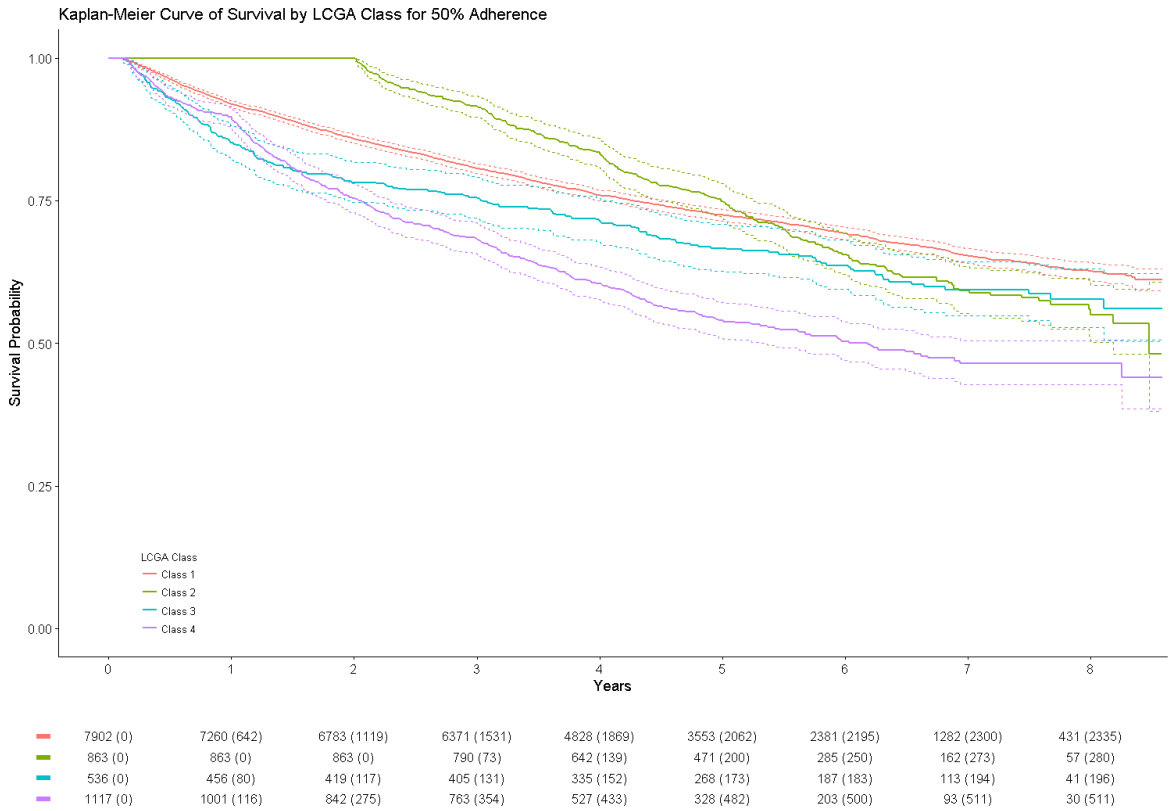


Figure 6.8: Kaplan-Meier Curve for Mortality by Predicted 50% Adherence LCGA Class Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the 50% adherence LCGA model. Numbers in table are number at risk (cumulative number of events).

Table 6.13: Hazard Ratios of Mortality by Predicted 50% Adherence LCGA Class

	HR (95% CI)	p-value
Class 1 (ref)	1.00	
Class 2	0.99 (0.87, 1.12)	<0.0001
Class 3	1.26 (1.09, 1.46)	
Class 4	1.84 (1.67, 2.02)	

HR = Hazard Ratio, CI = Confidence Interval. Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the 50% adherence LCGA model. Models fitted are Cox proportional hazards and were unadjusted for further covariates.

The hazard ratios from a Cox regression model are presented in Table 6.13 and illustrate the differing risks of mortality between the assigned classes. Furthermore, this also found that these latent classes were significantly associated with mortality risk within this population at the 5% level. Full model details, including the Schoenfeld residual plot, are provided in Appendix T1.

Taking the consistently adherent group (class 1) as a reference, both classes 3 and 4 had an increased risk, whilst the risk of mortality in class 2 was approximate to the risk observed in class 1. Specifically, patients in class 4 were

at the greatest risk of mortality, with an associated hazard 84% higher than for patients assigned to class 1. For patients in class 3, the associated hazard for mortality was 26% higher. There was no significant difference in the mortality hazards between classes 2 and 1, with the hazard ratio suggesting that there was a 1% lower risk for class 2. However, this lack of significant difference for these two classes is likely to have occurred due to a violation of the proportional hazards assumption of the Cox regression model. This violation was evident in the earlier Kaplan-Meier (Figure 6.8) where the two lines for these classes crossed. Therefore, this HR from the Cox model (comparing class 1 and 2) is included for completeness only and should not be taken as indicative of the true underlying trends illustrated in the Kaplan-Meier curve.

6.5.3.2 Mortality by Predicted LCGA Classes for 80% Adherence

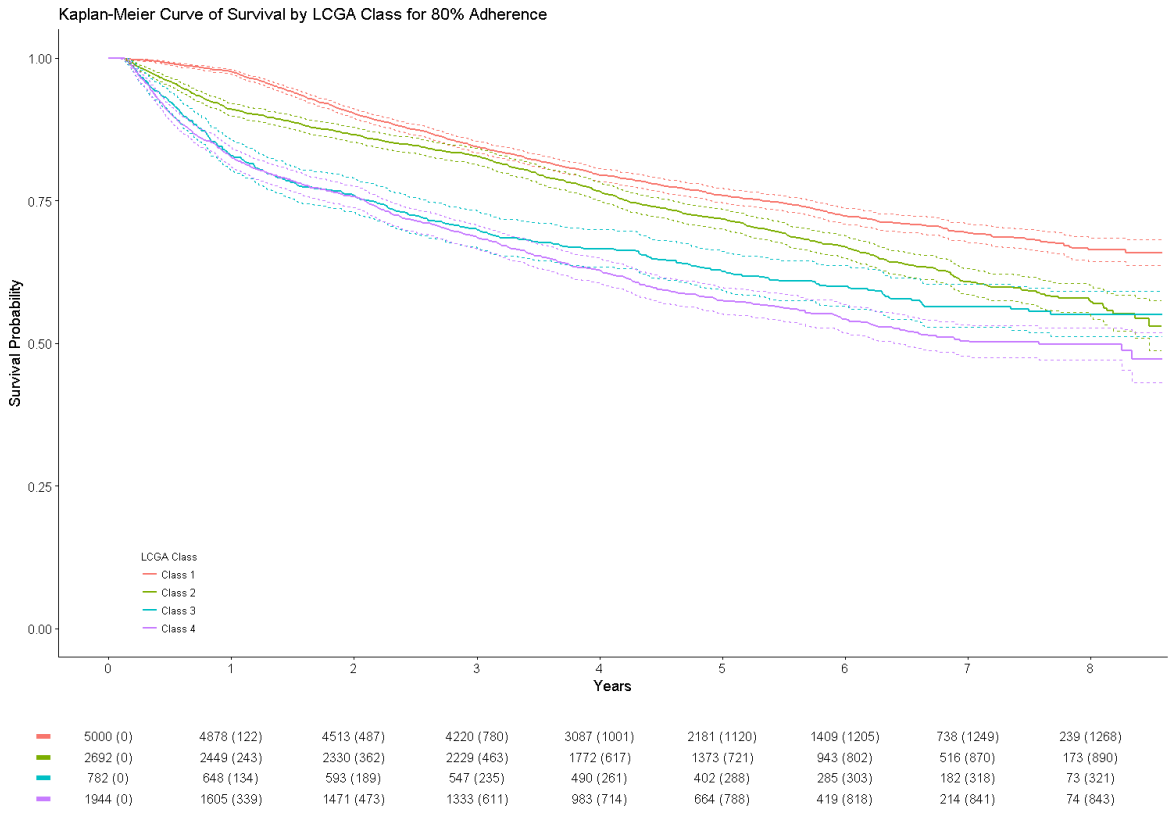


Figure 6.9: Kaplan-Meier Curve for Mortality by Predicted 80% Adherence LCGA Class. Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the 80% adherence LCGA model. Numbers in table are number at risk (cumulative number of events).

For the 80% adherence LCGA, the Kaplan-Meier revealed similar survival probabilities for classes 1, 3 and 4 (Figure 6.9) to those seen in the previous 50% adherence model. However, in this instance, those assigned to class 2 (high starting adherence with a large decrease) had survival probabilities between

classes 1 and 3 throughout follow up. Therefore, those assigned to class 1, whose adherence was fairly consistent following their baseline MI, had the optimal survival of all the classes in this model. The survival probabilities of those assigned to classes 3 (improved adherence) and 4 (low and decreasing adherence) were initially similar for the first three years of follow up, with survival only marginally improving for class 3 during the later years of data compared to class 4.

Table 6.14: Hazard Ratios of Mortality by Predicted 50% Adherence LCGA Class

	HR (95% CI)	p-value
Class 1 (ref)	1.00	
Class 2	1.30 (1.19, 1.42)	<0.0001
Class 3	1.71 (1.51, 1.93)	
Class 4	2.04 (1.87, 2.22)	

HR = Hazard Ratio, CI = Confidence Interval. Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the 80% adherence LCGA model. Models fitted are Cox proportional hazards and were unadjusted for further covariates.

These differences were also reflected in the hazard ratios from a Cox regression model, as shown in Table 6.14, with full details provided in Appendix T2. This regression found that there was a statistically significant association between mortality and the latent classes from this LCGA model at the 5% level. Using the consistently adherent group (class 1) as the reference group, patients assigned to each of the three other classes had a higher risk of mortality. Those assigned to class 2 (high initial and then decreasing adherence) had the smallest associated increase in the risk of mortality, with a hazard 30% greater than those in class 1. Meanwhile, those with increasing adherence over follow up (class 3) had a 71% increased risk of mortality than those consistently adherent. Class 4, low and decreasing proportions adherent, however, were most at risk, with an associated hazard more than twice that of those assigned to class 1.

6.6 Factors Associated with LDL Target Achievement

6.6.1 Patterns in Achieving LDL Targets (within patients)

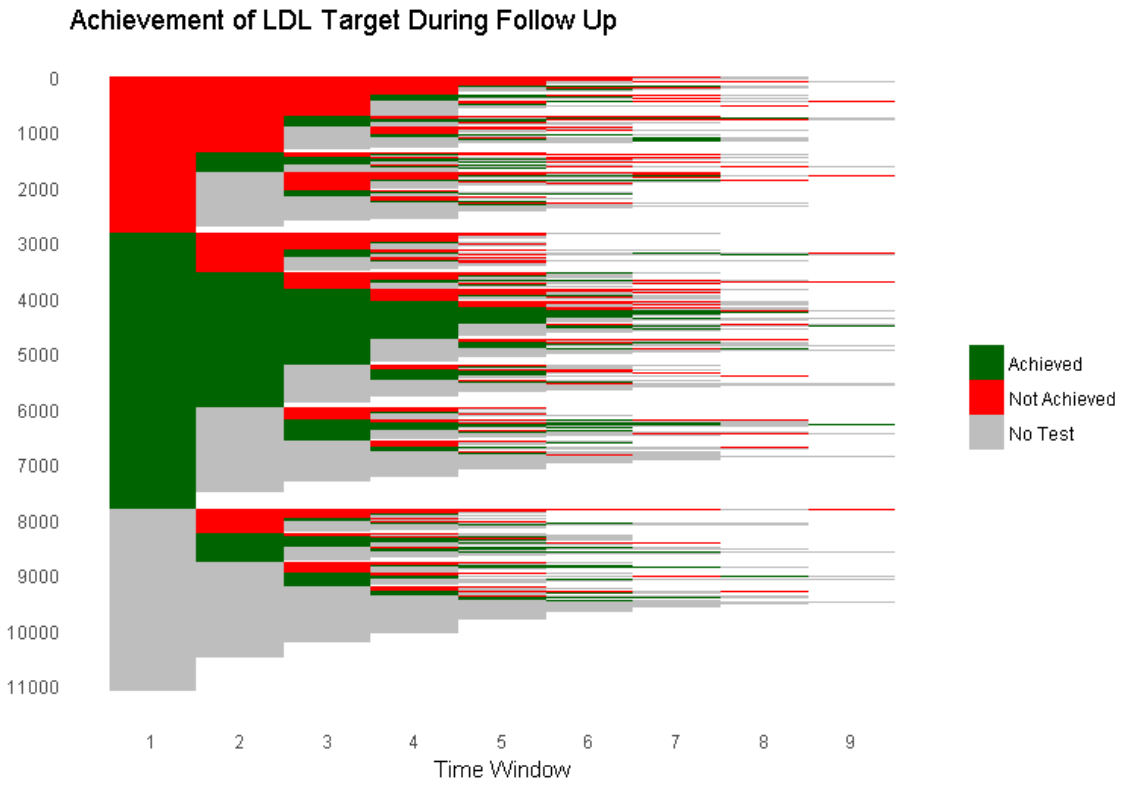


Figure 6.10 Individual Patterns of LDL Target Achievement During Follow Up (including incomplete years). Each time window was categorised by whether the patient achieved the LDL target of ≤1.8mmol/l, or whether no test result was available. Patient number on the y-axis is arbitrary, and patients were sorted by their LDL target achievement patterns before plotting, with lower patient numbers assigned to those with a longer period of not achieving the target.

In the first year after their baseline MI, approximately half of the patients were on target, with many remaining so throughout follow up, as shown in Figure 6.10. Similarly, many patients with no results in their first year of follow up also did not have results for the remainder. Of those that did have a test in subsequent years, similar proportions achieved and did not achieve the LDL target of ≤1.8mmol/l. For all years of follow up, patients with a lipid test result were more likely in the following year to have not had a test than having had a test result with a differing conclusion, regardless of whether they had achieved or not achieved the target in the current year. Furthermore, once one test was missed, only a minority of patients were subsequently tested again.

Table 6.15: Demographics by Consistency of Achieving Target LDL (LDL≤1.8mmol/l)

	All ≤1.8mmol/l 3,288 (34.6%)	Inconsistent 3,927 (41.3%)	Never ≤1.8mmol/l 2,289 (24.1%)
Gender			
Male	2,056 (62.5%)	2,537 (64.6%)	1,352 (59.1%)
Female	1,232 (37.5%)	1,390 (35.4%)	937 (40.9%)
Age at MI (years)			
Mean (SD)	68.3 (13.7)	62.9 (12.3)	65.5 (14.0)
Median	69.8	62.7	65.2
IQR	58.0 – 79.1	53.7 – 72.4	54.8 – 76.6
Range	19.3 – 102.2	20.5 – 99.3	22.5 – 98.1
SIMD 2012 Quintile (NHS GGC)			
1 (Most)	952 (29.2%)	872 (22.3%)	412 (18.2%)
2	738 (22.6%)	882 (22.6%)	551 (24.3%)
3	625 (19.1%)	813 (20.8%)	535 (23.6%)
4	488 (15.0%)	652 (16.7%)	454 (20.0%)
5 (Least)	461 (14.1%)	688 (17.6%)	317 (14.0%)
Missing	24	20	20
Diabetes at MI			
Any	771 (23.4%)	676 (17.2%)	315 (13.8%)
Type 1	33 (1.0%)	40 (1.0%)	32 (1.4%)
Type 2	738 (22.4%)	636 (16.2%)	283 (12.4%)
Prior MI	249 (7.6%)	224 (5.7%)	115 (5.0%)
1	184 (5.6%)	177 (4.5%)	95 (4.2%)
≥1	65 (2.0%)	47 (1.2%)	20 (0.9%)

Consistency does not allow for years where an LDL result is missing, e.g. a patient may only have LDL results every other year, but each time meet the target, and therefore would fall into all ≤1.8mmol/l group. Patients only included if at least one LDL result was available during their follow up (n=9,504).

Patients with at least one LDL result (n=9,504) were then categorised into those who only ever achieved the LDL target of ≤1.8mmol/l (35%), those who never achieved the target (24%), and those who achieved it inconsistently (41%). Those who consistently met the target had the highest average age, were more likely to reside in the most deprived areas, have a diagnosis of type 2 diabetes at their baseline MI, and have experienced a prior MI. In contrast, those who never met the target were the most likely to be female, most likely to reside in deprivation quintiles 2-4, and have a diagnosis of type 1 diabetes. Finally, patients who inconsistently achieved the target were younger and more likely to reside in the least deprived areas.

6.6.2 Time to Non-Target LDL

Kaplan-Meier plots (not presented) for the non-time-varying covariates considered in these models revealed the biggest differences were present by age category and deprivation quintile, where the oldest, and most deprived were less likely to become not on target. There was also a small difference noted by the prior MI status, with those with an MI before their baseline less likely to not meet the LDL target. There were no differences observed by sex.

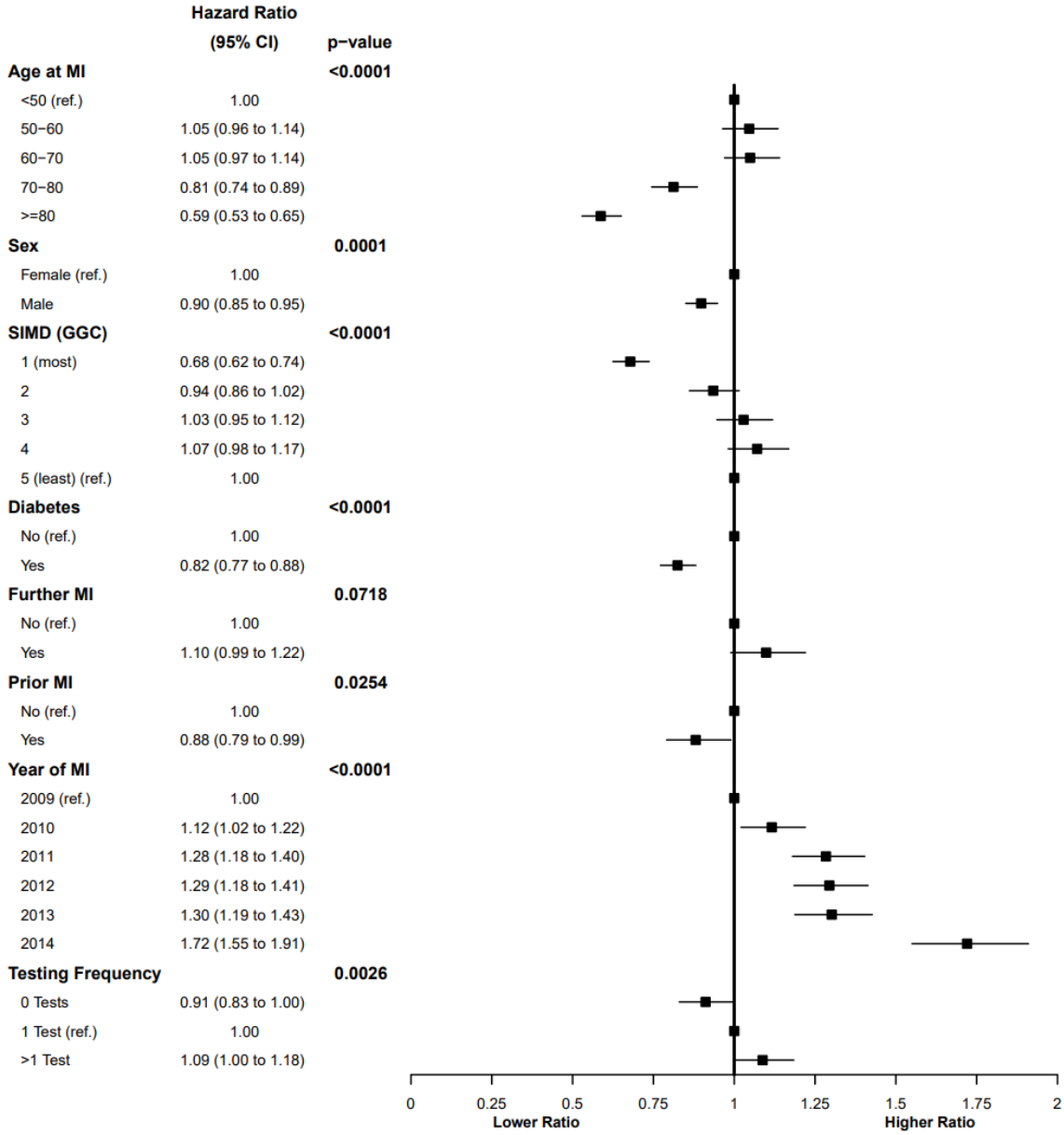


Figure 6.11: Time-Varying Cox Regression for Time to Non-Target LDL (>1.8mmol/l) with Lipid Testing Frequency Covariate Testing frequency was the number of lipid tests conducted in the previous year of follow up, with the recommended one test per year as the reference level. Testing frequency was a time-varying covariate with one test assumed for the first year of follow up. AIC for this model was 102802.8. Patients only included if ever prescribed statins during follow up.

Patients were only considered in these models (and subsequent analyses in the remainder of this chapter for non-target LDL) if they were prescribed statins at any point during their follow up. The optimal model for time to LDL>1.8mmol/l contained all the variables considered: age at MI, sex, deprivation, diagnosis of diabetes (time-varying), further MI in the last year (time-varying), prior MI to the baseline, and year of MI. The relationship with age was almost an inverse to that which had been observed with adherence, with those over the age of 70 increasingly less likely to become off-target, and those under 70 years more likely. However, consistent with the earlier adherence models, males were less likely to have non-target LDL than females. There was also a curved association with deprivation, where those residing in the most deprived areas were less likely to have LDL>1.8mmol/l. A diagnosis of diabetes and experiencing an MI before their baseline were both individually also associated with a decreased likelihood of not meeting the LDL target. A further MI in the last year, however, increased the likelihood of not meeting this target, although this did not reach statistical significance despite its inclusion in the optimal model. Finally, those with more recent MIs were more likely to have non-target LDL. These patterns remained consistent, although with slightly greater uncertainty, in a sensitivity analysis where time to non-target was measured from the time the LDL target was first achieved by the patient during follow up (results not presented). In a separate sensitivity analysis (not presented), the removal of the year of MI from the final model resulted in the hazard ratio for further MI in the last year increasing marginally and consequently reaching statistical significance, whilst the effects of all other covariates were not substantially altered.

The final model is shown in Figure 6.11, with full details provided in Appendix U. As was the case with the adherence models, the addition of the testing frequency as a time-varying covariate did not significantly alter any of these associations but did improve the fit of the model very slightly (AIC decreased from 102810.7 to 102802.8). Testing frequency in the previous year was significantly associated with time to not meeting the LDL target, with those over-tested more likely to not meet the target, and those under-tested less likely.

6.7 Latent Class Analysis for Target LDL Achievement

6.7.1 Optimal Model for Target LDL Achievement

The LCA model for target LDL of ≤1.8mmol/l used data from 16,215 time windows from 6,274 patients. To determine class membership, the following variables were considered: age at MI, sex, deprivation quintile, whether a patient had experienced an MI before their baseline, the year of their baseline MI, whether the patient's adherence was ≥50% the previous year, and whether the patient's LDL was ≤1.8mmol/l the previous year. Analyses were run using adherence in the previous year with both thresholds, with nearly identical models produced for both thresholds. The LCA using the 50% adherence threshold for the previous year demonstrated marginally superior diagnostic value over the one using the 80% threshold, so is included here, with full details for both optimal models in Appendix V. The optimal model (using the previous 50% adherence), which converged after 144 iterations, had two classes with predicted class shares of 43.2% and 56.8% respectively.

As expected, the largest differences in the conditional class probabilities were between whether the target of ≤1.8mmol/l had been met the previous year, with a high proportion achieving this in class 1, and a lower proportion in class 2. There was also a higher proportion of patients with adherence ≥50% in the previous year in class 1. There were also marginally higher proportions of patients from the most deprived quintile, with a diagnosis of diabetes, further MIs in the last year, and MIs before their baseline in class 1. In class 2, there was a slightly higher number of females and those under the age of 70 years than in class 1.

Table 6.16: Predicted Class Assignments for Optimal LDL Target LCA Model by Achievement of LDL Target

	Meet	Not Meet	Total
Class 1 (More)	7009	0	7009
Class 2 (Less)	1397	7809	9206
Total	8406	7809	16215

Time windows assigned to classes using predicted modal assignment. Only time windows with complete data available included in the LCA model. Likelihood of classes to meet the LDL target of ≤1.8mmol/l was ascertained through multinomial regression coefficients produced during a one-step model fitting process.

Comparing the predicted class assignments with whether an LDL≤1.8mmol/l was achieved, showed that all patients assigned to class 1 did so (Table 6.16). This resulted in 100% positive predictive value, and 100% specificity as all of the time windows where the target was not met were assigned to class 2 (Table 6.17). However, around 1,400 time windows assigned to class 2 also met the target. Nevertheless, this model still results in reasonable sensitivity and negative predictive value, with 83.4% of all time windows meeting the target identified, and 84.8% of time windows assigned to class 2 failing to achieve the target of ≤1.8mmol/l.

Sensitivity was the only diagnostic measure affected by the bootstrapping of this model, and increased by 0.4% to 83.8%, although its associated confidence interval contained the value observed in the original model. The estimated negative predictive value remained the same following bootstrapping, and for reasons expanded on in Chapter 6.2.4, the 100% values for both specificity and positive predictive value were also unaffected during the bootstrapping.

Table 6.17: Diagnostic Table for Optimal LCA Predictions for Target LDL

	1 vs 2	
	LCA	**Bootstrap**
Sensitivity	83.4%	83.8% (83.1%, 84.5%)
Specificity	100.0%	100.0% (100.0%, 100.0%)
Positive Predictive Value	100.0%	100.0% (100.0%, 100.0%)
Negative Predictive Value	84.8%	84.8% (84.0%, 85.6%)

LCA = Latent Class Analysis; CI = Confidence Interval. LCA diagnostics calculated using numbers presented in Table 6.16. Diagnostics were bootstrapped 5,000 times with confidence intervals ascertained using the percentile method. Within each bootstrap, one time window per patient and its predicted class was selected, and cross-tabulations and diagnostic measures were calculated as on the full set of time windows above.

Table 6.18: Average Posterior Probabilities for Optimal LCA for Target LDL

Class	Ave PP
1	0.9322
2	0.9431

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all time windows modally assigned to that class (i.e. only includes the time windows where the posterior probability was highest for that class).

The distinction between classes, as measured by relative entropy, was higher than had been observed within the adherence LCA models, with a value of 0.7486. Furthermore, the average posterior probabilities (shown in Table 6.18)

were above 0.9 for both classes. This suggested that the two classes were well separated from each other due to the high levels of certainty in the assignment of time windows to classes.

6.7.2 Mortality by Predicted Class

The results for the associations between mortality and assigned class in the second and third years of follow up were very similar (Table 6.19) and were similar when the LCA model using previous adherence at the 80% threshold was used (Appendix X). Contrary to expectations, when comparing the raw proportions, there was a higher rate of mortality in patients assigned to class 1 (the most likely to meet the target) than in class 2 (the least likely). This was then also reflected in the hazard ratios from the Cox proportional hazards models, respectively. Specifically, in the second year of follow up, there was a 28% reduction in mortality for those in class 2 compared to class 1, and in the third year, this reduction was 27%. In both instances, however, this reduction was not statistically significant at the 5% level. Full details of these models are provided in Appendix W.

Table 6.19: Mortality by Predicted LDL Class (from Optimal LCA Model)

	Year 2		Year 3	
Classes	**1** (Most)	**2** (Least)	**1** (Most)	**2** (Least)
N	2,412	2,332	1,682	1,940
Events (%)	84 (3.5%)	59 (2.5%)	46 (2.7%)	39 (2.0%)
Cox PH				
Hazard Ratio	1.00 (ref.)	0.72	1.00 (ref.)	0.73
95% CI	--	(0.52, 1.01)	--	(0.48, 1.12)
		***p-value:* 0.0527**		***p-value:* 0.1500**

CI = Confidence Interval, PH = Proportional Hazards, ref = reference category. Patients only included if a predicted class assignment was available for that year of follow up from the target LDL LCA model. Models fitted are unadjusted for further covariates.

However, this increased risk of mortality may not be due to the failure of the LCA to identify those most at risk. In years 2 and 3 of follow up, there were also a higher number of deaths among those who had achieved the target compared to those who had not. In year 2, 12 more patients died who met the LDL target than who were assigned to the class most likely to meet the target, and in year 3, this number was seven. Therefore, in these cases, if the results of the LCA

were used to make decisions for these patients, they would have been assigned to class 2 and would have received their annual plasma lipid test. However, their LDL would have been ≤1.8mmol/l, and therefore it is unlikely that any further intervention would have been made.

6.8 Latent Class Growth Analysis for Target LDL Achievement

6.8.1 Optimal Model for Target LDL Achievement

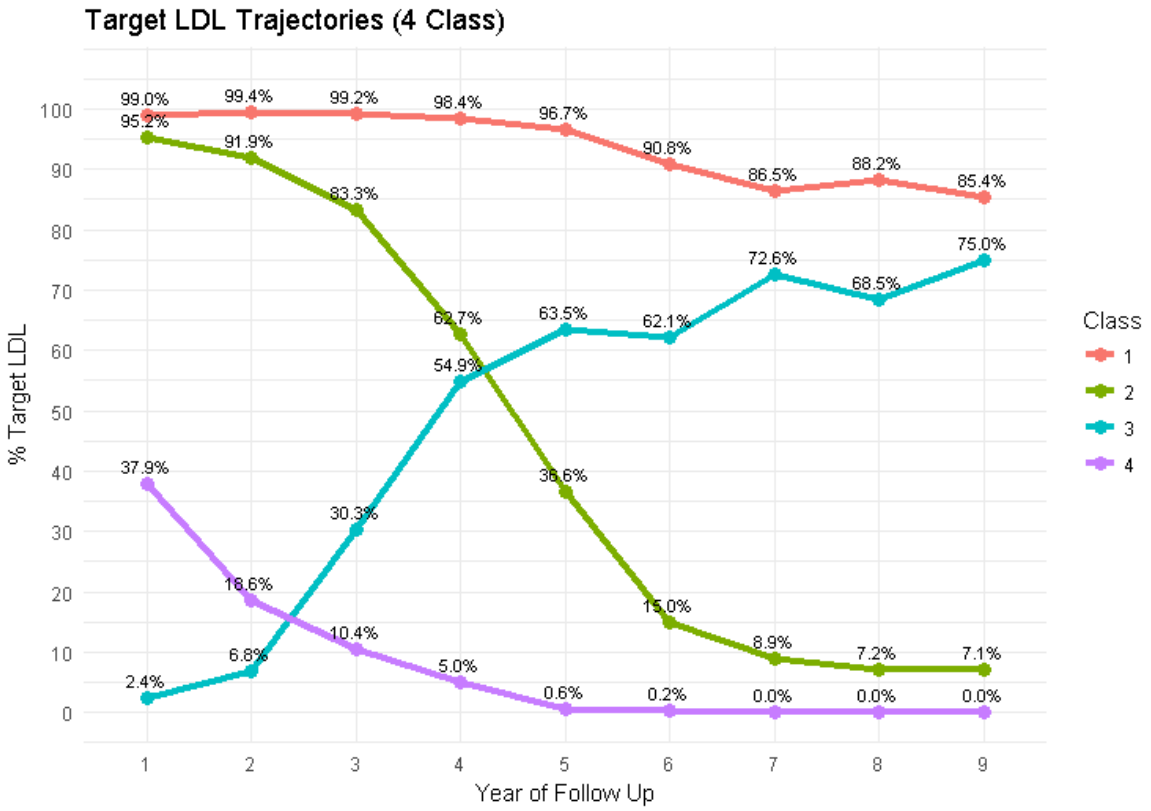


Figure 6.12: Trajectory Curves for % at Target LDL by Assigned Class from Optimal LCGA Model. Class was predicted for a patient using modal assignment. Percentages calculated using the number of patients with data available for that year of follow up. As patients may not be tested each year of follow up, this could lead to their inclusion in non-consecutive years until the end of their follow up or death.

For the LCGA model, the variables considered for the class-membership component of the model were those captured at baseline and identified earlier in the chapter as part of the optimal time-varying cox model for time to non-target LDL (Chapter 6.6.2). These were age at MI, sex, deprivation quintile, MI before baseline, and year of MI. The model was generated using data from 9,079 patients, and as with the adherence LCGA models, the optimal model contained four classes and had similar trajectory patterns (as shown in Figure 6.12).

However, there was little to no association between a patient's adherence trajectory and their target LDL trajectory. Furthermore, the estimated class shares were also very different, with just 17% assigned to class 1, and 37%, 14%, and 32% assigned to class 2, 3, and 4 respectively. The full details for this model are available in Appendix Y, including a comparison of model diagnostics for 2, 3, and 4 class models.

The two classes with declining proportions of patients achieving the LDL target over follow up (classes 2 and 4) have some similar characteristics. Both have higher proportions of females and higher proportions of patients under the age of 60 years at baseline, with class 4 (the least on target group) having the highest of both. Those assigned to classes 2 and 4 were also more likely to have experienced their baseline MIs later in the period of interest. Indeed, over one-fifth of those assigned to class 4 had their baseline MI in 2014, and for both classes more than half of the patients had their baseline MI in or after 2012. However, one of the biggest differences between these two groups was in terms of their deprivation profile. Those assigned to class 2 (high percentage on target at the start) had higher proportions of patients in the two most deprived quintiles, whereas those assigned to class 4 (a third on target at the start) were more likely to reside in the least deprived quintile (of all four classes). Patients in class 4 were also the least likely to have experienced an MI before baseline across the four groups.

In contrast, the group with a consistently high percentage of patients on target (class 1) had the highest percentage of patients with MIs before their baseline, over double that of class 4. This group were also more likely to be male and were older with a high proportion of patients over the age of 70 at the time of their baseline admission. They were also more likely to reside in the most deprived quintile and would have a longer average duration of follow up with high numbers of baseline MIs occurring in 2009 and 2010. Class 3 (whose percentage on target started low and increased substantially) shares some similar characteristics with this class; namely, higher proportions of males, and earlier baseline years, with many admissions occurring in 2010 and 2011. However, whilst they were likely to be older than those assigned to classes 2 and 4, they were slightly younger than class 1 with a quarter of the class aged

between 60 and 70 years. They were also less deprived with the majority of patients residing in the second, third, and fourth quintiles of deprivation, and had lower rates of MIs before their baseline admission.

Table 6.20: Average Posterior Probabilities for Optimal LCGA for 80% Adherence

Class	Ave PP
1	0.7115
2	0.6634
3	0.7719
4	0.7584

Ave PP = Average Posterior Probability. Ave PP for a class is the mean of the posterior class probabilities for all patients modally assigned to that class (i.e. only includes the patients where the posterior probability was highest for that class).

However, the relative entropy for this model was lower than the LCGA models for adherence, with a value of 0.5257. This would suggest that there was a limited distinction between the classes within the model and consequently there may be an increased risk of the misclassification of patients. This was also supported by the average posterior probabilities (Table 6.20) where three out of the four classes were above 0.7. The lowest of these probabilities, 0.6634, was for the group of patients who were predominantly on target at the start of follow up before declining substantially in likelihood (class 2). As this is below 0.7, this could be an indicator that there may be greater uncertainty surrounding the correct allocation of patients to this group.

There was also reasonable sensitivity to mortality within the classes in this model. This could, in part, be due to the inclusion of year of MI within the class-membership model with higher rates of mortality likely in classes with earlier MI dates. Nonetheless, in a sensitivity analysis where patients who died during follow up were excluded, the trajectories remained similar for the four classes, but 18.7% of included patients were assigned to a different class.

6.8.2 Mortality by Predicted Class

Figure 6.13 compares the overall survival probabilities for the classes from the LCGA model for target LDL trajectories. Patients assigned to class 4, which had the worst target LDL trajectory, had the highest survival probability throughout follow up. This was consistent with class 2, where those assigned to this other

decreasing trajectory had the next highest survival probabilities. This was in contrast to those assigned to the consistently on target class (class 1) and the increasingly on target class (class 3), which both had similar survival probabilities throughout follow up and the poorest survival curves of the classes assigned in this LCGA model.

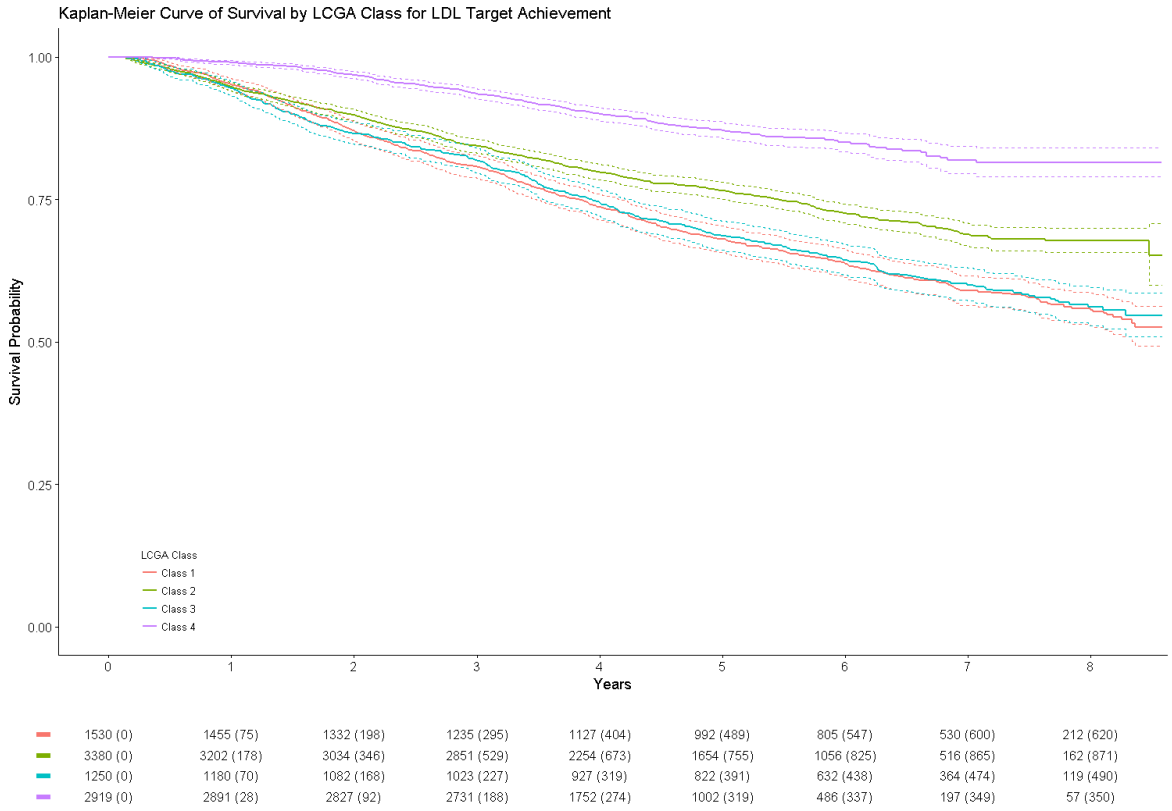


Figure 6.13: Kaplan-Meier Curve for Mortality by Predicted 50% Adherence LCGA Class. Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the target LDL LCGA model. Numbers in table are number at risk (cumulative number of events).

These patterns were also evident in the hazard ratios from a Cox regression model looking at the association between these assigned classes and their association with mortality. In this model, where the hazard ratios are presented in Table 6.21 and the full details are available in Appendix Z, the latent classes were found to be statistically significant in the risk of mortality at the 5% level.

As with the adherence LCGA models, patients assigned to the consistently on target group (class 1) were considered to be the reference group. Patients assigned to classes 2 and 4 (the two decreasingly on target groups) had the lower associated risks of mortality, compared to the reference group (29% lower and 65% lower respectively). In line with expectations from the earlier Kaplan-Meier

curve, the associated hazard of mortality from this Cox regression model for class 3 (increasingly likely to meet target) was similar to that of class 1. In particular, the associated hazard ratio was 0.98, conferring a nonsignificant 2% reduced likelihood of mortality for patients in class 3.

Table 6.21: Hazard Ratios of Mortality by Predicted Target LDL LCGA Class

	HR (95% CI)	p-value
Class 1 (ref)	1.00	<0.0001
Class 2	0.71 (0.64, 0.79)	
Class 3	0.98 (0.87, 1.10)	
Class 4	0.35 (0.30, 0.40)	

HR = Hazard Ratio, CI = Confidence Interval. Class was predicted for patients using modal assignment. Patients only included if a predicted class assignment was available from the target LDL LCGA model. Models fitted are Cox proportional hazards and were unadjusted for further covariates.

6.9 Discussion

6.9.1 Identification of Potential Patients for Reduced Lipid Monitoring

Within this post MI cohort from NHS GGC, the analysis in this chapter has shown that simple clinical and demographic factors, as well as the last year's adherence and lipid data, are associated with future adherence and meeting lipid targets. Using this routine clinical information, it is possible to identify patients for whom an annual lipid test could be considered unnecessary. All patients predicted to meet the 50% adherence threshold in a given time window (accounting for 85% of the years of follow up available) did so, and all time windows where patients were expected to meet the LDL target of ≤1.8mmol/l (43% of all years available) also did so. For adherence, this was supported further by survival models where those predicted to not have adherence ≥50% in the second and third years of follow up had a higher risk of mortality, with smaller effect sizes observed for adherence ≥80% adherence, giving these findings face validity. Furthermore, using patient trajectories identified in LCGA, patients assigned to the class with consistent adherence throughout follow up also had optimal survival probabilities. This analysis raises the possibility that annual lipid monitoring could be targeted in secondary prevention, potentially saving GP time, lab costs, and patient time in a sizable group of the most adherent patients.

To date, attempts to personalise testing schedules within chronic diseases have been limited to screening programmes outside of cardiovascular disease, thus emphasising the novelty of this approach. A systematic review published in 2016 seeking to evaluate the evidence for such risk-based approaches in the use of diabetic retinopathy screening found that RCTs were needed to draw reliable conclusions, whilst others have raised concerns about the possibilities of missed interventions should the interval between checks be extended significantly (McGhee, Harding and Wong, 2012; Lund *et al.*, 2016; Taylor-Phillips *et al.*, 2016). However, before such trials can take place to test reduced monitoring, models first need to exist to identify such patients. This was the case for one such risk score derived from data in Liverpool using electronic records, that has now been tested in an RCT and safely reduced the number of appointments needed by 40% with comparable risks between routine annual monitoring (Eleuteri *et al.*, 2017; Broadbent *et al.*, 2020). The analysis in this chapter, therefore, sought only to provide some foundations towards any implementation of such personalisation.

One of the main considerations in putting this data into context is what annual lipid tests are trying to achieve: adherence or cholesterol levels. The current SIGN guidelines do not specify a lipid target. Instead, it is stated that a reduction of 1 mmol/l or a 40% reduction in non-HDL would suggest adequate adherence and lifestyle changes following commencement of statin therapy (SIGN, 2017, p. 44), suggesting that adherence may be the primary purpose of such a test. Throughout this thesis, two adherence thresholds have been considered, 50% and 80%. Of these, the 50% adherence threshold, equivalent to alternate-day dosing, has been shown in this chapter to be relatively easy to predict and can be done so with high accuracy. Furthermore, in clinical practice, such a treatment regimen may be recommended in patients who experience side effects of statin treatment, as the duration of the lipid-lowering effect exceeds its half-life (Jordanov and Abou Assi, 2015). Regardless, using adherence as a marker would also yield practical benefits to the identification of patients whose test may be unnecessary as if this were ever to be implemented, monitoring lipid targets among patients who do not have tests would be difficult to achieve.

Another limitation of clinical use of lipid targets in this context is the direction of the, albeit not statistically significant, associations between predicted LCA LDL classes and survival. Higher rates of mortality were also observed in those that had met the LDL target that year. Consequently, the results of the LCA are not suggestive of an increased risk of mortality due to the predictions themselves. Indeed, a smaller proportion of patients died who were predicted to meet the target than those who had met the target in both the second and third years of follow up. This was also supported by the results of the LCGA for LDL, where the classes with the lowest proportions of patients achieving the LDL targets having the greatest survival probabilities. One explanation could lie in the causes of these deaths, if such causes result in lower cholesterol levels, although the small numbers in the groups with the cause of death information available prevented such comparisons being investigated. This also meant that in the case of the LCGA, repeating the Cox regression analysis by excluding those with deaths in the first two years to attempt to address issues of reverse causality was also not possible. Additionally, however, when looking at the characteristics of those predicted to meet the target in the LCA and consistently meeting it in the LCGA, it is clear that these groups have higher rates of cardiovascular risk factors such as being male, older, having a diagnosis of diabetes, and experiencing further and previous MIs. Therefore, the presence of these may have led to more intense risk management in these patients resulting in the confounding of these associations.

This confounding could also offer an insight into some of the discrepancy between adherence and the achievement of lipid targets. Previous research has shown that although statin adherence is reasonable in secondary prevention cohorts, this does not fully account for the achievement of cholesterol targets with many adherent patients not meeting the recommended LDL levels despite excellent adherence (Nordstrom *et al.*, 2015; Danese *et al.*, 2017). This has also been observed in earlier chapters, in this cohort (Chapter 4 and Chapter 5). Additionally, in this chapter there was little association between a patient's class trajectory for adherence and their class trajectory for achieving lipid targets, suggesting that these two outcomes are capturing different elements of a patient's risk. Given the lack of association between predicted classes and mortality, this further enforces the idea that the use of lipid targets in terms of

the effective identification of a suitable subgroup may be limited at a population level.

Furthermore, the time pressures on primary care physicians and the burden of risk scoring for various conditions should be recognised (Challener, Prokop and Abu-Saleh, 2019). The identification process developed in this chapter is not a manual risk scoring process, but an automated system that could be implemented using EMIS or VISION software in primary care, incorporating encashed prescription data. This means GPs would have real-time data informing them whether a lipid test is required or not as part of the visit, although they would also be able to request tests if they feel this is indicated. This may be particularly relevant in the context of other comorbidities that were not captured within the current cohort. Indeed, a diabetes diagnosis was only identified reliably using the diabetes register, but other comorbidities which may impact a patient's ability to adhere to medication, such as dementia (Smith *et al.*, 2017), were not routinely or reliably captured in GP local enhanced services data and were therefore not incorporated in the present analysis. Nevertheless, the additional burden required to implement this approach should be viewed in the context of health care resources saved through performing unnecessary tests.

6.9.2 Strengths and Limitations

External validation of these results is required. Nonetheless, given that the demographics of this cohort have been shown to broadly in line with other similar cohorts within the literature (Chapter 4), it is anticipated that this subgroup of patients is also likely to be present and reasonably accurately identified. However, the exact size of this subgroup may differ, as rates of adherence and LDL target achievement were slightly higher within this cohort than in other cohorts.

One limitation of the initial Cox regression models to identify the factors associated with non-adherence or non-achievement of the LDL target was that no interactions between the covariates were considered. For example, a significant interaction between sex and age was likely, as females experience their MIs at an older age (Mosca, Barrett-Connor and Kass Wenger, 2011) and are

less likely to be prescribed statins (Eindhoven *et al.*, 2018). This could have partly explained the increased likelihood of non-adherence at an older age observed within this cohort and the literature (Hope *et al.*, 2019). Therefore, stratification by sex could prove beneficial in future research into factors associated with non-adherence and non-target lipids. However, this should not have substantially impacted the results of the LCA or LCGA as these methods are well suited to data where high-order interactions are likely to be present (Vermunt and Magidson, 2003).

Nevertheless, LCA models were only run for years of follow up where all data was available for the necessary covariates, and LCGA was only run in patients where baseline covariates were available, and the outcome data was captured at least once during the patient's follow up. For LCA, this may not be a problem if the patterns of missing data within patients' follow up in this cohort are representative and similar to other cohorts.

Moreover, for the LCA models, the independence assumption of observations was violated as multiple years of follow up were included per patient. Despite this though, internal bootstrapping of the model predictions did little to alter the model's diagnostic accuracy measures and other measures of fit suggested that the overall fit of the final models was reasonable. However, methods which allow for changing covariates (which LCGA does not), such as Latent Transition Analysis should be considered in future research into this area.

In the LCGA models, missing data is assumed to be missing at random by the model, which may not have been the case in these data. For example, the percentage who were adherent may increase in a class due to patients who are non-adherent being more likely to die earlier within the follow-up period. To assess this possibility, sensitivity analyses excluding patients who died during follow up were conducted. These found similar trajectories of patient groups for all models, although with the exception of the 50% adherence model where less than 1% changed class, around a fifth of patients were reassigned, suggesting there may be some influence of mortality here. This may be explained by the inclusion of the year of MI within the class membership component of the model, with those experiencing later baseline MIs less likely to experience mortality before the end of follow up. Given that removal of the year of MI did not

substantially impact the associations observed between the other covariates and time to non-adherence or non-target LDL in the Cox regression, further analysis without year of MI could be considered, but other methods, such as joint modelling, may be needed to fully account for this.

Furthermore, despite the low percentage change in the 50% adherence model, when survival among the classes was compared, one class (class 2) did have an unusual survival curve with no events in the first two years, resulting in a violation of the proportional hazards assumption in the subsequent Cox model implemented. The reasons for this pattern were unclear, but the small class size means that this is likely the result of clinical differences in baseline health arising through chance, and the relatively small differences in class characteristics would make it difficult to identify such a group in other cohorts. Indeed, the relatively small differences in the conditional class probabilities, particularly in the LCGA, may also make the accurate allocation of patients to latent groups more challenging when validating results in other cohorts.

As indicated above, a key strength of this analysis is that to date and our knowledge, no previous research has attempted to identify a group of patients who may not require an annual lipid test as part of their annual review within a cardiovascular prevention setting. Although the exact purpose of a lipid test in this context remains unclear, if its purpose is to see if sufficiently reduced lipids have been achieved then using the LCA predictions, potentially 40% of tests might not be needed, and if it is to ascertain a patient's adherence, this could be as high as 85% if a 50% threshold is deemed adequate. However, it is important to emphasise that this approach refers only to the lipid test component of a patient's annual review. Other risk management tests, such as for blood pressure and diabetes, that form part of the review would still be needed, and therefore, attendance at appointments should still be required. Furthermore, as no lipid tests in the year before was associated with an increased likelihood of becoming non-adherent, it is unclear whether it is the engagement with health services that sustains adherence or vice versa, and therefore the act of attending an appointment may have an important role in CVD risk management. Nevertheless, lowering the number of lipid tests

conducted as part of these annual reviews would still have an impact on the financial burden and primary care workload.

6.9.3 Conclusions

In this chapter, analyses have identified factors associated with becoming non-adherent or having non-target LDL and used these to identify a group of patients within the post MI population who could be considered for reduced lipid testing with reasonable levels of accuracy. Additionally, the associations between predicted classes for adherence using the 50% threshold and mortality are similar to those observed between adherence and mortality in the previous chapters, suggesting that such a prediction is likely to be safe. No significant associations were observed between class predictions for achieving the LDL target of ≤1.8mmol/l and mortality, although the hazard ratio was suggestive of an increased risk for those on target. However, as this was also present amongst those achieving the target, this indicated that the predictions did not pose an additional risk to the patient. Nevertheless, the associations with adherence and their face validity, and the lack of an association between those predicted to meet the LDL target and those predicted to be adherent, suggested that the prediction of a patient's adherence alone could be sufficient to ascertain whether a lipid test is required when considering patient outcomes.

However, whilst this is encouraging, further research is needed to externally validate these findings in other populations with established cardiovascular disease and to evaluate the impact. Additionally, other statistical models may be needed to improve the robustness of these predictions to increase the certainty in the accuracy of the identification process. Nonetheless, this chapter has demonstrated that, in principle, it could be possible to personalise the frequency of lipid testing using routinely collected data already available in patients' electronic health records.

Chapter 7 Final Discussion

7.1 Summary of Chapters

The contents of this thesis have sought to address the aims and objectives outlined in Chapter 1.8. In Chapter 2, the systematic review laid the groundwork for the subsequent chapters by presenting and comparing the international guidelines for lipid-lowering in secondary prevention populations. Through this, it was clear that although the evidence base for the use of statins in this population was well-established, this was not the case for the use of lipid targets or the frequency that such patients should be monitored. For the former, targets were often stated with values derived from those used within clinical trials for lipid-lowering therapies, and in the latter, any recommendations were often supported only by the clinical opinion of the creators (Brown, Welsh and Logue, 2020). This lays the groundwork for epidemiological work, exploring whether existing opinion-driven clinical targets are frequently achieved, whether achieving targets has important consequences for clinical outcomes, and whether continuous monitoring of targets on an annual basis is necessary for all patients.

In the remainder of the thesis, data from the NHS GGC Safe Haven was used, with Chapter 3 looking at the derivation of the post MI population within that region and subsequent chapters focussing on the remaining aims of the thesis. Specifically, this population contained all patients in NHS GGC who experienced a non-fatal MI between 2009 and 2014. This is therefore a locally representative population of post MI patients, and inferences about the prevalence of meeting or missing targets are likely to be more widely valid. Patients were followed up until July 2017 or death, and data captured included baseline demographics, hospital admissions, death certificates, blood test results, diabetes status (from the diabetes register), and prescribing records for lipid-lowering therapies.

Following this derivation, the second aim sought to understand the demographics, adherence, and achievement of cholesterol-lowering targets of this population was addressed in Chapter 4. This found that the demographics of the 11,110 patients included in the post MI population were broadly similar to those observed in other cohorts outlined in the literature, with patients more

likely to be male, reside in more deprived areas, and had an average age at the time of MI of 67 years. Furthermore, the prevalence of diabetes within this population was also similar to other secondary prevention cohorts which focussed on patients admitted with acute cardiovascular events. This further supports the idea that these patients are representative of those seen in wider clinical practice. The size of the cohort, duration of follow up, and the use of real-world lipid results in repeated annual time windows, are less common in the literature, and represent significant strengths of the dataset.

In terms of average statin adherence, there was a higher proportion of patients with adherence <80%, and a lower proportion with adherence <50%, when compared to other similar cohorts. Length of follow up was likely to have played some role in the former, with adherence known to decrease as the duration increased (Naderi, Bestwick and Wald, 2012; Khunti *et al.*, 2018), and the high-risk nature of the cohort and the lack of prescription charge also likely to increase adherence in the case of the latter (Naderi, Bestwick and Wald, 2012; Leslie, McCowan and Pell, 2018). Achievement of lipid targets, both the 2016 ESC target of LDL≤1.8mmol/l (Piepoli *et al.*, 2016) and the 2014 NICE target of ≥40% reduction in non-HDL (NICE, 2014), was higher than observed in other cohorts, although achievement in this instance was defined by whether the target was achieved at any point during follow up, rather than a specific time following an event or statin initiation.

Associations between statin adherence, achievements of cholesterol targets, further MIs and mortality were explored in Chapter 5 and found that non-adherence and failure to achieve lipid targets were strongly associated with all-cause mortality within this population and similar patterns were also observed for death due to circulatory causes. However, their associations with further MIs resulting in hospitalisation were not significant in either direction, although confounding and temporal elements could explain this uncertainty. Indeed, adherence was calculated as an average over the duration of follow up (of both before and after further events), and patients were categorised as achieving lipid targets if they did so at any point during their follow up. Nevertheless, an association between average statin adherence and the achievement of lipid targets was also found, expanding on the associations with mortality and in line

with the mechanism of action of statins. However, as observed in other cohorts, adherence could not fully account for the achievement of lipid targets with not all adherent patients achieving lipid targets, and vice versa. Since adherence itself has a clear association with hard clinical outcomes, and those who achieve at least 50% adherence are significantly more likely to meet lipid targets, it is perhaps valid to query whether continuous monitoring of lipid targets is necessary in all patients. For example, for patients who have previously met the target, and continue to be adherent, there may not be any changes to clinical decisions that could be made with further lipid measurements.

The final two aims were investigated in Chapter 6. First, associations between demographic factors, previous adherence, and previous lipid results were examined. Few demographic factors were significantly associated with adherence, whilst for the LDL target of ≤1.8mmol/l, all demographic factors considered were included in the final model. With these factors determined, the final aim was addressed. This sought to identify patients who would not require any intervention in response to an annual lipid test and could therefore be tested less frequently. This process identified a significant proportion of patients assigned to classes most likely to meet the LDL target who were also extremely likely to have met the target, using latent class analysis. When using the 50% adherence threshold, similar accuracy was also demonstrated (both had 100% PPV) with higher sensitivity than observed with LDL models (98.6% vs 83.4%, respectively). Furthermore, there were no significant associations between predicted LDL classes and mortality, but strong associations were observed between those predicted to have adherence ≥50% and mortality, and therefore adherence may make an effective and practical marker for accurately identifying patients for reduced lipid testing.

7.2 Strengths and Limitations

One strength of this thesis is the data used in Chapters 3-6, which is from a large, contemporary, real-world cohort with numerous variables captured through patient interactions with all aspects of the NHS in the GGC area, including hospital admissions, lab results, and dispensed prescriptions. Following the derivation of the cohort, the description of the characteristics of this population allowed comparisons to be drawn with other similar cohorts. The

similarities of this cohort to other observational secondary prevention cohorts in the literature, as discussed in previous chapters, provide some external validation to these results. Nonetheless, further validation is required, particularly for the novel findings within Chapter 6 before widespread implementation is considered in clinical practice. For instance, CPRD data could be used to validate and expand on the results from Chapter 6.

Additionally, this cohort could also be used to demonstrate the extent to which the current, and previous, Scottish guidelines (SIGN, 2007, 2017) were being adhered to regarding the frequency of lipid monitoring. In particular, this highlighted a higher rate of testing in the year immediately following a non-fatal MI, which is likely to be the result of the initiation or modification of lipid-lowering therapy prescribed (Danese *et al.*, 2017). Additionally, only half of the patients with data available at each year of follow up had the recommended one test per year, and the proportion who received no tests in each year of follow up increased as the time from baseline MI also increased. However, whilst there was also some decrease in the proportion adherent as follow up increased, this was to a lesser extent, suggesting that some patients may have simply stopped attending reviews only.

Another possible explanation is that patient's annual reviews were still occurring, but a lipid test was not taken during them, and therefore decisions regarding the necessity of the test are already taking place in clinical practice based on expert opinion. This could not be verified within the variables captured within the data, as the presence of a lipid test result was the only indicator that such an annual review had taken place. Furthermore, not all of the variables contained within the original extracted data were routinely captured and consequently variables that were included had to be selected carefully. For example, all blood test results are automatically included in SCI store from the lab systems, meaning all tests that took place were recorded. However, blood pressure, another important cardiovascular risk factor that is likely to be checked at an annual review, may have been checked by the GP at a review, but was not consistently recorded into the enhanced services data for the entire post MI population. Similarly, with the exception of a diabetes diagnosis which could be reliably derived through the diabetes register, a patient's comorbidities

could not be reliably attained within the data available. These comorbidities, such as severe mental illness and dementia, could have a significant impact on the patient's cardiovascular risk directly and on the patient's ability to adhere to regular medications (Hippisley-Cox, Coupland and Brindle, 2017; Smith *et al.*, 2017; Hope *et al.*, 2019). Therefore, the scope of this project was limited to only the lipid test component of the annual review, rather than the review itself, with a view that a GP could still request the test if they believed it was indicated. Nevertheless, reducing the number of tests safely using the methods identified in this thesis could still be associated with a significant reduction in financial costs, as discussed below (Chapter 7.3).

The analytical methods employed on this observational real-world cohort means that it is difficult to draw causal inferences regarding the associations presented in this thesis, as has been discussed throughout. Both a patient's adherence and plasma lipid levels are influenced by a plethora of factors, and the plausible and established interactions between these and each other have likely resulted in confounding and mediating in the various relationships that have been presented in this thesis. For example, a patient's LDL cholesterol level is likely to be a mediator in the relationship between adherence and mortality, and consequently may have accounted for a non-negligible proportion of this observed relationship.

7.3 Application to Clinical Practice and Next Steps

The results in this thesis have shown that it is likely that a reasonable proportion of lipid tests in any given year would not require intervention in response to it, with the exact proportion dependent on whether the test's purpose is to assess lipids or adherence. For reasons discussed in Chapter 6, adherence, specifically using a 50% threshold, may be the preferred marker for a test's necessity, and in this population, its use translated to around 85% of tests being considered unnecessary. However, annual reviews of such patients are likely to involve more than simply a blood test, and therefore, these results do not suggest that such a review may also be always unnecessary. Nevertheless, removing the need for a blood test may facilitate a transition, which has been accelerated by the COVID-19 pandemic, to more virtual and remote reviews of patients, especially

as adherence could be monitored using prescribing records and home blood pressure monitoring has become more widely available.

However, reducing the number of tests conducted will still reduce the burden of biochemistry tests and the time required for processing the samples and results by healthcare professionals. As an approximation, using figures from the 2018-19 NHS National Cost Collection figures for England, the average cost of an individual lipid test (including phlebotomy) is £5 (National Health Service, 2019a), and with an estimated 1.4 million survivors of MI in the UK each year (British Heart Foundation, 2020b) who should be receiving an annual lipid test under current guidance (NICE, 2014; SIGN, 2017), this would be equivalent to a saving of £5.95million per year. In addition, it is possible that repeat appointments might be avoided, if for instance, a separate appointment is needed for the test, and another for the interpretation of the results and other CVD risk factors. If such repeat appointments, which are estimated to cost £30 each (National Health Service, 2019b), can be avoided, then this would translate into a further savings of £35.7million. However, these are clearly illustrative costs and are solely from the NHS perspective. Therefore the full impact of these results, including opportunity costs whereby more time would be available to focus on those patients most at risk, would need to be considered in a full cost-effectiveness modelling study.

Moreover, before implementation, several steps would be needed to confirm that this identification process is safe and cost-effective. Consequently, the analysis contained within this thesis provides only the groundwork to assess the potential viability that such a group could be identified. The first step towards implementation, therefore, is to validate these findings within other similar datasets containing survivors of MIs. Furthermore, as the post MI population accounts for only a proportion of the secondary prevention population, validation or the derivation of similar models would be needed in other subgroups to increase the overall cost-effectiveness benefits of reduced testing. A similar approach could also be considered in primary prevention populations, although some differences in population characteristics could result in reduced accuracy using the models from this thesis, and therefore separate models may need to be developed. Moreover, the proportion whose tests could be considered

unnecessary may well be smaller in this population as adherence has been shown to be lower in primary prevention (Naderi, Bestwick and Wald, 2012).

Nevertheless, whilst validation in other observational cohorts is important, cost-effectiveness studies and the safety of this approach should be formally tested in an RCT, before clinical implementation is considered. One approach to the latter could be similar to that used in the area of diabetic retinopathy, which has tested a risk score developed using the 12,000 patients from the Liverpool Diabetic Eye Screening Programme in a two-arm, parallel-assignment equivalence RCT (Eleuteri *et al.*, 2017; Broadbent *et al.*, 2020). This trial, which followed up patients for just over two years, randomised participants to either annual reviews as recommended by current guidelines, or to variable-interval screening using the risk calculator. Attendance at appointments and detection of advancing retinopathy rates were the primary and secondary outcomes and were formally tested using equivalence and non-inferiority testing, thereby testing the feasibility and safety of such an approach is comparable to current schedules (Broadbent *et al.*, 2019, 2020). In the case of the results presented in this thesis, possible durations of the intervals between tests that would be safe and cost-effective would also need to be determined, with biennial testing considered in the first instance to coincide with review appointments.

7.4 Final Conclusions

This thesis has found that whilst it is consistent across many guidelines, the recommendation for annual monitoring within cardiovascular disease risk management has little robust evidence to support it nor any other monitoring schedules. The remainder of the thesis has focussed on a subpopulation of the secondary prevention population, survivors of MIs, in the NHS GGC region which included data from over 11,000 patients. Descriptive statistics showed that this cohort's demographics, adherence, and lipid control was similar to other observational cohorts, and validated associations between adherence, and lipid control, with mortality. This also highlighted that although not all patients achieve sufficiently decreased lipid levels or satisfactory statin adherence, a significant proportion of patients did. Finally, this thesis identified factors associated with non-adherence and used latent class methods to accurately

identify a significant group of patients for whom an annual lipid test could be considered unnecessary.

However, whilst these latter findings do equate to large potential cost savings, it is important to emphasise that this thesis sought only to assess the potential of a personalised approach to the lipid testing component of the annual review. Consequently, the results presented will need further validation in different cohorts and other high-risk populations, and cost-effectiveness studies and more robust evidence, such as an RCT as outlined above, will be needed before implementation in clinical practice.

Appendices

Appendix A Handling of Duplicate Statin Prescriptions

Appendix A1 Method for the Extraction of Start and End Dates

This method should be run three times – firstly for the statins, then ezetimibe, then fibrates, before moving on to the final stage. Specific information for each drug class is detailed in each step. The following prescriptions should be extracted for each time of running through:

Statins	**Ezetimibe**	**Fibrates**
Atorvastatin	Ezetimibe	Bezafibrate
Fluvastatin	Simvastatin & Ezetimibe	Ciprofibrate
Pravastatin		Fenofibrate
Simvastatin		Gemfibrozil
Simvastatin & Ezetimibe		Fenofibrate/Simvastatin
Fenofibrate/Simvastatin		
Source:	**Source:**	**Source:** All except
All from Prescription data	Ezetimibe from Prescriptions	combination from Prescriptions.
	Combination from already sorted statin file	Combination from already sorted statin file.

1. Order by SafeHavenID, dispensed date, and prescribed date.
2. Calculate the prescription end date (which is the first day that the individual should have run out of drug) as the date dispensed plus the days supplied (dispensed quantity/number of doses per day). Day dispensed is used over date prescribed as this is the day that the individual first had access to the drugs.
    a. For ***statins,*** assume one dose daily. Four entries did not work due to a comma erroneously contained in the quantity variable “1,200”. Visual inspection of the data, and the other records for this individual, confirmed the correct value for this should be 200, and this was corrected.
    b. For ***ezetimibe,*** this should be done for the **ezetimibe only prescriptions** (as the combination dates will have already been adjusted for other statin courses). Assume one dose daily.
    c. For ***fibrates,*** this should be done for the **fibrate only prescriptions** (as the combination dates will have already been adjusted for other statin courses). Most fibrates are taken as one dose per day, however, ***Gemfibrozil*** is assumed to be taken three times per day, and ***Bezafibrate 200mg*** is assumed to be taken twice daily, so adjustments should be made for these medications.
3. Remove any prescriptions issued after death. Set any prescription end dates currently after the date of death to be the date of death (which assumes no medication was taken on the day of death) and adjust quantities dispensed to reflect this (= Quantity – (Prescription End Date – Updated End Date)).
4. Identify those with multiple prescriptions dispensed on the same day. Use the subsequent separate method to resolve these before continuing.
5. Identify new courses for an individual.

   a. Create variables containing previous drug (and dose) dispensed, and end date for the previous prescription, for the individual (with first entry set to NA).
   b. Create a new course indicator variable which is set to 1 (and 0, otherwise) if:
      i. The dispensed date is more than six months (182 days) since the end date of the previous prescription OR
      ii. The drug (or dose) dispensed is different from the previously dispensed drug (and dose).
6. Create a running total for each individual to assign each course and the prescriptions contained within each course a course number. Use this to create a course ID with the format: SafeHavenID – CourseNumber.
7. Identify the last prescription of a course with an indicator variable (1, if last; 0, otherwise).
8. Determine the start and end dates for each course:
   a. Use the indicator variable to extract the relevant observations into two datasets; start and end.
   b. In the start dataset, rename the dispensed date as the start date.
   c. In the end dataset, rename the prescription end date as the end date.
9. Merge the start and end dates by SafeHavenID and Course ID.
10. Order by SafeHavenID and course number.
11. Remove course overlaps:
   a. Create a variable containing the next start date (set to NA if new SafeHavenID).
   b. If the next start date is before the course end date, set end date equal to the day of next start date to prevent overlap.
12. Check no end dates for a course are prior to the same courses start date.
13. To extract all relevant prescriptions, merge start and end dates by SafeHavenID and course ID and keep if dispensed date falls within window. If prescription end date is after course end date, set prescription end date equal to course end date and adjust quantities (= Quantity – (Prescription End Date – Updated End Date)).

Appendix A2 Method for the Removal of Multiple Prescriptions for the Same Dispensed Dates

This should be performed prior to implementing steps 5 onwards in the above method, and repeated for each group of drugs.

1. Create an ID for the prescription equal to: SafeHavenID – Dispenseddate
2. Extract all IDs which have duplicates into a separate dataset, and all those without duplicates into a different dataset (which won't be required until the end).
3. Remove complete duplicates from the duplicate dataset (i.e. same prescribed date and drug and dose too)
4. Extract the now single entries from the duplicates dataset ready to be remerged and remove them from the duplicates dataset.
5. Aggregate totals for those dispensed multiple prescriptions for the same drug and dose on the same day:
   a. Extract the first entry for each batch of prescriptions (this will be the first prescribed one of the multiple prescriptions all dispensed the same day). However, as the dispensed date is the date primarily used, the selection of the entry is largely immaterial.

   b. Calculate the total quantity dispensed for each dispensed date and set the dispensed quantity equal to this for the first entries extracted. Remove the other prescriptions which the total was derived from, as this single entry will replace them all.
   c. ***(For all statins and non-combination Ezetimibe and Fibrates only)*** Recalculate the end date for the single entries with new totals and if end date is after death, set end date to date of death, and readjust quantities (= Quantity – (Prescription End Date – Updated End Date)). After implementation on statins, combination drugs should not contain duplicates, however, be sure in the Ezetimibe and Fibrate cases that the end dates for the combination drugs are not adjusted at this stage as they have already been adjusted for other statin courses.
   d. Once these single entries have been recombined with the duplicates, and entries which now have a unique prescription ID (i.e. no other drugs and doses dispensed on same day also) can be extracted and remerged with non-duplicates.
6. ***(For all statins ONLY)*** Determine drugs taken prior to and post the remaining duplicate entries:
   a. Extract all non-duplicate prescriptions for those with a duplicate entry.
   b. In the duplicate dataset, rename the dispensed date as the duplicate date, and merge the duplicate date to the data set with all prescriptions for these individuals.
   c. Split this dataset into two: Those collected prior to the duplicate date and those collected after. Remove any duplicates arisen as part of the merging process in both datasets.
   d. Keep the last prescription dispensed prior to, and the first prescription dispensed after, the duplicate date(s) in their respective datasets.
   e. Rename the drugs as the prior and post drugs respectively and merge the drug (and dose) to the duplicate dataset.
7. ***(For all statins and fibrates ONLY)*** If the drug (and dose) prescribed after the duplicate date is the same as one of the drugs dispensed in the duplicate, keep the record in the duplicate that matches the post duplicate drug and discard all others for that ID. **For fibrates,** due to the small numbers affected, this was done by hand.
8. ***(For all statins and fibrates ONLY)*** If the prior drug (and dose) does not match any of the drugs dispensed (and is not NA), but the post drug is equal to NA, then keep the different one. This rule retains some duplicates, but these can be removed at a later stage in the process. Remove all others for those prescription IDs. **For fibrates,** due to the small numbers affected, this was done by hand.
9. ***(For fibrates ONLY)*** If the prior and the post are the same, keep the different one and remove all others.
10. For those prescribed only one drug but at different doses:
   a. ***(For all statins ONLY)*** Extract those who were prescribed different doses of the same drug only. Calculate the maximum strength of the drug prescribed for each prescription ID. Keep the record which matches the max for each prescription to be remerged and remove all others (i.e. keep the highest dose). For ***combination*** drugs, as strength is captured in a different variable, implement same rule but for the correct variable.
   b. ***(For Fenofibrate only)*** Given the limited numbers, by hand:
      i. If one of the duplicates is for maximum tolerated dose 267 mg, keep this record and remove all others.

ii. If one of the duplicates is for 160 mg dose (different release mechanism), keep the 67 mg dose and discard all others.
iii. No other duplicates remained in this data.

c. ***(For Gemfibrozil only)*** Take the record corresponding to the maximum tolerated dose (600 mg) and discard all others.
d. For ***Bezafibrate*** and ***Ciprofibrate*** a similar approach should be taken but was not the case in this data.

11. For remaining duplicates, a hierarchy is established with regards to which records should be kept.
    a. For ***statins,*** the basic idea is that the highest dose of the highest intensity statin should be kept, except in cases of Fluvastatin and Pravastatin. The basic rules were shown below and should be implemented in this order to avoid incorrect selection:
        i. If one of the duplicates is Fluvastatin, keep the highest dose Fluvastatin and discard all others from that duplicate date.
        ii. If one of the duplicates is Pravastatin, keep the highest dose Pravastatin and discard all others from that duplicate date.
        iii. If one of the duplicates is Rosuvastatin, keep the highest dose Rosuvastatin and discard all others from that duplicate date.
        iv. If one of the duplicates is Atorvastatin, keep the highest dose Atorvastatin and discard all others from that duplicate date.
        v. If one of the duplicates is a Simvastatin combination, keep the highest dose of Simvastatin (whether in combination or not) and discard all others from that duplicate date. If simvastatin dose is the same, keep the combination one.
        vi. Otherwise, assume highest dose Simvastatin.
    b. For ***Ezetimibe,*** the combination drug represents a higher dose of treatment than Ezetimibe on its own so the combination drug should be retained, and others discarded.
    c. For ***fibrates,*** no other duplicates remained so a hierarchy was not needed.
12. Combine all entries to be remerged and check for duplicates. If duplicates remain, repeat steps 6 – 11. If none, all duplicate dispensing dates have been removed and previous method can be resumed.

Appendix A3 Statin Combinations: Method for adding adjusted dates back into Statin Datasets

Once all dates have been adjusted to prevent any overlap within the drug groups, the following method should be used to ensure that the dates in the statin dataset are corrected (as the Ezetimibe and Fibrate groups were corrected as part of the sequence).

1. Extract the dates and the prescriptions for Simvastatin & Ezetimibe (from the ezetimibe datasets) and Fenofibrate/Simvastatin (from the fibrate datasets), into a dates and prescriptions datasets.
2. Remove the dates and the prescriptions from the statin datasets for the combination drugs.
3. Combine the combination dates with the statin dates, and the prescriptions similarly.
4. Order the statin dates dataset by SafeHavenID, start date, and end date.
5. Order the statin prescriptions dataset by SafeHavenID, dispensed date, and prescribed date.

6. Correct the course numbering system in the dates dataset and merge the numbering with the prescriptions dataset. Correct the course ID with the format: SafeHavenID – CourseNumber.

Appendix B Calculating Medication Possession Ratio (MPR)

1. Identify the dates for the first non-fatal MIs of the individual within the time window of interest (1st January 2009 to 31st July 2014, inclusive), from the SMR01 file.
2. Set the first start date to be equal to this date in a new variable. The end date is then this date plus 364 days. Run an adjustment of +1 if this time window includes 29th February 2012 or 2016 as these are the two leap years that fall within the follow up period.
3. The next year starts the day after the previous end date. Repeat this until there are nine start and end dates for each individual as this is the maximum duration of follow up (up to 31st July 2017, inclusive). At each stage of creating an end date, adjust for the leap year as in step 2.
4. Transpose the data set so start and end dates are now in long form, with each window labelled. Merge with the individuals date of death.
5. Remove time windows which start after, and including, 1st August 2017, or which start on or after date of death.
6. Create two indicator variables (1,0). One should be used to indicate if the individual dies during the time interval, and the other should be used to indicate if the end of follow up (1st August 2017) occurs during the time interval.
7. If a death occurs in the time window, set the end date as equal to the date of death. After this, for all time windows, if the end date is equal to the 1st August 2017 or after, set the end date to the 31st July 2017.
8. In a separate dataset, merge the time windows with the dates of MIs occurring in follow up. Keep time windows where a follow up MI occurs during the time window and create an indicator variable (by setting it equal to one). Keep a list of SafeHavenID, time window number and the indicator variable, and then find unique records (as individuals may have more than one MI in the time window).
9. Merge this indicator variable into the time window data. Set the indicator variable equal to 0 for records which did not exist in the list devised in step 8.
10. Combine this time window dataset with the file of individual prescriptions. Order by SafeHavenID, time window number, and dispensed date.
11. Restrict to the prescriptions where the calculated end date is on or after the time window start date, and the dispensed date is prior to the time window end date.
12. Create a time window ID consisting of the SafeHavenID – time window.
13. Identify prescriptions which had started prior to the time window and those which would carry on after the time window with indicator (1,0) variables.
14. Format the quantity dispensed variable to be numeric, and adjust the quantity dispensed according to the number that should have already been administered prior to the time window start date, and the number that should be remaining after the time window end date (for each prescription). Remove prescriptions where the total is less than zero as a result of this calculation.
15. Calculate the number of prescriptions dispensed for each unique time window in a separate dataset.
16. Sum the total quantity of doses dispensed for each unique time window in a separate dataset.
17. Calculate the length of each time window (end date – start date +1). The +1 is added as it is assumed that a dose is taken on both the start and the end date.
18. Subset the prescription data to contain only the ID variables, the start and end dates, and the length of the time window. Remove complete duplicates.

19. Combine this subsetted data with the datasets created in 15 and 16, merging by time window ID. Then order the data by SafeHavenID and time window ID.
20. Calculate the MPR as the (total quantity/length of time window) *100.
21. Merge with the original time window list, as this dataset would only contain those where at least one prescription is dispensed.
22. Re-establish the time window ID and calculate the duration as in step 17. For those with missing numbers of prescriptions, and days of doses dispensed, set these equal to zero. (These would have been the time windows that were dropped during the merge with prescription data. They should be included in the analysis though, as post MI, all individuals should be prescribed a statin, and therefore their adherence should be 0.)
23. Calculate MPR as zero for the time windows where no doses were dispensed.
24. Finally, set the MPR to be missing for those time windows where a death or the end of follow up occurs (i.e. a full year of data is not available).

Appendix C Handling Duplicate Lipid Profile Results

The multiple components of the lipid profiles represented the same observation, so were transposed from the long form (where the observations collected for each of the components are listed one below the other) to the wide form (where the observations are combined onto the same entry in the dataset). If duplicates are present for any of the values, this can result in exacerbated duplication following transposition. A systematic method was employed to remove both the exacerbated duplication and handle instances of genuine duplication.

The components of the lipid profile test used are:

- Total Cholesterol
- Total:HDL Cholesterol
- HDL Cholesterol
- Triglycerides
- LDL Cholesterol

Whilst transposition of the dataset could be performed using a transpose command, due to the size of the dataset, for this analysis this was performed using the following method instead. If the dataset has already been transposed, skip to step 5.

1. Ensure the long dataset contains all the components of the lipid profiles listed above. In the SafeHaven, these have several names and checks for variations were conducted throughout the extraction of relevant data. For this analysis, the names of tests included:

Total Cholesterol	**Total:HDL Cholesterol**	**HDL Cholesterol**
Cholesterol Cholestrol Cholesterol UC	Chol/HDL Ratio Chol/HDL Ratio (for risk assess) CHOL/HDL Ratio Chol/HDL Ratio	HDL Chol HDL Cholesterol HDL-Cholesterol HDL-Cholesterol UC HDL CHOL HDL Cholesterol UC HDL Chol (mmol/l)

Triglycerides	**LDL Cholesterol**
Triglyceride Triglycerides Triglycerides UC	LDL (calculated) LDL Chol LDL Chol (calc) LDL-Chol calculated LDL-Cholesterol UC LDL Cholesterol LDL Cholesterol * LDL-Cholest (calc'd)

2. Check all results are reported in the same unit. If this is not the case, either convert all to the same unit, or ensure that the units are also transposed (and that they are consistently reported, i.e. no different levels of spaces or cases).
3. Create a dataset for each of the components of the lipid profile and rename the result to indicate the component (e.g. "LDL Result") and drop any remaining

unnecessary variables. Remove any complete duplicates to limit subsequent exacerbation

4. Combine each of the five components by SafeHavenID and the datetime to create a wide version of the dataset. Create a sample ID formatted as: SafeHavenID – SampleDateTime.
5. Remove any inconsistent naming before removing any complete record duplicates.
6. Identify those that are duplicates and create two separate datasets; one for all samples with duplicates, and one for samples with no duplicates.
7. Within the duplicates data set, calculate the LDL and Total:HDL Cholesterol ratio for each of the combinations. If both match for a record, extract these entries and discard all remaining duplicates for these samples.
8. If, in the remaining duplicated samples, there is a complete zero entry for a sample, extract all records for these samples into a separate dataset. The most complete row for each sample should then be retained, with all other records removed. All records from these samples should then be removed from the remaining duplicates.
9. Divide the remaining samples into three datasets: (a) the records with triglycerides too high to calculate LDL, and with LDL equal to zero or NA; (b) the records with the same sample IDs as the first group, but where the criteria isn't met; and (c) those samples where this is not the case.
10. In dataset (a), if the Total:HDL matches, keep the record that matches and discard the remaining records for that sample in (a). Due to rounding and the calculation method in some programs, the two numbers may not match due to a floating point error. Therefore, if subsequently, there is a match within 0.05 for the Total:HDL, these records should also be retained, with others with the same sample ID removed (only in (a)). All remaining records in (a) should also be kept.
11. In dataset (b), remove all records with triglycerides too high to calculate LDL as these will have been captured in (a). Records with triglycerides or LDL equal to zero should also be removed. Finally, records with matching Total:HDL and LDL matching after the 0.05 adjustment for floating point error should also be kept, with all other records removed.
12. Identify those in dataset (c) to keep where Total:HDL matches, and remove all records in (c) for the same samples.
13. Identify those where the Total:HDL matches after 0.05 adjustment for floating point error, and discard all remaining records for these samples from the dataset. Within those that do match, identify those that match LDL and discard all remaining records for these samples. Within those remaining, keep those with matching LDL after floating point error adjustment of 0.05 (discarding those from the same sample), and subsequently remaining records where LDL is not equal to NA.
14. Within those where Total:HDL did not match after adjustment of 0.05 for floating point error, keep those that matched LDL after floating point error adjustment, and those that did not but were not from the same samples where a match was achieved.
15. Combine all observations to be kept. Identify duplicates in this dataset again. Many of these duplicates are likely to have been present in the original long form, with some unable to be distinguished by the previous steps.
16. Repeat step 6.
17. Within the duplicates set, extract the min and maximum value for each of the components.

18. Calculate mean and mean excluding zeros of the minimum and maximum value (removing NAs where applicable). If the mean value is not equal to zero, keep the one excluding zeros, otherwise keep the one equal to zero.
19. Keep these means and rename them as the original variables, removing duplicates if created through the creation process, before combining with the non-duplicate set.

Appendix D Calculating Baseline Change in Non-HDL Cholesterol

1. Extract the lipid profile results for those within the post MI population, and merge with the date of the first MI of interest. Order the results by SafeHaven ID and sample date.
2. Restrict further to tests where the sample date is before the date of admission with the MI.
3. Extract the statin prescriptions for those within the post MI population and merge with the date of the first MI of interest. Order the results by SafeHaven ID and dispensed date. Restrict further to prescriptions where the dispensed date is before the date of admission with the MI.
4. Merge the lipid profile and statin prescription datasets by SafeHaven ID. This will result in the creation of many duplicates.
5. Extract lipid profile tests that were collected after the prescriptions dispensed date but before or on the prescription end date. Remove all records with the sample IDs in this list from the combined dataset.
6. Repeat step 5 but for all lipid profile tests that were collected within six months of the prescription end date.
7. Restrict the lipid profile results file to those with sample IDs which were remaining in the combined dataset or for individuals who had had no statin prescriptions prior to their MI
8. Order by SafeHaven ID and sample date once more. Keep only the results of the last sample per individual.
9. Calculate the non HDL by subtracting the HDL from the total cholesterol value.
10. Relabel all the components of the lipid profile to indicate that they are the pre-MI results.
11. For the lipids post MI (lipid results collected in the population on or after the admission date), calculate the non HDL for the samples.
12. Merge the baseline samples with the post MI samples.
13. Calculate the percentage change for each of the components including the non HDL component as: ((sample – baseline)/baseline)*100.
14. If the baseline component is equal to zero, this may autogenerate the percentage change as infinity or NaN (Not A Number). Set these cases equal to missing for each of the six components.

Appendix E Extracting Type and Date of Diagnosis of Diabetes

The date of diagnosis is in the same observation line as the type of diabetes meaning that observations relating to date and type did not have to be matched. The date of diagnosis is also the date of entry in subsequent patient observations, so the first date in the dataset for each patient is taken as the date of diagnosis. Due to the numerous variations of the types of diabetes given within the data, the definitions are collapsed to Type 1, Type 2, Not Diabetes and Not Known. The original descriptions and their mappings are shown below. Those considered to not have diabetes were subsequently removed from the diabetes dataset. Once this was complete, the method below could be implemented.

Type 1	**Type 2**
Latent autoimmune diabetes of adulthood	Diabetes in Remission
Maternally inherited diabetes and deafness	Diabetes Resolved
Neonatal Diabetes	Maturity Onset Diabetes of youth
T1DM	Other Type of Diabetes
	Secondary – Disease
	Secondary – Drug Induced
	Secondary – Pancreatic Pathology

Not Diabetes	**Not Known**
Current Gestational Diabetes	Not Known
History of Gestational Diabetes	
Impaired Fasting Glucose	
Impaired Glucose Tolerance	
Not Diabetic	
Stress Induced Hyperglycaemia (transitory)	

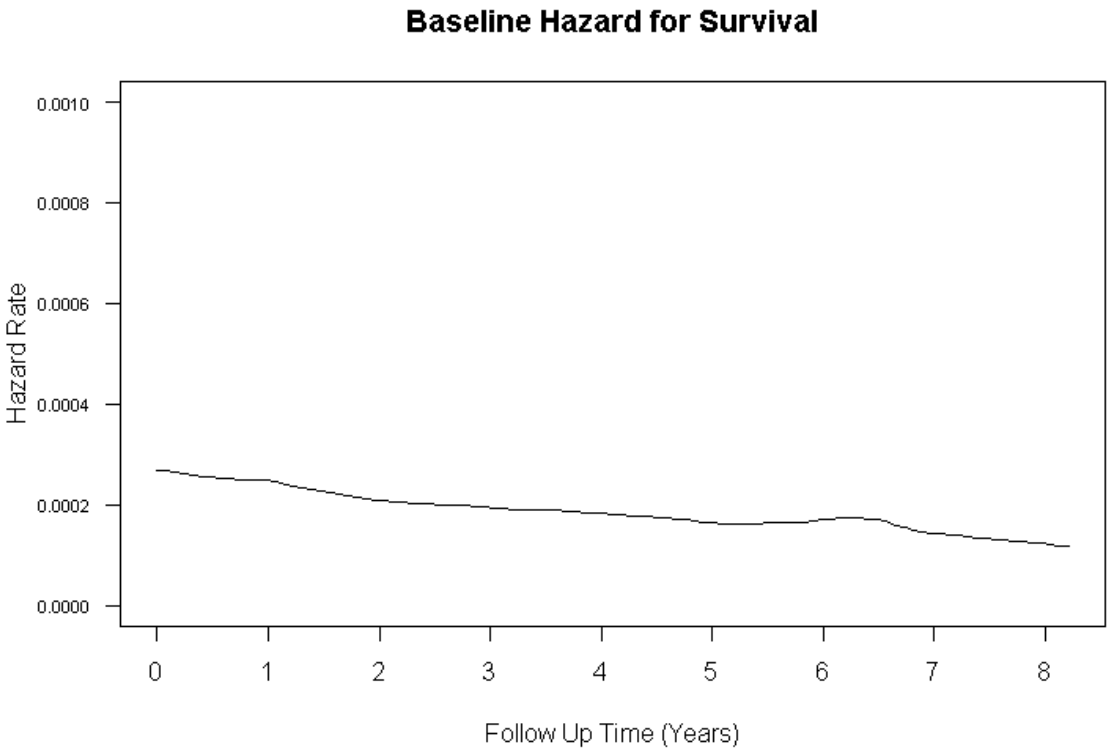
1. Extract valid dates of entry of the information:
    a. Calculate age at date of entry of the type of diabetes (date of birth available in the demography file).
    b. Identify entries occurring prior to date of birth.
    c. Calculate the number of entries prior to date of birth per patient.
    d. Calculate the total number of type of diabetes entries per patient.
    e. Combine these values with the SafeHaven ID into one dataset and remove duplicates.
    f. Merge this dataset with the original type of diabetes dataset.
    g. Keep only the observations with date of entry after birth, or where the total number of entries is equal to the number collected prior to date of birth.
    h. For entries remaining with age less than zero, set the age to NA.
2. Determine type of diabetes for the majority of the individuals:
    a. Using the categories in the previous table, for each individual, calculate the number of times they have been diagnosed with Type 1, Type 2, or Not Known, and combine the totals in a separate dataset.
    b. For most individuals, diabetes type can then be determined using the following rules:
        i. If all entries are Not Known, then Type = 2.
        ii. If only entries for Type 1 (or Not Known), Type = 1.
        If only entries for Type 2 (or Not Known), Type = 2.
        iii. If no Not Known entries and more for Type 1 than Type 2, Type = 1.

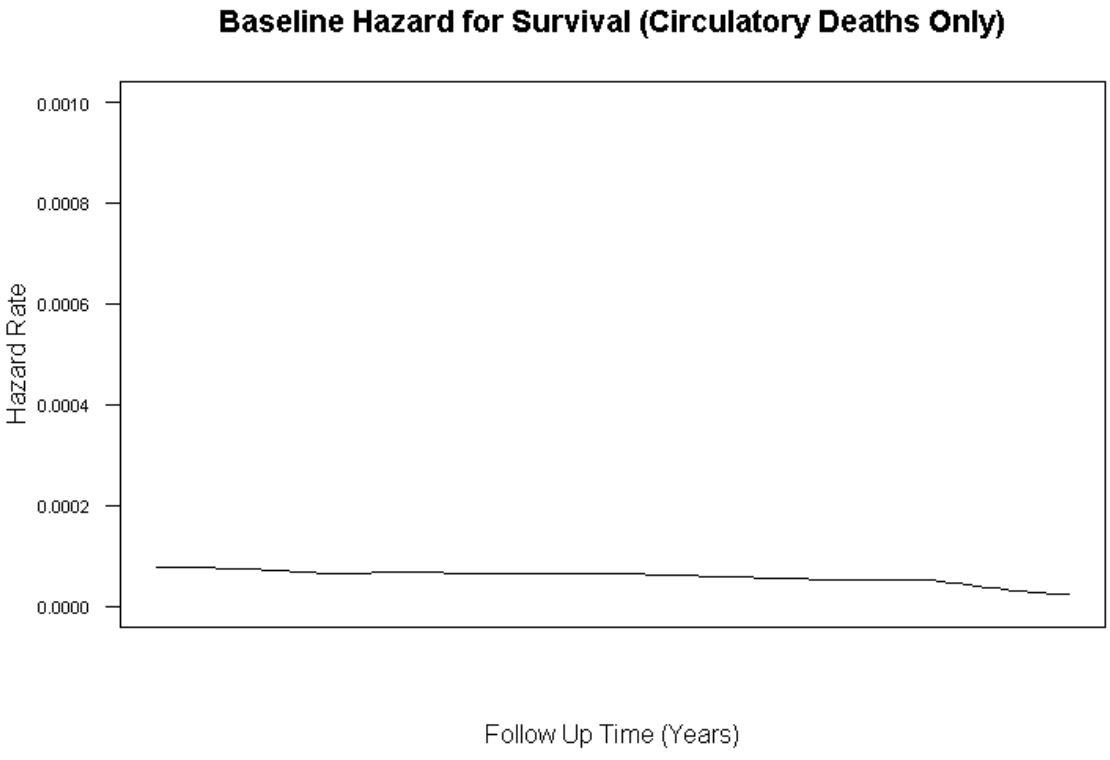
If no Not Known entries and more for Type 2 than Type 1, Type = 2.

iv. If more Type 1 entries than Type 2 entries and Not Known entries combined, Type =1.
If more Type 2 entries than Type 1 entries and Not Known entries combined, Type = 2.

3. Extract the first date of diagnosis:
    a. Find the first observation for each individual in list of valid dates and combine with type of diagnosis determined in last step.
4. Determine the remainder of the missing information for an individual's types of diabetes:
    a. Use the individual's age at first diagnosis to implement the following rules for those with missing type of diabetes:
        i. If age less than 20 years, Type = 1.
        ii. If age greater than or equal to 20 years, Type = 2.
5. Adjust those with Type 2 diabetes and less than 20 years at first diagnosis to be Type 1.

Appendix F Baseline Hazards from Kaplan-Meiers for Overall Mortality (All and Circulatory Deaths)





Appendix G Logistic Regression Models for Further Hospitalised MIs by Lipid Target Achievement

Appendix G1 Model for Further Hospitalised MIs by Non-HDL Target

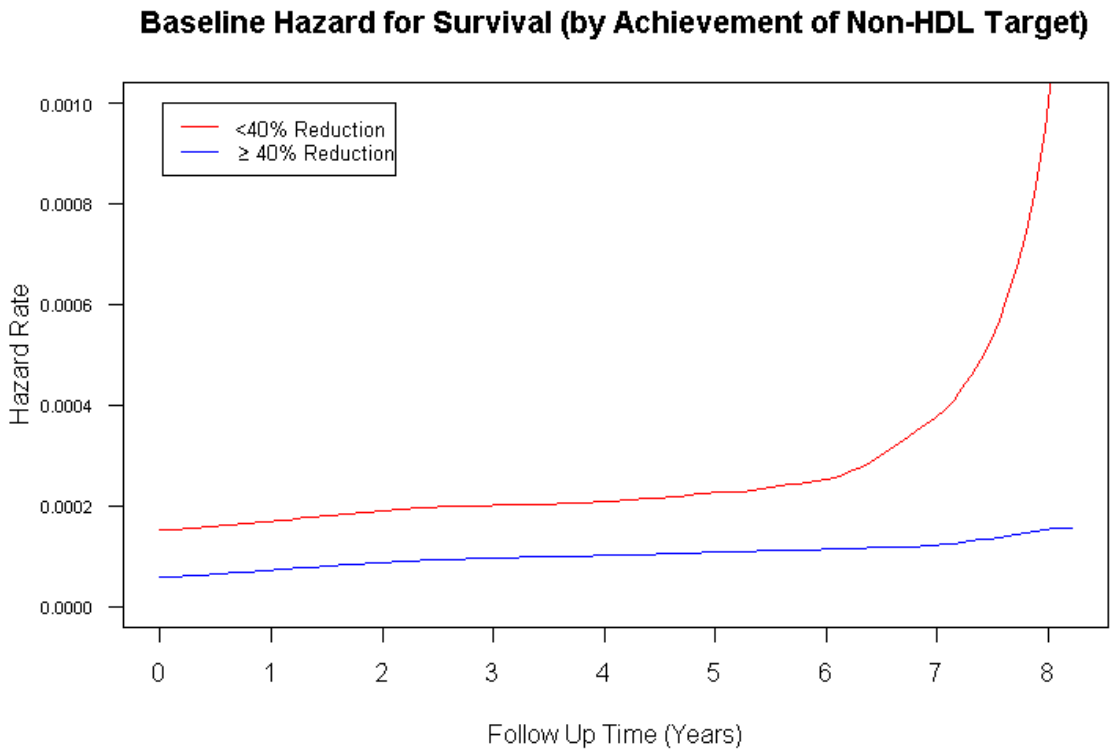
Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Non-HDL Target		
Achieved (ref.)	1.00	0.0721
Not Achieved	1.15 (0.99, 1.33)	
Adjusted Model		
Non-HDL Target		
Achieved (ref.)	1.00	0.0561
Not Achieved	1.16 (1.00, 1.35)	
Sex		
Female (ref.)	1.00	0.9147
Male	1.01 (0.86, 1.18)	
Age at MI (years)	1.00 (0.99, 1.01)	0.8183
SIMD (NHS GGC)		
1 [most]	0.86 (0.67, 1.10)	0.0288
2	1.01 (0.79, 1.28)	
3	0.90 (0.70, 1.15)	
4	0.69 (0.53, 0.91)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	0.1650
2010	0.98 (0.71, 1.36)	
2011	0.96 (0.70, 1.32)	
2012	1.22 (0.90, 1.66)	
2013	1.23 (0.91, 1.67)	
2014	1.11 (0.79, 1.55)	

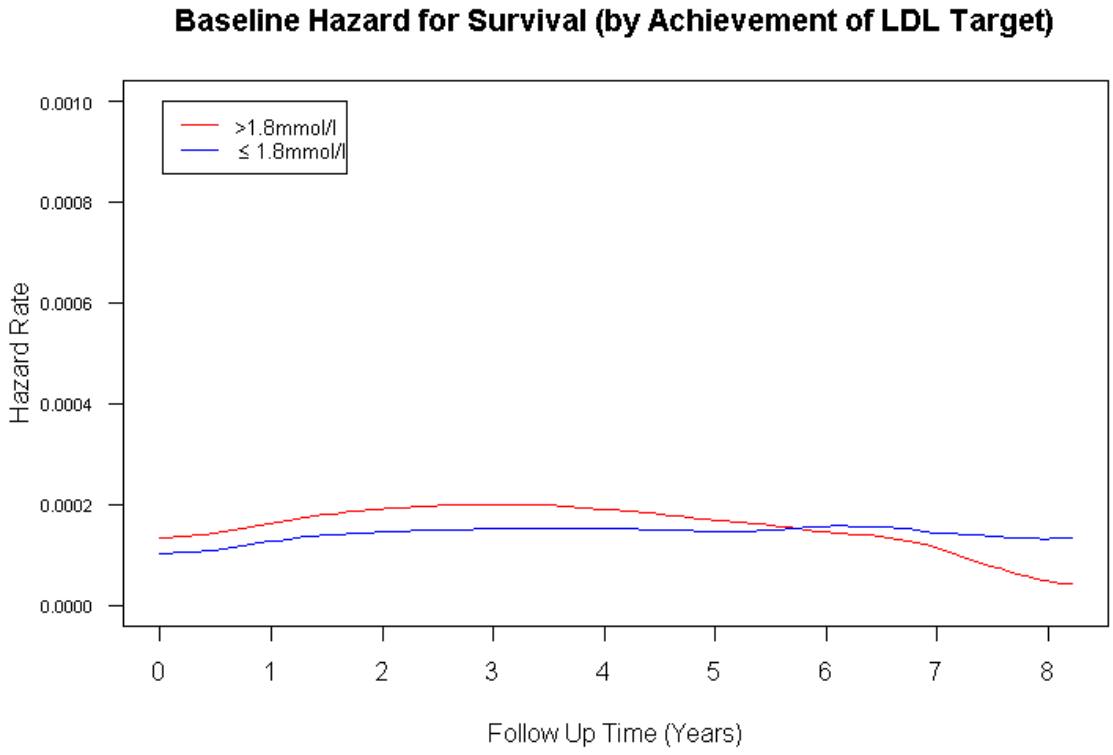
Appendix G2 Model for Further Hospitalised MIs by LDL Target

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
LDL Target		
Achieved (ref.)	1.00	0.4302
Not Achieved	0.96 (0.86, 1.06)	
Adjusted Model		
LDL Target		
Achieved (ref.)	1.00	0.3336
Not Achieved	0.95 (0.85, 1.06)	
Sex		
Female (ref.)	1.00	0.6709
Male	0.98 (0.89, 1.08)	
Age at MI (years)	1.00 (1.00, 1.01)	0.3928
SIMD (NHS GGC)		
1 [most]	1.17 (1.00, 1.35)	0.0611
2	1.05 (0.91, 1.22)	
3	1.09 (0.94, 1.27)	
4	0.95 (0.81, 1.12)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	1.00 (0.85, 1.17)	
2011	1.07 (0.91, 1.25)	
2012	1.38 (1.18, 1.61)	
2013	1.32 (1.13, 1.54)	
2014	1.23 (1.02, 1.48)	

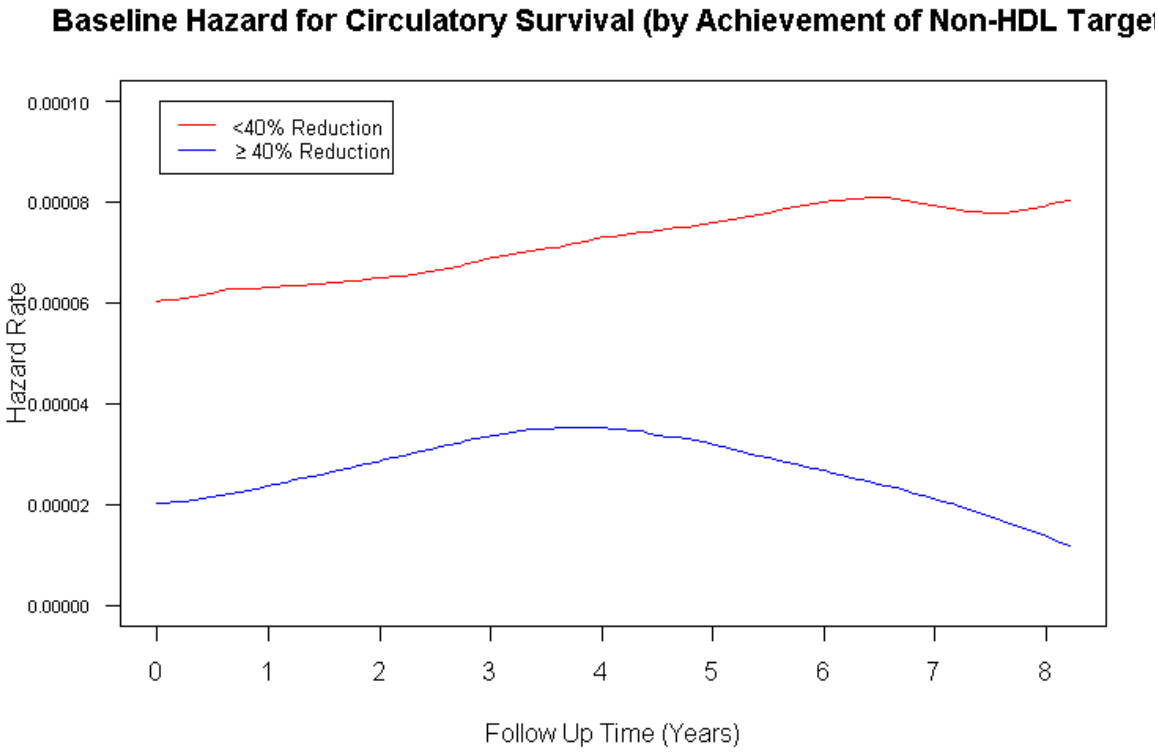
Appendix H Baseline Hazards from Kaplan-Meiers for Mortality by Lipid Target Achievement

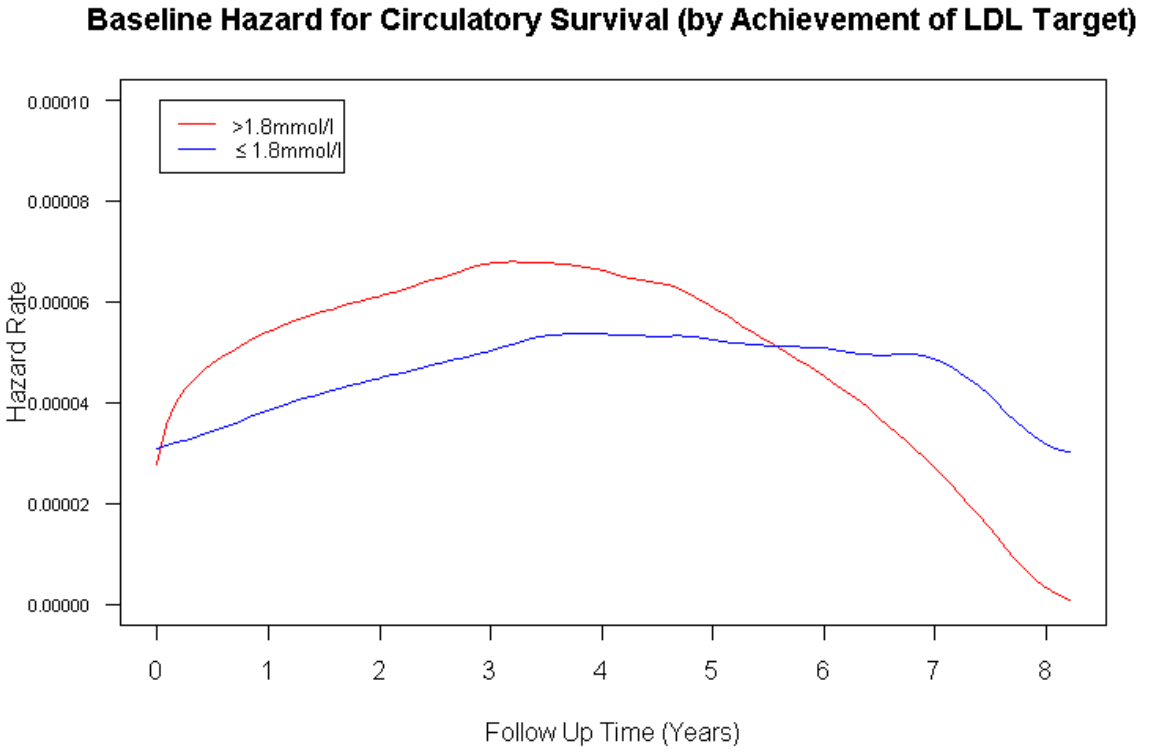
Appendix H1 Baseline Hazards for Overall Mortality





Appendix H2 Baseline Hazards for Circulatory Mortality





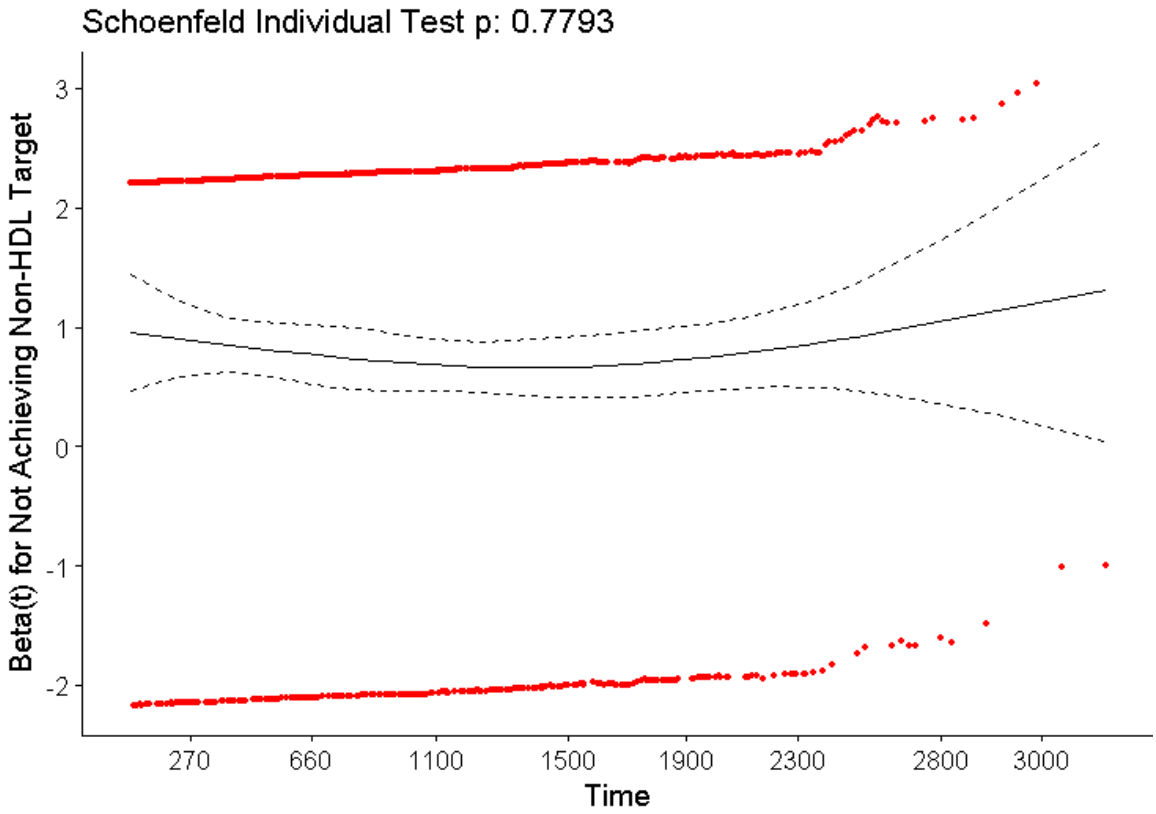
Appendix I Cox Regression Models for Mortality by Lipid Target Achievement

Appendix I1 Model for Overall Mortality by Non-HDL Target

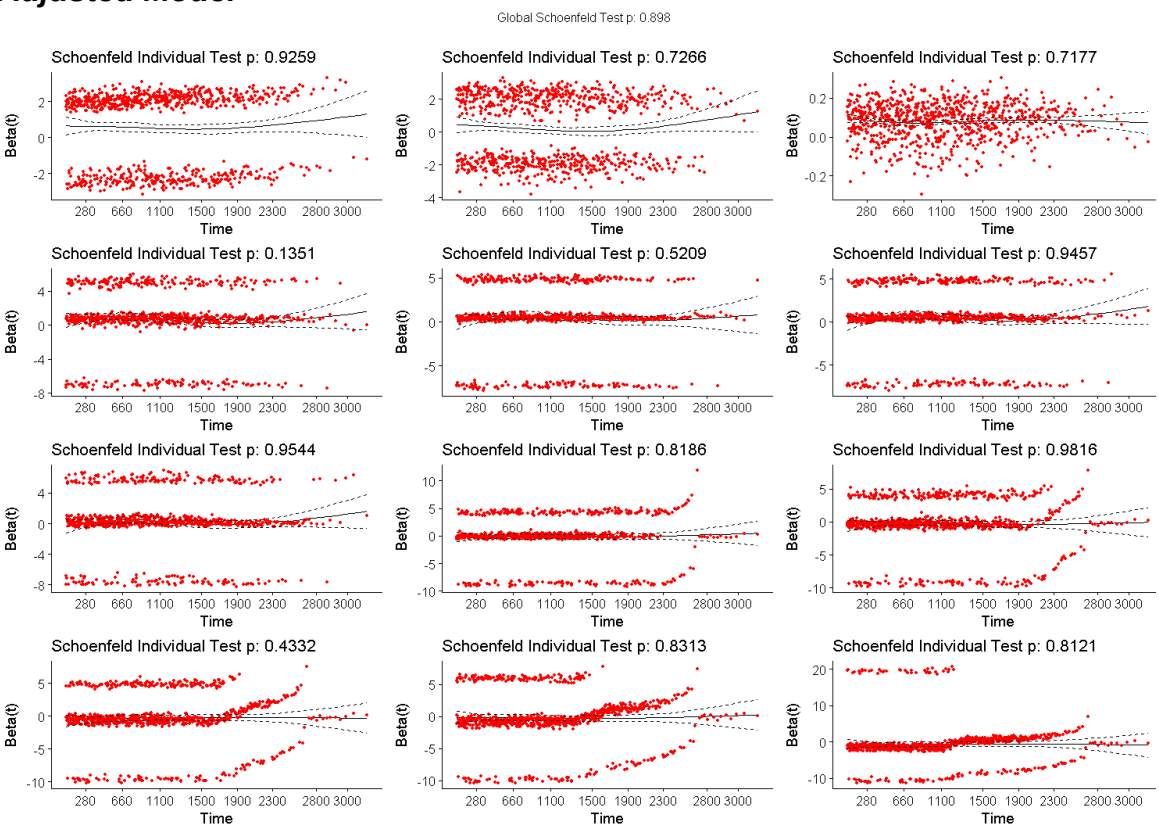
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Non-HDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	2.16 (1.87, 2.49)		0.7793
Adjusted Model			
Non-HDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.76 (1.52, 2.03)		0.9259
Sex			
Female (ref.)	1.00	0.0014	
Male	1.26 (1.09, 1.46)		0.7266
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.7177
SIMD (NHS GGC)			
1 [most]	1.95 (1.53, 2.49)		0.1351
2	1.56 (1.23, 1.99)		0.5209
3	1.60 (1.26, 2.03)	<0.0001	0.9457
4	1.22 (0.95, 1.57)		0.9544
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	0.90 (0.70, 1.16)		0.8186
2011	0.79 (0.61, 1.02)	0.0441	0.9816
2012	0.76 (0.58, 0.99)		0.4332
2013	0.85 (0.64, 1.11)		0.8313
2014	0.57 (0.39, 0.83)		0.8121
Global PH Test			0.8980

Appendix I2 Schoenfeld Residuals for Model for Overall Mortality by Non-HDL Target

Unadjusted Model



Adjusted Model

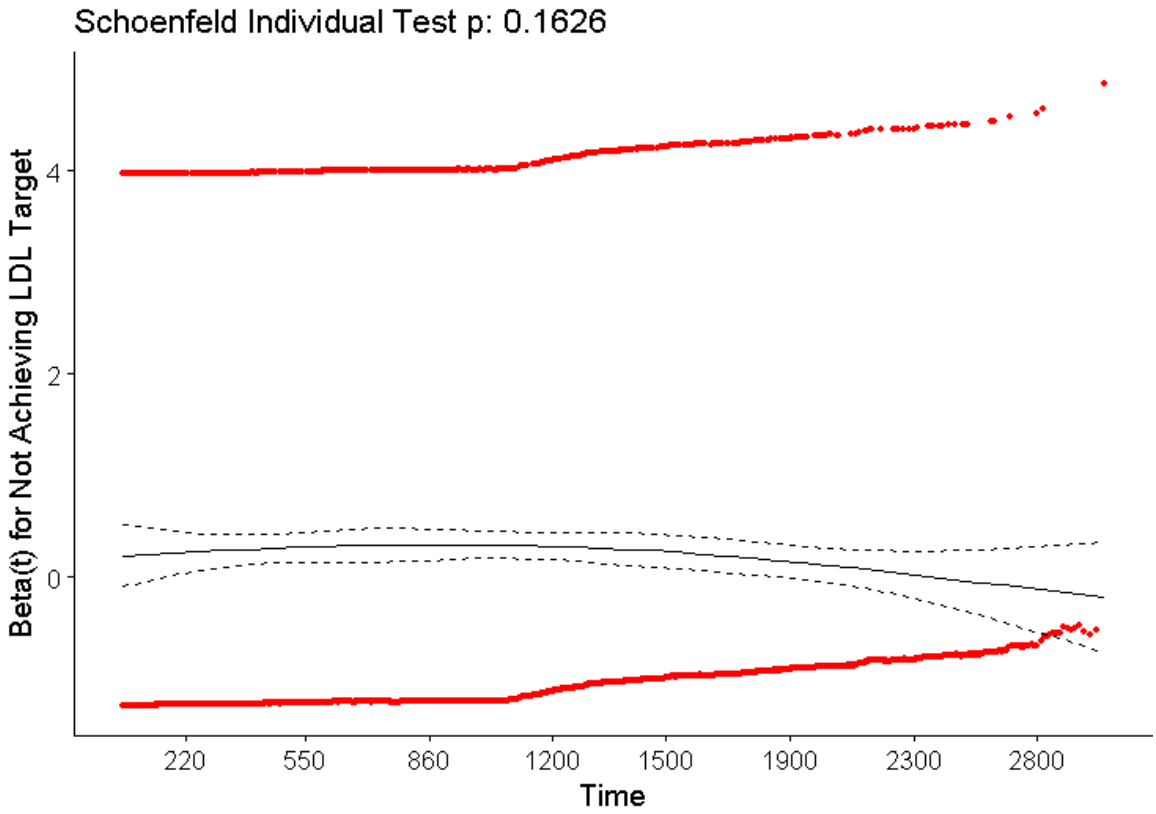


Appendix I3 Model for Overall Mortality by LDL Target

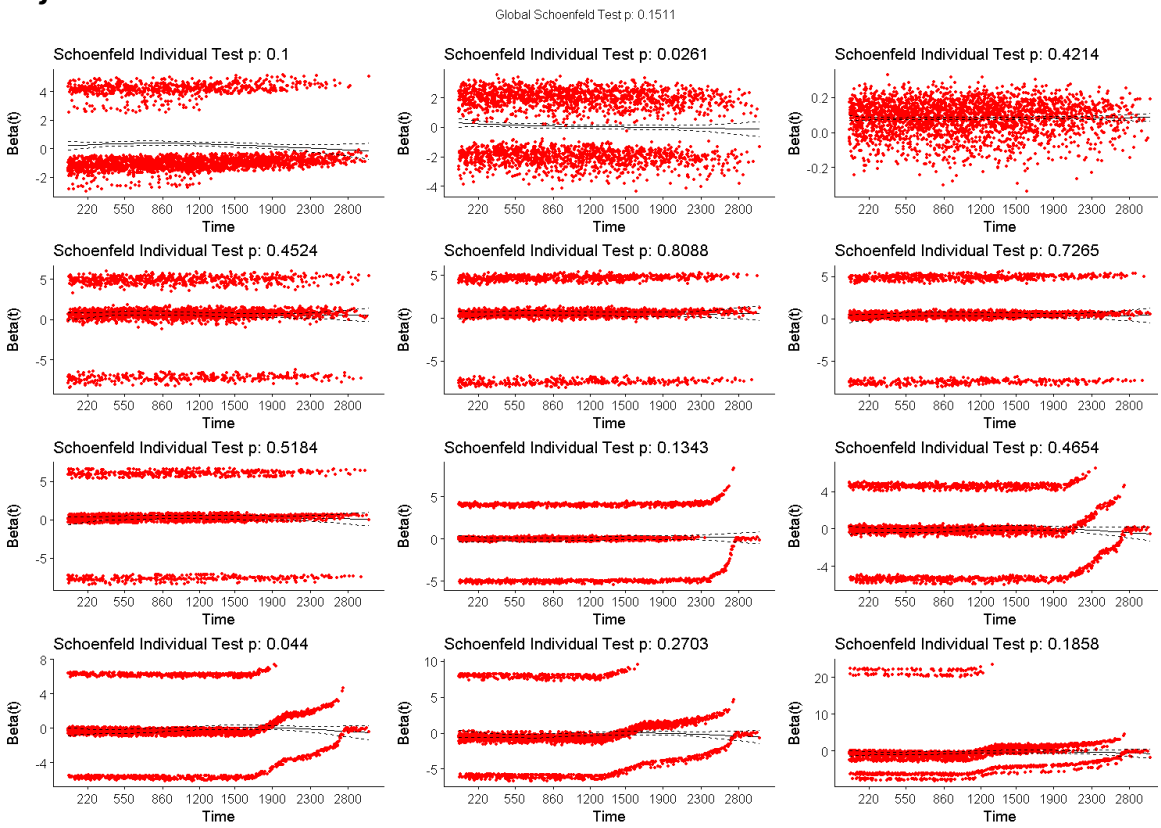
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
LDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.27 (1.16, 1.39)		0.1626
Adjusted Model			
LDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.32 (1.20, 1.44)		0.1000
Sex			
Female (ref.)	1.00	0.0361	
Male	1.09 (1.01, 1.18)		0.0261
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.4214
SIMD (NHS GGC)			
1 [most]	1.92 (1.68, 2.20)		0.4524
2	1.66 (1.45, 1.90)		0.8088
3	1.57 (1.37, 1.80)	<0.0001	0.7265
4	1.32 (1.14, 1.52)		0.5184
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	1.01 (0.90, 1.14)		0.1343
2011	0.94 (0.83, 1.06)	<0.0001	0.4654
2012	0.87 (0.76, 0.99)		0.0440
2013	0.83 (0.72, 0.96)		0.2703
2014	0.64 (0.52, 0.79)		0.1858
Global PH Test			0.1511

Appendix I4 Schoenfeld Residuals for Model for Overall Mortality by LDL Target

Unadjusted Model



Adjusted Model

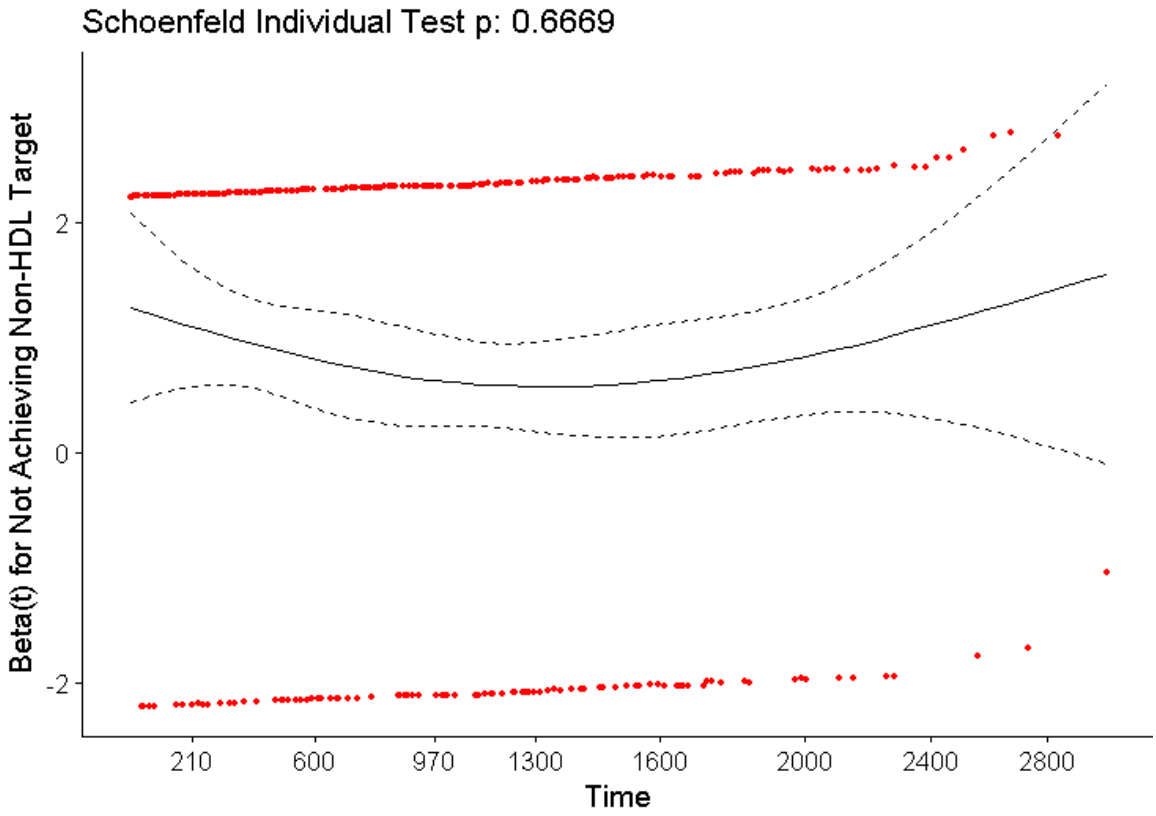


Appendix I5 Model for Circulatory Mortality by Non-HDL Target

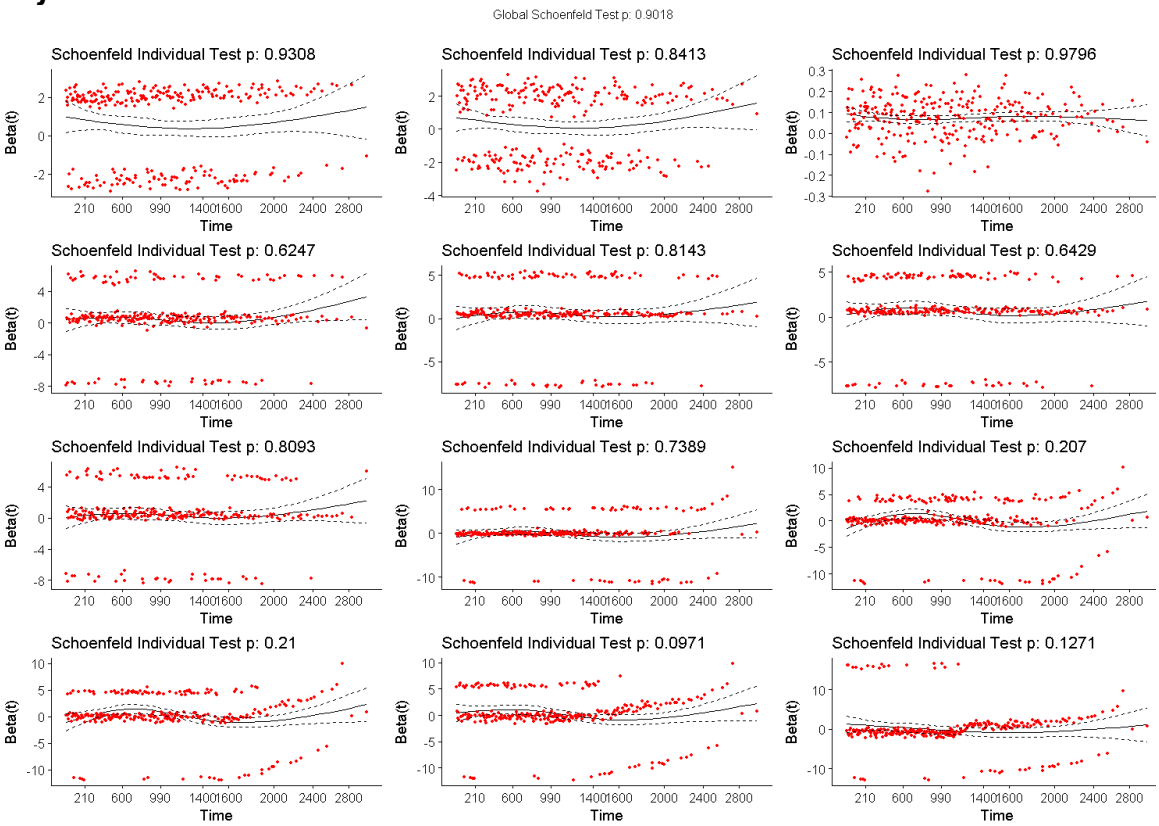
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Non-HDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	2.24 (1.74, 2.87)		0.6669
Adjusted Model			
Non-HDL Target			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.82 (1.41, 2.34)		0.9308
Sex			
Female (ref.)	1.00	0.0101	
Male	1.38 (1.08, 1.78)		0.8413
Age at MI (years)	1.08 (1.07, 1.09)	<0.0001	0.9796
SIMD (NHS GGC)			
1 [most]	1.66 (1.07, 2.58)		0.6247
2	1.65 (1.08, 2.52)		0.8143
3	1.91 (1.26, 2.90)	0.0286	0.6429
4	1.46 (0.95, 2.26)		0.8093
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	0.90 (0.55, 1.46)		0.7389
2011	1.14 (0.71, 1.83)	0.4666	0.2070
2012	1.22 (0.76, 1.98)		0.2100
2013	1.35 (0.82, 2.23)		0.0971
2014	1.00 (0.53, 1.90)		0.1271
Global PH Test			0.9018

Appendix I6 Schoenfeld Residuals for Model for Circulatory Mortality by Non-HDL Target

Unadjusted Model



Adjusted Model

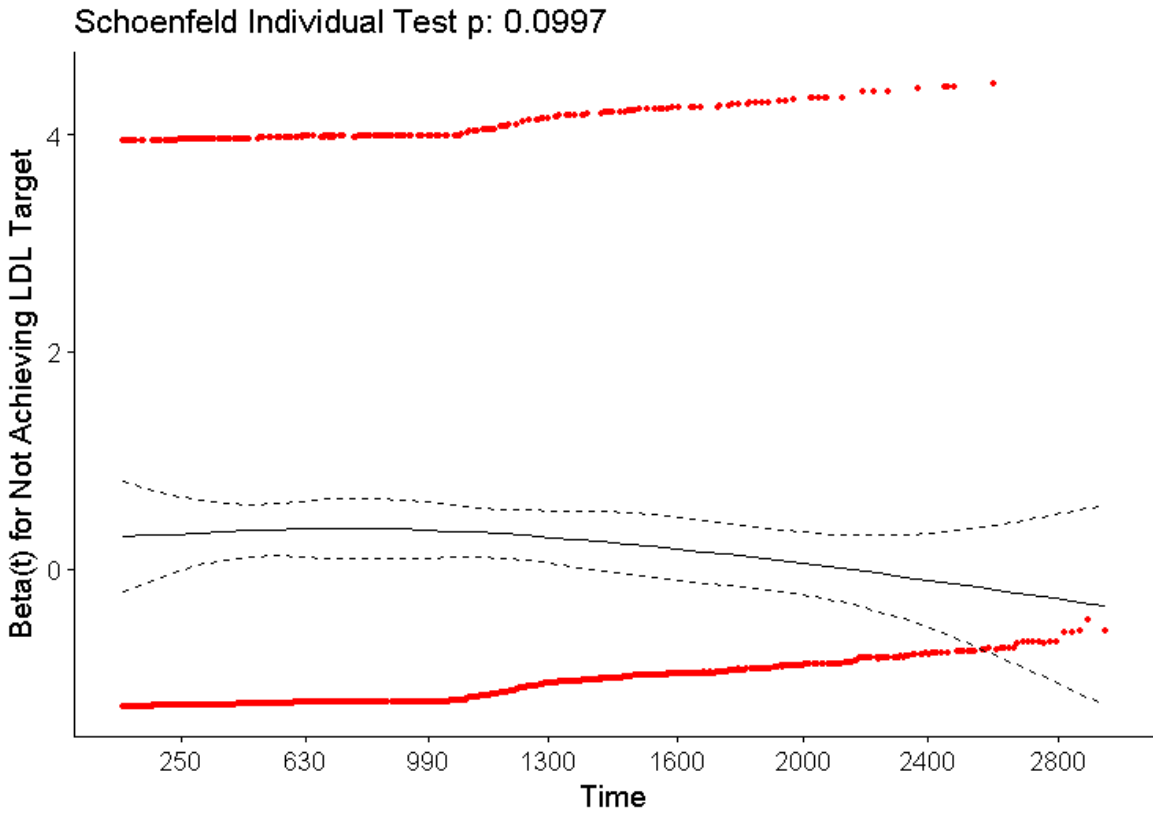


Appendix I7 Model for Circulatory Mortality by LDL Target

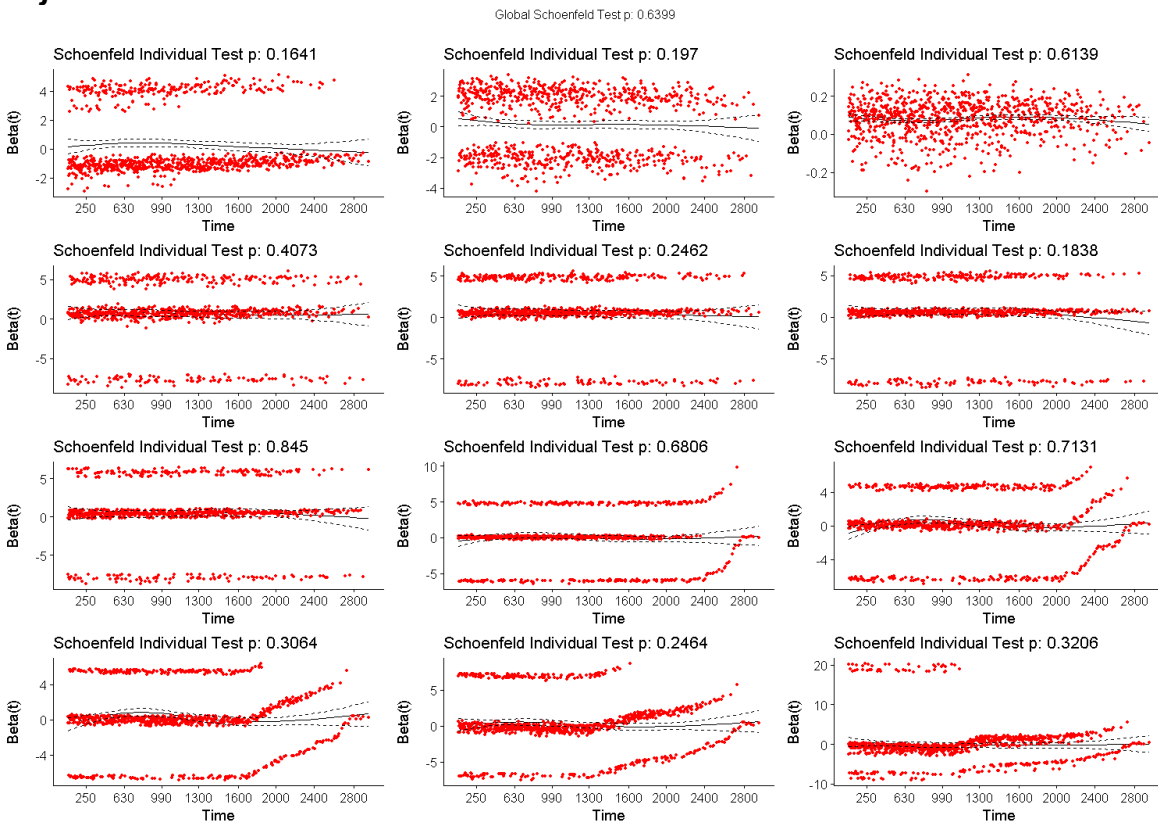
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
LDL Target			
Achieved (ref.)	1.00	0.0011	
Not Achieved	1.30 (1.11, 1.51)		0.0997
Adjusted Model			
LDL Target			
Achieved (ref.)	1.00	0.0012	
Not Achieved	1.30 (1.11, 1.52)		0.1641
Sex			
Female (ref.)	1.00	0.0062	
Male	1.22 (1.06, 1.40)		0.1970
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.6139
SIMD (NHS GGC)			
1 [most]	2.01 (1.58, 2.56)		0.4073
2	1.71 (1.34, 2.18)		0.2462
3	1.75 (1.38, 2.23)	<0.0001	0.1838
4	1.53 (1.19, 1.97)		0.8450
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	1.04 (0.83, 1.30)		0.6806
2011	1.20 (0.96, 1.50)	0.0570	0.7131
2012	1.33 (1.06, 1.68)		0.3064
2013	1.31 (1.02, 1.68)		0.2464
2014	0.98 (0.69, 1.39)		0.3206
Global PH Test			0.6399

Appendix I8 Schoenfeld Residuals for Model for Circulatory Mortality by LDL Target

Unadjusted Model



Adjusted Model



Appendix J Logistic Regression Models for Further Hospitalised MIs by Average Statin Adherence

Appendix J1 Model for Further Hospitalised MIs by Continuous Adherence

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Continuous Adherence		
10% Decrease	1.00 (0.98, 1.01)	0.5179
Adjusted Model		
Continuous Adherence		
10% Decrease	1.00 (0.98, 1.01)	0.5775
Sex		
Female (ref.)	1.00	0.9147
Male	0.99 (0.91, 1.09)	
Age at MI (years)	1.00 (1.00, 1.00)	0.5852
SIMD (NHS GGC)		
1 [most]	1.16 (1.00, 1.34)	
2	1.06 (0.92, 1.23)	
3	1.08 (0.93, 1.26)	0.0610
4	0.95 (0.81, 1.11)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	
2010	0.98 (0.83, 1.14)	
2011	1.03 (0.88, 1.20)	<0.0001
2012	1.33 (1.14, 1.54)	
2013	1.26 (1.08, 1.47)	
2014	1.12 (0.94, 1.34)	

Appendix J2 Model for Further Hospitalised MIs by Average Adherence (50%)

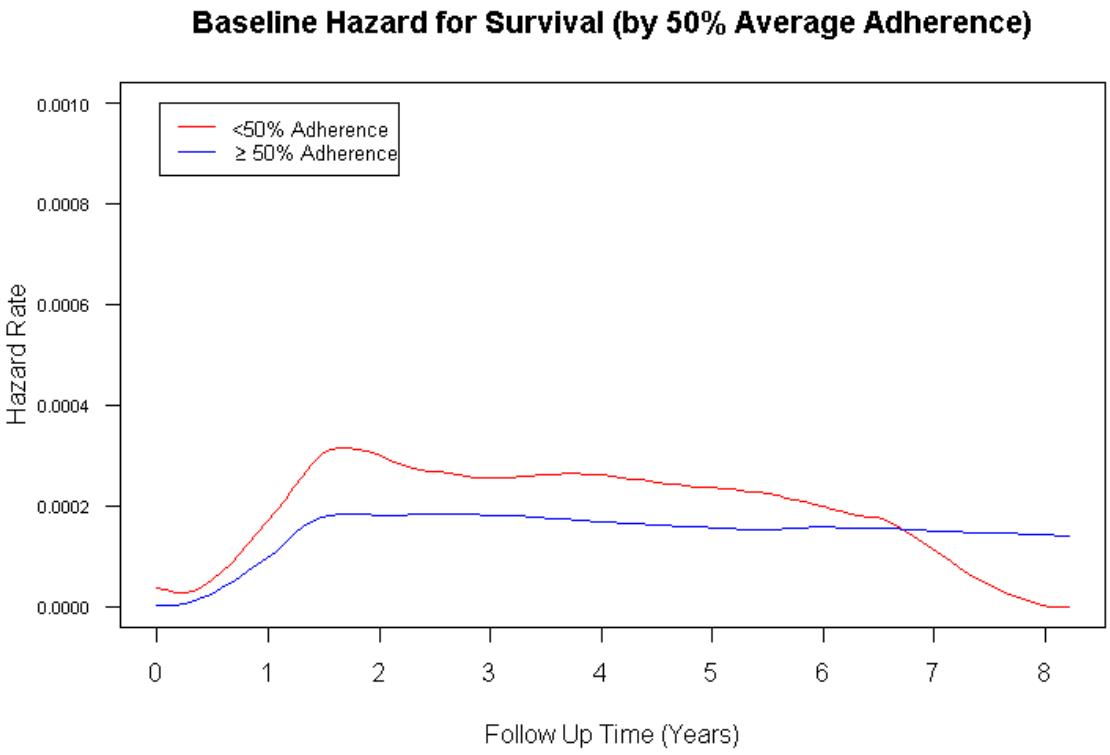
Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	0.0734
Not Achieved	0.89 (0.79, 1.01)	
Adjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	0.0748
Not Achieved	0.89 (0.79, 1.01)	
Sex		
Female (ref.)	1.00	0.8644
Male	0.99 (0.90, 1.09)	
Age at MI (years)	1.00 (1.00, 1.00)	0.6342
SIMD (NHS GGC)		
1 [most]	1.16 (1.00, 1.34)	0.0653
2	1.06 (0.91, 1.23)	
3	1.08 (0.93, 1.26)	
4	0.95 (0.81, 1.11)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	0.98 (0.83, 1.14)	
2011	1.03 (0.88, 1.20)	
2012	1.33 (1.15, 1.55)	
2013	1.26 (1.08, 1.47)	
2014	1.12 (0.94, 1.34)	

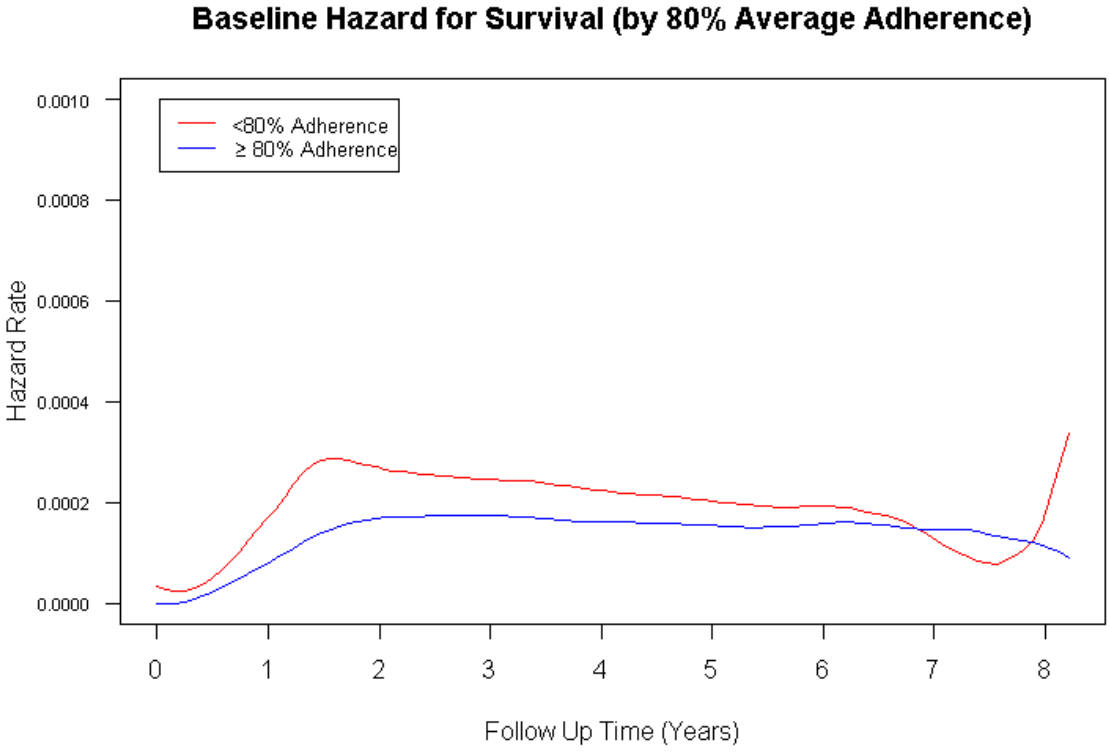
Appendix J3 Model for Further Hospitalised MIs by Average Adherence (80%)

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	0.1651
Not Achieved	1.07 (0.97, 1.17)	
Adjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	0.1520
Not Achieved	1.07 (0.97, 1.18)	
Sex		
Female (ref.)	1.00	0.9927
Male	1.00 (0.91, 1.10)	
Age at MI (years)	1.00 (1.00, 1.00)	0.5457
SIMD (NHS GGC)		
1 [most]	1.16 (1.00, 1.34)	0.0612
2	1.06 (0.92, 1.23)	
3	1.08 (0.93, 1.26)	
4	0.95 (0.81, 1.11)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	0.98 (0.83, 1.14)	
2011	1.03 (0.88, 1.20)	
2012	1.33 (1.14, 1.54)	
2013	1.27 (1.09, 1.47)	
2014	1.12 (0.94, 1.34)	

Appendix K Baseline Hazards from Kaplan-Meiers for Mortality by Average Statin Adherence

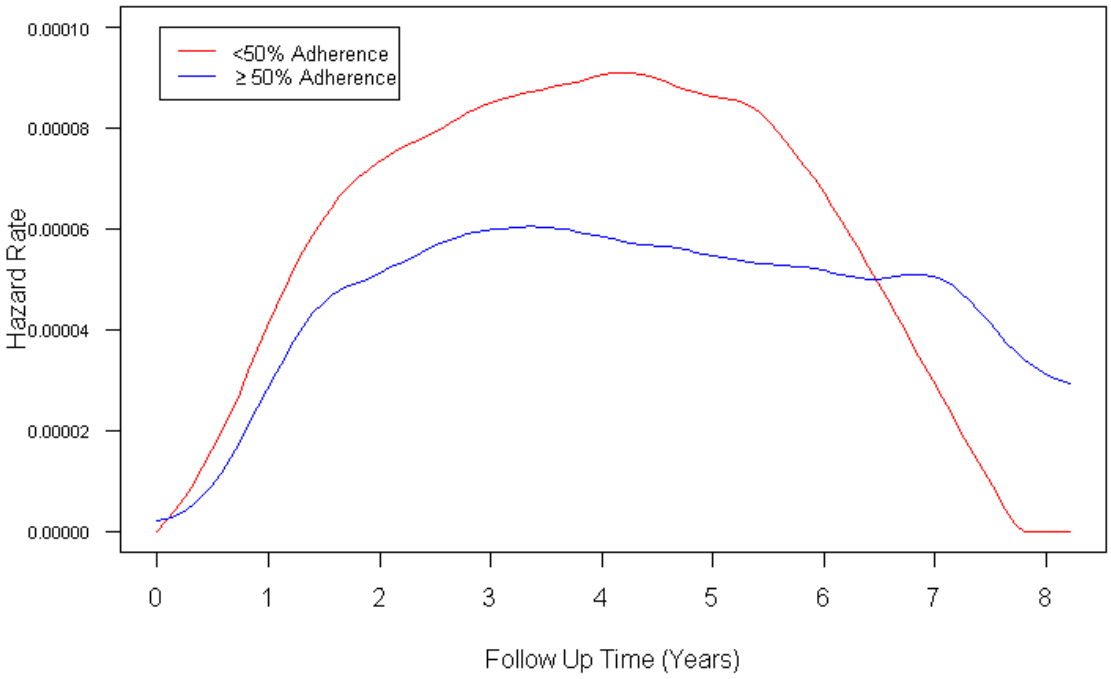
Appendix K1 Baseline Hazards for Overall Mortality



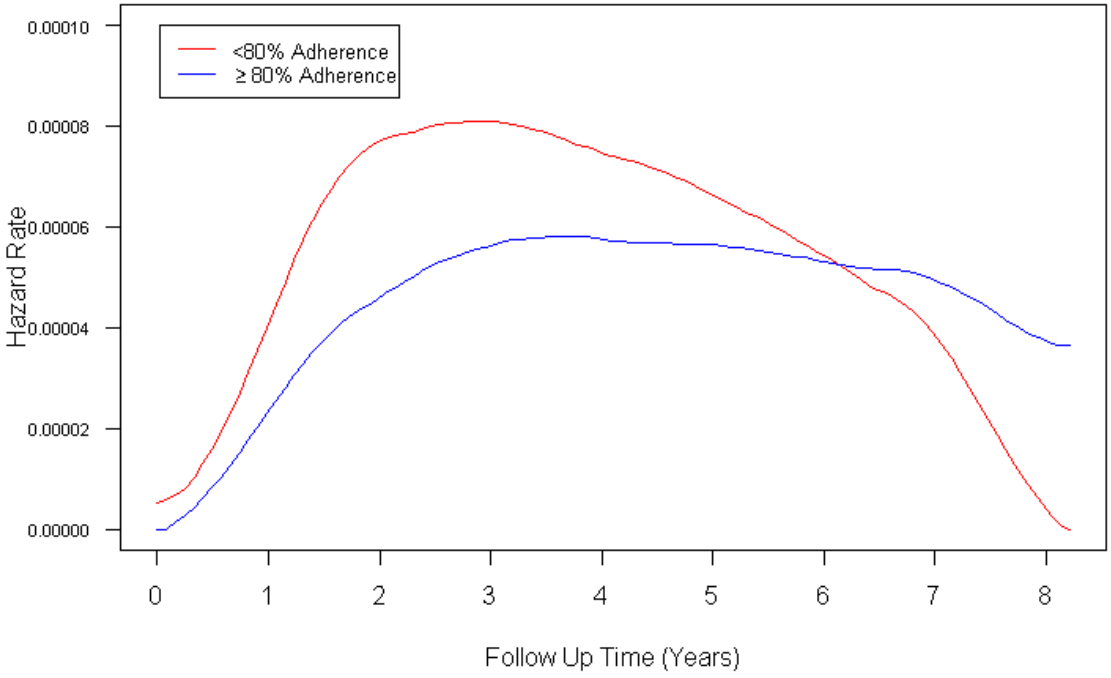


Appendix K2 Baseline Hazards for Circulatory Mortality

Baseline Hazard for Circulatory Survival (by 50% Average Adherence)



Baseline Hazard for Circulatory Survival (by 80% Average Adherence)



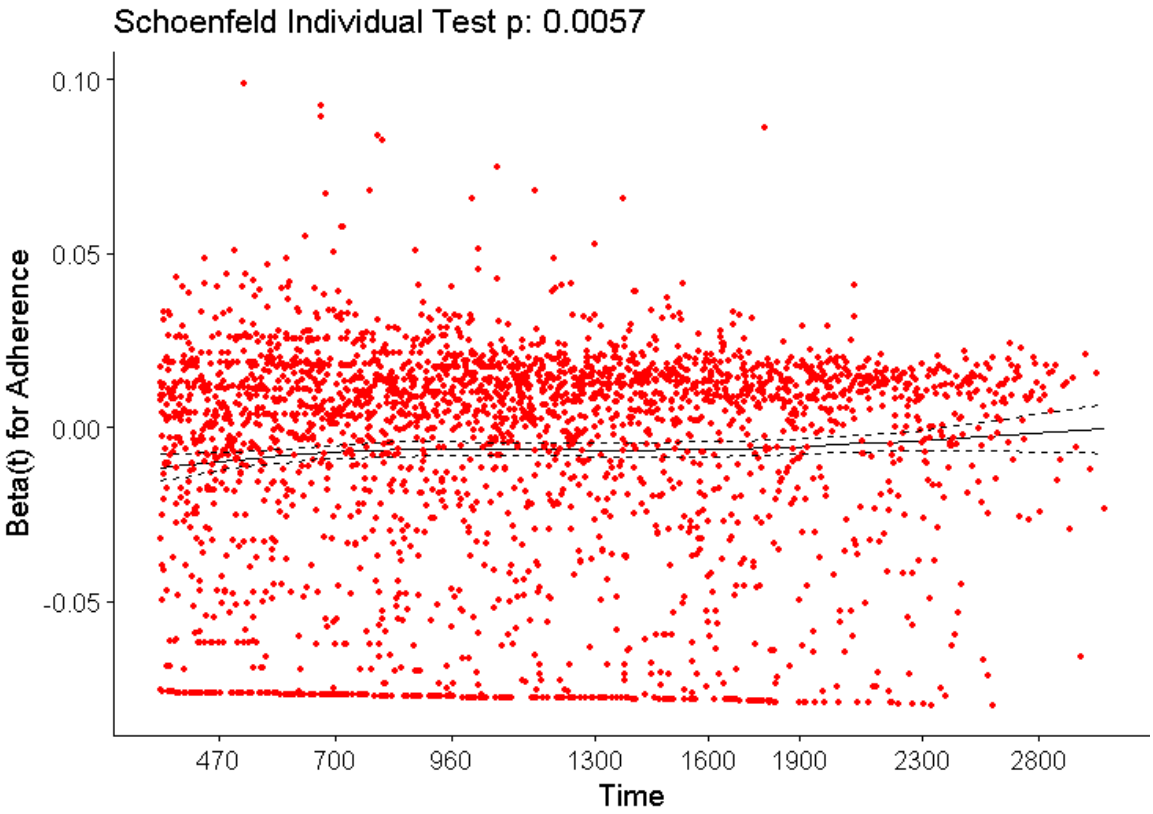
Appendix L Cox Regression Models for Mortality by Average Statin Adherence

Appendix L1 Model for Overall Mortality by Continuous Adherence

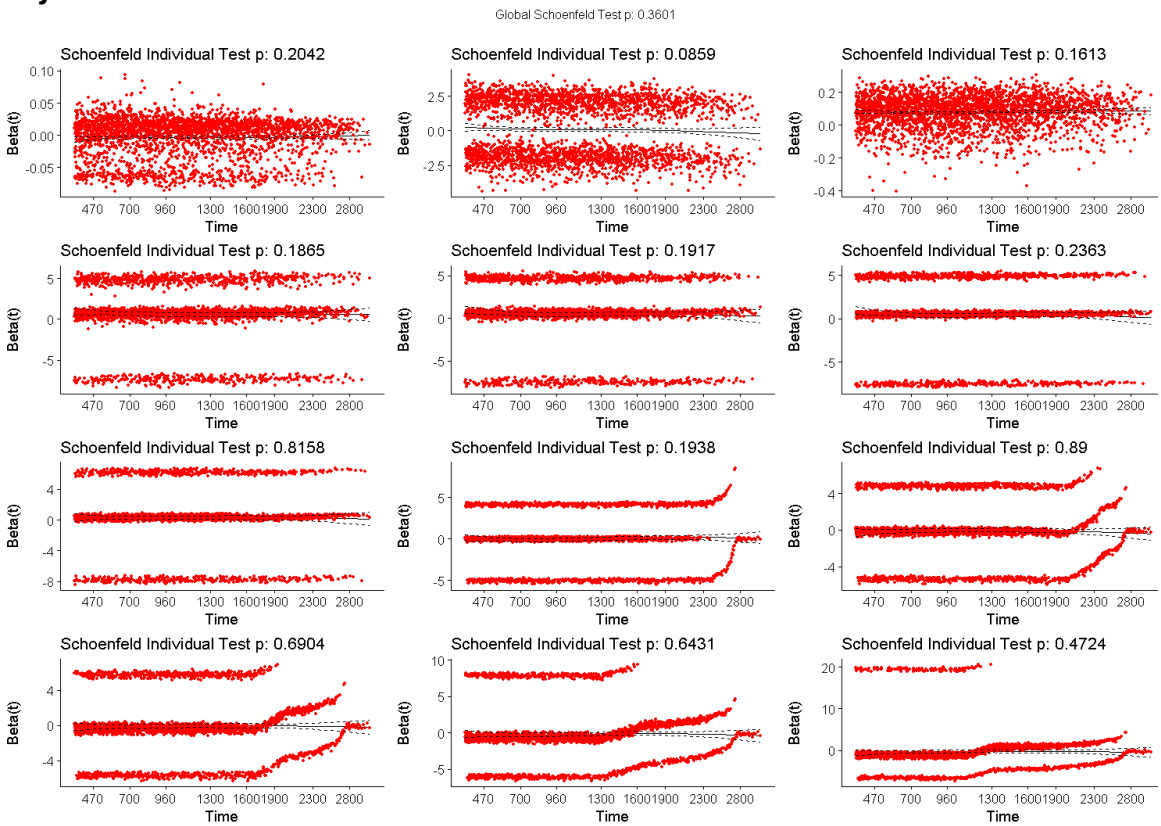
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Continuous Adherence			
10% Decrease	1.07 (1.06, 1.08)	<0.0001	0.0057
Adjusted Model			
Continuous Adherence			
10% Decrease	1.03 (1.02, 1.05)	<0.0001	0.2042
Sex			
Female (ref.)	1.00	0.0343	
Male	1.09 (1.01, 1.18)		0.0859
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.1613
SIMD (NHS GGC)			
1 [most]	1.91 (1.67, 2.18)		0.1865
2	1.70 (1.49, 1.94)		0.1917
3	1.63 (1.43, 1.86)	<0.0001	0.2363
4	1.34 (1.16, 1.54)		0.8158
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	1.01 (0.90, 1.13)		0.1938
2011	0.93 (0.82, 1.04)	0.0031	0.8900
2012	0.93 (0.82, 1.06)		0.6904
2013	0.83 (0.72, 0.96)		0.6431
2014	0.73 (0.60, 0.89)		0.4724
Global PH Test			0.3601

Appendix L2 Schoenfeld Residuals for Model for Overall Mortality by Continuous Adherence

Unadjusted Model



Adjusted Model

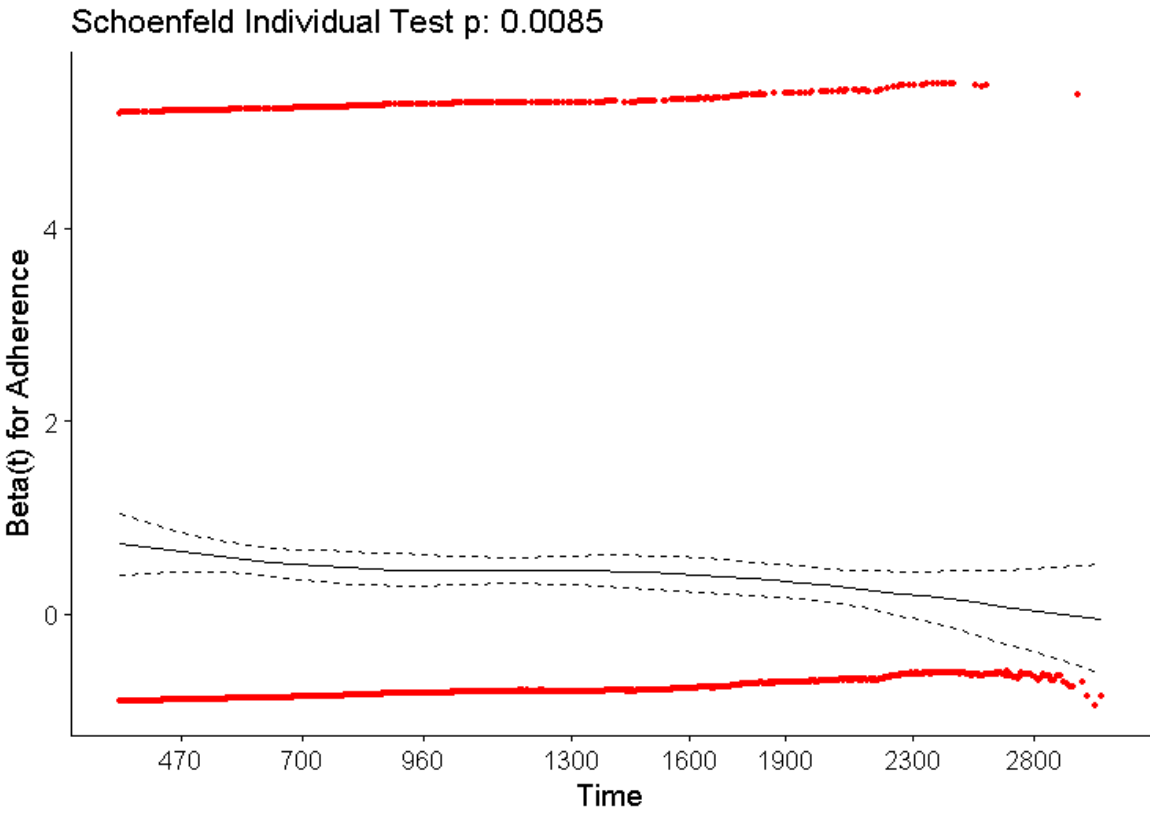


Appendix L3 Model for Overall Mortality by Average Adherence (50%)

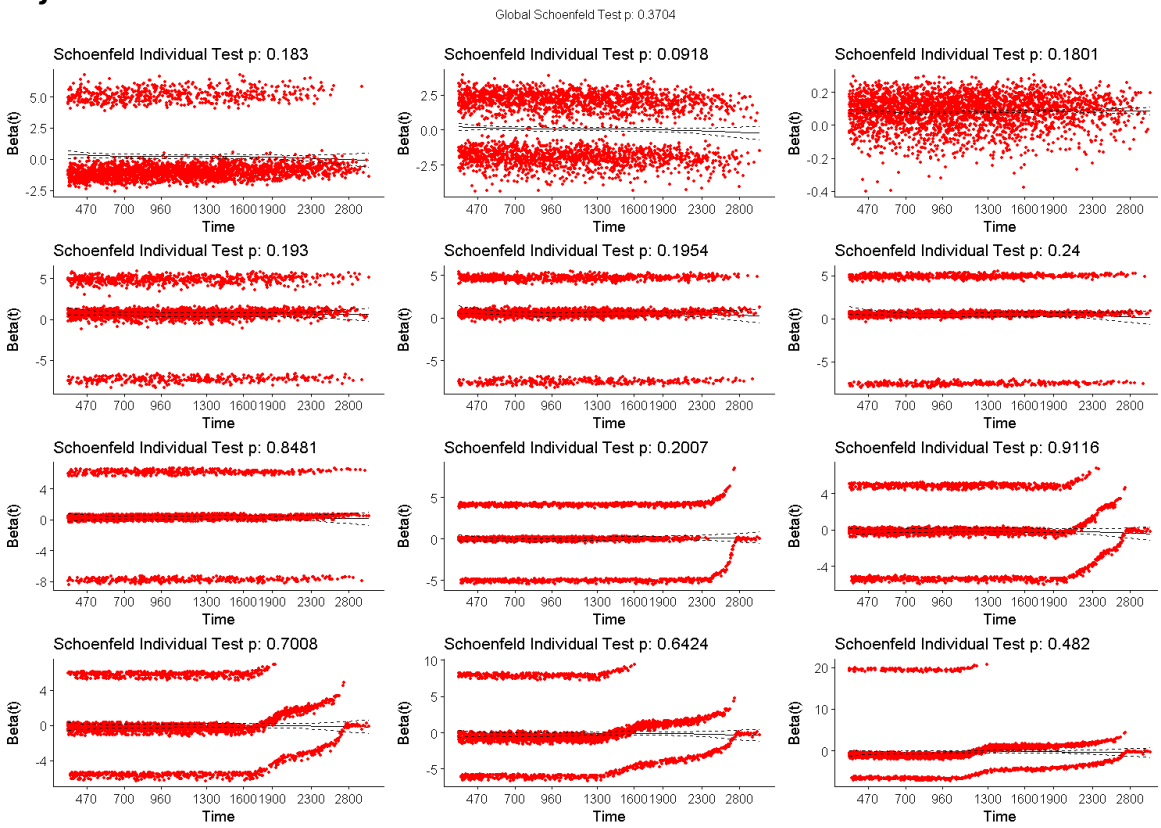
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Average Adherence (50%)			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.57 (1.44, 1.73)		0.0085
Adjusted Model			
Average Adherence (50%)			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.24 (1.13, 1.37)		0.1830
Sex			
Female (ref.)	1.00	0.0448	
Male	1.08 (1.00, 1.17)		0.0918
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.1801
SIMD (NHS GGC)			
1 [most]	1.91 (1.67, 2.18)		0.1930
2	1.69 (1.48, 1.93)		0.1954
3	1.63 (1.43, 1.86)	<0.0001	0.2400
4	1.34 (1.17, 1.54)		0.8481
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	1.01 (0.90, 1.13)		0.2007
2011	0.93 (0.82, 1.04)	0.0033	0.9116
2012	0.93 (0.82, 1.06)		0.7008
2013	0.83 (0.72, 0.96)		0.6424
2014	0.73 (0.61, 0.89)		0.4820
Global PH Test			0.3704

Appendix L4 Schoenfeld Residuals for Model for Overall Mortality by Average Adherence (50%)

Unadjusted Model



Adjusted Model

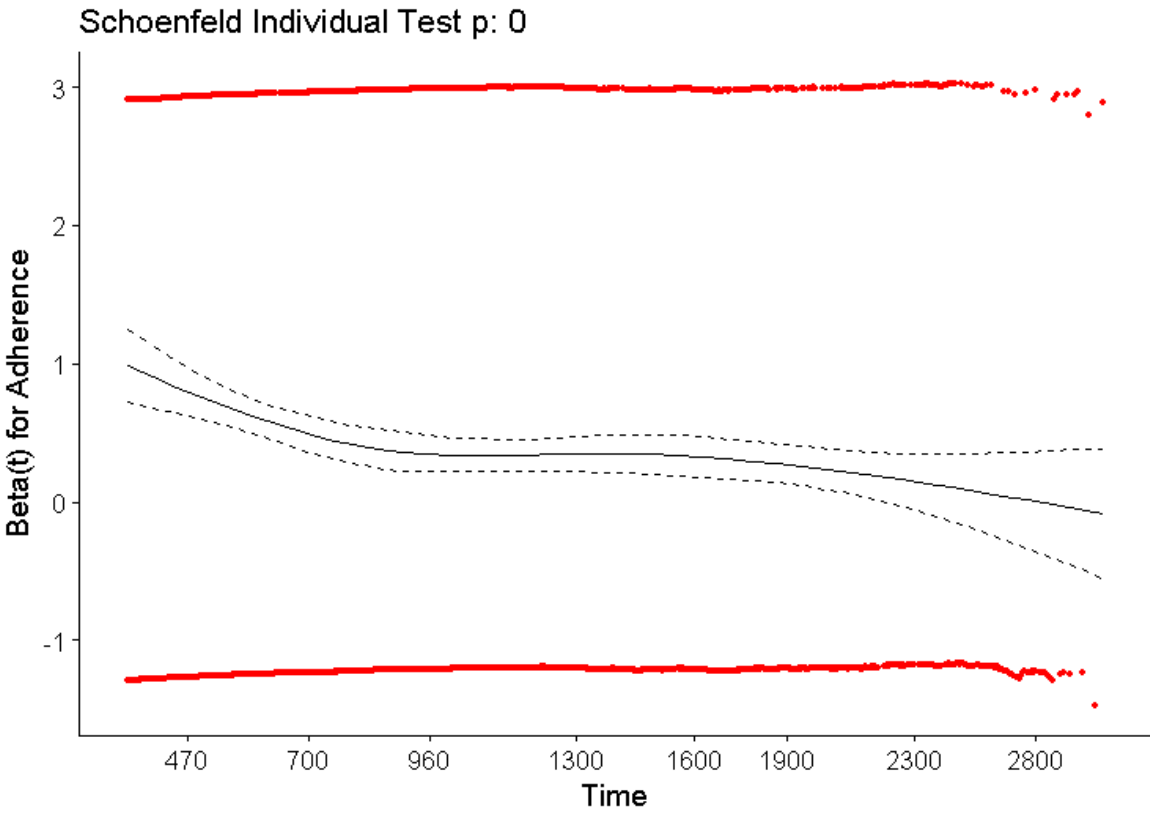


Appendix L5 Model for Overall Mortality by Average Adherence (80%)

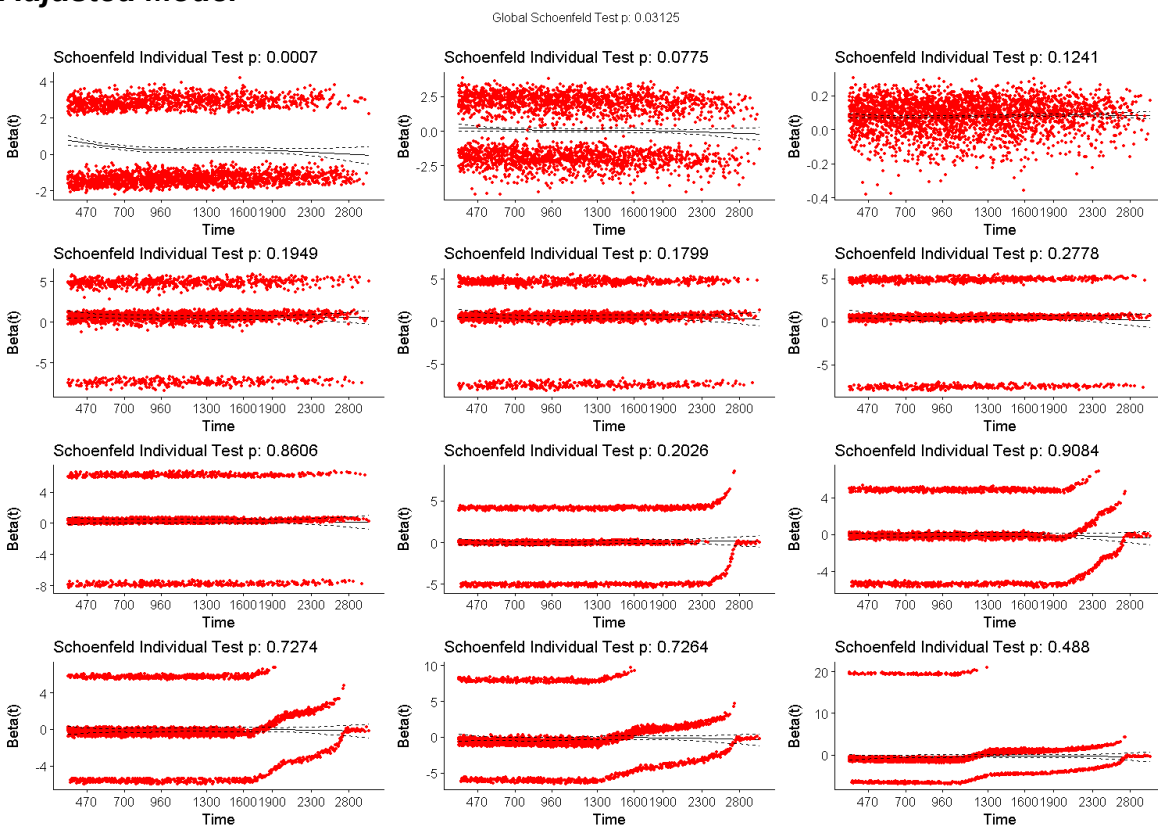
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Average Adherence (80%)			
Achieved (ref.)	1.00	<0.0001	<0.0001
Not Achieved	1.53 (1.41, 1.65)		
Adjusted Model			
Average Adherence (80%)			
Achieved (ref.)	1.00	<0.0001	0.0008
Not Achieved	1.37 (1.26, 1.48)		
Sex			
Female (ref.)	1.00	0.0337	0.0775
Male	1.09 (1.01, 1.18)		
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.1241
SIMD (NHS GGC)			
1 [most]	1.90 (1.67, 2.17)	<0.0001	0.1949
2	1.70 (1.49, 1.94)		0.1799
3	1.62 (1.42, 1.86)		0.2778
4	1.34 (1.16, 1.54)		0.8606
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00	0.0079	
2010	1.01 (0.90, 1.14)		0.2026
2011	0.93 (0.83, 1.05)		0.9084
2012	0.94 (0.83, 1.07)		0.7274
2013	0.85 (0.74, 0.98)		0.7264
2014	0.74 (0.61, 0.90)		0.4880
Global PH Test			0.0313

Appendix L6 Schoenfeld Residuals for Model for Overall Mortality by Average Adherence (80%)

Unadjusted Model



Adjusted Model

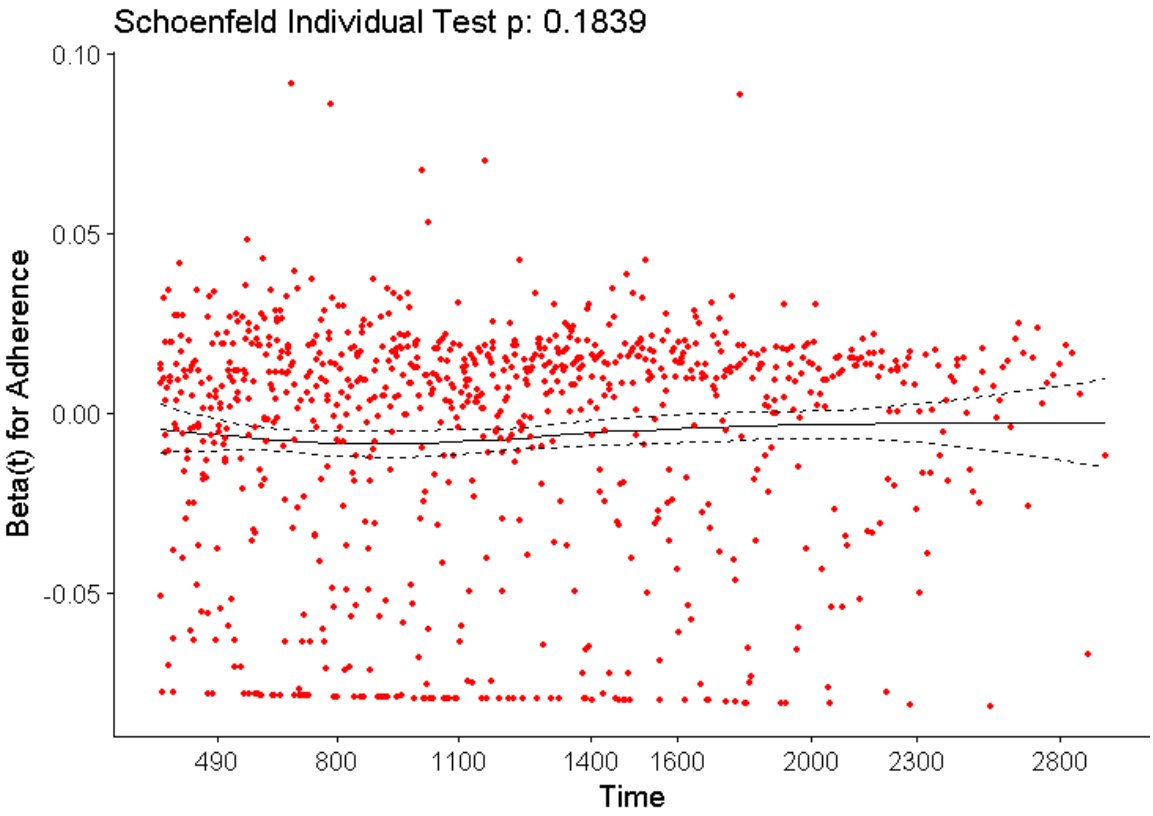


Appendix L7 Model for Circulatory Mortality by Continuous Adherence

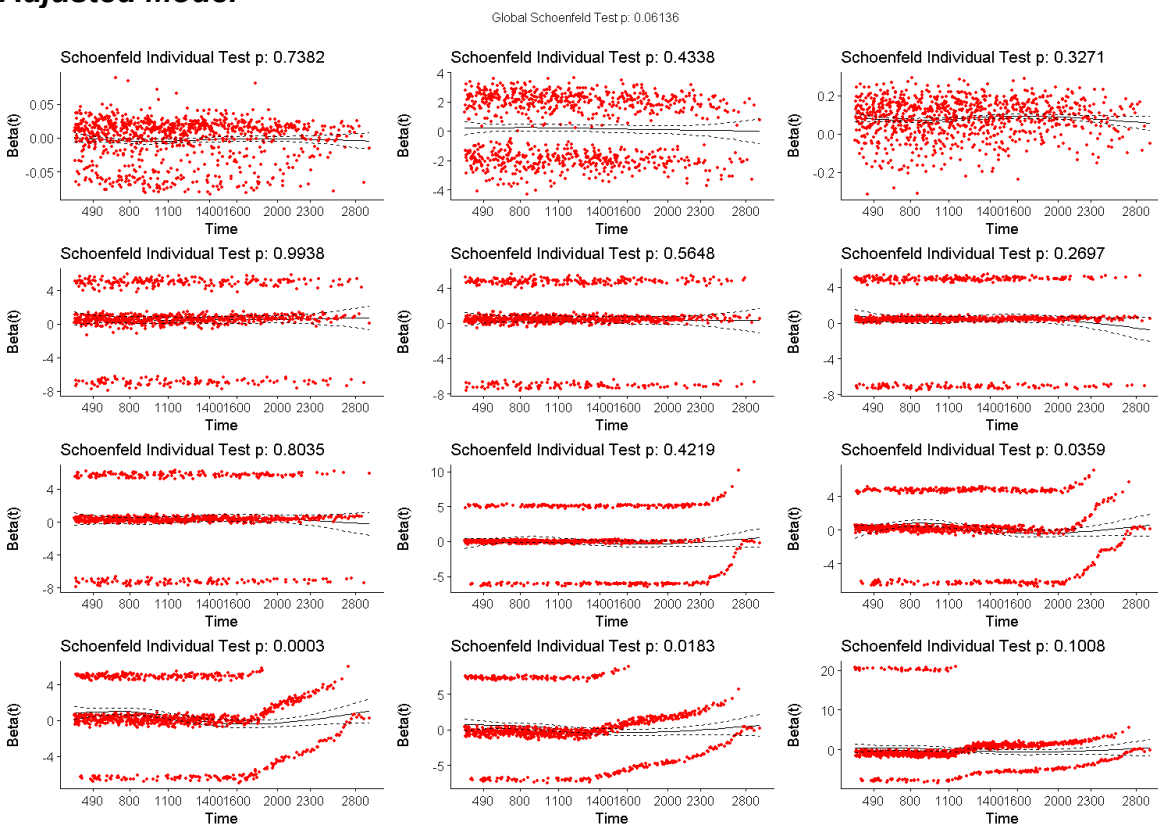
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Continuous Adherence			
10% Decrease	1.06 (1.04, 1.08)	<0.0001	0.1839
Adjusted Model			
Continuous Adherence			
10% Decrease	1.03 (1.01, 1.05)	0.0072	0.7382
Sex			
Female (ref.)	1.00	0.0039	
Male	1.23 (1.07, 1.41)		0.4338
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.3271
SIMD (NHS GGC)			
1 [most]	1.70 (1.35, 2.14)		0.9938
2	1.56 (1.25, 1.96)		0.5648
3	1.53 (1.22, 1.92)	<0.0001	0.2697
4	1.36 (1.08, 1.73)		0.8035
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	0.98 (0.79, 1.22)		0.4219
2011	1.23 (0.99, 1.53)	0.0005	0.0359
2012	1.51 (1.21, 1.88)		0.0003
2013	1.22 (0.95, 1.56)		0.0183
2014	0.96 (0.68, 1.36)		0.1008
Global PH Test			0.0614

Appendix L8 Schoenfeld Residuals for Model for Circulatory Mortality by Continuous Adherence

Unadjusted Model



Adjusted Model

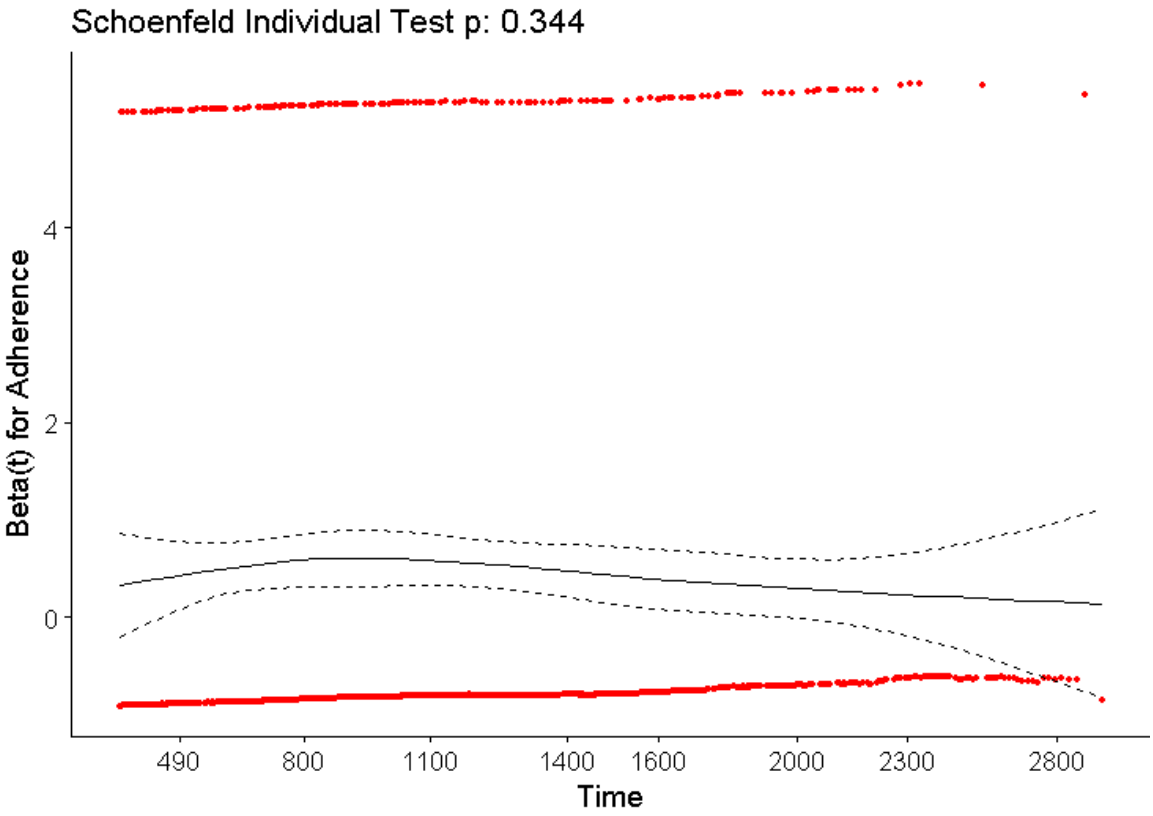


Appendix L9 Model for Circulatory Mortality by Average Adherence (50%)

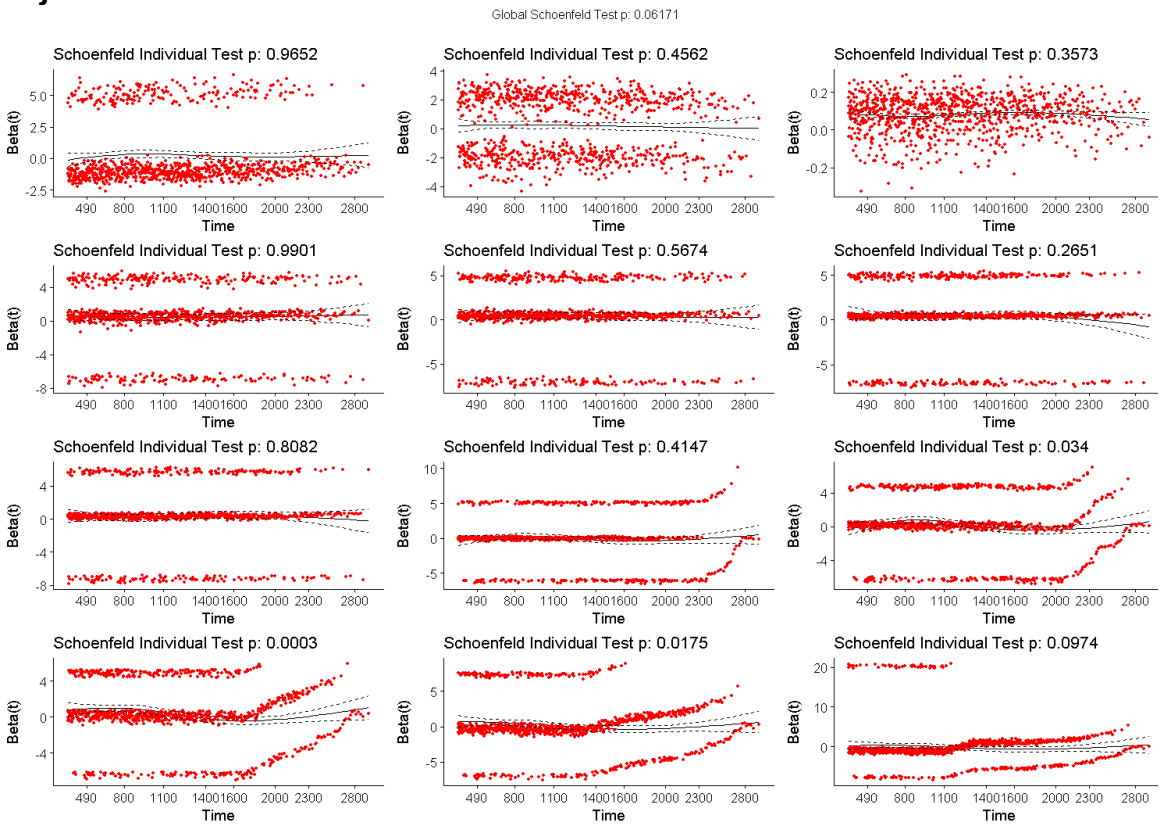
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Average Adherence (50%)			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.59 (1.35, 1.87)		0.3440
Adjusted Model			
Average Adherence (50%)			
Achieved (ref.)	1.00	0.0116	
Not Achieved	1.24 (1.05, 1.46)		0.9652
Sex			
Female (ref.)	1.00	0.0042	
Male	1.22 (1.07, 1.40)		0.4562
Age at MI (years)	1.09 (1.08, 1.09)	<0.0001	0.3573
SIMD (NHS GGC)			
1 [most]	1.70 (1.36, 2.14)		0.9901
2	1.56 (1.24, 1.95)		0.5674
3	1.53 (1.22, 1.92)	<0.0001	0.2651
4	1.37 (1.08, 1.73)		0.8082
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	0.98 (0.79, 1.22)		0.4147
2011	1.23 (0.99, 1.53)	0.0006	0.0340
2012	1.50 (1.21, 1.88)		0.0003
2013	1.22 (0.95, 1.56)		0.0175
2014	0.96 (0.68, 1.36)		0.0974
Global PH Test			0.0617

Appendix L10 Schoenfeld Residuals for Model for Circulatory Mortality by Average Adherence (50%)

Unadjusted Model



Adjusted Model

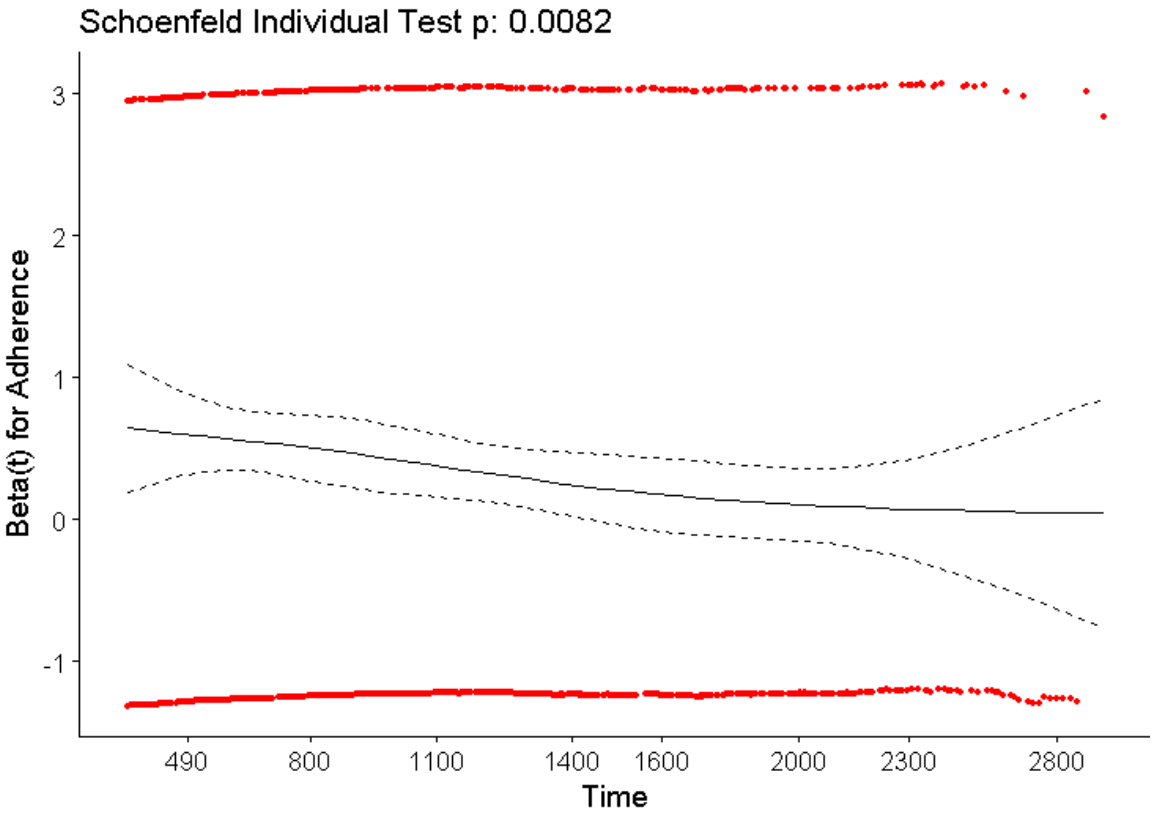


Appendix L11 Model for Circulatory Mortality by Average Adherence (80%)

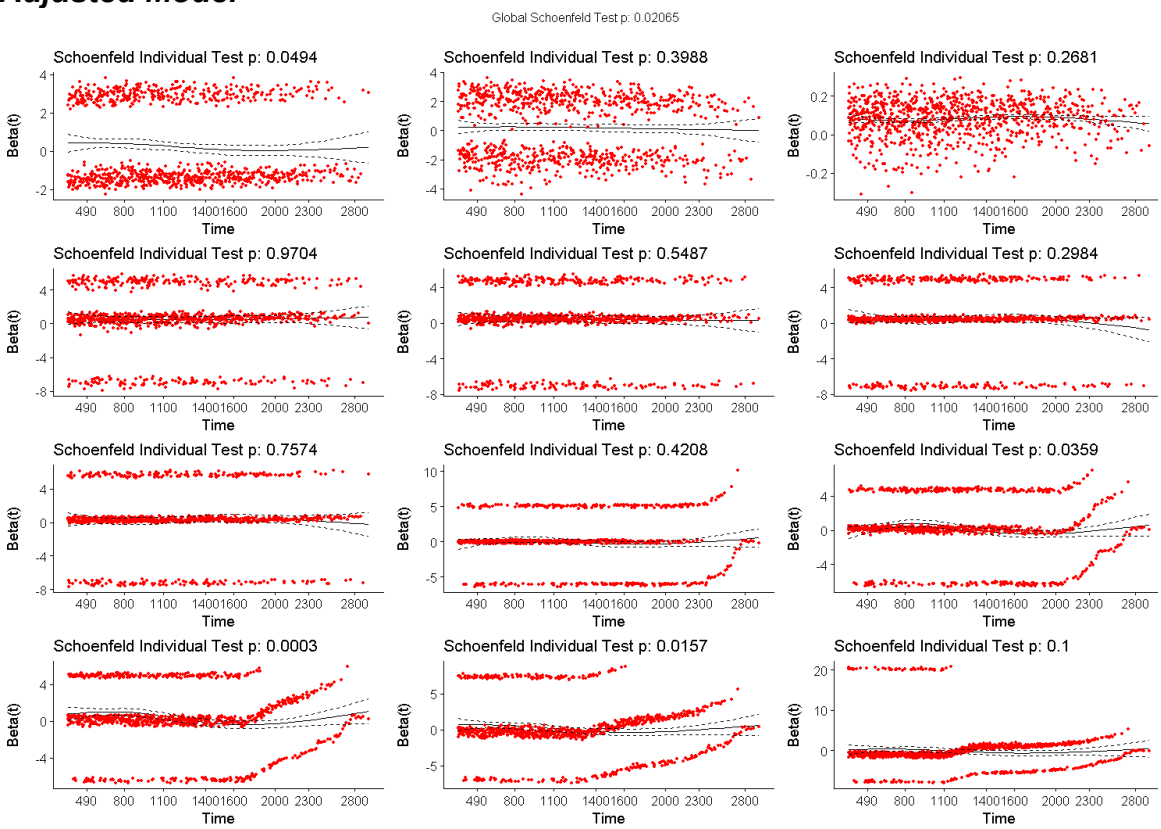
Model/Variable	Hazard Ratio (95% CI)	p-value	PH Test
Unadjusted Model			
Average Adherence (80%)			
Achieved (ref.)	1.00	<0.0001	
Not Achieved	1.42 (1.24, 1.63)		0.0082
Adjusted Model			
Average Adherence (80%)			
Achieved (ref.)	1.00	0.0003	
Not Achieved	1.29 (1.12, 1.48)		0.0494
Sex			
Female (ref.)	1.00	0.0039	
Male	1.23 (1.07, 1.41)		0.3988
Age at MI (years)	1.08 (1.08, 1.09)	<0.0001	0.2681
SIMD (NHS GGC)			
1 [most]	1.70 (1.35, 2.14)		0.9704
2	1.56 (1.25, 1.96)		0.5487
3	1.52 (1.21, 1.91)	<0.0001	0.2984
4	1.36 (1.08, 1.73)		0.7574
5 [least] (ref.)	1.00		
Year of MI			
2009 (ref.)	1.00		
2010	0.98 (0.79, 1.23)		0.4208
2011	1.24 (1.00, 1.54)	0.0004	0.0359
2012	1.52 (1.22, 1.90)		0.0003
2013	1.24 (0.97, 1.59)		0.0157
2014	0.97 (0.69, 1.38)		0.1000
Global PH Test			0.0207

Appendix L12 Schoenfeld Residuals for Model for Circulatory Mortality by Average Adherence (80%)

Unadjusted Model



Adjusted Model



Appendix M Logistic Regression Models for Not Achieving Target Non-HDL by Average Statin Adherence

Appendix M1 Model for Not Achieving Target Non-HDL by Continuous Adherence

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Continuous Adherence		
10% Decrease	1.24 (1.20, 1.27)	<0.0001
Adjusted Model		
Continuous Adherence		
10% Decrease	1.23 (1.20, 1.26)	<0.0001
Sex		
Female (ref.)	1.00	0.0062
Male	0.81 (0.69, 0.94)	
Age at MI (years)	1.01 (1.01, 1.02)	<0.0001
SIMD (NHS GGC)		
1 [most]	1.34 (1.05, 1.71)	
2	1.24 (0.97, 1.58)	
3	1.21 (0.94, 1.54)	0.2136
4	1.22 (0.95, 1.58)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	
2010	1.12 (0.82, 1.54)	
2011	1.12 (0.83, 1.51)	0.9033
2012	1.19 (0.88, 1.60)	
2013	1.16 (0.86, 1.56)	
2014	1.18 (0.85, 1.65)	

Appendix M2 Model for Not Achieving Target Non-HDL by Average Adherence (50%)

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	4.76 (3.86, 5.88)	
Adjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	4.47 (3.62, 5.54)	
Sex		
Female (ref.)	1.00	0.0046
Male	0.80 (0.69, 0.93)	
Age at MI (years)	1.01 (1.01, 1.02)	<0.0001
SIMD (NHS GGC)		
1 [most]	1.32 (1.04, 1.68)	
2	1.24 (0.97, 1.58)	
3	1.23 (0.96, 1.56)	0.2405
4	1.21 (0.94, 1.57)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	
2010	1.13 (0.82, 1.54)	
2011	1.12 (0.83, 1.51)	0.9455
2012	1.18 (0.88, 1.58)	
2013	1.13 (0.84, 1.51)	
2014	1.12 (0.81, 1.55)	

Appendix M3 Model for Not Achieving Target Non-HDL by Average Adherence (80%)

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	3.02 (2.59, 3.52)	
Adjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	2.98 (2.55, 3.48)	
Sex		
Female (ref.)	1.00	0.0014
Male	0.78 (0.67, 0.91)	
Age at MI (years)	1.02 (1.01, 1.02)	<0.0001
SIMD (NHS GGC)		
1 [most]	1.26 (0.99, 1.60)	
2	1.22 (0.96, 1.55)	
3	1.17 (0.92, 1.49)	0.3969
4	1.21 (0.94, 1.56)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	
2010	1.19 (0.87, 1.62)	
2011	1.14 (0.85, 1.54)	0.7523
2012	1.25 (0.93, 1.67)	
2013	1.23 (0.91, 1.65)	
2014	1.23 (0.89, 1.70)	

Appendix N Logistic Regression Models for Not Achieving Target LDL by Average Statin Adherence

Appendix N1 Model for Not Achieving Target LDL by Continuous Adherence

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Continuous Adherence		
10% Decrease	1.10 (1.08, 1.12)	<0.0001
Adjusted Model		
Continuous Adherence		
10% Decrease	1.11 (1.09, 1.12)	<0.0001
Sex		
Female (ref.)	1.00	0.0002
Male	0.82 (0.73, 0.91)	
Age at MI (years)	1.00 (0.99, 1.00)	0.0285
SIMD (NHS GGC)		
1 [most]	0.81 (0.68, 0.97)	<0.0001
2	1.23 (1.04, 1.45)	
3	1.28 (1.09, 1.52)	
4	1.52 (1.28, 1.81)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	1.11 (0.92, 1.34)	
2011	1.39 (1.16, 1.66)	
2012	1.40 (1.17, 1.68)	
2013	1.67 (1.39, 1.99)	
2014	4.14 (3.41, 5.02)	

Appendix N2 Model for Not Achieving Target LDL by Average Adherence (50%)

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	1.98 (1.74, 2.24)	
Adjusted Model		
Average Adherence (50%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	2.03 (1.78, 2.31)	
Sex		
Female (ref.)	1.00	<0.0001
Male	0.81 (0.73, 0.90)	
Age at MI (years)	1.00 (0.99, 1.00)	0.0306
SIMD (NHS GGC)		
1 [most]	0.81 (0.68, 0.97)	<0.0001
2	1.22 (1.04, 1.45)	
3	1.29 (1.09, 1.53)	
4	1.51 (1.27, 1.80)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	1.11 (0.92, 1.33)	
2011	1.38 (1.15, 1.65)	
2012	1.39 (1.16, 1.66)	
2013	1.63 (1.37, 1.96)	
2014	4.05 (3.34, 4.91)	

Appendix N3 Model for Not Achieving Target LDL by Average Adherence (80%)

Model/Variable	Odds Ratio (95% CI)	p-value
Unadjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	1.60 (1.45, 1.78)	
Adjusted Model		
Average Adherence (80%)		
Achieved (ref.)	1.00	<0.0001
Not Achieved	1.70 (1.53, 1.88)	
Sex		
Female (ref.)	1.00	<0.0001
Male	0.80 (0.72, 0.89)	
Age at MI (years)	1.00 (0.99, 1.00)	0.0757
SIMD (NHS GGC)		
1 [most]	0.80 (0.67, 0.95)	<0.0001
2	1.21 (1.02, 1.42)	
3	1.27 (1.07, 1.50)	
4	1.51 (1.27, 1.79)	
5 [least] (ref.)	1.00	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	1.12 (0.93, 1.35)	
2011	1.40 (1.17, 1.68)	
2012	1.43 (1.19, 1.71)	
2013	1.70 (1.42, 2.03)	
2014	4.16 (3.43, 5.05)	

Appendix O Cox Regression Models for Time to Non-Adherence with Lipid Testing Frequency

Appendix O1 Optimal Model for Time to Non-Adherence (50%) with Lipid Testing Frequency

Model

Variable	Hazard Ratio (95% CI)	p-value
Sex		
Female (ref.)	1.00	<0.0001
Male	0.82 (0.76-0.89)	
Age at MI (years)		
<50 (ref.)	1.00	
50-60	0.84 (0.74-0.95)	
60-70	0.78 (0.69-0.88)	<0.0001
70-80	1.01 (0.90-1.13)	
≥80	1.33 (1.18-1.51)	
SIMD (NHS GGC)		
1 [most]	1.02 (0.91-1.15)	
2	0.97 (0.87-1.09)	
3	1.12 (1.00-1.26)	0.1118
4	1.04 (0.92-1.18)	
5 [least] (ref.)	1.00	
Testing Frequency		
0 Tests	1.59 (1.43-1.78)	
1 Test (ref.)	1.00	<0.0001
>1 Test	1.13 (1.01-1.25)	
AIC	55302.72	

Appendix O2 Optimal Model for Time to Non-Adherence (80%) with Lipid Testing Frequency

Model

Variable	Hazard Ratio (95% CI)	p-value
Sex		
Female (ref.)	1.00	0.0018
Male	0.92 (0.87-0.97)	
Age at MI (years)		
<50 (ref.)	1.00	<0.0001
50-60	0.89 (0.82-0.97)	
60-70	0.81 (0.74-0.88)	
70-80	0.93 (0.86-1.02)	
≥80	1.12 (1.03-1.23)	
SIMD (NHS GGC)		
1 [most]	1.09 (1.01-1.19)	0.0352
2	1.06 (0.98-1.15)	
3	1.09 (1.00-1.19)	
4	0.99 (0.90-1.08)	
5 [least] (ref.)	1.00	
Further MI Last Year		
No (ref.)	1.00	0.0737
Yes	1.11 (0.99-1.25)	
Year of MI		
2009 (ref.)	1.00	<0.0001
2010	0.95 (0.87-1.03)	
2011	0.89 (0.82-0.97)	
2012	0.87 (0.80-0.94)	
2013	0.78 (0.72-0.85)	
2014	0.79 (0.71-0.88)	
Testing Frequency		
0 Tests	1.26 (1.14-1.40)	<0.0001
1 Test (ref.)	1.00	
>1 Test	0.99 (0.91-1.08)	
AIC	108472.8	

Appendix P Optimal Latent Class Analysis for Adherence

Appendix P1 Model for 50% Adherence

Conditional Class Probabilities

	Class 1 (Most)	**Class 2**	**Class 3**	**Class 4** (Least)
Sex				
Male	0.3127	0.8983	0.7127	0.5822
Female	0.6873	0.1017	0.2873	0.4178
Age at MI (years)				
<50	0.0000	0.1276	0.2848	0.1597
50-60	0.1034	0.2800	0.3743	0.2422
60-70	0.2385	0.3381	0.2399	0.2173
70-80	0.4091	0.1923	0.1010	0.2476
≥80	0.2490	0.0620	0.0000	0.1331
SIMD (NHS GGC)				
1 (Most)	0.1924	0.0000	0.4967	0.1980
2	0.2654	0.1570	0.2755	0.2199
3	0.2091	0.2181	0.1704	0.2316
4	0.1584	0.2882	0.0573	0.1772
5 (Least)	0.1747	0.3367	0.0000	0.1733
50% Adherent Last Year				
No	0.0137	0.0156	0.0312	0.5992
Yes	0.9863	0.9844	0.9688	0.4008
LDL on Target Last Year				
No	0.4096	0.4654	0.4164	0.5706
Yes	0.5904	0.5346	0.5836	0.4294
Est. Class Pop. Shares	0.2424	0.3080	0.3001	0.1494
Pred. Class Memberships	0.2169	0.3339	0.3070	0.1421
Maximum Log-Likelihood				-123644.0
BIC				247797.4
AIC				247388.1
# Iterations to Converge				795 (of 7000)
Tolerance				$1e^{-7}$

Regression Coefficients

Regression	Coefficient	SE(Coef)	exp(Coef)	p-value
2 vs 1				
Intercept	1.37290	0.04256	3.94678	<0.001
Meeting Target	-1.13353	0.04256	0.32189	<0.001
3 vs 1				
Intercept	3.42169	0.03478	30.62112	<0.001
Meeting Target	-3.20819	0.03478	0.04043	<0.001
4 vs 1				
Intercept	20.63205	0.11316	912820669.2	<0.001
Meeting Target	-23.17646	0.11316	<0.00001	<0.001

Model Comparisons

# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
2	-124597.2	24	249438.8	0.9116
3	-123856.3	37	248089.5	0.7488*
4	-123644.0	50	247797.4	0.6015

*To calculate relative entropy, those with posterior predictions of zero were set to the smallest non-zero value in the matrix, $3.063207e^{-322}$

Appendix P2 Model for 80% Adherence

Conditional Class Probabilities

	Class 1 (Most)	**Class 2**	**Class 3**	**Class 4** (Least)
Sex				
Male	0.5503	1.0000	0.2827	0.6162
Female	0.4497	0.0000	0.7173	0.3838
Age at MI (years)				
<50	0.0275	0.2406	0.1635	0.1837
50-60	0.1603	0.3700	0.2193	0.2841
60-70	0.2874	0.2861	0.2429	0.2130
70-80	0.3358	0.1033	0.2620	0.2112
≥80	0.1890	0.0000	0.1123	0.1080
SIMD (NHS GGC)				
1 (Most)	0.0000	0.2669	0.5183	0.2191
2	0.2012	0.2186	0.2914	0.2219
3	0.2301	0.1914	0.1589	0.2330
4	0.2665	0.1664	0.0314	0.1655
5 (Least)	0.3022	0.1568	0.0000	0.1604
Further MI Last Year				
No	0.9283	0.9320	0.9085	0.9159
Yes	0.0717	0.0680	0.0915	0.0841
Year of MI				
2009	0.2025	0.2397	0.1912	0.2260
2010	0.2210	0.2332	0.2133	0.2103
2011	0.1954	0.1934	0.2157	0.2032
2012	0.1681	0.1503	0.1834	0.1841
2013	0.1414	0.1188	0.1339	0.1180
2014	0.0716	0.0646	0.0625	0.0585
80% Adherent Last Year				
No	0.1151	0.1275	0.1228	0.9979
Yes	0.8849	0.8725	0.8772	0.0021
LDL on Target Last Year				
No	0.4676	0.4217	0.3814	0.5801
Yes	0.5324	0.5783	0.6186	0.4199
Est. Class Pop. Shares	0.3171	0.3260	0.2010	0.1558
Pred. Class Memberships	0.2971	0.3975	0.1418	0.1636
Maximum Log-Likelihood				-183050.0
BIC				366853.9
AIC				366248.1
# Iterations to Converge				2638 (of 7000)
Tolerance				$1e^{-7}$

Regression Coefficients

Regression	Coefficient	SE(Coef)	Exp(Coef)	p-value
2 vs 1				
Intercept	0.03208	0.08460	1.03260	0.705
Meeting 80% Target	-0.00512	0.07300	0.99489	0.944
3 vs 1				
Intercept	-0.39313	0.10055	0.67494	<0.001
Meeting 80% Target	-0.07210	0.08198	0.93044	0.379
4 vs 1				
Intercept	1.34134	0.08663	3.82416	<0.001
Meeting 80% Target	-5.21688	0.64308	0.00524	<0.001

Model Comparisons

# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
2	-184046.6	36	368460.0	0.7753
3	-183354.4	55	367269.0	0.7367
4	-183050.0	74	366853.9	0.6231

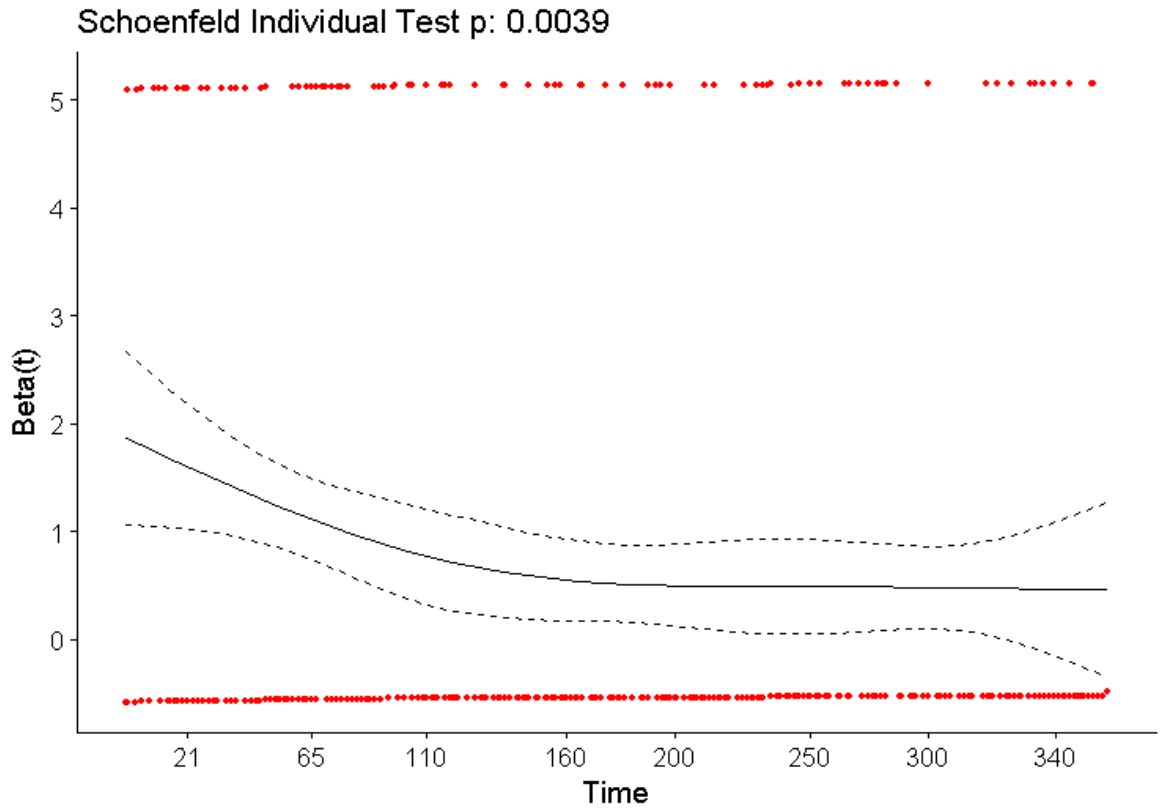
Appendix Q Models for Mortality by Predicted Class from Latent Class Analysis (50% Adherence)

Appendix Q1 Models for Year 2

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1,2,3 [Most] (ref.)	1.00	<0.0001	
4 [Least]	2.13 (1.70, 2.67)		
Global PH Test			0.0039

Schoenfeld Residual Plot

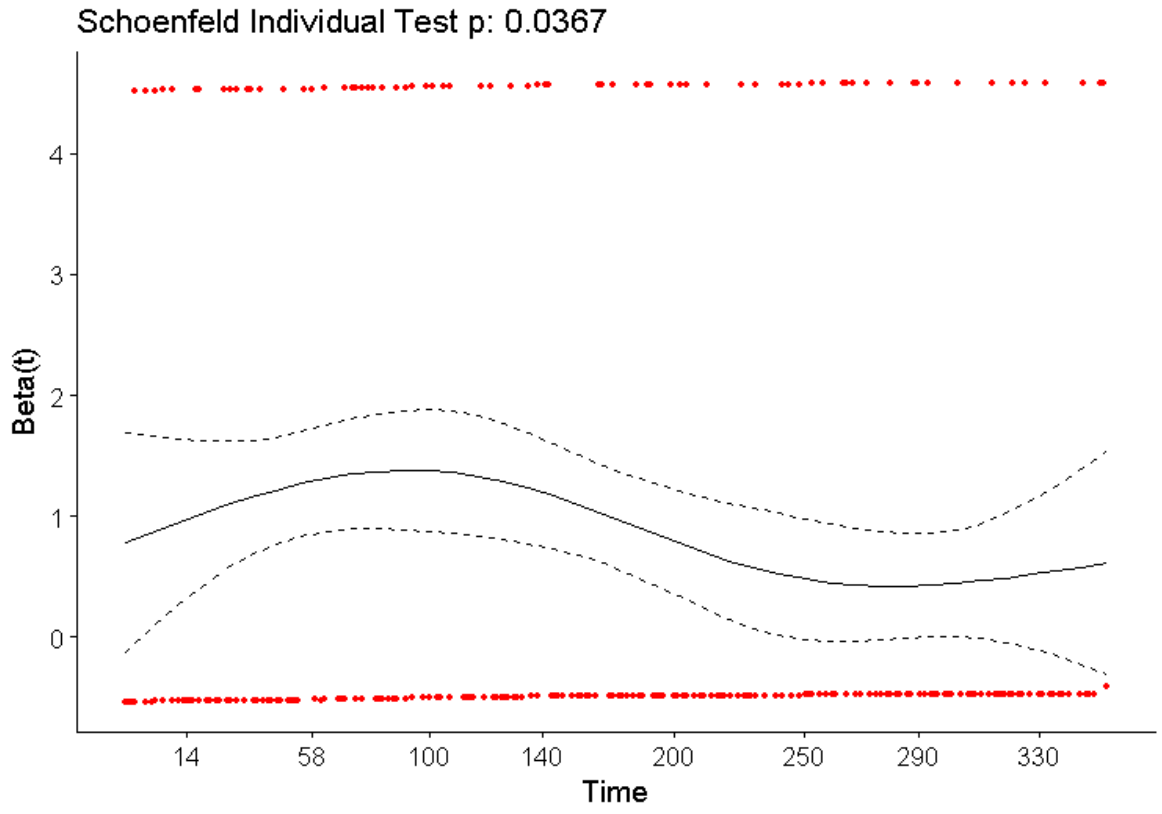


Appendix Q2 Models for Year 3

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1,2,3 [Most] (ref.)	1.00	<0.0001	
4 [Least]	2.40 (1.86,3.11)		
Global PH Test			0.0367

Schoenfeld Residual Plot



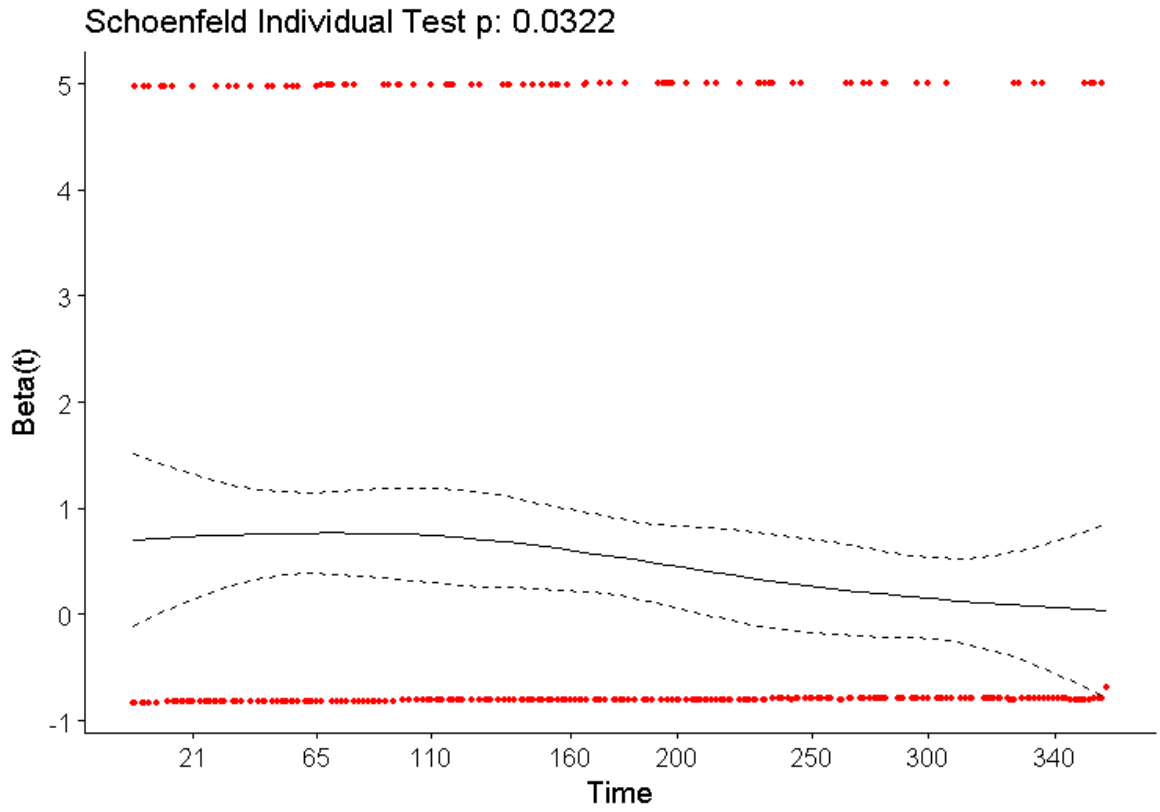
Appendix R Models for Mortality by Predicted Class from Latent Class Analysis (80% Adherence)

Appendix R1 Models for Year 2

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1,2,3 [Most] (ref.)	1.00	0.0001	
4 [Least]	1.61 (1.28, 2.03)		
Global PH Test			0.0322

Schoenfeld Residual Plot

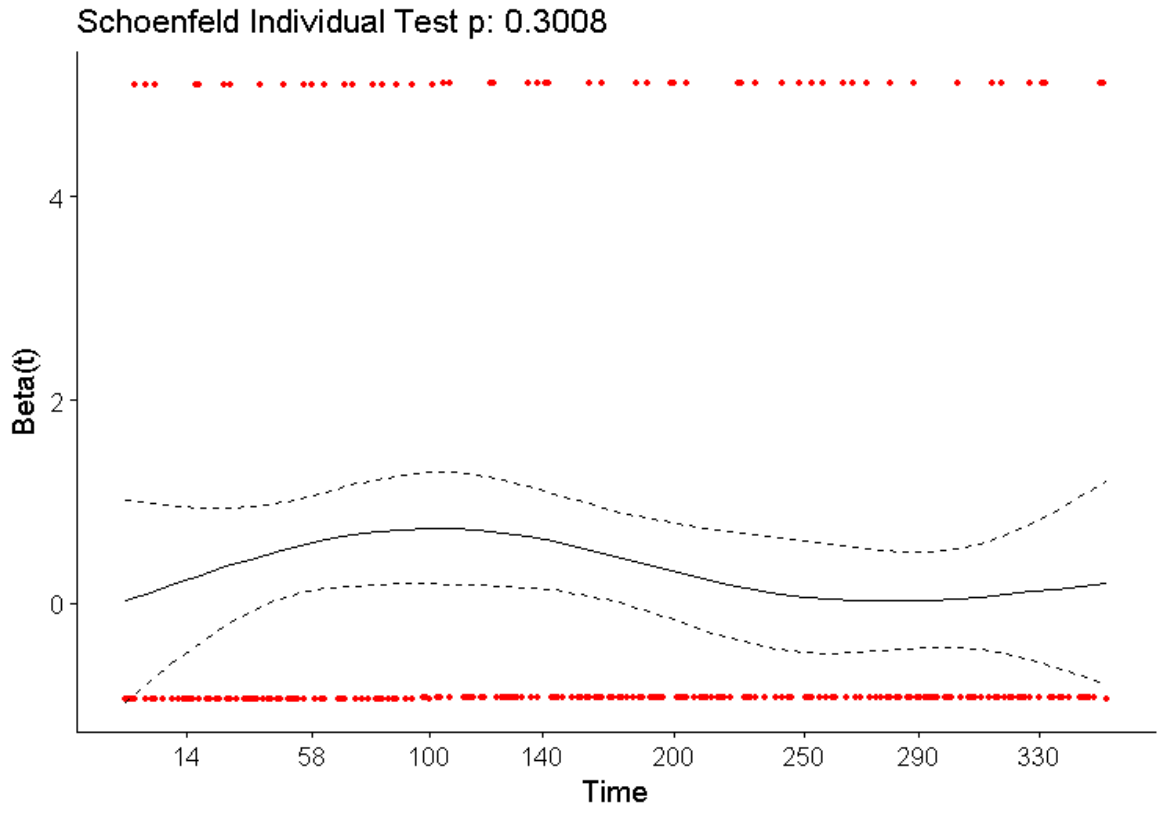


Appendix R2 Models for Year 3

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1,2,3 [Most] (ref.)	1.00	0.0239	
4 [Least]	1.40 (1.06, 1.86)		
Global PH Test			0.3008

Schoenfeld Residual Plot



Appendix S Optimal Latent Class Growth Analyses for Time to Non-Adherence

Appendix S1 Model for 50% Adherence

Class Proportions

	Class 1	Class 2	Class 3	Class 4
	7,902 (75.8%)	863 (8.3%)	536 (5.1%)	1,117 (10.7%)
Sex				
Male	5,017 (63.5%)	494 (57.2%)	338 (63.1%)	554 (49.6%)
Female	2,885 (36.5%)	369 (42.8%)	198 (36.9%)	563 (50.4%)
Age at MI (years)				
<50	1,014 (12.8%)	102 (11.8%)	100 (18.7%)	145 (13.0%)
50-60	1,780 (22.5%)	180 (20.9%)	128 (23.9%)	156 (14.0%)
60-70	1,916 (24.3%)	184 (21.3%)	116 (21.6%)	194 (17.4%)
70-80	1,877 (23.8%)	221 (25.6%)	125 (23.3%)	282 (25.2%)
≥80	1,315 (16.6%)	176 (20.4%)	67 (12.5%)	340 (30.4%)
SIMD (NHS GGC)				
1 (Most)	1,940 (24.6%)	190 (22.0%)	191 (35.6%)	196 (17.5%)
2	1,850 (23.4%)	204 (23.6%)	133 (24.8%)	234 (20.9%)
3	1,599 (20.2%)	175 (20.3%)	111 (20.7%)	269 (24.1%)
4	1,327 (16.8%)	140 (16.2%)	60 (11.2%)	230 (20.6%)
5 (Least)	1,186 (15.0%)	154 (17.8%)	41 (7.6%)	188 (16.8%)

Model Comparisons

# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
1	-22341.17	2	44700.86	1.0000
2	-16386.10	8	32846.21	0.8289
3	-15883.88	14	31897.27	0.7605
4	-15503.65	20	31192.33	0.7639

Appendix S2 Model for 80% Adherence

Class Proportions

	Class 1	Class 2	Class 3	Class 4
	5,000 (48.0%)	2,692 (25.8%)	782 (7.5%)	1,944 (18.7%)
Sex				
Male	3,155 (63.1%)	1,676 (62.3%)	540 (69.1%)	1,032 (53.1%)
Female	1,845 (36.9%)	1,106 (37.7%)	242 (31.0%)	912 (46.9%)
Age at MI (years)				
<50	599 (12.0%)	344 (12.8%)	131 (16.8%)	287 (14.8%)
50-60	1,121 (22.4%)	548 (20.4%)	206 (26.3%)	369 (19.0%)
60-70	1,292 (25.8%)	600 (22.3%)	174 (22.3%)	344 (17.7%)
70-80	1,238 (24.8%)	660 (24.5%)	173 (22.1%)	434 (22.3%)
≥80	750 (15.0%)	540 (20.1%)	98 (12.5%)	510 (26.2%)
SIMD (NHS GGC)				
1 (Most)	1,131 (22.6%)	705 (26.2%)	266 (34.0%)	415 (21.3%)
2	1,153 (23.1%)	661 (24.6%)	204 (26.1%)	403 (20.7%)
3	1,022 (20.4%)	520 (19.3%)	155 (19.8%)	457 (23.5%)
4	884 (17.7%)	420 (15.6%)	92 (11.8%)	361 (18.6%)
5 (Least)	810 (16.2%)	386 (14.3%)	65 (8.3%)	308 (15.8%)
Year of MI				
2009	658 (13.2%)	521 (19.4%)	249 (31.8%)	237 (12.2%)
2010	820 (16.4%)	564 (21.0%)	200 (25.6%)	292 (15.0%)
2011	913 (18.3%)	533 (19.8%)	136 (17.4%)	369 (19.0%)
2012	955 (19.1%)	450 (16.7%)	116 (14.8%)	445 (22.9%)
2013	1,042 (20.8%)	411 (15.3%)	62 (7.9%)	374 (19.2%)
2014	612 (12.2%)	213 (7.9%)	19 (2.4%)	227 (11.7%)

Model Comparisons

# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
1	-32551.84	2	65122.20	1.0000
2	-26763.31	9	53609.87	0.7662
3	-26327.46	16	52802.95	0.5951
4	-26023.90	23	52260.57	0.5914

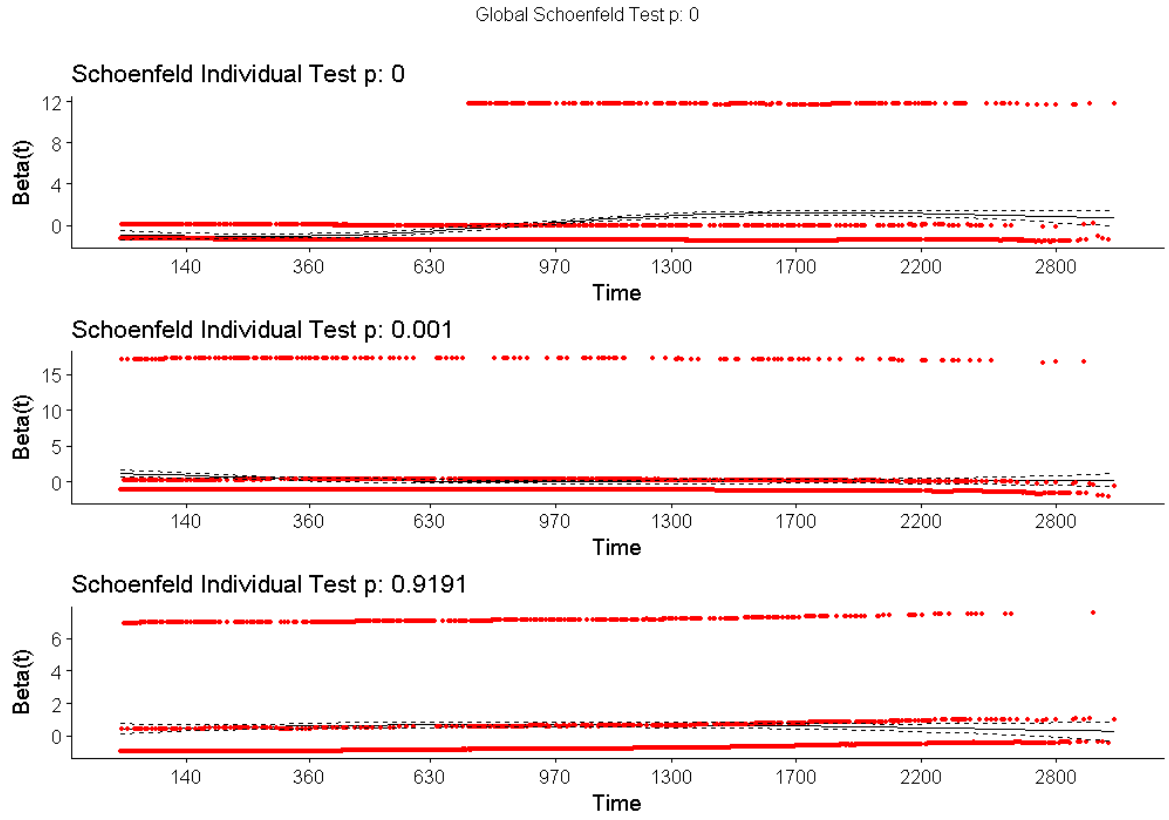
Appendix T Cox Models for Mortality by Predicted Class from Latent Class Growth Analysis (Adherence)

Appendix T1 Model from 50% Adherence Analysis

Model

	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 (ref.)	1.00		
2	0.99 (0.87, 1.12)	<0.0001	<0.0001
3	1.26 (1.09, 1.46)		0.0010
4	1.84 (1.67, 2.02)		0.9191
Global PH Test			<0.0001

Schoenfeld Residual Plot

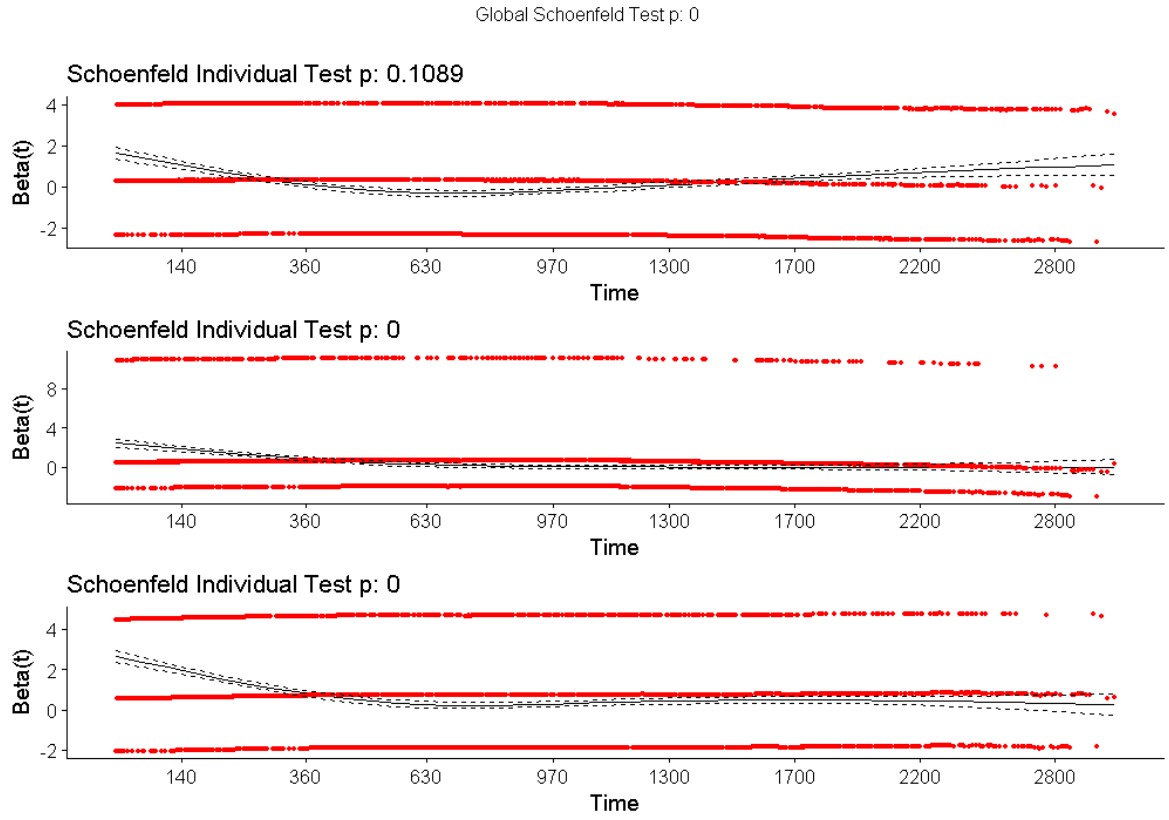


Appendix T2 Model from 80% Adherence Analysis

Model

	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 (ref.)	1.00		
2	1.30 (1.19, 1.42)	<0.0001	0.1089
3	1.71 (1.51, 1.93)		<0.0001
4	2.04 (1.87, 2.22)		<0.0001
Global PH Test			<0.0001

Schoenfeld Residual Plot



Appendix U Cox Regression Model for Time to Non-Target LDL with Lipid Testing Frequency

Model

Variable	Hazard Ratio (95% CI)	p-value
Sex		
Female (ref.)	1.00	0.0001
Male	0.90 (0.85-0.95)	
Age at MI (years)		
<50 (ref.)	1.00	
50-60	1.05 (0.96-1.14)	
60-70	1.05 (0.97-1.14)	<0.0001
70-80	0.81 (0.74-0.89)	
≥80	0.59 (0.53-0.65)	
SIMD (NHS GGC)		
1 [most]	0.68 (0.62-0.74)	
2	0.94 (0.86-1.02)	
3	1.03 (0.95-1.12)	<0.0001
4	1.07 (0.98-1.17)	
5 [least] (ref.)	1.00	
Diabetes Diagnosis		
No (ref.)	1.00	<0.0001
Yes	0.82 (0.77-0.88)	
Further MI Last Year		
No (ref.)	1.00	0.0718
Yes	1.10 (0.99-1.22)	
Previous MI		
No (ref.)	1.00	0.0254
Yes	0.88 (0.79-0.99)	
Year of MI		
2009 (ref.)	1.00	
2010	1.12 (1.02-1.22)	
2011	1.28 (1.18-1.40)	<0.0001
2012	1.29 (1.18-1.41)	
2013	1.30 (1.19-1.43)	
2014	1.72 (1.55-1.91)	
Testing Frequency		
0 Tests	0.91 (0.83-1.00)	
1 Test (ref.)	1.00	0.0026
>1 Test	1.09 (1.00-1.18)	
AIC	102802.8	

Appendix V Optimal Latent Class Analysis for Target LDL

Appendix V1 Model for Target LDL Model (Using 50% Adherence)

Conditional Class Probabilities

	Class 1 (More)	**Class 2** (Less)
Sex		
Male	0.6836	0.6407
Female	0.3164	0.3593
Age at MI (years)		
<50	0.1335	0.1494
50-60	0.2501	0.2680
60-70	0.2707	0.2885
70-80	0.2452	0.2161
≥80	0.1004	0.0780
SIMD (NHS GGC)		
1 (Most)	0.2646	0.1760
2	0.2264	0.2222
3	0.1784	0.2225
4	0.1506	0.1886
5 (Least)	0.1800	0.1907
Diabetes		
No	0.6902	0.7717
Yes	0.3098	0.2283
Further MI Last Year		
No	0.9159	0.9282
Yes	0.0841	0.0718
Previous MI		
No	0.9234	0.9460
Yes	0.0766	0.0540
Year of MI		
2009	0.2914	0.1919
2010	0.2594	0.2061
2011	0.1934	0.2054
2012	0.1362	0.1840
2013	0.0814	0.1435
2014	0.0383	0.0691
50% Adherent Last Year		
No	0.0569	0.1282
Yes	0.9431	0.8718
LDL on Target Last Year		
No	0.0741	0.7113
Yes	0.9259	0.2887

Class Shares and Model Fit

	Class 1 (More)	**Class 2** (Less)
Est. Class Pop. Shares	0.4353	0.5647
Pred. Class Memberships	0.4323	0.5677
Maximum Log-Likelihood		-120047.9
BIC		240483.6
AIC		240175.9
# Iterations to Converge		144 (of 7000)
Tolerance		$1e^{-7}$

Regression Coefficients

Regression	**Coefficient**	**SE(Coef)**	**Exp(Coef)**	**p-value**
1 vs 2				
Intercept	-4.48151	1.26060	0.01132	<0.001
Meeting LDL Target	6.06149	1.32112	429.01419	<0.001

Model Comparisons

# Class	**Log-Likelihood**	**# Paras**	**BIC**	**Relative Entropy**
2	-120047.9	40	240483.6	0.7486
3	-119724.2	61	240039.8	0.6665
4	-119548.5	82	239891.9	0.6419

Appendix V2 Model for Target LDL Model (Using 80% Adherence)

Conditional Class Probabilities

	Class 1 (More)	**Class 2** (Less)
Age at MI (years)		
<50	0.1328	0.1497
50-60	0.2494	0.2684
60-70	0.2712	0.2880
70-80	0.2458	0.2160
≥80	0.1009	0.0778
Sex		
Male	0.6831	0.6415
Female	0.3169	0.3585
SIMD (NHS GGC)		
1 (Most)	0.2637	0.1775
2	0.2266	0.2220
3	0.1786	0.2220
4	0.1513	0.1878
5 (Least)	0.1798	0.1908
Diabetes		
No	0.6902	0.7710
Yes	0.3098	0.2290
Further MI Last Year		
No	0.9162	0.9279
Yes	0.0838	0.0721
Previous MI		
No	0.9236	0.9456
Yes	0.0764	0.0544
Year of MI		
2009	0.2917	0.1926
2010	0.2597	0.2064
2011	0.1936	0.2052
2012	0.1358	0.1838
2013	0.0809	0.1432
2014	0.0384	0.0688
80% Adherent Last Year		
No	0.2001	0.2858
Yes	0.7999	0.7142
LDL on Target Last Year		
No	0.0759	0.7041
Yes	0.9241	0.2959

Class Shares and Model Fit

	Class 1 (More)	Class 2 (Less)
Est. Class Pop. Shares	0.4301	0.5699
Pred. Class Memberships	0.4318	0.5682
Maximum Log-Likelihood		-124008.2
BIC		248404.1
AIC		248096.3
# Iterations to Converge		185 (of 7000)
Tolerance		$1e^{-7}$

Regression Coefficients

Regression	Coefficient	SE(Coef)	Exp(Coef)	p-value
1 vs 2				
Intercept	-15.36150	0.05436	<0.00001	<0.001
Meeting LDL Target	16.94450	0.05436	22850875.74	<0.001

Model Comparisons

# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
2	-124008.2	40	248404.1	0.7796
3	-123775.5	61	248142.3	0.7495
4	-123511.6	82	247818.1	0.6486

Predicted Class Assignments by Achievement of LDL Target

	Meet	Not Meet	Total
Class 1 (More)	7002	0	7002
Class 2 (Less)	1404	7809	9213
Total	8406	7809	16215

Diagnostic Accuracy Measures

	1 vs 2	
	LCA	**Bootstrap**
Sensitivity	83.3%	84.2% (83.5%, 85.0%)
Specificity	100.0%	100.0% (100.0%, 100.0%)
Positive Predictive Value	100.0%	100.0% (100.0%, 100.0%)
Negative Predictive Value	84.8%	85.3% (84.5%, 86.1%)

Average Posterior Probabilities

Class	Ave PP
1	0.9315
2	0.9510

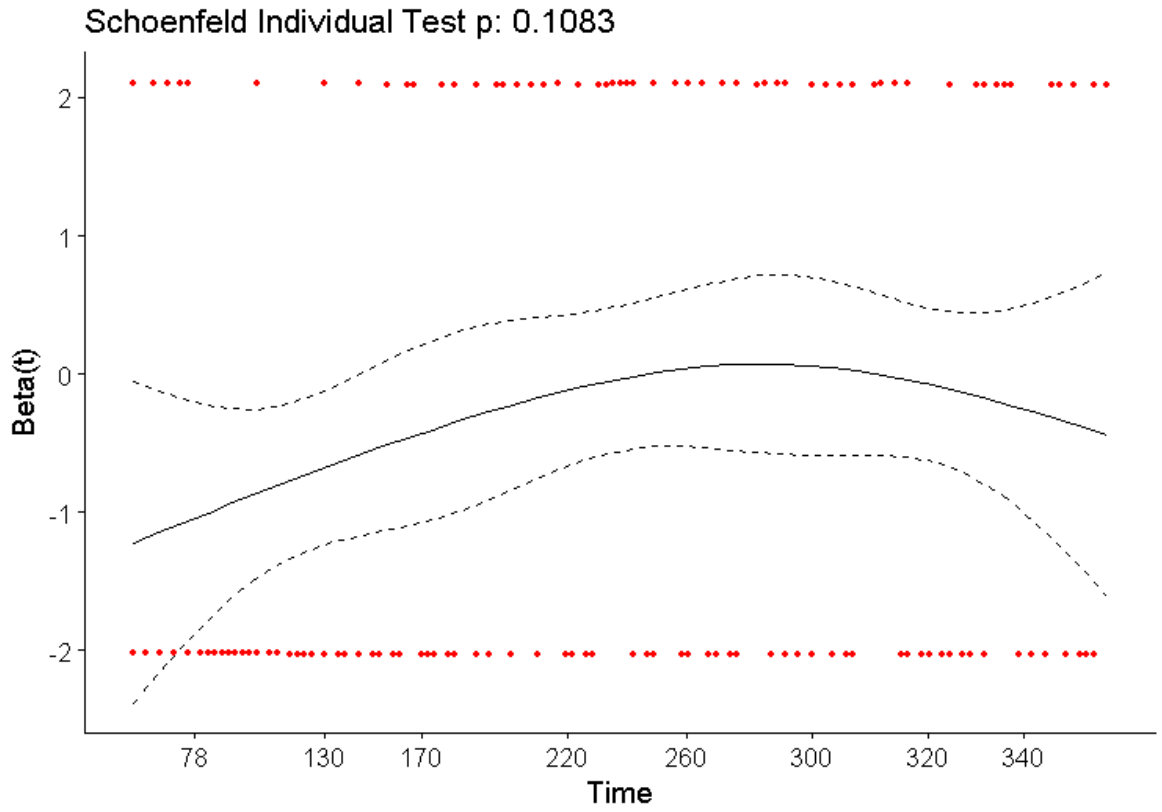
Appendix W Models for Mortality by Predicted Class from Latent Class Analysis (using 50% Adherence)

Appendix W1 Models for Year 2

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 [More] (ref.)	1.00	0.0527	
2 [Less]	0.72 (0.52, 1.01)		
Global PH Test			0.1083

Schoenfeld Residual Plot

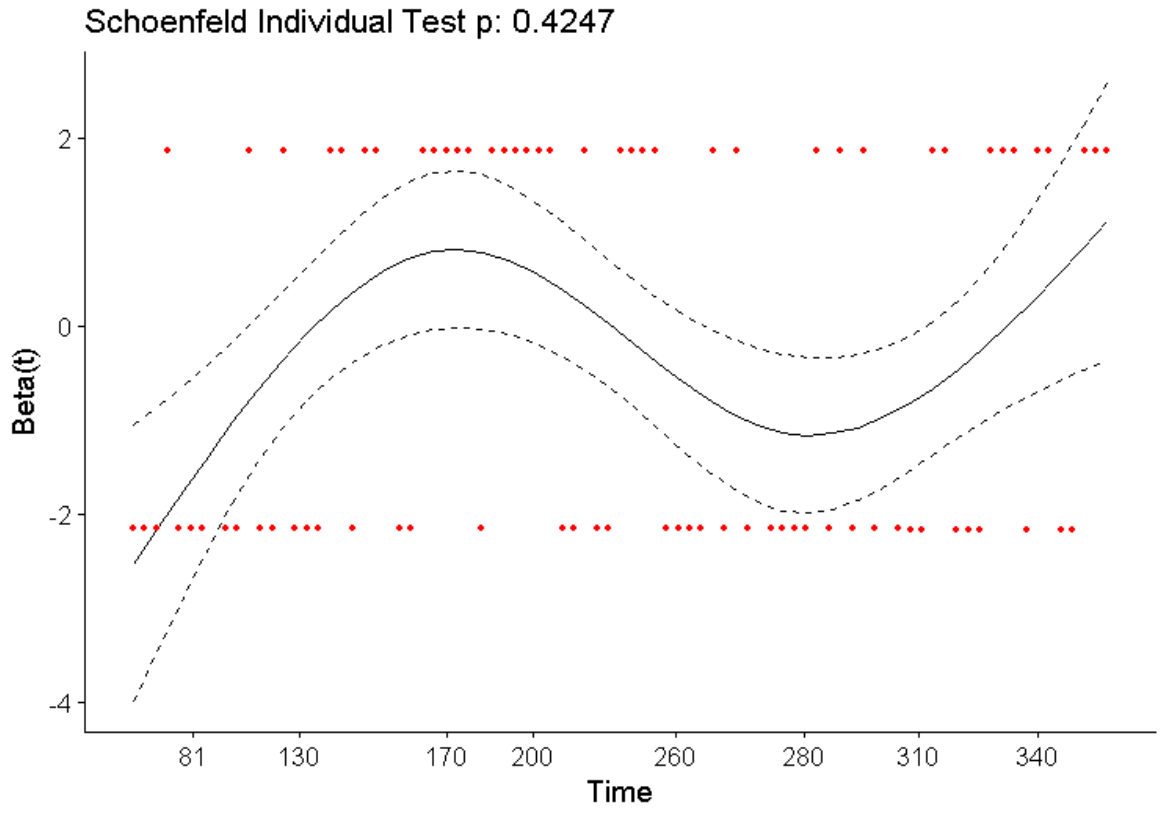


Appendix W2 Models for Year 3

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 [More] (ref.)	1.00	0.1500	
2 [Less]	0.73 (0.48,1.12)		
Global PH Test			0.4247

Schoenfeld Residual Plot



Appendix X Models for Mortality by Predicted Class from Latent Class Analysis (using 80% Adherence)

Appendix X1 Models for Year 2

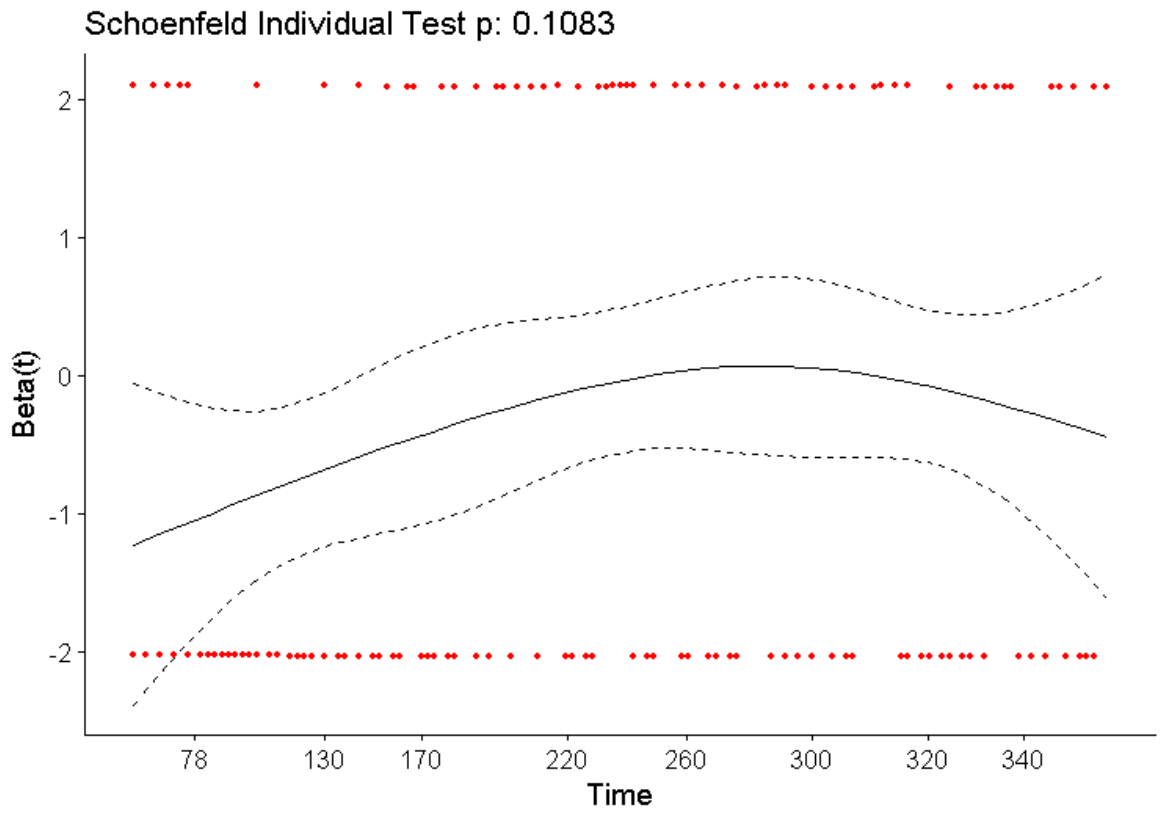
Cross-tabulation

	Class 1 (Most)	Class 2 (Least)
N	2,408	2,336
Events (%)	84 (3.5%)	59 (2.5%)

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 [More] (ref.)	1.00	0.0502	
2 [Less]	0.72 (0.52, 1.00)		
Global PH Test			0.1083

Schoenfeld Residual Plot



Appendix X2 Models for Year 3

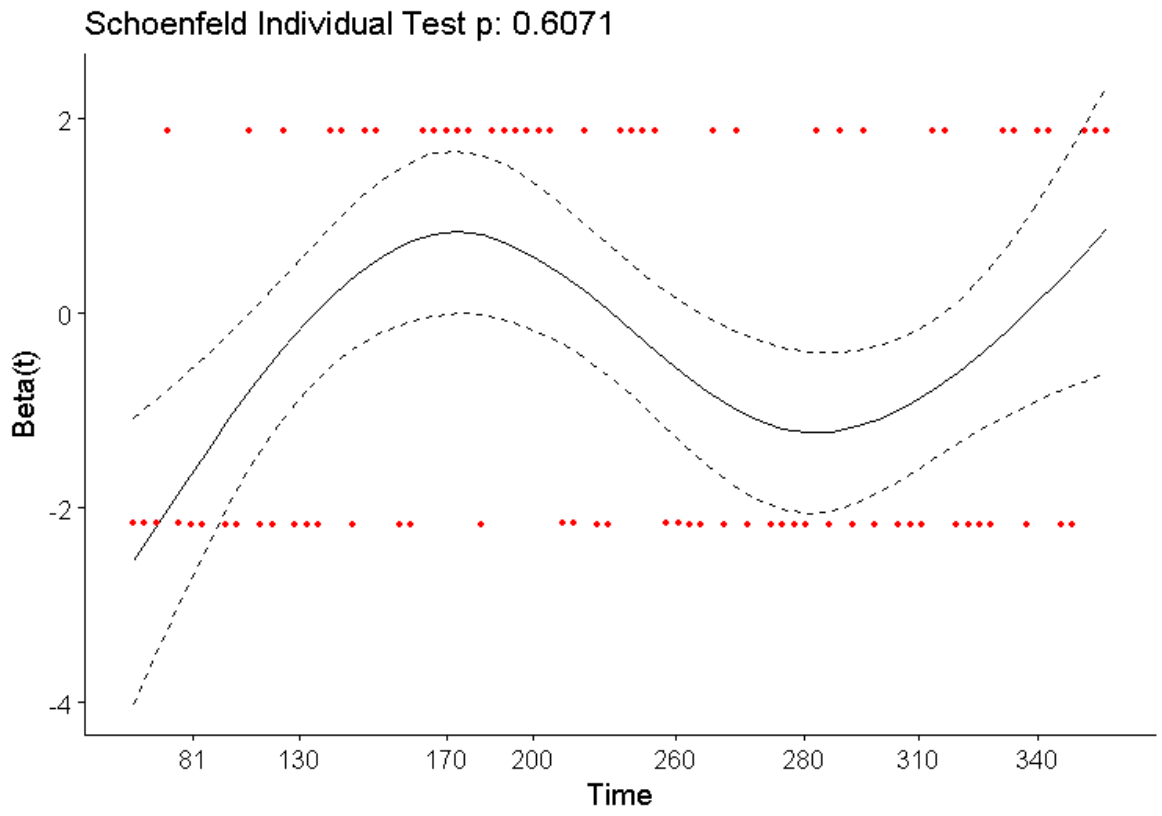
Cross-tabulation

	Class 1 (Most)	Class 2 (Least)
N	1,681	1,941
Events (%)	47 (2.8%)	38 (2.0%)

Cox Regression

Variable	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 [More] (ref.)	1.00	0.0964	
2 [Less]	0.70 (0.45, 1.07)		
Global PH Test			0.6071

Schoenfeld Residual Plot



Appendix Y Optimal Latent Class Growth Analysis for Time to Non-Target LDL

Class Proportions

	Class 1	Class 2	Class 3	Class 4
	1,530 (16.9%)	3,380 (37.2%)	1,250 (13.8%)	2,919 (32.2%)
Sex				
Male	1,050 (68.6%)	2,150 (63.6%)	842 (67.4%)	1,688 (57.8%)
Female	480 (31.4%)	1,230 (36.4%)	408 (32.6%)	1,231 (42.2%)
Age at MI (years)				
<50	138 (9.0%)	514 (15.2%)	97 (7.8%)	523 (17.9%)
50-60	262 (17.1%)	818 (24.2%)	218 (17.4%)	823 (28.2%)
60-70	338 (22.1%)	775 (22.9%)	332 (26.6%)	773 (26.5%)
70-80	444 (29.0%)	808 (23.9%)	333 (26.6%)	559 (19.2%)
≥80	348 (22.8%)	465 (13.8%)	270 (21.6%)	241 (8.3%)
SIMD (NHS GGC)				
1 (Most)	448 (29.3%)	971 (28.7%)	166 (13.3%)	576 (19.7%)
2	363 (23.7%)	764 (22.6%)	325 (26.0%)	655 (22.4%)
3	276 (18.0%)	647 (19.1%)	321 (25.7%)	649 (22.2%)
4	216 (14.1%)	496 (14.7%)	263 (21.0%)	552 (18.9%)
5 (Least)	227 (14.8%)	502 (14.9%)	175 (14.0%)	487 (16.7%)
Year of MI				
2009	607 (39.7%)	351 (10.4%)	385 (30.8%)	111 (3.8%)
2010	555 (36.3%)	531 (15.7%)	401 (32.1%)	196 (6.7%)
2011	249 (16.3%)	757 (22.4%)	330 (26.4%)	403 (13.8%)
2012	107 (7.0%)	776 (23.0%)	115 (9.2%)	690 (23.6%)
2013	12 (0.8%)	695 (20.6%)	19 (1.5%)	894 (30.6%)
2014	0 (0.0%)	270 (8.0%)	0 (0.0%)	625 (21.4%)
Previous MI				
Yes	166 (10.8%)	204 (6.0%)	77 (6.2%)	117 (4.0%)
No	1,364 (89.2%)	3,176 (94.0%)	1,173 (93.8%)	2,802 (96.0%)

Model Comparisons

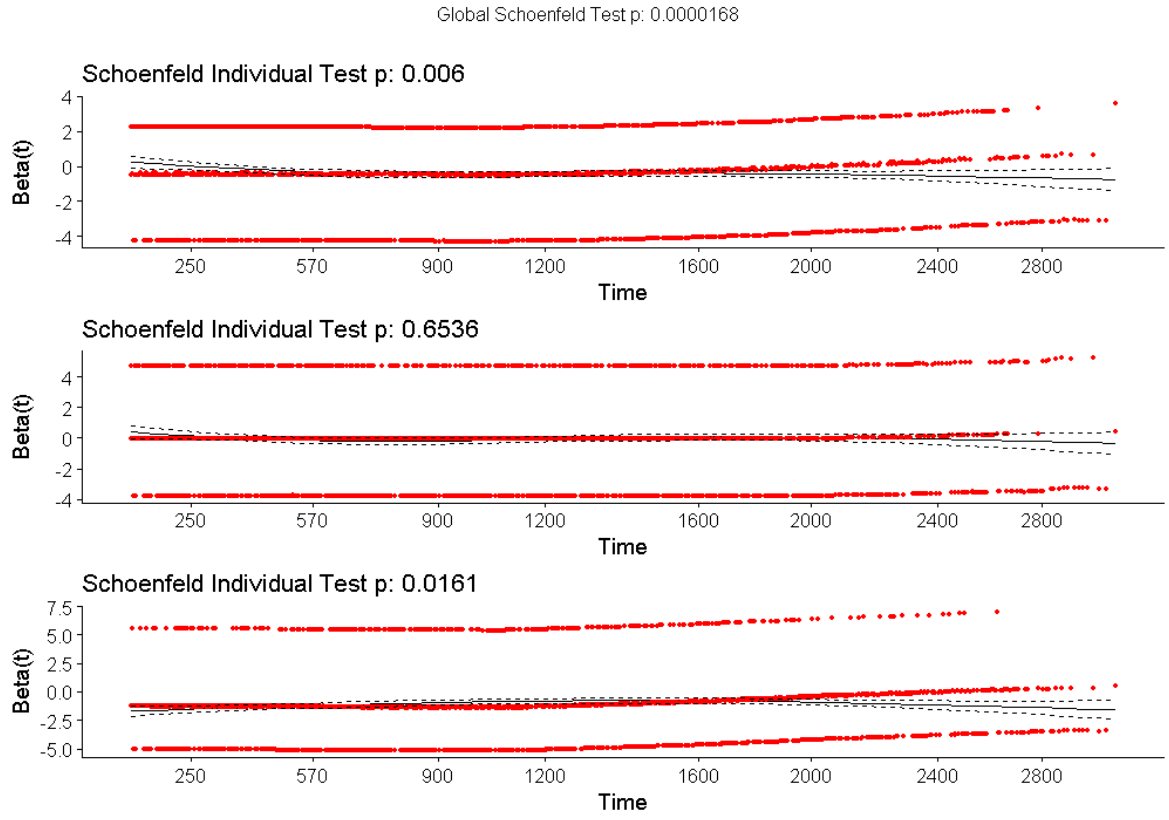
# Class	Log-Likelihood	# Paras	BIC	Relative Entropy
1	-20069.76	2	40157.75	1.0000
2	-17938.69	10	35968.52	0.5851
3	-17402.87	18	34969.80	0.6124
4	-17133.81	26	34504.58	0.5257

Appendix Z Cox Model for Mortality by Predicted Class from Latent Class Growth Analysis (Target LDL)

Model

	Hazard Ratio (95% CI)	p-value	PH Test
Class			
1 (ref.)	1.00		
2	0.71 (0.64, 0.79)	<0.0001	0.0060
3	0.98 (0.87, 1.10)		0.6536
4	0.35 (0.30, 0.40)		0.0161
Global PH Test			<0.0001

Schoenfeld Residual Plot



Appendix AA Posters

Appendix AA1 Association of Physicians of Great Britain & Ireland Annual Meeting

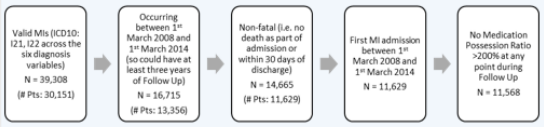
Glasgow's Post Myocardial Infarction Population: Demographics, Statin Adherence and Cholesterol Targets

Rosemary Brown[1], Sarah Wild[2], Jim Lewsey[3], Paul Welsh[1], Jennifer Logue[1]
[1]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [2]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK [3]Institute of Health & Wellbeing, University of Glasgow, Glasgow, UK

Background

- In Scotland, patients >40 years are classified as being at high or low risk of a CVD event in the next ten years.
- Patients with established CVD, are automatically considered high risk.
- The 2007 & 2017 SIGN guidelines state that
  - statins should be offered regardless of the patient's plasma lipid concentration
  - annual monitoring of lipid concentration represents good practice.
- Annual review should also encourage positive lifestyle changes, adherence to medication, and identify and discuss other CVD risk factors.
- This analysis examined the demographics of post MI patients in Glasgow, their estimated adherence to statin medication and whether the targets for cholesterol were achieved.

Methods



- Data was extracted from the NHS Greater Glasgow and Clyde Safe Haven
- All patients with an SMR01 record for a non-fatal myocardial infarction (MI) (ICD10: I21, I22) occurring between 1st March 2008 and 1st March 2014.
- Patient's demographics, lipid plasma concentration results, and statin prescription records up until 1st March 2017 (or death) were also extracted.

- Statin adherence was estimated for each year post MI by calculating the medication possession ratio (MPR) from the dispensing records, and then averaged across the years for each patient.
- Lipid plasma concentrations were averaged for each year post MI. Patients were considered to have met a cholesterol target if the target was met for any of the years of their follow up.
- The cholesterol targets of interest were:
  - 40% reduction in non-HDL cholesterol (SIGN Guidelines)[1]
  - LDL cholesterol <1.8 mmol/l (ESC Guidelines)[2]

Results: Demographics

Table 1: Baseline Demographics

Characteristic	Post MI Population (N=11,568)
Age at MI (years)	66.9 (13.9)
Gender	
Male	7,002 (60.5%)
Female	4,566 (39.5%)
SIMD 2012 Quintile	
1 (Most)	4,963 (43.2%)
2	2,106 (18.3%)
3	1,449 (12.6%)
4	1,357 (11.8%)
5 (Least)	1,614 (14.0%)
Missing	79
Diabetes at MI	
Type 1	124 (1.1%)
Type 2	2,033 (17.6%)
All	2,157 (18.6%)
Duration of Diabetes at MI (years)	11.5 (9.6)
Number of MIs (before end of FU)	
1	8,740 (75.6%)
≥1	2,828 (24.4%)
Deaths	4,053 (35.0%)
MPR Calculable	10,469 (90.5%)
% Non-HDL Reduction Calculable	3,883 (33.6%)
≥1 Post MI LDL Result Available	9,939 (85.9%)

Values are raw Ns and percentages for categorical variables and means and standard deviations for continuous variables.

Results: Estimated Adherence and Cholesterol Target Reduction

Table 2: Associations between Estimated Adherence with Cholesterol Targets and Mortality (50% Threshold)

	MPR <50%	MPR ≥50%	Total
Reduction in non-HDL			
<40%	494 (75.9%)	1,332 (43.6%)	1,826
≥40%	157 (24.1%)	1,723 (56.4%)	1,880
Total	651	3,055	3,706
			X^2, p<0.001
LDL Cholesterol			
> 1.8 mmol/l	444 (32.8%)	1,635 (20.1%)	2,079
≤1.8 mmol/l	909 (67.2%)	6,485 (79.9%)	7,394
Total	1,353	8,120	9,473
			X^2, p<0.001
Mortality			
Dead	638 (38.8%)	2,316 (26.2%)	2,954
Alive	1,008 (61.2%)	6,507 (73.8%)	7,515
Total	1,646	8,823	10,469
			X^2, p<0.001

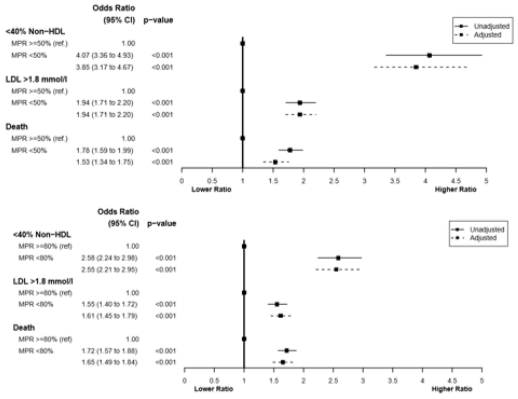
Percentages are calculated within columns

Table 3: Associations between Estimated Adherence with Cholesterol Targets and Mortality (80% Threshold)

	MPR <80%	MPR ≥80%	Total
Reduction in non-HDL			
<40%	808 (64.6%)	1,018 (41.4%)	1,826
≥40%	442 (35.4%)	1,438 (58.6%)	1,880
Total	1,250	2,456	3,706
			X^2, p<0.001
LDL Cholesterol			
> 1.8 mmol/l	789 (27.4%)	1,290 (19.6%)	2,079
≤1.8 mmol/l	2,089 (72.6%)	5,305 (80.4%)	7,394
Total	2,878	6,595	9,473
			X^2, p<0.001
Mortality			
Dead	1,206 (35.9%)	1,748 (24.6%)	2,954
Alive	2,154 (64.1%)	5,361 (75.4%)	7,515
Total	3,360	7,109	10,469
			X^2, p<0.001

Percentages are calculates within columns

Figure 1: Odds Ratios of Associations between Estimated Adherence with Cholesterol Targets and Mortality



Adjusted analyses were adjusted for age, sex, year of MI and SIMD 2012 deprivation quintile

Conclusions

- In the post MI population of NHS GG&C, patients with lower statin adherence
  - are less likely to achieve cholesterol targets.
  - had higher mortality.
- However, a large proportion of patients achieved targets and had excellent adherence, questioning the need for repeated lipid plasma measurements.
- Future analysis on this cohort will seek to identify whether patients with high adherence and well controlled cholesterol levels could receive less frequent lipid plasma concentration monitoring.

References

1. SIGN. SIGN 149: Risk estimation and the prevention of cardiovascular disease. 2017;(June). http://www.sign.ac.uk/assets/sign149.pdf.
2. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2016;37(29):2315-2381.

University of Glasgow · The University of Edinburgh · MRC Medical Research Council

Appendix AA2 Farr Institute Symposium and Innovation Workshop

Personalised Blood Testing Schedules in Chronic Disease Management

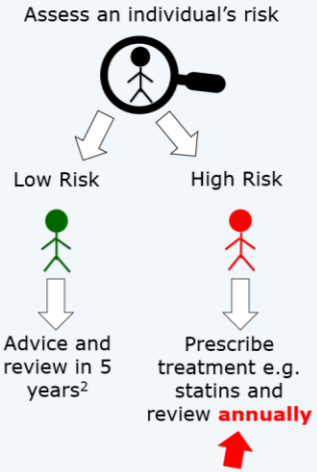
Rosemary Brown, University of Glasgow, r.brown.4@research.gla.ac.uk
Supervisors: Dr Jennifer Logue[1], Dr Paul Welsh[1], Dr Jim Lewsey[1], Prof Sarah Wild[2]
(1) University of Glasgow, (2) University of Edinburgh

Cardiovascular Disease (CVD)

Refers to conditions affecting the circulation of blood, with the most common being Coronary Heart Disease and Stroke
~670,000 people in Scotland aged 16+ diagnosed.[1]

Fig 1. Current Monitoring

SIGN (2017) guidelines for risk estimation and prevention of CVD[2] recommend:



This frequency is based solely on the guideline creators' experience.

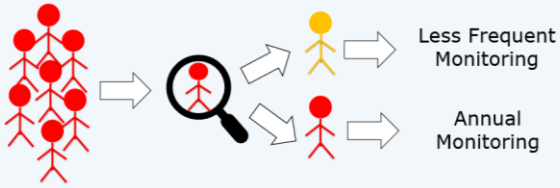
Current Problem

At a population level, this presents a problem. Blood tests for chronic diseases, including CVD, currently account for more than 50% of all blood tests in the UK.[3] The number of blood tests requested is increasing by 10% each year.[4]

Previous Research

Looker et al (2013) used data from the Diabetic Retinopathy Screening programme to identify that individuals with two prior clear scans could be considered for screening every two years, instead of the currently recommended annual review.[5]

Fig 2. Potential Solution



Develop algorithms to personalise the blood testing frequency for those at high risk, which:

Reduces unnecessary tests
AND
Does not increase cardiovascular events

Project Aims and Progress to Date

Use data from the NHS Greater Glasgow and Clyde Safe Haven including blood test results, hospital admissions, other comorbidities, and medication adherence to build an algorithm to personalise the frequency of blood testing. The potential cost savings in its implementation would then be modelled.

To date, work has focussed on data cleaning, with a view to first determining if the patterns of testing reflect the recommendations in the guidelines in the near future.

References
1. Brown, L. *et al. The Scottish Health Survey 2015*. **1**, (2015).; 2. SIGN. SIGN 149: Risk estimation and the prevention of cardiovascular disease. (2017).; 3. Department of Health. *Report of the review of NHS pathology services in England*. (2010).; 4. Smellie, W. S. A. Demand management and test request rationalization. *Ann. Clin. Biochem.* **49**, 323–336 (2012).; 5. Looker, H. C. *et al*. Predicted impact of extending the screening interval for diabetic retinopathy: The Scottish Diabetic Retinopathy Screening programme. *Diabetologia* 56, 1716–1725 (2013).

University of Glasgow | MRC Medical Research Council

References

(Australian) Stroke Foundation (2017) *Clinical Guidelines for Stroke Management 2017*. Available at: www.informme.org.au.

AGREE Next Steps Consortium (2010) 'AGREE II: Advancing guideline development, reporting and evaluation in healthcare', *CMAJ*, 182, pp. 839–842. Available at: https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf.

Agresti, A. (2012) 'Other Mixture Models for Discrete Data', in Agresti, A. (ed.) *Categorical Data Analysis*. 3rd edn. New Jersey: John Wiley & Sons. Available at: https://ebookcentral.proquest.com/lib/gla/reader.action?docID=1168529# (Accessed: 8 December 2020).

Akioyamen, L. E. *et al.* (2017) 'Estimating the prevalence of heterozygous familial hypercholesterolaemia: A systematic review and meta-analysis', *BMJ Open*, 7(9). doi: 10.1136/bmjopen-2017-016461.

Albers, G. W. *et al.* (2010) *Textbook of Stroke Medicine*. 2nd edn, *Textbook of stroke medicine*. 2nd edn. Edited by M. Brainin and W.-D. Heiss. Cambridge: Cambridge University Press.

Allahyari, A. *et al.* (2020) 'Application of the 2019 ESC/EAS dyslipidaemia guidelines to nationwide data of patients with a recent myocardial infarction: a simulation study', *European Heart Journal*, ehaa034. doi: 10.1093/eurheartj/ehaa034.

Anderson, T. J. *et al.* (2016) '2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult.', *The Canadian journal of cardiology*. England, 32(11), pp. 1263–1282. doi: https://dx.doi.org/10.1016/j.cjca.2016.07.510.

Appelman, Y. *et al.* (2014) 'Sex differences in cardiovascular risk factors and disease prevention', *Atherosclerosis*, 241(1), pp. 211–218. doi: 10.1016/j.atherosclerosis.2015.01.027.

Arnett, D. K. *et al.* (2019) '2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines', *Journal of the American College of Cardiology*. Elsevier USA, 74(10), pp. e177–e232. doi: 10.1016/j.jacc.2019.03.010.

Aspelund, T. *et al.* (2011) 'Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy', *Diabetologia*. Springer, 54, pp. 2525–2532. doi: 10.1007/s00125-011-2257-7.

ASSIGN (2014) *ASSIGN Score: Estimate the Risk*. Available at: http://assign-score.com/estimate-the-risk/ (Accessed: 17 November 2017).

Austin, P. C., Lee, D. S. and Fine, J. P. (2016) 'Introduction to the Analysis of Survival Data in the Presence of Competing Risks', *Circulation*. Lippincott Williams and Wilkins, 133(6), pp. 601–609. doi: 10.1161/CIRCULATIONAHA.115.017719.

Aviña-Zubieta, J. A. *et al.* (2008) 'Risk of cardiovascular mortality in patients with rheumatoid arthritis: A meta-analysis of observational studies', *Arthritis Care and Research*, 59(12), pp. 1690–1697. doi: 10.1002/art.24092.

Aviña-Zubieta, J. A. *et al.* (2012) 'Risk of incident cardiovascular events in patients with rheumatoid arthritis: A meta-analysis of observational studies', *Annals of the Rheumatic Diseases*, 71(9), pp. 1524–1529. doi: 10.1136/annrheumdis-2011-200726.

Bains, I. *et al.* (2019) 'Clinical impact and cost-effectiveness of primary cytology versus human papillomavirus testing for cervical cancer screening in England', *International Journal of Gynecological Cancer*. BMJ Publishing Group, 29(4), pp. 669–675. doi: 10.1136/ijgc-2018-000161.

Baltzer, N. *et al.* (2017) 'Risk stratification in cervical cancer screening by complete screening history: Applying bioinformatics to a general screening population', *International Journal of Cancer*. Wiley-Liss Inc., 141(1), pp. 200–209. doi: 10.1002/ijc.30725.

Barker, I. *et al.* (2018) 'Self-management capability in patients with long-term conditions is associated with reduced healthcare utilisation across a whole health economy: Cross-sectional analysis of electronic health records', *BMJ Quality and Safety*. BMJ Publishing Group, 27(12), pp. 989–999. doi: 10.1136/bmjqs-2017-007635.

Barter, P. J. *et al.* (2007) 'Effects of torcetrapib in patients at high risk for coronary events', *New England Journal of Medicine*, 357(21), pp. 2109–2122. doi: 10.1056/NEJMoa0706628.

Bartlett, J. *et al.* (2016) 'Is Isolated Low High-Density Lipoprotein Cholesterol a Cardiovascular Disease Risk Factor? New Insights From the Framingham Offspring Study.', *Circulation. Cardiovascular quality and outcomes*, 9(3), pp. 206–212. doi: 10.1161/CIRCOUTCOMES.115.002436.

Batelaan, N. M. *et al.* (2016) 'Anxiety and new onset of cardiovascular disease: Critical review and meta-analysis', *British Journal of Psychiatry*, 208(3), pp. 223–231. doi: 10.1192/bjp.bp.114.156554.

Benner, J. S. *et al.* (2004) 'Follow-up lipid tests and physician visits are associated with improved adherence to statin therapy', *PharmacoEconomics*. Springer, 22, pp. 13–23. doi: 10.2165/00019053-200422003-00003.

Bentzon, J. F. *et al.* (2014) 'Mechanisms of plaque formation and rupture', *Circulation Research*, 114(12), pp. 1852–1866. doi: 10.1161/CIRCRESAHA.114.302721.

Berlin, K. S., Parra, G. R. and Williams, N. A. (2014) 'An Introduction to Latent Variable Mixture Modeling (Part 2): Longitudinal Latent Class Growth Analysis and Growth Mixture Models', *Journal of Pediatric Psychology*. Oxford Academic, 39(2), pp. 188–203. doi: 10.1093/jpepsy/jst085.

Boekholdt, S. M. *et al.* (2014) 'Very low levels of atherogenic lipoproteins and the risk for cardiovascular events: A meta-analysis of statin trials', *Journal of the American College of Cardiology*. Elsevier USA, 64(5), pp. 485–494. doi: 10.1016/j.jacc.2014.02.615.

Bolck, A., Croon, M. and Hagenaars, J. (2004) ‘Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators’, *Political Analysis*, 12(1), pp. 3–27. doi: 10.1093/pan/mph001.

Bowman, L. *et al.* (2017) ‘Effects of anacetrapib in patients with atherosclerotic vascular disease’, *New England Journal of Medicine*, 377(13), pp. 1217–1227. doi: 10.1056/NEJMoa1706444.

British Heart Foundation (2018) *UK Factsheet*. Available at: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics.

British Heart Foundation (2020a) *BHF CVD Statistics Factsheet - Scotland*. Available at: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics/heart-statistics-publications/cardiovascular-disease-statistics-2018.

British Heart Foundation (2020b) ‘BHF CVD Statistics Factsheet - UK’, *British Heart Foundation*, pp. 1–21.

Broadbent, D. M. *et al.* (2019) ‘Individualised screening for diabetic retinopathy: The ISDR study - Rationale, design and methodology for a randomised controlled trial comparing annual and individualised risk-based variable-interval screening’, *BMJ Open*. BMJ Publishing Group, 9(6), p. e025788. doi: 10.1136/bmjopen-2018-025788.

Broadbent, D. M. *et al.* (2020) ‘Safety and cost-effectiveness of individualised screening for diabetic retinopathy: the ISDR open-label, equivalence RCT’, *Diabetologia*. Springer Science and Business Media Deutschland GmbH, 64(1), pp. 56–69. doi: 10.1007/s00125-020-05313-2.

Brown, L. *et al.* (2015) *The Scottish Health Survey 2015*.

Brown, M. T. *et al.* (2016) ‘Medication Adherence: Truth and Consequences’, *American Journal of the Medical Sciences*. Elsevier B.V., 351(4), pp. 387–399. doi: 10.1016/j.amjms.2016.01.010.

Brown, R. E., Welsh, P. and Logue, J. (2020) ‘Systematic review of clinical

guidelines for lipid lowering in the secondary prevention of cardiovascular disease events', *Open Heart*, 7(2), p. e001396. doi: 10.1136/openhrt-2020-001396.

Brown, R., Welsh, P. and Logue, J. (2018) *Systematic review of clinical guidelines regarding the risk factor assessment for the secondary prevention of cardiovascular disease events.*, *PROSPERO CRD42018098582*. Available at: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018098582.

Brunström, M. and Carlberg, B. (2018) 'Association of blood pressure lowering with mortality and cardiovascular disease across blood pressure levels: A systematic review and meta-analysis', *JAMA Internal Medicine*, 178(1), pp. 28–36. doi: 10.1001/jamainternmed.2017.6015.

Bryer, A. *et al.* (2010) 'South African guideline for management of ischaemic stroke and transient ischaemic attack 2010: a guideline from the South African Stroke Society (SASS) and the SASS Writing Committee.', *South African medical journal*. South Africa, 100(11), pp. 747–778. Available at: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med6&NEWS=N&AN=21081029.

Bull, K. and Spiegelhalter, D. J. (1997) 'Tutorial in Biostatistics Survival Analysis in Observational Studies', *Statistics in Medicine*. John Wiley & Sons, 16, pp. 1041–1074. doi: doi.org/10.1002/(SICI)1097-0258(19970515)16:9<1041::AID-SIM506>3.0.CO;2-F.

Bundy, J. D. *et al.* (2017) 'Systolic blood pressure reduction and risk of cardiovascular disease and mortality: A systematic review and network meta-analysis', *JAMA Cardiology*, 2(7), pp. 775–781. doi: 10.1001/jamacardio.2017.1421.

Bunker, S. J. *et al.* (2003) '"Stress" and coronary heart disease: psychosocial risk factors', *Medical Journal of Australia*, 178(6), pp. 272–276. doi: 10.5694/j.1326-5377.2003.tb05193.x.

Burke, G. M. *et al.* (2017) 'Temporal Associations Between Smoking and

Cardiovascular Disease, 1971 to 2006 (from the Framingham Heart Study)', *American Journal of Cardiology*, 120(10), pp. 1787–1791. doi: 10.1016/j.amjcard.2017.07.087.

Cannon, C. P. *et al.* (2006) 'Meta-Analysis of Cardiovascular Outcomes Trials Comparing Intensive Versus Moderate Statin Therapy', *Journal of the American College of Cardiology*, 48(3), pp. 438–445. doi: 10.1016/j.jacc.2006.04.070.

Carney, R. M. *et al.* (2004) 'Depression and Late Mortality After Myocardial Infarction in the Enhancing Recovery in Coronary Heart Disease (ENRICHD) Study', *Psychosomatic Medicine*, 66(4), pp. 466–474. doi: 10.1097/01.psy.0000133362.75075.a6.

Catalogue of Bias Collaboration *et al.* (2018) *Confounding*, *Catalogue of Biases*. Available at: https://catalogofbias.org/biases/confounding/ (Accessed: 15 July 2020).

Catapano, A. L. *et al.* (2016) '2016 ESC/EAS Guidelines for the Management of Dyslipidaemias', *European Heart Journal*, 37(39), pp. 2999-3058l. doi: 10.1093/eurheartj/ehw272.

Chalk, D. *et al.* (2012) 'Can the retinal screening interval be safely increased to 2 years for type 2 diabetic patients without retinopathy?', *Diabetes Care*. American Diabetes Association, 35(8), pp. 1663–1668. doi: 10.2337/dc11-2282.

Challener, D. W., Prokop, L. J. and Abu-Saleh, O. (2019) 'The Proliferation of Reports on Clinical Scoring Systems: Issues about Uptake and Clinical Utility', *JAMA*, 321(24), pp. 2405–2406. doi: 10.1001/jama.2019.5284.

Chami, N. *et al.* (2017) 'Rates of inappropriate laboratory test utilization in Ontario', *Clinical Biochemistry*. Elsevier Inc., 50(15), pp. 822–827. doi: 10.1016/j.clinbiochem.2017.05.004.

Chen, S. T. *et al.* (2019) 'Long-term statin adherence in patients after hospital discharge for new onset of atherosclerotic cardiovascular disease: A population-based study of real world prescriptions in Taiwan', *BMC Cardiovascular*

Disorders, 19(62). doi: 10.1186/s12872-019-1032-4.

Cheung, B. M. *et al.* (2017) '2016 Consensus statement on prevention of atherosclerotic cardiovascular disease in the Hong Kong population.', *Hong Kong Medical Journal*. China, 23(2), pp. 191–201. doi: https://dx.doi.org/10.12809/hkmj165045.

Chew, D. P. *et al.* (2016) 'National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand: Australian clinical guidelines for the management of acute coronary syndromes 2016.', *Heart, Lung & Circulation*. Australia, 25, pp. 895–951. Available at: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med8&NEWS=N&AN=27465769.

Cholesterol Treatment Trialists' (CTT) Collaboration (2010) 'Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170 000 participants in 26 randomised trials', *The Lancet*. Elsevier Ltd, 376(9753), pp. 1670–1681. doi: 10.1016/S0140-6736(10)61350-5.

Cholesterol Treatment Trialists' (CTT) Collaborators (2012) '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*, 380(9841), pp. 581–590. doi: 10.1016/S0140-6736(12)60367-5.

Chowdhury, R. *et al.* (2013) 'Adherence to cardiovascular therapy: A meta-analysis of prevalence and clinical consequences', *European Heart Journal*, 34(38), pp. 2940–2948. doi: 10.1093/eurheartj/eht295.

Chronic Kidney Disease Prognosis Consortium (2010) 'Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis', *The Lancet*, 375(9731), pp. 2073–2081. doi: 10.1016/S0140-6736(10)60674-5.

Chronic Kidney Disease Prognosis Consortium (2011) 'Lower estimated glomerular filtration rate and higher albuminuria are associated with all-cause and cardiovascular mortality. A collaborative meta-analysis of high-risk

population cohorts', *Kidney International*, 79(12), pp. 1341–1352. doi: 10.1038/ki.2010.536.

Cohen, J. (1960) 'A Coefficient of Agreement for Nominal Scales', *Educational and Psychological Measurement*. Sage PublicationsSage CA: Thousand Oaks, CA, 20(1), pp. 37–46. doi: 10.1177/001316446002000104.

Collins, R. *et al.* (2016) 'Interpretation of the evidence for the efficacy and safety of statin therapy', *The Lancet*. Elsevier Ltd, 388(10059), pp. 2532–2561. doi: 10.1016/S0140-6736(16)31357-5.

Cooney, M. T. *et al.* (2009) 'HDL cholesterol protects against cardiovascular disease in both genders, at all ages and at all levels of risk', *Atherosclerosis*, 206(2), pp. 611–616. doi: 10.1016/j.atherosclerosis.2009.02.041.

Coscia Requena, C., Muriel, A. and Peñuelas, O. (2018) 'Analysis of causality from observational studies and its application in clinical research in Intensive Care Medicine', *Medicina Intensiva (English Edition)*. Elsevier BV, 42(5), pp. 292–300. doi: 10.1016/j.medine.2018.01.010.

Costanzo, S. *et al.* (2010) 'Alcohol Consumption and Mortality in Patients With Cardiovascular Disease', *Journal of the American College of Cardiology*, 55(13), pp. 1339 LP – 1347. doi: 10.1016/j.jacc.2010.01.006.

Cox, D. R. (1972) 'Regression Models and Life-Tables', *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), pp. 187–220. doi: 10.2307/2985181.

Danese, M. D. *et al.* (2017) 'Management of lipid-lowering therapy in patients with cardiovascular events in the UK: A retrospective cohort study', *BMJ Open*, 7(5), pp. 1–10. doi: 10.1136/bmjopen-2016-013851.

Department of Health (2010) *Report of the review of NHS pathology services in England*. Available at: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/

dh_091984.pdf.

Dijkstra, M. G. *et al.* (2016) 'Safety of extending screening intervals beyond five years in cervical screening programmes with testing for high risk human papillomavirus: 14 year follow-up of population based randomised cohort in the Netherlands', *BMJ (Clinical research ed.)*. British Medical Journal Publishing Group, 355, p. i4924. doi: 10.1136/bmj.i4924.

Doll, H. *et al.* (2011) 'The rise of cholesterol testing: How much is unnecessary?', *British Journal of General Practice*. British Journal of General Practice, 61(583), pp. e81–e88. doi: 10.3399/bjgp11X556245.

Donnan, P. T., MacDonald, T. M. and Morris, A. D. (2002) 'Adherence to prescribed oral hypoglycaemic medication in a population of patients with Type 2 diabetes: A retrospective cohort study', *Diabetic Medicine*. doi: 10.1046/j.1464-5491.2002.00689.x.

Donnelly, L. A. *et al.* (2008) 'Long-term adherence to statin treatment in diabetes', *Diabetic Medicine*. doi: 10.1111/j.1464-5491.2008.02476.x.

Eindhoven, D. C. *et al.* (2018) 'Age and gender differences in medical adherence after myocardial infarction: Women do not receive optimal treatment – The Netherlands claims database', *European Journal of Preventive Cardiology*. SAGE Publications Inc., 25(2), pp. 181–189. doi: 10.1177/2047487317744363.

Eleuteri, A. *et al.* (2017) 'Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine', *Diabetologia*. Springer Verlag, 60, pp. 2174–2182. doi: 10.1007/s00125-017-4386-0.

Estruch, R. *et al.* (2013) 'Primary prevention of cardiovascular disease with a Mediterranean diet', *New England Journal of Medicine*, 368(14), pp. 1279–1290. doi: 10.1056/NEJMoa1200303.

Estruch, R. *et al.* (2018) 'Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts', *New*

England Journal of Medicine, 378(25), p. e34. doi: 10.1056/NEJMoa1800389.

Evans, J., Horton-Szar, D. and Newby, D. E. (2015) *Crash Course Cardiovascular Disease*. Revised 4t. Edited by D. Horton-Szar, J. Bowes, and H. Leng. Elsevier Ltd.

Expert Dyslipidemia Panel of the International Atherosclerosis Society Panel Members (2014) 'An International Atherosclerosis Society Position Paper: global recommendations for the management of dyslipidemia', *Journal of Clinical Lipidology*. Edited by A. H. Grundy SM Barter P, Bersot TP, Betteridge DJ, Carmena R, Cuevas A, Davidson MH, Genest J, Kesaniemi YA, Sadikot S, Santos RD, Susekov AV, Sy RG, Laletokgozoglu S, Watts GF, Zhao D. United States, 8(1), pp. 29–60. doi: https://dx.doi.org/10.1016/j.jacl.2013.12.005.

Fisher, L. D. and Lin, D. Y. (1999) 'Time-Dependent Covariates in the Cox Proportional-Hazards Regression Model', *Annual Review of Public Health*, 20, pp. 145–157. doi: https://doi.org/10.1146/annurev.publhealth.20.1.145.

Ford, I. *et al.* (2007) 'Long-term follow-up of the West of Scotland coronary prevention study', *New England Journal of Medicine*, 357, pp. 1477–1486. doi: 10.1056/NEJMoa065994.

Ford, I. *et al.* (2016) 'Long-term safety and efficacy of lowering low-density lipoprotein cholesterol with statin therapy 20-year follow-up of west of Scotland coronary prevention study', *Circulation*, 133(11), pp. 1073–1080. doi: 10.1161/CIRCULATIONAHA.115.019014.

Gansevoort, R. T. *et al.* (2013) 'Chronic kidney disease and cardiovascular risk: Epidemiology, mechanisms, and prevention', *The Lancet*, 382(9889), pp. 339–352. doi: 10.1016/S0140-6736(13)60595-4.

Gibson, J. *et al.* (2015) *Eighth National GP Worklife Survey*. Manchester. Available at: http://www.population-health.manchester.ac.uk/healtheconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/EighthNationalGPWorklifeSurveyreport.pdf (Accessed: 23 March 2020).

Goldberg, D. (1991) 'What Every Computer Scientist Should Know About Floating-Point Arithmetic', *ACM Computing Surveys (CSUR)*, 23(1), pp. 5–48. doi: 10.1145/103162.103163.

Golden Jubilee National Hospital (2019) *Golden Jubilee National Hospital: About Us*. Available at: https://www.nhsgoldenjubilee.co.uk/about/ (Accessed: 3 February 2019).

Gordon, M. and Lumley, T. (2020) 'forestplot: Advanced Forest Plot Using "grid" Graphics in R'. Available at: https://cran.r-project.org/web/packages/forestplot/index.html (Accessed: 15 January 2021).

Grambsch, P. M. and Therneau, T. M. (1994) 'Proportional Hazards Tests and Diagnostics Based on Weighted Residuals', *Biometrika*. JSTOR, 81(3), pp. 515–526. doi: 10.2307/2337123.

Grundy, S. M. *et al.* (2019) '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', *Circulation*, 139(25), pp. E1082–E1143. doi: 10.1161/CIR.0000000000000625.

Gutierrez, J. *et al.* (2012) 'Statin therapy in the prevention of recurrent cardiovascular events a sex-based meta-analysis', *Archives of Internal Medicine*. American Medical Association, 172(12), pp. 909–919. doi: 10.1001/archinternmed.2012.2145.

Hajati, F. *et al.* (2018) 'Patterns and trends of potentially inappropriate high-density lipoprotein cholesterol testing in Australian adults at high risk of cardiovascular disease from 2008 to 2014: analysis of linked individual patient data from the Australian Medicare Benefits S', *BMJ Open*. England, 8, p. e019041. doi: https://dx.doi.org/10.1136/bmjopen-2017-019041.

Hansson, G. K. (2005) 'Mechanisms of disease: Inflammation, atherosclerosis, and coronary artery disease', *New England Journal of Medicine*, 352, pp. 1685–

1695. doi: 10.1056/NEJMra043430.

Harrell, F. E. (2015a) ‘Cox Proportional Hazards Regression Model’, in *Regression Modeling Strategies*. 2nd edn. Springer, Cham, pp. 475–519. doi: 10.1007/978-3-319-19425-7_20.

Harrell, F. E. (2015b) ‘Introduction to Survival Analysis’, in *Regression Modeling Strategies*. 2nd edn. Springer, Cham, pp. 399–422. doi: 10.1007/978-3-319-19425-7_17.

Health and Social Care Information Centre (2014) *METHODOLOGY TO CREATE PROVIDER AND CIP SPELLS FROM HES APC DATA*, *The National Archives*. Available at: https://webarchive.nationalarchives.gov.uk/20180328130852tf_/http://content.digital.nhs.uk/media/11859/Provider-Spells-Methodology/pdf/Spells_Methodology.pdf/ (Accessed: 5 July 2018).

Van Der Heijden, A. A. W. A. *et al.* (2014) ‘Validation of a model to estimate personalised screening frequency to monitor diabetic retinopathy’, *Diabetologia*. Springer Verlag, 57, pp. 1332–1338. doi: 10.1007/s00125-014-3246-4.

Helwig, N. E. (2017) *Bootstrap Confidence Intervals*. Available at: http://users.stat.umn.edu/~helwig/notes/bootci-Notes.pdf (Accessed: 9 December 2020).

Herdy, A. H. *et al.* (2014) ‘South American guidelines for cardiovascular disease prevention and rehabilitation.’, *Arquivos brasileiros de cardiologia*. Brazil, 103(2 Suppl 1), pp. 1–31. Available at: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med8&NEWS=N&AN=25387466.

Herrett, E. *et al.* (2015) ‘Data Resource Profile: Clinical Practice Research Datalink (CPRD)’, *International Journal of Epidemiology*, 44(3), pp. 827–836. doi: 10.1093/ije/dyv098.

Hess, K., Gentleman, R. and Winsemius, D. (2019) ‘Hazard Function Estimation

in Survival Analysis: “muhaz” Package’. Available at: https://cran.r-project.org/web/packages/muhaz/index.html.

Hippisley-Cox, J. *et al.* (2008) ‘Predicting cardiovascular risk in England and Wales: Prospective derivation and validation of QRISK2’, *BMJ*, 336, p. 1475. doi: 10.1136/bmj.39609.449676.25.

Hippisley-Cox, J., Coupland, C. and Brindle, P. (2017) ‘Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: Prospective cohort study’, *BMJ*, 357, p. j2099. doi: 10.1136/bmj.j2099.

History & Maps (2015) *Glasgow Postcode Map*, *La Historia con Mapa*. Available at: http://www.lahistoriaconmapas.com/atlas/country-map193/glasgow-postcode-map.htm (Accessed: 1 April 2019).

Ho, P. M., Bryson, C. L. and Rumsfeld, J. S. (2009) ‘Medication adherence: Its importance in cardiovascular outcomes’, *Circulation*, 119(23), pp. 3028–3035. doi: 10.1161/CIRCULATIONAHA.108.768986.

Hobbs, F. D. R. *et al.* (2016) ‘Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14’, *The Lancet*. Lancet Publishing Group, 387(10035), pp. 2323–2330. doi: 10.1016/S0140-6736(16)00620-6.

Holmes, M. V. *et al.* (2014) ‘Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data’, *BMJ*, 349, p. g4164. doi: 10.1136/bmj.g4164.

Hope, H. F. *et al.* (2019) ‘Systematic review of the predictors of statin adherence for the primary prevention of cardiovascular disease’, *PLoS ONE*, 14(1). doi: 10.1371/journal.pone.0201196.

Information Services Division (2010) *Prescribing Information System Fields Document*, *Prescribing Information System Fields (1993 - Present)*. Available at: https://www.isdscotland.org/Health-Topics/Prescribing-and-

Medicines/Prescribing-Datamarts/docs/PIS_fields_for_researchers_v5_eDRIS Guidance.pdf (Accessed: 3 May 2018).

Information Services Division (2016) *Quality and Outcomes Framework: 2015/16 QOF Prevalence Data*, *Quality and Outcomes Framework*. Available at: https://www.isdscotland.org/Health-Topics/General-Practice/Quality-And-Outcomes-Framework/2015-16/Register-and-prevalence-data.asp (Accessed: 18 June 2020).

Information Services Division (2017) 'Scottish Heart Disease Statistics', (February). Available at: http://www.isdscotland.org/Health-Topics/Heart-Disease/Publications/2017-02-21/2017-02-21-Heart-Disease-Report.pdf.

Institute of Medicine (US) (2011) 'Clinical Practice Guidelines We Can Trust: Introduction', in Graham, R. et al. (eds) *Clinical Practice Guidelines We Can Trust*. 1st edn. Washington (DC): National Academies Press (US), p. Introduction. Available at: https://www.ncbi.nlm.nih.gov/books/NBK209539/.

International Diabetes Federation Guideline Development Group (2014) 'Global guideline for type 2 diabetes.', *Diabetes Research and Clinical Practice*. Edited by B.-N. H. Aschner P Bennett P, Boulton A, Colagiuri R, Colagiuri S, Franz M, Gadsby R, Gagliardino JJ, Home P, McGill M, Manley S, Marshall S, Mbanya JC, Neil A, Ramachandran A, Ramaiya K, Roglic G, Schaper N, Siminerio L, Sinclair A, Snoek F, Van Crombrugge P, Ve. Ireland, 104(1), pp. 1–52. doi: https://dx.doi.org/10.1016/j.diabres.2012.10.001.

Jacobson, T. A. *et al*. (2015) 'National Lipid Association recommendations for patient-centered management of dyslipidemia: Part 1 - full report', *Journal of Clinical Lipidology*. United States, 9(2), pp. 129–169. doi: https://dx.doi.org/10.1016/j.jacl.2014.07.007.

JBS3 Board (2014) 'Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3).', *Heart*. Edited by S. N. Deanfield J Simpson I, Wood D, Bradbury K, Fox K, Boon N, Winocour P, Feher M, Doherty P, Buckley J, Jennings C, Keenan J, Weissberg P, Knapton M, Williams B, Poulter N, Stevens P, Taal M, Grant P, Home P, Sattar N, Rees A, Neely D, Wheeler D,

Goldsmith. England, 100, pp. ii1–ii67. doi: https://dx.doi.org/10.1136/heartjnl-2014-305693.

Jordanov, M. S. and Abou Assi, H. (2015) 'Statins', in Yassine, H. (ed.) *Lipid Management: From Basics to Clinic*. 1st edn. Springer International Publishing, pp. 123–147. doi: 10.1007/978-3-319-11161-2_8.

Jousilahti, P. *et al.* (1999) 'Sex, Age, Cardiovascular Risk Factors, and Coronary Heart Disease', *Circulation*, 99(9), pp. 1165–1172. doi: 10.1161/01.cir.99.9.1165.

Jung, T. and Wickrama, K. A. S. (2008) 'An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling', *Social and Personality Psychology Compass*, 2(1), pp. 302–317. doi: 10.1111/j.1751-9004.2007.00054.x.

Kannel, W. B. (1996) 'Blood pressure as a cardiovascular risk factor: Prevention and treatment', *JAMA*, 275(20), pp. 1571–1576. doi: 10.1001/jama.275.20.1571.

Kassambara, A. *et al.* (2020) 'Drawing Survival Curves using "ggplot2"'. Available at: https://cran.r-project.org/package=survminer.

Katz, M. H. (2006) *Study Design and Statistical Analysis: A Practical Guide for Clinicians*, *Study Design and Statistical Analysis: A Practical Guide for Clinicians*. University of California, San Francisco: Cambridge University Press. doi: 10.1017/CBO9780511616761.

Kent, P., Jensen, R. K. and Kongsted, A. (2014) 'A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB', *BMC Medical Research Methodology*, 14, p. 113. doi: 10.1186/1471-2288-14-113.

Khanji, M. Y. *et al.* (2016) 'Cardiovascular Risk Assessment', *Annals of Internal Medicine*, 165(10), p. 713. doi: 10.7326/M16-1110.

Khanji, M. Y. *et al.* (2018) 'Lifestyle advice and interventions for cardiovascular risk reduction: A systematic review of guidelines.', *International journal of*

cardiology. Netherlands ([Comment in: Int J Cardiol. 2018 Jul 15;263:156-157; PMID: 29754913 [https://www.ncbi.nlm.nih.gov/pubmed/29754913]]), 263, pp. 142–151. doi: https://dx.doi.org/10.1016/j.ijcard.2018.02.094.

Khunti, K. *et al.* (2018) 'Association of a Combined Measure of Adherence and Treatment Intensity With Cardiovascular Outcomes in Patients With Atherosclerosis or Other Cardiovascular Risk Factors Treated With Statins and/or Ezetimibe', *JAMA Network Open*, 1(8). doi: 10.1001/jamanetworkopen.2018.5554.

Kim, E. S. *et al.* (2014) 'Life satisfaction and frequency of doctor visits', *Psychosomatic Medicine*. NIH Public Access, 76(1), pp. 86–93. doi: 10.1097/PSY.0000000000000024.

Kinoshita, M. *et al.* (2018) 'Japan Atherosclerosis Society (JAS) Guidelines for Prevention of Atherosclerotic Cardiovascular Diseases 2017', *Journal of Atherosclerosis and Thrombosis*, 25(9), pp. 846–984. doi: 10.5551/jat.GL2017.

Kinra, S. *et al.* (2003) 'Is maternal transmission of coronary heart disease risk stronger than paternal transmission?', *Heart*, 89, pp. 834–838. doi: 10.1136/heart.89.8.834.

Kirkwood, B. R. and Sterne, J. A. C. (2003) *Essential Medical Statistics*. 2nd edn, *Medical statistics*. 2nd edn. Blackwell Publishing.

Kleinbaum, D. G. and Klein, M. (2012) 'The Cox Proportional Hazards Model and Its Characteristics', in *Survival Analysis: A self-Learning Text*. 3rd edn. New York, NY: Springer Science and Business Media, pp. 97–159. doi: 10.1007/978-1-4419-6646-9_3.

Klug, E. *et al.* (2018) 'South African dyslipidaemia guideline consensus statement: 2018 update A joint statement from the South African Heart Association (SA Heart) and the Lipid and Atherosclerosis Society of Southern Africa (LASSA).', *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde*. South Africa, 108(11b), pp. 973–1000. Available at: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med15&NEWS=N

&AN=30421699.

Krueger, K. *et al*. (2018) ‘In search of a standard when analyzing medication adherence in patients with heart failure using claims data: a systematic review’, *Heart Failure Reviews*. Springer New York LLC, 23(1), pp. 63–71. doi: 10.1007/s10741-017-9656-x.

Kyrgiou, M. *et al*. (2016) ‘Personalised management of women with cervical abnormalities using a clinical decision support scoring system’, *Gynecologic Oncology*. Academic Press Inc., 141(1), pp. 29–35. doi: 10.1016/j.ygyno.2015.12.032.

Lam, W. Y. and Fresco, P. (2015) ‘Medication Adherence Measures: An Overview’, *BioMed Research International*. Hindawi Publishing Corporation, 2015(217047), pp. 1–12. doi: 10.1155/2015/217047.

Landis, J. R. and Koch, G. G. (1977) ‘The Measurement of Observer Agreement for Categorical Data’, *Biometrics*. JSTOR, 33(1), p. 174. doi: 10.2307/2529310.

Lang, S. J. *et al*. (2016) ‘Impact of socioeconomic deprivation on screening for cardiovascular disease risk in a primary prevention population: A cross-sectional study’, *BMJ Open*, 6, p. e009984. doi: 10.1136/bmjopen-2015-009984.

Lassenius, M. I. *et al*. (2020) ‘Cardiovascular event rates increase after each recurrence and associate with poor statin adherence’, *European Journal of Preventive Cardiology*. SAGE Publications Inc., p. 2047487320904334. doi: 10.1177/2047487320904334.

Lavikainen, P. *et al*. (2016) ‘Statin adherence and risk of acute cardiovascular events among women: A cohort study accounting for time-dependent confounding affected by previous adherence’, *BMJ Open*. doi: 10.1136/bmjopen-2016-011306.

Lee, S. *et al*. (2019) ‘Genetic testing for familial hypercholesterolemia: Impact on diagnosis, treatment and cardiovascular risk.’, *European Journal of Preventive Cardiology*. SAGE Publications Inc., 26(12), pp. 1262–1270. doi:

10.1177/2047487319829746.

Leibowitz, M. *et al.* (2017) 'Targeting LDL Cholesterol: Beyond Absolute Goals Toward Personalized Risk', *Current Cardiology Reports*. Current Medicine Group LLC 1, 19, p. 52. doi: 10.1007/s11886-017-0858-6.

Leskelä, R. L. *et al.* (2020) 'Outcomes of lipid control in secondary prevention of coronary artery disease in Finland: A 24-month follow-up after acute coronary syndrome', *Atherosclerosis*. Elsevier Ireland Ltd, 296, pp. 4–10. doi: 10.1016/j.atherosclerosis.2020.01.018.

Leslie, K. H., McCowan, C. and Pell, J. P. (2018) 'Adherence to cardiovascular medication: a review of systematic reviews', *Journal of Public Health*. Advance Access Publication, 41(1), pp. 84–94. doi: 10.1093/pubmed/fdy088.

Levine, G. N. *et al.* (2016) '2016 ACC/AHA Guideline Focused Update on Duration of Dual Antiplatelet Therapy in Patients With Coronary Artery Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines', *Journal of the American College of Cardiology*. Elsevier USA, 68(10), pp. 1082–1115. doi: 10.1016/j.jacc.2016.03.513.

Li, Y.-H. *et al.* (2018) '2018 Guidelines of the Taiwan Society of Cardiology, Taiwan Society of Emergency Medicine and Taiwan Society of Cardiovascular Interventions for the management of non ST-segment elevation acute coronary syndrome.', *Journal of the Formosan Medical Association = Taiwan yi zhi*. Singapore, 117(9), pp. 766–790. doi: https://dx.doi.org/10.1016/j.jfma.2018.06.002.

Lim, S. S. *et al.* (2012) 'A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010', *The Lancet*, 380(9859), pp. 2224–2260. doi: 10.1016/S0140-6736(12)61766-8.

Linzer, D. A. and Lewis, J. B. (2011) 'poLCA: An R Package for Polytomous Variable Latent Class Analysis. R Package'. Journal of Statistical Software, pp. 1–29. Available at: http://www.jstatsoft.org/v42/i10/.

Linzer, D. and Lewis, J. (2016) *Package 'poLCA' Manual*. Available at: https://cran.r-project.org/web/packages/poLCA/poLCA.pdf (Accessed: 8 December 2020).

Lloyd-Jones, D. M. *et al.* (2004) 'Parental Cardiovascular Disease as a Risk Factor for Cardiovascular Disease in Middle-aged Adults: A Prospective Study of Parents and Offspring', *JAMA*, 291(18), pp. 2204–2211. doi: 10.1001/jama.291.18.2204.

Lloyd, T. *et al.* (2019) 'Effect on secondary care of providing enhanced support to residential and nursing home residents: A subgroup analysis of a retrospective matched cohort study', *BMJ Quality and Safety*. BMJ Publishing Group, 28(7), pp. 534–546. doi: 10.1136/bmjqs-2018-009130.

Looker, H. C. *et al.* (2013) 'Predicted impact of extending the screening interval for diabetic retinopathy: The Scottish Diabetic Retinopathy Screening programme', *Diabetologia*, 56(8), pp. 1716–1725. doi: 10.1007/s00125-013-2928-7.

Low Wang, C. C. *et al.* (2016) 'Clinical update: Cardiovascular disease in diabetes mellitus', *Circulation*, 133(24), pp. 2459–2502. doi: 10.1161/CIRCULATIONAHA.116.022194.

Lund, S. H. *et al.* (2016) 'Individualised risk assessment for diabetic retinopathy and optimisation of screening intervals: A scientific approach to reducing healthcare costs', *British Journal of Ophthalmology*. BMJ Publishing Group, 100(5), pp. 683–687. doi: 10.1136/bjophthalmol-2015-307341.

Mach, F. *et al.* (2020) '2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid modification to reduce cardiovascular risk', *European Heart Journal*, 41(1), pp. 111–188. doi: 10.1093/eurheartj/ehz455.

Masyn, K. E. (2013) 'Latent Class Analysis and Finite Mixture Modeling', in Little, T. D. (ed.) *The Oxford Handbook of Quantitative Methods*. 1st edn. New York: Oxford University Press, pp. 551–611. Available at: https://www.statmodel.com/download/Masyn_2013.pdf (Accessed: 7 December 2020).

Mathers, C. D. and Loncar, D. (2006) 'Projections of global mortality and burden of disease from 2002 to 2030', *PLoS Medicine*, 3(11), pp. 2011–2030. doi: 10.1371/journal.pmed.0030442.

McGhee, S., Harding, S. P. and Wong, D. (2012) 'Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy by Aspelund et al. (2011) Diabetologia 54:2525-2532', *Graefe's Archive for Clinical and Experimental Ophthalmology*. Springer, pp. 477–478. doi: 10.1007/s00417-012-1939-2.

McNamee, R. (2003) 'Confounding and confounders', *Occupational and Environmental Medicine*. BMJ Publishing Group Ltd, 60(3), pp. 227–234. doi: 10.1136/oem.60.3.227.

Morgen, E. K. and Naugler, C. (2015) 'Inappropriate Repeats of Six Common Tests in a Canadian City: A Population Cohort Study Within a Laboratory Informatics Framework', *American Journal of Clinical Pathology*. American Society of Clinical Pathologists, 144(5), pp. 704–712. doi: 10.1309/AJCPYXDAUS2F8XJY.

Morrison, C. *et al.* (1997) 'Effect of socioeconomic group on incidence of, management of, and survival after myocardial infarction and coronary death: Analysis of community coronary event register', *British Medical Journal*, 314, p. 541. doi: 10.1136/bmj.314.7080.541.

Mosca, L., Barrett-Connor, E. and Kass Wenger, N. (2011) 'Sex/gender differences in cardiovascular disease prevention: What a difference a decade makes', *Circulation*, 124(19), pp. 2145–2154. doi: 10.1161/CIRCULATIONAHA.110.968792.

Mundi, S. *et al.* (2018) 'Endothelial permeability, LDL deposition, and cardiovascular risk factors-A review', *Cardiovascular Research*, 114(1), pp. 35–52. doi: 10.1093/cvr/cvx226.

Naderi, S. H., Bestwick, J. P. and Wald, D. S. (2012) 'Adherence to drugs that prevent cardiovascular disease: Meta-analysis on 376,162 patients', *American*

Journal of Medicine, 125(9), pp. 882-887.e1. doi: 10.1016/j.amjmed.2011.12.013.

National Health Service (2019a) *National Cost Collection for the NHS England 2018-19*. Available at: https://www.england.nhs.uk/national-cost-collection/ (Accessed: 13 January 2021).

National Health Service (2019b) *NHS England: Missed GP appointments costing NHS millions*. Available at: https://www.england.nhs.uk/2019/01/missed-gp-appointments-costing-nhs-millions/ (Accessed: 22 January 2021).

National Institute for Health and Care Excellence (2015) *National Clinical Guideline Centre: Preoperative Tests*. Available at: https://www.nice.org.uk/guidance/NG45/documents/guideline-appendices-13 (Accessed: 20 March 2020).

National Records of Scotland (2017) *Vital Events Reference Tables 2017 Section 6: Deaths - Causes*. Available at: https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/vital-events-reference-tables/2017/section-6-death-causes (Accessed: 28 January 2021).

National Records of Scotland (2019) 'Archived Postcode Extract | National Records of Scotland', *Archived Postcode Extract*. Scottish Government. Available at: https://www.nrscotland.gov.uk/statistics-and-data/geography/our-products/scottish-postcode-directory/archived-postcode-extract (Accessed: 13 October 2020).

National Records Scotland (2014a) *MEDICAL CERTIFICATE OF CAUSE OF DEATH*, *National Records Scotland Statistics*. Available at: https://www.nrscotland.gov.uk/files/statistics/death-certificates/f11-mccd-from-6-aug-14.pdf (Accessed: 22 April 2018).

National Records Scotland (2014b) *NRS: Mid-2014 Small Area Population Estimates Scotland*. Available at: https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-estimates/2001-

based-special-area-population-estimates/small-area-population-estimates/mid-2014/detailed-data-zone-tables (Accessed: 18 June 2020).

Naylor, A. R. *et al.* (2018) 'Management of Atherosclerotic Carotid and Vertebral Artery Disease: 2017 Clinical Practice Guidelines of the European Society for Vascular Surgery (ESVS)', *European Journal of Vascular and Endovascular Surgery*. Elsevier Ltd, 55(1), pp. 3–81. doi: doi.org/10.1016/j.ejvs.2017.06.021.

Neaton, J. D. *et al.* (1992) 'Serum Cholesterol Level and Mortality Findings for Men Screened in the Multiple Risk Factor Intervention Trial', *Archives of Internal Medicine*, 152(7), pp. 1490–1500. doi: 10.1001/archinte.1992.00400190110021.

New Zealand Ministry of Health (2018) *Cardiovascular Disease Risk Assessment and Management for Primary Care*. Available at: https://www.health.govt.nz/system/files/documents/publications/cvd-risk-assessment-and-management-for-primary-care-v2.pdf.

NHS (2018) *NHS: Statins*. Available at: https://www.nhs.uk/conditions/statins/ (Accessed: 1 April 2019).

NHS Digital (2018) *Integrated Care Programme: Standardising Health Records*. Available at: https://digital.nhs.uk/about-nhs-digital/our-work/transforming-health-and-care-through-technology/integrated-care-domain-d/integrated-care-programme# (Accessed: 18 March 2020).

NHS Greater Glasgow and Clyde (2015a) *ESSENTIAL NHS INFORMATION ABOUT HOSPITAL CLOSURES AFFECTING YOU, NHS Greater Glasgow and Clyde Information Leaflets*. Available at: https://www.shb.scot.nhs.uk/departments/childhealth/NewHospitalLeaflet.pdf (Accessed: 17 April 2018).

NHS Greater Glasgow and Clyde (2015b) *NHS GGC: Director of Public Health Report 2015-2017; Chapter 1*. Available at: https://www.nhsggc.org.uk/media/234486/nhsggc_ph_dphreport2015_population_of_nhsggc.pdf.

NHS Greater Glasgow and Clyde (2015c) *NHS GGC Director of Public Health Report 2015-2017; Chapter 5*. Available at: https://www.nhsggc.org.uk/media/234515/nhsggc_ph_dphreport2015_health_related_behaviours.pdf.

NHS Inform Scotland (2020) *Prescription charges and exemptions| NHS inform*. Available at: https://www.nhsinform.scot/care-support-and-rights/nhs-services/pharmacy/prescription-charges-and-exemptions (Accessed: 4 November 2020).

NHS Research Scotland (2015a) *Electronic Health Records*. Available at: http://www.nhsresearchscotland.org.uk/research-in-scotland/data/sub-page-4 (Accessed: 18 March 2020).

NHS Research Scotland (2015b) *Safe Havens*. Available at: https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens (Accessed: 18 March 2020).

NICE (2013) *Myocardial infarction: cardiac rehabilitation and prevention of further cardiovascular disease (CG172)*.

NICE (2014) *Cardiovascular disease: risk assessment and reduction, including lipid modification (CG181)*.

NICE (2019a) *Hypertension in adults: Diagnosis and management (NG136)*. Available at: https://www.nice.org.uk/guidance/ng136/resources/hypertension-in-adults-diagnosis-and-management-pdf-66141722710213.

NICE (2019b) *Treatment Steps for Hypertension: NG136, Nice Guidance Pathways*. Available at: https://pathways.nice.org.uk/pathways/hypertension (Accessed: 16 January 2020).

Nordestgaard, B. G. *et al.* (2013) 'Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society.', *European heart journal*. England, 34(45), pp.

3478–90a. doi: https://dx.doi.org/10.1093/eurheartj/eht273.

Nordstrom, B. L. *et al.* (2015) 'Treatment patterns and lipid levels among patients with high-risk atherosclerotic CVD in the UK', *British Journal of Cardiology*. doi: 10.5837/bjc.2015.034.

O'Sullivan, J. W. *et al.* (2018) 'Temporal trends in use of tests in UK primary care, 2000-15: Retrospective analysis of 250 million tests', *The BMJ*, 363, p. k4666. doi: 10.1136/bmj.k4666.

Oberski, D. (2015) *Latent class analysis*. Available at: http://daob.nl/wp-content/uploads/2015/07/ESRA-course-slides.pdf (Accessed: 8 December 2020).

Ohm, J. *et al.* (2019) 'Lipid levels achieved after a first myocardial infarction and the prediction of recurrent atherosclerotic cardiovascular disease', *International Journal of Cardiology*. Elsevier Ireland Ltd, 296, pp. 1–7. doi: 10.1016/j.ijcard.2019.07.001.

OpenStreetMap (2020a) *Mercator - OpenStreetMap Wiki*, *OpenStreetMap Wiki*. Available at: https://wiki.openstreetmap.org/wiki/Mercator (Accessed: 13 October 2020).

OpenStreetMap (2020b) *OpenStreetMap*. Available at: https://www.openstreetmap.org/#map=6/50.868/-5.559 (Accessed: 13 October 2020).

Osterberg, L. and Blaschke, T. (2005) 'Adherence to Medication', *New England Journal of Medicine*, 353(5), pp. 487–497. doi: 10.1056/NEJMra050100.

Pallazola, V. *et al.* (2018) *Was the Juice Worth the Squeeze? Understanding the New 2018 AHA/ACC Cholesterol Guideline - American College of Cardiology*, *American College of Cardiology: Latest in Cardiology*. Available at: https://www.acc.org/latest-in-cardiology/articles/2018/11/14/10/48/was-the-juice-worth-the-squeeze (Accessed: 25 March 2020).

Pashayan, N. *et al.* (2018) 'Cost-effectiveness and Benefit-to-Harm Ratio of Risk-

Stratified Screening for Breast Cancer A Life-Table Model', *JAMA Oncology*. American Medical Association, 4(11), pp. 1504–1510. doi: 10.1001/jamaoncol.2018.1901.

Pell, J. P. *et al.* (2008) 'Smoke-free Legislation and Hospitalizations for Acute Coronary Syndrome', *New England Journal of Medicine*. Massachussetts Medical Society, 359, pp. 482–491. doi: 10.1056/NEJMsa0706740.

Perak, A. M. *et al.* (2016) 'Long-term risk of atherosclerotic cardiovascular disease in US adults with the familial hypercholesterolemia phenotype', *Circulation*, 134(1), pp. 9–19. doi: 10.1161/CIRCULATIONAHA.116.022335.

Perera, R. *et al.* (2015) 'Optimal strategies for monitoring lipid levels in patients at risk or with cardiovascular disease: A systematic review with statistical and cost-effectiveness modelling', *Health Technology Assessment*, 19(100). doi: 10.3310/hta191000.

Piepoli, M. F. *et al.* (2016) '2016 European Guidelines on cardiovascular disease prevention in clinical practice', *European Heart Journal*, 37(29), pp. 2315–2381. doi: 10.1093/eurheartj/ehw106.

Preiss, D. *et al.* (2011) 'Risk of Incident Diabetes With Intensive-Dose Compared With Moderate-Dose Statin Therapy', *JAMA*, 305(24), pp. 2556–2564. doi: 10.1001/jama.2011.860.

Price, J. F. *et al.* (1999) 'Relationship between smoking and cardiovascular risk factors in the development of peripheral arterial disease and coronary artery disease. Edinburgh Artery Study', *European Heart Journal*, 20(5), pp. 344–353. doi: 10.1053/euhj.1998.1194.

Proust-Lima, C. *et al.* (2020) *Package 'lcmm' Manual*. doi: 10.18637/jss.v078.i02.

Proust-Lima, Cecile, Philipps, V. and Liquet, B. (2017) 'Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: R Package'. Journal of Statistical Software, pp. 1–56. Available at: 10.18637/jss.v078.i02.

Proust-Lima, Cécile, Philipps, V. and Liquet, B. (2017) ‘Estimation of extended mixed models using latent classes and latent processes: The R package lcmm’, *Journal of Statistical Software*. American Statistical Association, 78(2), pp. 1–56. doi: 10.18637/jss.v078.i02.

R-Bloggers (2016a) *Cox Model Assumptions | R-bloggers*, *Easy Guides, R-Bloggers*. Available at: https://www.r-bloggers.com/cox-model-assumptions/ (Accessed: 21 July 2020).

R-Bloggers (2016b) *Survival Analysis Basics | R-bloggers*, *Easy Guides, R-Bloggers*. Available at: https://www.r-bloggers.com/survival-analysis-basics/ (Accessed: 17 July 2020).

R Core Team (2018) ‘R: A Language and Environment for Statistical Computing’. Vienna, Austria. Available at: http://www.r-project.org/.

Raebel, M. A. *et al.* (2014) ‘Standardizing Terminology and Definitions of Medication Adherence and Persistence in Research employing Electronic Databases’, *Medical Care*, 51(8 Supplement 3), pp. 11–21. doi: 10.1097/MLR.0b013e31829b1d2a.Standardizing.

Rawshani, Aidin *et al.* (2017) ‘Mortality and cardiovascular disease in type 1 and type 2 diabetes’, *New England Journal of Medicine*, 376(15), pp. 1407–1418. doi: 10.1056/NEJMoa1608664.

Rawshani, Aidin *et al.* (2018) ‘Risk factors, mortality, and cardiovascular outcomes in patients with type 2 diabetes’, *New England Journal of Medicine*, 379(7), pp. 633–644. doi: 10.1056/NEJMoa1800256.

Rawshani, Araz *et al.* (2018) ‘Excess mortality and cardiovascular disease in young adults with type 1 diabetes in relation to age at onset: a nationwide, register-based cohort study’, *The Lancet*. Lancet Publishing Group, 392(10146), pp. 477–486. doi: 10.1016/S0140-6736(18)31506-X.

Read, S. H. *et al.* (2016) ‘Trends in type 2 diabetes incidence and mortality in Scotland between 2004 and 2013’, *Diabetologia*. Springer Verlag, 59(10), pp.

2106–2113. doi: 10.1007/s00125-016-4054-9.

Read, S. H. *et al.* (2019) 'Trends in incidence and case fatality of acute myocardial infarction, angina and coronary revascularisation in people with and without type 2 diabetes in Scotland between 2006 and 2015', *Diabetologia*. Springer Verlag, 62(3), pp. 418–425. doi: 10.1007/s00125-018-4796-7.

Rees, K. *et al.* (2013) '"Mediterranean" dietary pattern for the primary prevention of cardiovascular disease', *Cochrane Database of Systematic Reviews*, (8). doi: 10.1002/14651858.CD009825.pub2.

Rees, K. *et al.* (2019) 'Mediterranean-style diet for the primary and secondary prevention of cardiovascular disease', *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd, (3). doi: 10.1002/14651858.CD009825.pub3.

Reiner, Z. *et al.* (2016) 'Lipid lowering drug therapy in patients with coronary heart disease from 24 European countries--Findings from the EUROASPIRE IV survey.', *Atherosclerosis*. Ireland, 246, pp. 243–250. doi: https://dx.doi.org/10.1016/j.atherosclerosis.2016.01.018.

Rodriguez, F. *et al.* (2017) 'Association between intensity of statin therapy and mortality in patients with atherosclerotic cardiovascular disease', *JAMA Cardiology*, 4(3), pp. 206–213. doi: 10.1001/jamacardio.2016.4052.

Román, M. *et al.* (2017) 'Cumulative risk of breast cancer screening outcomes according to the presence of previous benign breast disease and family history of breast cancer: Supporting personalised screening', *British Journal of Cancer*. Nature Publishing Group, 116, pp. 1480–1485. doi: 10.1038/bjc.2017.107.

Ronksley, P. E. *et al.* (2011) 'Association of alcohol consumption with selected cardiovascular disease outcomes: A systematic review and meta-analysis', *BMJ*, 342, p. d671. doi: 10.1136/bmj.d671.

Roth, G. A. *et al.* (2017) 'Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015', *Journal of the American College of Cardiology*, 70(1), pp. 1–25. doi: 10.1016/j.jacc.2017.04.052.

Ryan, T. P. *et al*. (2017) 'Medication adherence, medical record accuracy, and medication exposure in real-world patients using comprehensive medication monitoring', *PloS One*. Edited by C. Leroyer. Public Library of Science, 12(9), p. e0185471. doi: 10.1371/journal.pone.0185471.

Sabatine, M. S. *et al*. (2017) 'Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease', *New England Journal of Medicine*. Massachussetts Medical Society, 376(18), pp. 1713–1722. doi: 10.1056/NEJMoa1615664.

Sattar, N. *et al*. (2010) 'Statins and risk of incident diabetes: A collaborative meta-analysis of randomised statin trials', *Lancet*, 375(9716), pp. 735–742. doi: 10.1016/S0140-6736(09)61965-6.

Sattar, N. and Preiss, D. (2017) 'Reverse Causality in Cardiovascular Epidemiological Research: More Common Than Imagined?', *Circulation*. Lippincott Williams and Wilkins, pp. 2369–2372. doi: 10.1161/CIRCULATIONAHA.117.028307.

Scheibl, F. *et al*. (2019) 'When Frail Older People Relocate in Very Old Age, Who Makes the Decision?', *Innovation in Aging*, 3(4), p. igz030. doi: 10.1093/geroni/igz030.

Schoenfeld, D. (1982) 'Partial residuals for the proportional hazards regression model', *Biometrika*, 69(1), pp. 239–241. doi: doi.org/10.1093/biomet/69.1.239.

Schwartz, G. G. *et al*. (2012) 'Effects of dalcetrapib in patients with a recent acute coronary syndrome', *New England Journal of Medicine*, 367(22), pp. 2089–2099. doi: 10.1056/NEJMoa1206797.

Schwartz, G. G. *et al*. (2018) 'Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome', *New England Journal of Medicine*. Massachussetts Medical Society, 379(22), pp. 2097–2107. doi: 10.1056/NEJMoa1801174.

ScotPHO (2016) *ScotPHO: Diabetes: estimates of number of cases by NHS Board*. Available at: https://www.scotpho.org.uk/health-wellbeing-and-disease/diabetes/data/nhs-boards/ (Accessed: 18 June 2020).

Scottish Care Information Diabetes Collaboration (2015) *SCI-Diabetes*. Available at: http://www.sci-diabetes.scot.nhs.uk/ (Accessed: 3 April 2019).

Scottish Government (2013) *High Level Summary of Statistics Trend: Scottish Index of Multiple Deprivation*, *Scottish Government Statistics*. Available at: https://www2.gov.scot/Topics/Statistics/Browse/Social-Welfare/TrendSIMD (Accessed: 16 October 2018).

Scottish Government (2015) *SIMD 2012 Health Board level quintiles and deciles*, *Scottish Government Statistics*. Available at: https://www2.gov.scot/Topics/Statistics/SIMD/localHBquin09 (Accessed: 16 October 2018).

Scottish Government (2018) *Scottish Index of Multiple Deprivation*, *Scottish Government Statistics*. Available at: https://www2.gov.scot/Topics/Statistics/SIMD (Accessed: 16 October 2018).

Shaper, A. G., Wannamethee, G. and Walker, M. (1988) 'Alcohol and Mortality in British Men: Explaining the U-Shaped Curve', *The Lancet*. Elsevier, 332(8623), pp. 1267–1273. doi: 10.1016/S0140-6736(88)92890-5.

Sheetz, K. H. and Nathan, H. (2020) 'Methods for Enhancing Causal Inference in Observational Studies', in Dimick, J. B. and Lubitz, C. C. (eds) *Health Services Research*. 2nd edn. Switzerland: Springer, Cham, pp. 185–197. doi: 10.1007/978-3-030-28357-5_16.

Shepherd, J. *et al.* (1995) 'Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia', *New England Journal of Medicine*, 333, pp. 1301–1308. doi: 10.1056/NEJM199511163332001.

Shetty, J. *et al.* (2016) 'Adherence to antiepileptic drugs in children with epilepsy in a Scottish population cohort', *Developmental Medicine and Child Neurology*. doi: 10.1111/dmcn.12942.

SIGN (2007) *SIGN 97: Risk estimation and the prevention of CHD: Estimating Cardiovascular Risk*. Available at:

http://www.sign.ac.uk/guidelines/fulltext/97/section3.html (Accessed: 17 November 2017).

SIGN (2013) *SIGN 129: Antithrombotics: indications and management*. Available at: http://www.sign.ac.uk/pdf/SIGN129.pdf.

SIGN (2017) 'SIGN 149: Risk estimation and the prevention of cardiovascular disease', (June). Available at: http://www.sign.ac.uk/assets/sign149.pdf.

Skyler, J. S. *et al.* (2009) 'Intensive glycemic control and the prevention of cardiovascular events: Implications of the ACCORD, ADVANCE, and VA diabetes trials', *Diabetes Care*, 32(1), pp. 187–192. doi: 10.2337/dc08-9026.

Smellie, W. S. A. (2012) 'Demand management and test request rationalization', *Annals of Clinical Biochemistry*, 49(4), pp. 323–336. doi: 10.1258/acb.2011.011149.

Smith, D. *et al.* (2017) 'A systematic review of medication nonadherence in persons with dementia or cognitive impairment', *PLoS ONE*, 12(2), p. e0170651. doi: 10.1371/journal.pone.0170651.

Smith, G. D. *et al.* (1992) 'Plasma Cholesterol Concentration and Mortality: The Whitehall Study', *JAMA: The Journal of the American Medical Association*, 267(1), pp. 70–76. doi: 10.1001/jama.1992.03480010078028.

Smolina, K. *et al.* (2015) 'Sex Disparities in Post-Acute Myocardial Infarction Pharmacologic Treatment Initiation and Adherence: Problem for Young Women', *Circulation: Cardiovascular Quality and Outcomes*. Lippincott Williams and Wilkins, 8(6), pp. 586–592. doi: 10.1161/CIRCOUTCOMES.115.001987.

Sniderman, A. D. and Furberg, C. D. (2008) 'Age as a modifiable risk factor for cardiovascular disease', *The Lancet*, 371(9623), pp. 1547–1549. doi: 10.1016/S0140-6736(08)60313-X.

Soto-Pedre, E., Pinies, J. A. and Hernaez-Ortega, M. C. (2015) 'External validation of a risk assessment model to adjust the frequency of eye-screening

visits in patients with diabetes mellitus', *Journal of Diabetes and its Complications*. Elsevier Inc., 29(4), pp. 508–511. doi: 10.1016/j.jdiacomp.2014.12.020.

Sperber, C. M., Samarasinghe, S. R. and Lomax, G. P. (2017) 'An upper and lower bound of the Medication Possession Ratio', *Patient Preference and Adherence*, 11, pp. 1469–1478. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5590759/.

Stephenson, J. J. *et al.* (2018) 'Comparison of claims vs patient-reported adherence measures and associated outcomes among patients with nonvalvular atrial fibrillation using oral anticoagulant therapy', *Patient Preference and Adherence*. Dove Medical Press Ltd., 12, pp. 105–117. doi: 10.2147/PPA.S148697.

Stone, M. (1977) 'An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion', *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), pp. 44–47. Available at: https://www.jstor.org/stable/2984877?seq=1#metadata_info_tab_contents (Accessed: 7 January 2021).

Stone, N. J. *et al.* (2014) '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.', *Circulation*. United States ([Erratum appears in Circulation. 2015 Dec 22;132(25):e396; PMID: 26667225], [Erratum appears in Circulation. 2014 Jun 24;129(25 Suppl 2):S46-8]), 129(25 Pt B), pp. 2889–2934. doi: https://dx.doi.org/10.1161/01.cir.0000437738.63853.7a.

Sutherland, J. J. *et al.* (2018) 'Assessment of Patient Medication Adherence, Medical Record Accuracy, and Medication Blood Concentrations for Prescription and Over-the-Counter Medications', *JAMA network open*. NLM (Medline), 1(7), p. e184196. doi: 10.1001/jamanetworkopen.2018.4196.

Tai, E. S. *et al.* (2017) 'Ministry of Health Clinical Practice Guidelines: Lipids.', *Singapore medical journal*. Singapore, 58(3), pp. 155–166. doi: https://dx.doi.org/10.11622/smedj.2017018.

Taylor-Phillips, S. *et al.* (2016) 'Extending the diabetic retinopathy screening interval beyond 1 year: Systematic review', *British Journal of Ophthalmology*. BMJ Publishing Group, 100, pp. 105–114. doi: 10.1136/bjophthalmol-2014-305938.

The Blood Pressure Lowering Treatment Trialists' Collaboration (2014) 'Blood pressure-lowering treatment based on cardiovascular risk: A meta-analysis of individual patient data', *The Lancet*, 384(9943), pp. 591–598. doi: 10.1016/S0140-6736(14)61212-5.

The Emerging Risk Factors Collaboration (2009) 'Major lipids, apolipoproteins, and risk of vascular disease', *JAMA*, 302(18), pp. 1993–2000. doi: 10.1001/jama.2009.1619.

The Health Foundation (2018) *Data Linkage: joining up the dots to improve patient care*. Available at: https://www.health.org.uk/newsletter-feature/data-linkage-joining-up-the-dots-to-improve-patient-care (Accessed: 18 March 2020).

The Scottish Government (2015) *eHealth Strategy 2014 - 2017*. Available at: https://www.gov.scot/binaries/content/documents/govscot/publications/advice-and-guidance/2015/03/ehealth-strategy-2014-2017/documents/ehealth-strategy-2014-2017/ehealth-strategy-2014-2017/govscot%3Adocument/00472754.pdf (Accessed: 19 March 2020).

The Scottish Government (2018) *Scotland's Digital Health and Care Strategy*. Available at: https://www.gov.scot/publications/scotlands-digital-health-care-strategy-enabling-connecting-empowering/pages/10/ (Accessed: 22 September 2020).

The WOSCOPS Study Group (1995) 'Screening experience and baseline characteristics in the West of Scotland coronary prevention study', *The American Journal of Cardiology*, 76(7), pp. 485–491. doi: 10.1016/S0002-9149(99)80135-0.

Therneau, T., Crowson, C. and Atkinson, E. (2020) *Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model*. Available at:

https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf (Accessed: 11 December 2020).

Therneau, T. M. (2020) ‘A Package for Survival Analysis in R’. Available at: https://cran.r-project.org/package=survival.

Thomas, R. L. *et al.* (2012) ‘Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the diabetic retinopathy screening service for Wales: Retrospective analysis’, *BMJ (Online)*. British Medical Journal Publishing Group, 344, p. e874. doi: 10.1136/bmj.e874.

Thompson, M. and Walter, F. (2016) ‘Increases in general practice workload in England’, *The Lancet*. Lancet Publishing Group, pp. 2270–2272. doi: 10.1016/S0140-6736(16)00743-1.

Thygesen, K. *et al.* (2007) ‘Universal definition of myocardial infarction: Kristian Thygesen, Joseph S. Alpert and Harvey D. White on behalf of the Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction’, *European Heart Journal*. doi: 10.1093/eurheartj/ehm355.

Thygesen, K. *et al.* (2012) ‘Third universal definition of myocardial infarction.’, *Nature reviews. Cardiology*. Edited by A. J. S. Thygesen K White HD, Jaffe AS, Katus HA, Apple FS, Lindahl B, Morrow DA, Chaitman BR, Clemmensen PM, Johanson P, Hod H, Underwood R, Bax JJ, Bonow RO, Pinto F, Gibbons RJ, Fox KA, Atar D, Newby L, Galvani M, Hamm CW, Uretsky BF, Steg G, Wijns W, Bassand. England, 9(11), pp. 620–633. doi: https://dx.doi.org/10.1038/nrcardio.2012.122.

Tibuakuu, M. *et al.* (2020) ‘Patient-reported Outcomes, Resource Utilization, and Healthcare Expenditures Across the 3 Major Atherosclerotic Diseases: The Medical Expenditure Panel Survey (MEPS)’, *Circulation: Cardiovascular Quality and Outcomes*. Ovid Technologies (Wolters Kluwer Health), 13;A25(Suppl_1). doi: 10.1161/hcq.13.suppl_1.25.

Toplak, H. *et al.* (2016) ‘Austrian Lipid Consensus on the management of metabolic lipid disorders to prevent vascular complications: A joint position

statement issued by eight medical societies. 2016 update.', *Wiener klinische Wochenschrift*. Austria, 128 Suppl, pp. S216-28. doi: https://dx.doi.org/10.1007/s00508-016-0993-x.

Tunstall-Pedoe, H. and Woodward, M. (2006) 'By neglecting deprivation, cardiovascular risk scoring will exacerbate social gradients in disease', *Heart*, 92, pp. 307–310. doi: 10.1136/hrt.2005.077289.

Vasan, R. S. *et al.* (2001) 'Impact of high-normal blood pressure on the risk of cardiovascular disease', *New England Journal of Medicine*, 345(18), pp. 1291–1297. doi: 10.1056/NEJMoa003417.

Vermunt, J. K. and Magidson, J. (2003) 'Latent class models for classification', *Computational Statistics and Data Analysis*, 41(3–4), pp. 531–537. doi: 10.1016/S0167-9473(02)00179-2.

Vilaprinyo, E. *et al.* (2014) 'Cost-Effectiveness and Harm-Benefit Analyses of Risk-Based Screening Strategies for Breast Cancer', *PLoS ONE*. Edited by A. Sapino. Public Library of Science, 9(2), p. e86858. doi: 10.1371/journal.pone.0086858.

Vitale, C., Mendelsohn, M. E. and Rosano, G. M. C. (2009) 'Gender differences in the cardiovascular effect of sex hormones', *Nature Reviews Cardiology*, 6, pp. 532–542. doi: 10.1038/nrcardio.2009.105.

Vupputuri, S. *et al.* (2016) 'LDL cholesterol response and statin adherence among high-risk patients initiating treatment', *Journal of Managed Care*, 22(3). Available at: https://www.ajmc.com/journals/issue/2016/2016-vol22-n3/ldl-cholesterol-response-and-statin-adherence-among-high-risk-patients-initiating-treatment.

Wanberg, C. R. (2012) *The Oxford Handbook of Organizational Socialization*, *The Oxford Handbook of Organizational Socialization*. doi: 10.1093/oxfordhb/9780199763672.001.0001.

Wang, Y., Liu, M. and Pu, C. (2017) '2014 Chinese guidelines for secondary

prevention of ischemic stroke and transient ischemic attack.', *International journal of stroke : official journal of the International Stroke Society*. United States, 12(3), pp. 302–320. doi: https://dx.doi.org/10.1177/1747493017694391.

Ward, N. C., Watts, G. F. and Eckel, R. H. (2019) 'Statin Toxicity: Mechanistic Insights and Clinical Implications', *Circulation Research*. Lippincott Williams and Wilkins, 124(2), pp. 328–350. doi: 10.1161/CIRCRESAHA.118.312782.

Wei, L., Fahey, T. and MacDonald, T. M. (2008) 'Adherence to statin or aspirin or both in patients with established cardiovascular disease: Exploring healthy behaviour vs. drug effects and 10-year follow-up of outcome', *British Journal of Clinical Pharmacology*. doi: 10.1111/j.1365-2125.2008.03212.x.

Wellcome Trust (2015) *Enabling Data Linkage to Maximise the Value of Public Health Research Data: full report*. London. Available at: https://wellcome.org/sites/default/files/enabling-data-linkage-to-maximise-value-of-public-health-research-data-phrdf-mar15.pdf (Accessed: 22 September 2020).

Welsh, C. *et al.* (2019) 'Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease: Data from UK Biobank', *Circulation*. doi: 10.1161/CIRCULATIONAHA.119.041149.

Welsh, C. *et al.* (2020) 'Glycated Hemoglobin, Prediabetes, and the Links to Cardiovascular Disease: Data From UK Biobank', *Diabetes Care*, 43(2), p. dc191683. doi: 10.2337/dc19-1683.

Whelton, P. K. *et al.* (2018) '2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Pr', *Journal of the American College of Cardiology*. Elsevier USA, 71(19), pp. e127–e248. doi: 10.1016/j.jacc.2017.11.006.

Whittaker, E. *et al.* (2020) 'Socio-economic differences in cardiovascular disease

risk factor prevalence in people with type 2 diabetes in Scotland: a cross-sectional study', *Diabetic Medicine*. Blackwell Publishing Ltd, 37(8), pp. 1395–1402. doi: 10.1111/dme.14297.

Wilkins, E. *et al.* (2017) 'European Cardiovascular Disease Statistics 2017 edition', *European Heart Network, Brussels*, p. 192. doi: 978-2-9537898-1-2.

Wilson, A. *et al.* (2017) *High Risk Prevention Strategies*. obo. doi: 10.1093/OBO/9780199756797-0159.

Woodward, M., Brindle, P. and Tunstall-Pedoe, H. (2007) 'Adding social deprivation and family history to cardiovascular risk assessment: The ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC)', *Heart*, 93, pp. 172–176. doi: 10.1136/hrt.2006.108167.

World Health Organization (2007) *Prevention of Cardiovascular Disease: Guideline for assessment and management of cardiovascular risk*, *World Health Organization*. doi: 10.1093/innovait/inr119.

World Health Organization (2016) *ICD-10 Version: 2016*. Available at: https://icd.who.int/browse10/2016/en (Accessed: 9 April 2018).

World Health Organization (2017a) *Cardiovascular Diseases: Prevention of Recurrences of Myocardial Infarction and Stroke Study;The PREMISE programme*. Available at: http://www.who.int/cardiovascular_diseases/priorities/secondary_prevention/country/en/index1.html (Accessed: 11 January 2018).

World Health Organization (2017b) *Cardiovascular Diseases (CVDs): Fact Sheet*. Available at: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (Accessed: 3 October 2019).

World Health Organization (2018) *Alcohol*. Available at: https://www.who.int/news-room/fact-sheets/detail/alcohol (Accessed: 11 March 2020).

World Health Organization (2020) *World Health Statistics 2020: Monitoring Health for the SDGs*. Geneva. Available at: https://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf (Accessed: 22 September 2020).

Xu, R. (2020) *Lecture 7: Time-dependent Covariates in Cox Regression*. San Diego, US. Available at: http://www.math.ucsd.edu/~rxu/math284/slect7.pdf (Accessed: 11 December 2020).

Xu, T. *et al.* (2017) 'Statin Adherence and the Risk of Stroke: A Dose-Response Meta-Analysis', *CNS Drugs*, 31, pp. 263–271. doi: 10.1007/s40263-017-0420-5.

Yarnell, J. *et al.* (2003) 'Family history, longevity, and risk of coronary heart disease: The PRIME study', *International Journal of Epidemiology*, 32(1), pp. 71–77. doi: 10.1093/ije/dyg038.

Yin, C. *et al.* (2016) 'Artery tertiary lymphoid organs: Powerhouses of atherosclerosis immunity', *Frontiers in Immunology*, 7. doi: 10.3389/fimmu.2016.00387.

Yusuf, S. *et al.* (2004) 'Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study', *Lancet*, 364(9438), pp. 937–952. doi: 10.1016/S0140-6736(04)17018-9.

Zhang, Z. *et al.* (2018) 'Time-varying covariates and coefficients in Cox regression models', *Annals of Translational Medicine*. AME Publishing Company, 6(7), pp. 121–121. doi: 10.21037/atm.2018.02.12.