# Cardiovascular risk prediction in the acute care setting: a mixed methods evaluation using machine learning, real world evidence and qualitative methods

A thesis submitted to The University of Manchester for the Degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2023

Charles Reynard

School of Medical Sciences

Division of Cardiovascular Sciences

# Contents

**Word count 46,105**

# List of figures

# List of tables

# List of acronyms

| | |
|---|---|
| 95% CI | 95% confidence interval |
| ACS | Acute coronary syndromes |
| ADA | American Diabetes Association |
| AIC | Akaike information criterion |
| AMI | Acute myocardial infarction |
| APC | Admitted patient care |
| AUC | Area under the curve |
| BGTH | Burnley General Teaching Hospital |
| BIC | Bayesian information criterion |
| BP | Blood pressure |
| CHD | Coronary heart disease |
| CKD | Chronic kidney disease |
| CPH | Cox proportional hazards model |
| CPM | Clinical prediction model |
| cTn | Cardiac troponin |
| cTnT | Cardiac troponin T |
| CVD | Cardiovascular disease |
| DBP | Diastolic blood pressure |
| EBCD | Experience based co-design |
| ED | Emergency department |
| EDACS | Emergency Department Assessment of Chest Pain Score |
| eGFR | Estimated glomerular filtration rate |
| EWS | Early warning score |
| GCS | Glasgow coma scale |
| Hb | Haemoglobin |
| HDL | High density lipo-proteins |
| HDL | High density lipo-proteins |
| HES | Hospital episode statistics |
| HIGH STEACS | High-sensitivity troponin in the evaluation of patients with suspected acute coronary syndrome |
| HR | Heart rate |
| hs-cTn | High sensitivity cardiac troponin |
| hs-cTnI | High sensitivity troponin I |
| hs-cTnT | High sensitivity troponin T |
| HTN | Hypertension |
| IFCC | International Federation of Clinical Chemistry and Laboratory Medicine |
| IMD | Index of multiple deprivation |
| LDL | Low density lipo-proteins |
| LDL | Low density lipo-proteins |
| LoD | Limit of detection |
| LT CVD | Long term cardiovascular disease |
| MACCE | Major adverse cardiac and cerebral events |
| MACE | Major adverse cardiovascular event |
| MRI | Manchester Royal Infirmary |
| NHS | National Health Service |
| NHSD | National Health Service Digital |
| NICE | National Institute for Health And Care Excellence |
| NSTEMI | Non-ST elevation myocardial infarction |

| | |
|---|---|
| O/E | Observed over expected ratio |
| PPI | Patient and public involvement |
| QoF | Quality and outcomes framework |
| RBTH | Royal Blackburn Teaching Hospital |
| RCS | Restricted cubic spline |
| REM | Random effects model |
| REML | Restricted maximum likelihood method |
| RR | Respiratory rate |
| RUI | Rural urban index |
| SBP | Systolic blood pressure |
| STEMI | ST elevation myocardial infarction |
| T2 DM | Type 2 diabetes mellitus |
| TMACS | Troponin only Manchester Acute Coronary Syndrome rule out strategy |
| TMC | Trial management committee |
| TRIPOD | Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis |
| TSC | Trial steering committee |
| UK | United Kingdom |
| URL | Upper reference limit |
| USA | United States of America |
| WBC | White blood cell |
| WHO | World Health Organisations |

# Abstract

## Background

Emergency Department (ED) clinicians use clinical prediction models (CPMs) to assist with the diagnosis of acute myocardial infarction (AMI). CPMs are also used in primary care to predict future cardiovascular disease (CVD). Patient care in the ED could be improved by updating existing CPMs for AMI diagnosis or developing new care pathways to predict future CVD.

## Aims

To update a CPM for AMI diagnosis (the Troponin-only Manchester Acute Coronary Syndromes CPM) and co-develop a novel care pathway for the prediction of CVD in acute care.

## Methods

I used real-world data from three centres, linked to a national dataset, to evaluate TMACS. I used three methods to update the CPM: recalibration, model extension and dynamic Bayesian updating.

The novel care pathway for CVD prediction was co-developed with a mixed methods approach. This was informed by a systematic review and meta-analysis of potential prognostic factors. I then studied the prognostic value of factors of interest using a retrospective cohort linked to a national dataset. Finally I conducted 41 semi-structured interviews using a co-production framework to construct a novel care pathway.

## Results

TMACS demonstrated good discrimination with an area under the curve (AUC) of 0.88 (95% CI 0.86 to 0.89), but calibration had deteriorated with a calibration in the large (CITL) of -3.93 (95% CI -4.12 to -3.74). Dynamic updating demonstrated favourable model characteristics with an AUC of 0.87 (95% CI 0.85 to 0.89) and CITL of -1.45 (95%CI -1.63 to -1.27).

Several routinely collected variables were found to be predictive of future CVD. Framingham CPM demonstrated favourable prognostic model characteristics in external validation. The qualitative research led to a co-produced care pathway based on five themes including loci of clinical responsibility, poor communication, avoidance of pandemic hospitals, focused EM care, and automation of preventative EM care.

## Conclusion

Calibration drift affected TMACS, and this should be a cautionary tale for other deployed CPMs. However dynamic updating did successfully counter this and could provide a sustainable solution if the digital infrastructure is available. Routinely collected data in the ED is predictive of long term CVD, and a care pathway appears possible. A feasibility randomised control trial should be conducted to further assess this intervention.

## Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

## Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and they have given the University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made. Presentation of Theses Policy You are required to submit your thesis electronically Page 10 of 24

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=2442 0), in any relevant Thesis restriction declarations deposited in the University Library, the University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in the University's policy on Presentation of Theses.

# Dedication

This research would not have been possible without my funders the National Institute for Health Research, Manchester University NHS Foundation Trust, and the Royal College of Emergency Medicine. Thank you to them and the opportunity they granted me.

Thank you to all my supervisors Professor Anthony Heagerty, Professor Evan Kontopantelis, Dr Glen Martin and Dr Brian McMillan for their guidance and expertise. Professor Richard Body as my primary supervisor and mentor thank you for your faith, motivation and many wise words.

I am deeply grateful to the participants in the research and the time they gave for this project. Thank you also to the patient representatives who aided the project. I am also grateful to my co-authors for their time and efforts with the publications.

Most of all thank you to my wife for her support and sacrifices to help me complete this work.

# The Author

## Qualifications

Bachelor of Medicine, Bachelor of Surgery – Hull York Medical School 2015

Master of Research in medical sciences – Hull York Medical School 2014

## Grants and Awards

Clinical Doctoral Research Fellowship, National Institute for Health Research

Data-Driven Impact grant, Manchester University NHS Foundation Trust

Research Impact Scholarship, University of Manchester

Annual Project grant, Royal College of Emergency Medicine

President's Doctoral Scholar Award 2018-2021, University of Manchester

## Achievements from this thesis

### Publications

Reynard, C., Jenkins, D., Martin, G.P., Kontopantelis, E. and Body, R., 2022. Is your clinical prediction model past its sell by date?. *Emergency Medicine Journal*, 39(12), pp.956-958.

Reynard, C., McMillan, B., Jafar, A., Heagerty, A., Martin, G.P., Kontopantelis, E. and Body, R., 2022. Long-term cardiovascular risk prediction in the emergency department: a mixed-methods study protocol. *BMJ open*, 12(4), p.e054311.

Reynard, C., van den Berg, P., Oliver, G., Naguib, M.P., Sammut-Powell, C., McMillan, B., Heagerty, A. and Body, R., 2022. The prognostic value of emergency department measured hypertension: a systematic review and meta-analysis. *Academic Emergency Medicine*, 29(3), pp.344-353.

Reynard, C., Martin, G.P., Kontopantelis, E., Jenkins, D.A., Heagerty, A., McMillan, B., Jafar, A., Garlapati, R. and Body, R., 2021. Advanced cardiovascular risk prediction in the emergency department: updating a clinical prediction model–a large database study protocol. *Diagnostic and Prognostic Research*, 5, pp.1-7.

### Conferences

2022 , Community for analytical measurement science annual conference, dynamic updating models

2019, Oral Presentation at the Royal College of Emergency Medicine annual scientific conference, Hypertension in the Emergency Department

# Chapter 1 Introduction

## 1.1 Preface

Patients attending the Emergency Department (ED) with chest pain account for around 6% of all visits (1). However even with recent advances in care pathways and national guidelines prior to the pandemic it was the most common indication for emergency admission to hospital in England and Wales (1–3). Research has demonstrated that less than 15% of patients with suspected acute myocardial infarction (AMI) transpire to have it (4). This highlights the need for better diagnostics, and clinical prediction models (CPM) have great potential to address this need by optimising the use of diagnostic information from different modalities. CPMs can improve the care of patients and the use of finite resources in the United Kingdom's (UK) National Health Service (NHS). Several CPMs for AMI already exist including the Troponin Only Manchester Acute Coronary Syndrome algorithm (TMACS), and the HEART and Emergency Department Assessment of Chest Pain Score (EDACS) (5–7). While they are widely used in practice, there is a risk that their diagnostic performance may not be maintained as time passes. Calibration drift can impact model performance over time due to changing patient demographics, and clinical care pathways (8). There is a need to understand whether there is calibration drift, and whether machine learning approaches could be used to counter it by refining and improving the existing CPMs as more data becomes available. This can help ensure the resources used to derive, validate and deploy CPMs are not squandered and allowing patients to continue benefiting from them.

Not only can CPMs be updated, but as new clinical needs develop their use cases can be expanded, such as for long term cardiovascular disease (CVD) outcomes. In a pilot randomised controlled trial of the original version of the TMACS decision aid, participants who received care guided by the decision aid were asked to complete patient satisfaction questionnaires. The findings identified that patients felt our current approach was suboptimal with regard to advising on "ways to avoid illness

and stay healthy" (9). This suggests that patients want their physicians to not just identify their current diagnosis, but also mitigate risk of ongoing health issues after leaving the ED.

ED attendees tend to engage less frequently with primary care (10,11). Therefore, they may be missing opportunities for primary and secondary prevention, meaning that their ED attendance is a valuable opportunity to identify those at high risk of incident CVD and intervene. There is scope for acute care to leverage it's high footfall and predict long term CVD outcomes.

In this PhD I aimed to improve CVD risk prediction in the acute care setting. I sought to do this by building an acute care pathway for long term CVD. I also aimed to evaluate the TMACS model performance and improve it by updating the algorithm.

In this chapter I provide background information on cardiovascular disease, evolving cardiovascular disease diagnostics (including troponin and clinical prediction models), and novel research methodologies

## 1.2 Cardiovascular disease

CVD can present to acute care in a multitude of diverse ways, including a facial droop or dysarthria (distorted articulation of speech) indicative of a stroke, a severe headache from a hypertensive crisis, acute tearing abdominal pain indicative of a dissecting aortic aneurysm, or central crushing chest pain radiating to the arms demonstrating an AMI (noting that radiation to either or both arms can occur in AMI) (12).

CVD encompasses many pathologies, ranging from coronary artery disease to ischaemic cerebral vascular accidents (Table 1.1). The underlying pathophysiology involves either the heart or body-wide vasculature. This process can be driven by atherosclerosis, where arteries undergo subendothelial remodelling (atherogenesis). Subsequent inflammation leads to plaque formation consisting of lipids, elastin and collagen. The net effect is narrowing of the vessels lumen (stenosis), this reduces the blood supply to the end organ and creates turbulent flow increasing the likelihood of clot formation.

Atherogenesis takes place over several years however acute events can occur. Atherosclerotic plaques can rapidly rupture and form a thrombus causing complete occlusion or a critical stenosis, such as is seen in the coronary arteries in an AMI.

There are known risk factors for cardiovascular disease (Table 1.2), some of which can be mitigated by an intervention. Screening for these factors has been historically conducted by primary care in high-risk individuals (e.g., over forty health check). If a risk factor is positively identified then an intervention is recommended, either by lifestyle modification or medication. It is estimated that primary prevention to reduce blood pressure can reduce CVD by 16% (target <140/90 mmHg or <130/80 mmHg with diabetes), or treatment of dyslipidaemia reduces CVD by 37% (13–15). When treated in combination the risk reduction can rise to greater than 50% (13). However, because important risk factors for CVD such as hypertension and dyslipidaemia are usually asymptomatic, identification of at-risk individuals can be challenging unless they regularly attend primary care appointments.

There is an enormous human and economic cost associated with cardiovascular disease. In terms of mortality, CVD accounts for more than one quarter of all deaths in the UK (16). In monetary terms, CVD has a total cost to the economy of £19 billion a year (17). In terms of morbidity, there are approximately 7 million people living with CVD in the UK (16). Locally the problem is amplified, national data from 2018-2020 show that Greater Manchester had the highest age standardised death rate from coronary heart disease (CHD) at 161.8 per 100,000 population, significantly higher than the national average of 64.1 per 100,000 (18).

| Site | Pathology |
|---|---|
| Central Nervous System | Vascular Dementia<br>Ischaemic Stroke<br>Haemorrhagic Stroke<br>Cerebral Aneurysm<br>Subarachnoid haemorrhage |
| Heart | Coronary Artery Disease<br>Dysrhythmia<br>Ischaemic Heart Disease<br>Heart Failure<br>Valvular Disease<br>Acute Myocardial Infarction |
| Vascular System | Peripheral artery disease<br>Peripheral venous insufficiency<br>Aortic aneurysm<br>Hypertension |

*Table 1.1- Summary of the composite pathologies of cardiovascular disease (CVD).*

| | |
|---|---|
| Non-modifiable | Increasing age (19) |
| | Male (19) |
| | 1st degree relative with first CVD event <60 (19) |
| | Polygenic risk (20) |
| | Chronic Kidney Disease (21) |
| | Low socio-economic group (19) |
| | Previous CVD (22) |
| | Ethnicity (23) |
| Modifiable | Obesity (19) |
| | Hypertension (24) |
| | Diabetes Mellitus (24) |
| | Dyslipidaemia (19) |
| | Physical inactivity (25) |
| | Smoker (19) |

*Table 1.2 Risk factors for cardiovascular disease*

### 1.2.1 Acute Myocardial Infarction

Ischaemic heart disease incorporates both acute coronary syndromes (ACS), including AMI and stable angina. It is defined as a mismatch between oxygen supply and demand to the myocardium, the majority of which is a result of atherosclerotic stenosis of the coronary arteries. Stable angina (or stable coronary artery disease) is defined by the European Society of Cardiology (ESC) as "*reversible myocardial demand/supply mismatch, related to ischaemia or hypoxia, which are usually inducible by exercise"* and is associated with stable plaques and stable symptoms (26). In contrast ACS are inherently associated with unstable plaques and unstable symptoms and the demand/supply mismatch often causes irreversible injury to the myocardium.

ACS incorporates unstable angina, non-ST elevation myocardial infarction and ST elevation myocardial infarction (STEMI). Cardiac troponin, a biomarker of myocardial injury, is a cornerstone of this diagnosis. The fourth universal definition for type one myocardial infarction requires a rise or fall of cardiac troponin values with at least one above the 99$^{th}$ percentile upper reference limit and another finding (27). The other finding can include either: symptoms in keeping with AMI, a newly ischaemic electrocardiogram (ECG), imaging evidence consistent with new cardiac ischaemia, and evidence of a coronary thrombus via angiography (27). The diagnosis of STEMI is made using the electrocardiogram (ECG). Diagnostic ST elevation on the ECG is defined by the European Society of Cardiology (ESC) as elevation in two contiguous leads ≥to 2.5mm in men under 40 years old, ≥2mm in men over 40 years old and ≥1.5mm in women (28). Either ST or non-ST myocardial infarction can be further specified into five types of myocardial infarction. Differentiating between type 1 and 2 MI is important and these again have been defined by ESC (27); type 1 is an acute occlusion of the coronary artery due to an atherosclerotic plaque rupture with thrombus. Type 2 is delineated as an acute non-cardiac stressor (e.g., massive haemorrhage) that results in a reduced oxygen supply and/ or increased demand resulting in myocardial infarction. Type 3 MI is defined as a sudden unexpected

cardiac death often with symptoms suggestive of myocardial infarction. Type 4 is a MI associated with percutaneous coronary intervention and type 5 is an MI associated with cardiac surgery.

## 1.3 Evolving diagnostics

In recent years there have been tremendous advances in diagnostic technology that can allow the diagnosis of AMI to be 'ruled out' for many patients within a few hours of arriving in the ED. While many of the diagnostic algorithms that have been developed have excellent sensitivity and negative predictive value, none have perfect accuracy (5,7,29–31). There is therefore always a small risk that patients who are discharged from the ED will develop a major adverse cardiac event (MACE) in the near future.

In an international survey, only 40% of emergency physicians would accept a 1% probability that a patient they discharge will develop a MACE within 30 days (9). This may reduce adherence to validated 'rule out' algorithms. Indeed, in a recent large cluster randomized controlled trial, patients who were randomized to receive care guided by the HEART score were no more likely to be discharged early than patients who received standard care (29).

### 1.3.1 High sensitivity cardiac troponin assay (hs-cTn)

Troponin is 52 kilo-dalton protein complex that is composed of troponin C, I and T protein sub-units (32). Troponin I was identified as a useful marker for cardiac necrosis and myocardial infarction in 1987 (33). Later this was extended to troponin T leading to their successfully incorporation into clinical care pathways with contemporary cardiac troponin assays (34).

To describe these assays several analytical metrics are used. The limit of blank (LoB) is the highest apparent reading from an assay when a blank sample containing no target analyte is present (35). The limit of detection (LoD) is the lowest concentration of the target analyte detectable when it can be consistently differentiated from the LoB (35). The 99$^{th}$ centile or upper reference limit (URL) is the 99$^{th}$ centile cTn result in a healthy population, recently this has been amended to include sex specific URLs (36).

From the initial assays a new generation of assay was developed. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) defines a hs-cTn assay as one that can detect cTn below the 99th centile and above the LoD in at least 50% of healthy subjects (i.e. detect a normal cTn level) (37). It must also be sufficiently precise to have a co-efficient of variation of less than 10% meaning that two samples tested twice should have two results within 10% of each other (37).

Hs-cTn is currently an integral part of the diagnosis of AMI, however alternate strategies have been proposed that can enable an early rule out from the hs-cTn alone (27,38). The UK's National Institute for Health and Care Excellence recommends two early rule out strategies for NSTEMI using only hs-cTn. NICE proposes using a sample taken at presentation with a hs-cTn threshold at or near the LoD (38). Alternatively they recommend a serial sampling strategy with a sample taken at presentation and the second sample taken up to 3 hours later with a threshold of the URL or around the LoD (38).

The LoD strategy has been investigated previously and demonstrated to have a sensitivity of 97.4% (95% CI 94.9 – 98.7) and a specificity of 42.4% (95% CI 31.2 – 54.5) allowing the rapid rule out of 33.9% of patients (39). When this strategy was examined in a randomised control trial the length of stay was not found to be statistically significantly different from pre-existing care pathways (30).

Another trial examining the LoD strategy was the High-Sensitivity Troponin in the Evaluation of Patients With Acute Coronary Syndrome algorithm trial (High-STEACs) (31,40). High-STEACs used a threshold of <5ng/l on a 0hr sample and the demonstrated a high sensitivity of 97.7% (95% CI 97.3 – 98.1%) with a specificity of 64.0% (95% CI 63.6 -64.5) (31,40). This allowed for a proportion ruled out of 58.4% and demonstrated a modest reduction in length of stay of 3.3 hours) (31,40). The ESC have also put forward thresholds near the LoD which are specific to each manufacturer, their strategy was found to have a sensitivity of 98.4% (95% CI 95.1 - 99.5), specificity of 91.2% (95% CI 86.0 – 94.6) and a proportion ruled out of 50-55% (41).

High-STEACs, like ESC and the LoD strategy (LoD), do not provide a risk prediction that is continuous (0-100%) which can limit their clinical utility (39–41).

### 1.3.2 Clinical Prediction Models

CPMs are used in a wide variety of healthcare settings from hospital admissions (42), long term CVD outcomes (19) to risk of acute kidney injury (43). They are considered diagnostic if the disease status of the patient at the time the predictors are measured is of interest. An example of a diagnostic model would be a CPM used to detect subarachnoid haemorrhages in patients with undifferentiated headaches. This would be considered a diagnostic CPM as the target condition exists when the predictors would be measured. In contrast a prognostic model attempts to predict an outcome which is not present at the time of predictor measurement, an example here is long term CVD models. Such models measure the predictors long before the outcome of interest is a possibility. Diagnostic and prognostic models are prone to biases, diagnostic models can suffer from partial verification where the outcome is not fully verified (44). Whereas prognostic models must be wary of censoring where an outcome is yet to occur by the end of time horizon window for measurement (44).

To improve the quality of CPMs reporting guidelines have been developed called the “Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis” (TRIPOD) (45). In it they describe what parameters should be reported in order to allow the reader to make an informed judgement on the CPM presented. This includes reporting measures of model performance for both discrimination and calibration. Discrimination is the ability of a model to differentiate a positive outcome from a negative outcome, this can be commonly measured using the c-statistic (also called area under the receiver operator curve) (46). The calibration of a CPM is how close the predicted risk is to the actual or observed risk (46). If they are similar the model is considered well calibrated.

The TRIPOD guidelines also comment on the importance of model validation, this is where the performance of the model is assessed separately from the derivation. Internal validation uses the

same data as the model was derived in, but external validation uses new external data to that of the derivation. The latter is considered more robust and can reveal overly optimistic derivation or internal validation performance estimates (47). As such it is an important step CPMs prior to clinical use.

### 1.3.3 Troponin Only Manchester Acute Coronary Syndrome algorithm (TMACS)

The Troponin-only Manchester Acute Coronary Syndromes (TMACS) CPM is in active clinical use to risk stratify patients presenting to the ED with suspected ACS (5). The CPM was derived by logistic regression and uses seven variables including details of a patient's history, physical examination, initial ECG and the hs-cTn concentration measured at the time of arrival in the ED. These variables are inputted into the CPM which predicts the probability that a patient has ACS. Of note, none of the five 'Framingham' risk factors for CVD (diabetes, hypertension, hyperlipidaemia, tobacco smoking and a family history of premature cardiovascular disease) are included as predictors in the TMACS model because they were not found to be independent predictors of ACS. Their poor predictive value for ACS in acute settings has been previously shown (48). This may be because patients and clinicians have a heightened awareness of the cardiovascular risk associated with these risk factors and therefore have a lower threshold for attending the ED or undertaking investigations. Further, the most dangerous modifiable risk factors are arguably those that were previously unrecognised and therefore untreated.

Based on the calculated probability of ACS, the TMACS algorithm stratifies patients into four risk groups. 'Very low risk' patients (40% of the total) are considered to have ACS 'ruled out' and are eligible for immediate discharge. 'Low risk' patients require a second hc-cTn test but could be managed in an ambulatory care unit. 'Moderate risk' patients are likely to require further investigation (including serial hs-cTn testing and cardiac imaging). 'High risk' patients can be considered to have ACS 'ruled in' and can be immediately referred to a cardiologist for specialist treatment.

TMACS has been extensively externally validated including 1,459 patients in the UK (5), 1,244 patients from Australasia (49), Thailand (50) and Norway (51). Each validation demonstrated clinically desirable model performance. A pilot randomized controlled trial of an early version of TMACS demonstrated that a greater number of patients were discharged within four hours of arrival compared to normal care (52). TMACS was implemented into clinical practice at Manchester Royal Infirmary in June 2016 and was subsequently chosen as an exemplar project by Health Innovation Manchester to be rolled out regionally at 12 hospitals.

Several other CPMs have been derived and validated to aid in the early diagnosis of ACS in ED patients with suspected cardiac chest pain (5,7,29–31). These CPMs all include hs-cTn as either a categorised or continuous variable, however the other predictor variables vary (Figure 1.1). The common goal of these CPM's is to rule out ACS in the ED rapidly and safely with one hs-Tn result. Each has been synthesised with different methods leading to heterogenous clinical input variables and differing statistical strengths (Figure 1.1). Three CPMs were compared in the same cohort by Body et al allowing for direct comparison of performance (Table 1.3) (53). TMACS had a sensitivity of 99.2% (95% CI 95.7-100.0) and a specificity of 53.3 (95% CI 49.9 56.7) ruling out 46.5% of patients (53). Whilst EDACS ruled out 48.3% of patients it had a lower sensitivity of 96.2% (95% CI 92.2 - 99.4) (53). HEART had a lower sensitivity and proportion ruled out of 91.8% (95% CI 85.0 -96.2) and 34.9% respectively (53).

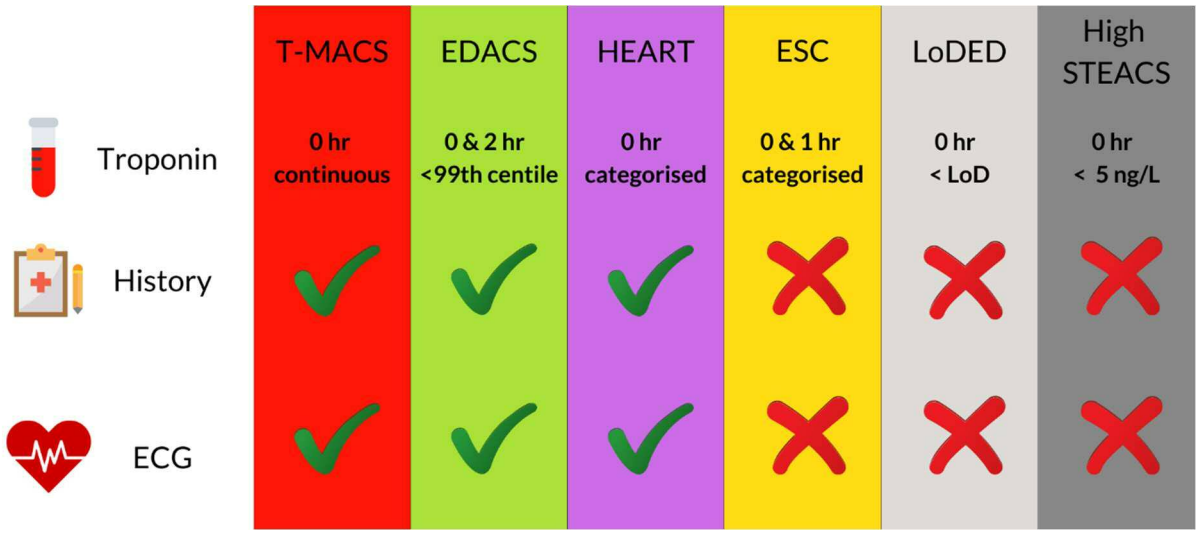



Figure 1.1 Variables included in clinical prediction models for acute coronary syndromes. EDACS – Emergency Department Assessment of Chest Pain Score, TMACS – Troponin-only Manchester Acute Coronary Syndrome rule out strategy, ESC – European society of cardiology -algorithm, LoD – limit of detection algorithm, High-STEACs - High-Sensitivity Troponin in the Evaluation of Patients With Acute Coronary Syndrome algorithm (5,7,29–31). ESC, LoDED and High STEACS included ECG and symptomatic criteria for entry into their care pathway but did not include it within the algorithm itself.

| CPM | Sensitivity (95% CI) | Specificity (95% CI) | Proportion ruled out |
|---|---|---|---|
| EDACS | 96.2 (92.2 -99.4) | 55.1 (51.4 -58.7) | 48.3% |
| HEART | 91.8 (85.0 -96.2) | 38.6 (35.1 -42.2) | 34.9% |
| TMACS | 99.2 (95.7 -100) | 53.3 (49.9 -56.7) | 46.5% |

*Table 1.3 Characteristics of clinical prediction models for acute myocardial infarction when compared in the same cohort from Body et al (53). CPM – clinical prediction model, EDACS – Emergency Department Assessment of Chest Pain Score, TMACS – Troponin-only Manchester Acute Coronary Syndrome rule out strategy.*

### 1.3.4 Predicting Long term Cardiovascular Disease

There is a lineage of successfully implemented CVD risk prediction models for primary care. An early achievement was the Framingham score based on a longitudinal cohort study of 5,127 patients from the United States of America with data collection starting in 1948 (54). In 1961, Kannel et al identified that the incidence of CHD was higher in patients with certain characteristics (55). In the subgroup who were male, 40-59 years of age, and had a total serum cholesterol of greater than 245 mg/100ml the incidence rate of CHD was 120.3 per 1,000 population versus the expected rate of 71.8 per 1,000 (55). Similarly for definite hypertension in men between 40-59 years of age the incidence of CHD was 123.7 per 1,000 of the population when it was expected to be 77.8 per 1,000 (55). In participants who were male between 40-59 and had two or more risk factors (high total serum cholesterol, definite hypertension and evidence of left ventricular hypertrophy on the ECG) the incidence was 226.4 per 1,000 of the population when the expected incidence was only 83.0 per 1,000 (55). This started the journey for long term cardiovascular risk prediction and the importance of multiple predictors. The Framingham project published several risk scores including one for coronary heart disease which gained widespread adoption, and was used within clinical care (56). The risk factors within the Framingham score were age, sex, smoking status, total cholesterol, high density lipoprotein, systolic blood pressure and diabetes mellitus (56). The resultant model had favourable performance with a c statistic of 0.73-0.77 for predicting the outcome coronary artery disease (comprised of angina pectoris, AMI, coronary insufficiency and cardiovascular related death) (56).

Most recently a series of CPMs called 'QRisk' for predicting the development of CVD have been derived using the electronic records of 1.28 million patients in the UK. The original QRisk model was derived in 2007, QRISK-2® in 2008 and QRISK-3® in 2017 (57–59). Each iteration brought improvements in model characteristics, with respective c-statistics of 0.788, 0.817/0.792 (male/female), and 0.880/0.858 (male/female) (58,59). With these improvements came new complexity in the form of variables that included severe mental illness, atrial fibrillation, chronic

kidney disease, family history of AMI, migraines and more. This lineage of CPMs were adopted into NICE guidelines for the assessment of CVD risk, they formed part of a protocol that included pharmaceutical interventions (60). If the 10-year predicted risk of CVD exceeds 10%, the National Institute for Health and Care Excellence (NICE) recommends that statin therapy should be considered (61). A range of other measures (advice on smoking cessation, weight loss, diet, exercise and review of comorbidities) are also recommended.

This tool could also be used in the ED because most of the data required to calculate the risk of incident CVD are already routinely collected (59). Applying a prognostic tool such as QRISK-3 in the ED may help to identify patients who are at high risk of CVD who may otherwise have been unidentified. By identifying these patients in the ED, we may not only enable patients to be better informed, but we could prevent more incident CVD by detecting at-risk patients who may not have been seen in primary care.

This may present an excellent opportunity for emergency physicians to identify a large population of patients at high risk of future CVD, at a time when they are likely to be particularly receptive.

### 1.3.5 Calibration Drift

Calibration drift is a reported phenomenon where a model's performance in the domain of calibration diminishes over time (8). It is important in CPMs deployed in clinical practice as a mis-calibrated model leads to inaccurate risk prediction and can enable wrong decisions to be made. This effect could be exaggerated in a clinical pathway with risk thresholds where the incorrect intervention is selected as the patient is incorrectly designated to a risk group.

Calibration drift has been observed by Hickey et al when analysing a CPM predicting mortality after cardiothoracic surgery (EUROSCORE) (62). They demonstrated a divergence between predicted and observed risk over time, a drift in calibration. The earliest observed mortality was 4.1% and the EUROSCORE predicted mortality was 5.6%, this gave an observed/expected mortality ratio of 0.73 (O/E ratio) (62). In a perfectly calibrated model, the O/E ratio would be 1, in a CPM that was

overestimating risk the O/E ratio would be less than 1 and if underestimating greater than 1. When a decade had passed in the data the analysis was repeated, and the observed mortality was observed to have dropped to 2.8% however the predicted mortality had risen to 7.6% causing the O/E ratio to drift to 0.37 (62). The main driver of this drift was thought to be an improvement in surgical techniques and a more elderly population.

An analysis of CPMs for the outcome acute kidney injury identified a drift with the O/E ratio moving from 1.0 to 0.8. (63). Similar again for CPMs predicting hospital mortality drifted as well from 1.0 to 0.75. (64). These drifts were in part driven by decreasing incidence for the CPM's outcome, and they demonstrate the breadth of conditions.

The causes of calibration drift are specific to each CPM and the population it is deployed within. The key drivers that can effect most CPMs are the older but healthier populations and new diagnostic and therapeutic interventions (62–64). The common cause of these drivers is the static nature of CPMs. Whilst the CPM, its variables, their coefficients and intercepts remain static the population it is deployed in changes (65). This is exacerbated by the time-lag between model derivation and implementation with models requiring observation cohort studies, derivation and validation all prior to implementation.

## 1.4 Novel research methodologies

### 1.4.1 Updating Clinical Prediction Models

Statistical and machine learning methods have previously been proposed to update models. Some methods can overcome the time delay by enabling automation where rederivation and validation can take place without active input. This dynamic process can allow the CPM to adapt to changes in the population's demographics and updates to clinical practice. This offers improvements over historic methods of updating algorithms; it can be done quickly, continually and more accurately. Several different methods for updating clinical prediction models have been suggested (66); regression coefficient updating, meta-modelling and Bayesian dynamic updating. Regression

coefficient updating is a targeted method where only the individual coefficients previously in the CPM are adjusted from a static analysis. Meta-modelling is the amalgamation of several clinical prediction models into one again using one static singular further re-derivation (67–69). Bayesian dynamic enables updating and derivation after every patient encounter, once the correct implementation has been achieved it can theoretically continuously re-derive with new data (70,71). Siregar et al reviewed different updating methods and they found regression coefficient updating and dynamic updating to have a similar improvement on CPMs (72). Their analysis of dynamic updating suggested for smaller subgroups a Bayesian approach may yield greater improvements in accuracy (72). Strobl et al updated a prostate cancer CPM, and they again found that different updating methods had similar benefits for model performance (73). The only exception was random forest regression which was significantly worse.

### 1.4.2 Co-production with patients

Batalden comments in his essay that the patient health care relationship is not a simple consumer provider relationship (74,75). It is a dynamic interaction where the consumer can exert an influence upon the provider. This approach has been championed by service improvement initiatives, where the identification of the area for improvement (the idea genesis) is conducted by researchers and patients in partnership. Batalden proposes that through co-production this relationship can continue beyond idea genesis and into improvement and success measurement. Various different methodologies for this process have been proposed; most widely adopted is the experience-based co-design (EBCD).

EBCD is a design used by several different initiatives, it encompasses four stages; Capture, Understand, Improve and Measure (76). The capture phase is where ideas are generated and prominent issues for resolution are condensed, then in understand the selected issues are extensively mapped. In the improve phase solutions are conceived to the issues in the first phases then in the measure phase the implementation of the proposed solution is checked for improvement. This process is synergistic with national guidance for developing complex

interventions from the Medical Research Council and the National Institute for Health Research (77). This framework describes four activities that are conducted to successfully develop and deploy an intervention, these include: develop the intervention, assess feasibility, implement and evaluate (77). They are clustered around core elements or principles which include: context, develop/refine programme theory, engage stakeholders, refine intervention and more (77). The development of any new care pathway could adopt the principles of both EBCD and complex intervention guidance.

## 1.5 The Emergency Department Population

There were 23.4 million patients presentations to the UK's emergency departments (ED) in 2016, and this has been increasing by 10% each year (1). When these patient interactions represent the majority of a patients contact with the NHS, we cannot ignore their long-term health any longer.

The increase in ED attendance represents a paradigm shift in the way patients are accessing health care, this has caused many previously noted problems however it also presents opportunity. Patients who attend EDs have often tried and failed to access primary care (78,79), therefore those that do attend EDs may be a portion of society underserved by primary care. Furthermore, in the drive for NHS system wide efficiency, we must maximise the use of each patient interaction with the health service. Considering the large amount of data already collected and stored routinely by EDs around the UK, this presents a fantastic opportunity to predict and intervene on cardiovascular risk. Preventative medicine is not a new concept to emergency medicine, and it has been researched and implemented successfully before (80). Patients demand and expect clinical staff in the acute care setting to have tools to inform them of their long term cardiovascular risk (9).

While in the ED, all patients with suspected ACS will have vital signs recorded. However, these data are not currently used to identify patients at risk of CVD, which represents an important missed opportunity. Previous research has demonstrated that 76.4% of patients with hypertension in the ED had persistently elevated blood pressure in the community setting (81). Contrary to widespread belief,

hypertension in the ED cannot be wholly attributed to pain or anxiety. Recent studies have shown that hypertension measured in the emergency department is predictive of 10-year major adverse cardiovascular outcomes (82). Not only does predictive data exist in the ED but patients are also potentially more receptive to long-term CVD advice during an ED attendance (83). This interest in long term CVD health has been demonstrated in a survey and also a qualitative study (52,84). Ferry et al interviewed 50 patients with suspected ACS and discovered an emergent theme of "Approaches to future health", where patients were considering their long term CVD risk and were more open to advice about disease prevention (84).

In an ED setting, CVD risk prediction and modification has been trialled in the UK and the USA and has mainly been conducted in observational units. Katz et al used health belief model constructs during and after a presentation to ED with cardiac chest pain (85). They measured readiness to change and behaviours around modifiable cardiovascular risk factors and demonstrated improved behaviour.

Katz et al also conducted a randomised control trial of a lifestyle counselling intervention (one hour face to face and two thirty minute telephone follow ups), which also demonstrated an improvement in all lifestyle risk factors (86). Despite this, challenges remain. Collins et al conducted an RCT of hypertension follow up in the UK ED setting. The patient population included ambulatory patients over the age of 35 years old who had their blood pressure taken as part of their routine care (87). The intervention group were given a counselling session, information booklet and a letter sent to their GP, whilst the control group were not (87). The study demonstrated no difference in follow up via primary care, with no difference in lifestyle advice and no new anti-hypertensives prescribed across intervention or control group. This result was replicated in the United States where Julliard et al conducted a retrospective database review (88); they found that 32% of hypertensive patients in the ED were referred to primary care but worse still only 8.6% of hypertensive patients attended an appointment. Perhaps the most concerning aspect is that none of the referrals were primarily for hypertension follow up, they were all for other concurrent medical conditions. Schwalm et al

examined strategies to prevent CVD and highlighted the role of non-physicians in the identification and treatment of patients. It is feasible that in the United Kingdom this could involve health-care assistants, specialist nurses, or pharmacists.

In the literature the pragmatic system factors have been explored; these are summarised in Table 1.4. These range from concerns over resources to opportunities of the teachable moment (82,83).

| **Concerns** | Reference |
|---|---|
| Rapid lowering of BP may be harmful, and therefore the initiation of treatment in the ED may not be advisable | Julliard (88) |
| There is a fundamental question as to whether it is the ED's role to conduct preventative medicine. | Lee et al (89) |
| There is concern that the ED does not have the resources to conduct any additional activities. | Lee et al (89) |
| Any CVD risk prediction would incorporate a blood pressure reading, and there is a belief that ED hypertension is related to anxiety | Tanabe et al (83) |
| For any ED CVD risk modification, the long term follow-up of patients would be essential. A barrier to this continuity is that patients would have to independently research a GP and book/register | Collins et al (87) |
| There is evidence that patients who consulted in the ED for CVD risk medication unfortunately have poor compliance with follow-up instructions. | Collins et al (87) |
| **Opportunities** | |
| There is a subset of the population who receive most of their care from the ED | Tanabe et al (83) |
| When a patient presents to the ED with chest pain, it is thought they are then more receptive to CVD advice | Tanabe et al (83) |
| As a patient becomes older their compliance with CVD risk advice increases. | Murray et al (90) |
| The opportunity for non-clinical advisors has been explored successfully, making CVD advice more cost-effective. | Schwalm et al (91) |

*Table 1.4 Pragmatic system factors for a long-term cardiovascular disease (CVD) prevention pathway in the emergency department (ED).*

## 1.6 Objectives of the thesis

The two primary objectives of this thesis are listed below with secondary objectives listed in bullet points underneath.

1. To develop a care pathway for the prediction of long-term CVD in the acute care setting.

    - To identify the feasibility of measuring and using common risk factors for CVD in the Emergency Department setting.
    - External validation of prognostic models for long term CVD outcomes.
    - To create a co-produced care pathway for long term CVD.

2. To validate, assess and update the TMACS CPM.

    - To evaluate the diagnostic accuracy of the TMACS CPM for acute myocardial infarction in a real-world cohort of patients.
    - To assess the variation in performance of TMACS over time.
    - Identify the optimal ML methods for ongoing updating of TMACS.

## 1.7 Thesis Structure

This thesis is divided into seven chapters (including this introduction chapter). Chapters two through four describe work to develop a care pathway for the prediction of long-term CVD in the acute care setting. Chapter two is a systematic review and meta-analysis of prognostic factors measured in the ED and their feasibility for use in long term CVD prediction. Chapter three is a retrospective cohort study where real-world evidence from the ED is linked with NHS Digital data to assess the prognostic characteristics of routinely measured emergency department data. Chapter four describes a qualitative study using co-production methods to produce a long-term CVD care pathway for the acute care setting. Chapter five and six include research to achieve the second objective of the thesis validating, assessing and updating TMACS. Chapter five is a validation of TMACS in a large real world evidence cohort, where it is assessed for calibration drift. Chapter six then seeks to update and improve TMACS with a different method and identify the optimal approach. Chapter seven is the final chapter where the thesis is summarised and overall findings are concluded.

# Chapter 2 A systematic review of risk factors for long term cardiovascular disease in the acute care setting

## 2.1 Background

It is well-established that a number of factors significantly increase the risk of future cardiovascular disease (CVD). That risk can be mitigated through medical intervention (14). The 23.4 million patients who attend Emergency Departments (EDs) each year in England and Wales are those who attend ED are likely to have tried and failed to access primary care, reducing important opportunities for primary prevention (1,78,79). Prognostic factors and CPMs have been extensively studied for primary care populations predicting the outcome of long term cardiovascular disease but there is a paucity of research for the acute care setting (54,92).

I conducted four systematic reviews and meta-analyses, exploring the predictive value of four established prognostic factors in the ED population and the feasibility of their use for primary prevention in this setting. These included hypertension (HTN), chronic kidney disease (CKD), type two diabetes mellitus (T2DM), and dyslipidaemia (19). If I found them to be of prognostic value in the ED setting, I then planned to carry them forward in my analysis of long-term CVD prognostic risk factors from EPR data.

## 2.2 Published work

I led the work on these systematic reviews however other researchers acted as second reviewers for abstracts, full texts and data extraction. Dr Patricia van den Berg (PB) was a second reviewer for the hypertension review, Dr Govind Oliver (GO) was a second reviewer for type two diabetes mellitus and dyslipidaemia, and Dr Mina Naguib (MN) was the second reviewer for chronic kidney disease. Professor Richard Body was the third adjudicator if there was disagreement. Despite my leadership of these systematic reviews, I have adopted the first-person plural in this text to reflect the wider team.

The hypertension systematic review and meta-analysis described in this chapter has been published (93).

Reynard, C., van den Berg, P., Oliver, G., Naguib, M.P., Sammut-Powell, C., McMillan, B., Heagerty, A. and Body, R., 2022. The prognostic value of emergency department measured hypertension: a systematic review and meta-analysis. *Academic Emergency Medicine*, *29*(3), pp.344-353.

## 2.3 Objectives

(a) To identify the prevalence, confirmation and follow up rate for common prognostic factors for CVD in the Emergency Department setting.

(b) To determine whether those risk factors retain their prognostic value for long term CVD outcomes over time periods greater than or equal to one year in the ED setting.

This systematic review and meta-analysis was conducted in accordance with Preferred reporting items for systematic reviews and meta-analysis guidelines and was also registered on PROSPERO [CRD42018110517] (94,95). The systematic review looks back twenty years rather than the ten originally specified. This decision was taken as there has been little change in the technology measuring ED blood pressure and as such we felt that the longer time period would continue to be relevant to the present day.

## 2.4 Methods

We targeted two populations for investigation: firstly the general ED population (‘regardless of presenting complaint), who could be screened for long-term CVD risk; secondly, patients attending the ED with suspected cardiac chest pain, where there may be opportunities to exploit a ‘teachable moment’ in a patient group that may be particularly receptive to a screening programme (85).

### 2.4.1 Search question

We used the population, factor, outcome framework (PFO), to build the research questions for four systematic reviews around different factors (96).

**Population** –general ED population or ED patients with chest pain of suspected cardiac aetiology.

**Factor**- hypertension (HTN), and dyslipidaemia, chronic kidney disease (CKD), and type 2 diabetes mellitus (T2DM).

**Outcome** – long term cardiovascular outcomes (CVD)

**Timing** -outcome occurring beyond or equal 12 months

**Setting** – Emergency Departments

### 2.4.2 Criteria for considering studies for this review

Study inclusion criteria:

- studies that match the PFO framework or may be applicable
- that have been conducted in the last 20 years
- randomised control trials
- prospective observational trials
- retrospective studies
- conference abstracts
- published and in press studies
- English language

#### *2.4.2.1 Targeted Population*

Two patient groups were of interest for this review: (a) undifferentiated patients attending the ED; and (b) patients presenting with chest pain suspicious of cardiac aetiology. These populations were selected are they were felt to be the ideal targets for future screening programmes. Undifferentiated chest pain suspicious of cardiac aetiology is defined as clinician suspicion or entry into a cardiac chest pain care pathway.

#### *2.4.2.2 Prognostic factors considered*

Blood pressure – measured in the emergency department. It was envisaged that the most widely available variable would be dichotomous around the common clinically significant threshold of 140/90 mmHg (97). It is recommended that patients have serial readings, involving periods of rest or ambulatory measurements. It was not envisaged that this would be universal in the Emergency Departments but we planned to attempt sensitivity analysis of the different measurement techniques if possible, to attempt to detect any effect it may have had on measurement (98).

Dyslipidaemia – this includes low density lipo-protein (LDL), and high density lipo-protein (HDL). This is measured from blood serum drawn during presentation from the patient. It is presented in mmol/l or mg/dl. There is international variability in the clinically significant threshold for dyslipidaemia and is frequently incorporated directly into cardiovascular prognostic models. Therefore, we planned to adopt a pragmatic approach on the thresholds, and measure for introduction of error via sensitivity analysis if feasible.

Chronic kidney disease (CKD) – this is routinely measured using the Modification of Diet in Renal Disease study equation (MDRD) (99), or the Chronic Kidney Disease Epidemiology collaboration (CKD-EPI) (100) which both use an equation based on age, sex, race and creatinine for estimation of estimated glomerular filtration rate (eGFR) in the units of ml/min/1.73m$^2$. The measurement is a continuous variable, however it is frequently interpreted and presented as an ordinal variable in categories of CKD stage 1-5. we sought to analyse it as a continuous variable where possible.

Type 2 diabetes mellitus (T2DM) – the definition of T2DM varies again internationally, the World Health Organisation recommends that serial samples are taken for asymptomatic patients and only in patients with symptoms of diabetes can it be diagnosed with a single blood test (101,102). This serum blood test can consist of HbA1c, fasting plasma glucose or non-fasting plasma glucose, they are considered suspicious of T2DM at levels of 48 mmol/mol (6.5%), 11 mmol/l, 7 mmol/l respectively. The American Diabetes Association recommends that these two test can be taken from the same sample (103). We will be pragmatic in our proposed meta-analysis and examine both definitions and any other encountered.

### 2.4.3 Outcome measurement

*Feasibility outcomes* – point prevalence, persistence and attendance for follow-up for the prognostic factors.

*Prognosis* - ability to predict cardiovascular outcomes including acute myocardial infarction, coronary revascularisation, coronary heart disease, angina, stroke, transient ischaemic attack, cerebrovascular event (ischaemic or haemorrhagic) and death (all cause and CVD). The timescale was greater than one year from the index event.

### 2.4.4 Information Sources

The following information sources were searched independently by two researchers.

- Electronic databases: Medline and Embase
- World Health Organisation – International Clinical Trials Registry Platform
- British Library Thesis
- Secondary Reference checking of papers selected for inclusion

Experts in the field: Professor Richard Body and Professor Tony Heagerty

### 2.4.5 Search Strategy – Medline & Embase

This was reviewed by two independent researchers and is reported in the appendix.

## 2.5 Data Collection

### 2.5.1 Selection of studies

The studies returned from the initial searches had their title and abstract reviewed by two independent reviewers (CR for all, PB for HTN, GO for T2DM and Lipids, MN for CKD). Selected papers then underwent full text review by again by two independent researchers, when disagreement occurred between researchers a third independent researcher reviewed the paper and agreement was met by consensus (RB).

### 2.5.2 Data Extraction and Management

The electronic search strategies were stored electronically on the Ovid interface, bespoke electronic study and data collection forms were created for the purpose of the systematic review. These forms enabled blinding of each researcher's review.

### 2.5.3 Risk of bias in individual studies

We used the QUIPS tool to analyse the risk of bias in each prognostic factor study (104), reporting by each domain: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis.

## 2.6 Data Synthesis

### 2.6.1 Meta-analysis plan

When the outcome data was suitable and comparable, we conducted a meta-analysis using a random effect model and exact likelihood method outlined by Hamza et al for binomial distributions (105). Statistics were appropriately transformed as needed. We used R, including packages metafor, meta and ggplot2 (106–109). We also used the $I^2$ statistic to assess heterogeneity, $I^2$ quantifies the impact intra-study heterogeneity on the pooled results (110).

### 2.6.2 Subgroup analysis and sensitivity analysis

We conducted a sub-group analysis of blood pressure based on initial repetition of BP measurement and acuity of presenting patients. Examining dyslipidaemia, we also conducted a sensitivity analysis based on the differing diagnostic thresholds, similar for T2DM.

## 2.7 Results for the prognostic factor hypertension

### 2.7.1 Study selection

We ran the electronic database search on the 27/10/20 our electronic database search we identified 1,072 papers, 725 remained after removal of duplicates. 53 were selected for full text review after 27 were excluded, 26 were included for review. Exclusions were based on outcome, population and duplicates (Figure 2.1). The authors of six of the studies identified were contacted for more information, including the 4 trial protocols identified and two studies which did not contain sufficient data to be included (111–115). Unfortunately, no replies were received in time for publication. No additional papers were identified from grey literature searches.

### 2.7.2 Study Characteristics

The characteristics of the 27 included studies can be found in Table 2.1 and Table 2.2, most studies reported outcomes relevant to the feasibility outcomes.

Eighteen were from the USA, two from the UK, one from Australia, Canada, South Korea, Israel, Iran, Sweden and Hong Kong. 11 studies were prospective and the remainder retrospective. The sample size varied from 701,952,422 to 88 (116,117), the implications of which are discussed later. Prevalence estimates were available from fifteen studies (81,87,88,116,118–128). Eight studies had outcomes for and the persistence of hypertension within the ED (118,120,123,127,129–132) and 12 outside of ED (81,83,87,88,118,125,127,132–136). 8 studies reported attendance rates for follow up appointments (87,88,118,126,129,132,134,135) and only three analysed the prognostic characteristics of ED hypertension as a risk factor for long-term cardiovascular disease (82,128,137). All studies included broad ED populations, none were specific to chest pain.

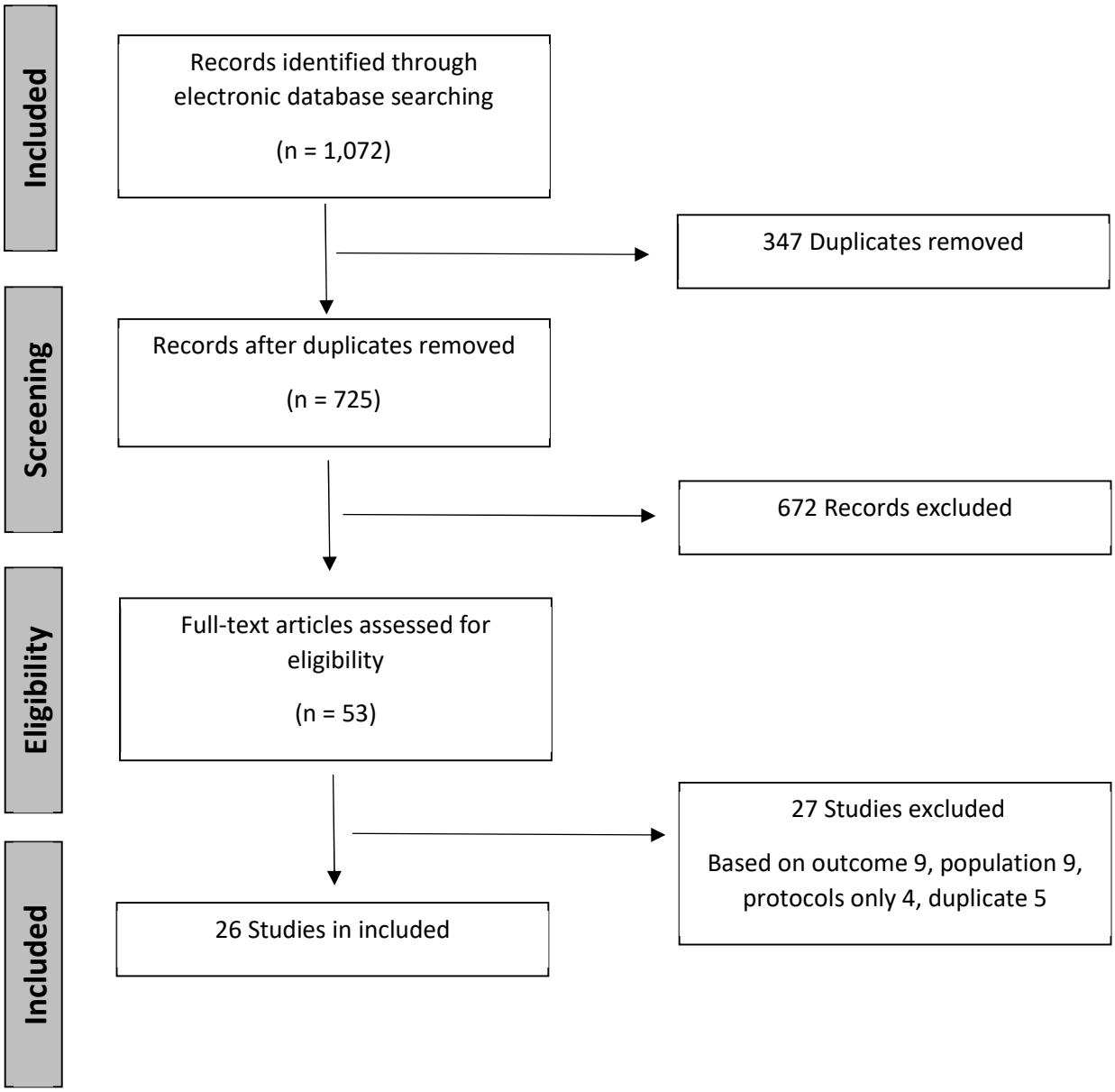



*Figure 2.1 Flowchart of blood pressure study selection process*

| Author, Country, Year | Study Design | N | Measurement of BP | Recruitment Period | Follow up | Outcome Measure | Risk of bias |
|---|---|---|---|---|---|---|---|
| Adhikari, USA, 2016 (131) | R | 179 | ED | 2011 | N/A | BP advice | M |
| Backer, USA, 2003 (132) | P | 407 | 3 x ED | 2000 | 6 months | F/U BP | M |
| Baumann, USA, 2007 (120) | P | 991 | 3 x Researcher | 2004 | N/A | Descriptive | M |
| Baumann, USA, 2009 (123) | R | 4245 | ED | 2005-2006 | N/A | Repeat ED BP | M |
| Cienki, USA, 2013 (130) | R | 1000 | ED | 2005-2010 | N/A | BP advice | M |
| Cline, USA, 2006 (134) | R | 1391 | EPR | 2003-2004 | 3 months | Secondary care F/U | M |
| Collins, UK, 2008 (87) | P | 765 | 2 x Researcher | 2005-2006 | 6 months | Primary care F/U | M |
| Dolatabadi, Iran, 2014 (125) | P | 346 | 2 x ED | 2009 - 2010 | 30 days | Prevalence | M |
| Fleming, UK, 2005 (81) | P | 991 | 2 x Researcher | 2004 | 30 days | Sustained HTN | M |
| Goldberg, USA, 2017 (127) | P | 151 | 3 x Researcher | 2014-2017 | 2 weeks | Sustained HTN | M |
| Julliard, USA, 2012 (88) | R | 662 | EPR | 2009 | 3 months | Primary care F/U | M |
| Karras, USA, 2005 (118) | P | 7238 | 1 X ED | 2002 | 30 days | Primary care F/U | M |
| Lee, South Korea, 2018 (82) | R | 262.927 | EPR | 2002-2013 | 10 years | MACE | M |

*Table 2.1 Characteristics of the included studies examining hypertension (part one of two), R = retrospective, P = prospective, EPR = electronic patient record, N/A = not applicable, F/U = follow up, HTN = hypertension, BP = blood pressure H = high risk of bias M = moderate risk of bias, L = Low risk of bias.*

| Author, Country, Year | Study Design | N | Measurement of BP | Recruitment Period | Follow up | Outcome Measure | Risk of bias |
|---|---|---|---|---|---|---|---|
| Masood, Canada, 2016 (137) | R | 206,147 | EPR | 2002-2012 | 2 years | All-cause mortality | M |
| McNaughton, USA, 2015 (116) | R | 701,952,422 | EPR | 2006-2012 | N/A | Descriptive | M |
| Meurer, USA, 2019 (136) | P | 201 | 1 x ED 1 x self-reported | 2014-2015 | 3 months | Sustained HTN | M |
| Oras, Sweden, 2020 (128) | R | 300,193 | EPR | 2010-2016 | 6 yeas | MACE | M |
| Shah, USA, 2011 (124) | R | 601 | EPR | 2009-2010 | N/A | F/U | M |
| Shiber-Ofer, Israel, 2015 (133) | P | 195 | ED | 2009-2010 | 5 years | F/U BP | L |
| Souffront, USA, 2016 (129) | R | 2367 | EPR | 2014-2015 | N/A | N/A | M |
| Svenson, USA, 2008 (121) | R | 2821 | EPR | 2006 | 1 year | Descriptive | M |
| Tan, Australia, 2013 (126) | P | 534 | 1 x Researcher | 2010-2011 | 5 weeks | Descriptive | M |
| Tanabe, USA, 2004 (117) | R | 88 | ED | 2001 | N/A | Descriptive | M |
| Tanabe, USA, 2008 (83) | P | 175 | 2 x ED | 2005-2006 | 1 week | Sustained HTN | L |
| Tilman, USA, 2007 (119) | R | 1574 | EPR | 2004 | N/A | BP advice | M |
| Tsoi, Hong Kong, 2012 (135) | P | 245 | 2 x ED | 2010 | 2 weeks | Primary care F/U | M |
| Umscheid, USA, 2008 (122) | R | 2061 | EPR | 2005 | N/A | Descriptive | M |

*Table 2.2 Characteristics of the included studies examining hypertension (part two of two), R = retrospective, P = prospective, EPR = electronic patient record, N/A = not applicable, F/U = follow-up, HTN = hypertension, BP = blood pressure H = high risk of bias M = moderate risk of bias, L = Low risk of bias.*

### 2.7.3 Risk of bias within studies

We used the QUIPS tool to assess the risk of bias, most studies were rated as a moderate overall risk of bias (Table 2.3). No studies were found to have a high risk and two were found to have a low risk of bias (83,133). The domain which qualified studies as moderate risk was mostly the study participation domain for inclusion criteria. Four studies adjusted for anxiety and fear as confounders, however they all demonstrated a persistence of hypertensive readings despite the presence of these factors. Some studies accounted for the potential confounders of fear and anxiety (81,83,131,132). A moderate risk of bias was assigned to this domain if potential confounders were not accounted for.

| Author, Year | Domain 1 Population | Domain 2 Attrition | Domain 3 Factor | Domain 4 Outcome | Domain 5 Confounding | Domain 6 Analysis | Overall |
|---|---|---|---|---|---|---|---|
| Adhikari, 2016 (131) | Moderate | Low | Moderate | N/A | Moderate | Low | Moderate |
| Backer, 2003 (132) | Moderate | Low | Low | Low | Moderate | Low | Moderate |
| Baumann, 2007 (120) | Low | Low | Low | Low | Low | Moderate | Moderate |
| Baumann, 2009 (123) | Moderate | Low | Moderate | N/A | Moderate | Moderate | Moderate |
| Cienki, 2013 (130) | Low | Low | Moderate | Moderate | Moderate | Moderate | Moderate |
| Cline, 2006 (134) | Moderate | Moderate | Moderate | Moderate | Moderate | Moderate | Moderate |
| Collins, 2008 (87) | Moderate | Low | Low | Low | Low | Moderate | Moderate |
| Dolatabadi, 2014(125) | Low | Moderate | Low | Low | Low | Moderate | Moderate |
| Fleming, 2005 (81) | Low | Moderate | Low | Low | Low | Low | Moderate |
| Goldberg, 2017 (127) | Low | Low | Low | Low | Moderate | Low | Moderate |
| Julliard, 2012 (88) | Moderate | Moderate | Low | Moderate | Moderate | Low | Moderate |
| Karras, 2005 (118) | Low | Low | Moderate | Moderate | Moderate | Moderate | Moderate |
| Lee, 2018 (82) | Moderate | Moderate | Moderate | Low | Moderate | Low | Moderate |
| Masood, 2016 (137) | Moderate | Moderate | Moderate | Moderate | Moderate | Low | Moderate |
| McNaughton, 2015 (116) | Moderate | Moderate | Moderate | N/A | Moderate | Moderate | Moderate |
| Meurer, 2019 (136) | Low | Moderate | Moderate | Low | Low | Low | Moderate |
| Oras, Sweden, 2020 (128) | Low | Low | Moderate | Moderate | Moderate | Low | Moderate |
| Shah, 2011 (124) | Moderate | Low | Moderate | Moderate | Moderate | Moderate | Moderate |
| Shiber-Ofer, 2015 (133) | Low | Low | Low | Low | Low | Low | Low |
| Souffront, 2016 (129) | Moderate | Low | Low | Moderate | Low | Moderate | Moderate |
| Svenson, 2007 (121) | Moderate | Moderate | Low | Low | Moderate | Low | Moderate |
| Tan, 2013 (126) | Low | Low | Low | Low | Low | Moderate | Moderate |
| Tanabe, 2004 (117) | Moderate | Low | Moderate | Moderate | Moderate | Moderate | Moderate |
| Tanabe, 2008 (83) | Low | Low | Low | Low | Low | Low | Low |
| Tilman, 2006 (119) | Low | Low | Low | Low | Moderate | Low | Moderate |
| Tsoi, 2012 (135) | Low | Low | Low | Low | Low | Moderate | Moderate |
| Umscheid, 2008 (122) | Moderate | Low | Low | Moderate | Low | Low | Moderate |

*Table 2.3- Risk of bias in studies examining hypertension. This was conducted using the QUIPs tool (20). Domain 1 – Study Participation, domain 2 – Study Attrition, domain 3 – Prognostic factor measurement, domain 4 – outcome measurement, domain 5 – study confounding, domain 6 – Statistical analysis and reporting*

### 2.7.4 Extracted outcomes

Across studies, the prevalence hypertension in the ED (defined by a single reading of ≥ 140/90 mmHg) ranged from 16% to 62% (119,127). The pooled prevalence of hypertension was 31% (95% CI 25 – 37%). This pooled estimate was unchanged in a sensitivity analysis where the largest study was removed; 31% (95% CI 25 – 39%) (Figure 2.2) (116). Perhaps due to the inclusion of studies with different designs and in different settings, we estimated very high heterogeneity, with an $I^2$ of 99.7. Heterogeneity has also been noted to be higher due to the inclusion of more studies in a meta-analysis which is relevant here (138). We conducted further sensitivity analyses to see the effect of different subgroups of studies on the pooled results. However, they did not demonstrate a significant association with the result (subgroups included: blood pressure measurement method, risk of bias, prospective vs retrospective and risk of bias). Furthermore, having a pooled estimate is most informative for clinicians. The width of the 95% confidence intervals reflects the heterogeneity and even the lower 95% confidence interval is a clinically meaningful prevalence.

Of those with an initial hypertensive reading, hypertension was found to persist throughout the ED visit in 62% of patients, overall (pooled estimate, 95% CI 56 – 68%), with a range from 44 to 77% ($I^2$ 0.84, indicating high heterogeneity; see Figure 2.3). (127,131). The lowest rate of persistence was from Goldberg et al at 44%, this could potentially be due to the automated BP measuring device used which averages 5 readings taken over 5 minutes. Following discharge from the ED, hypertension persisted in 50% of patients (pooled estimated, 95% CI 38 – 61%, $I^2$ 90.2%, see Figure 2.4), with a range of 26 to 88%.

There were substantial differences in the proportion of patients who attended follow-up after hypertensive readings in the ED, with estimates ranging from 2% to 65% (I2 97.3%) (129,132). The meta-analysis found a pooled estimate for attending a follow up appointment for ED hypertension to be 0.41 (95% CI 0.23 – 0.62), see Figure 2.5. Again, a sensitivity analysis was conducted to examine the effect of subgroups on the heterogeneity, it appeared to be driven in part by prospective vs retrospective methodologies.

The prognostic characteristics of hypertensive readings for long term cardiovascular outcomes were identified in three large retrospective studies from Canada, South Korea and Sweden (82,128,137). The Canadian study examined ED patients who had a coded diagnosis of hypertension and their outcomes at two years. Of the 206,147 patients 3.59% (95% CI 3.51-3.68) died and 9.3% had a complication of hypertension (control data was not presented) (137). The incidence of major adverse cardiac events (MACE) was examined in a South Korean database comprised of 262,972 low acuity attendances. The association of coded hypertension with MACE was examined and at 0-3 years follow up the hazard ratio (HR) was 4.25 (95% CI 3.83 - 4.71), at 4-6 years follow the HR was 3.65 (95% CI 3.14 - 4.24), and at 7-10 years follow up the HR was 3.20 (95% CI 2.50 - 4.11) (82). Incident atherosclerotic CVD was examined in a Swedish retrospective study of 300,193 ED (128). Over a mean follow up time of 3.5 years a systolic blood pressure of 140-159 mmHg had a HR of 1.24 (95% CI 1.09-1.41), 160-179mmHg had a HR of 1.62 (95% CI 1.42-1.85) and >180mmHg had a HR of 2.02 (95% CI 1.75-2.33) (128). This demonstrates an exposure-response relationship for atherosclerotic CVD and increasing grades of systolic hypertension. A meta-analysis was not conducted combining the Swedish and South Korean studies due to the differing outcome definition. The South Korean study included ACS, revascularisation, stroke, heart failure, pacemaker insertion and cardiovascular death (137). Whereas the Swedish study included AMI, stroke, and cardiovascular death (128).

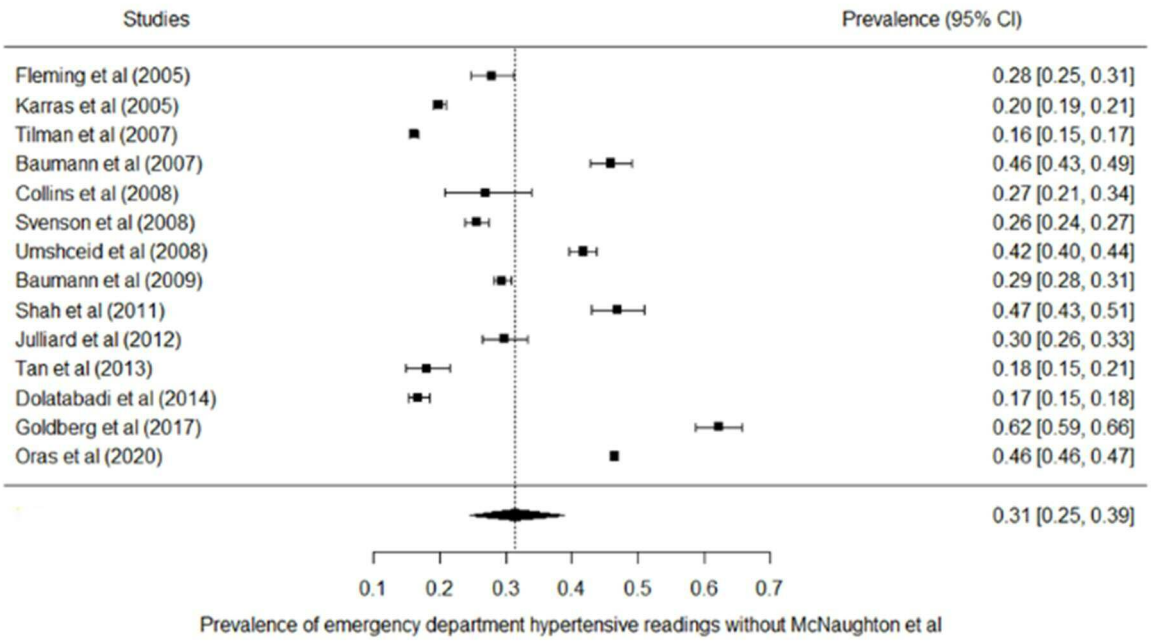



*Figure 2.2 Meta-analysis forest plot of the prevalence of hypertensive readings in the emergency department*

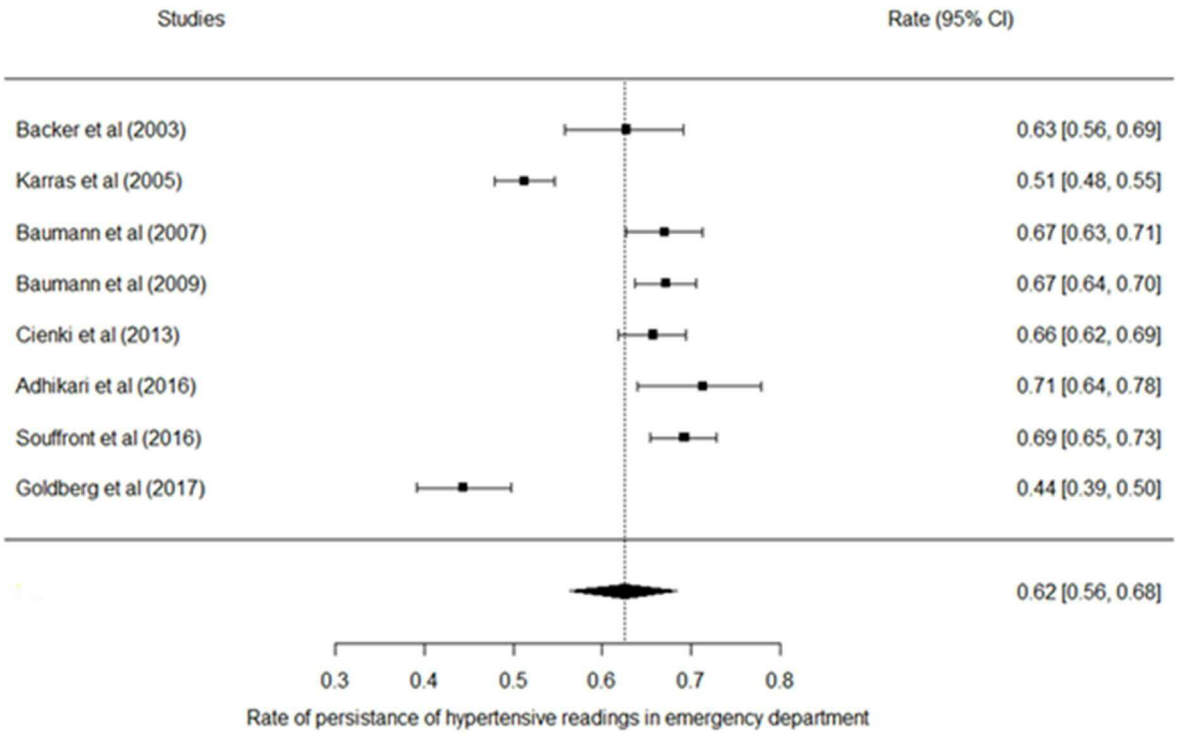



*Figure 2.3 Meta-analysis forest plot of the persistence of hypertensive readings in the emergency department*

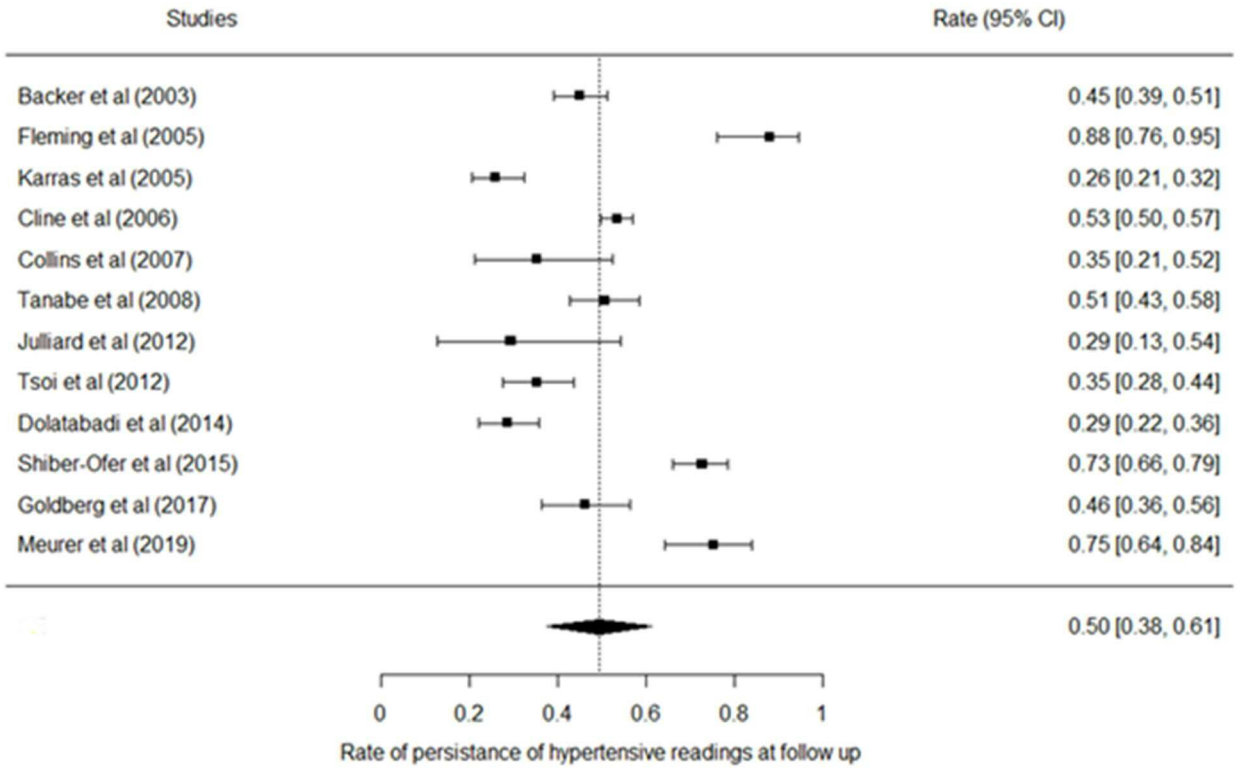



*Figure 2.4 Meta-analysis forest plot of the persistence of hypertensive readings at follow up appointments*

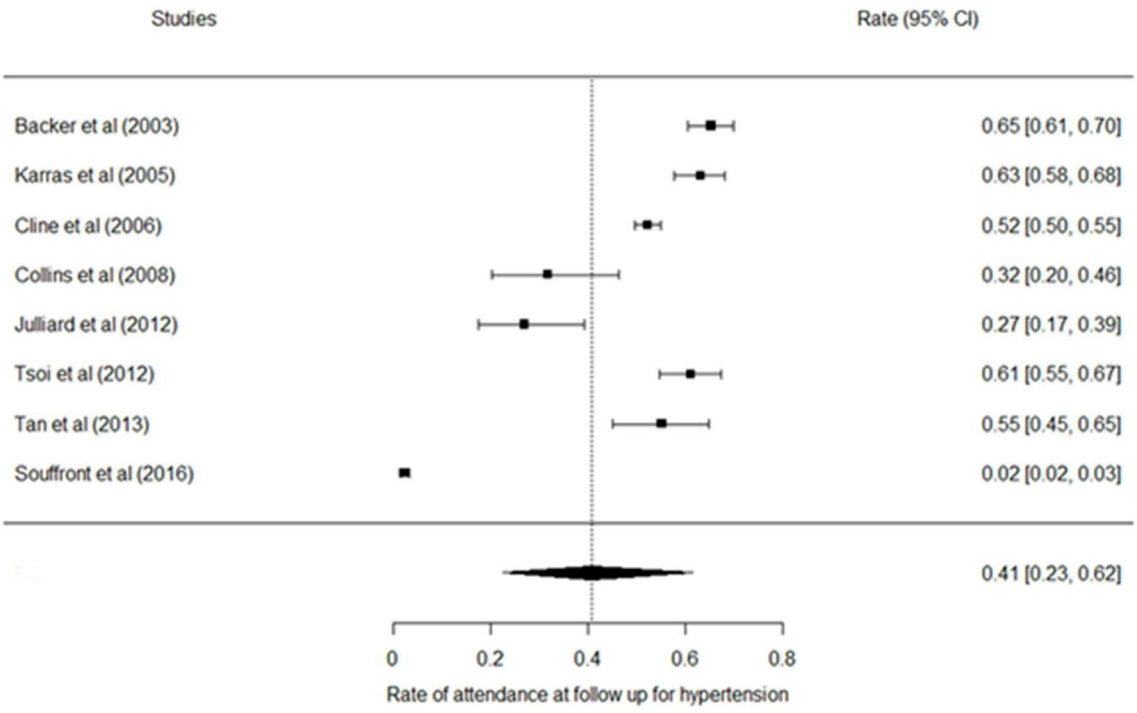



*Figure 2.5 Meta-analysis forest plot of the rate of attendance at hypertension follow up clinics*

## 2.8 Results for the prognostic factor type two diabetes mellitus

### 2.8.1 Search Results

Searches were conducted on the 09/10/2021 and we identified 969 studies, of which 779 were unique. After title and abstract review 21 underwent full text review and 10 were selected for inclusion (Figure 2.6).

### 2.8.2 Study characteristics

The majority of the studies were conducted in the USA and two diagnostic criteria were encountered for T2DM: the World Health Organisation (WHO) definition (102) and American Diabetes Association (ADA) definition (103). Eight of the 9 studies were prospective, and the sample sizes ranged from 101 to 2652 (Table 2.4). All studies were from general ED cohorts except those with outcomes for those reporting prognostic characteristics for CVD.

### 2.8.3 Risk of bias

A moderate risk of bias was identified in all but two studies, with Silvermann et al (2006) obtaining a low score (139) and Silvermann et al (2011) obtaining a high score (140). Most studies had a high participant attrition rate and failed to account for confounders that may have influenced blood glucose such as physiological stress. Full details are given in Supplementary Table 8.1.

### 2.8.4 Prevalence of type two diabetes mellitus

Two definitions of the target condition were present amongst the identified studies: WHO and ADA definitions. Both used similar thresholds, but the WHO definition explicitly requires the two abnormal results to be from two separate samples drawn at different times. These different definitions were examined separately in meta-analysis. The studies using the WHO definition screened for T2DM in a two-phase process using either a blood glucose or HbA1c as a first screen. If this led to a positive result, the patients underwent further confirmatory investigations. The reported prevalence is for confirmed cases of T2DM by the end of the diagnostic pathway which included follow-up visits.

Six studies reported against the ADA definition (140–145). Each study included general ED patients based on elevated HbA1c and serum glucose levels in the ED, but they used different glucose thresholds (5.5, 6.0, 7.0 mmol/L). On meta-analysis, the pooled prevalence of confirmed T2DM was 30% (95% CI 18-46%), the I2 value was 70.3% (Figure 2.7). Seven studies were identified that took serial samples over separate days thereby enabling the analysis of prevalence as per the WHO definition (139–143,145,146). An exact likelihood random effects model was used that gave an estimated prevalence of 18% (95% CI 10 – 29%, I2 91.14%, Supplementary Figure 8.1).

### 2.8.5 Adherence to follow-up for type two diabetes mellitus

Five studies reported the proportion of patients who were identified as at risk of type 2 diabetes mellitus in the ED, defined by elevated HbA1c or serum glucose, who attended follow-up (139,141–145), observed in the USA, Australia and the UK. The pooled estimate of the proportion of patients attending follow-up appointments was 0.55 (95% CI 0.25 - 0.82). The $I^2$ value was 98.3% demonstrating significant heterogeneity across these studies (Supplementary Figure 8.2).

### 2.8.6 Prognostic value of type two diabetes mellitus

Only two studies reported the association between previously diagnosed T2DM and long-term cardiovascular outcomes: Farkouh et al and Sanchis et al (147,148). Farkouh et al retrospectively examined 2,271 patients presenting with chest pain for a median follow-up of 13.6 years in Canada (147). The adjusted hazard ratio (HR) for T2DM to predict mortality was found to be 2.45 (95% CI 2.14-2.81) and being under 50 years old interacted with T2DM, resulting in an overall HR 4.7 (95% CI 1.56-14.1). Farkouh et al also examined a composite outcome for major adverse cardiac and cerebral events which included death, AMI, stroke, or revascularisation. The HR for this composite outcome was 3.19 (95% CI 1.65 – 6.18). Sanchis et al prospectively analysed 1,011 patients presenting with chest pain across one year, with the primary outcomes of all-cause mortality and acute myocardial infarction. The hazard ratio for those with T2DM was reported as 2.3 (95%CI 1.4-3.8) for this composite outcome. A meta-analysis was not conducted due to the small number of studies.

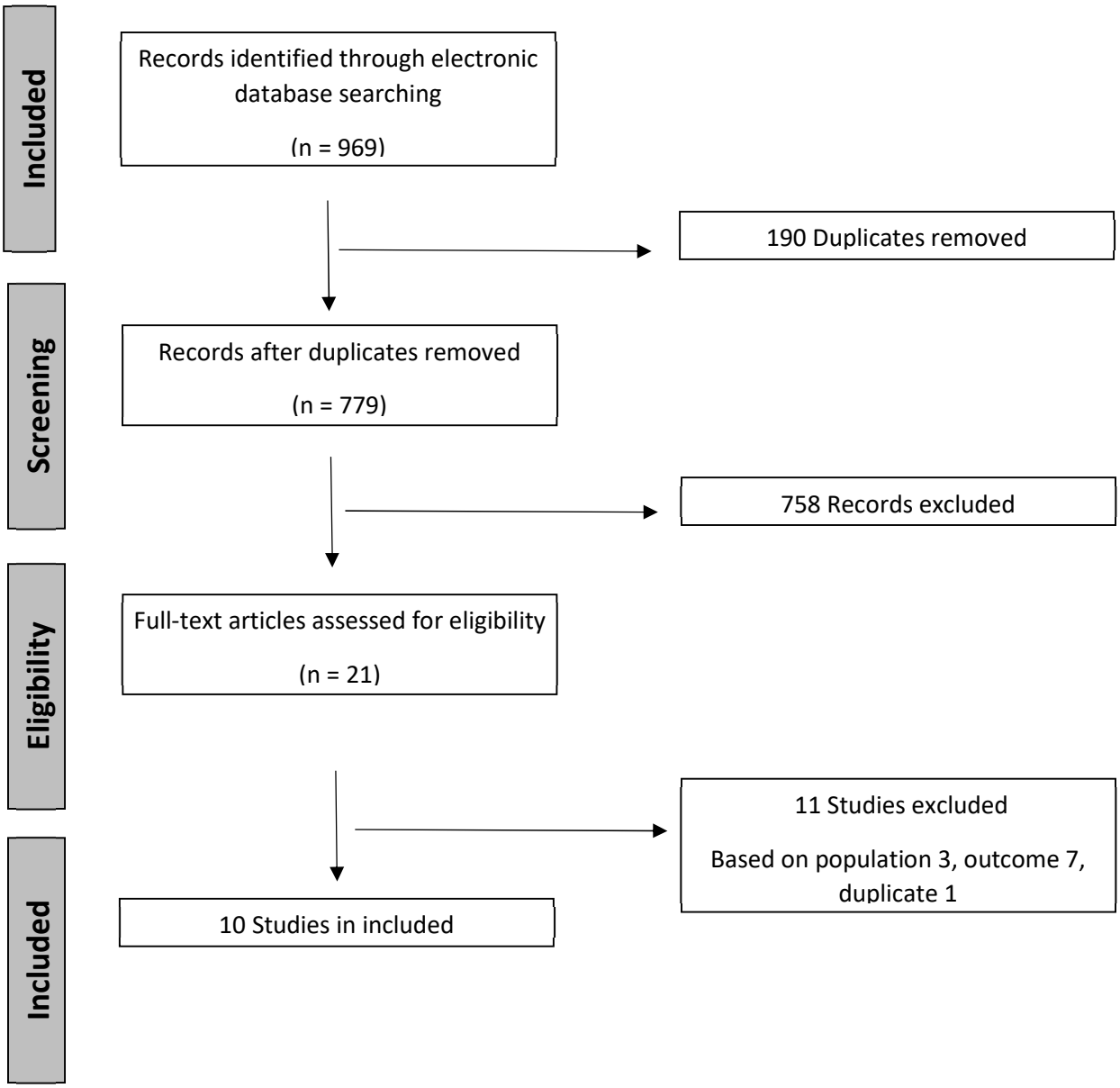



*Figure 2.6 Flowchart of type 2 diabetes mellitus study selection process*

| Author, Country, Year | Study Design | N | Glycaemic Threshold (mmol) | Recruitment Period | Follow up | Outcome Measure | DM definition | Population | Risk of bias |
|---|---|---|---|---|---|---|---|---|---|
| Hng, Australia 2016 (146) | P | 2652 | 5.5 | NR | N/A | T2DM Prevalence | HbA1c | General | M |
| Jelinek, Australia, 2010 (145) | P | 725 | 6.0 | 2007 - 2008 | 4 weeks | T2DM Prevalence | OGTT | General | M |
| Charfen, USA, 2009 (142) | P | 528 | 7.8 | 2004-2006 | 6 weeks | T2DM Prevalence | HbA1c & PBG | General | M |
| Hewat, UK, 2009 (144) | P | 101 | 5.5 | 2006-2007 | 2 months | T2DM Prevalence | NR | General | M |
| George, UK, 2005 (141) | P | 500 | 7.0 | 2003 | 1 week | T2DM Prevalence | Fasting CBG | General | M |
| Silverman, USA, 2006 (139) | P | 541 | 6.1 | 2006 | N/A | Elevated HbA1c | N/A | General | L |
| Ginde, USA, 2008 (143) | P | 265 | N/A | 2007 | 3 months | T2DM Prevalence | OGTT | General | M |
| Silverman, USA 2011 (140) | P | 618 | N/A | 2005-2007 | 55 days | T2DM Prevalence | OGTT | General | H |
| Farkouh, USA, 2009 (147) | R | 2271 | N/A | 1985-1992 | 7.5 years | Mortality / MACE | N/A | Chest-pain | M |
| Sanchis, USA, 2008 (148) | P | 1011 | N/A | 2001-2006 | 1 year | Death or AMI | N/A | Chest-pain | M |

*Table 2.4 - Characteristics of the included studies for the prognostic factor type 2 diabetes mellitus, R = retrospective, P = prospective, MACE = major adverse cardiac event, AMI = acute myocardial infarction, CBG = capillary blood glucose, OGTT = oral glucose tolerance test, H = high risk of bias M = moderate risk of bias, L = Low risk of bias. Glycaemic threshold = the threshold above which further confirmatory testing was conducted.*

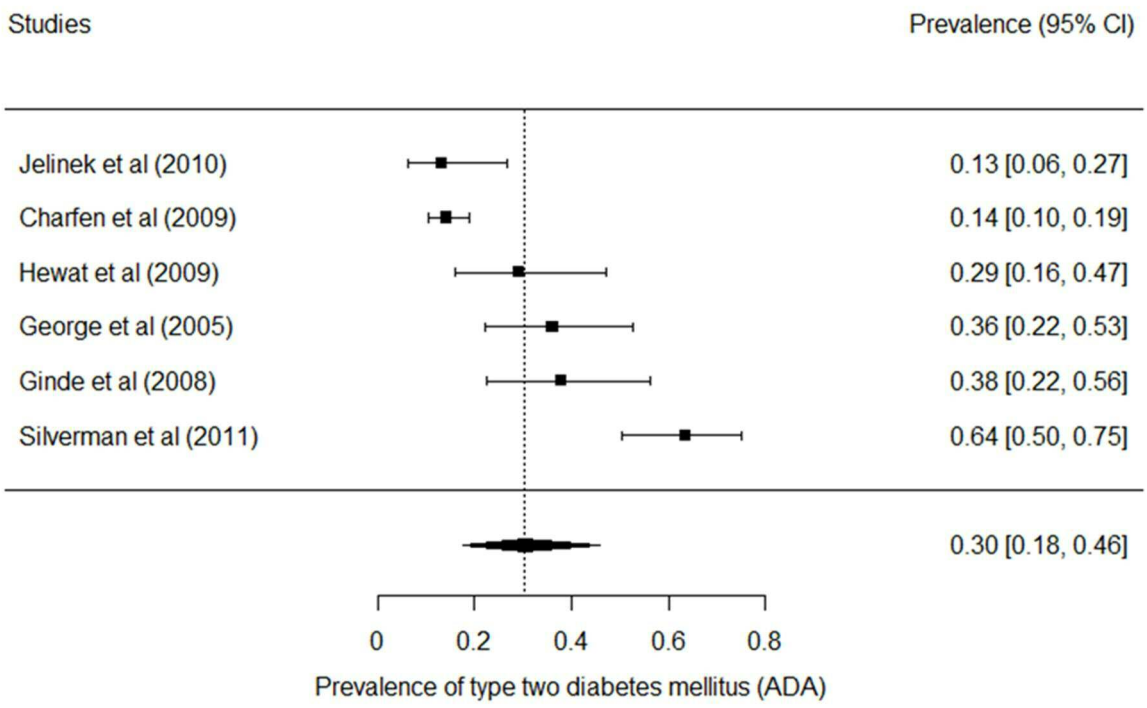



*Figure 2.7 -Meta- analysis forest plot of prevalence of type two diabetes mellitus by American diabetes association definition. This includes only the general ED population has no other results were identified.*

## 2.9 Results for the prognostic factor dyslipidaemia

### 2.9.1 Search Results

We identified 723 papers in the database searches, consisting of 719 unique (Figure 2.8). Seven were selected for full text review, resulting in 3 further exclusions due to the sample population, leaving 4 included for review (126,149–151).

### 2.9.2 Study characteristics

Of the four studies included, two were from the USA (150,151) and two were from Australia (126,149) (Table 2.5). Only Tan et al's study recruited undifferentiated ED patients (126), with the remainder focusing on a suspected ACS population. Three studies reported total cholesterol but used a variety of different thresholds (126,149,150) ranging from >5.0mmol/L ((149)), to >6.0mmol/L (126). The baseline population prevalence of dyslipidaemia was reported to be similar at 51.5 and 51 percent in each country (149,150). High density lipo-protein was reported in two trials (149,150), and low density lipo-protein levels were reported in two studies (149,151).

### 2.9.3 Risk of bias

Elder et al scored moderate on the domain study attrition due to 51% of patients being missed and thereby potentially incorporating bias into the findings. The remaining three studies had a low risk of bias with only the prevalence of dyslipidaemia reported. Full details are given in Supplementary Table 8.2.

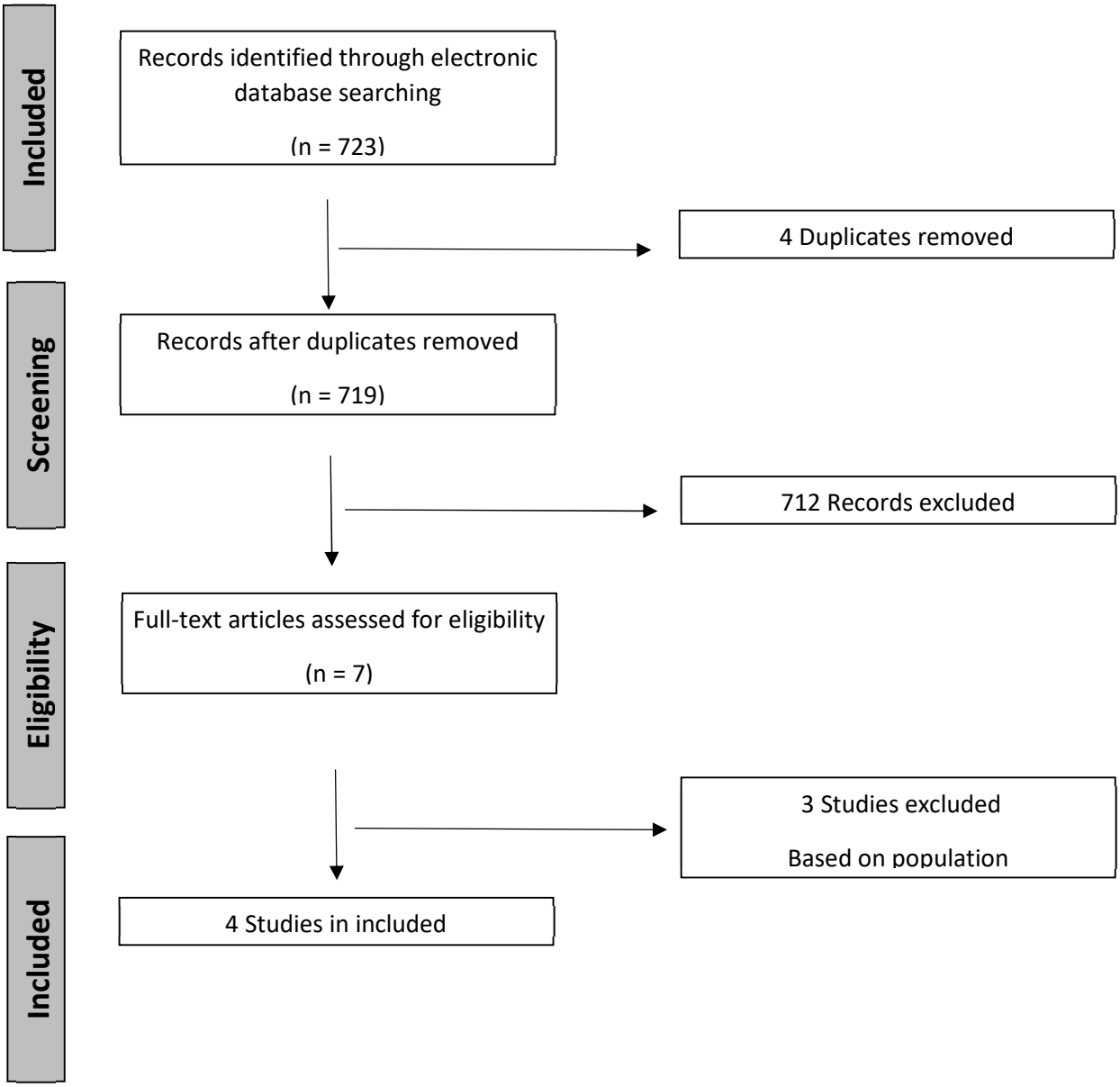



*Figure 2.8 Flowchart of dyslipidaemia study selection process*

| Author, Country, Year | Study Design | N | Recruitment Period | Components of dyslipidaemia | Total Cholesterol (mmol/L) | Outcome Measure | Population | Risk of bias |
|---|---|---|---|---|---|---|---|---|
| Elder, Australia, 2006 (149) | P | 185 | 2003-2004 | TC, LDL, HDL, Apo B | > 5.0 | Prevalence | Chest-pain | M |
| Diercks, USA, 2002 (150) | P | 606 | 1999 | TC, HDL | > 5.2 | Prevalence | Chest-pain | L |
| Chandra, USA, 2002 (151) | P | 112 | 2001 | LDL | > 4.9 | Prevalence | Chest-pain | L |
| Tan, Australia, 2013 (126) | P | 827 | 2010-2011 | TC | > 6.0 | Prevalence | General | L |

*Table 2.5 Characteristics of the included studies for dyslipidaemia search, R = retrospective, P = prospective, TC = total cholesterol, LDL = low density lipo-protein, HDL = high density lipo-protein, Apo B = Apolipoprotein B, H = high risk of bias M = moderate*

### 2.9.4 Prevalence of dyslipidaemia

A direct comparison of the prevalence of dyslipidaemia was not possible due to differing thresholds and tests. Total cholesterol levels were abnormal in 21% and 57% of patients respectively, across three studies. Elder et al examined any patient with non-traumatic chest pain and had the highest prevalence of abnormal total cholesterol (>5mmol threshold) at 57% (149). Diercks et al examined low risk patients admitted to a chest pain evaluation unit and found a prevalence of abnormal total cholesterol to be 41% (>5mmol threshold) (150). Tan et al found the lowest prevalence of abnormal total cholesterol levels when examining all ED attendees without a "substantial illness", it was 21% using a higher threshold of >6mmol (126). Whereas high LDL was noted in 25% to 52% of patients and high HDL in 11 to 24%. Because of the variations in the thresholds used to diagnose dyslipidaemia between studies, meta-analysis was not felt to be appropriate (Table 2.6).

| | Number of studies | Pooled estimate (95% CI) | Min | Max |
|---|---|---|---|---|
| Prevalence in ED setting | | | | |
| **T2DM** | 10 | - | - | - |
| ADA | 6 | 0.3 (0.18 - 0.46) | 0.13 | 0.64 |
| WHO | 7 | 0.18 (0.1 - 0.29) | 0.11 | 0.38 |
| **Dyslipidaemia*** | 4 | - | - | - |
| Total cholesterol | 2 | - | 0.21 | 0.57 |
| LDL | 2 | - | 0.25 | 0.52 |
| HDL | 2 | - | 0.11 | 0.24 |
| Follow-up attendance | | | | |
| T2DM | 5 | 0.55 (0.25 - 0.82) | 0.23 | 0.99 |

*Table 2.6 Summary of systematic review findings – ADA – American Diabetes Association definition, WHO – World Health Organisation definition, LDL – low density lipo-protein, HDL – High density lipo-protein, T2DM – Type two diabetes mellitus, *- some studies reported multiple outcomes*

## 2.10 Results for the prognostic factor chronic kidney disease

### 2.10.1 Study selection

Through our electronic database search, we identified 259 papers, 197 remained after removal of duplicates. Five were selected for full text review after 3 were excluded, 2 were included for review (148,152). Exclusions were based on outcome, population and duplicates (Figure 2.9).

### 2.10.2 Study Characteristics

Both studies focused on chest pain populations, with Sanchis et al conducting a prospective analysis in a low risk chest pain population of 1,011 patients (no ST deviation on ECG or abnormal troponin result)(148). It was conducted in a single Spanish hospital and included patients from 2001- 2006. Chaikriangkrai et al ran a retrospective analysis of chest pain patients in whom a computed tomography coronary angiogram had been done (152). It included 949 patients, from a single USA centre between 2005 – 2008 (152). Each study used different outcomes over different follow up periods. Sanchis et al recorded one year rate of all-cause mortality or AMI and Chaikriangkrai examined major adverse cardiac events over a median of 5.3 years (148,152). Both converted the continuous variable of renal function into a dichotomous one, thereby losing statistical power, estimated to be equivalent to discarding a third of the data (153).

### 2.10.3 Risk of bias

Risk of bias was assessed using the QUIPS tool (104), this is reported in Supplementary Table 8.3. Chaikriangkrai et al was given a high-risk bias due to a very large study attrition rate. Only 949 out of a possible 5066 patients were included. This was mainly due to a lack of CT coronary angiography being conducted, however unfortunately there was no comparison of these two groups to quantify the effect of the missing data.

The only successfully extracted outcome were for the prognostic effect of CKD on long term CVD outcomes.

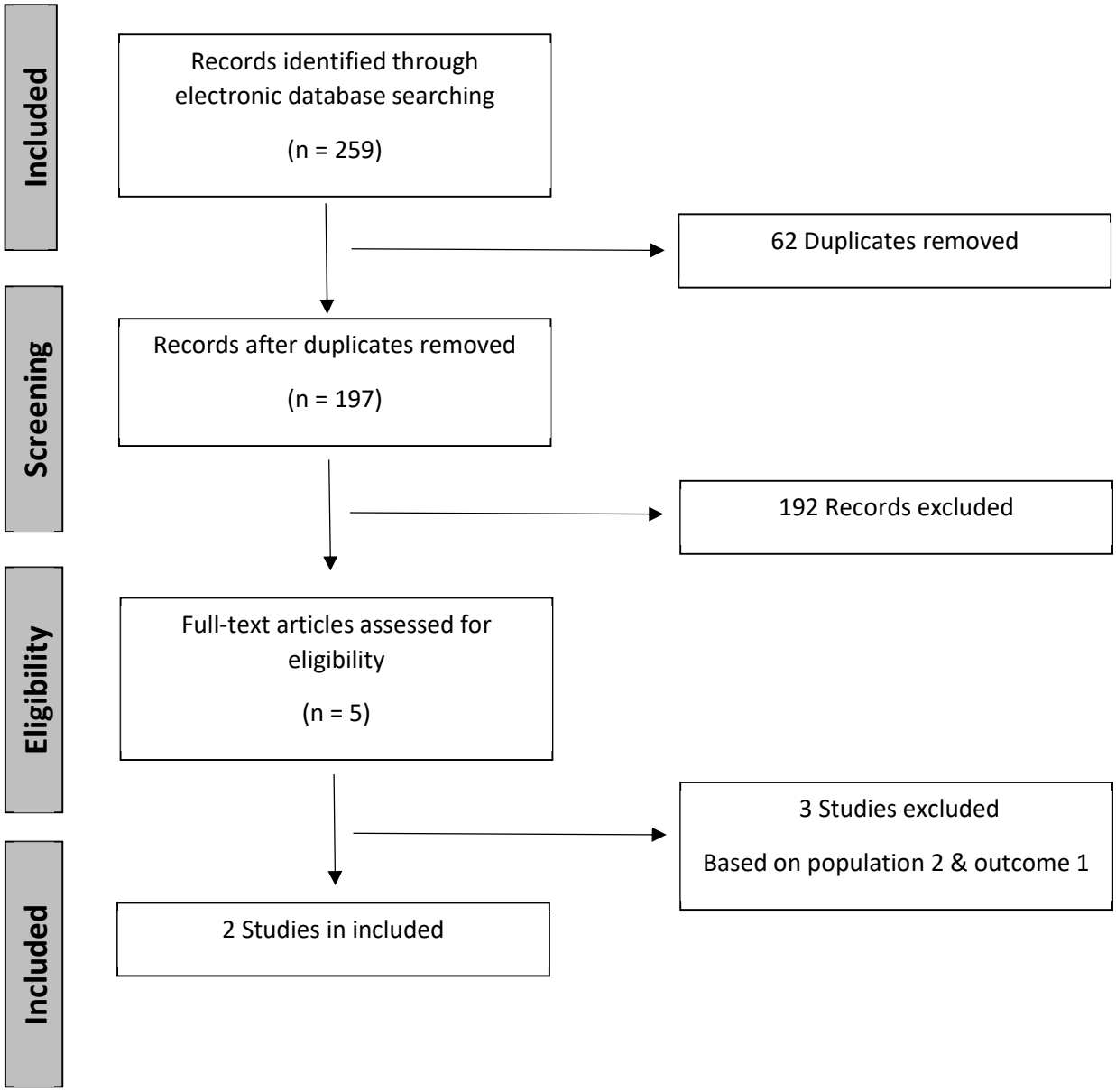



*Figure 2.9 - Flowchart of chronic kidney disease study selection process*

### 2.10.4 Findings

Sanchis et al did not find CKD to be a significant individual prognosticator in a cox regression analysis for the composite outcome of all cause death or AMI at one year (148). Chaikriangkrai et al demonstrated a significant effect of CKD on MACE with a multivariable Cox regression analysis; hazard ratio 10.18 (95% CI 4.24 - 24.25) (152). Examining the Kaplan-Meier curves the hazard ratio only significantly differs beyond one year.

## 2.11 Discussion

The detection of CVD prognostic factors where it is unrelated to the presenting complaint could be considered preventative medicine. Such public health programmes in the emergency department have been argued for before. Bernstein et al described it as a "critical juxtaposition", highlighting the large need of the ED population contrasted by a specialty that does not classically take a role in primary prevention (154).

This work has demonstrated that screening patients in the ED would reveal a high prevalence of prognostic factors for future cardiovascular disease. It is often believed that hypertension in the ED may be caused by pain or anxiety, leading clinicians to downplay its significance. However, this systematic review has also identified a substantial rate of persistence beyond the emergency department and that it was prognostically important. We also found that T2DM and dyslipidaemia were prevalent in the ED population, particularly those presenting with chest pain, further indicating the potential for targeted CVD risk factor screening.

For hypertension the prevalence was the most reported however it consisted of mainly moderately biased studies. The meta-analysis estimate for the prevalence of ED hypertension was 0.31 (95% CI 0.25 – 0.37). Furthermore pain was adjusted for in several of the studies and does not appear to have a significant effect on the readings.(81,88,117,117,121,122,131–133). These hypertensive readings persist within ED, and at follow up (pooled proportion of 0.67 and 0.55 respectively). This

suggests that observing hypertension within the ED can not only be indicative of persistent hypertension  but there was evidence that hypertensive ED readings were associated with an increased risk of long term CVD outcomes (82,137).

A standard approach for confirming hypertension might be serial measurements between the ED, outpatient clinics or primary care. Unfortunately, we demonstrated that services might only expect an attendance rate of 41%. To capture these patients in a care pathway this confirmation could be conducted in the ED, thereby avoiding loss to follow up. Such a service would be feasible as per the review by Bowen et al due to the demonstrated prevalence and persistence (155).

Our review has identified several articles which indicate that the screening of T2DM and dyslipidaemia within patients presenting to the ED provides an opportunity for prevention, by early detection and management of risk factors for CVD. In particular, the pooled prevalence for T2DM was found to be 30% in the ED setting and was associated with an over 2-fold hazard on cardiac-related poor outcomes and death in multiple studies, indicating that there could be a missed opportunity in current practice.

The prevalence of the T2DM and dyslipidaemia favour a screening programme in the emergency department. However, more needs to be done to ensure retention in the care pathway; the proportion of patients with suspected T2DM attending outpatient follow up appointments was low with a pooled estimate of 55%. This may pose a challenge to any planned clinical pathway that would seek to improve care in this population. It still may be worthwhile given that studies were also found indicating an increased risk for long term CVD disease in the ED population (147,148).

For chronic kidney disease no prevalence estimates, or reliability outcomes were successfully identified.

Two studies examined the effect of CKD on cardiovascular outcomes. Sanchis et al found no effect , and Chaikriangkrai et al found to infer a higher risk of MACE with a hazard ratio of 10.18 (95% CI 4.24 – 24.25) (148,152). Between these two contrasting studies there are three main issues:

(i) Sanchis et al used a single serum creatinine measurement whereas Chaikriangkrai used an estimated glomerular filtration rate.
(ii) They examined different composite outcomes.
(iii) The increased risk was only observed by Chaikriangkrai et al beyond 12 months, which was the length of follow up for Sanchis et al.
(iv) Chaikriangkrai et al has a high risk of bias in QUIPs assessment (148,152).

A targeted screening programme could capitalise on a period of increased acceptability for health belief change during an acute episode (156). For example, when exploring screening in patients presenting with chest pain, those that had acute coronary syndromes excluded demonstrated a willingness to engage with modification of CVD risk factors (85). Consequently, there is potential for targeted screening of a population that would be more receptive to change.

These risk factors could also be used to inform a model for long term CVD, and once risk stratified the patient could go on to receive targeted interventions. This would represent an efficient use of healthcare visits, and a more holistic approach to the patient.

### 2.11.1 Limitations

The generalisability of the meta-analysis is limited by the degree of heterogeneity that was identified. This is likely due to varying practice by country, study type and treatment thresholds. It has also been proposed that heterogeneity is also associated with a large number of included studies in meta-analyses (138). Despite this the pooling estimate is most informative for clinicians and the heterogeneity is reflected in the 95% confidence intervals which are provided. The lower boundary of the 95% confidence intervals demonstrate a significant prevalence of these factors, therefore even with significant heterogeneity the factors are likely important.

## 2.12 Conclusions

Type two diabetes mellitus can be feasibly assessed in the ED, and there is early evidence that indicates that it is predictive of CVD diagnosis. Dyslipidaemia, across a range of definitions, appears to be prevalent in ED populations but further research is required to estimate its prognostic performance.

CKD measured by eGFR in the ED had sparse data and as such no firm conclusions can be drawn regarding the prognostic effect of chronic kidney disease in this population.

ED measured hypertension is prevalent, persistent, and predictive of long-term CVD outcomes. The American college of emergency physicians recommendations state that ED measured hypertension should be followed up by an outpatient appointment or immediate therapy initiated (157). This systematic review indicates that this is a well-founded recommendation. The care pathway requires careful consideration and planning given the potential for low attendance at follow up.

ED measured hypertension, T2DM and dyslipidaemia present an opportunity to further the 'critical juxtaposition' of preventative medicine in emergency medicine with little additional burden (154).

# Chapter 3 Evaluation of the long-term prognostic value of routinely collected data for patients presenting to the Emergency Department with chest pain

## 3.1 Background

In this chapter I describe the work that builds on the previous chapter. There I demonstrated that ED measured hypertension, CKD, T2DM and dyslipidaemia were prevalent and had favourable prognostic characteristics for long term CVD. However, there are evidence gaps, many potential predictors are yet to be evaluated. Therefore, building on my work in the prior chapter I sought to examine routinely collected ED data to confirm the predictive characteristics of predictors.

Diagnostic innovations in the suspected cardiac chest pain care pathway have been shown to demonstrate favourable prognostic characteristics for long term CVD outcomes (147,158,159). However, there is still an evidence gap to fully describe the prognostic characteristics of routinely collected data for the outcome long term CVD.

### 3.1.1 Additional methodological background

#### *3.1.1.1 Variable Selection*

To measure the prognostic characteristics of candidate variables they can be assessed within multivariable cox proportional hazards model which provides adjusted estimates. The variables that make up this model need to be selected and there are various methods to do this including expert opinion, univariate analysis, stepwise selection and least absolute shrinkage and selection operator analysis (LASSO) (160). Tibshirani proposed that LASSO variable section produced less variable models than other selection methods, however this has been disputed (161,162). When practically implemented in the statistical programme R, LASSO selection requires a package that can utilise it for variable selection. There is currently only one package that can conduct LASSO variable selection but it cannot handle other complex functions (such as restricted cubic splines) (160).

#### *3.1.1.2 Cox proportional hazard model*

Cox proportional hazard models require upon two underlying assumptions to be correct (i) the relationship between the variable and the outcome must be linear, and (ii) the relative hazard does not change over time (163).

##### Assessing the linearity of the outcome – candidate predictor relationship

To assess the linearity of the relationship between the variable of interest and the outcome simple plots of the potential predictor or interest versus the outcome can be used. There are also Martingale plots that can assess the linearity where a deviation from a best fit line of model residuals with a gradient of 1 is an indication of a non-linear relationship (164).

Restricted cubic splines (RCS) can be used to transform variables which have a non-linear relationship with the outcome. The spline is fitted to the data in a number of windows referred to as knots, the more complex the relationship the larger the number of knots that may be required (165). However, the more knots that are used the larger the risk of over-fitting. AIC metrics have previously been used for the selection of smoothing parameters (number of knots), Keele describes using univariate cox proportional hazard models with the untransformed variable and comparing the AIC of univariate RCS models (166).

##### Assessing the proportional hazards assumption

Schoenfeld plots are a graphical representation of the scaled Schoenfeld residuals from cox regression against time and are used to detect if the proportional hazards assumption has been maintained. Grambsch et al ran a series of simulations and demonstrated that plots with non-zero slopes indicate the assumption has been broken (167). Further assessment by Hess et al highlighted

the benefit of visual inspection of Schoenfeld plots to assess the integrity of the proportional hazards assumption (168).

#### *3.1.1.3 Missing data*

Missing data can be defined as missing completely at random, missing at random or missing not at random (169). Data that is missing completely at random has missing values that are not associated with those that are present and therefore cannot be predicted from them (169). The more common scenario is data that is missing at random where the missingness is associated with the present data and can be predicted from the data available, this allows for the prediction of the missing data from the present data (169). Finally there is missing not at random where the missing data is associated with an unmeasured exposure (169).

When data is assumed to be missing at random multiple imputation can be used to predict or infer the missing data from that the data which is present. Multiple imputation has been shown to have greater statistical power when the outcome variable of interest is included in the calculation to impute the missing data (170).

## 3.2 Aims / objectives

- To identify the feasibility of measuring and using common risk factors for CVD in the Emergency Department setting.
- External validation of prognostic models for long term CVD outcomes

## 3.3 Published work

Reynard, C., McMillan, B., Jafar, A., Heagerty, A., Martin, G.P., Kontopantelis, E. and Body, R., 2022. Long-term cardiovascular risk prediction in the emergency department: a mixed-methods study protocol. *BMJ open*, *12*(4), p.e054311.

## 3.4 Methods

### 3.4.1 Design and setting

I conducted a retrospective cohort study using data from Manchester Royal Infirmary (MRI), a large teaching hospital and a part of Manchester University NHS Foundation Trust. The annual ED attendance for 2021 was 104,449.

### 3.4.2 Study population

I included adult patients (>18 years) who had presented to the ED with a presenting complaint of 'chest pain' recorded by the triage nurse. Patients with an index admission diagnosis AMI were excluded (as per local and national coding).

Three distinct cohorts were identified. First, I included patients between January 1st 2009, to December 31st 2009, because this would enable us to obtain long-term (10-year) follow-up data. Second, I included patients between November 1st 2011 – October 31st 2012 because hs-cTnT assay was introduced in November 2011. This allowed me to study the prognostic value of high-sensitivity cardiac troponin with the longest possible follow-up period. Third, I included patients between July 1st 2016 – June 30th 2017 because the TMACS decision aid was implemented in June 2016. This allowed us to study the prognostic value of the TMACS decision aid for the longest possible follow-up period.

### 3.4.3 Data sources

I collated data from MRI from six different internal databases. This data was then linked by NHSD with their outcome data from the Hospital Episode Statistics database – Admitted Patient Care (HES APC) and the Civil Registry. The HES APC database records all diagnostic and intervention codes from every patient admission in the NHS, as well as some baseline demographic data. The civil registry records deaths, including cause and date of death. In compliance with data minimisation principles from NHSD only episodes with outcomes matching our definition of a cardiovascular event were returned. NHSD returned data for episodes up to ten years after the index event date (see table 3.1).

| **Outcome Data** | **Source** |
|---|---|
| Index event Date Time | Local EPR |
| Index event ICD-10 codes | Local EPR |
| Index event ICD-10-OPC codes | Local EPR |
| Subsequent event Date Time | HES APC |
| Subsequent event ICD-10 codes | HES APC |
| Subsequent event ICD-10-OPC codes | HES APC |
| Subsequent event treatment specialty | HES APC |
| Date of death | CR |
| Cause of death | CR |
| | |
| **Predictor Data** | |
| Age | Local EPR |
| Gender | Local EPR |
| Ethnicity | Local EPR/HES APC |
| Physiological observations | Local EPR |
| Triage Data | Local EPR |
| Time of departmental events | Local EPR |
| Laboratory investigations | Local EPR |
| TMACS data | Local EPR |
| Rural/urban indicator | HES APC |
| Indices of Deprivation | HES APC |

*Table 3.1 Data variables to be collected. EPR – local electronic patient record, HES APC – NHS Digital's Hospital Episode Statistics Admitted Patient Care dataset, CR – NHS Digital's Civil Registry , TMACS – troponin only acute coronary syndrome diagnostic algorithm (includes BP, sweating, crescendo angina, ECG ischaemia, troponin, pain radiating to the right arm or shoulder.)*

### 3.4.4 Outcomes

The outcome of cardiovascular disease was defined as a composite including international classification of disease -10 (ICD-10) coded diagnoses of angina pectoris, myocardial infarction, coronary artery revascularisation, ischaemic heart disease, atraumatic stroke, transient ischaemic attack, and cardiovascular mortality. The relevant coded diagnoses from the ICD-10 (171) were: I20-24, I60-64 and G45.9. The intervention codes from the office of population census and surveys version four (OPCS-4) were K40-50, K63 and K75 (172). Recorded deaths with a cause of death from the aforementioned diagnostic codes was deemed to be a cardiovascular disease death.

### 3.4.5 Sample size calculation

I ran a sample size calculation for the multivariable model that was to be used to provided adjusted hazard ratios for the predictors. This was conducted per the methodology by Riley et al (173,174). To populate the calculation, as the method suggests, potential model characteristics were extrapolated from a similar model. I used Q-Risk3 as a source of model parameters for the sample size calculation, specifically the AUC of 0.858 and incidence per 1,000 person years at 61.9 (19). This assumes a similar performance and incidence to that of the base model Q-Risk3. This was used to calculate the required sample size to derive a survival model with ten candidate predictors in each cohort. The maximum sample size was 3,255 patients. A further calculation was conducted to estimate the sample size required to derive a survival model with 20 candidate predictors. The sample size required per cohort 1 and 2 was 5580 and cohort 3 was 6509. The sample per cohort was expected to exceed this.

### 3.4.6 Statistical analysis

R and R studio were used for the statistical analysis, and the packages are outlined in the supplementary appendix (175,176).

I used descriptive statistics to summarise the cohorts, including either mean, standard deviation and range, or median and range. I selected candidate variables based on data availability and those that

were available in routine clinical care for the prognostic factor study. I sought to consolidate predictors based on collinearity to reduce the number of variables in the model.

#### *3.4.6.1 Missing data*

I prioritised the potential clinical implementation of any individual prognostic factor or CPM in this work. The source of the derivation data was the same digital eco-system where any models or prognostic factor prediction would be clinically used. As such any data missing in the derivation data would likely be missing in the clinical environment where a CPM would be used. I excluded data from model development that had missingness greater than 75%. This would also simplify implementation as these missing variables would have otherwise create a burden from imputation methods that may hamper any practical implementation. Variables with less than 75% missingness were included and missing variables were imputed with multiple imputation that included the outcome as this has been shown to benefit the performance (170). I used the argeImpute function from the Hmisc package for the imputation which utilises additive regression (177). Informative missingness was examined where there was a clinically plausible rationale such as a missing laboratory test being caused by a clinical decision not to conduct it (e.g. troponin). This was achieved by categorising the variable and adding a "missing" category. The variables representing the diagnostic innovations for which the cohorts were chosen were included regardless of the amount of missing data, this was due to likely informative missingness and that they were a focus of this work. This included conventional troponin T (cTnT), high sensitivity troponin T (hs-cTnT) and TMACS. eGFR has previously been shown to be strongly associated with raised cardiac troponins, so as not to miss this important interaction it was also included (178).

#### *3.4.6.2 Adjustment Model Derivation*

I utilised Cox proportional hazard (CPH) models for the prognostic factor study, the derivation was conducted in R with the Survival package (176,179). I assessed the proportionality assumption held with Schoenfeld plots (180). If the assumption was invalid, then I had planned to use time-varying interactions or flexible parametric survival models. Prior to inclusion in the model, I assessed the variables distribution by Joanes et al's skew metric, histograms and quantile-quantile plots (181). Variables were subsequently transformed if they were found to be significantly skewed, with improvements sought in the skewness metric. I assessed linearity between the predictor of interest and outcome using martingale plots and outcome vs predictor plots. If the martingale plots suggested non-linearity then a restricted cubic spline was considered. To assess if a restricted cubic spline (RCS) was preferrable, a univariable CPH model was created with the variable of interest (transformed if necessary), this was compared to univariable CPH models with the variable incorporated into a RCS with three, four and five knots. The models were assessed on Akaike information criterion (AIC) metric and Bayesian information criterion (BIC), given the loss of power with the increased number of knots only a significant improvement in AIC/BIC was deemed to merit inclusion of a more complicated RCS (more knots). Log odds plots were used to further assess if there was a significant difference between knots.

Co-linearity was assessed by visual inspection, and for continuous variables Pearson correlation coefficients. Variables that were to be included in the model that appeared correlated on visual and statistical inspection were considered for a colinear terms where a biologically plausible mechanism was apparent. This was supplemented by assessing variance inflation factors within the multivariable models.

#### *3.4.6.3 Prognostic Factor Evaluation*

For each prognostic factor I reported adjusted hazard ratios by using a multivariable CPH model, I planned for three adjustment methods. Firstly, using a single data set with imputed data, I derived a multi-variable model with variable selected by backwards selection on the AIC. In the second adjustment scenario the same multi-variable model was used but only cases where the predictor of interest was originally present (not imputed) were used. Finally, in the third adjustment, a multivariable model was created with all available variables included. For each adjustment model we reported discrimination (ability to differentiate cases from controls) with c-statistics and calibration (agreement between the observed and expected event rates) with intercept, slopes and calibration plots. I planned to use forest plots and survival curves to visualise the effect of each prognostic factor. Hazard ratio plots were to be used to visualise the variable effect of a non-linear predictor transformed with a restricted cubic spline.

#### *3.4.6.4 External Model Validation*

I considered four models considered for external validation in the emergency care population using the same database. They were Framingham, QRisk 1, QRisk 2, and ASSIGN (56,57,182,183). Once the database was compiled, I assessed which model the database had the least missing data for. That model was then validated. I considered the use of national or regional data sources to provide averages for variables that were missing from the study database.

I reported model performance in the domains of discrimination and calibration. For discrimination I reported AUC (or concordance) and for calibration I plotted calibration curves.

## 3.5 Results

### 3.5.1 Demographics

During data extraction there was substantial missing data noted at the beginning of cohort one (Jan 2009). After discussing with local data analysts, it transpired that the electronic patient record was replaced during this period which may have resulted in missing data. I opted to move the cohort window forward by 10 months to start in November 2009, this placed the cohort well beyond any disruption caused by EPR updates. Furthermore, due to delays from the pandemic and NHS Digital the data extract was not available as planned, this delay enabled ten years of follow-up data to be collected for the baseline cohort and up to 9.16 years for the hs-cTnT cohort.

The numbers of patients in cohorts 1-3 were 6055, 6089 and 6846 respectively, fulfilling our sample size criteria (table 4.2). In addition to the variables that represented the diagnostic innovations, routinely collected variables with less than 75% missing data (averaged across the three cohorts) were considered for inclusion. By these criteria, the following variables were included: age, gender, ethnicity, heart rate, systolic blood pressure, diastolic blood pressure, respiratory rate, Glasgow coma scale, temperature, oxygen saturation, early warning score, haemoglobin, white blood cell count, index of multiple deprivation, rural urban index and time of presentation (Figure 3.1). The distribution of the missingness demonstrated that the laboratory blood test results were frequently clustered, along with physiological observations (Figure 3.2). The time element of the presentation date-time variable was corrupted for the most recent 14.52% of patients. The source of the error was related to the data export formatting and did not affect the date aspect of the variable.

The demographics of each group were similar with the exception of ethnicity varying between cohorts, the proportion of patients who were white decreased from 51.30% to 44.73% whilst the proportion of patients who were of mixed ethnicity increased from 4.95% to 8.24%. The proportion of patients who were of Asian ethnicity fluctuated from 20.66%, to 23.19% to 18.03% year on year (Table 3.2 and Supplementary Table 8.4).

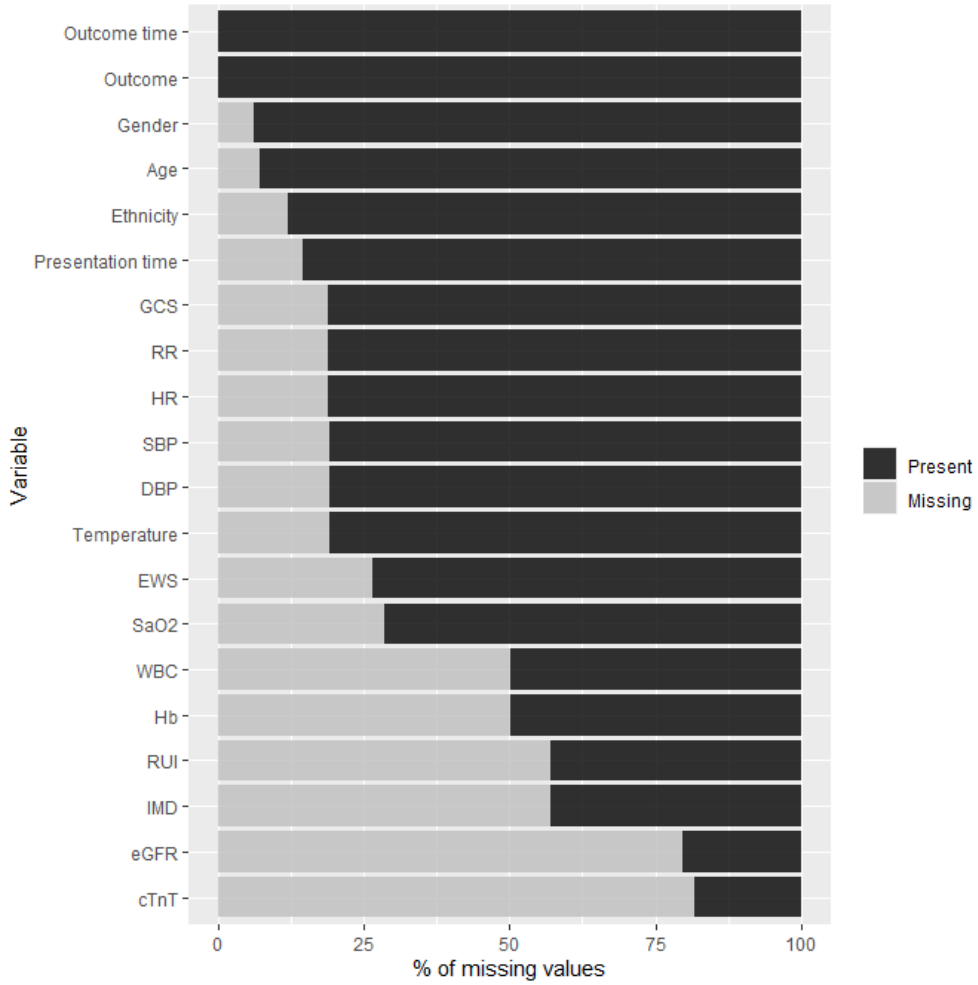



*Figure 3.1 Proportion missing for baseline variables. GCS – Glasgow comma scale, RR – respiratory rate, HR- heart rate, SBP – systolic blood pressure, DBP – diastolic blood pressure, EWS – early warning score, SaO2 – peripheral oxygen saturation, WBC – white blood cell count, Hb – haemoglobin, RUI – rural urban index, IMD – index of multiple deprivation, eGFR – estimated glomerular filtration rate, cTnT – cardiac troponin T*

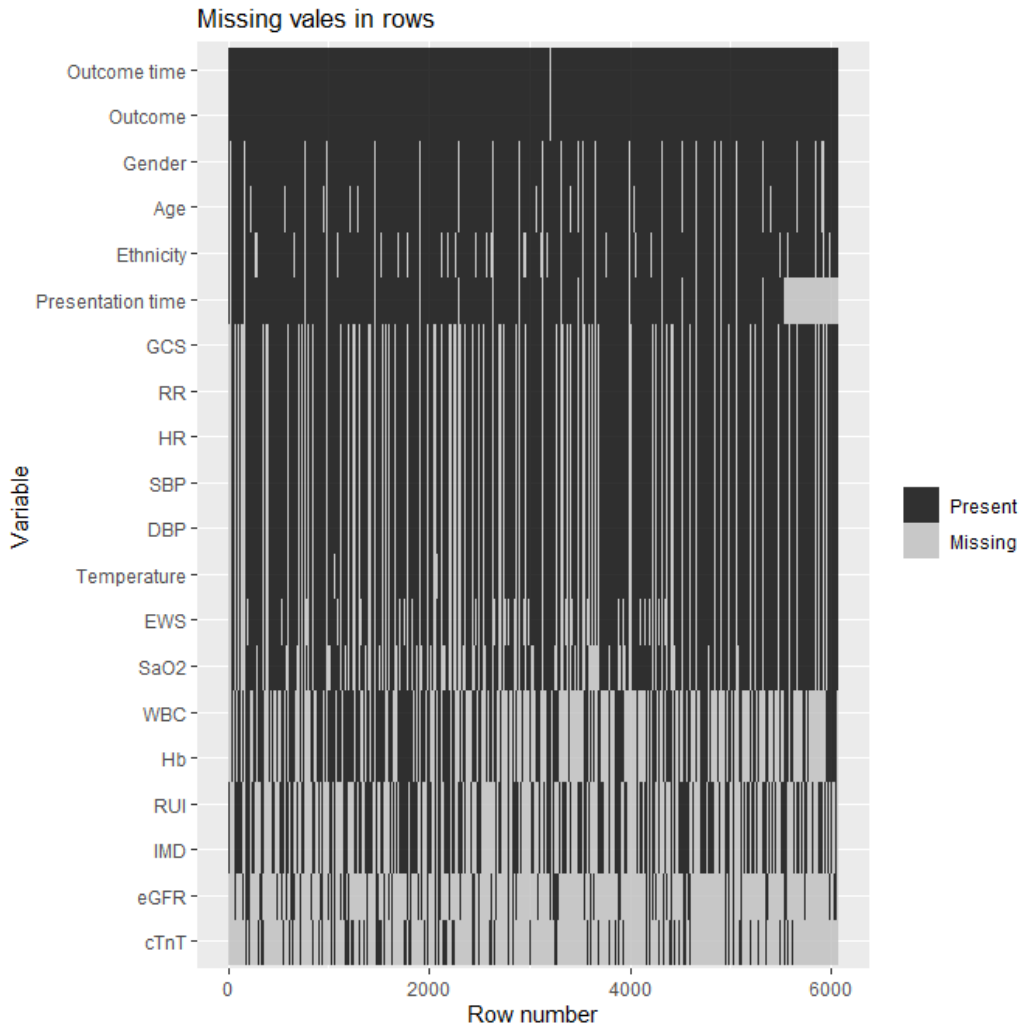



*Figure 3.2 Distribution of missing data for baselines variables. GCS – Glasgow comma scale, RR – respiratory rate, HR- heart rate, SBP – systolic blood pressure, DBP – diastolic blood pressure, EWS – early warning score, SaO2 – peripheral oxygen saturation, WBC – white blood cell count, Hb – haemoglobin, RUI – rural urban index, IMD – index of multiple deprivation, eGFR – estimated glomerular filtration rate, cTnT – cardiac troponin T*

| | **Cohort 1** | **Cohort 2** | **Cohort 3** |
|---|---|---|---|
| Time window | Nov 2009 – Nov 2010 | Nov 2011 – Nov 2012 | July 2016 – July 2017 |
| | | | |
| Total | 6,055 | 6,089 | 6,846 |
| | | | |
| Age (years) | | | |
| Mean | 45.3 | 43.4 | 44.5 |
| SD | 18.1 | 17.4 | 17.7 |
| Missing (n) | 14.5% (879) | 8.8% (534) | 7.0% (478) |
| | | | |
| Male | | | |
| Percentage (n) | 53.7% (3249) | 52.4% (3190) | 51.8% (3543) |
| Missing | 6.2% (378) | 7.4% (453) | 6.1% (419) |
| | | | |
| Ethnicity | | | |
| Asian (n) | 20.7% (1251) | 23.2% (1412) | 18.0% (1234) |
| Black (n) | 9.8% (595) | 11.4% (696) | 10.5% (719) |
| Mixed (n) | 5.0% (300) | 5.3% (321) | 8.2% (564) |
| Other (n) | 6.1% (371) | 7.1% (434) | 4.6% (316) |
| White (n) | 51.3% (3106) | 46.8% (2847) | 44.7% (3062) |
| Missing (n) | 7.1% (432) | 7.2% (439) | 7.5% (516) |
| | | | |
| cTnT (ng/ml) | | | |
| Mean | 0.1 | | |
| SD | 0.5 | | |
| Missing (n) | 81.9% (4956) | | |
| | | | |
| hs-cTnT (ng/L) | | | |
| Mean | | 11.6 | 12.4 |
| SD | | 65.4 | 67.1 |
| Missing (n) | | 62.7% (3788) | 41.7% (2852) |
| | | | |
| TMACS (%) | | | |
| Mean | | | 8.36 |
| SD | | | 17.07 |
| Missing (n) | | | 80.7% (5522) |
| | | | |
| Early Warning Score | | | |
| Median | 1 | 0 | 0 |
| SD | 0 – 10 | 0 – 10 | 0-11 |
| Missing (n) | 26.5% (1604) | 12.0% (732) | 9.1% (623) |
| | | | |

*Table 3.2 Demographic characteristics of cohorts. cTnT – cardiac troponin T, hs-cTnT – high sensitivity cardiac troponin T, TMACS – troponin only Manchester acute coronary syndrome rule out strategy*

### 3.5.2 Outcome Incidence

The incidence of the primary outcome (incident cardiovascular disease) steadily declined between cohorts, from 10.49% in cohort 1 to 5.30% in cohort 3. The declining length of follow up could be an explanation (Table 3.3). However, as the median time to event for the 10-year cohort was 2.05 years, and the most recent cohort had a follow up period of 4.43 years it might be expected that the majority of outcomes would have occurred. Between the 1st and 2nd cohort this change appears to be driven by a reduction in cardiovascular death and ICD-10 criteria. The ICD-10 codes that qualified a patient for the outcome varied over time (Table 3.4). “Unstable angina”, I20.0, was the most prevalent qualify ICD-10 code in cohort 1 but was third by cohort 3. “Angina Pectoris”, I20.9, started as the second most frequent qualifying code but in cohort 2 and 3 was the most frequent code. Similarly, “non-ST elevation myocardial infarction”, I21.4, was fourth in the rankings of qualifying codes in cohort 1 but was second by cohort 3. The OPCS intervention codes were more stable over time (Table 3.5). “Percutaneous transluminal balloon angioplasty and insertion of 1-2 drug-eluting stents into coronary artery”, K75.1, was the most frequent qualifying code for the intervention criteria across all cohorts. “Anastomosis of mammary artery to left anterior descending coronary artery” (K45.3), and “percutaneous transluminal balloon angioplasty and insertion of 3 or more drug-eluting stents into coronary artery” (K75.2) were consistently in the top 3, although their order varied.

| | Cohort 1 | Cohort 2 | Cohort 3 |
|---|---|---|---|
| Major adverse cardiac event (n) | 10.5% (635) | 7.8% (475) | 5.3% (363) |
| ICD-10 criteria | 8.1% (491) | 6.2% (380) | 4.2% (290) |
| OPCS criteria | 2.6% (156) | 2.4% (144) | 1.5% (101) |
| Death criteria | 0.7% (45) | 0.4% (25) | 0.2% (14) |
| Maximum follow up (years) | 10.0 | 9.2 | 4.4 |
| Median time to event (years) | 2.1 | 2.3 | 1.6 |

*Table 3.3 Outcome incidence. ICD-10 - International Classification of Diseases, Tenth Revision, OPCS - Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures*

| | Cohort 1 | Cohort 2 | Cohort 3 |
|---|---|---|---|
| Position (n) | | | |
| 1 | Unstable angina (151) | Stable angina (131) | Stable angina (117) |
| 2 | Stable angina (147) | Unstable angina (99) | NSTEMI (64) |
| 3 | AMI (45) | NSTEMI (54) | Unstable angina (42) |
| 4 | NSTEMI (36) | AMI (27) | Acute ischaemia (29) |
| 5 | Acute ischaemia (18) | Acute ischaemia (20) | STEMI (12) |

*Table 3.4 Top 5 ICD-10 outcome codes for positive cases. Stable Angina = angina pectoris, NSTEMI = Non-ST elevation (NSTEMI) myocardial infarction, AMI = Acute myocardial infarction, unspecified, Acute ischaemia = Acute ischemic heart disease, unspecified, STEMI = ST elevation (STEMI) myocardial infarction of inferior wall.*

| | Cohort 1 | Cohort 2 | Cohort 3 |
|---|---|---|---|
| Position (n) | | | |
| 1 | PCI + DES (84) | PCI + DES (84) | PCI + DES (62) |
| 2 | CABG (22) | PCI + >3 DES (32) | CABG (13) |
| 3 | PCI + >3 DES (22) | CABG (18) | PCI + >3 DES (8) |
| 4 | PCI + stents (7) | * | * |
| 5 | * | * | * |

*Table 3.5 Top 5 OPCS outcome codes for positive cases. PCI (percutaneous coronary intervention) + DES = Percutaneous transluminal balloon angioplasty and insertion of 1-2 drug-eluting stents into coronary artery, PCI + >3 DES = Percutaneous transluminal balloon angioplasty and insertion of 1-2 stents into coronary artery, PCI + stents = Percutaneous transluminal balloon angioplasty and insertion of 1-2 stents into coronary artery. * indicates censoring for small groups as per NHS Digital policy.*

### 3.5.3 Variable preparation

The potential variables were examined for distribution, linearity with the outcome, and co-linearity. In addition, the variables were assessed for the potential opportunity to cluster or reduce their number. This process laid out in the methods resulted in 8 variables undergoing transformation, including 5 using a RCS (due to non-linearity with the outcome) and 6 being removed through data clustering. This is summarised in Table 3.6 and the supplementary appendix section 3.6.1.

| Variable | Transformation | Data clustering | RCS |
|---|---|---|---|
| Age | | | 4 knots |
| Heart Rate | Logarithmic | | |
| SBP | Logarithmic | | 3 knots |
| DBP | | Clustered | |
| Temperature | | Clustered | |
| Respiratory Rate | Logarithmic | Clustered | |
| GCS | | Clustered | |
| $SaO_2$ | Exponential | Clustered | |
| EWS | Square root | | |
| WBC | Logarithmic | | 3 knots |
| Haemoglobin | | | |
| Ethnicity | Categorical | | |
| P-time | Categorised | | |
| IMD | | | |
| RUI | | | |
| eGFR | Cube | | |
| cTnT | Logarithmic | | |
| hs-cTnT | Logarithmic | Partially | 5 knots |
| TMACS | Logarithmic | | 5 knots |

| Co-linear terms |
|---|
| Age & SBP |
| Hb & eGFR |
| cTnT & RUI |
| hs-cTnT & age |
| hs-cTnT & eGFR |
| TMACS & age |
| TMACS & eGFR |

*Table 3.6 Summary of candidate predictors and their transformations and clustering – grey indicates exclusion after clustering. RCS- restricted cubic spline, SBP – systolic blood pressure, DBP – diastolic blood pressure, GCS – Glasgow comma scale, $SaO_2$ – oxygen saturations, EWS – early warning score, p-time – presentation time, IMD – index of multiple deprivation, RUI – rural urban index, eGFR – estimated glomerular filtration index, cTnT – cardiac troponin T, hs-cTnT – high sensitivity cardiac troponin T, TMACS – troponin only Manchester acute coronary syndrome algorithm*

### *3.5.3.1 The proportional hazard's assumption*

The proportional hazard assumption for each variable was examined with Schoenfeld plots, no variable's Beta(t) varied significantly over time. The Beta(t) always fell within the 95% confidence interval from the initial time point (Figure 3.3). The proportional hazard assumption was considered to have been maintained. Therefore, no time-varying interactions or flexible parametric survival models were used.

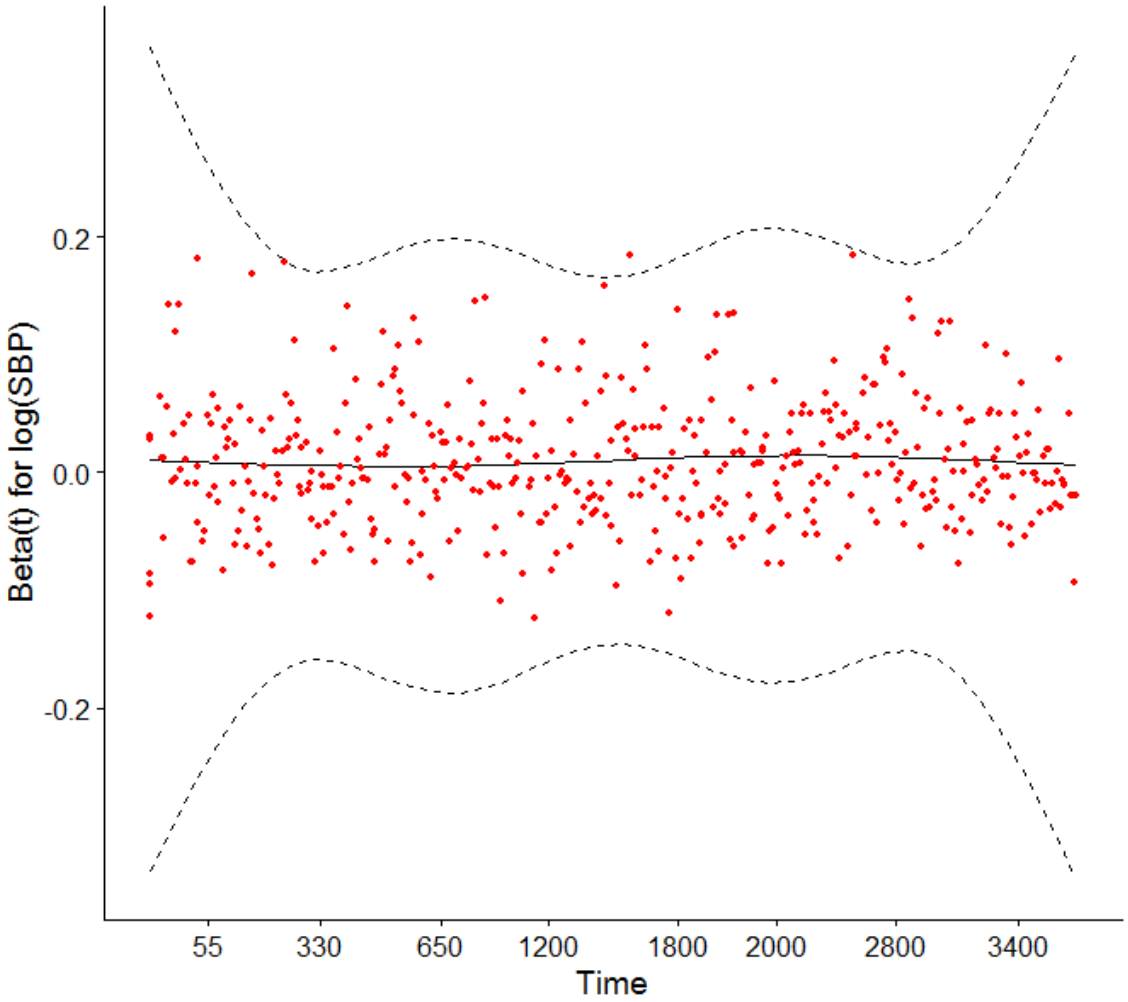



*Figure 3.3 Schoenfeld plot for systolic blood pressure*

### 3.5.4 Missing data

I used multiple imputation ensuring that the outcome was included in the imputation process (170), convergence plots were used to ensure suitability of the imputation. The possibility of informative missingness was considered with the variables that were tests within a care pathway that represented a clinical decision. These were cTnT, hs-cTnT and TMACS. To test for informative missingness these variables were categorised with missing included as the lowest level, the results are presented in the next section.

### 3.5.5 Prognostic factor evaluation

Cox models were derived to enable the production of adjusted estimates of each prognostic factor. Each prognostic factor was adjusted with three methods. The first method used a model derived using forward stepwise selection using the AIC metric where all the missing data was imputed. The second method used the same model but only cases where the candidate predictor was not missing were included (complete cases). The third method used a model that included all variables with no selection and all missing data was imputed.

#### *3.5.5.1 Cohort 1 – Baseline cohort - cTnT*

I derived the model through forwards stepwise variable selection via AIC metric. This model and the variables selected included cTnT, heart rate, WBC, age, gender, ethnicity, and time of presentation (Figure 3.4). The concordance (c-statistic) of the model was 0.79, the $R^2$ was 0.135 and the calibration slope was 1.00 . The variable cTnT had an unexpected hazard ratio of 0.82 (95% CI 0.69 – 0.97) and therefore I conducted an exploratory analysis to better understand this result. 85% of all cTnT values were 0.01 with only 71 results greater than this limit of blank (LoB). The incidence of the outcome was 17.67% (95% CI 15.39% – 20.11%) in those patients with a cTnT equal to the LoB and 12.68% (95% CI 5.96% - 22.7%) in those with a cTnT greater than the LoB. I also explored removing the unstable angina ICD-10 code (I20.9), however this did not change the hazard ratio significantly (HR 0.85 95%CI 0.71 – 1.00).

The hazard ratios from each multivariable model (adjustment scenario) varied slightly, the variables which were consistently statistically significant were gender, the WBC spline variables, part of the age spline, Asian ethnicity, other ethnicity (Table 3.7). Female gender had an adjusted hazard ratio of 0.57 (95% CI 0.48-0.67) demonstrating an association with the absence of the outcome (see survival curve – Figure 3.5). The same was true of the "Other" level of the categorical variable ethnicity with an adjusted hazard ratio of 0.63 (95% CI 0.41 - 0.97) (Figure 3.6). In contrast the level "Asian" was associated with the outcome through an adjusted hazard ratio of 1.25 (95% CI 1.04-1.51). The co-efficients for the RCS transformed variables are difficult to interpret as such I produced relative hazard plots which demonstrates the different hazard ratio produced across the variable range. Figure 3.7 demonstrates the varying hazard ratio for age as part of the stepwise selected multivariable model. There is a steep exponential increase in the relative hazard to 60 years of age which then decreased in shallow linear gradient. The varying hazard ratio for WBC due to the RCS transformation is demonstrated in Figure 3.8. The hazard ratio peaks over 2.0 with a log(WBC) value of 2.0 (exponentiated 7.4 $x10^9$/L), which is in the middle of the normal reference range. The hazard ratio decreases significantly above and below this value. The decreasing hazard ratio for cardiovascular events above a log(WBC) value of 2.0 may be a product of censoring. Patients with a high WBC may have an infection and possibly sepsis. As the mortality for sepsis has been estimated to be 24.4% some of these patients may have been censored (184).

Evening time of presentation was predictive across 2/3 adjusted scenarios with only the complete case scenario not finding statistical significance. cTnT was statistically significant in one adjustment scenario but as previously mentioned the HR of less than one was unexpected. None of the interaction terms were statistically significant. Heart rate, whilst not statistically significant, was selected for inclusion by the stepwise method indicating that it improved model performance. This is re-enforced by a p value that was just greater than the 0.05 threshold (0.05214).

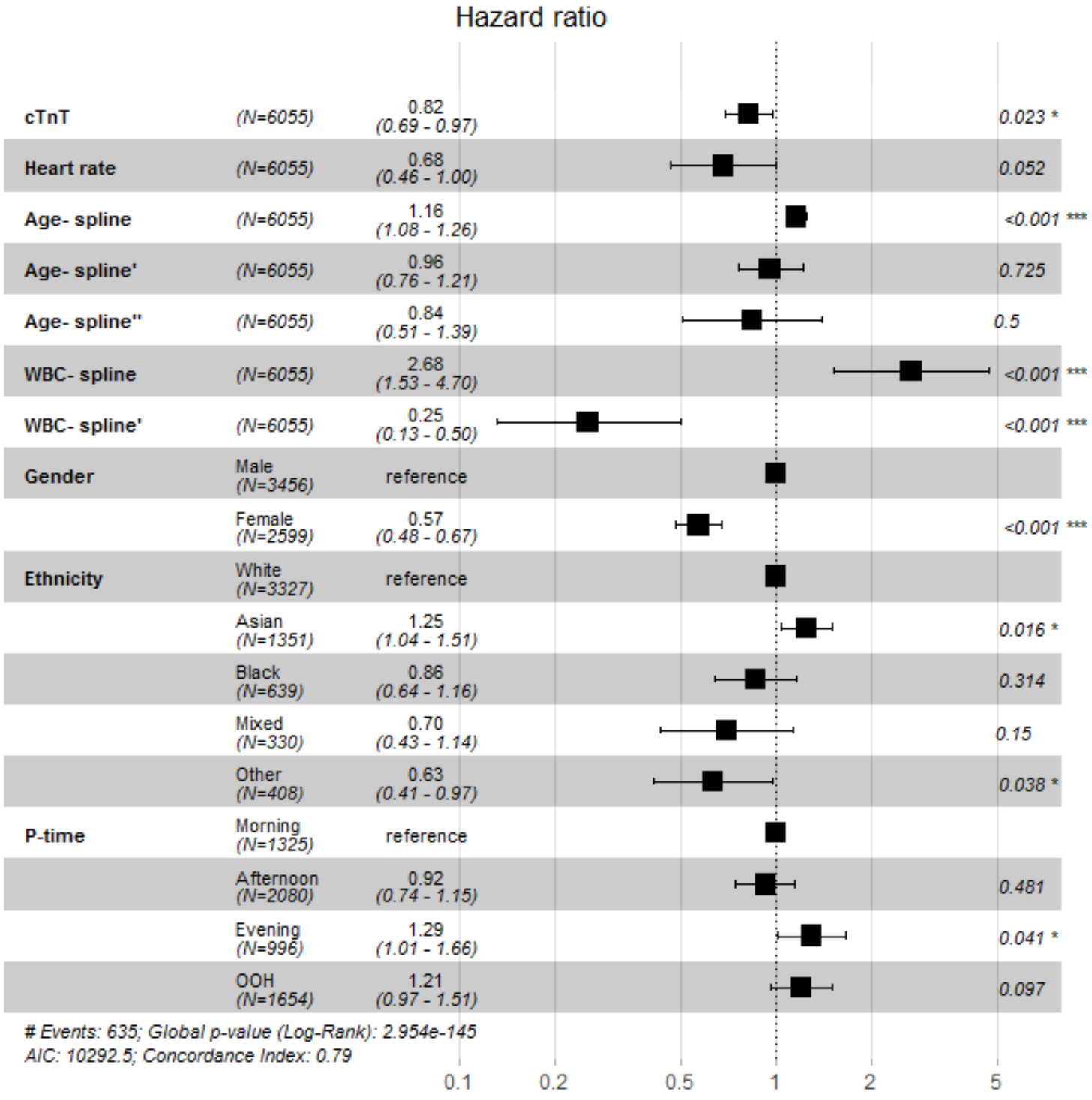



*Figure 3.4 - Forest plot of model selected by AIC criteria from cohort 1 – cTnT – cardiac troponin T, and P-time – presentation time*

| | Hazard Ratio | | |
|---|---|---|---|
| **Variable** | **Adjusted by selected model** | **Adjusted by selected model with complete cases** | **Adjusted by all-inclusive model** |
| cTnT | 0.82 (0.69 – 0.97)* | 0.71 (0.48 - 1.05) | 0.88 (0.70 – 1.12) |
| Heart Rate | 0.68 (0.46 – 1.00)† | 0.66 (0.43 – 1.02) | 0.68 (0.45 – 1.00) |
| WBC – spline 1 | 2.68 (1.53- 4.70)* | 4.37 (2.06 -9.25)* | 2.71 (1.54 - 4.74)* |
| WBC – spline 2 | 0.25 (0.13 – 0.50)* | 0.13 (0.05 – 0.31)* | 0.25 (0.13 – 0.49)* |
| Age – spline 1 | 1.63 (1.08 – 1.26)* | 1.15 (1.07 – 1.24)* | 1.16 (1.08 – 1.25)* |
| Age – spline 2 | 0.96 (0.76 – 1.21)* | 1.00 (0.79 – 1.26) | 0.96 (0.76 – 1.22) |
| Age – spline 3 | 0.84 (0.51 – 1.39) | 0.77 (0.47 – 1.27) | 0.83 (0.50 – 1.37) |
| Gender | | | |
| Male - reference | - | - | - |
| Female | 0.57 (0.48 – 0.67)* | 0.57 (0.49 – 0.68)* | 0.56 (0.47 - 0.67)* |
| Ethnicity | | | |
| Asian | 1.25 (1.04 – 1.51)* | 1.25 (1.04 – 1.51)* | 1.25 (1.03 – 1.51)* |
| Black | 0.86 (0.64 – 1.16) | 0.88 (0.65 -1.18) | 0.84 (0.62 – 1.14) |
| Mixed | 0.70 (0.43 -1.14) | 0.76 (0.47 – 1.24) | 0.70 (0.43 – 1.14) |
| Other | 0.63 (0.41 -0.97)* | 0.62 (0.40 – 0.98)* | 0.64 (0.41 – 0.98)* |
| White – reference | - | - | - |
| Time of presentation | | | |
| Morning - reference | - | - | - |
| Afternoon | 0.92 (0.74 – 1.15) | 0.94 (0.74 - 1.19) | 0.92 (0.74 – 1.14) |
| Evening | 1.30 (1.01 -1.66)* | 1.28 (0.97 – 1.68) | 1.29 (1.00 – 1.65)* |
| OOH | 1.21 (0.97 – 1.51) | 1.27 (1.00 – 1.63) | 1.20 (0.97 – 1.49) |
| EWS | 1.01 (0.89 – 1.15) | 1.03 (0.88 – 1.20) | 1.01 (0.89 – 1.15) |
| SBP – spline 1 | 1.03 (0.31- 3.45) | 0.92 (0.25 -3.38) | 1.03 (0.31 - 3.44) |
| SBP – spline 2 | 0.82 (0.22 – 2.91) | 1.03 (0.26 – 4.12) | 0.81 (0.23 – 2.91) |
| IMD | 1.00 (1.00 - 1.01) | 1.00 (1.00 - 1.01) | 1.00 (1.00 – 1.01) |
| Rural urban index | 1.02 (0.79 – 1.33) | 1.09 (0.83 – 1.42) | 1.04 (0.80 – 1.34) |
| Haemoglobin | 0.99 (0.94 -1.04) | 0.98 (0.92 – 1.04) | 1.04 (0.95 – 1.12) |
| eGFR | 1.00 (1.00 – 1.00) | 1.00 (1.00 – 1.00) | 1.00 (1.00 - 1.00) |
| Interaction terms | | | |
| Hb * eGFR | - | - | 1.00 (1.00 – 1.00) |
| Age * SBP | - | - | 1.01 (0.97 – 1.06) |
| cTnT * RUI | - | - | 0.65 (0.18 – 2.30) |

*Table 3.7 The hazard ratios for each predictor across each adjustment scenario in cohort 1 – * statistical significance with a p-value of <0.05 and † that the variable was selected for the model but was not statistically significant itself. SBP – systolic blood pressure, EWS – early warning score, IMD – index of multiple deprivation, RUI – rural urban index, eGFR – estimated glomerular filtration index, OOH – out of hours cTnT – cardiac troponin T, hs-cTnT – high sensitivity cardiac troponin T, TMACS – troponin only Manchester acute coronary syndrome algorithm*

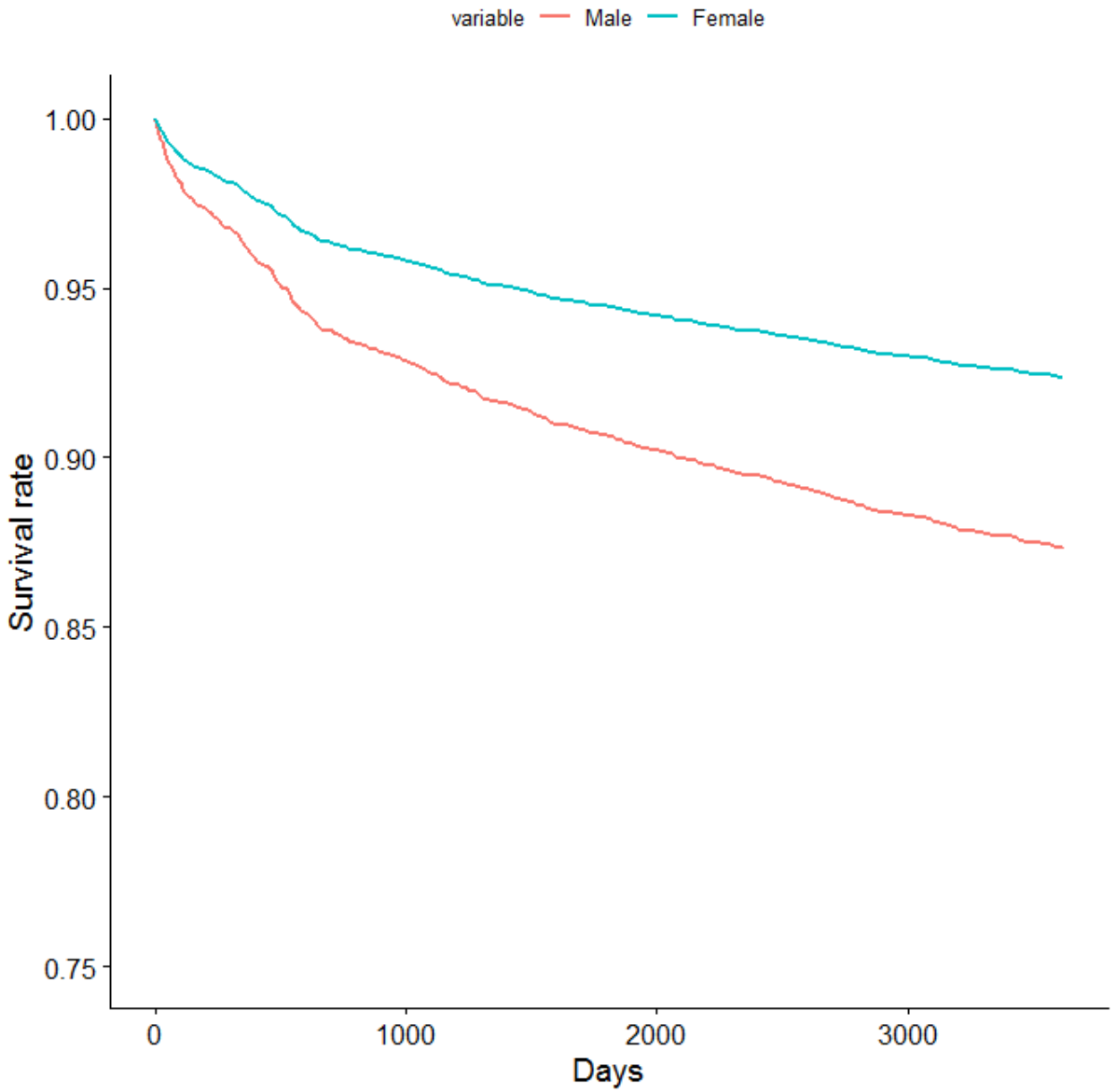



*Figure 3.5 Adjusted survival curve for gender*

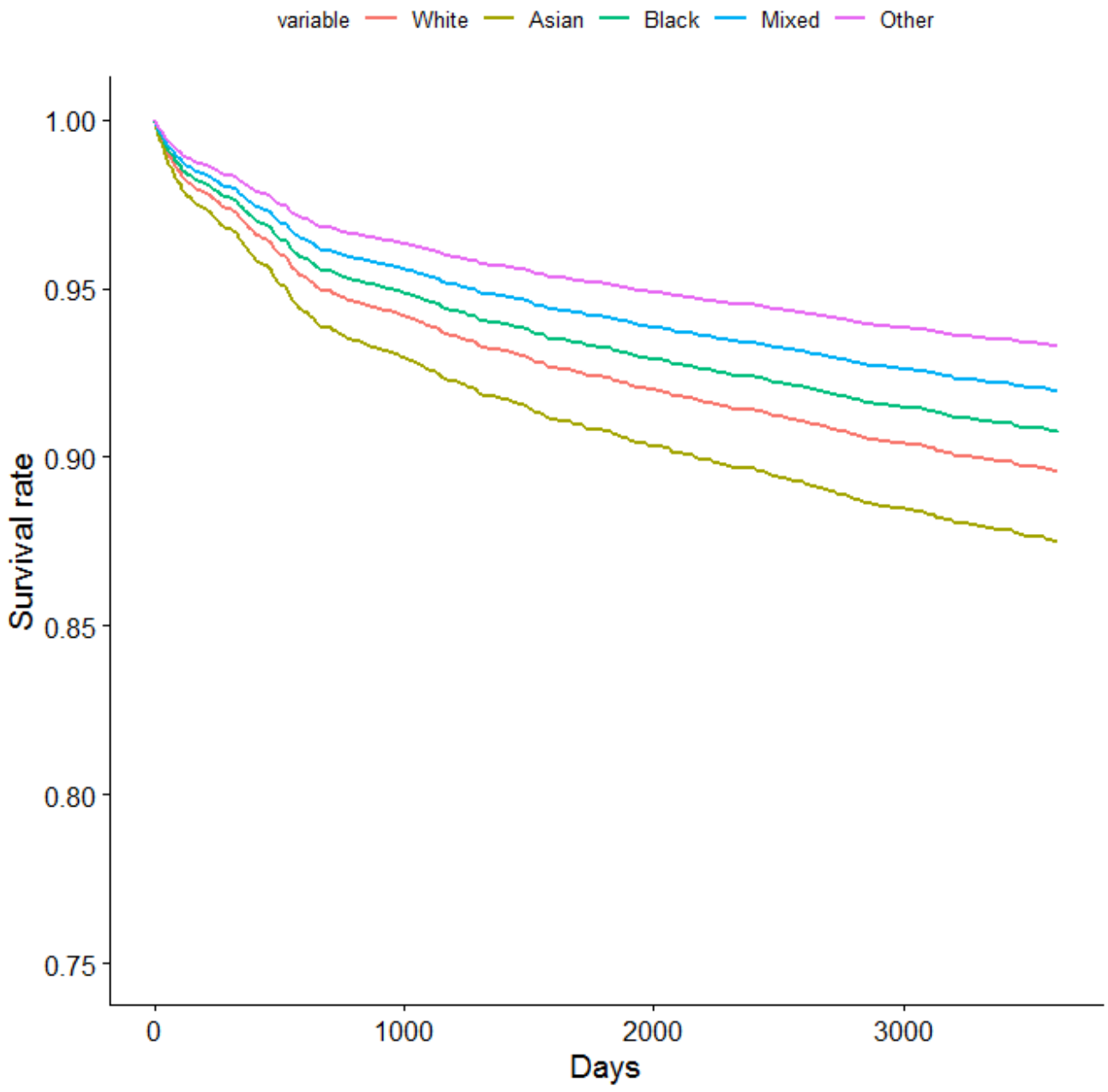



*Figure 3.6 Adjusted survival curve for ethnicity*

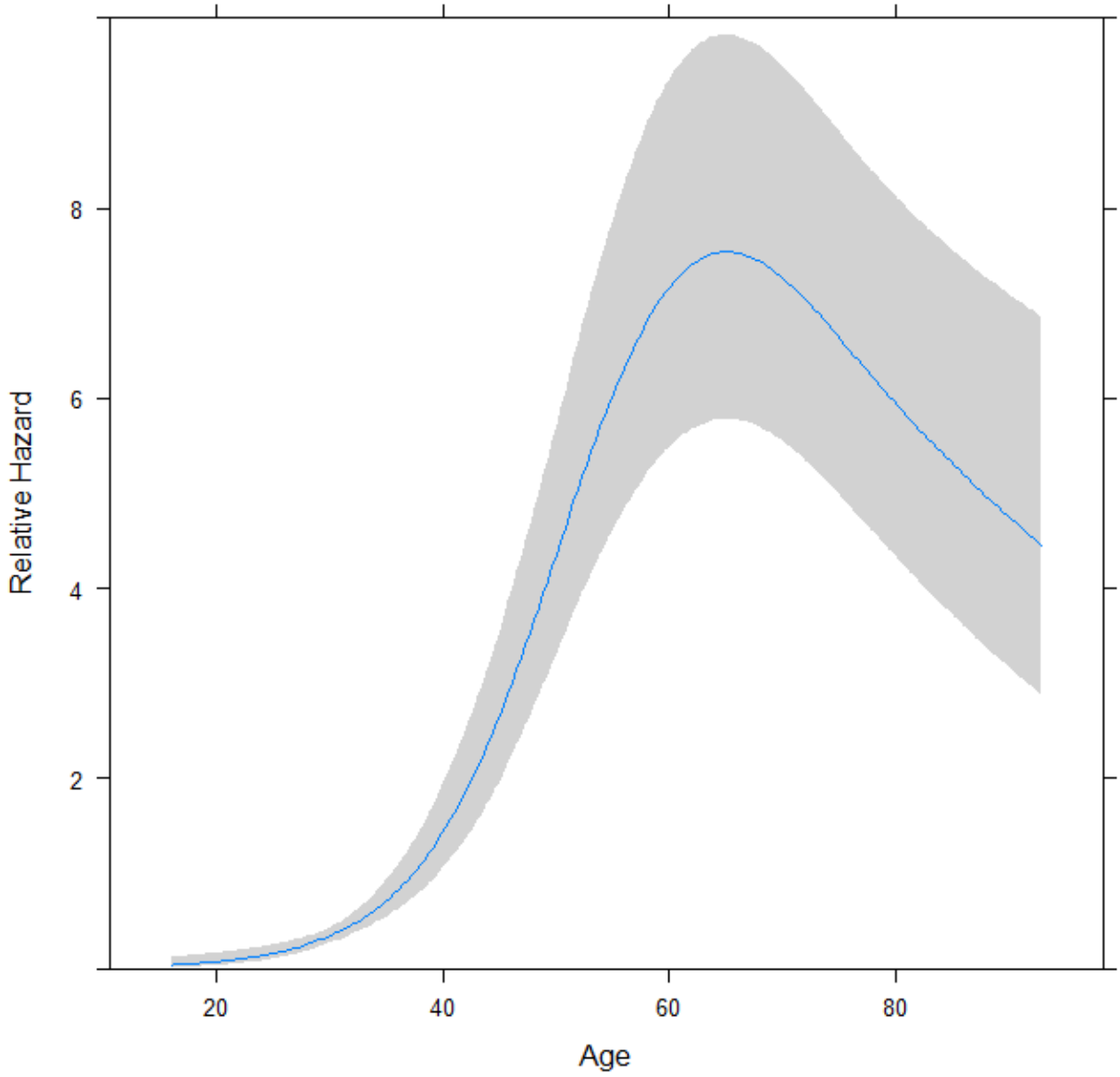



*Figure 3.7 Adjusted elative Hazard plot of the variable age transformed with a restricted cubic spline. The grey area represents the 95% confidence interval. Relative hazard is synonymous with hazard ratio. This is adjusted by the selected model.*

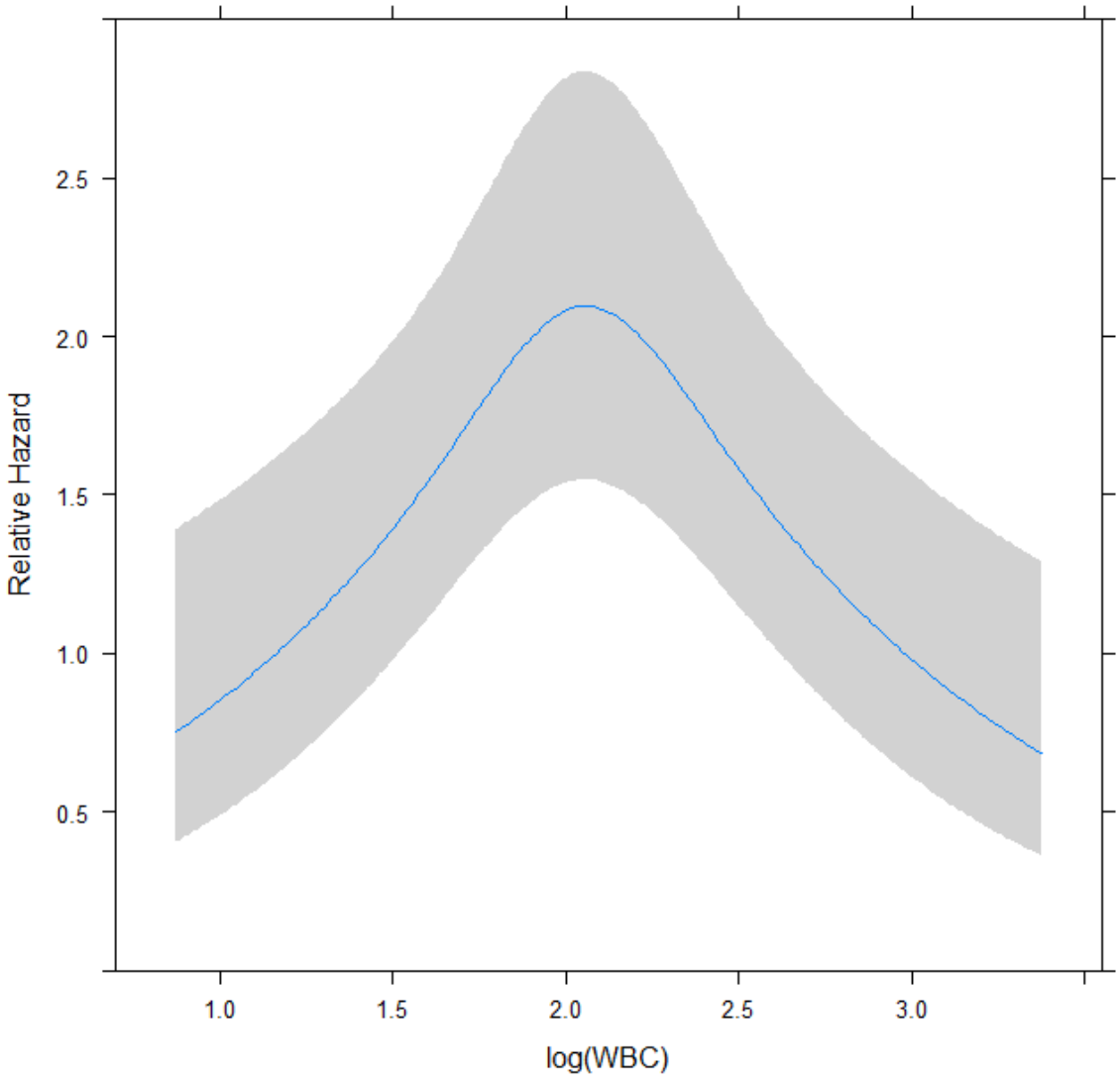



*Figure 3.8 - Adjusted relative Hazard of the variable logarithm of white blood cell count (WBC) passed through a restricted cubic spline. This is adjusted by the selected model.*

The possibility of informative missingness was also explored with cTnT. The variable was categorised into three levels: missing, less than or equal to the limit of blank (LoB) or greater than the LoB (high). LoB was used as the reference level in the analysis which identified that missing data was associated with an absence of an outcome shown by a hazard ratio of 0.73 (95% CI 0.61 – 0.86) (Table 3.8). cTnT that was deemed high did not have a statistically significant hazard ratio.

| | Hazard Ratio | | |
|---|---|---|---|
| Variable | Adjusted by selected model | Adjusted by selected model with complete cases | Adjusted by all-inclusive model |
| cTnT – categorised | | | |
| Missing | - | 0.73 (0.61 – 0.86)* | 0.73 (0.61 -0.87)* |
| LoB - reference | - | - | - |
| High | - | 0.58 (0.30 – 1.14) | 0.62 (0.31 - 1.23) |

*Table 3.8 Hazard ratios of categorised conventional cardiac troponin T (cTnT)*

### *3.5.5.2 Cohort 2 – hs-cTnT*

Again, I derived an adjustment model with forward stepwise selection of variable by AIC metric, this time including the potential variable hs-cTnT which had superseded cTnT from cohort 1. The results of this derivation differed from cohort 1 by only one included factor. Instead of the variable heart rate this new derivation included the interaction term hs-cTnT * age. This interaction term was associated with the absence of the outcome with a hazard ratio of 0.99 (95%CI 0.98 - 0.99) but was only statistically significant in one adjustment scenario (Table 3.9). When inspecting for informative missingness only the missing level of the categorical variable was statistically significant with a hazard ratio of 0.68 (95% CI 0.53 – 0.87). The coefficients of the RCS transformed hs-cTnT are difficult to interpret, whilst the individual coefficients were not consistently statistically significant they were included in the stepwise selection model indicating that they improved model performance. The varying hazard ratio is demonstrated in Figure 3.9, it shows a hazard ratio of <1

inflecting above one after hs-cTnT of 2.72ng/l with a steep gradient till 20.1ng/l where the hazard ratio continues to increase but at a slower rate. The confidence intervals are wide for this RCS, this is despite multiple imputation replacing missing variables. This is due to the clustering of results around the limit of blank, meaning that even when missing data is imputed there are still sparce results at the higher values. Informative missingness was examined by categorising hs-cTnT into missing, limit of detection (<3 ng/l), 99$^{th}$ centile (14 > x ≥ 3ng/l), and high (>14ng/l). A dose response relationship was seen with the hazard ratio climbing with each ascending category (HR 0.68, 0.94, 1.07). Only the missing level of the categorical variable was persistently statistically significant with an adjusted hazard ratio of 0.68 (95% CI 0.53 – 0.87). Interestingly the missing level was the lowest risk group, even more than the LoB group (Figure 3.10). This may indicate that the clinical decision to take a serum hs-cTnT is associated with an increased risk of cardiovascular disease even compared to those patients with an undetectable hs-cTnT.

| | Hazard Ratio | | |
|---|---|---|---|
| Variable | Adjusted by selected model | Adjusted by selected model with complete cases | Adjusted by all-inclusive model |
| | | | |
| hs-cTnT x age | 0.99 (0.98 - 0.99) * | 0.99 (0.99 -1.00) | 0.99 (0.98 -1.00) |
| hs-cTnT – spline 1 | 0.63 (0.19 – 2.09) | 0.69 (0.28 - 1.70) | 0.63 (0.25 – 1.61) |
| hs-cTnT – spline 2 | 1.01 x $10^5$ (4.44 – 2.31x$10^9$)* | 0.69 (0.28 -1.70) | 1.40 x $10^2$ (0.30 – 6.66x$10^4$) |
| hs-cTnT – spline 3 | 4.28 x $10^{-9}$ (1.31x$10^{-16}$ - 0.14)* | 1.55 x$10^{-4}$ (3.35x$10^{-9}$ – 7.14) | 2.81 x $10^{-4}$ (5.71 x $10^{-9}$ - 13.86) |
| | | | |
| hs-cTnT – categorised | | | |
| Missing | - | 0.68 (0.53 -0.87)* | 0.68 (0.53 -0.87)* |
| LoB – reference | - | - | - |
| 99$^{th}$ centile | - | 0.94 (0.72 -1.24) | 0.94 (0.72 -1.24) |
| High | - | 1.07 (0.74 -1.55) | 0.94 (0.64 - 1.37) |

*Table 3.9 Hazard ratios for hs-cTnT across adjustment scenarios. Hs-cTnT – high sensitivity cardiac troponin T, LoB – limit of blank. * denotes statistically significant result*

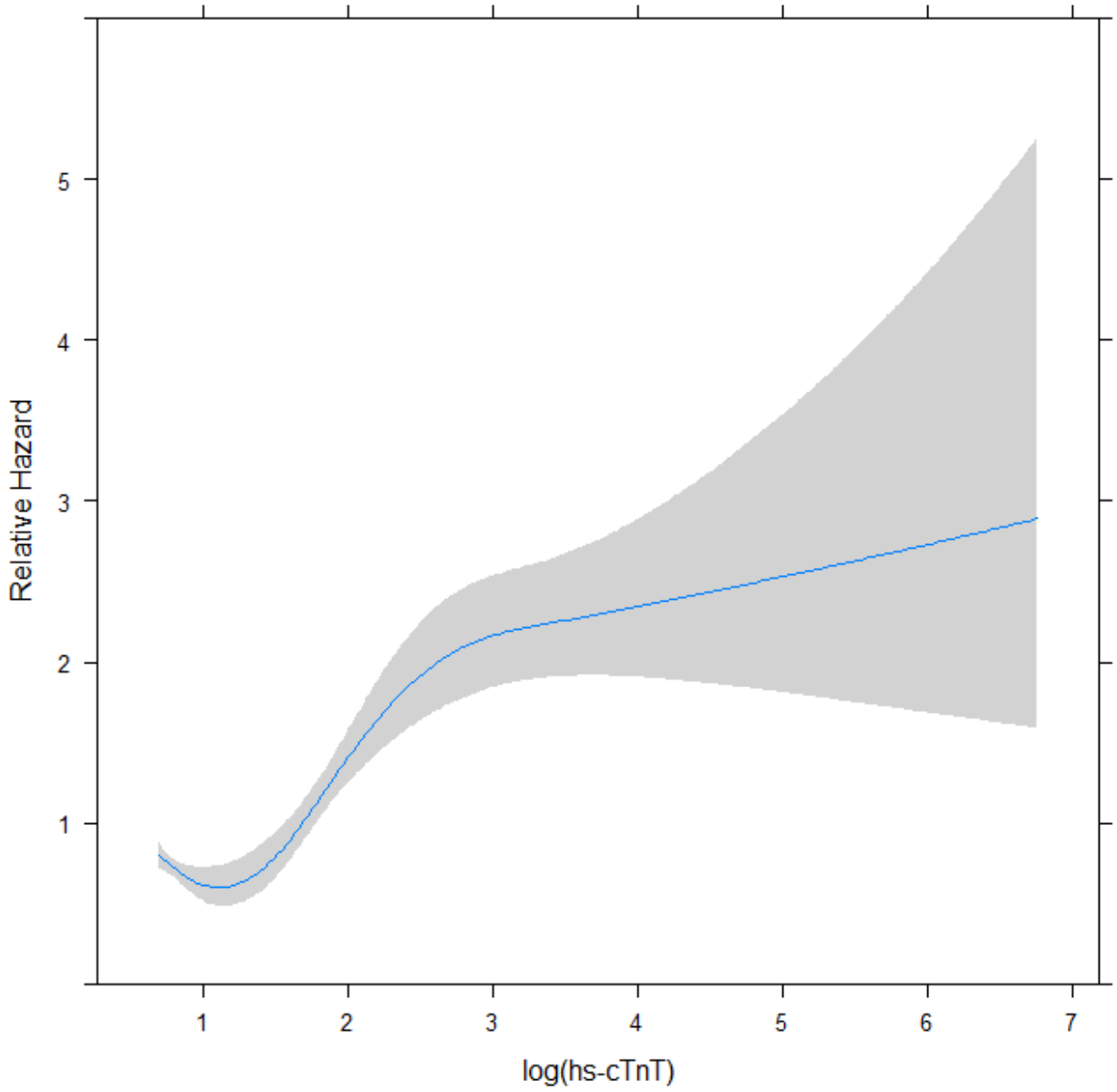



*Figure 3.9 Adjusted relative hazard of the restricted cubic spline transformation of the logarithmic of high sensitivity cardiac troponin T (hs-cTnT). This is adjusted by the selected model.*

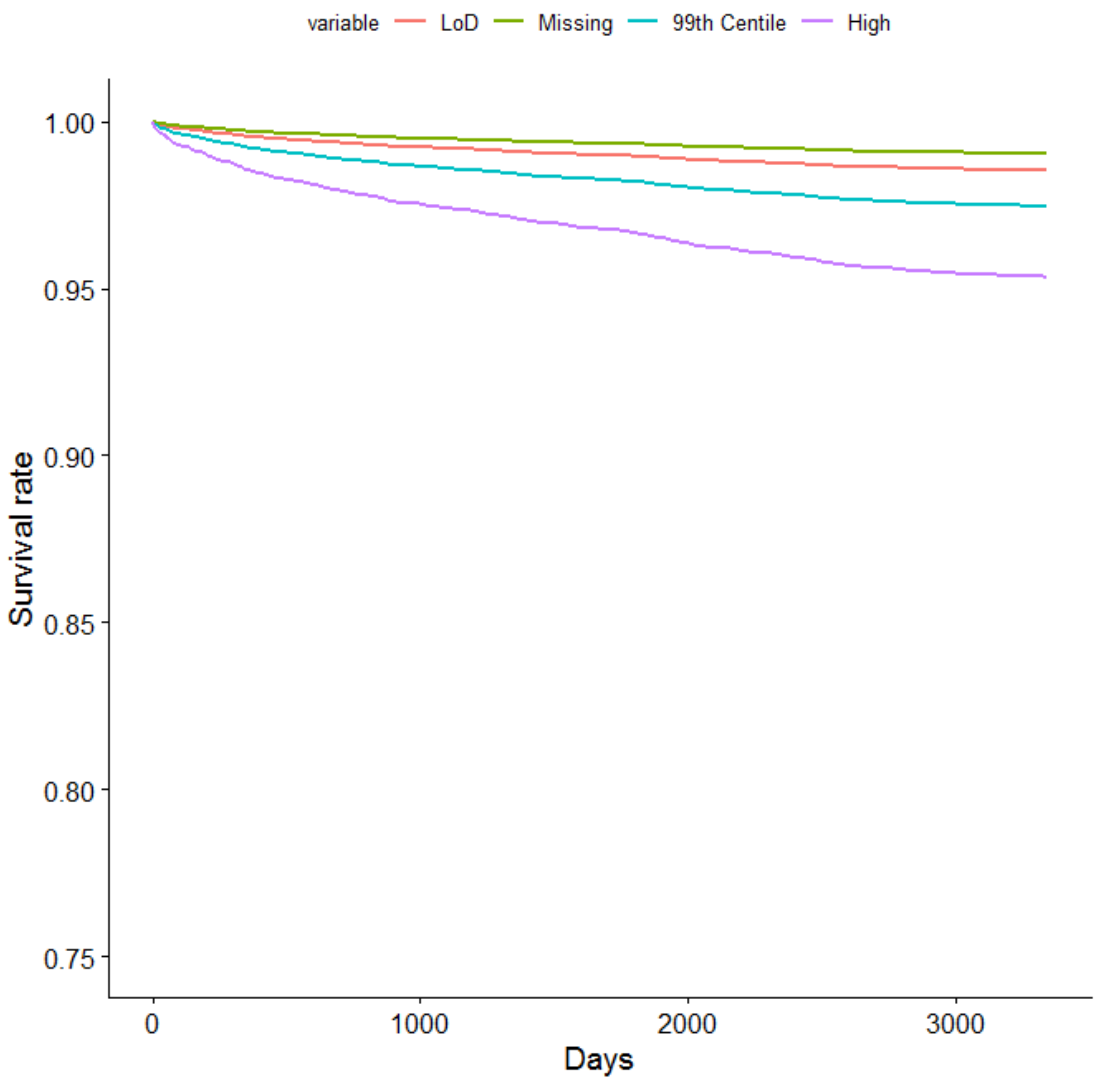



*Figure 3.10 Adjusted survival curve of categorised high sensitivity cardiac troponin T. LoD – limit of detection*

#### *3.5.5.3 Cohort 3 – TMACS*

TMACS was considered as a prognostic variable in cohort 3, again the spline terms are difficult to interpret with varying statistical significance across adjusted hazard ratios (Table 3.10). Furthermore the 95% confidence intervals are very broad implying increased uncertainty even in imputed scenarios. There was instability in the estimates of the spline coefficients across the three adjustment methods (spline term 3 HR 1.25 $x10^{150}$, $5.29x10^{-15}$, $3.69x10^{140}$). The complete case adjustment scenario was persistently discrepant with the other two fully imputed scenarios. This is likely in part due to the high number of knots in the RCS and the low number of complete cases that the imputed data could be based on. Despite this the RCS for TMACS was selected as part of the multivariable model, therefore it has a statistically significant effect as part of the wider model. Furthermore, when the spline terms are examined on a relative hazard plot it is apparent that the confidence intervals are not prohibitively broad in practice (Figure 3.11).

When the TMACS variable was considered for informative missingness it was categorised according to the clinical pathway risk groups (missing, very low, low, moderate and high risk). As with hs-cTnT a dose response relationship was seen with larger hazard ratios for cardiovascular disease with ascending risk groups. The missing level of categorised TMACS was associated with a statistically significant decreased risk of cardiovascular disease with a HR 0.51 (95% CI 0.36 - 0.71). The very low risk level was associated with a statistically significant but lower magnitude of risk reduction with a HR of 0.64 (95% CI 0.42 – 0.98). Neither moderate nor high risk levels were statistically significant as both had hazard ratios spanning 1. Only high risk had a HR point estimate greater than one; HR 1.77 (95% CI 0.54-5.76). This indicates a possible association with an increased risk of cardiovascular disease (Figure 3.12).

| | Hazard Ratio | | |
|---|---|---|---|
| Variable | Adjusted by selected model | Adjusted by selected model with complete cases | Adjusted by all-inclusive model |
| Log(TMACS) | | | |
| Spline term 1 | 231.40<br>(9.93 – $5.4x10^{3}$)* | 3.71<br>(0.03 – 401.50) | $1.59 x10^{2}$<br>(6.70 -$3.79x10^{3}$)* |
| Spline term 2 | $3.78 x10^{-101}$<br>($1.15 x10^{-181}$ – $1.23 x10^{-20}$)* | $5.62 x 10^{6}$<br>($1.96 x 10^{-124}$ – $1.61 x 10^{137}$) | $3.46x10^{-95}$<br>($1.08x10^{-175}$ –$1.11x10^{14}$) |
| Spline term 3 | $1.25 x10^{150}$<br>($2.24 x10^{24}$ – $6.93x10^{275}$)* | $5.29x10^{-15}$<br>($3.76 x10^{-218}$ - $7.43 x10^{188}$) | $3.69x10^{140}$<br>($6.50x10^{14}$ - $2.09x10^{266}$)* |
| Spline term 4 | $7.14 x10^{-50}$<br>($4.66x10^{-97}$ – 0.01)* | 21.07<br>($4.51 x10^{-67}$ - $4.48x10^{84}$) | $4.83x10^{-46}$<br>($2.93x10^{-93}$ - 79.63) |
| TMACS * Age | 1.00 (0.99 - 1.00) | 0.99 (0.97 – 1.00)* | - |
| | | | |
| TMACS – categorised | | | |
| Missing | - | 0.51 (0.36 - 0.71)* | 0.51 (0.36 – 0.71)* |
| Very Low Risk | - | 0.64 (0.42 – 0.98)* | 0.64 (0.42 - 0.98)* |
| Low Risk - reference | - | - | - |
| Moderate Risk | - | 0.89 (0.59 – 1.34) | 0.89 (0.59 – 1.34) |
| High Risk | - | 1.77 (0.54 – 5.76) | 1.77 (0.57 – 5.76) |

*Table 3.10 - Hazard ratios for TMACS across adjustment scenarios. * indicates statistical significance. TMACS risk categories are defined as very low risk <2%, 2%<=low risk<5%, 5%<moderate risk<95%, and high risk >95%. TMACS – troponin only Manchester acute coronary syndrome algorithm*

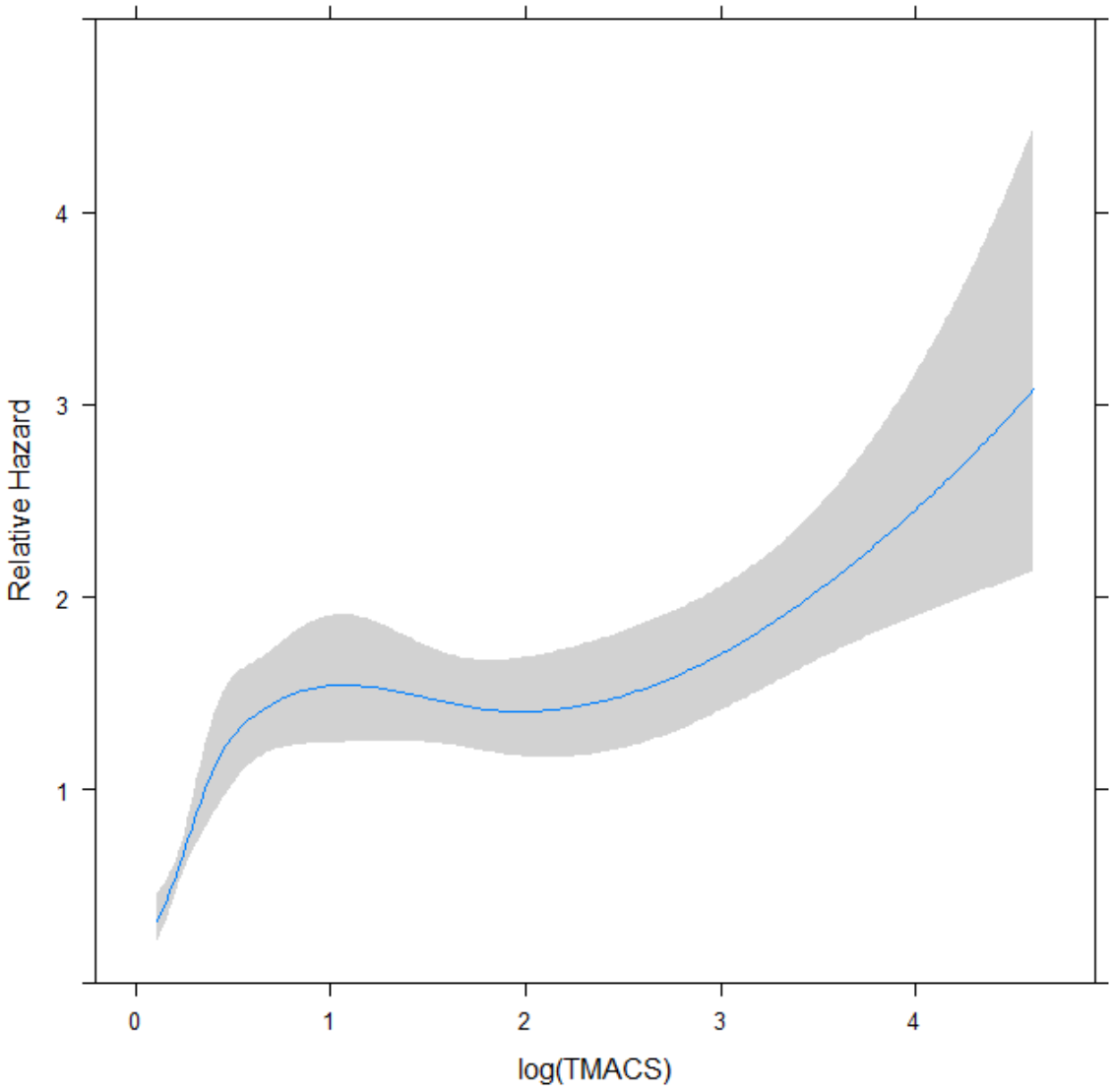



*Figure 3.11 Relative Hazard plot of log(TMACS) when transformed in a five-knot restricted cubic spline adjusted with the selected model. TMACS - Troponin only Manchester acute coronary syndrome algorithm*

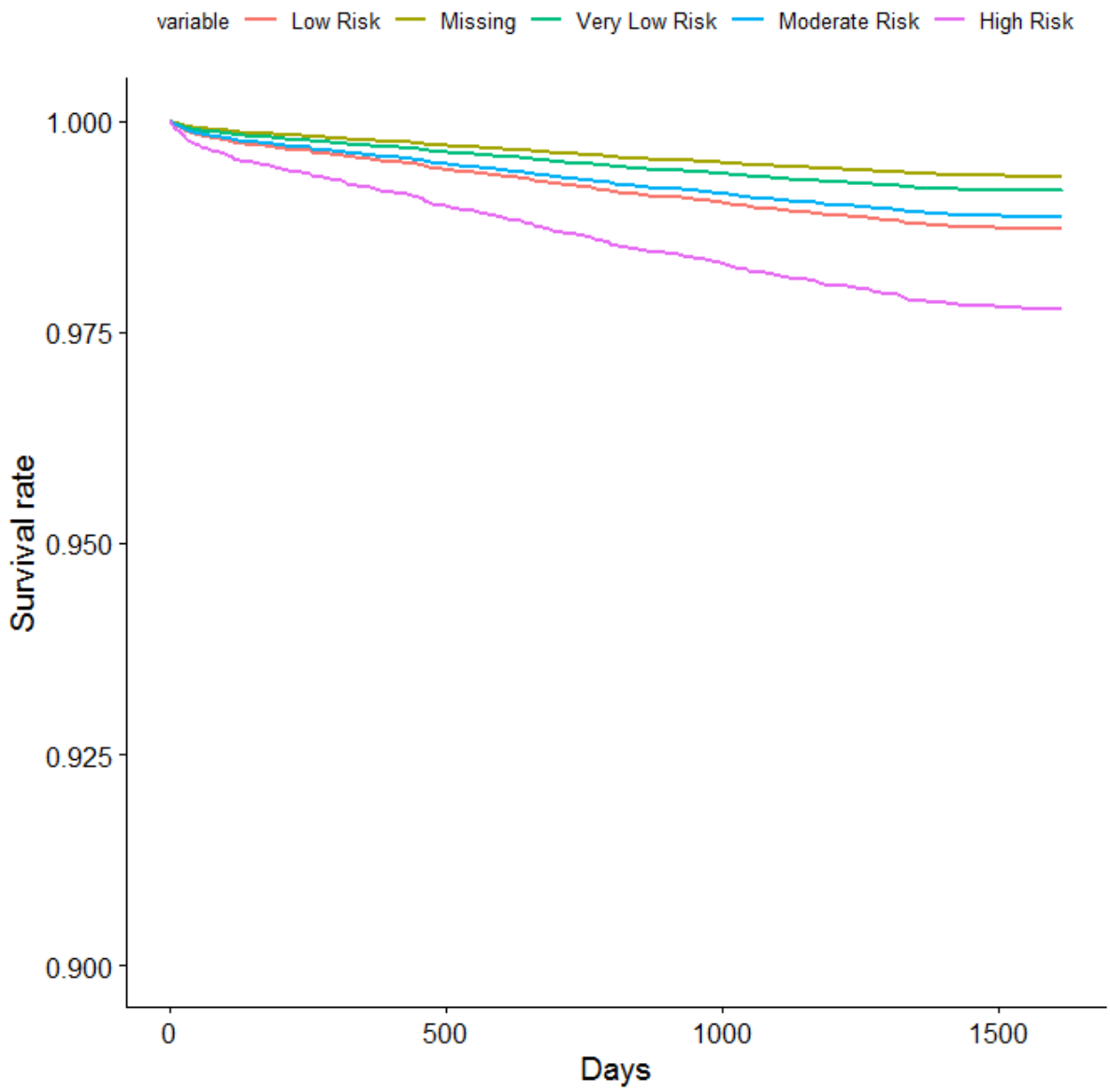



*Figure 3.12 - Adjusted survival curve for categorised TMACS. TMACS risk categories are defined as very low risk <2%, 2%<=low risk<5%, 5%<moderate risk<95%, and high risk >95%. TMACS – troponin only Manchester acute coronary syndrome algorithm*

### 3.5.6 External Validation of Pre-existing Models

I considered four models for external validation, Framingham, QRisk 1, QRisk 2, and ASSIGN (56,57,182,183). The Framingham model has a different outcome definition to the earlier work in this chapter and the other models considered for external validation. It did not include stroke or TIAs in the composite outcome.

I mapped the variables required for each model to what was available from our routinely collected data (Table 3.11). Data that was missing from our database represented data not collected as part of routine care. Whilst the Framingham model was the simplest and oldest model we had the least missing data for it. As such I opted to carry this model forwards for external validation.

Of the variables held within our database only SBP had been shown to vary in effect over time for our broader outcome. Given that Framingham's outcome was different and that this was the only variable showing such variation I opted to pool the three cohorts for the external validation. The missing variables (total cholesterol, high density lipo-protein, smoking and type two diabetes mellitus) were sourced from regional and national data. Regional data was used for prevalence of smoking and T2DM (18.2% and 6.5%) (185). No regional data was identified for TC or HDL, therefore national data was sourced from the SCORE project (TC - 6.3 mmol/l and HDL 1.1mmol/l) (186).

In the pooled cohort of 18,990 patients the mean predicted risk was 12.3% with a range of 0.02 to 88.56%. The concordance was 0.75 (95% CI 0.74 - 0.76), interestingly this was lower than the adjustment models which had concordances ranging from 0.79 – 0.85 (Supplementary Table 8.9). The Framingham concordance is likely more accurate as this is an external validation, as a pose to the adjustment models which are not.

| Variable | Framingham | QRisk1 | QRisk2 | ASSIGN |
|---|---|---|---|---|
| Age | | | | |
| Sex | | | | |
| SBP | | | | |
| Smoking | | | | |
| TC/HDL | | | | |
| BMI | | | | |
| Ethnicity | | | | |
| IMD | | | | |
| FMH | | | | |
| CKD | | | | |
| RA | | | | |
| AF | | | | |
| DM | | | | |
| Anti-hypertensives | | | | |

*Table 3.11 Required variables for different pre-existing long term cardiovascular disease models. TC/HDL – total cholesterol / high density lipoprotein, BMI – body mass index, IMD, indices of multiple deprivation, FMH – family medical history of cardiovascular disease <50 years old, CKD – chronic kidney disease, RA – rheumatoid arthritis, AF – atrial fibrillation, DM – diabetes mellitus. Green >75% data, Yellow <75%, Grey – absent, white not present in model of interest*

The calibration was visually inspected by a calibration plot (Figure 3.13). The lower 9 risk deciles appeared in close proximity to the perfectly calibrated line (y=x), but there was a deviation in the highest risk. This was markedly better than the calibration for the adjustment models which showed a similar pattern of mis-calibration for the high risk deciles but it was much more exaggerated (Supplementary Figure 8.6, Supplementary Figure 8.8, and Supplementary Figure 8.7). It may be possible for Framingham to be further improved by adding TMACS or hs-cTnT through model extension. The summary characteristics of the Framingham model should be considered more robust than the adjustment multivariable models as it is not being assessed upon the data it was derived from. As such the discrimination and calibration statistics are nor prone to over fitting.

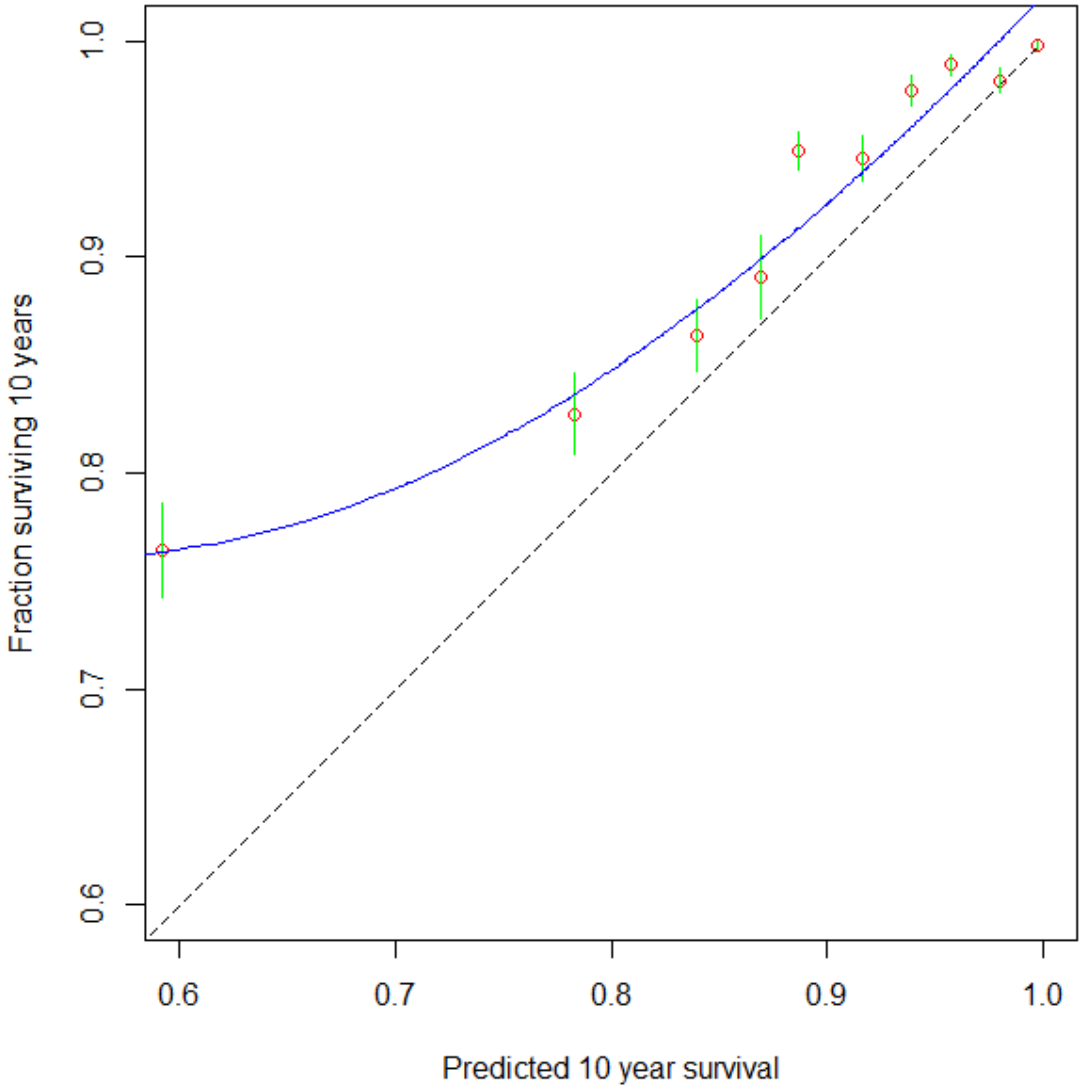



*Figure 3.13 Calibration plot of the Framingham model applied to all cohorts*

## 3.6 Discussion

I have evaluated the potential to use data that are routinely collected in the ED to predict the future occurrence of CVD among patients with suspected cardiac chest pain who do not have AMI. I found that cardiac troponin, using either a contemporary or a high-sensitivity assay, is a relatively poor predictor of long-term CVD. The TMACS decision aid had modest predictive value. However, the Framingham risk score appeared to have the greatest predictive value with a concordance of 0.75 (95% CI 0.74 - 0.76).

I found statistically significant predictors in the baseline cohort most consistently were WBC, age, gender and ethnicity. Heart rate and time of presentation were found to be either statistically significant predictors or statistically significant beneficial additions to the stepwise selected models. Predictors that have been included in primary care cardiovascular prediction models were not all found to be significant in this acute care population (19). Of note eGFR, SBP and IMD were not found to be significant predictors. This could be due in part to the acute population, particularly for eGFR and SBP. In my data it is not possible to distinguish between acute kidney injury and chronic kidney disease from the one eGFR result we have available. It is possible that if CKD could have been distinguished then an effect may have been seen. I have shown that a high blood pressure in the emergency department is confirmed as hypertension in the community in 0.50 of cases (93). As such it may be that a single hypertensive reading in the emergency department cannot prognosticate for long term cardiovascular outcomes, even within a multivariable model.

When cTnT was categorised, the missing level was found to have a hazard ratio of 0.73 (95%CI 0.61 - 0.86). This predictive missingness was present across the diagnostic innovations of hs-cTnT and TMACS. This consistency implies the predictive value is due to a clinical decision; to investigate the diagnosis of AMI. Further research could consider if this logic can be extended back to the factors that a clinician uses to judge the need for further investigation i.e. cardiac chest pain symptoms and examination findings. Across the three cohorts, the median time to event was 2.05, 2.25 and 1.55

years. This highlights a potential limitation of recent AMI rapid rule out pathways. In some studies the follow up is one year, this data would suggest this is insufficient and that a longer time period is required (187–189).

Hinton et al examined all hospital admissions and the prognostic value of hs-cTnI for the outcome of mortality at one year. These results demonstrated a HR of less than 3 rather than 17.86 from Hinton et al (190). The primary outcome is a key difference in these analyses. Firstly, Hinton et al did not report HR for CVD mortality only all-cause mortality, and secondly as a cause of death CVD only made up 13.4%. This makes direct comparison difficult. Moreover, my analysis looked at a composite outcome of CVD disease beyond death.

The overall performance of the multivariable models used to adjust the prognostic factors improved in each cohort with each diagnostics innovation. This work only derived these models and did not seek to validate them, however it is still apparent that the discrimination and calibration improved.

Interestingly the simple Framingham model was better calibrated in external validation, but the discrimination was lower than that of adjustment models. It is possible that the addition of TMACS or hs-cTnT, through model extension, could enhance the Framingham score further (191). The improvement in calibration could be due to several factors: additional variables, prospective data collection, or modified outcome. The additional cholesterol, smoking and T2DM variables may have significantly improved the model's calibration. The data collection for the Framingham study was prospective in an observational cohort study, it is possible the retrospective nature of my work using routinely collected data may be more prone to biases. Similar studies that have used routinely collected data have utilised much larger cohorts up to 7.89 million (19). It is possible this sample size is required to better derive new clinical prediction models with this type of data. Finally, the outcome may be the reason for the improvement in calibration. The absence of stroke may have driven some of the improvement in calibration, it is possible that stroke is harder to predict in this

cohort than the other constituent parts of the combined cardiovascular outcome. As such its removal may improve model performance.

### 3.6.1 Limitations

cTnT had an unexpected hazard ratio of <1. This is likely due to selection bias which is a limitation of this work. We sought to include patients presenting with chest pains as this was the largest populations laid out in our qualitative work. The practical implementation of this was to select patients by the "chest pain" category given by the triage nurses. Given my experience with TMACS, I expected to accrue roughly 3,600 patients with TMACS scores in the final cohort, however only 1,324 were present. This indicated that not all patients undergoing investigation for AMI were included in our data. This could be due to the subsequent clinical review where a history of chest pain may have been newly offered or denied thereby changing the diagnostic focus of their visit. However, given that less than half of TMACS patients were present, this may represent a database issue in the EPR.

Another limitation is the biases inherent with retrospective data from the electronic patient record, cohorts are prone to selection bias and confounding from unmeasured risk factors.

The data relies on reliable data linkage from NHS Digital, and it was not possible to distinguish between cases which did not successfully link and linked cases without a primary outcome. It is therefore not possible to quantify the extent of missing data due to absent or incorrect data linkage via the NHS Digital process.

## 3.7 Conclusion

The ED EPR does contain data that can be used to predict long term CVD outcomes. The improving diagnostic technology for AMI appeared to improve the prognostic characteristics. The available data suited the Framingham model, and it demonstrated acceptable prognostic characteristics, it could be improved with the additional of recent diagnostic innovations.

This work enhances the understanding around what is a good prognostic predictor or CPM from the ED for long term CVD. From this new understanding it is possible to build a care pathway to implement and utilise this knowledge.

# Chapter 4 Co-produced long term cardiovascular care pathway for acute care

## 4.1 Background

Previous research has shown that patients often attend the ED instead of engaging with primary care because of greater perceived accessibility and greater availability of diagnostic testing (192). Up to 37% of patients who attend EDs have unsuccessfully attempted to access primary care before their visit (79). There is evidence that access to transport and social support also influence engagement with primary care following discharge from hospital (193). Further, patients who attend EDs may have disproportionately high cardiovascular risk: one study found that 27% of patients in primary care have asymptomatic hypertension, versus 43.5% for patients who attend EDs (194).

There is also evidence that patients who attend the ED with chest pain would like more attention to be paid to their long-term risks. For example, in a pilot trial of the original version of the T-MACS decision aid, our research group administered a patient survey to research participants (9). Patients were asked to rate "Advice you got about ways to avoid illness and stay healthy" out of five the mean score from 127 patients was 2.78. This was by far the lowest score, of any polled rating. This viewpoint was expressed multiple times at two patient and public involvement meetings that have since been run by the research group. This concept of improving advice on ways to stay healthy has led us to how we could improve cardiovascular disease.

For patients diagnosed with hypertension in an ED, early engagement with follow-up in primary care is associated with greater compliance with treatment (195). Despite this, one study found that only 49% of patients who were found to be hypertensive in the ED had repeat measurements in primary care and only 4.6% received appropriate onward referrals (129).

Together, this highlights the need for (a) better identification of risk factors for cardiovascular disease in patients who attend the ED; and (b) more patient-centred systems to support successful engagement with follow-up to mitigate for and treat risk factors when present. With that in mind, careful consideration is required when designing the optimal pathway to intervene on long-term CVD risk. Co-production enables the designing of a pathway using a mixed methods approach with all stakeholders taking part in the design process.

## 4.2 Published work

Reynard, C., McMillan, B., Jafar, A., Heagerty, A., Martin, G.P., Kontopantelis, E. and Body, R., 2022. Long-term cardiovascular risk prediction in the emergency department: a mixed-methods study protocol. BMJ open, 12(4), p.e054311.

## 4.3 Aims and Objectives

The overarching aim of the work described in this chapter is to co-produce a care pathway to optimise the identification and management of CVD risk in the acute care setting. This was achieved through understanding the opportunities and barriers to implementing such a care pathway, creating and gathering feedback on prototype pathways and creating a final care pathway.

## 4.4 Methods

I adopted a co-production methodology using the qualitative technique of semi-structured interviews. This work took part in two waves; firstly, idea creation and design of potential solutions, then secondly a prototyping phase of those solutions. I envisaged that these would be prototypes of methods to disseminate information and interventions for cardiovascular disease.

This study gained ethical approved by a Research Ethics Committee and the Human Research Authority under reference 19/WA/0312.

### 4.4.1 Positionality and ethical paradigm

I believe that this project is a worthy endeavour and that its objectives can be achieved, this is likely an inherent position of any researcher who undertakes a project. My clinical work has predominantly been in Emergency Medicine. This means I am well versed in the operational aspects of this clinical area, but this could be perceived as a bias towards this specialty. I was cautious when undertaking interviews to not only maintain impartiality in my judgement but also the appearance impartiality so as not to influence the interviewee.

The work presented here embraces the principles of post-positivism, by which we understand that there is a single reality however it cannot be understood perfectly (196). There are biases that influence the data I collect to ascertain the reality, which should be minimised but are unavoidable. However, there are also constructivist principles that apply to this work, i.e., that reality is perceived through a subjective human mind and can therefore be best described as a social construct. It is possible to have valid conclusions that are different depending on the different subjective perspective. The co-produced pathway that I sought to create is an example: there is no single pathway that I am seeking to discover, but rather it will be socially constructed based on the subjective understanding of the relevant stakeholders. In this work an example might be that an intervention is deemed an unwise idea from a secondary care perspective and a fantastic one from a primary care perspective.

### 4.4.2 Study design and setting

I conducted two waves of semi-structured interviews according to the topic guides. The overall objective of the first wave was to co-design an improvement for heart disease care in ED. I invited consultants in Emergency Medicine, general practitioners, nurses, and patients. I used the topic guide in the supplementary appendix.

Participants from wave 1 were also invited to participate in wave 2, however this was not mandatory for inclusion.

In wave 2 I invited feedback upon prototype solutions developed from wave 1, seeking to elicit the participants opinions on the benefits or challenges of the different approaches.

Semi-structured interviews were again conducted with consultants in Emergency Medicine, general practitioners, nurses, and patients. A second topic guide was developed seeking feedback on initial themes from WP1 and on the prototype solutions (see supplementary appendix).

### 4.4.3 Sample Size

I planned to include at least four emergency medicine consultants, four general practitioners, four acute care nurses and eight patients.

I intended to continue until data saturation was met, I anticipated that this would be achieved with 20 interviews per wave but planned to continue if necessary. In wave 2 I iterated target recruitment by participants role to make it an equal 5 per group, this was due to early data saturation in the wave 1 patient cohort.

### 4.4.4 Selection of participants

The clinicians (ED doctors, GPs, and Nurses) were recruited from local NHS centres. Patients were recruited from the emergency departments at Manchester Royal Infirmary. Participant information sheets were be distributed with consent forms collected on the day of the interviews.

I offered the chance of participation to patients who were streamed to an ambulatory part of the emergency department and were awaiting blood results. I approached patients during what I anticipated to be a "teachable moment," when the intervention might be applied in practice, and was therefore most likely to yield findings that could be generalised.

Participants from wave 1 were invited to participate in wave 2, however this was not mandatory for inclusion. I anticipated that additional participants could be required if the sample size was not met, or data saturation was not met. If this was the case, then the same identification method as stipulated in wave 1 were used.

### 4.4.5 Inclusion Criteria

Participants were either emergency medicine consultants, general practitioners, nurses, or emergency department patients with low-risk chest pain.

These patients were deemed low risk by the standard care pathway based around the TMACS algorithm. Once deemed low risk they are then transferred to an ambulatory care unit where they receive further assessment.

### 4.4.6 Exclusion Criteria

- The participants could not attend at least one of the semi-structured interviews

- Not fluent in English language

- The ambulatory ward patient's clinical condition had deteriorated or was severe to the extent that participating in the research would (a) interfere in their clinical care, or (b) that participating would be too strenuous. This was be judged by the nursing staff on the ambulatory care unit, and myself prior to the interviewing the patients.

- Unwilling to take part

### 4.4.7 Consent

To ensure varied patient opinions I needed to conduct interviews during the "teachable moment," which required approaching patients about potential participation while they were waiting for blood

results in the Emergency Department or Ambulatory Care Unit. Patients were given time to read and digest the patient information sheet and to ask any questions. Whilst this approach did not give the patients the same time to consider the consent as the other participants it did have the important advantage of enabling the interviews to be conducted there and then, thereby not inconveniencing the patient by asking them to return.

Clinicians were given more than 24 hours to consider their consent to being involved in the study. This gave them time to read the participant information sheet and seek any clarification they needed. This was feasible as the interviews were be scheduled with 24 hours advanced notice.

### 4.4.8 Coding

I used multi-grounded theory described by Goldhulkl et al (197). This enables a mixture of deductive and inductive reasoning, using 'pure' grounded method to deduct and existing theory to induct creating a more robust analysis as a whole. The audio was transcribed by an external contractor and analysed using thematic analysis.

Two separate researchers coded select transcripts separately to ensure that no themes were misinterpreted or omitted. Each researcher was blinded to the process and if conflicting codes were identified then a third researcher adjudicated. This was to ensure the transferability of the findings.

### 4.4.9 Care pathway creation

From the first wave of semi-structured interviews concepts and themes were mapped to a logic model in a similar method to that of Smith et al (198). There were two logic models, one consisting of potential model inputs and a second with potential model outputs. The trial steering committee group was then involved which had members from emergency medicine, primary care and patient representatives. The logic models were reviewed by the trial steering committee in conjunction with the thematic analysis to collaboratively generate care pathways. This process was used to create three prototype care pathways for feedback at the second wave of semi-structured interviews.

## 4.5 Results

The interviews were conducted in the two waves and with sample sizes of 21 and 20. The first wave took place over four months from October 2020 to January 2021 and the second wave over three months from November 2021 to January 2022. These interviews took place in the COVID-19 pandemic, and the influence of this is discussed later.

| **Demographic Factor** | | Wave 1 | Wave 2 |
|---|---|---|---|
| Mean Age (SD) | | 49.2 (11.8) | 42.1 (14.0) |
| Gender (n) | Male | 66.7% (14) | 45%% (9) |
| | Female | 33.3% (7) | 55% (11) |
| Ethnicity (n) | White | 76.2% (16) | 76.2% (16) |
| | Black, Black British, Caribbean or African | 9.5% (2) | 9.5% (2) |
| | Mixed | 9.5% (2) | 4.8% (1) |
| | Asian or Asian British | 4.8% (1) | 4.8% (1) |
| Participant roles | Patient | 8 | 5 |
| | GP | 5 | 5 |
| | EMN | 4 | 5 |
| | EMC | 4 | 5 |
| Years' Experience (SD) | | 14.4 (13.3) | 7.1 (8.0) |
| Highest Educational achievement | Nil | 3 | 0 |
| | Secondary school | 0 | 4 |
| | Undergraduate | 16* | 10 |
| | Postgraduate | 1 | 6 |
| Average IMD Decile (SD) | | 3.3 (2.2)* | 4.8 (3.2) |

*Table 4.1 Demographics of participants in wave 1 of interviews. SD – standard deviation, GP – general practitioner, EMN – emergency medicine nurse, EMC – emergency medicine consultant. Secondary school education encompasses GCSEs, A-levels and MVQs. * Highest educational achievement and years' experience were data points added by amendment and were only collected part way through wave 1.*

I reached the target sample size in wave one across the different stakeholders (primary care, emergency medicine clinicians, nurses, and patients). The sample was in keeping with expected demographic factors of age, years' experience, education achievement and index of multiple deprivation (Table 4.1). The ethnicity of the professional cohorts was predominantly White British with only one participant being mixed African-Caribbean. Only five of the thirteen participants were

female. This is an under-representation of these groups. The patient cohort is mainly male (7/8) and four were White British (50%), two Black British, one Pakistani ethnicity and mixed Irish/central American black. This is a far more representative sample of the varied ethnic groups of Manchester's population, and whilst the majority were male this is representative of chest pain patients. An amendment was submitted on the advice of the trial steering committee to include highest educational achievement and indices of multiple deprivation. Unfortunately, due to study timelines it was not possible to include to wait for the amendment prior to starting the semi-structured interviews. The amendment and resulting data were therefore only collected from participant 14. The nurses and medical doctors were assumed to have an undergraduate degree (including diplomas) as this is a criterion for professional registration.

The wave two cohort consisted of 20 participants with an even representation of each role within the care pathway. The second wave was younger than the first with a mean age of 42.1 versus 49.1, on average the clinical staff were also less experienced with an average years' experience of 7.1 versus 14.1. The cohort had a higher proportion of women at 55% compared to 33.3% in the first cohort and the distribution of ethnicities were broadly similar. Highest educational achievement was different between wave 1 and 2, with the second cohort having more participants with postgraduate qualifications 6 vs 1 and no one with no educational qualifications or achievements. The average IMD for wave 2 was 4.8 which sits close to the middle decile for the country (5), implying that this was overall a representative cohort. Given the missing IMD data from the first wave a comparison is not possible.

### 4.5.1 Wave 1 Thematic analysis

Interviews were coded with multi-grounded theory in a thematic analysis. Prior knowledge of the subject matter was used to inform the interpretation of the interviews, which were coded and synthesised into themes. These themes were to be used to guide the creation of care pathways later. The interviews were also deconstructed into building blocks for a pathway (attached), this is presented in a series of lists from which different aspects can be selected to create a care pathway.

There were six main themes from the current version of the analysis: Loci of clinical responsibility, Poor communication, avoidance of pandemic hospitals, reductive EM care, acceptability of CVD care in EM and automation of preventative EM care. As this is an iterative process this was refined in the second wave of interviews.

#### *4.5.1.1 Loci of clinical responsibility*

The study participants repeatedly highlighted the importance of who was responsible for new clinical findings and their treatment. This incorporated two main sub themes: primary care to secondary care and patients vs the NHS. This was an important theme as it is the foundation of any clinical pathway.

##### Primary to secondary care

Whilst the overall sentiment of the interviews was in favour of some degree of preventative clinical pathway, precisely who should be responsible elicited emotive responses from some participants. This was also an emotive point for some professional participants as the potential pathways represented, to some, a change in definition of their role. EM clinicians were frustrated at the general lack of resource and time available for clinical care and felt that any further task would only exacerbate this issue. Some general practitioners felt that this was not the remit of emergency care and expressed concern that it was encroaching on an already embattled primary care identity.

A given example of the current loci of responsibility for patients in the ED was referral to rapid access cardiology clinics. Current practice is that patients requiring an elective cardiology investigation should be referred by EM to GPs who then refer to a cardiology clinic. Frequently this

request is a single line on a discharge summary. The logic for this that potentially GPs have access to the wider patient record and then can make a more informed referral. However the counter argument is that this is a system inefficiency that forces patients through a complicated system. In this process the loci of responsibility transitions from EM to primary care to cardiology. EM delegates responsibility to colleagues in primary care after their emergency visit is completed (normally for chest pain), at which point due to system factors the responsibility does not move directly to cardiology rather via general practice. It was noted by participants that every time the responsibility for patient care transitioned, information was known to be lost (e.g. a heart tracing), which may indicate that more information is also unknowingly lost. There are conflicting sentiments here (i) the emergency physicians are too busy to take it on and (ii) we shouldn't shift responsibility because information might be lost. It is clear that responsibility has to shift back to primary care because the ED cannot go on looking after patients' long-term CVD risk indefinitely (or at all, following the ED visit). Therefore, we need to minimise transitions in responsibility and ensure that systems are designed to minimise data loss at handover points.

Regarding risk factors for long term heart health all participants agreed that a positive result identified opportunistically in the ED should be followed up or actioned. Some clinicians felt that the clinician overseeing the acute episode should be ultimately responsible, whilst others felt it reasonable to merely act as a conduit for the results to pass to another service. Caution was noted that sometime clinicians are not aware of all services available, or if they are functional in part due to a lack of feedback.

Some patient participants deferred to the paternalistic medical model where they were happy to follow the clinician's instructions, thereby keeping the responsibility on the clinician. However, other patients felt that once their fellow patients had been given the knowledge and tools to themselves action a positive result, the onus should be on the patient to act. Clinicians echoed this, in one instance with frustration when an international comparison was drawn.

*"I think, the NHS is already brilliant at taking responsibility away from patients, we follow up a lot, we do a lot for patients compared to other countries."*

A GP participant quote regarding the theme loci of clinical responsibility.

Clinicians in general were in favour of the system reaching out to patients to investigate/treat a new risk factor. However, some did caution that this should be done in a limited way, as the result of these pathways may well be the patient being offered a therapy which requires them to comply autonomously. They were however keen that wherever the pathway's responsibility lay it was transparent and communicated clearly (see below quote).

*"responsibility is kind of important, but actually making sure that everybody is probably aware of what everyone's responsibility is."*

An Emergency Medicine Consultant's quote regarding the theme loci of clinical responsibility.

#### *4.5.1.2 Poor communication*

Communication throughout the healthcare system was described as sub-optimal. This theme was split into two further subthemes.

##### Primary and secondary care interface

GPs universally felt that the discharge summaries from the emergency department were not optimal, some were emotive at the suggestion of additional communication when the basics were not yet acceptable. Overall comments were that the summaries were short, poorly organised and the requested actions were of varying acceptability. Some GPs expressed a lack of trust in the accuracy of the information contained within the summaries. Two GP participants reported having recent experience of the EM clinical environment and expressed sympathy for the workload in EM, the remainder expressed strong feelings that this was not an optimal situation. It is possible that knowledge of the context enabled empathy between the specialties.

*"We know that communication between Emergency Departments and primary care are very poor."*

A quote from an Emergency Medicine Consultant describing communication between primary and secondary care.

EM clinicians agreed with GPs that the summaries required improvement. They expressed doubt that this could not be easily fixed due to a lack of time available to fill in a sufficiently detailed summary per patient. They hoped for an IT solution where work was not duplicated, and an automated system could create a summary from the existing clinical record. EM clinicians also noted that they did not know if a referral/summary had been received, was deemed appropriate, or actioned. This blind follow-up disincentivised clinicians from using follow up systems. GPs also expressed concern at the vulnerability of their care pathway due to the volume of the secondary care communication and the required processes to action and requests.

*"A lot of these emergency department letters, especially in today's brave new world where a lot of letters won't get seen by clinicians at all, but they'll get simply triaged by the admin teams in practices, they'll just get filed and no further action will be taken."*

A quote from a general practitioner describing the processing of communications from secondary care

*"… quality of the letters from the emergency department to GP which vary a lot depending on who writes and their understanding of the SHO, consultant, …"*

A quote from a general practitioner describing the range of quality of letters from secondary care

EM clinicians also noted the variation in GP practices and that they were not one single organisation that could be consulted. Instead, they were interacting with a series of individual businesses with unique characteristics.

### Acute condition diagnosis

Patients expressed frustration that they were not always told what the cause of their original acute presentation could be. They felt that often it was simply stated that it was not a heart attack, and that this was not further explored. This view was reinforced by the nursing participants who stated that they often found themselves explaining results to patients again, due to miscommunication. The source of this miscommunication was hypothesised to be a lack of time, and the potential for patients to have a reduced capacity to process the information in a stressful situation. Patients and

nurses also felt that there was a different primary focus for the patient and clinician. Patients wanted to know what the cause of their symptoms was, whereas clinicians were primarily focused on excluding high risk pathologies (such as a heart attack). These different priorities often went unnoticed early in the patient's journey and only became apparent at the conclusion when the patient was simply told they did not have a heart attack. GPs expressed an awareness of this situation and empathy for EM clinicians due to the clinical uncertainty, however they were keen for a provisional diagnosis to be made in EM prior to discharge.

*"… a possible explanation for why I got that such severe pain that was just out of the blue, rather than just, oh, it's nothing, don't worry about it, because brains don't work like that, and my brain certainly doesn't. "*

A quote from a patient explaining their concerns about why they got such severe pain,

#### *4.5.1.3 Avoidance of pandemic hospitals*

Patients stated that they wanted to avoid hospitals in the pandemic and that their thresholds for attending were substantially higher. They were anxious regarding catching COVID-19 whilst in hospital and were wary of overburdening the system. Clinicians noted the change in patient behaviour as well but reflected that it was had changed. Early in the pandemic it was felt that all but the most unwell patients stayed away from hospital, however later the normal patient mix returned. GPs noted that they were dealing with more patient presentations that would normally have been deemed worthy of attendance at an emergency department.

Clinicians universally expressed that there was a cohort of patients who did not attend hospital but should have, and that this was a problem that would become more apparent later for the health service.

#### *4.5.1.4 Reductive emergency medicine care*

In contrast to whole patient or holistic patient care, EM care was pervasively described as reductive. For example, it was focused on a condition or symptom such as focusing on a pneumonia whilst excluding other issues. This was seen to be because of time pressures and the desire for efficiency.

These characteristics of reductive care were viewed as both a positive and negative. Nurses felt this described their role was and it was the "ED lane", this was echoed by some GPs and EM doctors as exemplified in the below quotes.

*"I think the GP is going to take the care, this is a long-term chronic condition, which the patient is going to have to live with. So the GP needs to be the person to do that. I don't think that's A&Es role."*

Emergency Medicine Consultant quote

Not actioning this had negative connotations for some EM clinicians who stated it was patching a problem only for it to break in the future. Frustrations around this were shared by GPs particularly when it came to communication, they felt that often EM clinicians placed things outside their remit inappropriately by solely focusing on a single symptom or condition. EM's high-pressure environment and time focused clinical targets were thought to exacerbate this issue by placing reward on quick pathways and penalising slower versions.

#### 4.5.1.5 Acceptability of implementing measures to prevent long-term cardiovascular disease in the acute care setting

Clinical stakeholders felt that implementing measures that aim to prevent future cardiovascular disease would be broadly acceptable but there were some concerns expressed. EM clinicians were positive about the potential intervention but were cautious about what could be done in the ED and how this would be followed up over a prolonged period. GPs were broadly in favour if it was a low intensity activity that might signpost patients to traditional primary care prevention care pathways.

*"If they've got risks of heart disease, then they need to get followed up and they need to be started on the appropriate treatments and given appropriate health advice. And then probably followed up, not just once but over a period of time, to actually encourage them to modify their risks as best we can."*

Emergency Medicine Consultant quote

*"I don't see anything wrong with using it as an opportunity thing."*

General Practitioner quote

The ED nursing staff were similarly positive about the prospect. However, whilst they were in favour of the concept there was some concern that some patients may not be receptive to the concept of the teachable moment. Given the length of time that patients often wait they were concerned that they may be frustrated if a topic not directly relevant to their acute presentation was raised.

*"I guess what I'm trying to say is, lecturing a patient after they've spent four hours in A&E waiting for blood results and a chest x-ray I don't think is the right time and place."*

Emergency Medicine Nurse quote

Patients were firmly in favour of the focus on preventing future CVD. They were surprised that data collected that could be interpreted to predict CVD was not utilised. They did not express any of the concern that the other health care stakeholders had.

*"No, I think they should do it, I think they should be truthful with everything they say to you what's going to happen"*

Patient quote

#### *4.5.1.6 Automation of preventative care in EM*

The use of routinely collected data for preventative medicine was widely seen to be an efficient use of resources. The addition of further tests, such as HbA1c (to screen for type two diabetes mellitus) and a cholesterol panel, was also seen to be a welcome addition. Particularly as this addition may well avoid a patient visit in primary care. There was concern over ownership of the results which is discussed previously.

The ability to target a potentially under-served group whilst they were in a teachable moment was also seen to be a strength. Alternative strategies to identify a cohort previous hidden from preventative care strategies were thought to be likely too costly.

### 4.5.2 Logic model for care pathway generation

To build the prototype care pathways logic models were constructed from the semi-structured interviews. These were separated into care pathway inputs and outputs. Inputs included those things required to go into the care pathway such staff, risk factors and data (Supplementary Table 8.10). The outputs or potential outcomes of the care pathway including a referral or lifestyle advice which were placed in a separate logic model (Supplementary Table 8.11). I have summarised these logic models in the style from Bradbury et al in Figure 4.1 (199).

The prototype synthesis was conducted within the trial steering committee by collaboration based on the themes and sub-themes identified. Key themes that influenced the prototype care pathways were the loci of responsibility, poor communication and reductive EM care. Three pathways were distilled from the trial steering committee deliberations, an ideal care pathway, a streamlined care pathway and a worst possible care pathway (Figure 4.2, Figure 4.3, and Figure 4.4). The later care pathway was designed to stimulate discussion and to highlight factors to avoid.

The ideal care pathway prototype is detailed in Figure 4.2, it utilises automated data collection thereby avoiding any additional burden on clinical staff. Routinely collected data would be used to calculate future CVD risk, which would then be automatically pushed to the treating clinician. At this point the clinician would communicate the results to the patient, providing relevant leaflets and/or internet links for bespoke websites. The aim of this brief interaction would be to communicate the patient's future risk of CVD and discuss lifestyle factors that could begin to manage the risk. The output of the pathway would also be communicated to the patient's general practitioner, where a pre-agreed automated handover will take place and the patient is considered for further review.

This proposed care pathway and data flow is hypothesised to make the identification of patients who are at risk of future CVD feasible in a busy ED, thereby enabling primary prevention and a reduced future burden of CVD. This requires patients to engage in behavioural change (potentially including medication compliance) over a prolonged period. The concept of capitalising upon a

teachable moment when patients may be more amenable to interventions is central to this pathway. It warrants further investigation to understand the magnitude of this effect.

A disadvantage of the ideal pathway is that it requires a brief intervention by the treating emergency clinician, which conflicts with the sentiment that busy EM clinicians need to retain focus on the acute complaint. Therefore, an alternative streamlined care pathway was also distilled (Figure 4.3). This pathway is identical to the ideal pathway until it comes to communicating the results of the pathway to the patient. At discharge, the patient is told that the results will be communicated later to their primary care service. This allows for a warm handover between primary and secondary care whilst not burdening the ED clinician with an undertaking a brief intervention.

A 'worst possible' care pathway was also defined to incorporate the 'faux pas' highlighted in the interviews (Figure 4.4). This pathway required manual input from clinical staff to collect and input additional data. The patient was not informed of the process at any point of their acute presentation. Primary care was informed in a verbal hand over which would entail a time burden to acute care and potentially lead to miscommunication. In this scenario the care pathway has not been agreed between primary and secondary care and therefore the call will likely be unexpected. The patient, unaware of this process, only becomes aware if primary care contacts them to invite them for follow up. This negates the benefit of the teachable moment by disconnecting the long-term CVD consultation from the acute care episode.

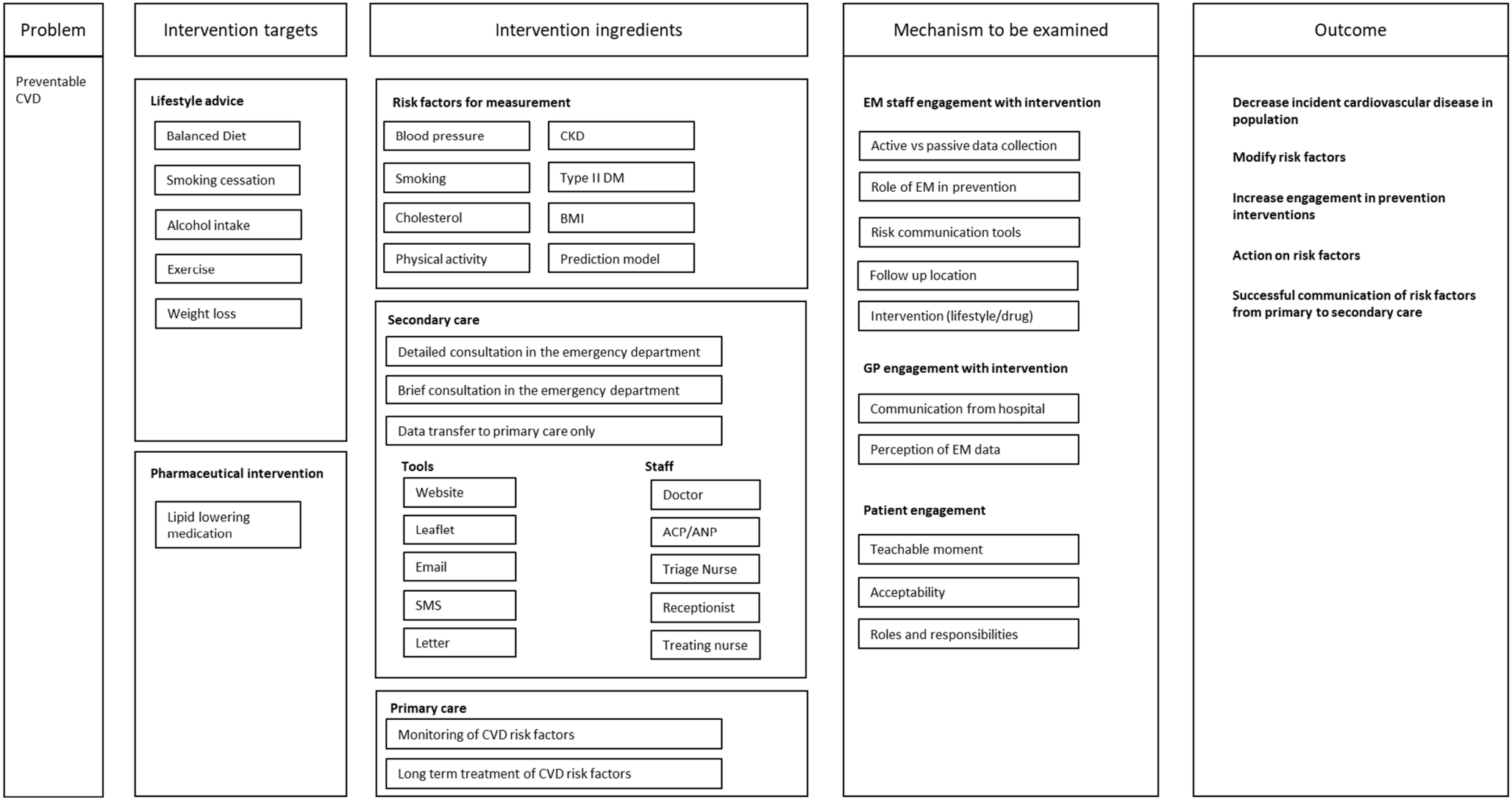



*Figure 4.1 Summary logic model of potential cardiovascular care pathway*

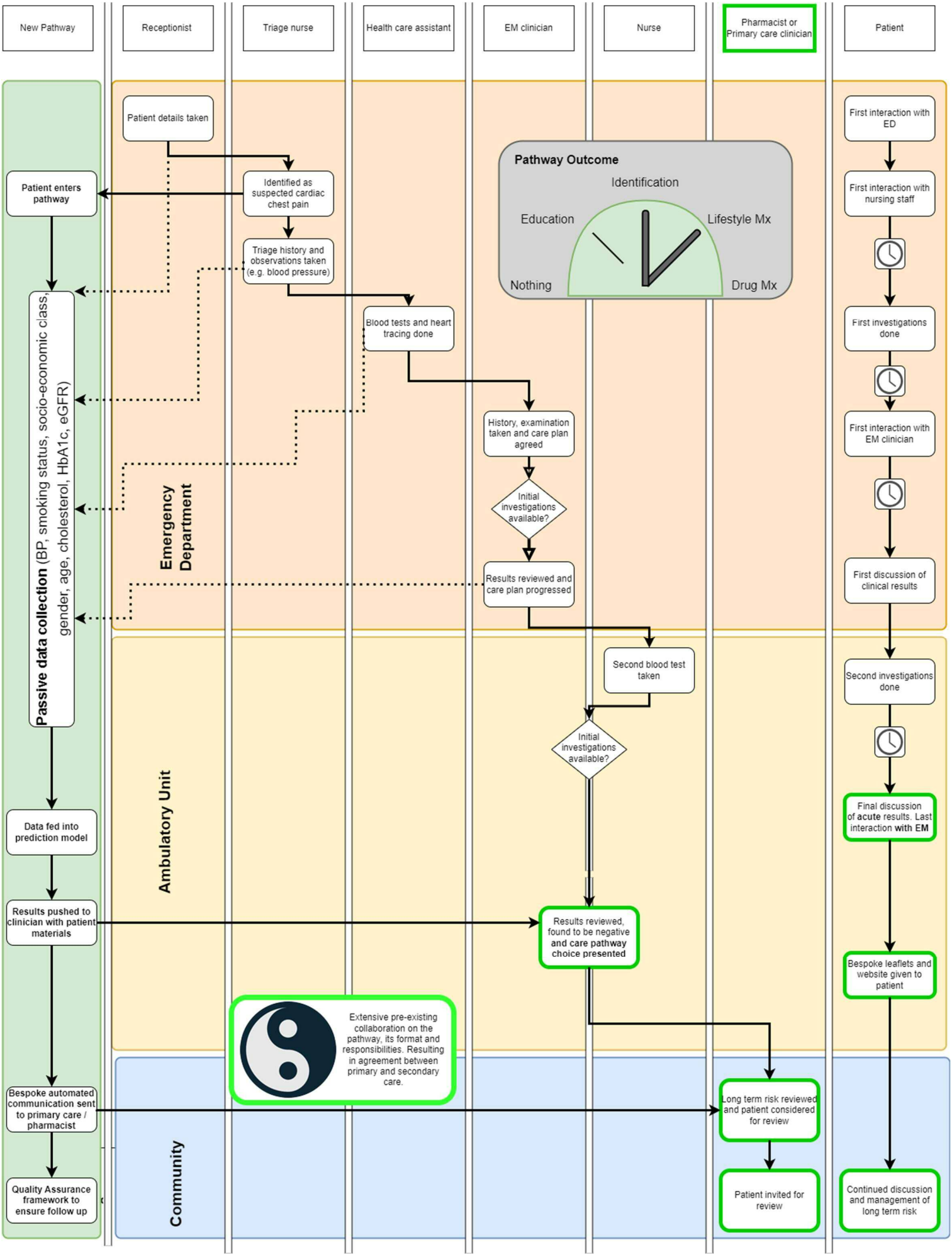



*Figure 4.2 Prototype care pathway – the ideal care pathway from the totality of the research*

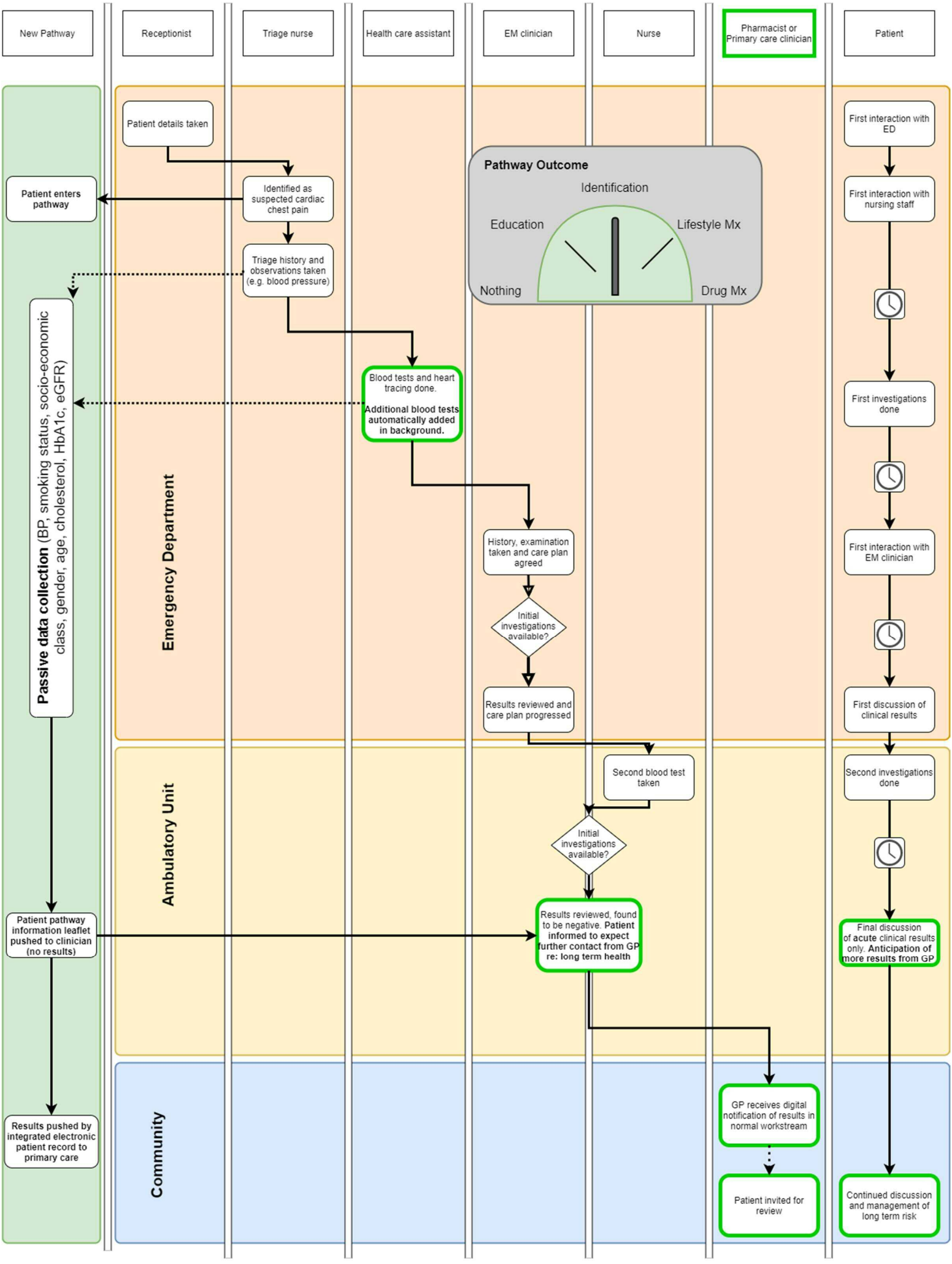



*Figure 4.3 Prototype care pathway – the streamlined care pathway reducing the activity in acute care to a minimum*

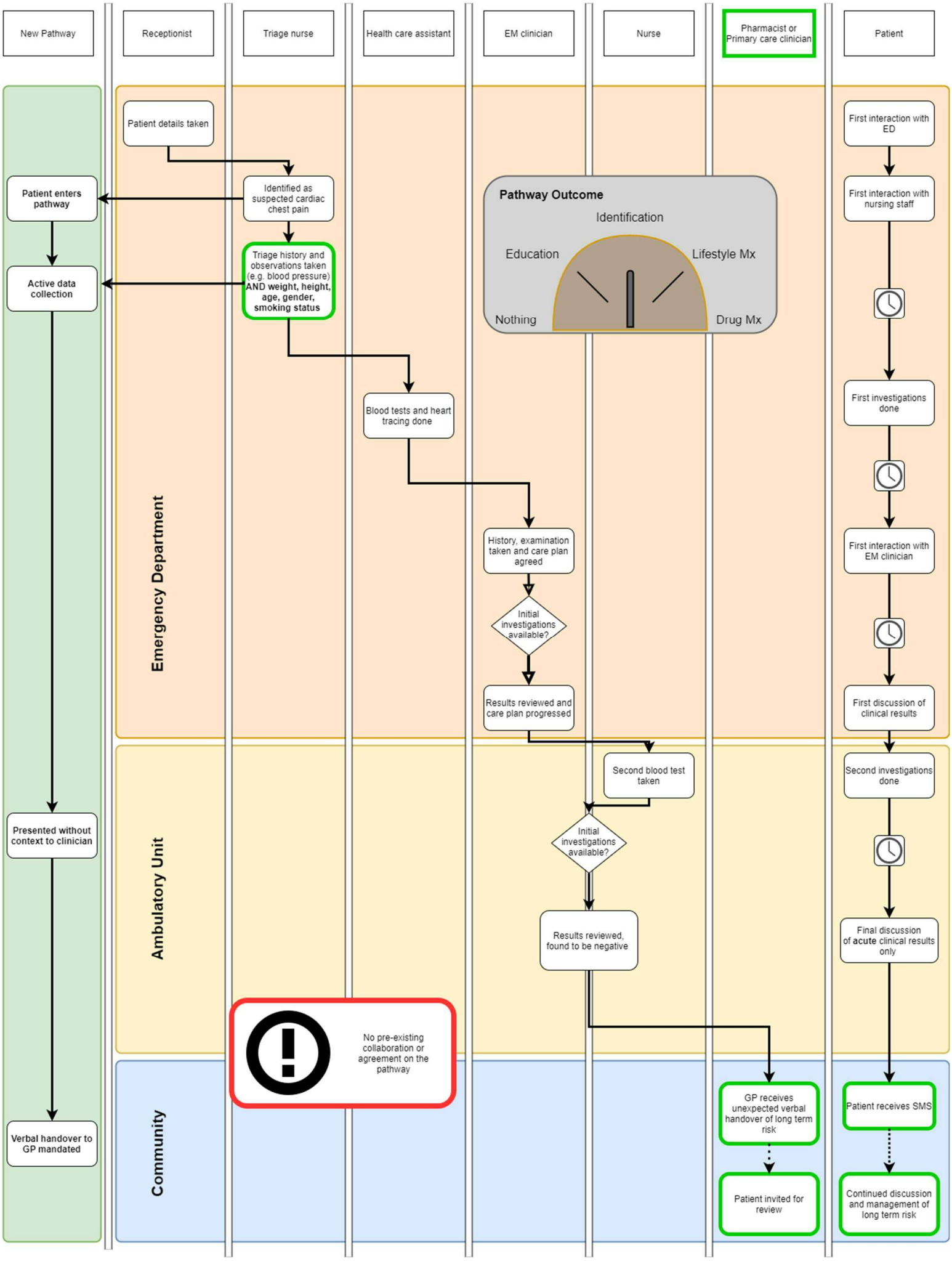



*Figure 4.4 Prototype care pathway – the worst possible care pathway from the totality of the research*

### 4.5.3 Wave 2: Finalisation of the co-produced care pathway

#### *4.5.3.1 Reflection on previous themes*

Loci of clinical responsibility was explored again with participants and again the broad concepts of the importance of transparency came out. There was reflection that some clinical practice was in part driven by concerns around medico-legal consequences. Below are two quotes that demonstrate this concept.

*"ultimately, you know, there's the consultant name on for that day"*

A quote from an Emergency Medicine Consultant regarding the medico-legal impact perspective on responsibility

*"I think ultimate responsibility for a test still has to lie with the person who does it."*

A quote from a GP regarding the medico-legal impact perspective on responsibility

The theme of poor communication was also reinforced in the second wave of interviews, for the primary/ secondary care interface and the clinician/EPR interface. That information was lost or corrupted when a clinician inputted into a computer system be it a discharge summary to the GP or a referral from primary care to the ED. Not only was communication through the clinical/EPR interface leading to information being lost but clinicians also felt that it was oversharing. Secondary care discharge summaries contained data that was not relevant to primary care and this increased the noise and made it harder to identify the signal or clinically pertinent information. The use of the EPR particularly for discharge summaries was not valued by secondary care staff, the acute patients were prioritised and this task with no immediacy to its nature was devalued.

*"I think in a working A&E department that's a low priority task and a lot of the time it's left until a later date when somebody can just go through and write all these letters"*

A quote from an Emergency Medicine nurse

*"A lot of information that can be communicated, I would imagine that the GPs suffer from two different information deluge of stuff that they don't need to know and a lack of stuff that they think that we do need to know."*

A quote from an Emergency Medicine consultant

#### *4.5.3.2 The dangers of "bolt-on" activities to core clinical work*

A new theme emerged from the second wave of interviews, and it focused on the dangers of 'bolt ons.' These were small additional burdens to a core piece of work, in this context it was used to describe the additional work created by an axillary pathway . There was frustration from primary care stakeholders that they had previously agreed to pathways that had changed over time and become more burdensome. This was on top of a background sentiment of already be used as the administrators of the NHS by some in secondary care. Interestingly this feeling of imposition was also felt by EM stakeholders either from within their own department or other hospital specialties.

*"all these pathways, they are presented in a way that's like, oh you can barely feel it, it'll be fine. And nearly always that's not the case. I'm a little bit sceptical that it'll be as easy as, [it says]"*

Quote from Emergency Medicine Consultant

*"It's the opposite of British cycling, you chip away tiny little fractions that in the end make the difference between winning and losing. Well in the NHS we do the exact opposite, we gradually make things less and less efficient until the whole thing falls apart because we've overloaded it."*

Quote from Emergency Medicine Consultant

*"there's a lot of local angst about agreed pathways which subtly change over time, to burden the GP with more work"*

Quote from General Practitioner

This anxiety translated into inertia around care pathway innovation, where the clinical stakeholders felt justified in their hesitancy for implementation. Primary care stakeholders repeated a narrative of being asked by secondary care to do activities that in their opinion should and were previously

routinely conducted in secondary care. In EM, an example of a well implemented bolt on was where the implementation was contingent on no extra EM resources being required.

#### *4.5.3.3 Feedback on pathways*

The feedback for the care pathway prototypes was guided by the structure described by Peters et al (200). This included acceptability, likely adoption, appropriateness, feasibility, and coverage.

The acceptability of the care pathways for the clinical stakeholders was contingent on the additional workload it would give them. All felt that the principle of improving long term CVD outcomes in this group was a worthwhile endeavour, however there was concern that any new care pathway might detract from other work. The EM clinical stakeholders demonstrated a preference for the stream-lined care pathway prototype, which reduced the tasks for EM. General practitioners expressed concern over the feasibility and appropriateness of these tasks being inevitably pushed to them. The sentiment that the problem should be fixed where it was identified was common. Patients found both the streamlined and optimal pathways acceptable, though they did express a preference for the optimal pathway where more activity took place in the ED.

The initial adoption of the pathway relied on the alignment between primary and secondary care as the pathway spanned both areas. A suggestion to facilitate this was a local clinical forum, which could be initiated prior to care pathway implementation. The EM nursing stakeholders highlighted the need for training in the subject area long term CVD if a lot of activity was expected in the ED. A primary care stakeholder suggested that it may be possible for some of the EM activity in this care pathway to contribute to quality and outcome framework payments in primary care. If it was possible to automate the sharing and coding from secondary to primary care EPRs then this may drive adoption from the primary care stakeholders.

The appropriateness of the pathway was not questioned, however the location and intensity of the intervention was. Primary care stakeholders expressed concern for more work to be passed from

secondary to primary care. EM stakeholders were concerned that this was another 'bolt-on,' and that it would divert resources from their core work reducing their ability to provide 'routine' clinical care. The streamlined care pathway was broadly acceptable to both groups of stakeholders, with the caveat from primary care that any communication and its integration into their systems was pivotal to its success. The optimal care pathway with a bespoke consultation at the point of discharge was seen as inappropriate by EM stakeholders given the current clinical pressures faced by EDs. There was an acceptance from primary care stakeholders that the eventual management of long-term CVD would fall to them, the point of contention was when that hand over happened.

The acceptability links to feasibility, as something unacceptable to staff is inherently not feasible. A key factor for the feasibility is the streamlined care pathway requiring an integrated EPR system across primary and secondary care. The ability of the NHS to deploy such a complicated digital solution was questioned when other far less difficult digital tasks were seen as unachieved. This digital solution allows the worst-case care pathway to be avoided. Specifically preventing additional tasks being added at triage, allowing the risk prediction to occur seamlessly in the background and the information transfer to primary care in a pre-agreed system and format that maximises utility.

The capacity for the pathways to function out of hours was a concern for stakeholders and the trial steering committee. EM nursing stakeholders were concerned that patients attending out of hours may not receive the same level of care from this pathway due to extra overnight pressures. The steering committee members were concerned regarding the lack of adaption for highlighted the need to adapt any patient-facing material for non-English speakers in the optimal care pathway. The streamlined care pathway was selected as being able to maximise coverage overnight due to its simplicity and to for non-English speakers any materials could be translated and the messaging from the clinician would be shorter and more easily translated in a time poor environment.

### 4.5.4 Final care pathway

The final care pathway was derived from the feedback and analysis of the wave two interviews (figure 4.4). EM clinicians made it clear that a pathway that added extra 'non-routine' activity would not be easily implemented and that for some it was deemed completely unfeasible. Given this red line for EM stakeholders an option was added that merged the optimal and stream-lined care pathways. This choice aims to alleviate the feeling of an imposed bolt-on for EM stakeholders but introducing an element of clinician control over the pathway. At the same time it allows for quick communication with the patient to maximise the potential of the teachable moment. Any materials would need to be translated into prevalent languages for the local area.

Similarly the prototype care pathways the CPM inputs were automated from the routinely collected data in the ED. Also, the pre-agreement and integration with primary care remained a pivotal part of the pathway.

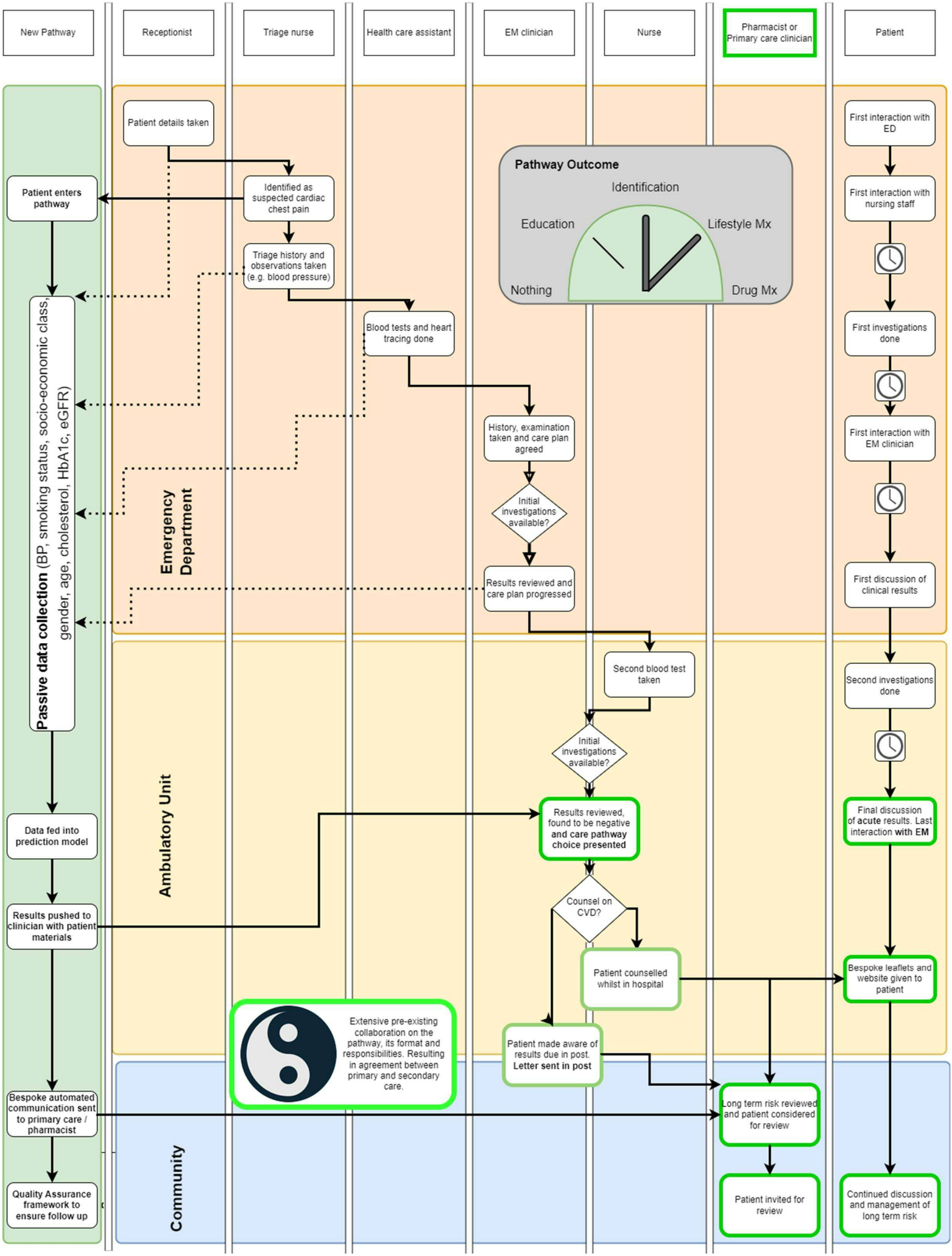



*Figure 4.5 Final proposed care pathway*

## 4.6 Discussion

In this chapter I have described the co-production of a long-term CVD care pathway using 41 semi-structured interviews across multiple stakeholder groups. The final care pathway has the potential to be successfully trialled and deployed without failing at the identified points of contention. It utilises routinely collected data to predict long term CVD risk relying on a digital solution for secondary care. The integration into primary care relies on good pre-implementation communication and agreement regarding the medium and content of any communication. This could be maximised if QoF targets could be achieved for primary care by the secondary care activity.

This work draws on elements of EBCD as described by The Point of Care Foundation (201). I did not conduct all the formal stages of EBCD such as a trigger film or celebration event. The framework seeks to overcome the asymmetry between clinicians and patients and it has been hypothesised that not following it fully could compromise this (202). A systematic review by Green et al proposed that to qualify as EBCD patients needed to be involved in the experience gathering and co-design phase (202). It is reasonable to argue that in this work the semi-structured interviews with patients formed the experience gathering portion. Furthermore the co-design phase could be satisfied by noting the important participation of two lay members in the steering trial steering committee, and the valuable feedback provided by patients in the second wave of interviews. It is important to note that the patients and clinicians involved in this work were all from the Greater Manchester area. This may limit the transferability, particularly to rural areas.

The patient desire to receive advice about how to stay healthy after leaving the ED has also been identified in previous qualitative work. Ferry et al conducted semi-structured interviews with patients before and after implementation of a rapid rule-out pathway for suspected ACS (84). They identified a theme on ‘approaches to future heart health’ with patients noting that they were likely to be particularly receptive to health prevention activities when attending the ED with chest pain.

This further confirms that there is an unmet need and untapped potential for a care pathway to address ongoing CVD risk in the teachable moment around chest pain attendances.

The MRC and NIHR have provided guidance for the development of complex interventions in healthcare (203). This package of work, combined with the outputs of the two preceding chapters, adheres to that guidance (Table 4.2). The process was planned and involved a multitude of stakeholders. I reviewed the evidence in systematic reviews and undertook primary data collection where necessary. I designed the intervention with thought to the future implementation.

| **NIHR & MRC complex intervention guidance theme** (203) | **Doctoral activity** |
|---|---|
| Plan the process | Doctoral planning/protocol publication |
| Involve stakeholders | Semi-structured interviews with GPs, patients, EM consultants and nurses. |
| Bring together a team and establish a decision-making process | Trial steering committee involving representatives from primary and secondary care as well as patient representatives. |
| Review published evidence & draw on existing theories | Conducted a systematic review and used multi-grounded theory to enable iterative approach in qualitative studies |
| Undertake primary data collection & understand the context | Co-production of the pathway made possible by semi-structured interviews, and prognostic factor study to understand what can be included. |
| Design the intervention and pay attention to future implementation | Co-produced care pathway iterated in part to enable sustainability. |

*Table 4.2 O'Cathain et al's complex intervention guidance themes mapped to conducted doctoral activity (199) .*

This work could be carried forward by a pilot randomised control trial (RCT). A stepped wedge cluster randomised control trial could be used to deploy the intervention and the primary outcome could be CVD risk factor modification in the primary care setting. A prognostic outcome (incident CVD identified at longer term follow up) could be examined however this would require a prolonged follow period. An interrupted time series could be conducted examining the period before and after the intervention and accounting for trends in outcomes of interest (204). This trial design is generally cheaper than RCTs and quicker to complete.

## 4.7 Conclusion

This chapter reports on the co-production of a care pathway to identify and manage ongoing CVD risk for patients who present to the ED with chest pain. The work is synergistic with the outputs for the systematic review and prognostic factor studies of chapter 2 and 3. In totality this work demonstrates the prognostic factors and CPMs that can best predict long term CVD in patients accessing acute care. Critically, it also defines a pathway to implement this for the benefit of patients.

# Chapter 5 Validation of TMACS across the GM-TMACS database

## 5.1 Background

The TMACS CPM has been in clinical use in the Greater Manchester region since 2016. It has been externally validated for patients with suspected cardiac chest pain (5,49–51). However, the CPM's performance has not been examined locally using routinely collected data, nor has it been examined longitudinally to see how the model performance changes over time.

Furthermore, there are now additional factors and subgroups that are of special interest for model performance. Since the derivation of TMACS age and sex dependent reference limits for high sensitivity troponin have been proposed, given that high sensitivity troponin is an integral part of TMACS the effect of these factors should be examined (205). A similar variation has also been shown with eGFR with lower kidney function being associated with higher high sensitivity troponins, this could affect the CPMs performance and also warrants further investigation (178). Ethnicity has been highlighted as another factor to consider when examining model performance, in the Framingham model the calibration and model performance was found to vary by gender and ethnicity (206,207). Another factor of interest is deprivation which has previously been shown to be associated with an increased prevalence of CVD  risk factors, and an increase in 10 year CVD mortality (208,209).

### 5.1.1 Model Performance Metrics

A CPM can have its diagnostic performance measured in two domains discrimination and calibration. Discrimination is the ability of the model to discern a positive outcome from a negative and calibration is the closeness of the model's predicted probability to the observed (actual) probability.

#### *Area under the curve*

Area under the curve is a measure of discrimination and can also be referred to as the area under the receiver operator curve, c-statistic, Harrell's c-statistics or concordance. The curve that AUC

refers to is created by plotting sensitivity against 1- specificity (210). An AUC of 1 describes a model with perfect discrimination and an AUC of 0.5 infers that the model had no predictive ability (210).

### *Calibration in the large*

Calibration in the large (CITL) is an overall measure of a model's calibration. This is where the mean predicted risk is compared to the outcome event rate (211,212). If a model had a mean predicted risk of 2% for a population and the event rate was also 2% the CITL would be 0, indicating good overall calibration. However if the model over estimated risk the CITL would be less than 0 and if it underestimated the risk it would be greater than 0.

### *Calibration plot and slope*

Calibration can be visualised by a calibration plot where the observed versus the predicted risk are displayed (212). Here the model's predictions can be grouped by decile or pre-allocated risk groups. This enables the visualisation of the model's calibration across different predicted and observed risks.

Calibration can also be assessed by measuring the calibration slope, the gradient of the points on the calibration plot (213). The best overall calibration would be a calibration slope of 1. A slope of less than one would that the risks predicted are too polarised, i.e. the high observed risk predictions are too high and the low observed risk predictions are too low (212).

### *Brier score*

The brier score is a metric that assesses both calibration and discrimination and is of interest in models with a binary outcome (214). The Brier score is calculated from the average squared error of the observed outcome compared to the predicted risk (215). The score ranges from 0 to 1, with 0 inferring the model can has no average error (perfect) and 1 the model has the maximum possible average error (worst).

## 5.2 Aims

To evaluate the diagnostic accuracy of the TMACS CPM for acute myocardial infarction in a real-world cohort of patients.

I. To evaluate the diagnostic accuracy of the TMACS decision aid for acute myocardial infarction in a real-world cohort of patients.
II. To assess the performance of TMACS in subgroups of interest including site of presentation, age quartiles, gender, ethnicity, kidney function clinical stage, deprivation quintile.
III. To assess the variation in performance of TMACS over time

## 5.3 Published work

Reynard, C., Martin, G.P., Kontopantelis, E., Jenkins, D.A., Heagerty, A., McMillan, B., Jafar, A., Garlapati, R. and Body, R., 2021. Advanced cardiovascular risk prediction in the emergency department: updating a clinical prediction model–a large database study protocol. *Diagnostic and Prognostic Research*, *5*(1), pp.1-7.

## 5.4 Methods

These methods and results are reported in keeping with the TRIPOD guidelines (45).

### 5.4.1 Design and setting

I used a multi-centre retrospective cohort study to assess the real-world performance of TMACS. Patients who presented with chest pain to the ED where the clinician suspected ACS were included. This was achieved by using sites where TMACS had been implemented in a care pathway to target these specific patients. These sites were the emergency departments at Manchester Royal Infirmary (MRI), Royal Blackburn Teaching Hospital (RBTH), and Burnley General Teaching Hospital (BGTH). The respective annual ED attendances in 2020 were 104,449, 104,009 and 44,519.

### 5.4.2 Study Population

Patients presenting to the ED with chest pain were included if the clinician suspected ACS and therefore used TMACS as part of the routine clinical pathway. Across the three clinical sites from June 2016 to October 2022 this was estimated to include approximately 14,000 patient episodes. All patients had two troponin tests recommended during the study period, which minimised verification

bias. Each site had a different troponin assay in operation, MRI used Roche's hs-cTnT, Blackburn used Siemens hs-cTnI and Burnley used an Abbott iStat point of care troponin assay.

### 5.4.3 Data Sources

At each centre, TMACS was implemented using a bespoke digital interface, which enabled the collection of data from consecutive patients who followed the pathway. At MRI, clinicians entered data directly into a macro-enabled spreadsheet (Microsoft Excel) at the point of care. This was the only way for clinicians to access departmental guidelines for management of patients with suspected acute coronary syndromes, helping to ensure complete data collection. Upon completion of data entry, all data were saved to a secure NHS server. The data were subsequently collated and linked to the hospital central laboratory database (to include results of all cardiac troponin and renal function testing).

At RBTH and BGTH, clinicians were similarly required to enter data directly into a TMACS calculator that was integrated with the electronic patient record. The data was to be cross-referenced with NHS Digital's Hospital Episode Statistics (HES) database and the civil registry. This would provide data for the outcomes of interest (AMI & MACE), even if the patient had not reattended the original hospital.

The MRI TMACS database required a validation step to check the compiling of the database had been successful. The macro-enabled spreadsheet outputted one fully populated spreadsheet per patient episode. However, clinicians were able to update their calculation by creating a newer version, which was also saved creating some duplicate entries. Given that duplicate entries are likely to be a result of corrections or newly available data, I selected the last prediction calculated by the clinician as the 'prime entry'. Validity for duplicate entries was checked using the principles of Weiskopf et al (216). Initially an automated check was planned reviewing electronically stored versus manually transcribed biochemistry data, then a random sample of 100 patient episodes underwent case note review. If the error rate was greater than 5% then all duplicate entries were to be reviewed (as per workflow in supplementary materials).

### 5.4.4 Outcomes

The primary outcome was a diagnosis of AMI within 30 days of the index date, including any subsequent admission. Patients were considered to have an AMI if they had an ICD-10 coded diagnosis present as a primary discharge code on any admission within 30 days. A sensitivity analysis was conducted to expand the outcome definition to any AMI code regardless of position. The relevant AMI codes included I21, I22 or I23 (see Supplementary Table 8.12). There was also a secondary composite outcome of major adverse cardiovascular events (MACE) within 30 days. MACE was defined as AMI, death and revascularisation. ICD-10 diagnostic codes included I21, I22, I23, I46, R96, R99. OPCS intervention codes included K40-50, K63, and K75 (see Supplementary Table 8.12).

### 5.4.5 Statistical analysis

I calculated the output of TMACS using the algorithm as it was used in clinical practice during the study period, which includes minor alterations to the originally published coefficients (unpublished – figure 5.1). I used the c-statistic to measure discrimination and confidence intervals were calculated using the method described by DeLong et al (217). Calibration was evaluated through calibration plots in conjunction with CITL and calibration slopes. The Brier score was used to capture overall model accuracy. I also observed the diagnostic characteristics longitudinally, splitting the cohort into annual quarters. Sensitivity of TMACS was calculated using the very low risk group threshold (<2%. Predicted risk). Specificity of TMACS was calculated using the high-risk group (> 95% predicted risk).

I pre-specified additional variables/subgroups of interest including age (quartile), gender, site, kidney function, ethnicity, and deprivation (indices of multiple deprivation – quintiles).

The model was also validated longitudinally to assess for any temporal change in its performance, this was done by quarter (January – March, April – June, July – September and October – December).

$$l = log_b \frac{p}{1-p} = 1.828x_e + 1.514x_a + 0.849x_r + 1.783x_v + 1.878x_s + 1.412x_h + 0.089x_t - 4.65$$

Figure 5.1 – The TMACS clinical prediction model. $l$ = log-odds of the primary outcome acute myocardial infarction, $x_e$ = presence of ECG ischaemia, $x_a$ = crescendo angina, $x_r$ = paint radiating to the right arm, $x_v$= pain associated with vomiting, $x_s$ = sweating observed, $x_h$ = hypotension, and $x_t$= is high sensitivity troponin T result on arrival.

## 5.5 Results

### 5.5.1 Cohort description

A total of 13,207 patient episodes were collated from MRI (10,710), BGTH (1,242) and RBTH (1,255) from June 2016 to October 2020. A validation step was undertaken for MRI data, given that data entry into the TMACS interface relied upon manual input. First, the troponin results entered into the TMACS calculator were compared with the central laboratory database. This identified an error rate of 2.9% of (n = 312). Because this was a real-world assessment of TMACS, these values were not corrected, although we ran a sensitivity analysis using the correct initial laboratory results. Second, 100 cases were randomly selected for case note review to compare the data taken from the TMACS database with the output of TMACS that was documented in patient case notes. No mismatches were identified. This did not meet the prespecified threshold of a 5% error rate to require a full set of case note reviews, therefore the data validation was deemed to have been passed.

Of the 13,207 patients, 81.2% were from MRI and 46.8% were male. The mean age was 56.4 and 31.9% were of white ethnicity (Table 5.1). The mean predicted probability was 13.7%. However, the most frequent risk category was ‘very low risk’ (calculated probability of AMI <2%), accounting for 42.3% of patients (Table 5.1). The high proportion of very low risk predictions caused a positive skew in the distribution of predicted risk (**Error! Reference source not found.**)

| Feature | N (%) |
|---|---|
| | |
| **Total number of patients** | 13207 |
| Manchester Royal Infirmary | 10710 (81.2%) |
| Burnley | 1242 (9.4%) |
| Blackburn | 1255 (9.5%) |
| | |
| | |
| Male | 6175 (46.8%) |
| Female | 4535 (34.4%) |
| Missing | 2497 (18.9%) |
| | |
| **Age** | |
| Mean | 56.38 |
| Standard Deviation | 16.91 |
| Range | 16-103 |
| Missing | 0 |
| | |
| **Ethnicity** | |
| Asian | 2435 (18.4%) |
| Black | 756 (5.7%) |
| Mixed | 472 (3.6%) |
| Other | 347 (2.6%) |
| White | 4213 (31.9%) |
| Missing | 4984 (37.7%) |
| | |
| **Original Predicted Probability** | |
| Mean | 13.68% |
| Range | 1.12 – 100.00 % |
| Standard Deviation | 24.29 |
| Missing | 13 |
| | |
| **Risk category** | |
| Very low risk (<2%) | 5580 (42.3%) |
| Low risk (2<=x>5%) | 2213 (16.8%) |
| Moderate risk (5<=x>95%) | 4897 (37.1%) |
| High risk (>95%) | 504 (3.8%) |
| | |
| **Index of Multiple Deprivation** | |
| Mean | 39.27 |
| Range | 1.03-81.76 |
| Standard deviation | 17.86 |
| Missing | 3408 |
| | |

*Table 5.1 TMACS Cohort Demographics*

There was confirmed linkage with NHSD in 74.76% of cases where information on the index admission was returned from NHSD (Table 5.2). The remaining 25.2% did not have any data returned, this is confirmation that no index event data was successfully linked however it is possible that beyond the index episode the care record was successfully searched but no further episodes found (i.e. no link to index episode but linked to general record with no primary outcome present). The NHSD linkage was only with the admitted patient care dataset, so for a successful link the patient had to be admitted. The 25.2% with no data returned is in keeping with the literature for the proportion of patients discharged directly from the Emergency Department, in whom we would not expect to see linkage with NHSD's admitted patient care dataset (218).

The incidence of the primary outcome was lower than originally anticipated in the cohort overall (3.7%), though the incidence of AMI was marginally higher at RBTH (4.3%) and BGTH (4.6%) (Table 5.3). MRI had lower rates of the primary outcome with an incidence of 3.6%. When any AMI coded in any position was considered the incidence rose to 4.2%. For MACE the overall incidence was 5.7% in the primary code position and 6.1% in any position.

| Group | N (%) |
|---|---|
| No data returned | 25.2% (3334) |
| Linked | 74.8% (9873) |
| Fully linked | 57.2% (7551) |
| Linked but date time error up to 24 hours | 15.6% (2055) |
| Linked but no index episode data | 2.0% (261) |
| Local outcome linkage | 0.1% (6) |

*Table 5.2 Linkage status between local electronic patient record data and national data held by NHS digital*

| Outcome group | Incidence |
|---|---|
| **Acute myocardial infarction (position 1)** | **3.7%** |

| | |
|---|---|
| MRI | 3.6% |
| Burnley | 4.6% |
| Blackburn | 4.3% |
| **Acute myocardial infarction (any position)** | **4.2%** |
| MRI | 4.1% |
| Burnley | 4.7% |
| Blackburn | 4.8% |
| **Major adverse cardiovascular event (position 1)** | **5.7%** |
| MRI | 5.5% |
| Burnley | 5.7% |
| Blackburn | 7.5% |
| **Major adverse cardiovascular event (any position)** | **6.1%** |
| MRI | 6.0% |
| Burnley | 5.8% |
| Blackburn | 7.9% |

*Table 5.3 Incidence for outcomes of interest. Position 1 denotes where the relevant codes were found in the primary position. Any position denotes where the codes were found at any position in the electronic patient record.*

### 5.5.2 Overall Model performance

TMACS demonstrated good discrimination for the primary outcome with an AUC of 0.88 (95% CI 0.86 to 0.89) (Figure 5.1). The sensitivity of TMACS to 'rule out' AMI (very low risk group versus all other risk groups) was 97.2% (95% CI 95.7 to 98.6) with a negative predictive value of 99.7% (95% CI 99.6 to 99.9%) (Table 5.4). There were 14 false negatives for the primary outcome (Table 5.5). These cases were examined, and a number were found to be likely erroneous due to issues with the outcome and predictions. Due to data privacy agreements, small subgroups cannot be reported, and as such it is not possible to specify the precise number of suspected erroneous false negatives or the directly attributable rationale. The rationales included incorrect troponin values and differing codes across local and national databases. If all suspected erroneous false negative cases are discounted then the sensitivity and NPV rise to 97.8% (95% CI 96.0 to 98.9%) and 99.9% (95% CI 99.7 to 99.9%), respectively. This results in an AUC of 0.88 (95% CI: 0.87 to 0.90).

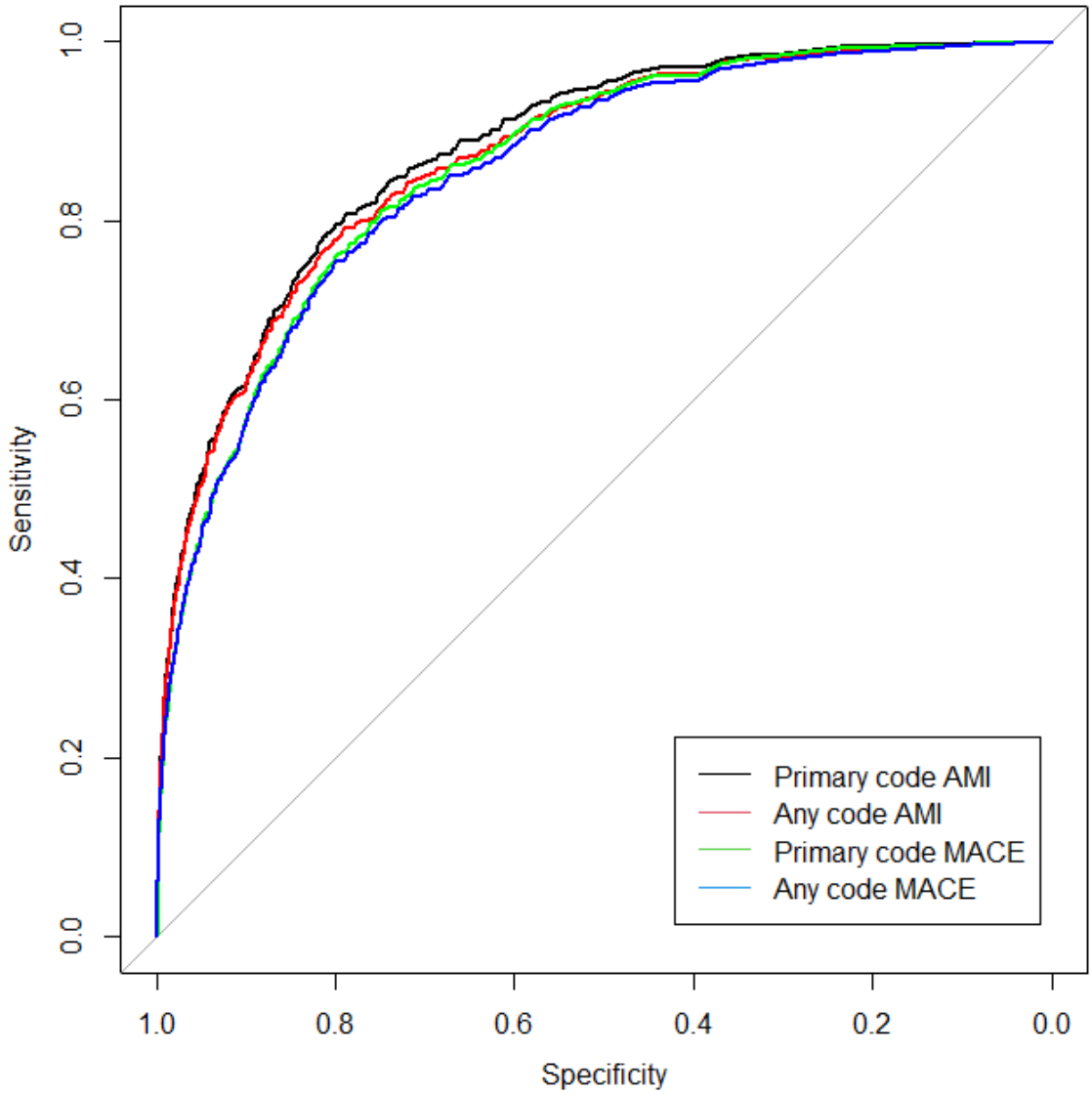



*Figure 5.1 Receiver operator curve for TMACS predicting different outcomes of interest. AMI – acute myocardial infarction, MACE – major adverse cardiac event*

| | Sensitivity VLR (95% CI) | Specificity HR (95% CI) | Negative predictive value VLR (95% CI) | Positive predictive value HR (95% CI) |
|---|---|---|---|---|
| **AMI (P1)** | 97.2 (95.7 to 98.6) | 40.1 (35.8 to 44.4) | 99.7 (99.6 to 99.9) | 97.7 (96.3 to 99.0) |
| **AMI (any)** | 96.4 (94.9 to 98.0) | 38.7(34.6 to 42.7) | 99.6 (99.5 to 99.8) | 97.3 (95.9 to 98.7) |
| **MACE (P1)** | 96.1 (94.8 to 97.5) | 32.0 (28.6 to 35.3) | 99.5 (99.3 to 99.6) | 96.0 (94.3 to 97.7) |
| **MACE (any)** | 95.4 (94.0 to 96.9) | 31.5 (28.3 to 34.7) | 99.3 (99.2 to 99.5) | 95.6 (93.8 to 97.4) |

*Table 5.4 Diagnostic characteristics for the different risk categories predicted by TMACS. VLR = Very low risk and HR = high risk, CI – confidence interval, AMI – acute myocardial infarction, MACE – major adverse cardiac event, P1 – qualifying code at position one only.*

| | Acute Myocardial infarction (position 1) | |
|---|---|---|
| | Present | Absent |
| Very low risk | 14 | 5566 |
| Greater than very low risk | 480 | 7134 |

*Table 5.5 2x2 diagnostic table for TMACS very low risk group versus the primary outcome – acute myocardial infarction coded in first position*

The calibration of TMACS demonstrated significant over prediction with an overall CITL of -3.93 (95% CI -4.12 to -3.74) and a O/E of 0.27 indicating significant overestimation of risk. The calibration slope was 0.05 which indicates significant miscalibration. This overestimation of risk is again shown in the calibration plot for the primary outcome in figure 5.8. The highest TMACS prediction decile substantially overestimated the risk, however the lower 7 deciles are clustered very tightly below 0.2 expected risk (2% predicted risk). When these deciles are examined closely, the calibration curve is more favourable (Figure 5.2 and Figure 5.3).

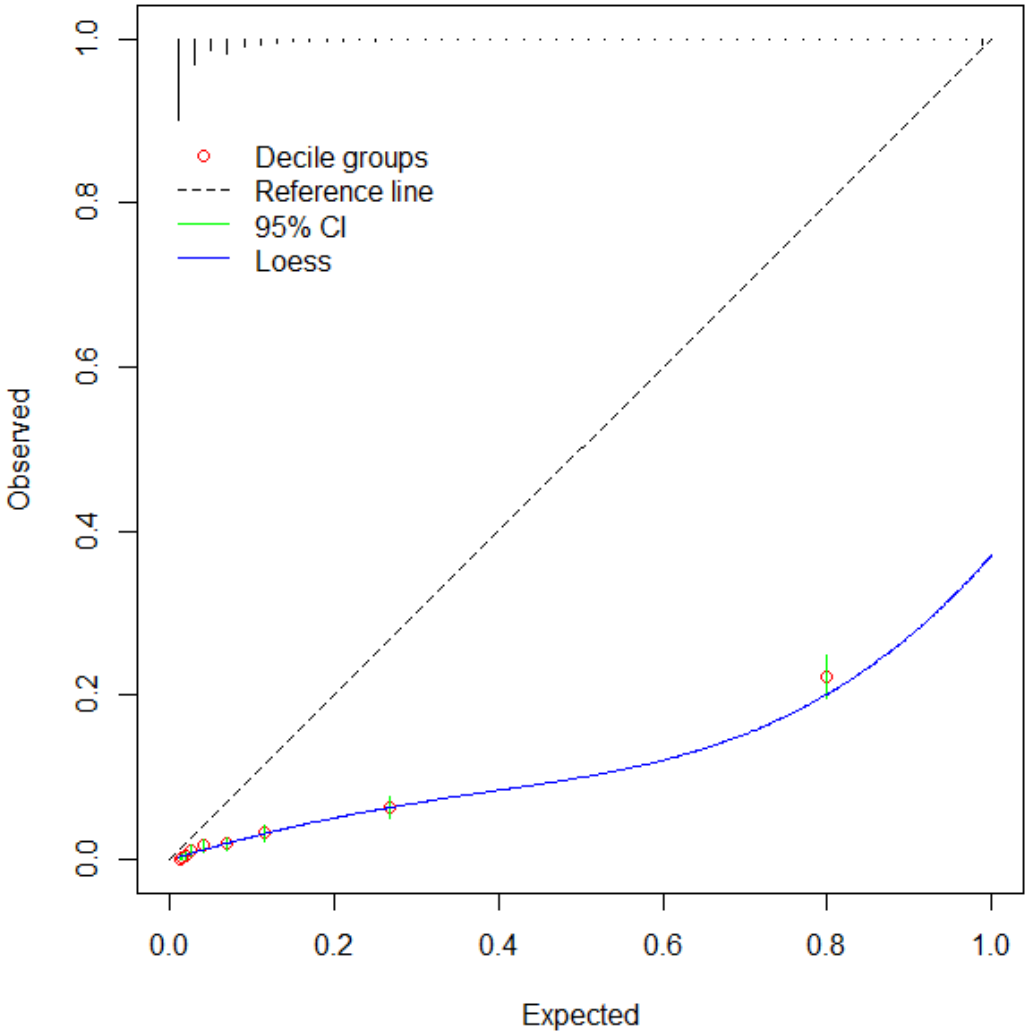



*Figure 5.2 Calibration plot for TMACS predicting the primary outcome. TMACS predictions were grouped into deciles and a Loess curve was fitted to the values.*

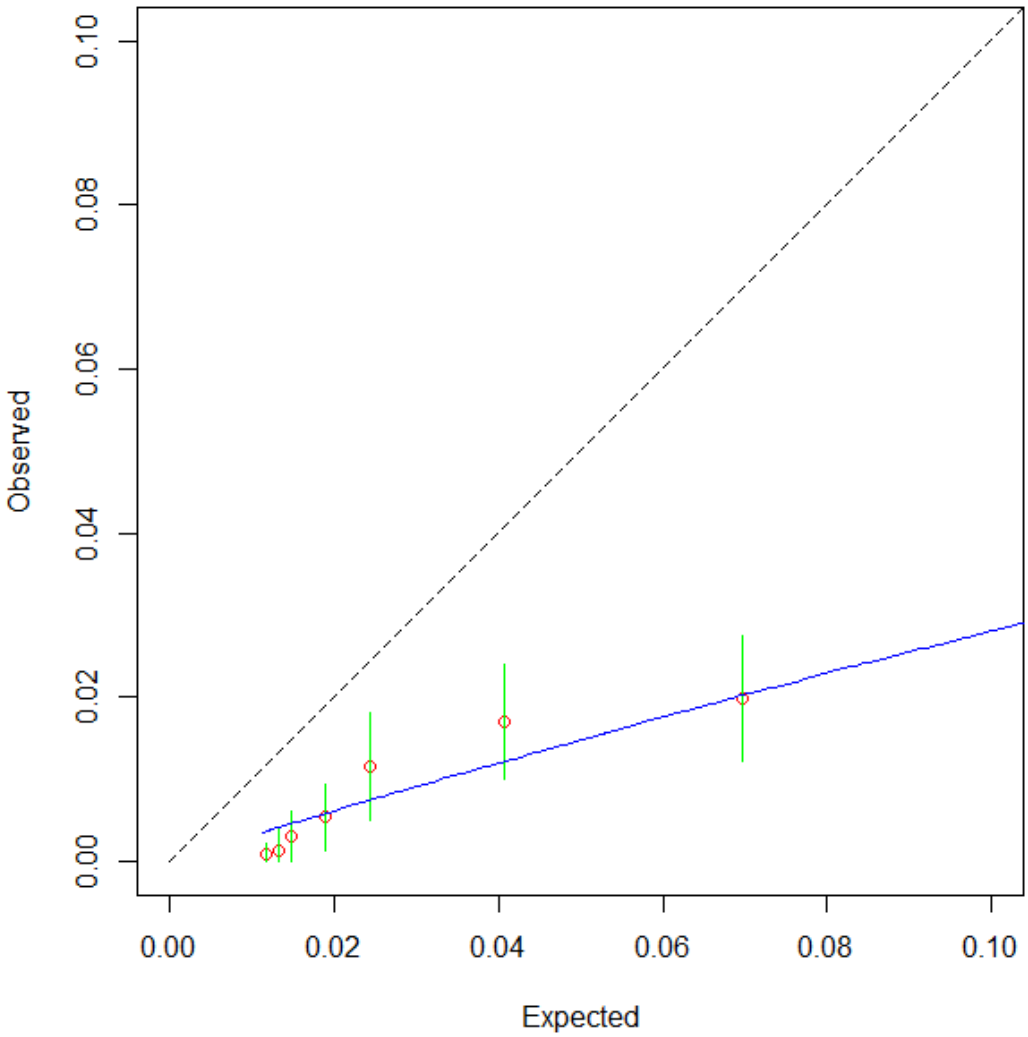



*Figure 5.3 Focused calibration for TMACS predicting the primary outcome. TMACS predictions were grouped into deciles and a Loess curve was fitted to the values.*

For the secondary outcome of MACE, the AUC was 0.86 (95% CI 0.85 to 0.87) and CITL -2.32 (95% CI -2.45 to -2.19). Using the broader outcome definition incorporating codes from any position more outcomes were identified. This decreased the performance of TMACS for the primary and secondary outcome across both discrimination and calibration metrics, AUC decreased by 0.09 for AMI and 0.06 for MACE (Table 5.6).

| | AUC (95% CI) | CITL (95% CI) | C-slope (95% CI) | Brier |
|---|---|---|---|---|
| AMI (P1) | 0.88 (0.86 to 0.89) | -3.93 (-4.12 to -3.74) | 0.05 (0.04 to 0.06) | 0.07 |
| AMI (any) | 0.87 (0.86 to 0.89) | -3.41 (-3.58 to -3.24) | 0.06 (0.05 to 0.07) | 0.07 |
| MACE (P1) | 0.86 (0.85 to 0.87) | -2.32 (-2.45 to -2.19) | 0.09 (0.07 to 0.10) | 0.07 |
| MACE (any) | 0.86 (0.84 to 0.87) | -2.07 (-2.20 to -1.95) | 0.09 (0.07 to 0.10) | 0.07 |

*Table 5.6 TMACS discrimination and calibration metrics for different outcomes. AMI (P1) – acute myocardial infarction position 1, AMI (any) acute myocardial infarction coded in any position, MACE (P1) – major adverse cardiac event in position 1, MACE (any) – major adverse cardiac event in any position, AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

### 5.5.3 Subgroups of interest

Six subgroups of interest were examined in the cohort: site of presentation, age quartiles, gender, ethnicity, kidney function clinical stage, deprivation quintile.

#### *Geography*

The diagnostic performance of TMACS at each site was examined and heterogeneity was identified. The AUC for MRI was lower than the pooled AUC at 0.87 (95% CI 0.85 to 0.89), in contrast to Blackburn and Burnley where this discrimination metric was higher at 0.91 (95% CI 0.86 to 0.95) and 0.92 (95% CI 0.89 to 0.96) (Table 5.7and Figure 5.4). The calibration for MRI was slightly better (but still shows miscalibration) than the pooled estimate with a CITL of -3.00 (95% CI -3.18 to -2.82). Blackburn and Burnley demonstrated poorer calibration with a lower CITL of -8.34 (95% CI -9.29 to -7.38) and -10.63 (95% CI -11.63 to 9.63) (Figure 5.5). The calibration slope for MRI was slightly improved over the pooled results at 0.20 vs 0.05, however this still demonstrate significant miscalibration.

| | AUC (95% CI) | CITL (95% CI) | C-slope (95% CI) | Brier |
|---|---|---|---|---|
| Overall | 0.88 (0.86 to 0.89) | -3.93 (-4.12 to -3.74) | 0.05 (0.04 to 0.06) | 0.07 |
| MRI | 0.87 (0.85 to 0.89) | -3.00 (-3.18 to -2.82) | 0.20 (0.17 to 0.22) | 0.06 |
| RBTH | 0.91 (0.86 to 0.95) | -8.34 (-9.29 to -7.38) | 0.01 (0.00 to 0.01) | 0.10 |
| BGTH | 0.92 (0.88 to 0.96) | -10.63 (-11.63 to -9.63) | 0.03 (0.02 to 0.04) | 0.12 |

*Table 5.7 TMACS discrimination and calibration metrics for sites. MRI - Manchester Royal Infirmary, RBTH - Royal Blackburn Teaching Hospital, BGTH - Burnley General Teaching Hospital, AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

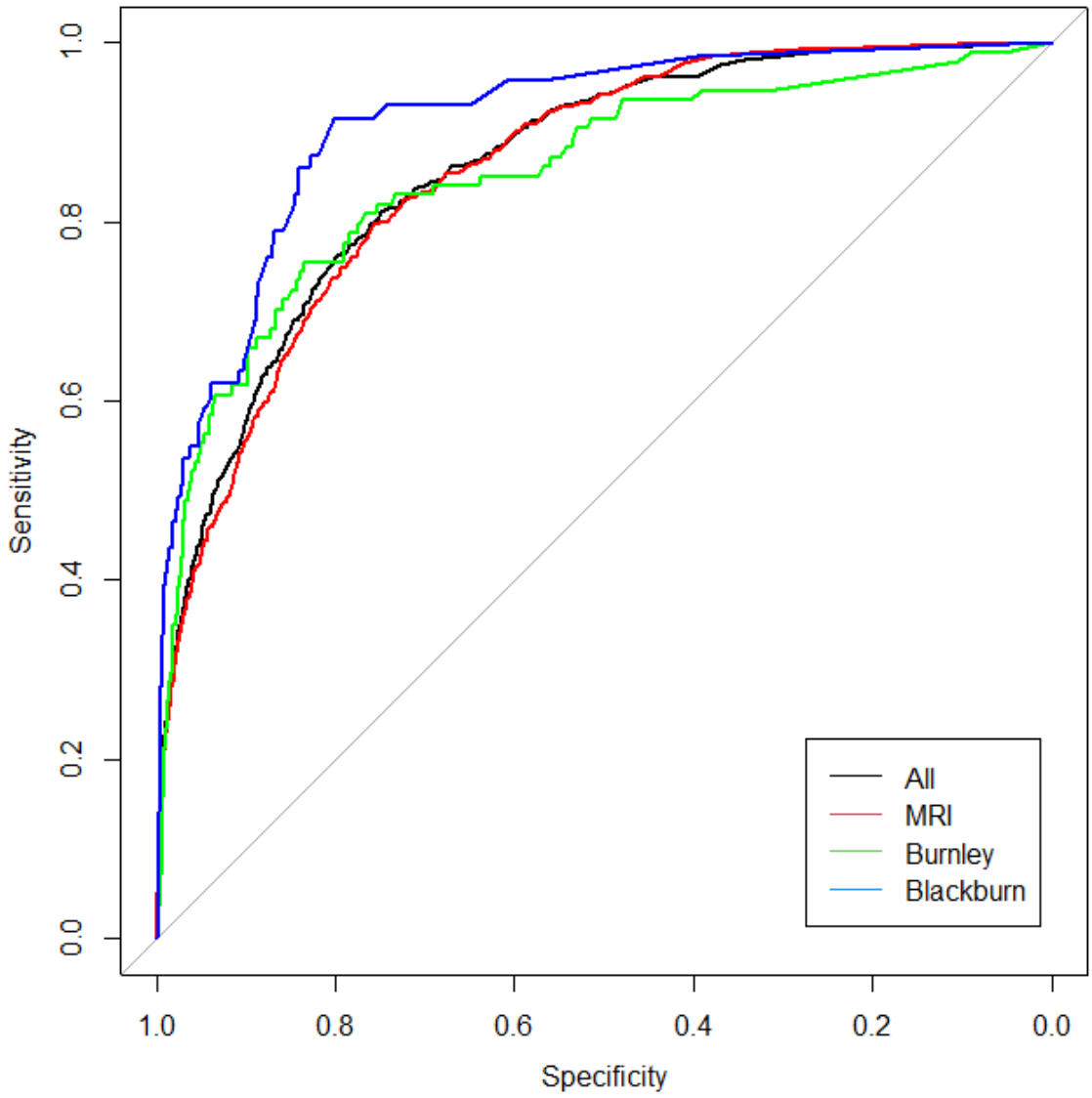



*Figure 5.4 Receiver operator curve for TMACS predicting the primary outcome separated into the different sites. MRI – Manchester Royal Infirmary, Burnley - Burnley General Teaching Hospital, and Blackburn - Royal Blackburn Teaching Hospital*

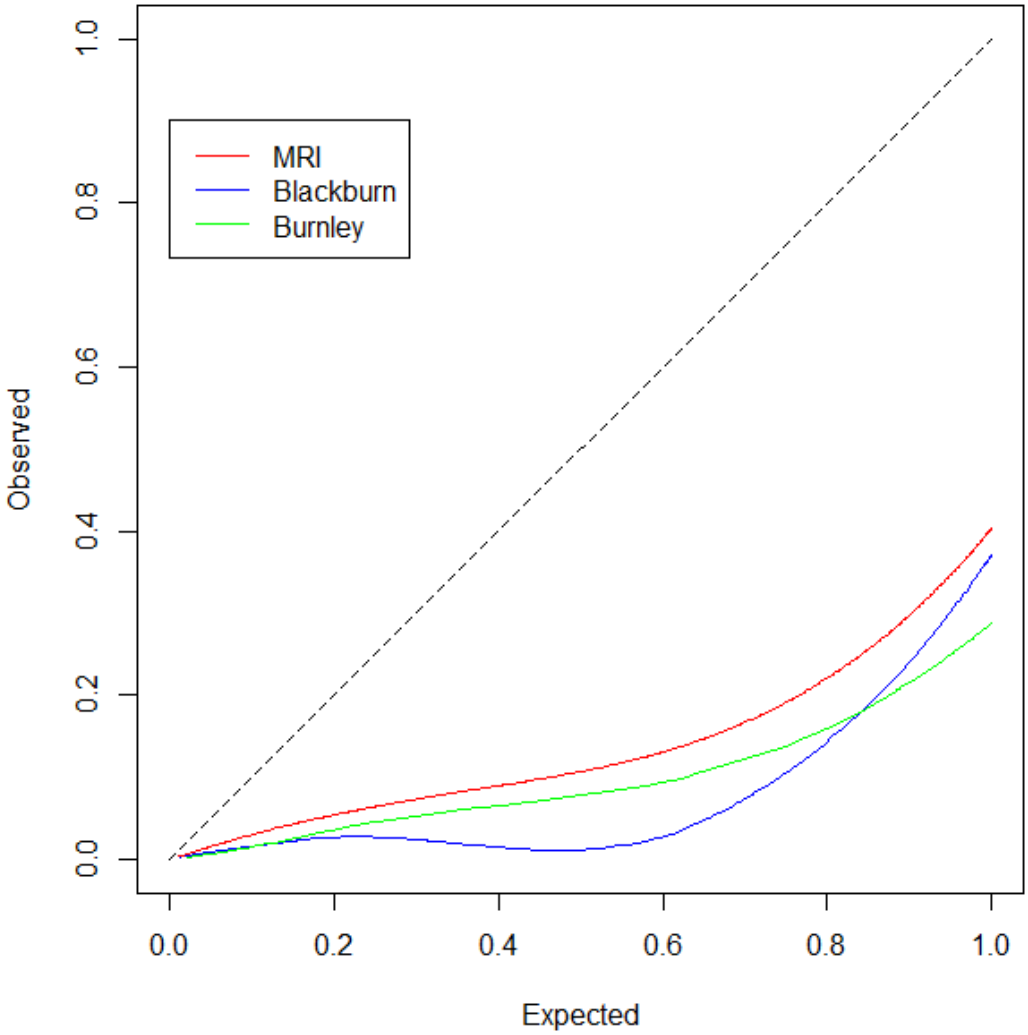



*Figure 5.5 Calibration plot for TMACS predicting the primary outcome by site. A Loess curve was fitted to the values of the different subgroups. MRI – Manchester Royal Infirmary, Burnley - Burnley General Teaching Hospital, and Blackburn - Royal Blackburn Teaching Hospital*

### *Age*

Age was examined in quartiles, for this cohort that created groups for the ages of <44 years old, 44 to 55, 56 to 69 and greater than 69. The discrimination deteriorated with increasing age, the first quartile had an AUC of 0.92 (95% CI 0.88 to 0.95) compared to the last quartile with an AUC of 0.85 (95% CI 0.82 to 0.88) (Table 5.8 and Figure 5.6). The calibration across the age quartiles was also found to be variable but not in the same manner as AUC. The second and third age quartile had lower CITL compared to the first and last quartile indicating better calibration. However, all quartiles demonstrated systemic overestimation of risk (Table 5.8 and Figure 5.7)

| **Age** | **AUC (95% CI)** | **CITL (95% CI)** | **C-slope (95% CI)** | **Brier** |
|---|---|---|---|---|
| Q1 | 0.92 (0.88 to 0.95) | -4.71 (-5.42 to 3.99) | 0.03 (0.01 to 0.05) | 0.03 |
| Q2 | 0.89 (0.86 to 0.91) | -3.35 (-3.73 to 2.96) | 0.29 (0.24 to 0.33) | 0.05 |
| Q3 | 0.86 (0.83 to 0.89) | -3.21 (-3.51 to -2.92) | 0.07 (0.05 to 0.09) | 0.08 |
| Q4 | 0.85 (0.82 to 0.88) | -4.83 (-5.19 to -4.48) | 0.04 (0.03 to 0.05) | 0.12 |

*Table 5.8 - TMACS discrimination and calibration metrics for age quartiles. AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

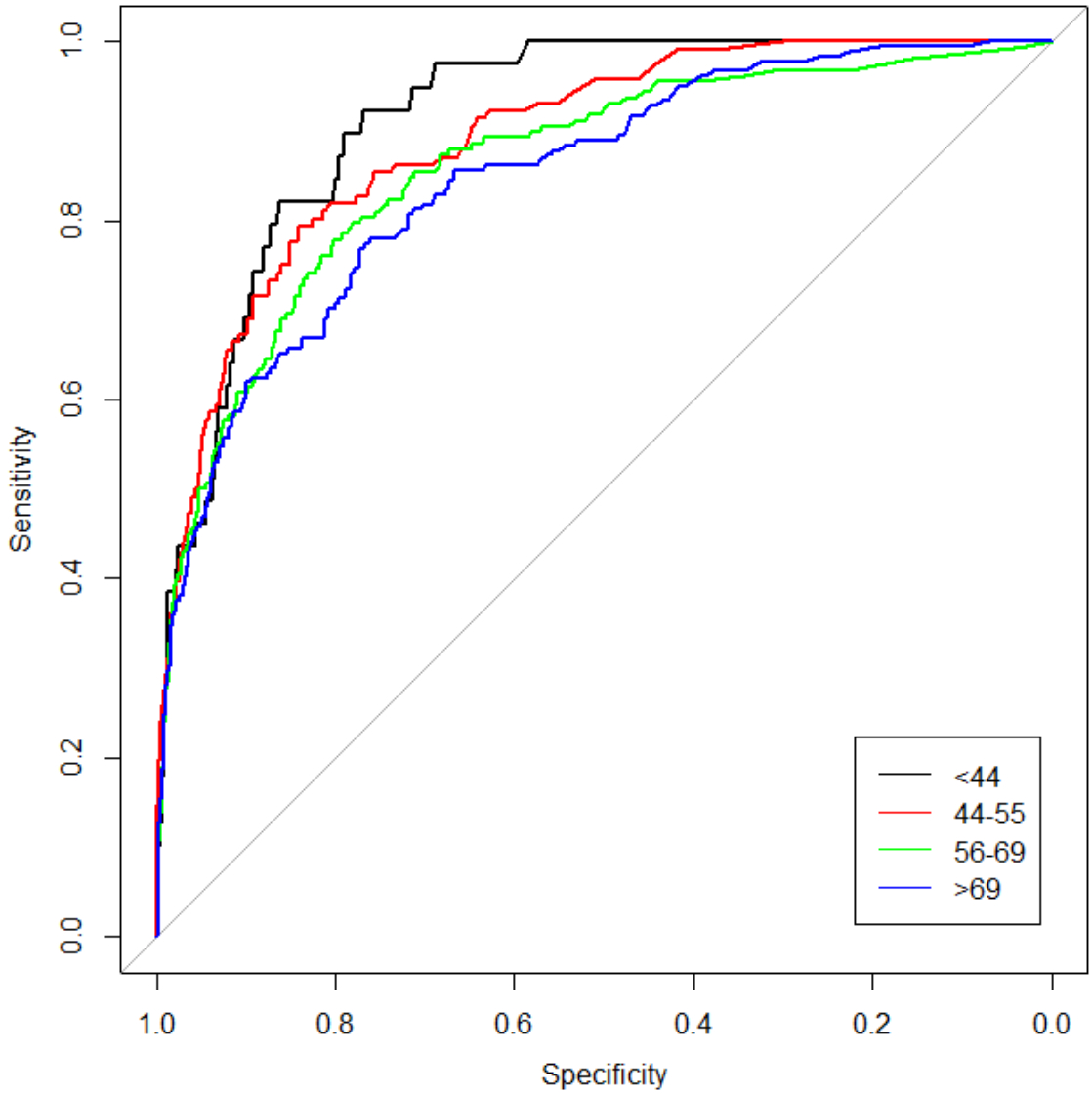



*Figure 5.6 Receiver operator curve for TMACS predicting the primary outcome separated into age quartile.*

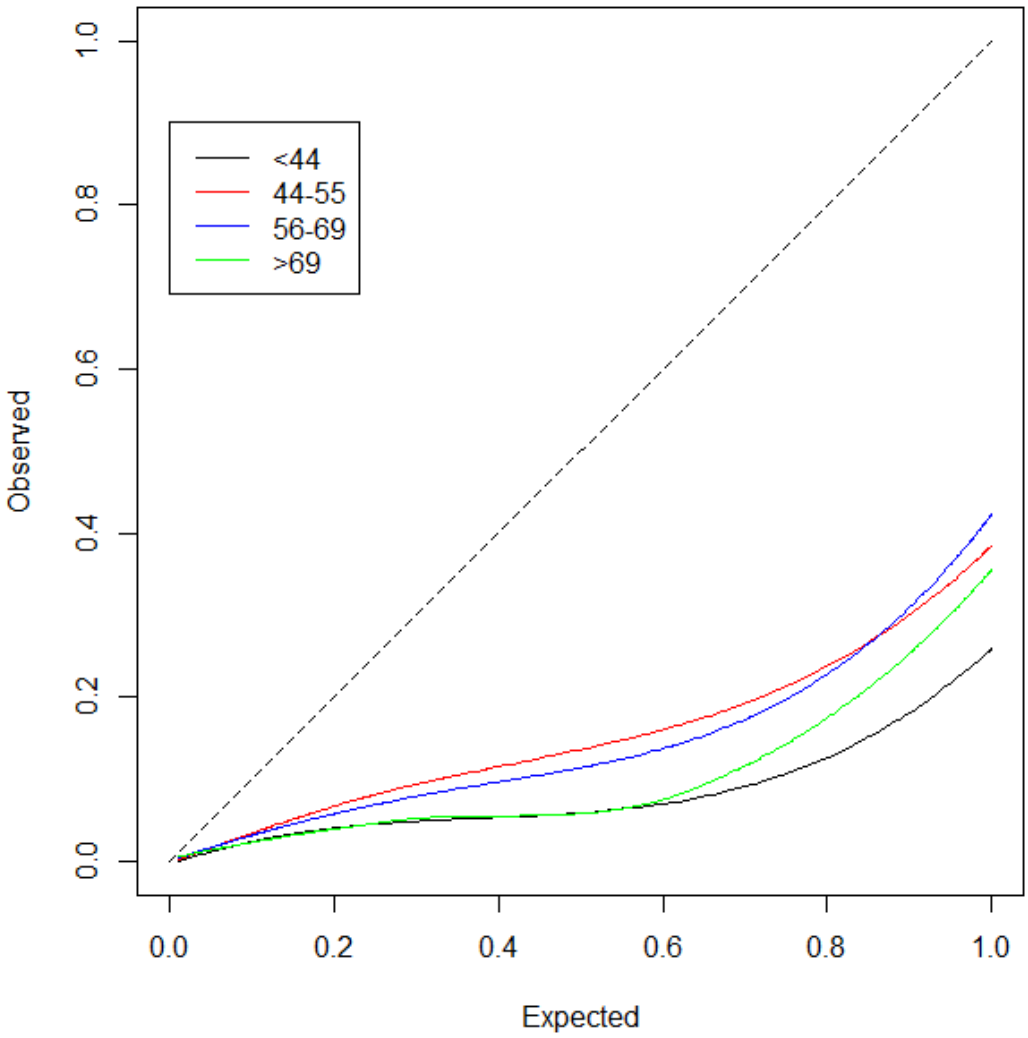



*Figure 5.7 Calibration plot for TMACS predicting the primary outcomes separated by age quartile. A Loess curve was fitted to the values of the different age quartiles*

### *Gender*

The AUC for females was 0.89 (95% CI 0.87 to 0.92) versus 0.87 for men (95% CI 0.85 to 0.89) (Table 5.9 and Figure 5.8). Conversely the calibration was better for men with a CITL of -3.58 (95% CI -3.81 to -3.35) (Figure 5.9). This was also reflected in the c-slope, however it still demonstrated significant miscalibration.

| Gender | AUC | CITL | C-slope | Brier |
|---|---|---|---|---|
| Male | 0.87 (0.85 to 0.89) | -3.58 (-3.81 to -3.35) | 0.10 (0.09 to 0.12) | 0.08 |
| Female | 0.89 (0.87 to 0.92) | -4.25 (-4.58 to -3.92) | 0.04 (0.02 to 0.05) | 0.06 |

*Table 5.9 TMACS discrimination and calibration metrics for age quartiles. AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

### *Ethnicity*

Ethnicity was examined in five groups: Asian, Black, Mixed, Other and White. The discrimination and calibration were variable across the groups with no group consistently better or worse across both domains (Table 5.10). Mixed ethnicity had the highest AUC of 0.92 (95% CI 0.87 to 0.97) whereas White ethnicity demonstrated the lowest with an AUC of 0.87 (95% CI 0.85 to 0.89) (Figure 5.10). CITL was lowest in those with Asian ethnicity at -2.52 (95% CI -2.80 to 2.24) and highest in Black ethnicity at -5.98 (95% CI 6.86 to -5.10). Similarly for c-slope Asian ethnicity were the least mis-calibrated group at 0.22. The calibration plot demonstrated that for mixed ethnicity the calibration across different predicted risks varied and was comparatively worse at higher predicted risks (Figure 5.11).

| Ethnicity | AUC | CITL | C-slope | Brier |
|---|---|---|---|---|
| Asian | 0.87 (0.84 to 0.90) | -2.52 (-2.80 to -2.24) | 0.22 (0.18 to 0.26) | 0.06 |
| Black | 0.90 (0.83 to 0.97) | -5.98 (-6.86 to -5.10) | 0.15 (0.07 to 0.22) | 0.06 |
| Mixed | 0.92 (0.87 to 0.97) | -3.71 (-4.60 to -2.81) | 0.10 (0.06 to 0.15) | 0.05 |
| Other | 0.92 (0.86 to 0.97) | -3.86 (-4.91 to -2.82) | 0.13 (0.05 to 0.22) | 0.06 |
| White | 0.87 (0.85 to 0.89) | -4.03 (-4.28 to -3.77) | 0.03 (0.02 to 0.04) | 0.08 |

*Table 5.10 TMACS discrimination and calibration metrics by ethnicity. AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

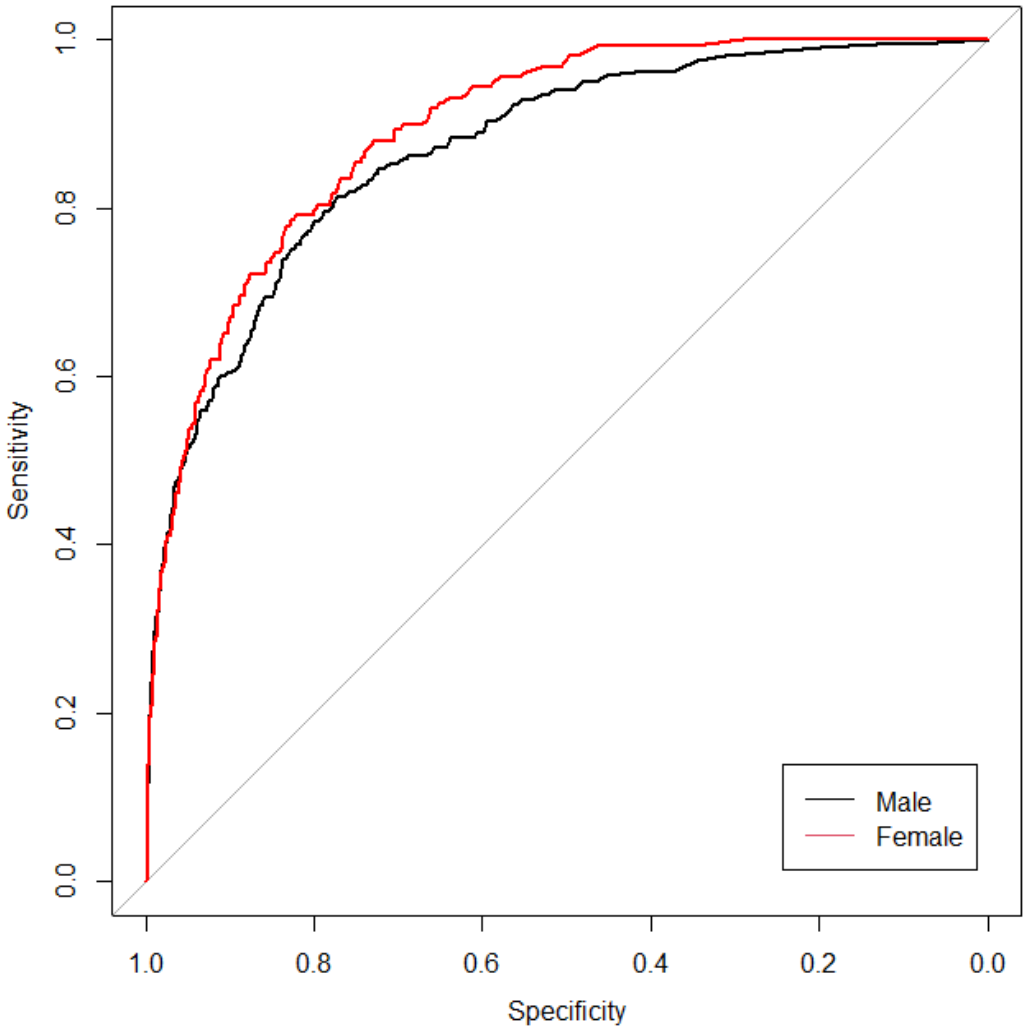



*Figure 5.8 Receiver operator curve for TMACS predicting the primary outcome separated by gender.*

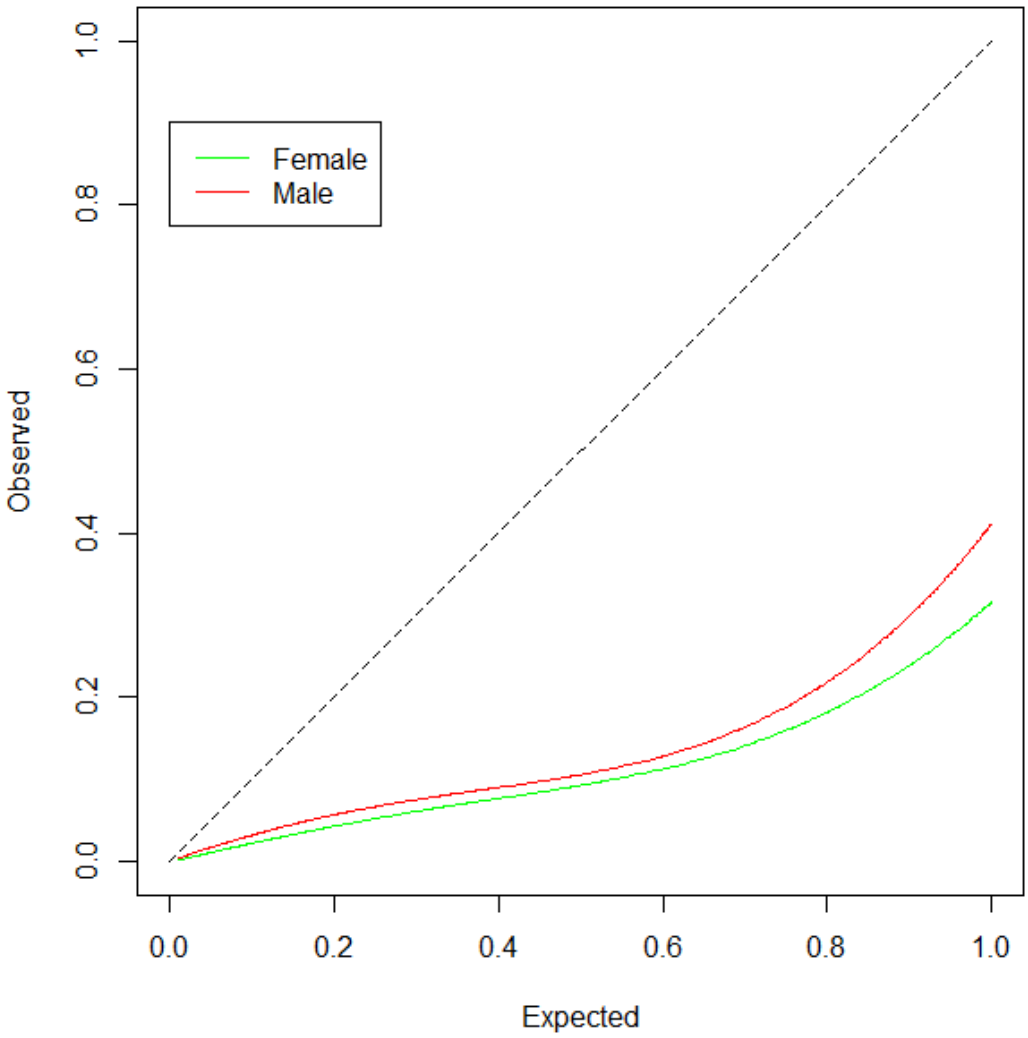



*Figure 5.9 Calibration plot for TMACS predicting the primary outcomes separated by gender. A Loess curve was fitted to the values*

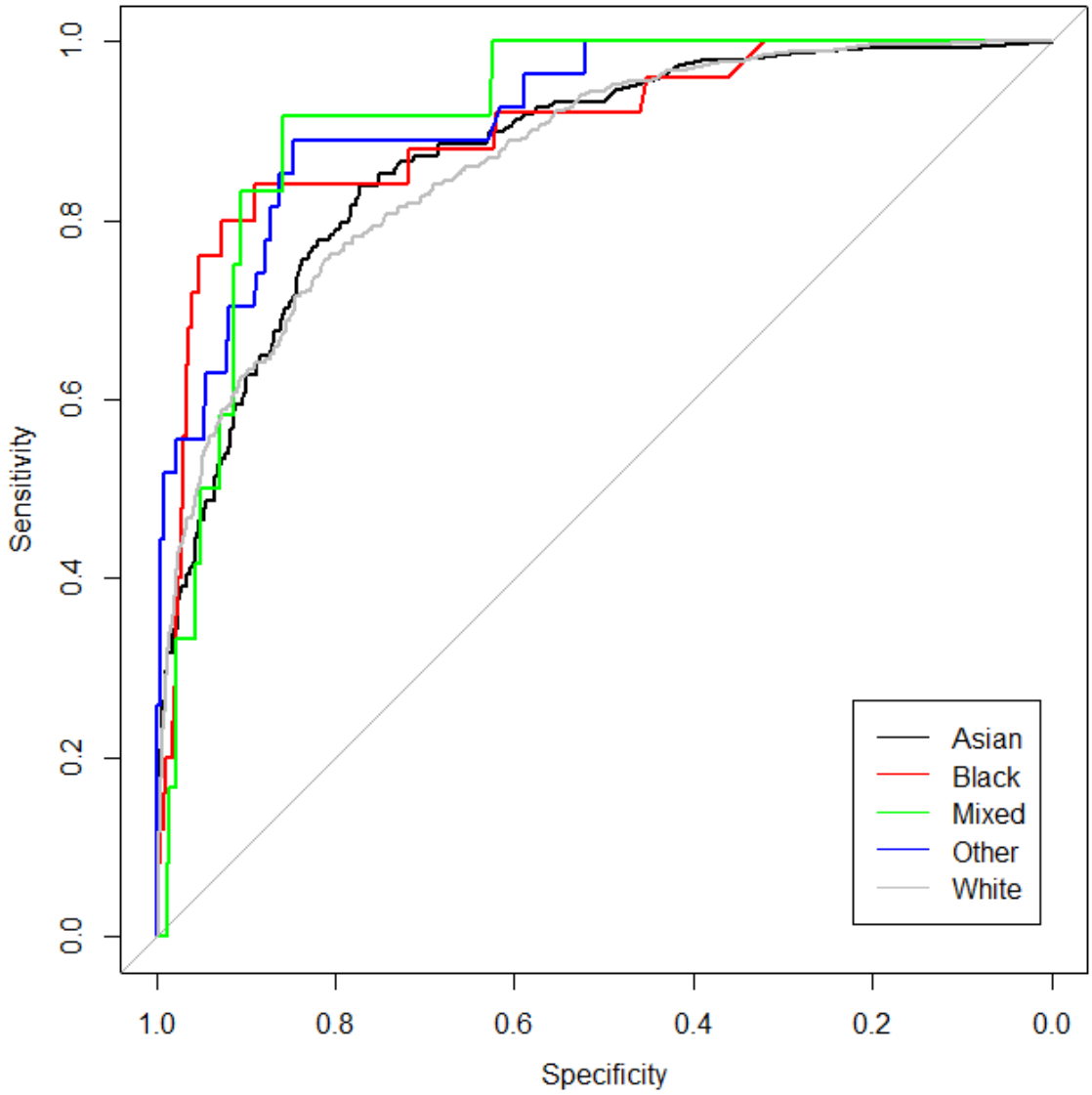



*Figure 5.10 - Receiver operator curve for TMACS predicting the primary outcome separated by ethnicity.*

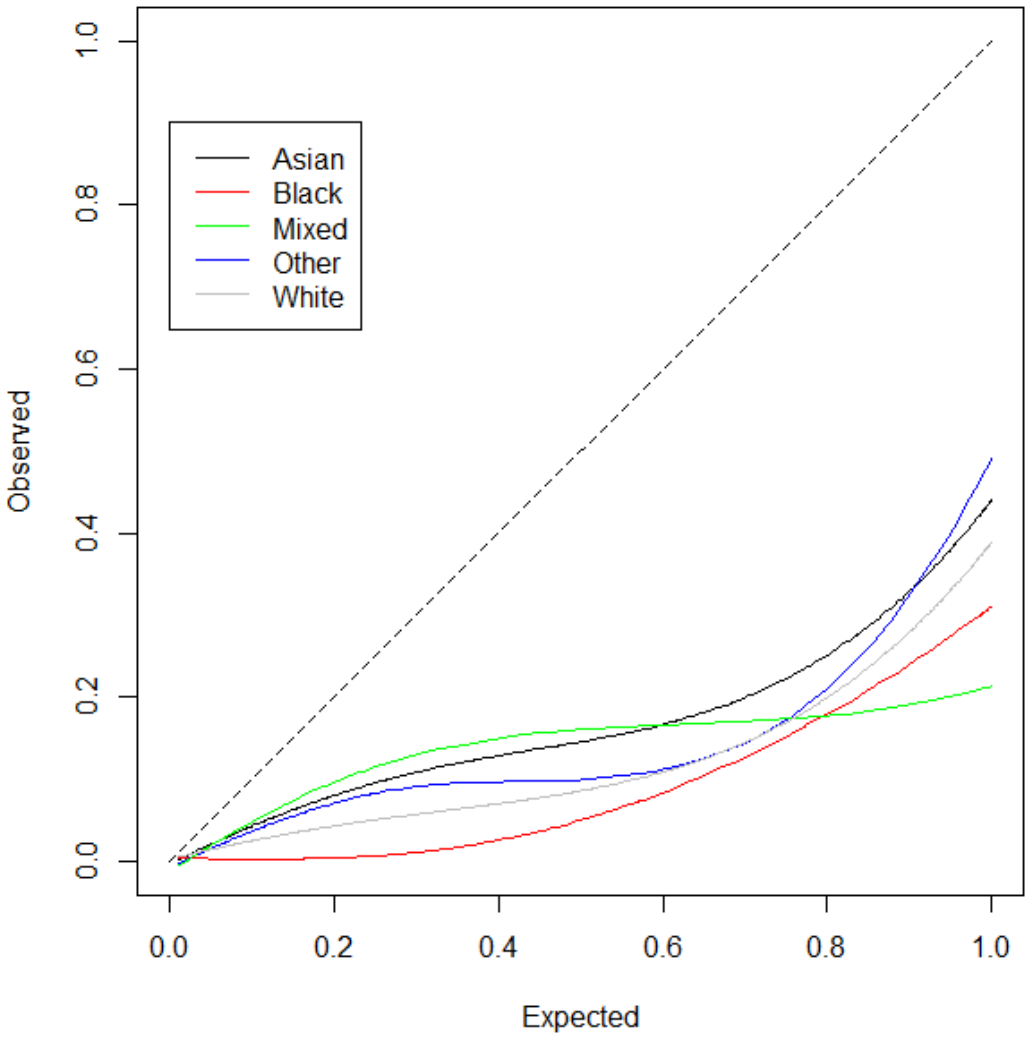



*Figure 5.11 - Calibration plot for TMACS predicting the primary outcomes separated by ethnicity. A Loess curve was fitted to the values*

### *Kidney function*

TMACS performance was examined by kidney function stratified by chronic kidney disease (CKD) classification as per the Kidney Disease: Improving global Outcomes guideline (219). In contrast to other subgroups the discrimination and calibration performance followed a similar pattern of deterioration as the kidney function worsened with increasing stage of chronic kidney disease (Table 5.11). The AUC in the CKD stage 1 group was 0.91 (95% CI 0.88 to 0.93) and it decreases to 0.78 (0.54 to 1.00) in the stage 5 group (Figure 5.12). Similarly, the CITL starts at -2.08 (95% CI -2.36 to -1.80) and decreases to -14.96 (95% CI 17.28 to 12.65). The calibration plot indicates that this hierarchy of calibration is persistent across predicted risk (Figure 5.13).

| **CKD stage** | **AUC** | **CITL** | **C-slope** | **Brier** |
|---|---|---|---|---|
| I (or <) | 0.91 (0.88 to 0.93) | -2.08 (-2.36 to -1.80) | 0.40 (0.35 to 0.46) | 0.03 |
| II | 0.87 (0.84 to 0.90) | -2.32 (-2.57 to -2.08) | 0.32 (0.27 to 0.37) | 0.06 |
| III | 0.81 (0.75 to 0.86) | -3.63 (-4.05 to -3.20) | 0.12 (0.08 to 0.16) | 0.12 |
| IV | 0.89 (0.80 to 0.98) | -7.08 (-8.59 to -5.56) | 0.21 (0.10 to 0.32) | 0.27 |
| V | 0.78 (0.54 to 1.00) | -14.96 (-17.28 to -12.65) | 0.05 (-0.00 to 0.11) | 0.56 |

*Table 5.11 TMACS discrimination and calibration metrics by Chronic Kidney Disease (CKD) stage. AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

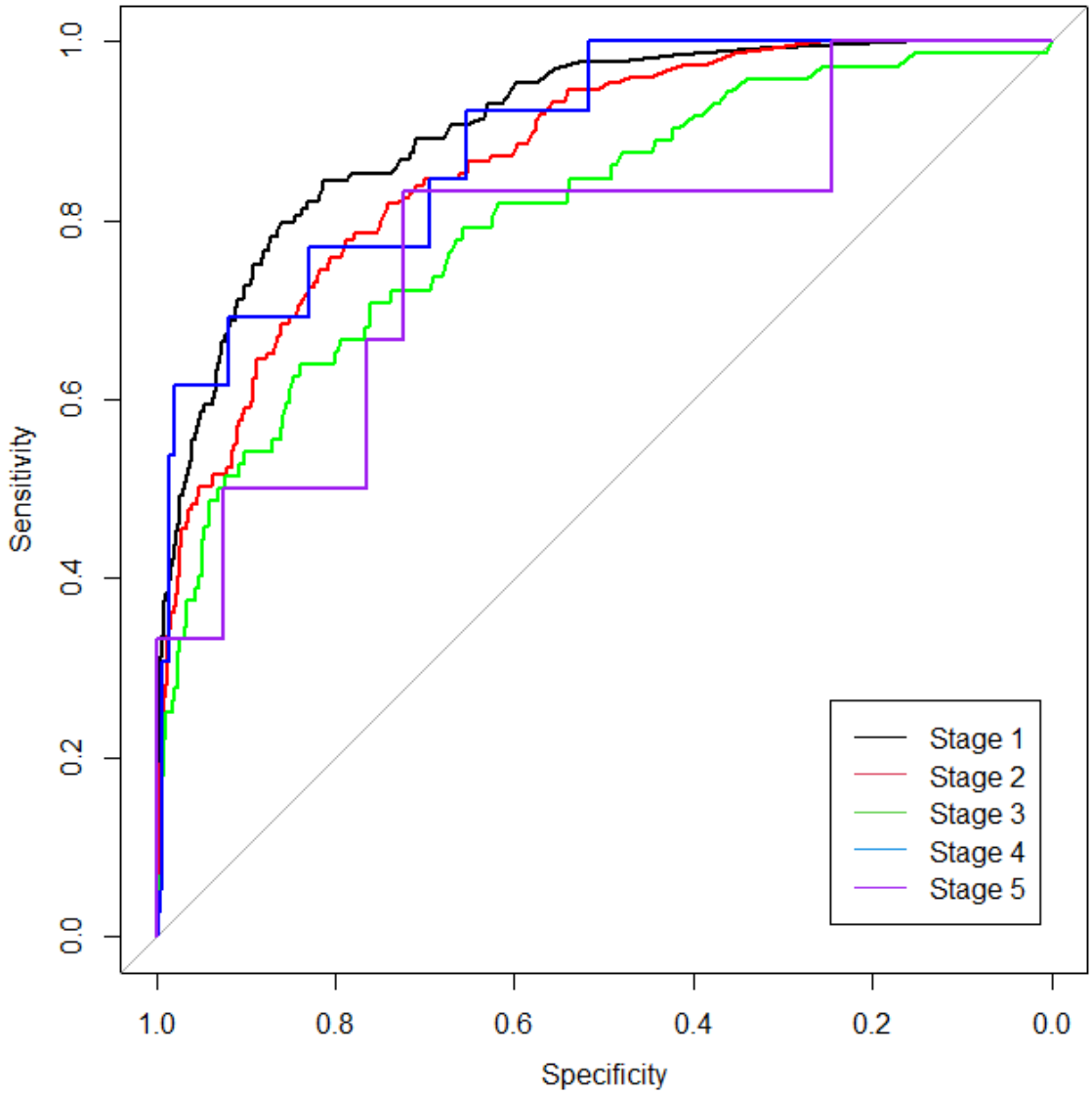



*Figure 5.12 Receiver operator curve for TMACS predicting the primary outcome separated by Chronic Kidney Disease Stage.*

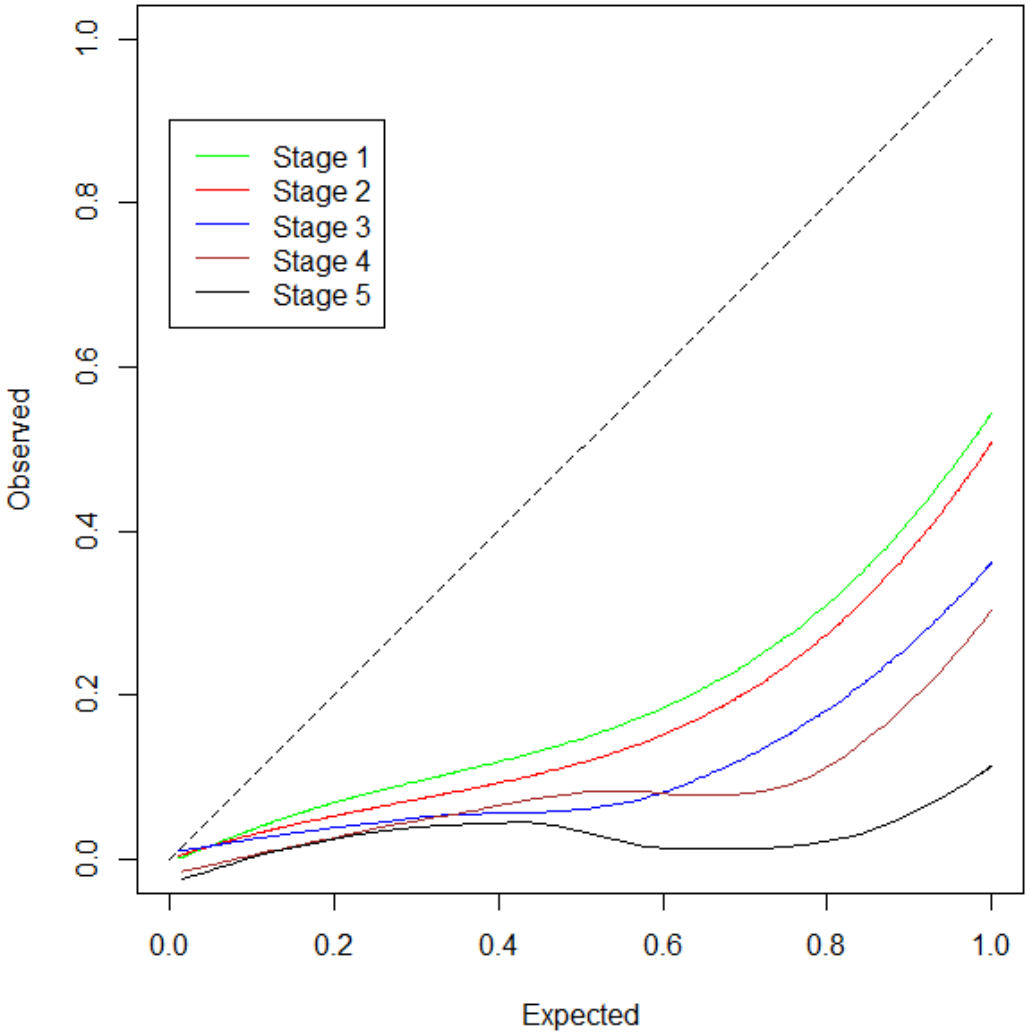



*Figure 5.13 Calibration plot for TMACS predicting the primary outcomes separated by Chronic Kidney Disease stage. A Loess curve was fitted to the values*

### *Deprivation*

The performance of TMACS was examined against quintiles of deprivation from the index of multiple deprivation (IMD) provided by NHSD. Discrimination trended to deteriorate with increasing depravation (Table 5.12). AUC in the first quintile was 0.89 (95% CI 0.86 to 0.92) and the highest fifth quintile it was 0.84 (95% CI 0.80 to 0.88) (Figure 5.14). CITL did not demonstrate a similar trend and demonstrated consistent overestimation of risk by TMACS, the comparative calibration across the different subgroups was relatively static across different predicted risks (Figure 5.15).

| IMD | AUC | CITL | C-slope | Brier |
|---|---|---|---|---|
| Q1 | 0.89 (0.86 to 0.92) | -3.44 (-3.85 to -3.04) | 0.02 (0.01 to 0.02) | 0.08 |
| Q2 | 0.85 (0.81 to 0.88) | -3.94 (-4.38 to -3.50) | 0.11 (0.08 to 0.14) | 0.09 |
| Q3 | 0.89 (0.86 to 0.92) | -2.97 (-3.36 to -2.58) | 0.01 (0.00 to 0.02) | 0.07 |
| Q4 | 0.87 (0.83 to 0.91) | -3.18 (-3.57 to -2.80) | 0.19 (0.14 to 0.23) | 0.08 |
| Q5 | 0.84 (0.80 to 0.88) | -3.30 (-3.70 to -2.91) | 0.14 (0.10 to 0.17) | 0.09 |

*Table 5.12 TMACS discrimination and calibration metrics by Index of Multiple Deprivation quintile. AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope.*

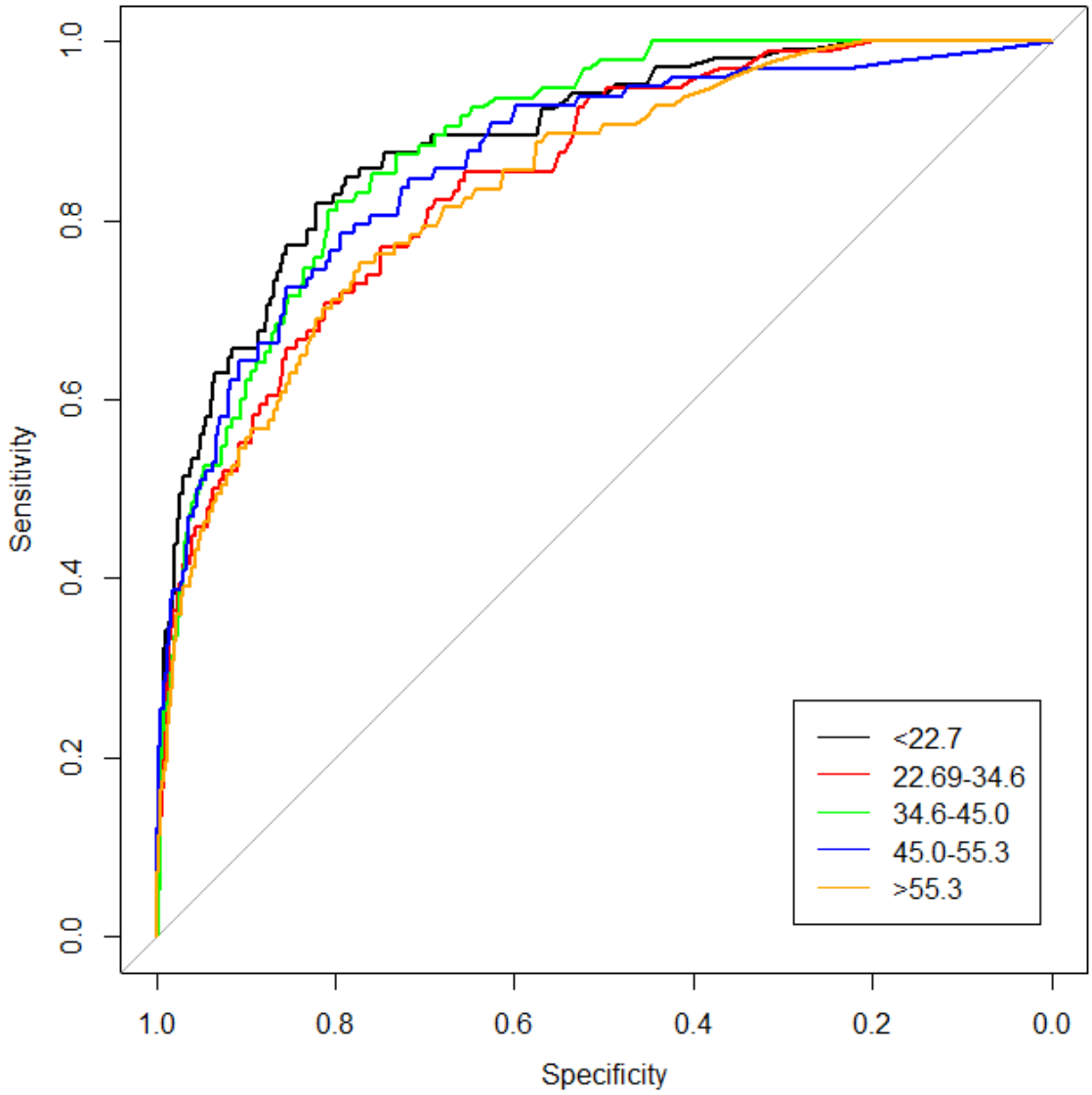



*Figure 5.14 Receiver operator curve for TMACS predicting the primary outcome separated by quintile of index of multiple deprivation*

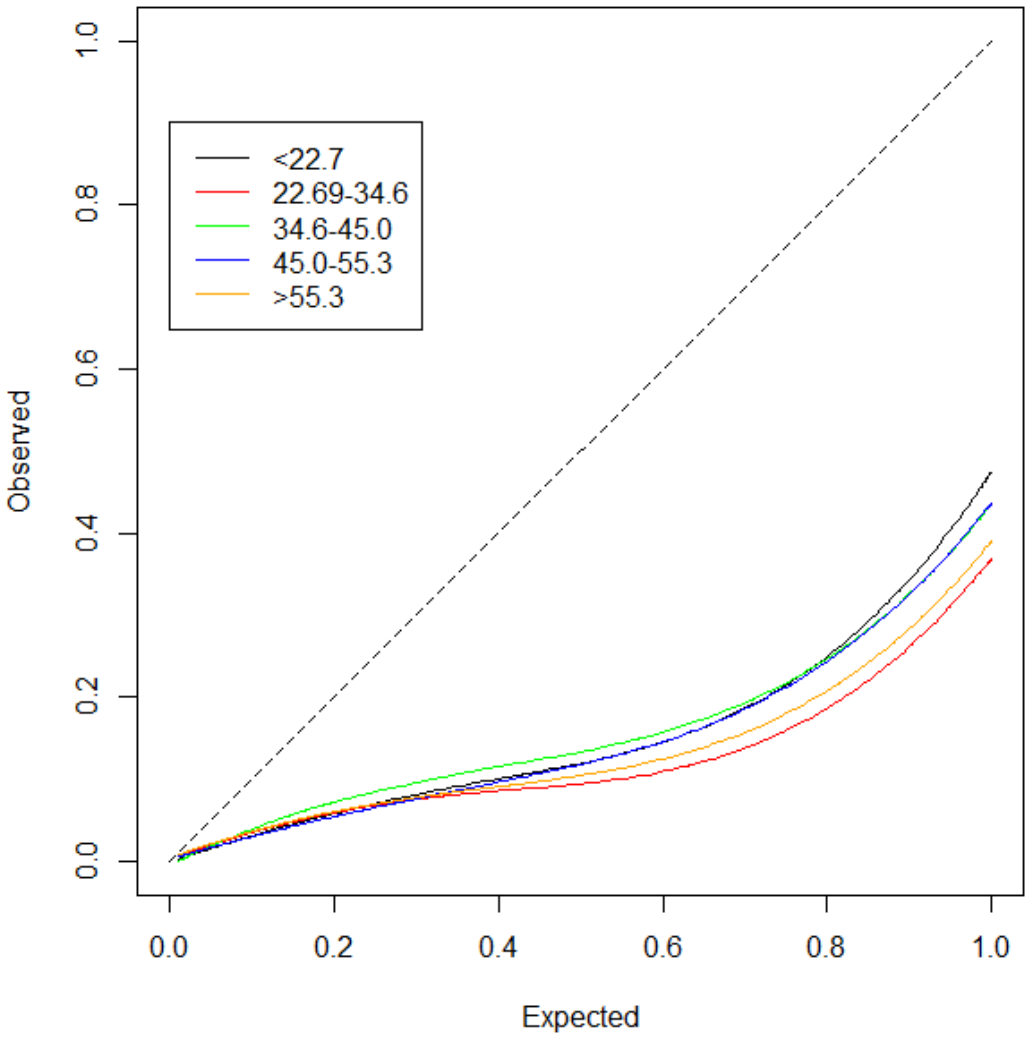



*Figure 5.15 Calibration plot for TMACS predicting the primary outcomes separated by quintile of index of multiple deprivation. A Loess curve was fitted to the values*

### 5.5.4 Longitudinal model performance

The incidence of the primary outcome varied substantially over time from 4% in Q2 2016 to 30% in Q4 2018. The secondary outcome of MACE followed the incidence trends of the primary outcome at a low of 5% in Q2 2016 and a high of 48% in Q4 2018 (Figure 5.16). The expanded outcome definition in the sensitivity analysis demonstrated a small increase in incidence which at its most was a 5% increase in incidence of the primary outcome (Q4 2018)

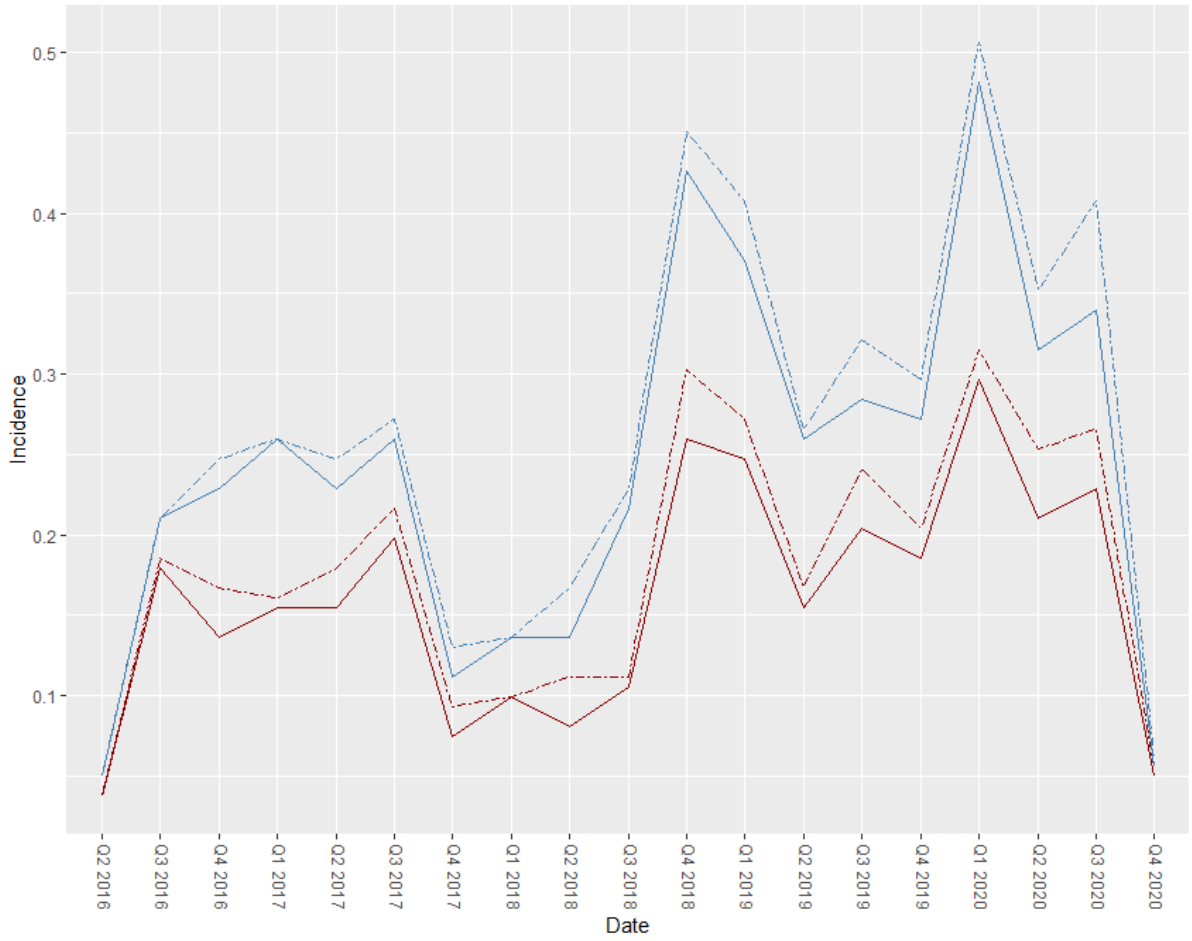



*Figure 5.16 Incidence of outcomes of interest over time. Red solid line - acute myocardial infarction in first code position, red dotted line – acute myocardial infarction in any position, blue solid line – major adverse cardiovascular event in first code position, blue dotted line -major adverse cardiovascular event in any position*

The performance of TMACS over time was examined, by yearly quarters. The calibration slope demonstrated a trend for decreasing over time, however there was significant variation. C-slope ranged from 0.51 in the first quarter to 0.10 in the last. CITL varied substantially over time, with a low of -2.10 (95% CI -2.68 to -1.53) in Q1 2017 to -7.05 (95% CI -8.00 to - 6.11) in Q1 2019 (Figure 5.18). A similar but lower magnitude variation was demonstrated in the AUC with a high of 0.94 (95% CI 0.91 to 0.97) in Q4 2016 and a low of 0.81 (95% CI 0.72 to 0.90) in Q2 2018 (Figure 5.19).

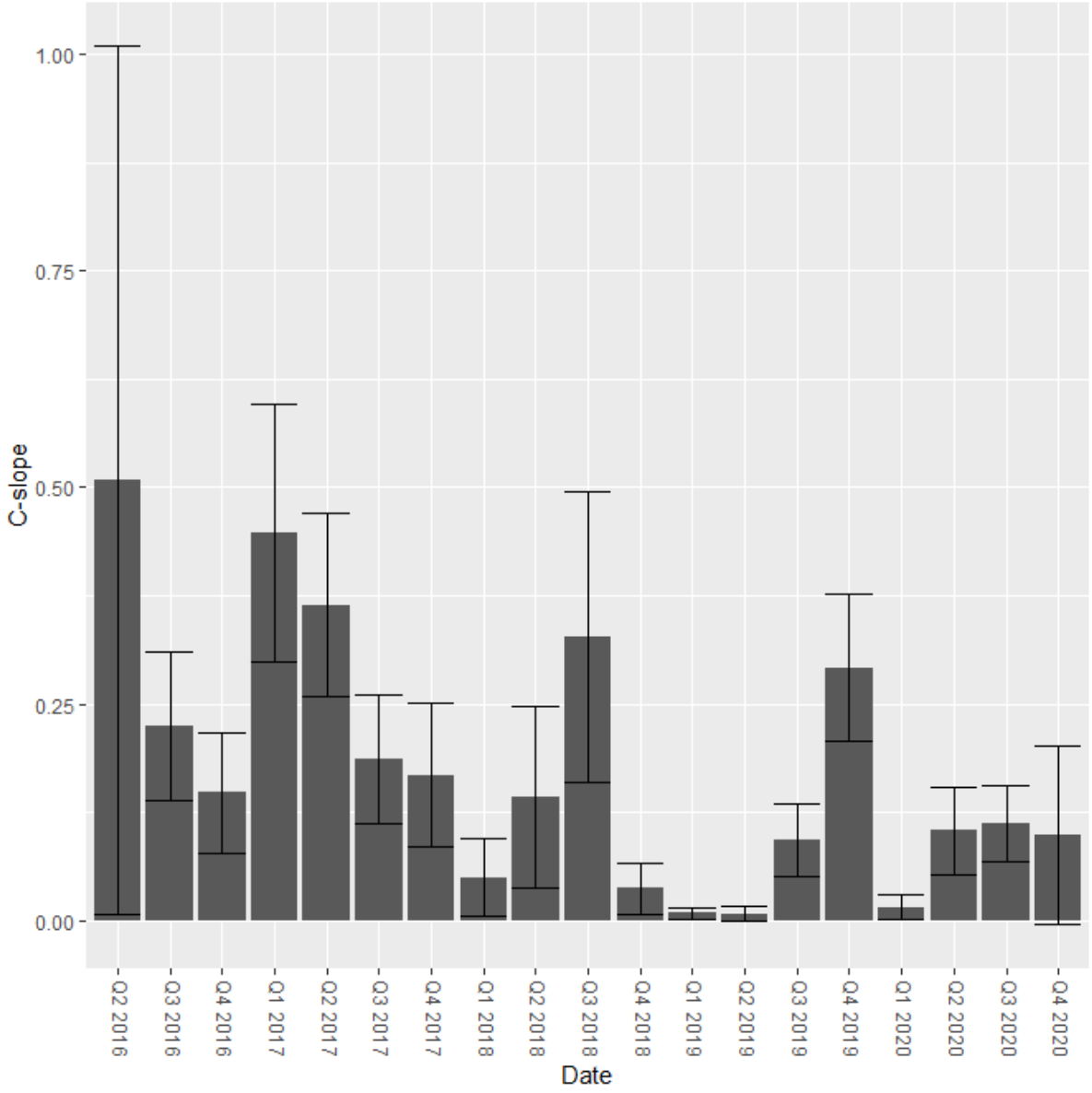



*Figure 5.17 Calibration slope of TMACS plotted against time - time measured in annual quarters. 95% confidence intervals are displayed.*

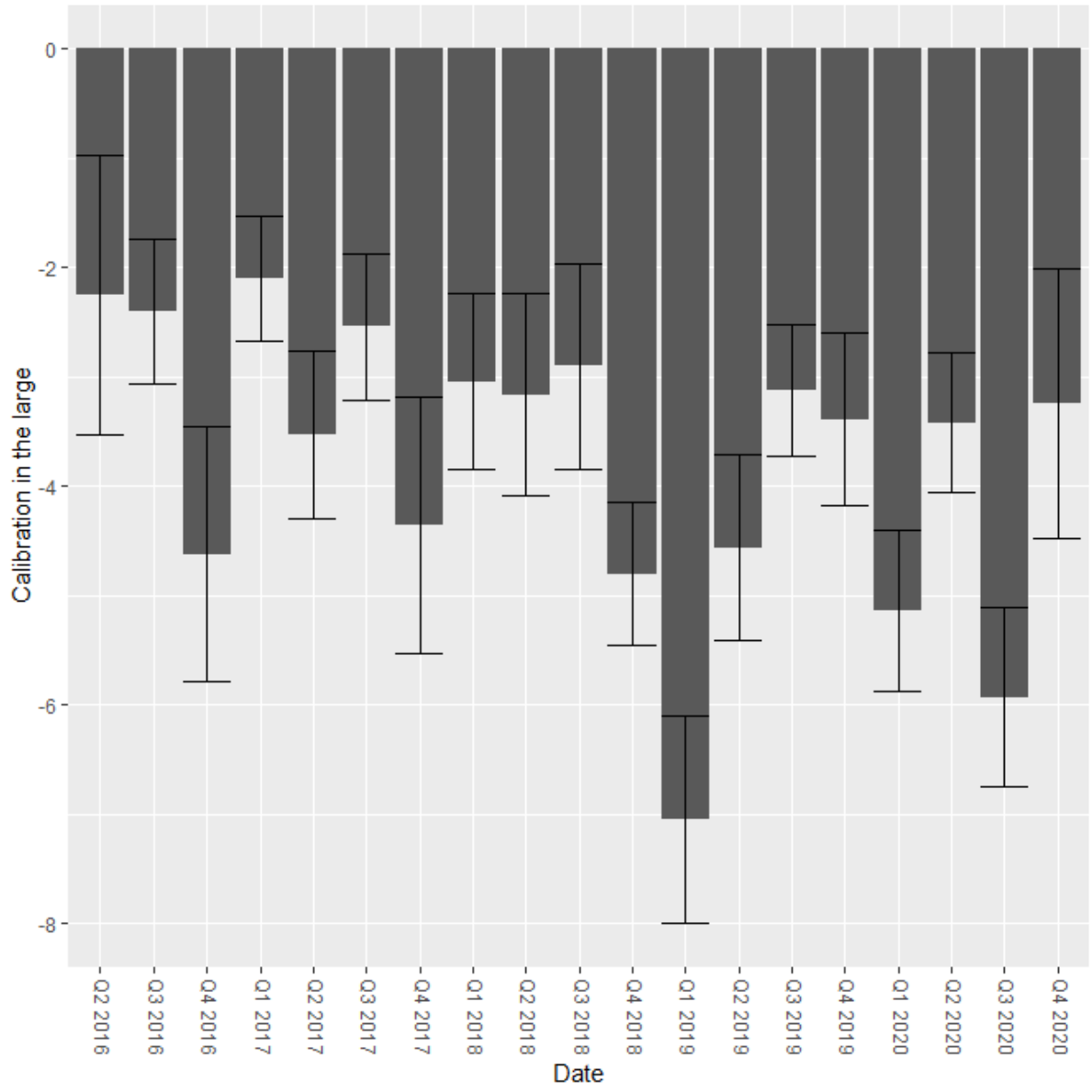



*Figure 5.18 Calibration in the large plotted against time - time measured in annual quarters. 95% confidence intervals are displayed.*

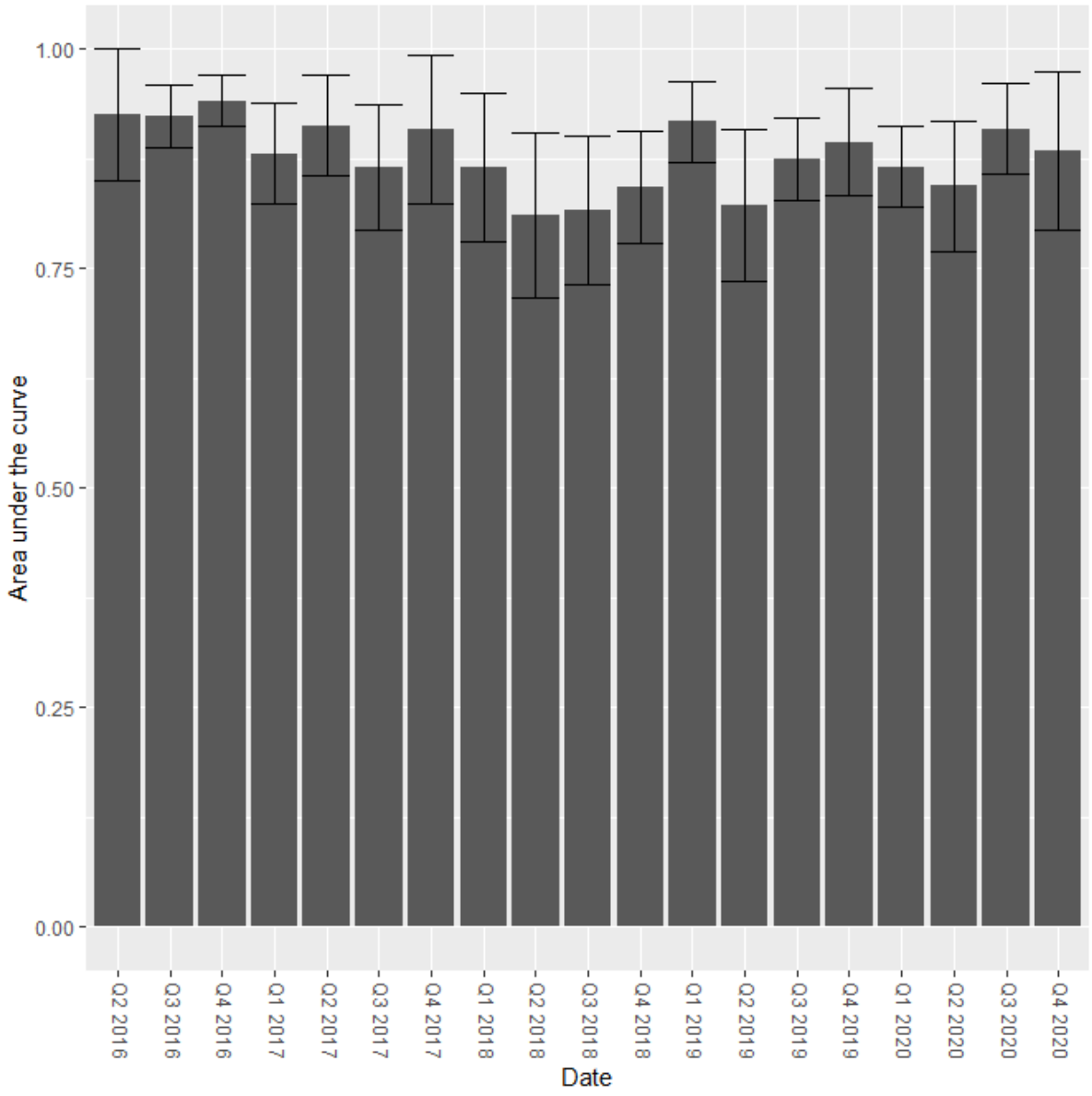



*Figure 5.19 Area under the curve plotted against time - time measured in annual quarters. 95% confidence intervals are displayed.*

## 5.6 Discussion

TMACS demonstrated good discrimination overall with an AUC of 0.88 (95% CI 0.86 – 0.89). This corresponded to a sensitivity for the very low risk group of 97.2% and a negative predictive value of 99.7%. However, if the correct hs-cTnT had been used and suspected coding errors were factored in the model performance is marginally improved with a sensitivity of 97.76% and an AUC of 0.88 (95% CI 0.86 to 0.90). This was a marginal drop in discriminatory performance compared to the original validation for the outcome of ACS which had a sensitivity of 98.1% (95% CI 95.2 to 99.5%) and an AUC of 0.91 (220). An external validation of TMACS predicting AMI demonstrated a sensitivity of 99.1% (95% CI 95.2% to 100%) and an AUC of 0.91 (95% CI 0.88 to 0.95), again highlighting how TMACS in this study appears to have deteriorated in discriminatory performance.

In this analysis I used the same co-efficient for the troponin variable across different troponin assays, as this has been consistent across various validation studies (221–223). Other CPMs have used cut-offs for each individual troponin assay, as specified by the manufacturers, such as the LoD or URL (30). An area for future research could be to further evaluate whether the accuracy of T-MACS would be optimised by re-calibrating the CPM using different commercially available troponin assays.

The three sites did not contribute equal numbers of participants to the database, with MRI contributing 81.2%, Blackburn 9.5%, Burnley 9.4%. This is despite MRI and Blackburn having similar attendance rates for the ED. This could be in part due to different compliance with the TMACS care pathway, different rates of chest pain attendances or differences in the data collection methods between sites.

Overall TMACS did not have favourable calibration results, with a very low CITL of -3.93 and low C-slope of 0.05 confirming the findings in the calibration plots that the model substantially overestimated risk. This could in part be related to the concentration of TMACS risk predictions in the very low risk group (**Error! Reference source not found.**) and the low overall incidence of 3.7%

compared to the 22.3% from the derivation trial (220). In the focused calibration plot TMACS appears visually to be well calibrated around the majority of the risk prediction deciles (Figure 5.2). This over-estimation of risk is not out of keeping with clinical practice where clinicians felt that they were only willing to miss 0.1 to 1% of AMI diagnoses (224). In this risk averse clinical setting a CPM that is very well calibrated in the lower risk groups but overestimates risk in the higher risk groups may be clinical advantageous. Future work could seek new methods to assess calibration when such caution is warranted. Calibration is an important CPM characteristic. However, how to describe or quantify it where over-estimation of risk is preferable is not yet clear. Loess calibration plots have been ranked as the most important test for calibration (when continuous predictors are present) (225,226). It may be that the visual inspection of a loess calibration plot is the appropriate test for such scenarios.

The incidence was lower than anticipated, this could be because of the change in use of the CPM. It could also be due to the reliance on clinical codes, it is possible that the clinical teams and coders are underdiagnosing AMI. Clinically adjudicating the outcomes could cast light on the cause of this low incidence and should be considered for future work.

The variation in performance by TMACS site is interesting. As TMACS was derived using data from MRI, it was probable that it was overfitted and could have struggled in predictions at different sites. However, the inverse appears to be the case regarding discrimination, with MRI demonstrating a lower AUC than RBTH and BGTH (0.87 vs 0.91/0.92). The calibration at different sites did follow the expected pattern with a better CITL for MRI than RBTH or BGTH.

TMACS demonstrated decreasing discriminatory performance with increasing age. Ethnicity highlighted further variation in TMACS with mixed ethnicity having the highest AUC (0.92) and Asian ethnicity the closest CITL to zero (-2.52). The Brier score however suggested that when accounting for discrimination and calibration, white ethnicity had better overall model performance (0.08).

When the cohort was examined by CKD stage, based off the initial urea and creatinine measurement, a consistent drop in performance across domains was demonstrated. This is likely related to two factors, firstly a portion of these patients will have established chronic kidney disease and as such likely have raised background levels of cardiac troponin circulating due in part to microvascular damage to the myocardium (227). The troponin variable in TMACS may be inappropriately calibrated to patients with normal renal function and not to this pathology. Secondly another cause for impaired renal function can be due to an acute physiological issue related to infection, it is possible that by examining by kidney function I selected a portion of patients who were more likely not to have a diagnosis of AMI. Thirdly If the kidney function test reflected chronic kidney disease then this would represent a cohort of patients who had a higher rate of cardiovascular disease as such this difference in background risk may have impacted the performance of the model (228).

These different subgroups represent different variables that could be incorporated into the model in future updating work. Furthermore, any further updating work should re-examine these subgroups to see if the performance across domains has improved.

The longitudinal performance of the model did demonstrate variability. The c-slope trended for a decrease overtime, but the CITL demonstrated no such visual trends in its variability. As such it appears that the model became mis-calibrated prior to the start of our cohort. Whilst this is indicative of calibration drift it could be more accurately described as a ‘jitter’, due to the substantial variation in performance without a clear trend. Incidence varied substantially through the year however it did not appear to correlate with the change performance of TMACS across discrimination and calibration domains. This is in contrast to the analysis conducted by Hinkley et al on a cardiovascular surgery CPM (229). They demonstrated a consistent drop in model calibration over time which was not evident from our model, however their outcome was death which is feasibly more objective and less open to misdiagnosis/coding than specific medical conditions such as AMI.

## 5.7 Conclusion

TMACS demonstrated good discriminatory performance and was well calibrated in the majority of its predicted risk deciles. There was a marginal drop in discriminatory model performance overall compared to the original internal and external validation. TMACS demonstrated substantial calibration drift overall. The cause of this could be due to outcome data quality issues or may represent an actual drop in performance. In various subgroups of interest, the model performance across discrimination and calibration domains varied, which was most prominent in kidney function subgroups. Calibration jitter was identified but no declining trend or drift in performance was evident. Updating TMACS may counter this jitter and improve model performance across subgroups of interest.

# Chapter 6 Updating and refining TMACS

## 6.1 Introduction

It is important to ensure that the performance of a CPM remains optimal once introduced into practice. There are several reasons for this, including (a) the need to verify performance in real-world cohorts and in different settings (ensuring external validity); (b) the need to re-calibrate based on changing technologies (e.g., new biomarker assays) or practices (e.g., ECG interpretation); (c) the need to ensure optimal performance in subgroups of interest (e.g., based on gender, ethnicity, comorbidities such as CKD); and (d) the need to ensure that performance doesn't deteriorate over time.

In the previous chapter I identified and examined how calibration drift was affecting TMACS. There are numerous methods available to update CPMs that have become mis-calibrated over time. In this chapter I focused on three methods to update TMACS, model recalibration, model extension and Bayesian dynamic updating.

The simplest is a model recalibration using a single static recalibration where the linear predictor is used as a prediction variable in a new model (66). Another possible method is model extension where additional variables are incorporated into the pre-existing model to create a new model (230). This would enable new variables that could represent new diagnostic technologies or newly identified predictors to be added to CPMs. These updating techniques do not allow for continual improvements, Bayesian dynamic updating enables the constant rederivation of a model after every prediction (66). By constantly updating the model it may prevent the model from deteriorating over time (calibration drift).

## 6.2 Aims

Identify the optimal machine learning methods for ongoing updating of TMACS

- model recalibration
- model extension
- Bayesian dynamic updating

## 6.3 Published work

Reynard, C., Martin, G.P., Kontopantelis, E., Jenkins, D.A., Heagerty, A., McMillan, B., Jafar, A., Garlapati, R. and Body, R., 2021. Advanced cardiovascular risk prediction in the emergency department: updating a clinical prediction model–a large database study protocol. *Diagnostic and Prognostic Research*, *5*(1), pp.1-7.

## 6.4 Methods

I used the cohort described in chapter 5 to examine the different methods for updating TMACS. In summary, this included patients presenting to the ED with chest pain where the clinician suspected ACS and used TMACS as a part of routine care. Data were collected across three clinical sites from June 2016 to October 2022. As part of routine clinical care all patients received two troponin tests. TMACS was digitally captured at each site and this data was supplemented with outcome data from NHS Digital. The primary outcome was a diagnosis of AMI within 30 days of the index attendance at the ED.

### 6.4.1 Data sources

TMACS data was collated from local registries that used routinely collected clinical data from NHS trust's where it was part of routine care. This included sites in Manchester (MRI), Burnley (RBTH) and Blackburn (BGTH). Subgroup data was extracted locally where possible and centrally from NHSD. The local databases were linked with NHSD to provide national coverage for outcome codes relating to AMI. Subgroup data was also provided by NHSD where possible.

### 6.4.2 Outcome

The primary outcome was a diagnosis of AMI within 30 days of the index date, including any subsequent admission. To achieve this primary discharge codes were examined for AMI ICD-10 codes including I21, I22 or I23 (see Supplementary Table 8.12).

### 6.4.3 Sample Size Calculation

I calculated the sample size required by using the more conservative of two methods, firstly the protocol described by Riley et al for model derivation (173) and secondly the rule of 10 primary outcome cases per variable that had become standard practice, with a plan to accept whichever calculation yielded a larger minimum sample size. I estimated the use of 17 variables which were dependant on data availability. It was anticipated that for the rule of ten we would need 170 primary cases. In unpublished service evaluations, it was found that the prevalence of AMI in the first 1,033 patients was 6.9%. Therefore. in order to achieve the rule of ten sample size we would require 2,464 patients. I also used Riley et al's sample size calculation using the original TMACS validation AUC of 0.90, the prevalence of AMI at 6.9% and the derivation of a model with 17 variables (173,231). This estimated that 727 was the minimum sample size required. The rule of ten sample size calculation exceeded Riley et al's method so I opted to be conservative and adopt it.

### 6.4.4 Recalibration

Initially I planned to update TMACS by recalibration, re-estimating both the intercept and a new overall calibration slope.

$$_{mr} = \hat{} + \widehat{_{o}} \; _{sq}$$

$$_{mr} = \hat{} + \hat{}_{0} \; _{TMACS} + \sum_{1,\ldots,7} \hat{}_{0} * (\; _{i,TMACS} x_i)$$

Equation 1: $_{mr}$ - model updated by recalibration, $\hat{}$ is the re-estimated intercept, $\widehat{_{o}}$ the new overall calibration slope, and $_{sq}$ – is the linear prediction of the TMACS model.

### 6.4.5 Extension

I sought to update the model by extending it and adding new variables to the original TMACS algorithm (see equation 2). These new variables were sourced from the available data and included the subgroups of interest: age, ethnicity, kidney function, IMD, and gender. The distribution of each variable was assessed visually and with a skew metric. If a skew was identified the variable was transformed. Secondly a linear relationship with the outcome was assessed by outcome variable plots, if non-linearity was suspected then splines were considered. The resulting transformed variables were then selected for the extended model using forward step wise selection. Co-linearity between the new candidate variables was assessed in co-linear plots. In a sensitivity analysis where co-linearity was suspected colinear terms were added and included in a further stepwise selected model.

$$_{e} = \hat{} + \sum_{1,\ldots,7} \; _{i} x_i + \sum_{j \epsilon s} \hat{}_{j} \, x_j$$

Equation 2: $_{e}$ - model updated by extension, $_{i}$ are the original coefficients for the original covariables, s is the new covariates and $\hat{}_{i}$their new coefficients.

### 6.4.6 Bayesian dynamic updating

I planned to use dynamic updating to update the T-MACs' intercepts and coefficients after each patient episode. During the course of this project we engaged with patient representatives who desired a trial period where the model would be monitored in real-time and after which it could progress to run autonomously. A 3-month probationary period was highlighted in the model to demonstrate a period when there would be human oversight. After this period we will update the model after every patient episode. The dynamic updating was conducted with recursive estimation via a prediction equation:

$$_t|\ ^{t-1} \sim \ (\ \hat{}_{t-1},\ _t)$$

is a dimensional vector of regression coefficients, $^{t-1}$is a set of past outcomes, t is a given time and $_t = \hat{\Sigma}_{t-1}/_t$ (71). $_t$ is the forgetting factor, that serves to down-weight previous observations by increasing the variance. It will be selected in order to enable the sample size to continue to meet the specifications laid out by Riley et al (232) and to ensure appropriate stability of the intercept and the coefficients for the variables.

This is passed through a Bayesian framework, the posterior is proportional to the product of the probability distribution at time t and t-1, giving.

$$(\ _t|\ ^t) \propto \ (\ _t|\ _t)\ (\ _t|\ ^{t-1}) \propto l\ \ l\ \ hoo\ \ x\ \ o$$

The pooling of data from different sites was to be assessed by dynamically updating each site individually and observing the variation and trends in the intercept and coefficients for the variables (betas). If a similar beta fingerprint could be identified between sites then pooling would be considered.

### 6.4.7 Model Characteristics

The new models were assessed with measures of discrimination and calibration, including AUC, CITL, C-slope and calibration plots. For the recalibrated and extended models' internal validation was

conducted by boot strap validation across 1,000 iterations. Sensitivity analyses were conducted to examine the effect of not pooling the data from different sites.

The dynamically updated model was validated using one-step ahead prediction. As the model was dynamic, I also sought to describe the changes in the model over time by reporting these summary statistics per annual quarter as well.

As before, subgroups of interest were examined including age, IMD, gender, kidney function and ethnicity. Reclassification was visualised by Sankey plots and new model intercepts/coefficients were reported.

The updating methods recalibration and dynamic updating only required the original TMACS variables and the outcome variable. There was no missing data for these data points. For model extension new variables were added whilst updating the model, multiple imputation was used to handle missing data.

## 6.5 Results

A total of 13,207 patient episodes were included in the database from MRI (10,710), BGTH (1,242) and RBTH (1,255) from June 2016 to October 2020. The incidence of the primary outcome was 3.7% overall, with higher rates found at RBTH (4.3%) and BGTH (4.6%) and slightly lower rates at MRI (3.6%.) Missing data for potential predictors in the model extension method is reported in Table 6.1.

| **Feature** | **N (%)** |
|---|---|
| | |
| **Total number of patients** | 13207 |
| Manchester Royal Infirmary | 10710 (81.2%) |
| Burnley | 1242 (9.4%) |
| Blackburn | 1255 (9.5%) |
| | |
| **Gender** | |
| Male | 6175 (46.8%) |
| Female | 4535 (34.4%) |
| Missing | 2497 (18.9%) |
| | |
| **Age** | |
| Mean | 56.38 |
| Standard Deviation | 16.91 |
| Range | 16-103 |
| Missing | 0 |
| | |
| **Ethnicity** | |
| Asian | 2435 (18.4%) |
| Black | 756 (5.7%) |
| Mixed | 472 (3.6%) |
| Other | 347 (2.6%) |
| White | 4213 (31.9%) |
| Missing | 4984 (37.7%) |
| | |
| **Estimated glomerular filtration rate** | |
| Mean | 84.00 |
| Range | 2.51-159.30 |
| Standard Deviation | 24.27 |
| Missing | 2793 |
| | |
| **Index of Multiple Deprivation** | |
| Mean | 39.27 |
| Range | 1.03-81.76 |
| Standard deviation | 17.86 |
| Missing | 3408 |

*Table 6.1 Cohort Demographics*

### 6.5.1 Recalibration

The recalibrated model had a new overall intercept of -3.22 and a new overall co-efficient of 0.05, which gave an AUC of 0.88 (95% CI 0.86 to 0.89). When adjusted for optimism by bootstrapping through 1,000 iterations the AUC was found to be 0.88 and the CITL 0.00 (Table 6.2) . The recalibration reclassified all patients who were originally very low risk, low risk and moderate risk to low risk (Figure 6.1). This reclassification made the calculation of the sensitivity for the very low risk group impossible, as such for the sensitivity statistic calculation a new predicted risk threshold was used. The threshold was selected that had properties similar to the original sensitivity, this yielded a threshold of <3.2% which had a sensitivity of 97.2% (95% CI 95.7 to 98.6) and a specificity of 100 % (95%CI 99.9 to 100.0). This risk threshold demonstrated a proportion ruled out of 42.3%.

I conducted a sensitivity analysis to see if recalibrating each site individually resulted in better model performance statistics. The AUC per site was unchanged between pooled and individually recalibrated site subgroups. As such the pooled analysis was adopted.

The calibration of the new model was also examined with a calibration plot (Figure 6.2), this demonstrated a different pattern for the predicted risk deciles compared to the original TMACS. The majority of the deciles sit under the dashed line (y=x) indicating risk is underestimated for the lower deciles. As expected in a recalibrated model the parametrically fitted line of best fit for the calibration plot demonstrated near perfect calibration (y=x). However, when examined with a non-parametrically fitted line, an issue is exemplified (Figure 6.3), significant overestimation is observed. It demonstrated that below 0.60 predicted risk the observed risk was higher i.e. underestimation of risk.

| | Recalibration | | |
|---|---|---|---|
| | **Derivation (95% CI)** | **Bootstrapping** | **Status Quo** |
| **AUC** | 0.88 (0.86 to 0.89) | 0.88 | 0.88 (0.86 to 0.89) |
| **CITL** | 0.00 (-0.09 to 0.09) | 0.00 | -3.93 (-4.12 to -3.74) |
| **C-slope** | 1.00 (0.82 to 1.12) | 0.03 | 0.05 (0.17 to 0.22) |

*Table 6.2 Recalibrated TMACS summary statistics. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*


*Figure 6.1 Sankey Diagram for reclassification of risk from the original predicted risk groups on the left to the newly predicted risk groups on the right from the recalibrated TMACS model.*

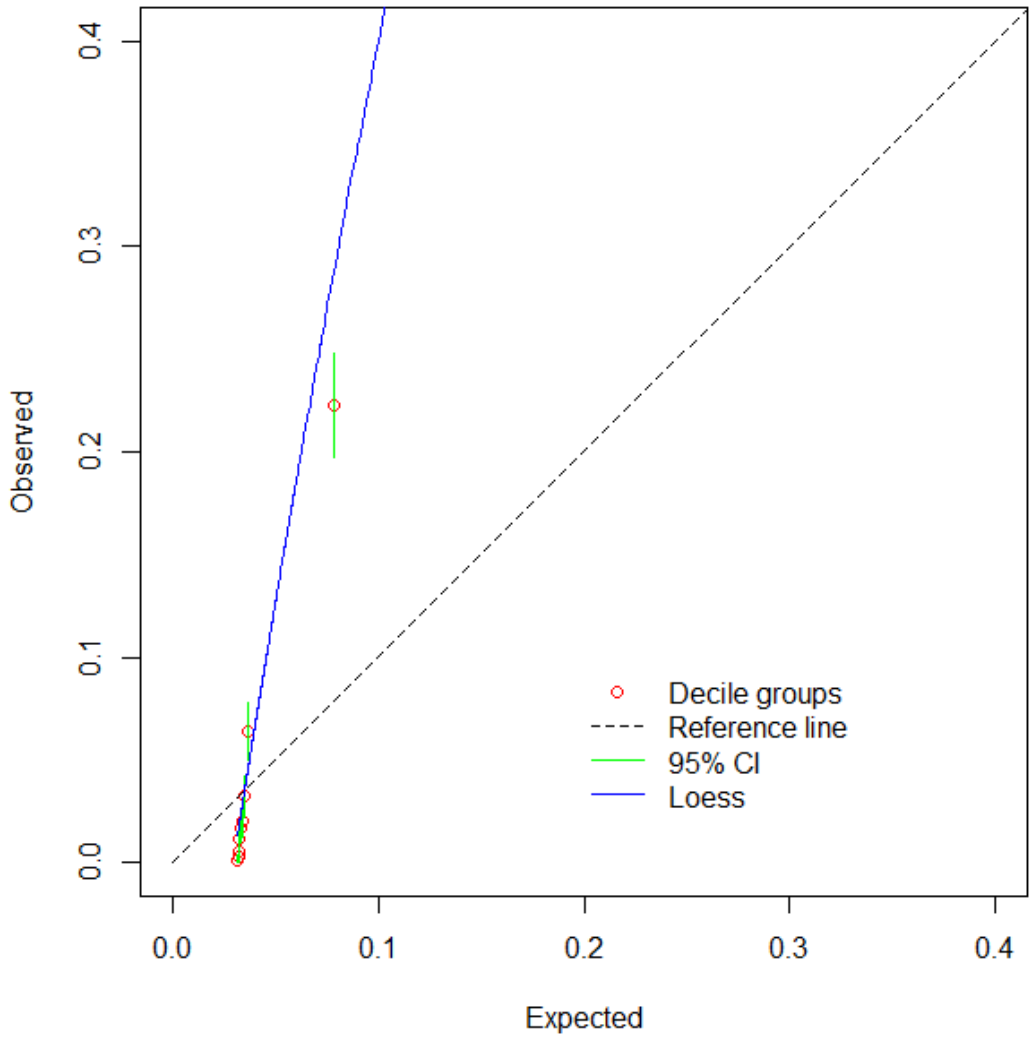



*Figure 6.2 - Focused calibration plot for the recalibrated TMACS CPM. Predicted risk deciles are plotted on an axis of observed vs expected risk. A non-parametric fitted line is applied using LOESS regression.*

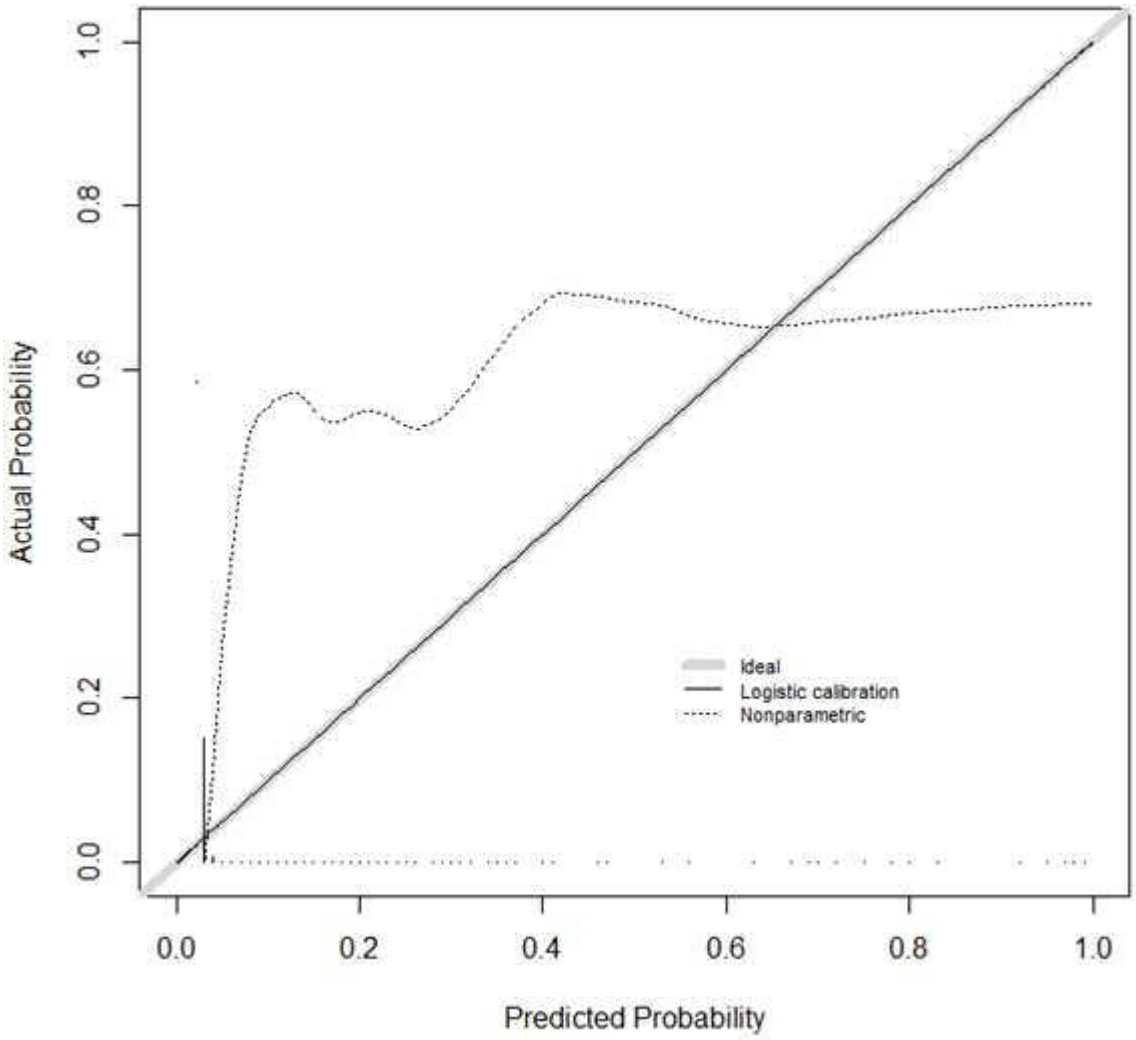



*Figure 6.3 Calibration plot for the recalibrated TMACS model demonstrating the comparison of a logistic vs non-parametric line of best fit.*

#### *6.5.1.1 Subgroup results*

### Site

When the recalibrated model performance was examined by site subgroups the AUC for MRI was below the pooled AUC result of 0.87 (95% CI 0.85 to 0.89) at 0.88 (95% CI 0.86 to 0.89). Blackburn and Burnley demonstrated more favourable numeric discriminatory characteristics with AUCs of 0.91 and 0.92, respectively (Table 6.3). In contrast the CITL for MRI was closer to 0 than that for Blackburn and Burnley, which was an improvement compared to the original statistic.

| | **Extension** | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo** |
| **AUC** | | | |
| MRI | 0.87 (0.85 to 0.89) | 0.87 | 0.87 (0.85 to 0.89) |
| Blackburn | 0.91 (0.86 to 0.95) | 0.91 | 0.91 (0.86 to 0.95) |
| Burnley | 0.92 (0.88 to 0.96) | 0.92 | 0.92 (0.88 to 0.96) |
| **CITL** | | | |
| MRI | 0.02 (-0.09 to 0.12) | 0.02 | -3.00 (-3.18 to -2.82) |
| Blackburn | -0.05(-0.37 to 0.24) | -0.04 | -8.34 (-9.29 to -7.38) |
| Burnley | -0.10 (-0.42 to 0.12) | -0.09 | -10.63 (-11.63 to -9.63) |

*Table 6.3 Recalibrated TMACS summary statistics by site. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*

### Gender

The discrimination by gender was numerically higher for females with an adjusted AUC of 0.89, however the CITL for females was -0.23 versus 0.17 for males (Table 6.4). Indicating slight over and under estimation of predicted risk respectively. However, in contrast to the original TMACS CPM the CITL was markedly improved.

| | **Recalibration** | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo** |
| **AUC** | | | |
| Male | 0.86 (0.83 to 0.88) | 0.86 | 0.87 (0.85 to 0.89) |
| Female | 0.89 (0.87 to 0.92) | 0.89 | 0.89 (0.87 to 0.92) |
| | | | |
| **CITL** | | | |
| Male | 0.17 (0.04 to 0.29) | 0.17 | -3.58 (-3.81 to -3.35) |
| Female | -0.23 (-0.42 to -0.06) | -0.23 | -4.25 (-4.58 to -3.92) |

*Table 6.4 Recalibrated TMACS summary statistics by gender. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*

### Age

The discriminatory performance numerically decreased with increasing age quartile according to AUC adjusted for optimism starting at 0.92 for those <44 years of age and ending at 0.85 for those >69 years of age (Table 6.5). The recalibrated model demonstrated an improved CITL compared to the original TMACS (status quo), but in contrast to discriminatory performance the calibration was worst in the younger age quartile.

| | Recalibration | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo** |
| **AUC** | | | |
| Age Q1 (<44) | 0.92 (0.88 to 0.95) | 0.92 | 0.92 (0.88 to 0.95) |
| Age Q2 (44-55) | 0.89 (0.86 to 0.92) | 0.89 | 0.89 (0.86 to 0.91) |
| Age Q3 (56-69) | 0.86 (0.83 to 0.89) | 0.86 | 0.86 (0.83 to 0.89) |
| Age Q4 (>69) | 0.85 (0.82 to 0.88) | 0.85 | 0.85 (0.82 to 0.88) |
| **CITL** | | | |
| Age Q1 (<44) | -1.06 (-1.40 to -0.75) | -1.05 | -4.71 (-5.42 to -3.99) |
| Age Q2 (44-55) | -0.02 ( -0.21 to 0.17) | - 0.01 | -3.35 (-3.73 to -2.96) |
| Age Q3 (56-69) | 0.23 (0.06 to 0.39) | 0.23 | -3.21 (-3.51 to -2.92) |
| Age Q4 (>69) | 0.28 (0.12 to 0.44) | 0.29 | -4.83 (-5.19 to -4.48) |

*Table 6.5 - Recalibrated TMACS summary statistics by age quartile. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*

### Ethnicity

Through the different ethnicities, variation was observed in discrimination and calibration performance of the recalibrated model. The highest AUC adjusted for optimism was for Mixed ethnicity at 0.92 compared to the lowest of 0.87 for Asian and White ethnicities (Table 6.6). In contrast the CITL did no demonstrate the same pattern with Black ethnicity having the largest magnitude CITL of -0.43 representing systematic over-estimated risk and Asian ethnicity had an AUC of 0.35 indicating an underestimation of risk. This was however a marked improvement to the calibration by ethnicity demonstrated in the original status quo TMACS CPM.

| | Recalibrated | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo** |
| **AUC** | | | |
| Asian | 0.87 (0.84 to 0.90) | 0.88 | 0.87 (0.84 to 0.90) |
| Black | 0.90 (0.83 to 0.97) | 0.90 | 0.90 (0.83 to 0.97) |
| Mixed | 0.92 (0.87 to 0.97) | 0.92 | 0.92 (0.87 to 0.97) |
| Other | 0.92 (0.86 to 0.97) | 0.91 | 0.92 (0.86 to 0.97) |
| White | 0.87 (0.85 to 0.89) | 0.86 | 0.87 (0.85 to 0.89) |
| **CITL** | | | |
| Asian | 0.35 (0.17 to 0.51) | 0.35 | -2.52 (-2.80 to -2.24) |
| Black | -0.44 (-0.86 to -0.06) | -0.43 | -5.98 (-6.86 to -5.10) |
| Mixed | -0.09 (-0.54 to 0.30) | -0.08 | -3.71 (-4.60 to -2.81) |
| Other | -0.29 (-0.92 to 0.24) | -0.29 | -3.86 (-4.91 to -2.82) |
| White | 0.12 (-0.01 to 0.24) | 0.12 | -4.03 (-4.28 to -3.77) |

*Table 6.6 Recalibrated TMACS summary statistics by ethnicity. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*

### Kidney Function

The discriminatory performance for kidney function strati was only minimally changed still demonstrating a decreasing trend for AUC from stage 1 to stage 5 with an optimism adjusted AUC of 0.90 – 0.79 (Table 6.7). Calibration was however markedly improved across each group with no clear trend apparent in contrast to the decreasing trend in the status quo model.

| | Recalibration | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo** |
| **AUC** | | | |
| Stage 1 | 0.91 (0.88 to 0.93) | 0.90 | 0.91 (0.88-0.93) |
| Stage 2 | 0.87 (0.84 to 0.90) | 0.87 | 0.87 (0.84-0.90) |
| Stage 3 | 0.81 (0.75 to 0.86) | 0.80 | 0.81 (0.75-0.86) |
| Stage 4 | 0.89 (0.79 to 0.98) | 0.89 | 0.89 (0.80-0.98) |
| Stage 5 | 0.78 (0.54 to 1.00) | 0.79 | 0.78 (0.54-1.00) |
| **CITL** | | | |
| Stage 1 | -0.22 (-0.40 to -0.05) | -0.21 | -2.08 (-2.36 to -1.80) |
| Stage 2 | 0.11 ( -0.06 to 0.27) | 0.12 | -2.32 (-2.57 to -2.08) |
| Stage 3 | 0.36 (0.12 to 0.60) | 0.36 | -3.63 (-4.05 to -3.20) |
| Stage 4 | 0.66 (0.03 to 1.19) | 0.68 | -7.08 (-8.59 to -5.56) |
| Stage 5 | -0.10 (-1.19 to 0.74) | -0.01 | -14.96 (-17.28 to -12.65) |

*Table 6.7 Recalibrated TMACS summary statistics by kidney function stage. AUC -area under the curve, CITL – calibration in the large, VLR – very low risk group, 95% CI – 95% confidence interval.*

### Index of multiple deprivation

The discrimination and calibration performance across IMD quintiles in the recalibrated model did not demonstrate any clear trends (Table 6.8). The AUC ranged from 0.88 to 0.84 which was a slight decrease from the status quo range of 0.89 – 0.84. The CITL was markedly improved across all IMQ quartiles ranging for 0.323 – 0.367 in comparison to the status quo results.

| | **Recalibration (95% CI)** | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapping** | **Status Quo (95% CI)** |
| **AUC** | | | |
| IMD Q1 | 0.88 (0.85 to 0.91) | 0.88 | 0.89 (0.86 to 0.92) |
| IMD Q2 | 0.86 (0.82 to 0.89) | 0.86 | 0.85 (0.81 to 0.88) |
| IMD Q3 | 0.89 (0.86 to 0.92) | 0.88 | 0.89 (0.86 to 0.92) |
| IMD Q4 | 0.87 (0.53 to 0.91) | 0.87 | 0.87 (0.83 to 0.91) |
| IMD Q5 | 0.84 (0.80 to 0.88) | 0.84 | 0.84 (0.80 to 0.88) |
| **CITL** | | | |
| IMD Q1 | 0.32 (0.11 to 0.51) | 0.32 | -3.44 (-3.85 to -3.04) |
| IMD Q2 | 0.30 (0.09 to 0.50) | 0.31 | -3.94 (-4.38 to -3.50) |
| IMD Q3 | 0.19 (-0.04 to 0.41) | 0.19 | -2.97 (-3.36 to -2.58) |
| IMD Q4 | 0.30 (0.08 to 0.51) | 0.31 | -3.18 (-3.57 to -2.80) |
| IMD Q5 | 0.36 (0.14 to 0.56) | 0.37 | -3.30 (-3.70 to -2.91) |

*Table 6.8 Recalibrated TMACS summary statistics by index of multiple deprivation quintile. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

### 6.5.2 Model Extension

The variables age, ethnicity, kidney function, IMD, and gender were considered for inclusion in an extended model. Initially the distribution of the variables were assessed visually and with the skewness metric (181). Kidney function (eGFR) was found to have a left skew which was most improved with a transformation by squaring. The variables relationship with the outcome was assessed for linearity. The linearity plot for age did not demonstrate a linear relationship, so the variable was transformed with a restricted cubic spline (Figure 6.4).

| **Variable** | **Skewness** | **Linear outcome relationship** | **Transformation** |
|---|---|---|---|
| Age | 0.01 | No | Spline |
| Kidney function | -0.62 | Yes | Squared |
| IMD | 0.42 | Yes | Nil |

*Table 6.9 Summary of variable transformations for inclusion in extended TMACS model.*

Forwards stepwise selection for additional variables resulted in all but IMD being added to the final model (Figure 6.5 and Supplementary Table 8.15). The resulting model had an AUC of 0.84 (95% CI 0.82 -0.86) which after adjustment for optimism was unchanged at 0.84 (Figure 6.6). The sensitivity for the very low risk group (by original <2% threshold) was 0.63 (95% CI 0.59 – 0.68) and specificity for the high risk group was 0.99 (95% CI 0.99 – 00.99). Calibration was assessed again with CITL which was 0.00 (95% CU -0.15 – 0.15) and after adjustment for optimism by bootstrapping the CITL was -0.01. The C-slope was 0.07 (95% CI 0.06 – 0.08) and after adjustment for optimism it was 0.39. This is indicative of miscalibration. This significant miscalibration is likely due to the linear predictor from the original TMACS being included. However. when the calibration was also assessed by calibration plots this demonstrated good calibration at the lowest predicted risk deciles, with slight over-estimation (Figure 6.7).

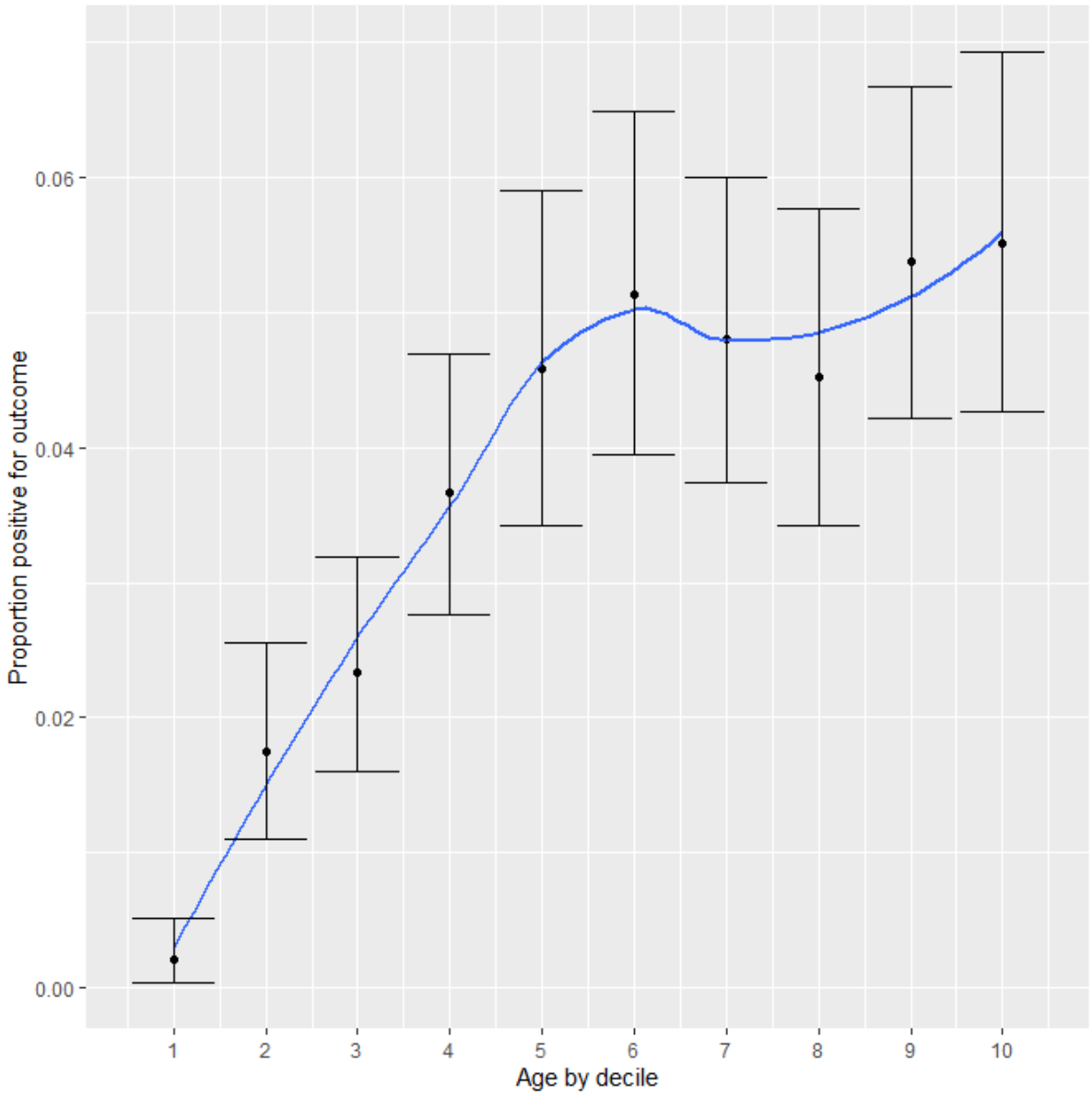



*Figure 6.4 Outcome variable linearity plot with the variable age plotted by decile against the proportion positive for the primary outcome.*

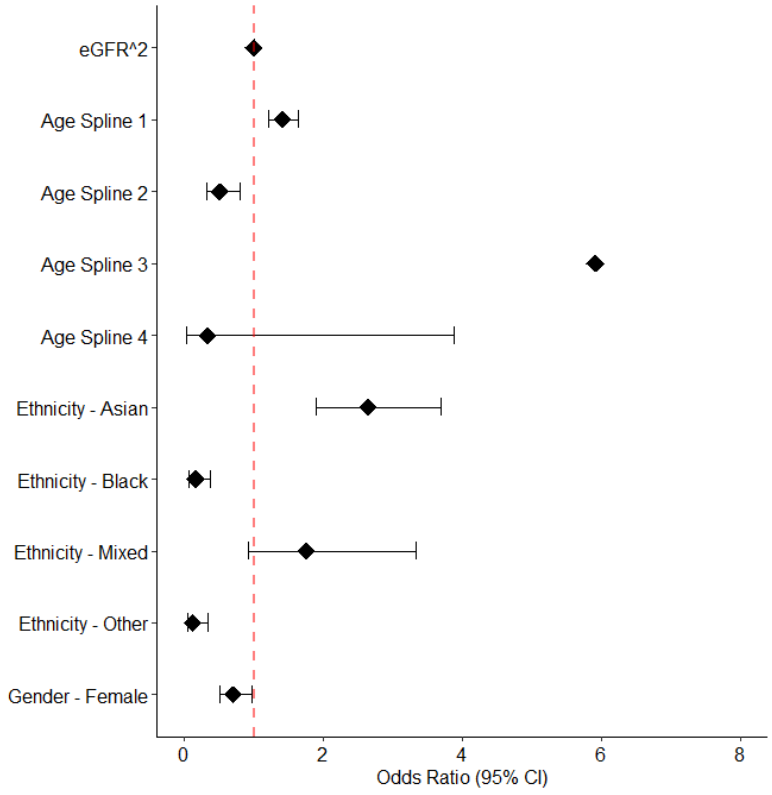



*Figure 6.5 Forest plot of variables included in the new extended TMACS clinical prediction model. TMACS's original coefficients remained unchanged. eGFR – estimated glomerular filtration rate (kidney function).*

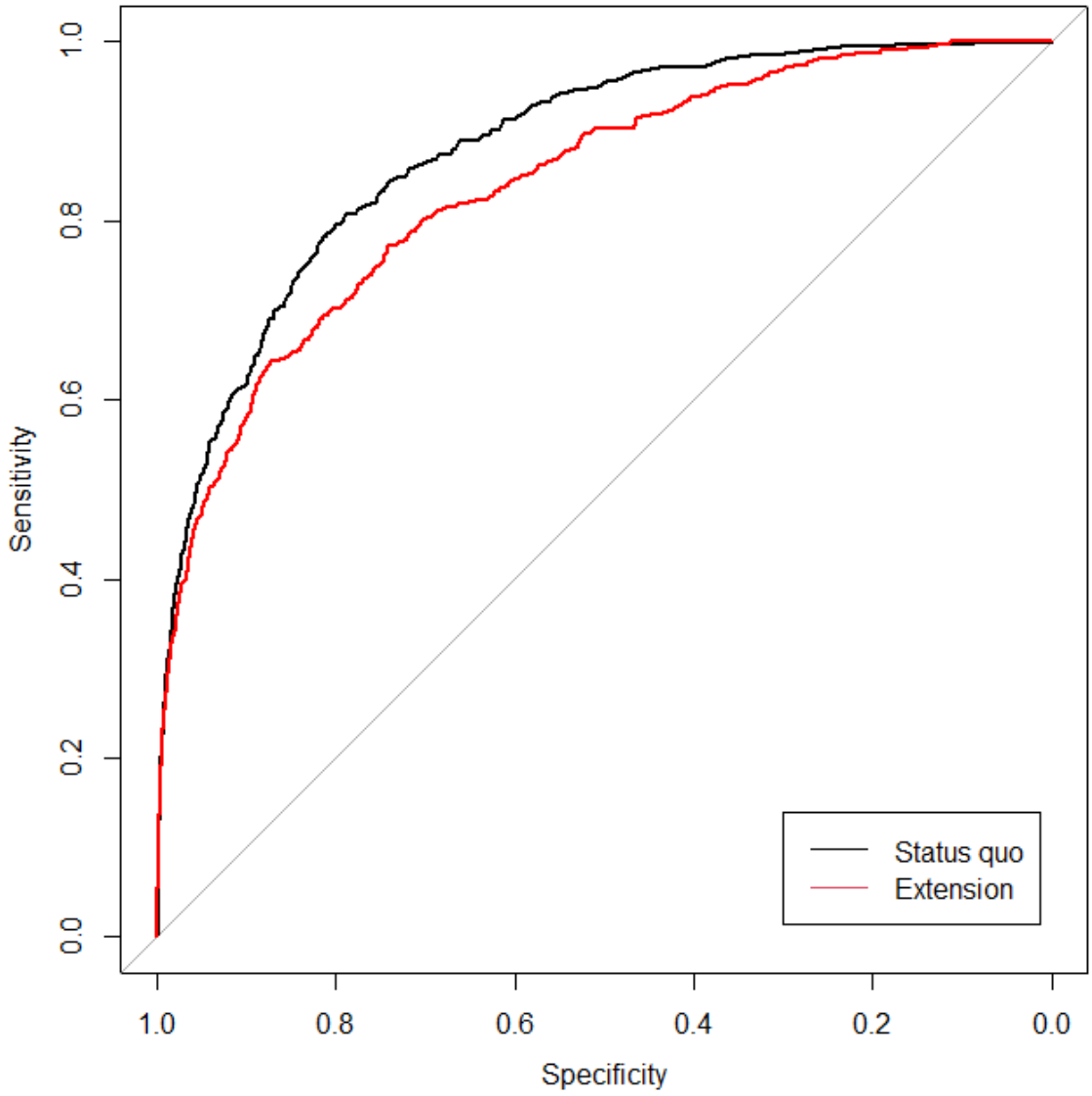



*Figure 6.6 - Receiver operator curve for extended and status quo TMACS clinical prediction models*

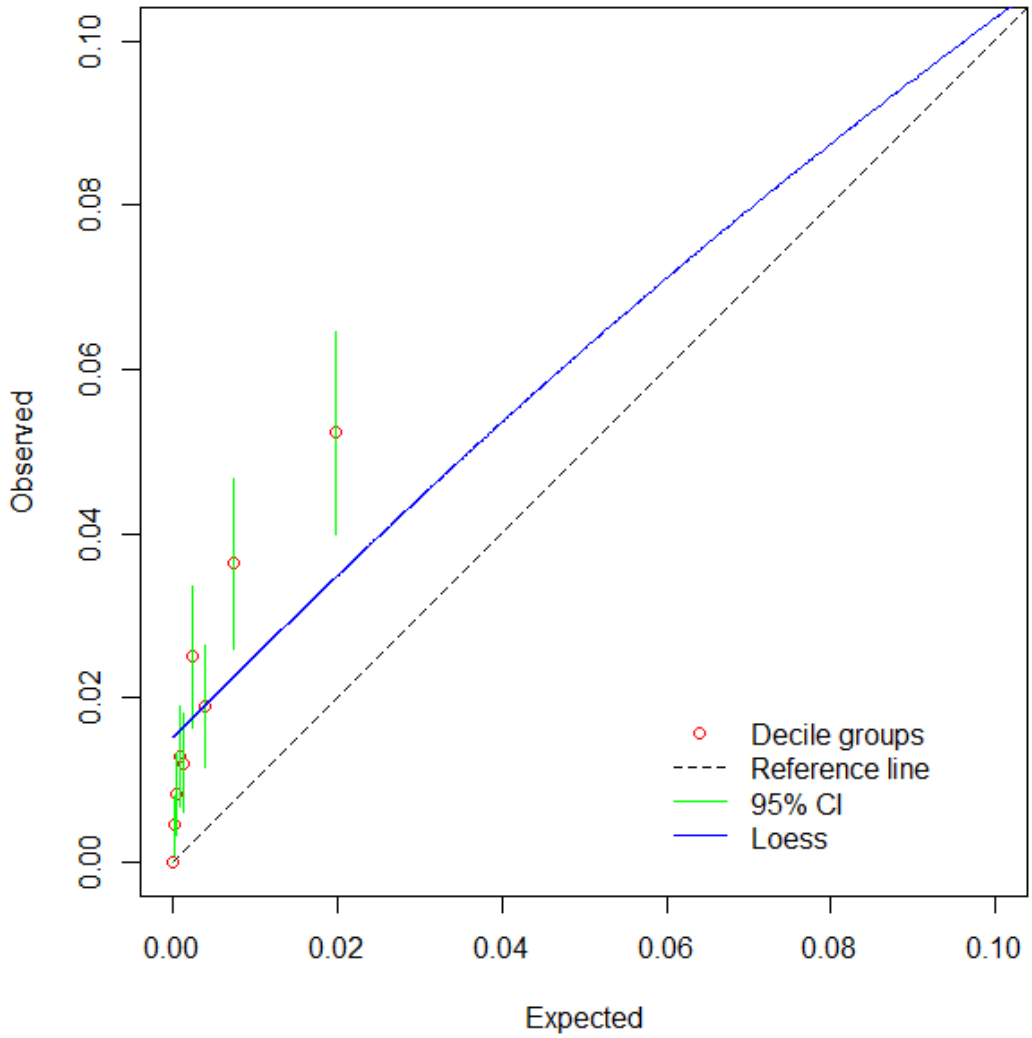



*Figure 6.7 - Focused calibration plot of the extended TMACS model plotted by predicted risk deciles. A Loess line of best fit is also plotted.*

The reclassification of patients in the new extended model was a more complex picture than that for recalibration (Figure 6.8). The overall trend was for the downgrading of risk categories across low risk, moderate risk and high risk groups. The majority of the original very low risk group were again classified as very low risk however some participants were reclassified upwards with less than or equal to 5 moving to low and moderate risk groups respectively (data censoring is mandated in small groups). For the original low risk group, the majority were reclassified downwards to the new very low risk group, however 1.85% kept their low-risk classification and six were reclassified upwards to moderate risk. The original moderate risk group was mainly reclassified to very low risk (70.64%), low risk (14.72%) and moderate risk (14.62%). Less than or equal to 5 participants were classified upwards to high risk. A large portion of the high risk group was again classified as high risk in the extended TMACS algorithm (45.13%), with 31.6% moving to moderate risk, 2.18% moving to low risk and 21.07% being reclassified to very low risk.

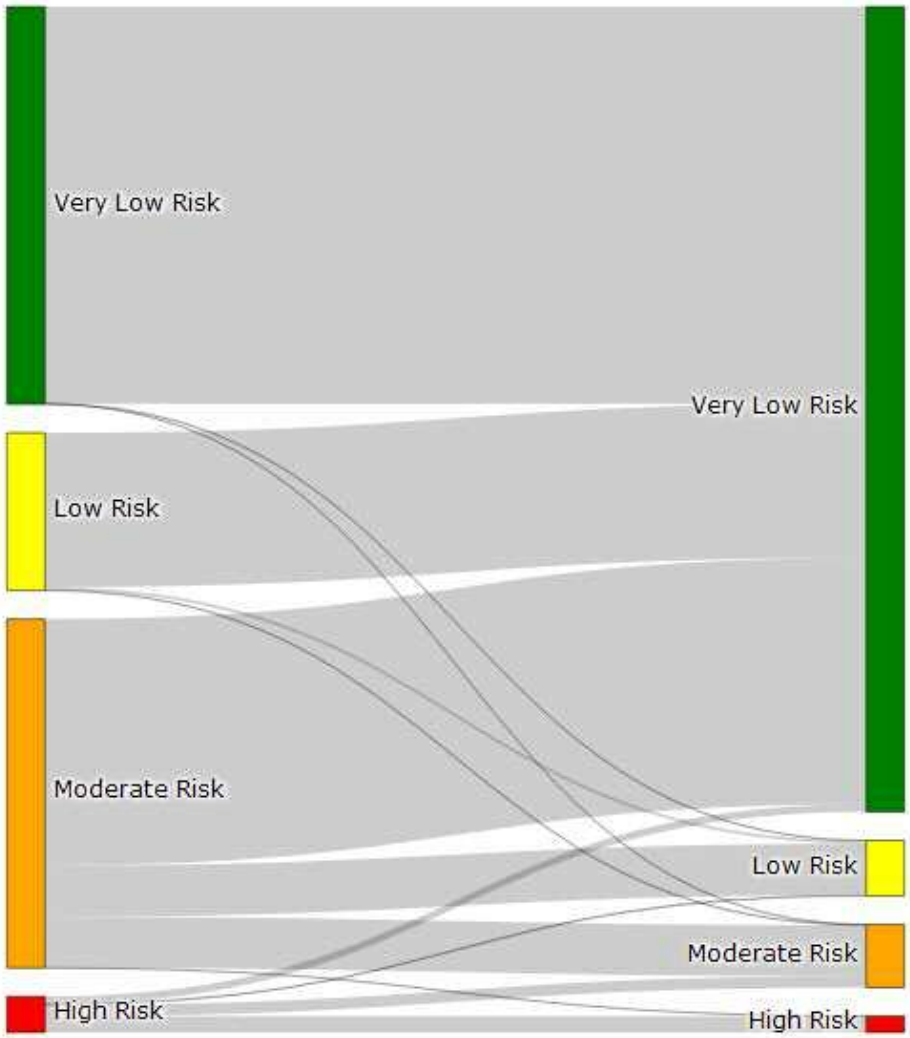



*Figure 6.8 Sankey Diagram for reclassification of risk from the original predicted risk groups on the left to the newly predicted risk groups on the right from the extended TMACS model.*

### 6.5.2.1 Sensitivity analysis

A sensitivity analysis was again conducted to examine the effect of not pooling the data from different sites and instead deriving at the site level. The AUC for the extended model for MRI, Burnley and Blackburn were 0.80 (95% CI 0.78 to 0.82), 0.76 (95% CI 0.70 to 0.82) and 0.84 (95% CI 0.79 to 0.90) respectively. The calibration measured by CITL was -0.06 (95%CI -0.18 to 0.06), 1.03 (0.72 to 1.32), and -0.41 (95% CI 0.78 to 0.06) respectively.

A further sensitivity analysis was conducted including co-linear terms into the pool of potential predictors prior to forward stepwise selection. After assessment by co-linear plots (example in figure 6.9) and consideration of biological plausibility. After this the following colinear terms were considered [ethnicity ∗ kidney function], [ethnicity ∗ troponin], [gender ∗ troponin], [kidney function ∗ troponin], and [kidney function ∗ age]. Following stepwise selection of the additional variables the following terms were selected: age, ethnicity, ethnicity*troponin, troponin*kidney function and gender*troponin. The final model characteristics were similar for domains of calibration with a CITL of 0.00 (95% CI -0.10 – 0.10) after adjustment for optimism was found to be -0.01. However, the AUC was 0.78 (95% CI 0.76 – 0.80) and 0.79 after adjustment for optimism, compared to 0.84 from the model extension with no co-linear terms.

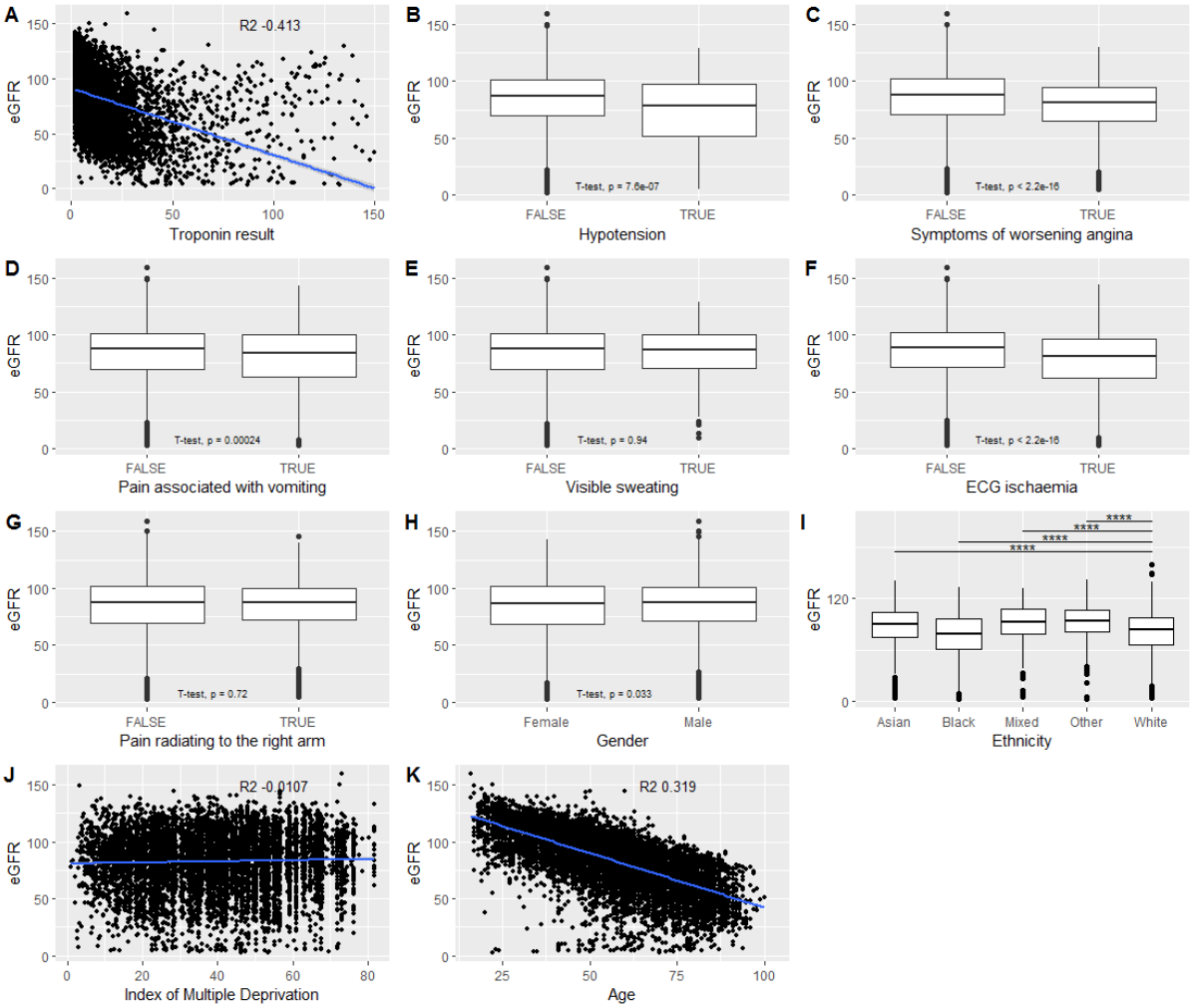



*Figure 6.9 Colinear plot of kidney function against other potential variables. eGFR – estimated glomerular filtration rate*

#### 6.5.2.2 Subgroup results

##### Site

The extended TMACS model decreased the discriminatory performance across the site subgroups with a lower AUC in all sites with Burnley falling from 0.92 to 0.77 (Table 6.10). The calibration was improved across each site with the CITL coming closer to 0. Burnley's CITL was -10.63 in the status quo model but in the extended model it had improved to 2.90

##### Gender

Similarly for the gender subgroups there was a drop in AUC for the extended TMACS model with the AUC for Males originally being 0.87 but dropping to 0.82 (Table 6.11). For Females the AUC decreased from 0.89 to 0.85 in the extended TMACS model. Calibration was improved across both male and female subgroups closing on 0 from -3.58 and -4.25 to -0.01 and -0.01 respectively.

| | **Extension** | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapped** | **Status Quo** |
| **AUC** | | | |
| MRI | 0.85 (0.83 to 0.87) | 0.85 | 0.87 (0.85 to 0.89) |
| Blackburn | 0.90 (0.86 to 0.95) | 0.91 | 0.91 (0.86 to 0.95) |
| Burnley | 0.77 (0.71 to 0.83) | 0.77 | 0.92 (0.88 to 0.96) |
| **CITL** | | | |
| MRI | -0.02 (-0.19 to 0.14) | -0.03 | -3.00 (-3.18 to -2.82) |
| Blackburn | -4.52 (-5.38 to -3.65) | -4.52 | -8.34 (-9.29 to -7.38) |
| Burnley | 2.90 (2.58 to 3.19) | 2.90 | -10.63 (-11.63 to -9.63) |

*Table 6.10 Extended TMACS summary statistics by index of clinical site AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | **Extension** | | |
|---|---|---|---|
| | **Derivation** | **Bootstrapped** | **Status Quo** |
| **AUC** | | | |
| Male | 0.82 (0.80 to 0.85) | 0.82 | 0.87 (0.85 to 0.89) |
| Female | 0.85 (0.82 to 0.88) | 0.85 | 0.89 (0.87 to 0.92) |
| **CITL** | | | |
| Male | 0.00 (-0.18 to 0.18) | -0.01 | -3.58 (-3.81 to -3.35) |
| Female | 0.00 (-0.26 to 0.26) | -0.01 | -4.25 (-4.58 to -3.92) |

*Table 6.11 Extended TMACS summary statistics by gender. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

#### Age

Age quartiles demonstrated a drop in discriminatory performance from the status quo TMACS to the extended model. The fourth age quartile dropping from 0.85 to 0.79 (Supplementary Table 8.16). However, the calibration again improved across all quartiles with the range of the CITL for the extended model being 0.652 to -0.180 compared to the status quo's range of -4.83 to -3.21.

#### Ethnicity, Age and Index of Multiple Deprivation

This pattern of improvement was also demonstrated in ethnicity and IMD subgroups with AUC decreasing but CITL improving (Supplementary Table 8.17 and Supplementary Table 8.19). The kidney function subgroups again demonstrated the same decrease in AUC but the CITL still demonstrated marked miscalibration (Supplementary Table 8.18). The stage five kidney function group had a CITL of -10.21 in the extended TMACS model from -14.96. The magnitude of the improvement for CITL from the extended to status quo model is similar across the subgroup.

### 6.5.3 Dynamic updating

#### *6.5.3.1 Forgetting factor selection (lambda)*

To proceed with the dynamic updating a lambda was required, to avoid confounding from pooling between sites for this lambda selection only the Manchester data was used. I initially used Riley et al's sample size calculation to calculate a suitable lambda as suggested in a doctoral thesis by Jenkins et al (233). This lambda is used in Figure 6.10 (lambda = 0.99876), this figure is a beta plot of the intercept and variable coefficients over time. It enables visualisation of the change of the model over time. There are dramatic changes in the coefficients with this lambda and these dramatic changes indicate instability that is not clinically acceptable. Particularly for the coefficients that cross 0 indicating a switch from positive to negative predictive properties (Figure 6.10– graph C,D,E,F,G and H). The lambdas up to 0.9999 were examined (at increments of 0.9990, 0.99925. 0.99950, and 0.99975), all except 0.9999 demonstrated the same instability (see supplementary figures). A

lambda of 0.9999 was therefore selected for model stability whilst still enabling the model to be dynamic (Figure 6.11).

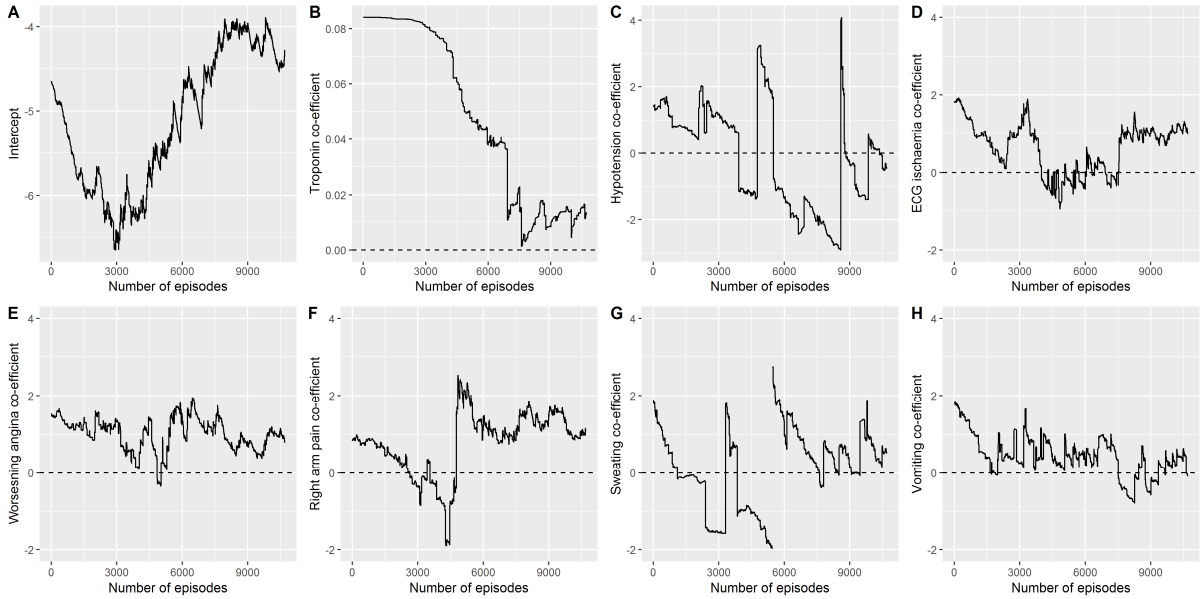



*Figure 6.10 Beta plot for lambda selection 0.99876 – Manchester data only used in Bayesian dynamic updating model. The intercept and coefficients of each variable are plotted on separate graphs each against time measured by number of episodes.*

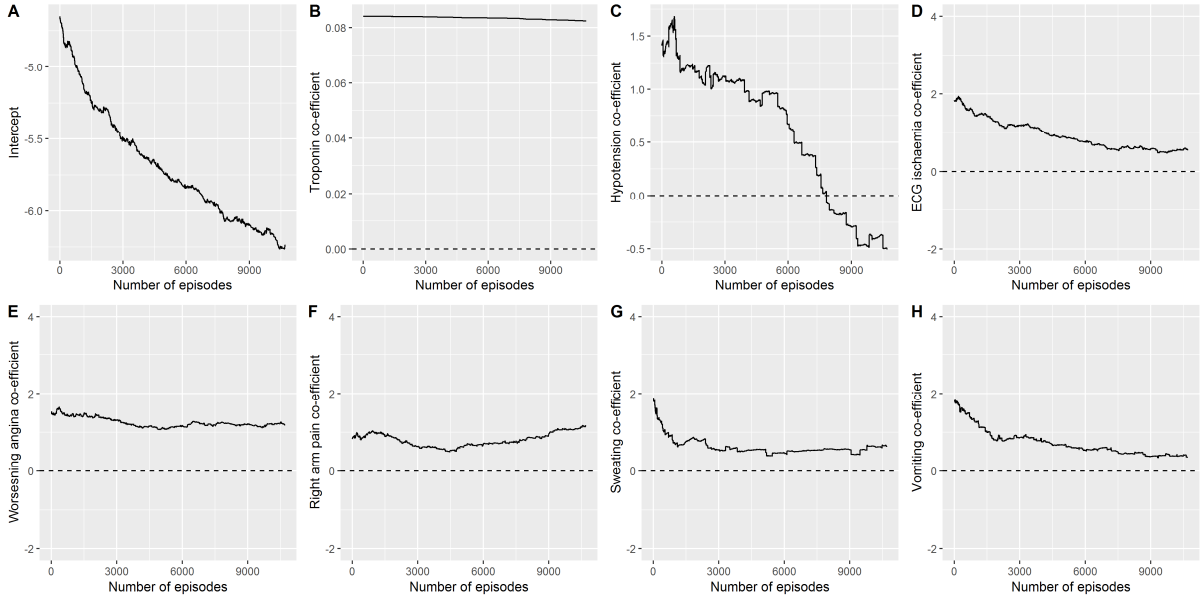



*Figure 6.11 Beta plot for lambda selection 0.9999 – Manchester data only used in Bayesian dynamic updating model. The intercept and coefficients of each variable are plotted on separate graphs each against time measured by number of episodes.*

### *6.5.3.2 Pooling review*

Using the selected lambda (0.9999) each site's data was used to create a dynamic updating model. The resulting beta plot was used to assess the suitability for pooling. Manchester and Blackburn demonstrated similar trends across the intercept and variable coefficients (Figure 6.11 & Figure 6.12). In contrast Burnley did not have the same trends of the other sites with the intercept demonstrating the near inverse trend and hypotension co-efficient not falling (Figure 6.13). Therefore I opted to pool Manchester and Blackburn's data and treat Burnley separately.

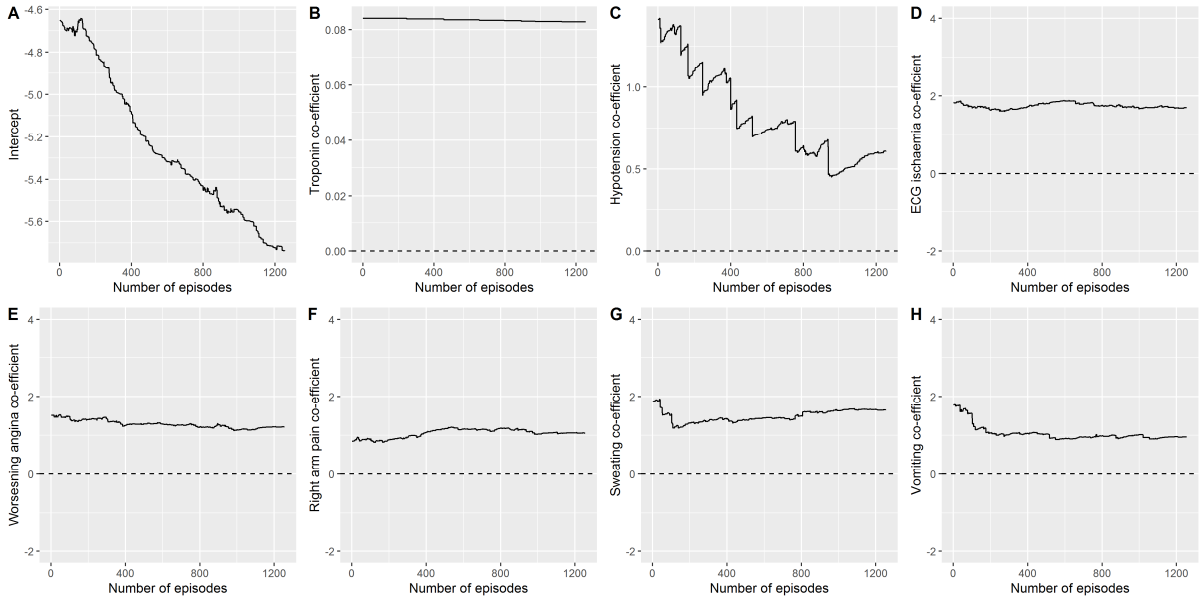



*Figure 6.12 Beta plot for Blackburn data only used in Bayesian dynamic updating model. Lambda = 0.9999. The intercept and coefficients of each variable are plotted on separate graphs each against time measured by number of episodes.*

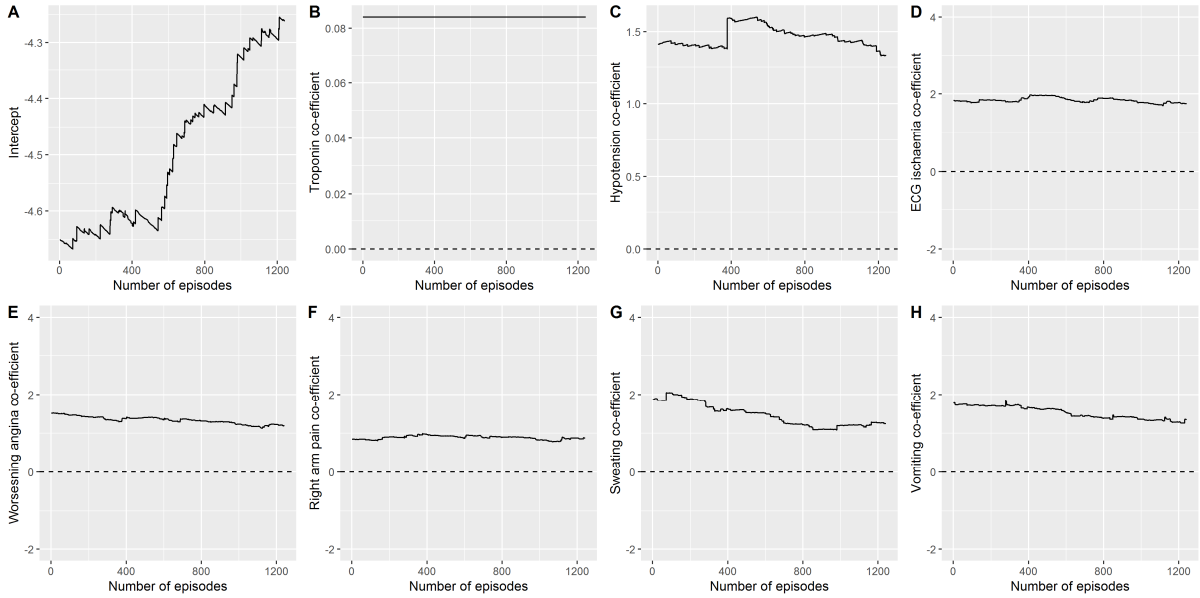



*Figure 6.13 Beta plot for Burnley data only used in Bayesian dynamic updating model. Lambda = 0.9999. The intercept and coefficients of each variable are plotted on separate graphs each against time measured by number of episodes*

### *6.5.3.3 Model Summary results*

The betas all varied overtime with the intercept and hypotension coefficient demonstrating the most substantial change from -4.65 to -6.48 and 1.412 to -0.82 respectively (Table 6.12). The patient representatives felt it was important to have a supervised period prior to a more automated one. Whilst this has no impact on the underlying statistical method used in this work and is a practicality for implementation, this 3-month supervised period is highlighted in (Figure 6.14) to assess the trends during this supervised period and after. During this period no beta demonstrated a dramatic change in trend after the supervised period. Sweating and vomiting coefficients became more stable after the supervised period.

| | **Initial** | **Final** |
|---|---|---|
| **Intercept** | -4.65 | -6.48 |
| **Coefficient** | | |
| **Troponin** | 0.084 | 0.079 |
| **Hypotension** | 1.412 | -0.82 |
| **ECG ischaemia** | 1.828 | 0.562 |
| **Worsening Angina** | 1.514 | 1.069 |
| **Right arm pain** | 0.849 | 1.098 |
| **Sweating** | 1.878 | 0.872 |
| **Vomiting** | 1.783 | 0.433 |

*Table 6.12 Comparison of TMACS model components from status quo to final dynamic update – for Manchester and Blackburn data only.*

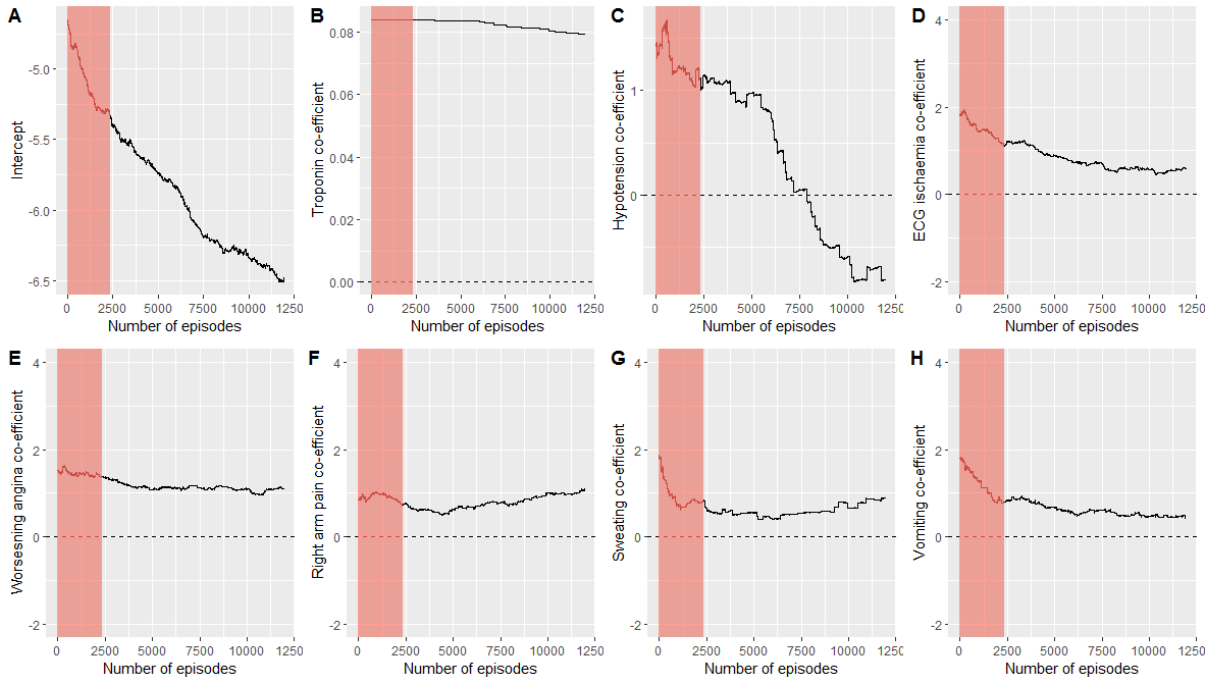



*Figure 6.14 - Beta plot for Manchester and Blackburn data only used in Bayesian dynamic updating model. Lambda = 0.9999. The intercept and coefficients of each variable are plotted on separate graphs each against time measured by number of episodes. A three-month intensive observation period is highlighted in red.*

The model derived by Bayesian dynamic updating from Manchester and Blackburn's pooled data had an AUC of 0.87 (95% CI 0.85 to 0.89), a CITL of -1.45 (95% CI -1.63 to -1.27) and a C-slope of 0.061 (95%CI 0.049 – 0.073) (Table 6.13) . Burnley's dynamically updated model had an AUC of 0.73 (95% CI 0.66 - 0.80) and a CITL of 0.47 (95% CI 0.18 to 0.76). The calibration of the model was also assessed with a calibration plot (Figure 6.15). This demonstrated good calibration especially at the lowest deciles. The sensitivity for the Manchester and Blackburn dynamic model was found to be 98.7% (95% CI 98.5 -98.9) and the specificity 80.7% (95% CI 77.0 to 84.4). For the dynamic updated model from the Burnley data, the sensitivity was 59.6% (95% CI 46.9 to 72.4) and the specificity was 89.3% (95% CI 87.5 to 91.0).

The cohort size for Burnley was not sufficient for detailed analysis by quarter and due to missing data subgroup analysis was not possible for gender or kidney function. The results therefore focus on the Manchester and Blackburn pooling.

| | Dynamic updating (95% CI) | Status Quo (95% CI) |
|---|---|---|
| **AUC** | | |
| MRI | 0.86 (0.84 to 0.88) | 0.87 (0.85 to 0.89) |
| Blackburn | 0.92 (0.89 to 0.95) | 0.91 (0.86 to 0.95) |
| Burnley | 0.73 (0.66 to 0.80) | 0.92 (0.88 to 0.96) |
| **CITL** | | |
| MRI | -1.15 (-1.32 to -0.98) | -3.00 (-3.18 to -2.82) |
| Blackburn | -5.87 (-6.79 to -4.94) | -8.34 (-9.29 to -7.38) |
| Burnley | 0.47 (0.18 to 0.76) | -10.63 (-11.63 to -9.63) |

*Table 6.13 Dynamically updated TMACS summary statistics by index of clinical site AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval. Manchester and Blackburn's results are from a pooled dynamically updated TMACS analysis.*

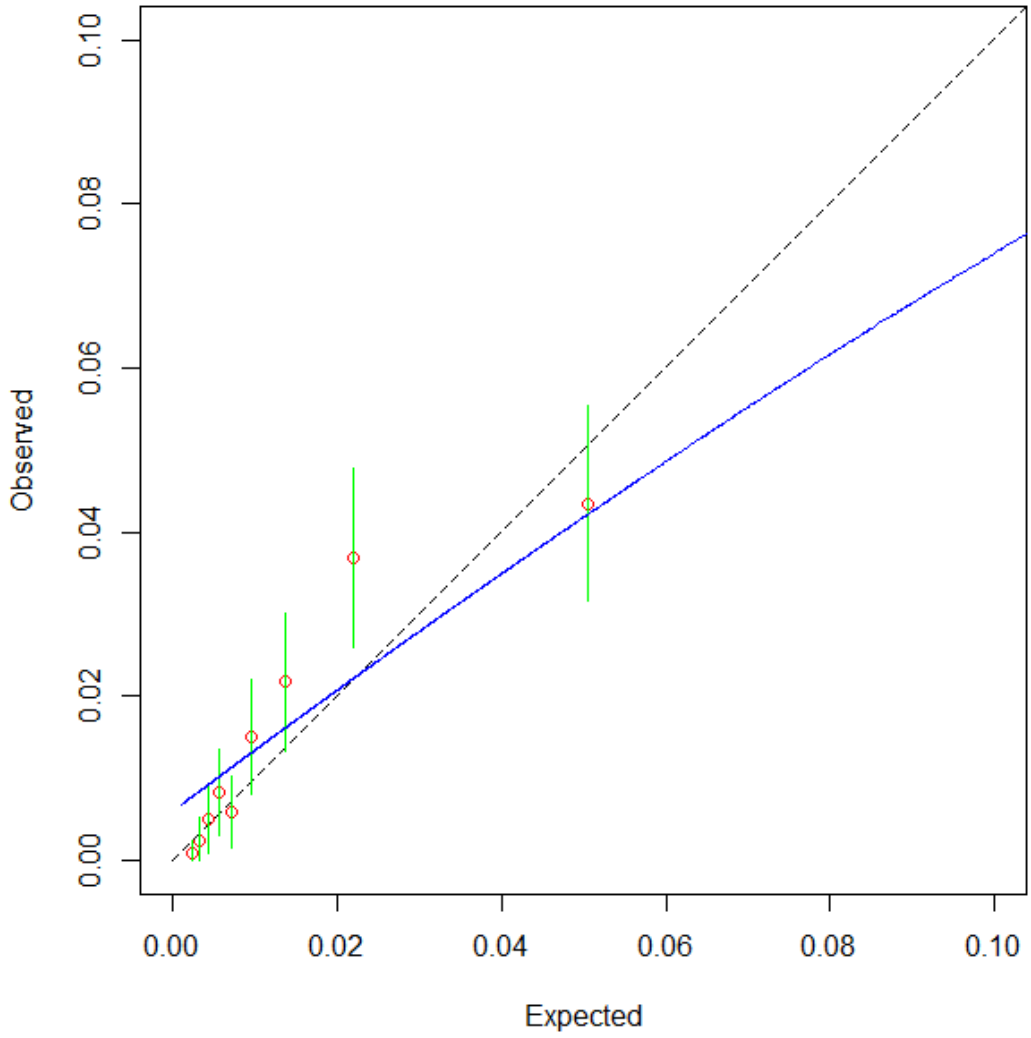



*Figure 6.15 Focused calibration plot of the dynamically updated TMACS model plotted by predicted risk deciles. A Loess line of best fit is also plotted. Manchester and Blackburn data only.*

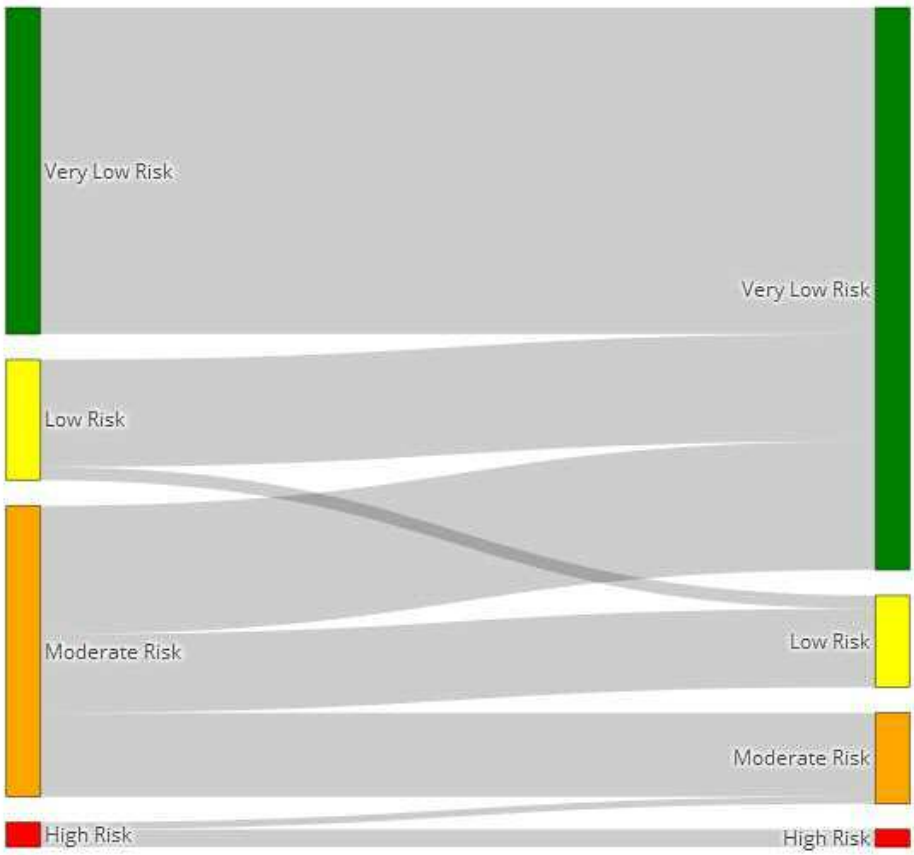



*Figure 6.16 Sankey Diagram for reclassification of risk from the original predicted risk groups on the left to the newly predicted risk groups on the right from the dynamically updated TMACS model. (Manchester and Blackburn data only)*

The reclassification of patient risk groups was predominantly a single step downgrading the group (Figure 6.16). The exception was the original moderate risk group in which a portion moved down two risk groups to very low risk. All the very low risk group were reclassified to the same group, 89.1% of the low risk group were classified as very low risk and the remaining 10.9% were classified as low risk again. 44.2% of the original moderate risk group were reclassified as VLR, 27.0% were classified as low risk and 28.8% were classified again as moderate risk. Of those originally classified high risk 71.8% remained high risk and 28.2% were reclassified to moderate risk.

As the model varied with time, I assessed the model performance per quarter to examine any change over time. Figure 6.17 shows the changing AUC over time starting a 0.93 (95% CI 0.85 to

1.00) and ending at 0.93 (95% CI 0.87 to 0.99). Whilst it ranged from a low of 0.78 (95% CI 0.67 to 0.89) in quarter 10 to 0.98 (95% CI 0.92 to 0.97) in quarter three no clear trend was visible, and most of the quarters had AUCs with overlapping confidence intervals.

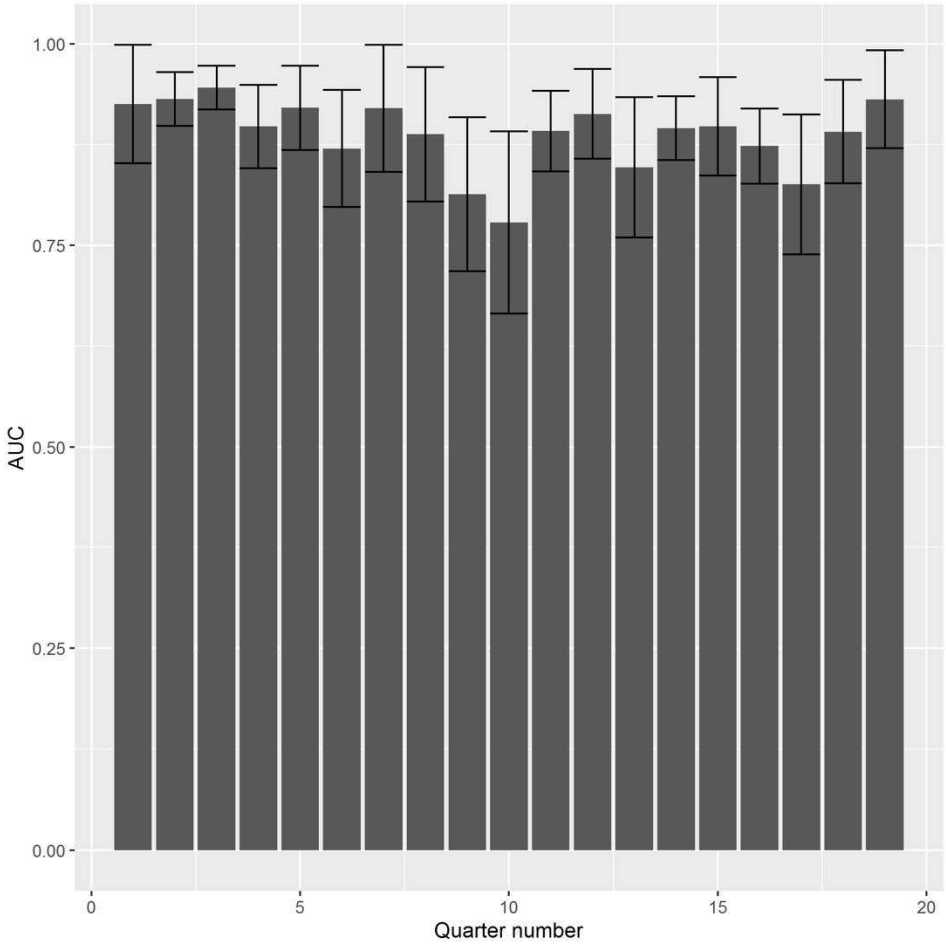



*Figure 6.17 Area under the curve of the dynamically updated TMACS model per quarter. Confidence intervals are displayed.*

Whilst no trend was identifiable from the AUC over time the CITL demonstrated a clear trend for improvement (Figure 6.18). Q1 CITL was – 2.16 (95% CI -3.45 to -0.87) and finished at 0.14 (95% CI -1.05 to -1.34).

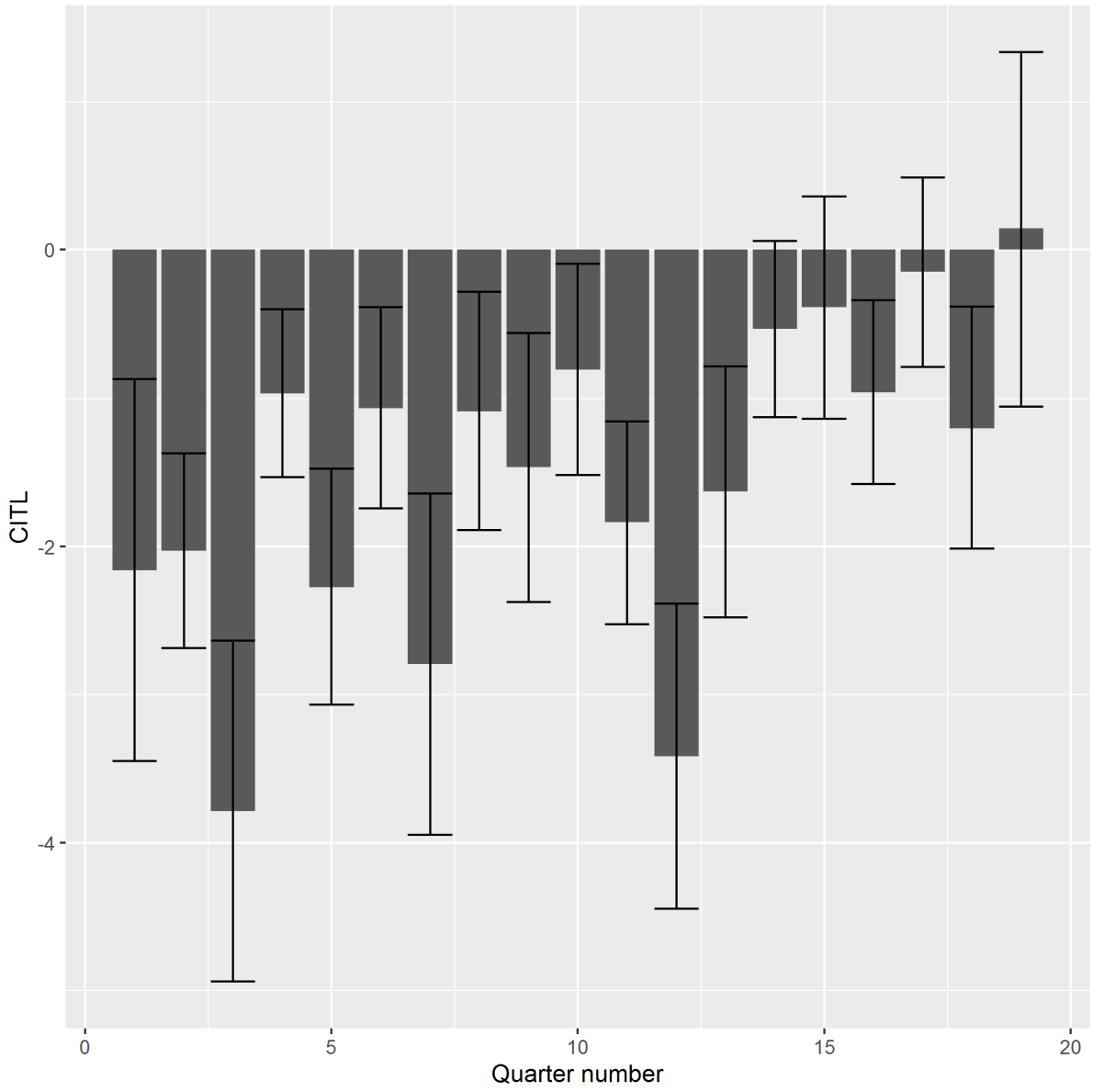



*Figure 6.18 Calibration in the large of the dynamically updated TMACS model per quarter. Confidence intervals are displayed.*

#### *6.5.3.4 Subgroups*

In the age subgroups, the AUC was calculated and the point estimate results were similar to the status quo results with a slight up to 0.02 decrease (Supplementary Table 8.21). The confidence intervals all overlapped, however there was a decrease of up to 0.03. The CITL across the age subgroups was improved in the dynamically updated model, there was still a slight overestimation of risk. A similar pattern was seen in the IMD, ethnicity and gender subgroups (Supplementary Table 8.22, Supplementary Table 8.23, and Supplementary Table 8.24). In the kidney function subgroups the AUC followed the pattern of a marginal decrease (Supplementary Table 8.25). The calibration whilst improved was still significantly mis-calibrated. There was a trend for worsening calibration with increasing CKD stage with stage groups 4 and 5 having a CITL of -5.53 (95% CI -7.13 to -3.93) and -13.26 (95% CI -15.49 to -11.03) respectively.

## 6.6 Discussion

In this analysis I successfully trialled three methods for updating a clinical prediction model. Calibration was improved across all methods however discrimination was impacted to varying degrees (Table 6.14).

Recalibration demonstrated marked improvements in calibration including across subgroups for kidney function. This was not improved for the other updating methods. This model updating technique had the best calibration results with a CITL and unadjusted c-slope. However, when calibration is examined with a non-parametric line of best fit or the c-slope was adjusted an overestimation of risk was apparent.

Model extension allowed new variables to be added to the model, however it did not improve the discriminatory performance. Whilst there were modest gains in calibration these were overshadowed by the benefits from recalibration. This method led to a general lowering of the risk groups, however some patients were reclassified to higher risk groups. I did not observe this with

any other of the updating methods and was driven by the addition of new variables. It unfortunately did not lead to a substantial improvement in discrimination and the improvements in calibration were better by other methods. The substantial miscalibration here is likely partly attributable to the forced addition of the original TMACS to the model.

The discrepancy between the parametric (C-Slope) and non-parametric assessment (LASO fitted lines of best fit) may indicate that in this CPM non-parametric assessment is more appropriate. It is possible that the skewed distribution of the predicted probabilities unduly effects the parametric measures of calibration.

| | **Updating method (95% CI)** | | | |
|---|---|---|---|---|
| | **Status Quo** | **Recalibration** | **Extension** | **Dynamic Updating** |
| **AUC** | 0.88 (0.86 to 0.89) | 0.88 (0.86 to 0.89) | 0.84 (0.82 to 0.86) | 0.87 (0.85 to 0.89) |
| *Adjusted* | | 0.88 | 0.84 | |
| **CITL** | -3.93 (-4.12 to -3.74) | 0.00 (-0.09 to 0.09) | 0.00 (-0.15 to 0.15) | -1.45 (-1.63 to -1.27) |
| *Adjusted* | | 0.00 | -0.01 | |
| **C-Slope** | 0.05 (0.17 to 0.22) | 1.00 (0.82 to 1.12) | 0.07 (0.06 to 0.08) | 0.06 (0.05 to 0.07) |
| *Adjusted* | | 0.03 | 0.39 | |

*Table 6.14 TMACS summary statistics across different updating methods. 95% CI – 95% confidence interval, AUC – area under the curve, CITL – calibration in the large, C-slope – calibration slope. Adjusted – adjusted for optimism with bootstrapping.*

Dynamic updating demonstrated fascinating results with maintained discriminatory characteristics and improved calibration. The model adapted over time, varying with the patient cohort that it was exposed to. The CITL steadily improved as the model learnt. The performance in subgroups was also improved across the domains of discrimination and calibration except kidney function. In this analysis only recalibration demonstrated a marked improvement in patients with poor kidney function (CKD stage >3). The dichotomised hypotension variable in the dynamically updated model changed from being positively predictive of the outcome to negatively. An exploratory analysis was conducted to check the linearity of the continuous form of this variable in relation to the outcome (Figure 6.19). In contrast to the original TMACS derivation cohort (unpublished data), SBP did not

demonstrate a linear relationship. Whilst the lower deciles did demonstrate an increased risk so did the upper most deciles. This currently dichotomised variable could be a target for future improvement.

Hickey et al previously reported a deterioration in the calibration of EUROSCORE, a CPM to predict mortality in patients undergoing cardiothoracic surgery (229). EUROSCORE II was an updated version of the original CPM, with re-derivation and extension to include new variables. Whilst forward stepwise selection was used as in our extended model, univariable regression was used to screen candidate variables and they also rederived the original model (234). The original model has an AUC of 0.79 and the updated model an AUC of 0.81 which was not a statistically significant different trend. Calibration was examined by comparing the expected vs observed mortality which was 3.95% vs 4.18% which was deemed to be an acceptable under-prediction (234). Dynamic updating was also trialled for this cohort and it was demonstrated to be possible and variables changed from being positively to negatively predictive as well, potentially indicating that this a phenomenon not specific to TMACS (235).

Dynamic prediction models have been thought to be able to counter the static nature of derivation or updating by continually updating (236). However, the required connected digital infrastructure to enable such a system is expensive and may require a more mature digital healthcare system.

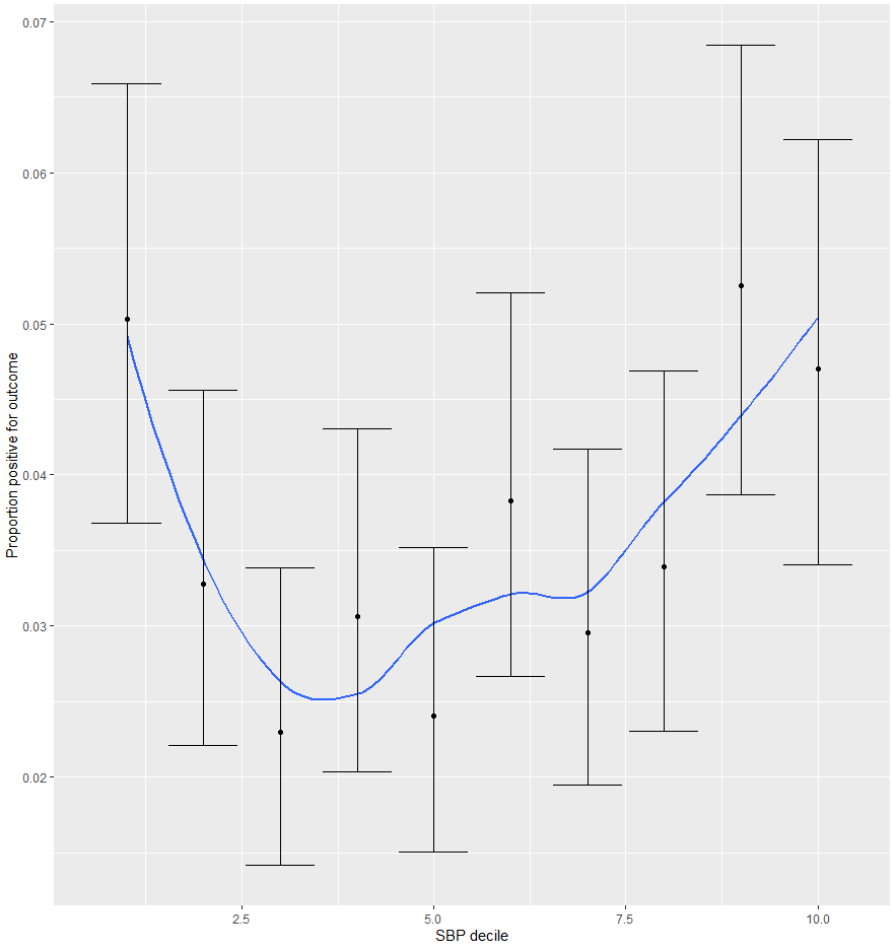



*Figure 6.19 - Linearity plot of Systolic blood pressure deciles (SBP) by the proportion of each decile positive for the primary outcome. Line of best fit by lasso and the 95% confidence intervals are plotted.*

## 6.7 Conclusion

The TMACS CPM can have its calibration updated and discrimination maintained by recalibration and dynamic updating and whilst model extension did improve calibration it did not maintain the discriminatory performance and performed worse across the subgroups of interest. The kidney function subgrouping revealed poor model calibration at CKD stage 3 and above, whilst this was improved most by recalibration it is a potential target for future work. Dynamic updating offers the prospect of a model that can adjust to changes in the population and if nested within the required digital infrastructure it could constantly update the CPM.

# Chapter 7 Conclusions

The overarching objective of improving ED care and outcomes for CVD in the long and short term has been achieved through two aims. Firstly, I have developed a new care pathway to predict future CVD in the acute care setting by developing the quantitative evidence base and the qualitative data for how to implement such a pathway in chapters 2 through 4. This work creates a care pathway that utilises routinely collected data with minimum impact of ED resources and takes advantage of the teachable moment in a patient's journey. Secondly TMACS has been validated in its largest cohort, allowing for a detailed analysis of its real-world diagnostic characteristics in chapters 5 and 6. This revealed relatively preserved diagnostic accuracy (in terms of sensitivity and NPV) but suboptimal calibration. The latter did not appear to be deteriorating progressively over time. Loss of calibration could be countered with model updating strategies, although this led to compromised sensitivity as a rule-out tool. Clinical CPMs that have been deployed in clinical practice should be assessed for similar deterioration in their performance characteristics.

## 7.1 Limitations

The limitations for this work are explored individually in each chapter. My thesis uses qualitative and quantitative data which is predominantly from the Greater Manchester area. When seeking to generalise the conclusions of my work others should examine if factors such as population demographics or clinical practices are similar enough to enable this. The quantitative elements of my thesis rely on the accuracy of diagnostic and interventional coding in local sites. Coles et al examined the accuracy of HES for detection of AMI in a cancer cohort (237). They found a discrepancy in STEMI admissions where 16.5% of admissions were not recorded in HES. The study did not seek to understand if these were false positives or true positives however it is a signal that warrants further investigation with clinical verification of coded AMI outcomes. The national data opt out is also another factor to consider when interpreting the quantitative data. Patients in the UK have the option to opt out of their data being shared for research purposes, the national average is 5.36% and

the northwest has an overall average of 5.9% (238). This may have removed positive CVD outcomes from this work and may artificially lower the primary outcome incidence. It is possible that patients who have opted out are systematically different to the remaining patients, which could mean that the models evaluated perform differently in those patients.

## 7.2 Long-term cardiovascular care

The long-term CVD care pathway relies on close co-ordination between primary and secondary care. The interface between these two parts of the patient's journey was highlighted as a risk for any potential pathway. More broadly, work is required to bring primary and secondary care stakeholders together to better understand their peers' perspectives. Patients expect the different parts of the health care system to be seamless in the way they operate, however, operationally there is room for more alignment. The strength of such a pathway is the use of routinely collected data to inform the risk of long-term CVD. Age, gender, WBC, ethnicity, hs-cTnT and TMACS all demonstrated favourable prognostic characteristics. In addition, informative missingness was identified for cardiac specific investigations (cTnT, hs-cTnT and TMACS), indicating that the decision not to investigate cardiac causes of chest pain was associated with a lower rate of CVD. More work could be conducted to better understand what is underlying this decision making, that may be prognosticating for long term outcomes (absence of symptoms etc.) The external validation of the Framingham model demonstrated accepted characteristics in the domains of discrimination and calibration. It may be that this model could be improved with the addition of hs-cTnT or TMACS. This work lays the foundation for a long-term CVD care pathway for acute care and whilst further work is required to assess the implementation of the proposed pathway my research has demonstrated that this warrants further investigation.

CPMs have been re-purposed from facilitating the early diagnosis of ACS in the ED to predicting long-term CVD. Farkouh et al examined patients who were triaged according to the Agency for Health Care Policy and Research AMI CPM (147). Their participants had a medium follow up of 7.3 years and

they found patients deemed high risk by the CPM had a hazard ratio of 2.45 (95% CI 1.67-3.58) for cardio and cerebrovascular events. Kavsak et al examined laboratory predictors including hs-cTnT and found sensitivities of  up to 98.5% for the prediction of CVD outcomes at one year (158). Etaher et al examined hs-cTnT in the elderly with suspected ACS with a median follow of 32 months, it was reported that regardless of index diagnosis there was an increased rate of mortality with higher baseline hs-cTnT levels (159). Whilst these models are not looking into the decade long follow up associated with primary care CVD CPMs, they demonstrate a potential.  Hinton et al in a multi-variate cox regression found that a hs-cTnI relative to upper reference limit >0.25, had a statistically significant hazard ratios of 2.09 (95% CI 1.40 – 3.13), 2.10 (95%CI 1.37 – 3.23), and 6.02 (95%CI 2.20 – 16.47) in inpatient, outpatient and ED settings respectively (190). Age and gender were also found to predictive in this setting.

## 7.3 Acute cardiovascular care

TMACS has been validated in a variety of settings previously and the AUC has been reported to be between 0.86 to 0.91 in validations (49,222). Due to the smaller sample size in previous prospective studies, its performance has not been evaluated in subgroups of interest. In contrast this work is the largest evaluation of T-MACS to date. It does however utilise coded diagnoses from the clinical record in contrast to prospective observational studies with carefully validated outcomes (220).

There was an apparent deterioration in the performance of TMACS between the prospective derivation clinical trial and this real-world evaluation (231). Direct comparison of the two trials is difficult due to the lack of adjudicated outcomes in the real-world evaluation. There was variation in the performance of the model within this real-world evaluation in domains of calibration and discrimination. There was variation in model performance across subgroups, particular kidney function. TMACS demonstrated calibration drift with miscalibration measured by CITL and c-slope. However more favourable calibration was identified when non-parametric lines were fitted to

calibration plots, particularly at the lower predicted risk deciles in which AMI would be ruled out by TMACS. This focus is useful, given the need for caution when treating suspected AMI, however it may be at the expense of better calibration at the higher predicted risk thresholds. The focus of future work could be to understand the most appropriate calibration metric for such a scenario. Further research could also look to adjudicate the coded outcomes so as to better align with the prospective cohort trial that derived the work (231).

The calibration of TMACS across different genders and ethnicities varied. The CITL for males was -3.58 (95% CI -3.81 to -3.35) compared to -4.25 (95% CI -4.58 to -3.92) for females. This is potentially due to the derivation population which was 61.2% male, however this was representative of the gender disparity in cardiovascular disease (220). The calibration by ethnicity also varied substantially, the Asian ethnicity subgroup had a CITL of -2.52, Mixed -3.71, Other -3.86, White -4.03 and Black -5.98. The ethnicity of the derivation cohort was not described so inferences cannot be made if this miscalibration is due to the derivation cohort. The prospective verification of these results will be important to better understand this apparent discrepancy in performance. It is possible that the CPM and its selected predictors vary in their diagnostic ability between ethnicities, this is of interest as different AMI symptom profiles have been noted previously in Canada (239). It is also possible that the outcome diagnostic and intervention codes are the source of the miscalibration.

Updating TMACS did not yield the same discrimination statistics as its original clinical trial validation, however, it did improve the calibration of the CPM. This improvement was seen across various subgroups. The static recalibration yielded substantial improvements in calibration, but it is a method that will require constant manual updating. The dynamic updating model yielded similar improvements but is a method that could be semi-autonomous with the pre-requisite digital infrastructure. Careful consideration is required for variables which change from being positively associated with the outcome to negative. In this work we allowed the variable to flip, however it

may be that where there is biological implausibility the variable is only permitted to become null (equal to zero).

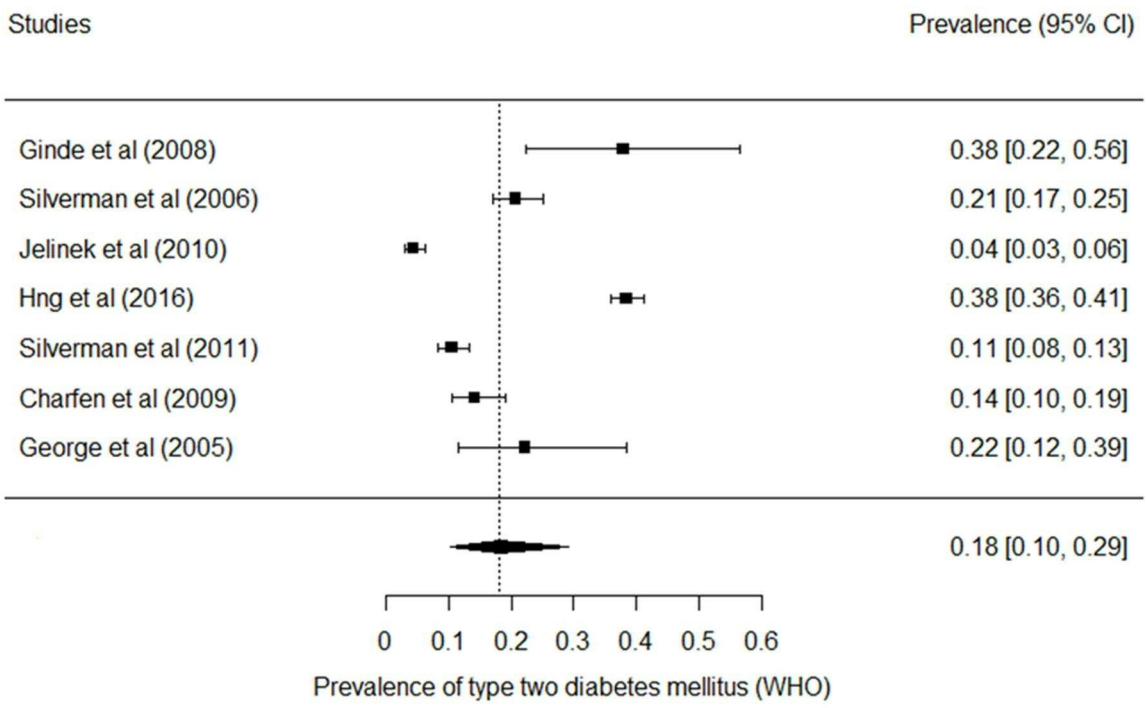
The in-depth assessment of a CPM highlights important areas for the future including the performance and improvement of models by ethnic group. This large RWE evaluation of a CPM also emphasises the importance of monitoring the performance of already implemented CPMs. Guidelines for the reporting of CPMs highlight external validation as an important step, now implemented CPMs are widespread and addition may be not just external but temporal validation (45). This could be a further validation at a post-implementation time point to assess the CPM's performance. Specifically to TMACS further formal health economic assessment is required to understand which updating method is viable, balancing the cost of each respective method versus the improvement in performance. This would have to assess costs of digital infrastructure required for dynamic updating versus the benefit of continual updating on clinical outcomes.

## 7.4 Personal development

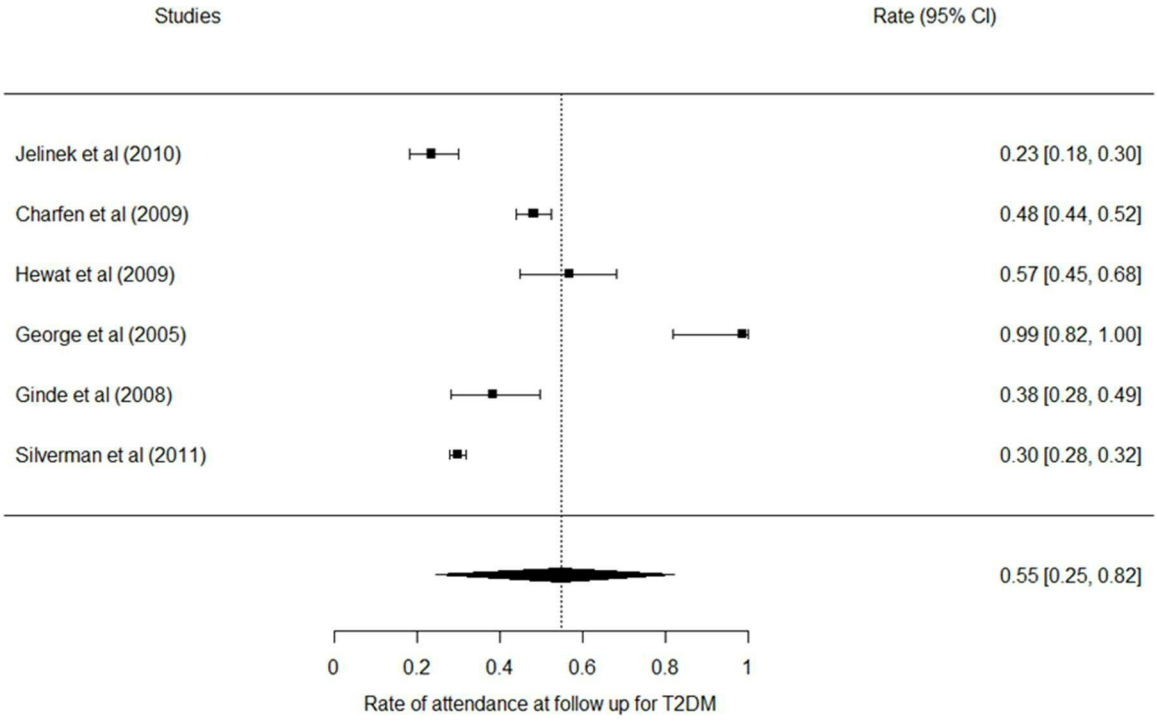
I have gained extensive experience through my PhD journey in qualitative and quantitative methodologies as well as practical operational knowledge required to run clinical trials. I have obtained new skills in qualitative methodologies and thematic analysis. I have also gained skills in real world evidence including database creation, cleaning, linking and machine learning methodologies. To deploy these newly developed skills I also successfully setup, gained approval and ran two clinical trials as chief investigator and principal investigator. This included complicated research ethics applications and data sharing agreements. This development has given me a broad transferable skill set.

# Chapter 8 Supplementary Materials

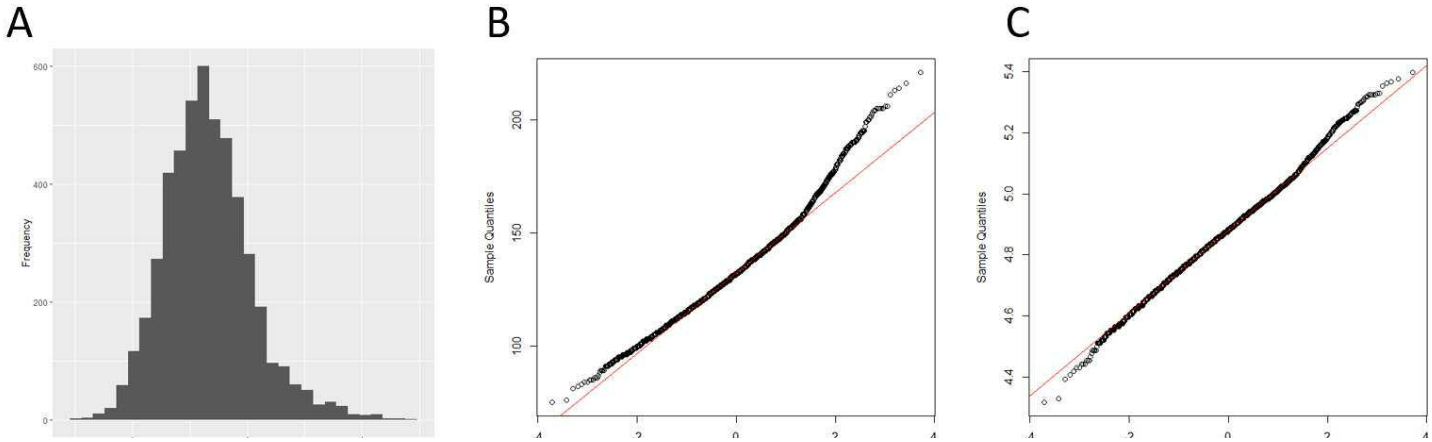
## 8.1 Supplementary figures



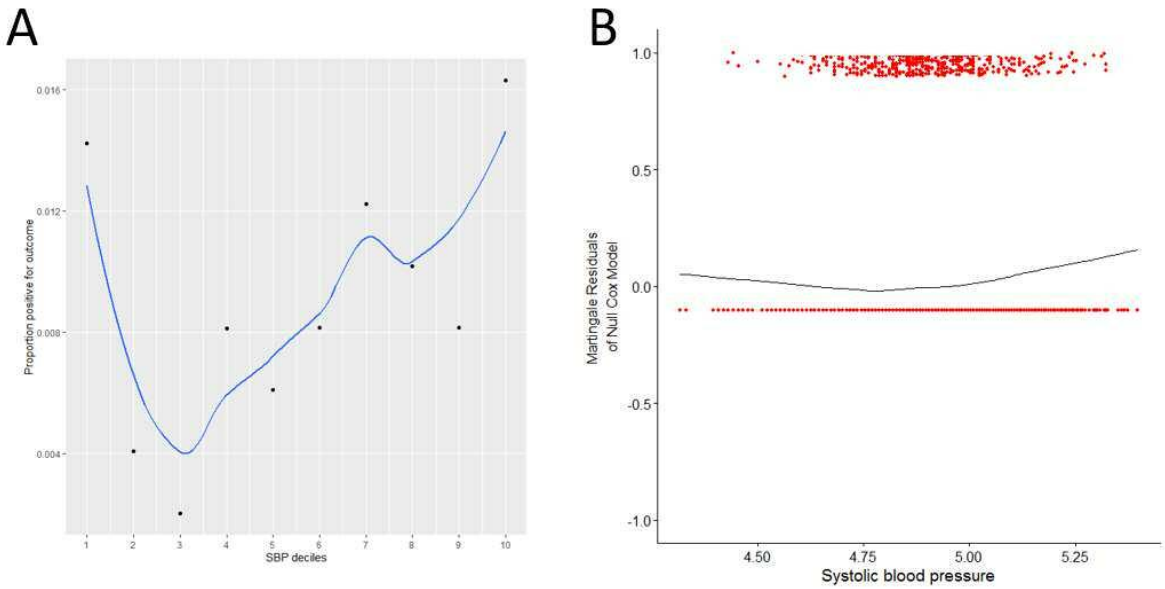

*Supplementary Figure 8.1 - Meta-analysis forest plot of the prevalence of type two diabetes in the emergency department as per the World Health Organisation definition (50–54,56,57).*



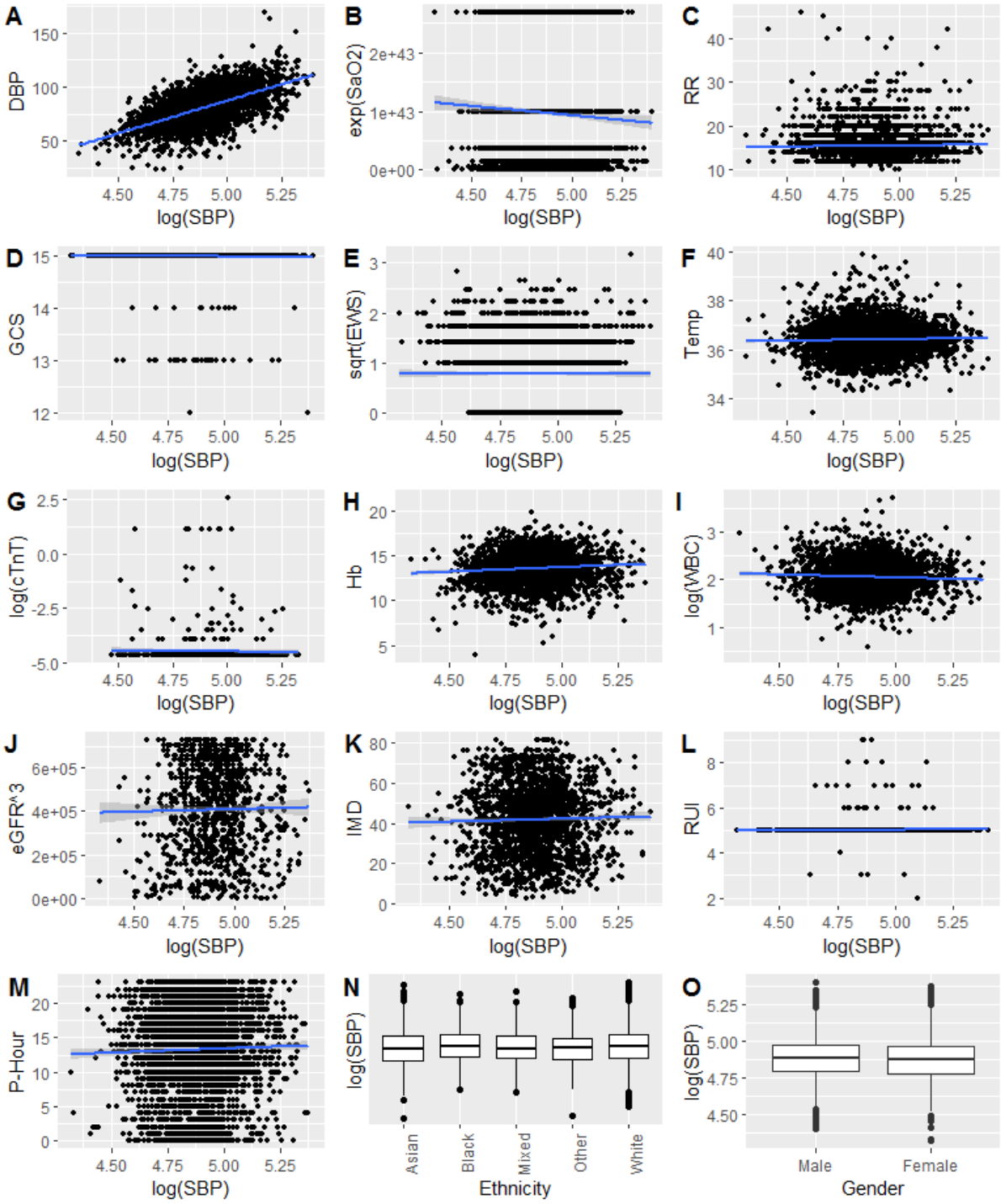

*Supplementary Figure 8.2 Meta-analysis forest plot for the proportion attendance at follow up appointments for type two diabetes (50,52–56)*



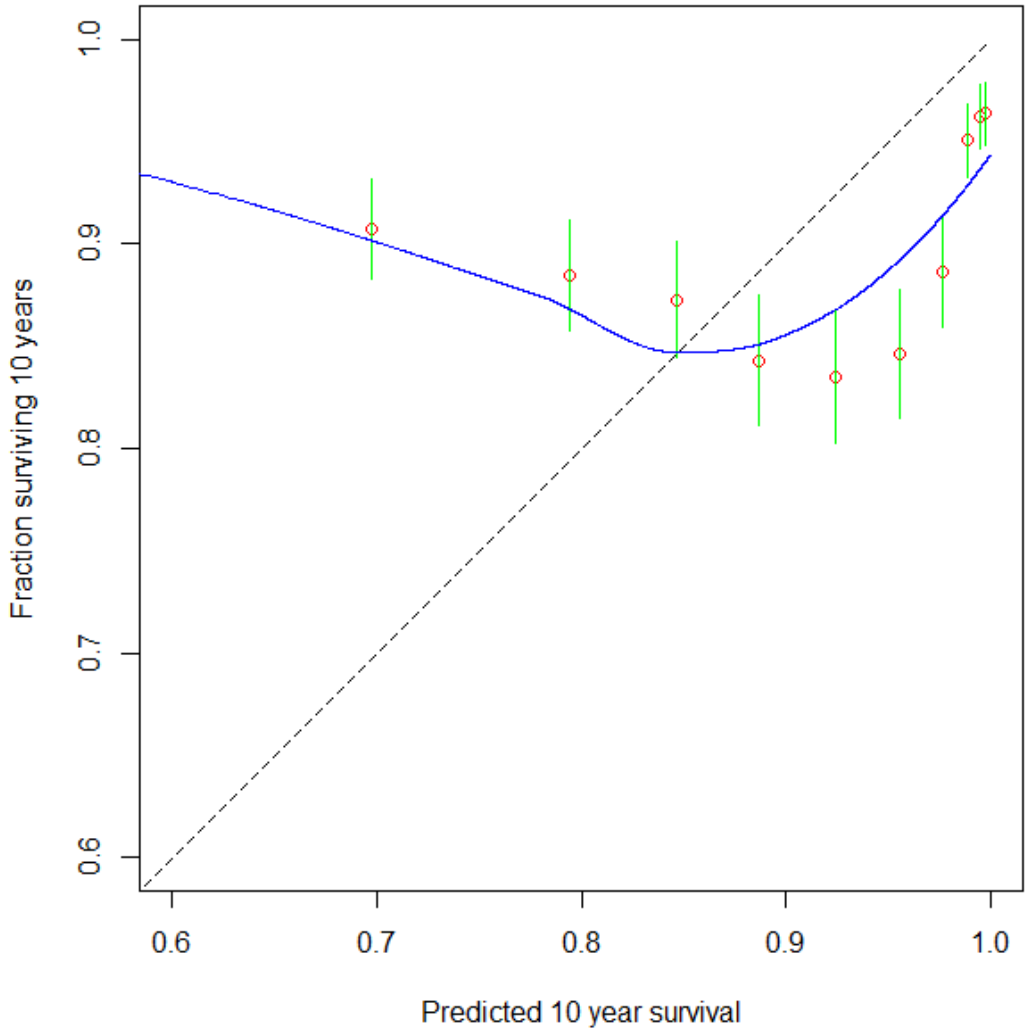

*Supplementary Figure 8.3 Skew check for systolic blood pressure variable. A – A histogram of the systolic blood pressure measured in millimetres of mercury, B – Quantile – Quantile plot of the untransformed systolic blood pressure variable, C- a Quantile-Quantile plot of the systolic blood pressure variable logarithmically transformed.*



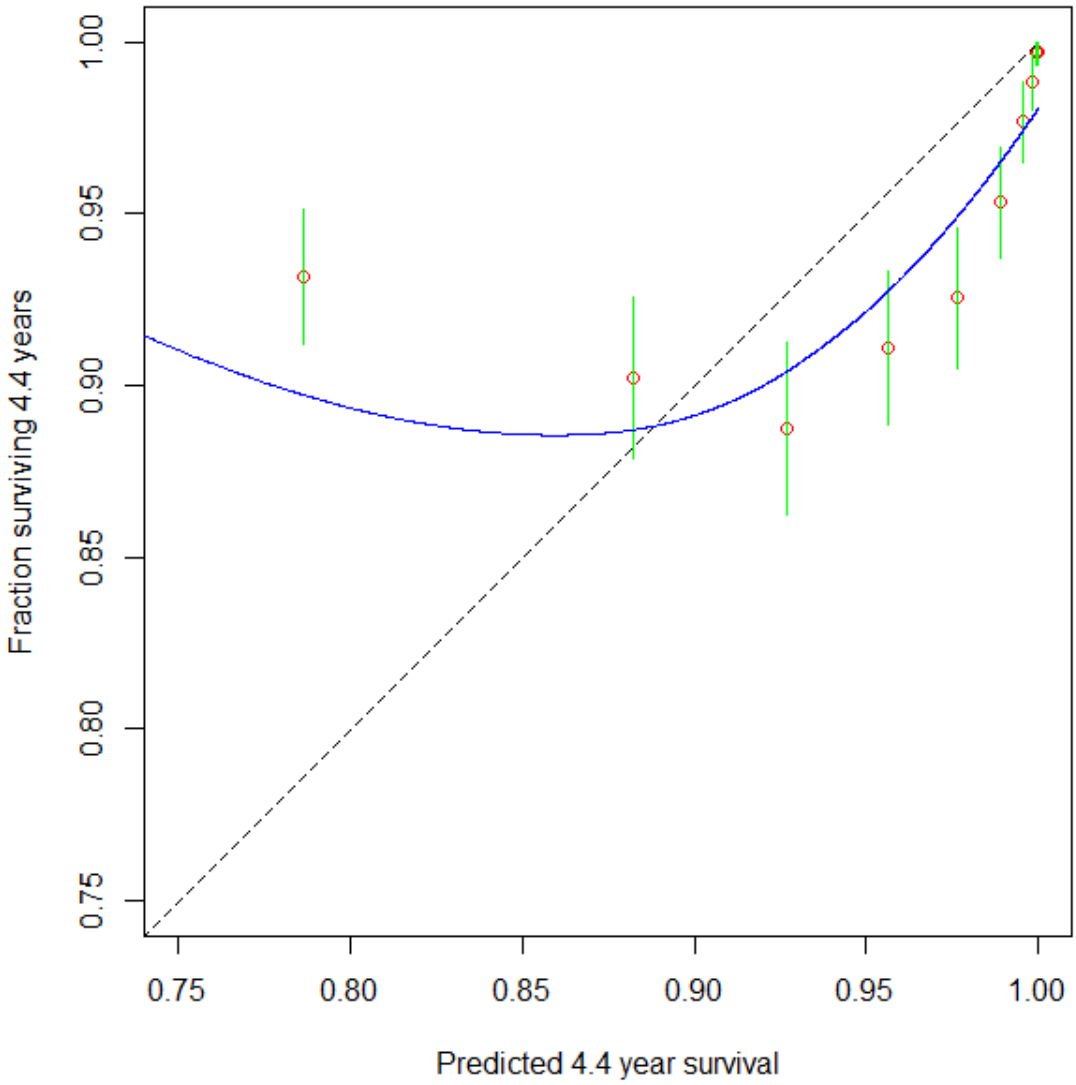

*Supplementary Figure 8.4 Linearity check for Systolic Blood Pressure (SBP). A – Plot of systolic blood pressure grouped by decile against the proportion with a positive outcome B – Martingale residual plot for systolic blood pressure.*



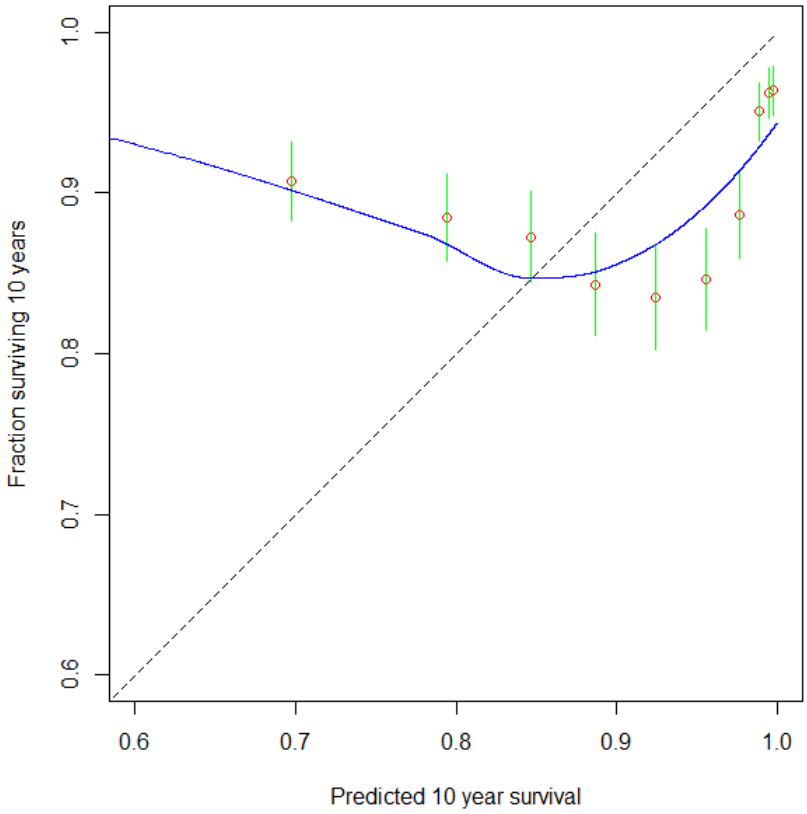

*Supplementary Figure 8.5 Colinear plot for systolic blood pressure*



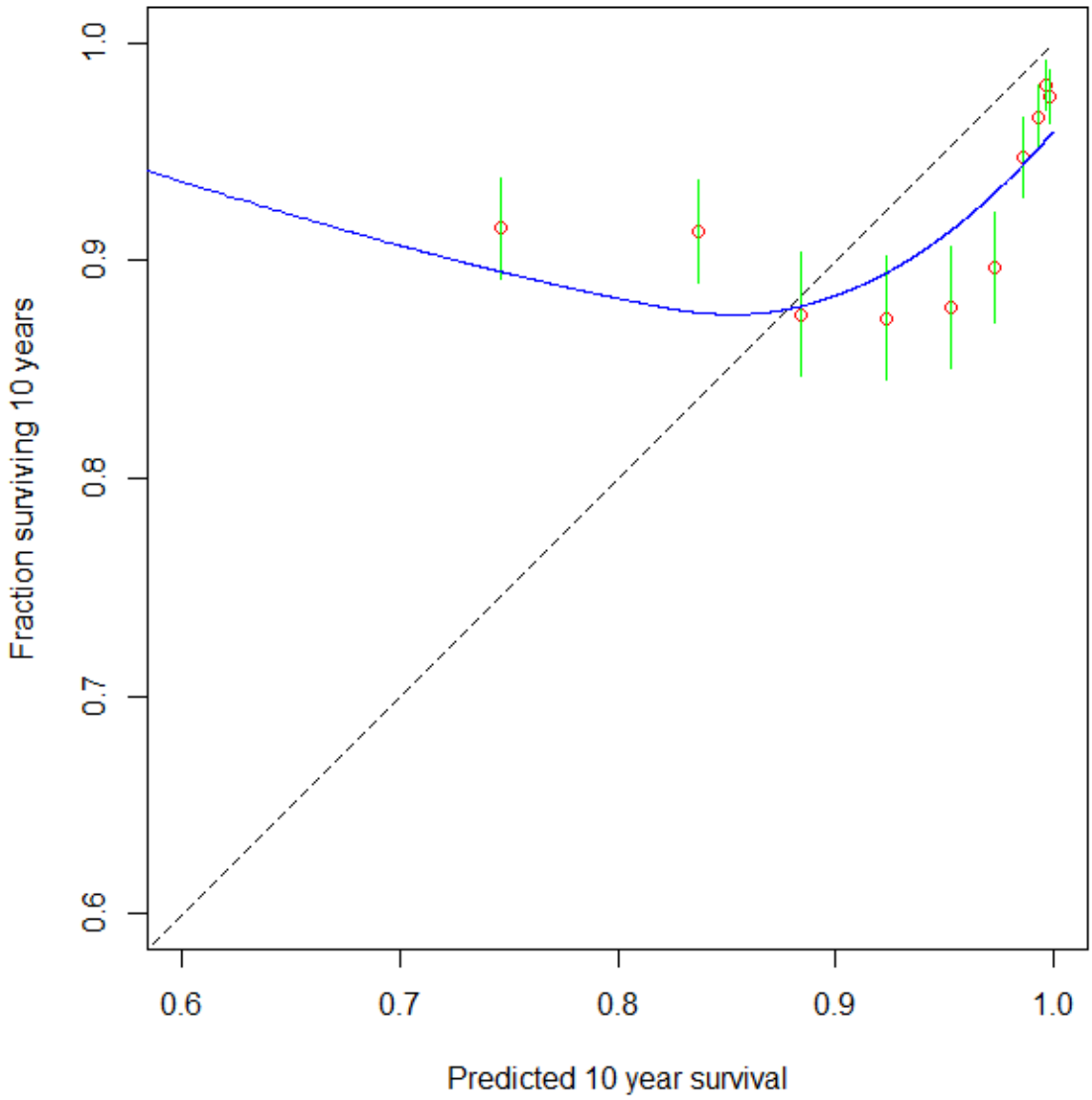

*Supplementary Figure 8.6 calibration plot of the multivariable cox model from cohort 1 selected by AIC statistic including conventional cardiac troponin T as a continuous variable.*




*Supplementary Figure 8.7 Calibration plot of the multivariable cox model from cohort 2 selected by AIC statistic including TMACS as a continuous variable*




*Supplementary Figure 8.8 Calibration plot of the multivariable cox model from cohort 2 selected by AIC statistic including high sensitivity cardiac troponin T as a continuous variable.*


Fraction surviving 10 years
1.0
0.9
0.8
0.7
0.6
0.6
0.7
0.8
0.9
1.0
Predicted 10 year survival

## 8.2 Supplementary tables

| Author, Country, Year | Domain 1 | Domain 2 | Domain 3 | Domain 4 | Domain 5 | Domain 6 | Overall |
|---|---|---|---|---|---|---|---|
| Hng, Australia 2016 (146) | Low | Low | Low | Moderate | Low | Low | Moderate |
| Jelinek, Australia, 2010 (145) | Low | Moderate | Low | Low | Moderate | Low | Moderate |
| Charfen, USA, 2009 (142) | Low | Moderate | Low | Low | Moderate | Low | Moderate |
| Hewat, UK, 2009 (144) | Low | Moderate | Low | Low | Moderate | Low | Moderate |
| George, UK, 2005 (141) | Low | Low | Low | Low | Moderate | Low | Moderate |
| Silverman, USA, 2006 (139) | Low | Low | Low | Low | Low | Low | Low |
| Ginde, USA, 2008 (143) | Low | Low | Low | Low | Moderate | Low | Moderate |
| Silverman, USA 2011 (140) | Low | High | Low | Low | Moderate | Moderate | High |
| Farkouh, USA, 2009 (147) | Moderate | Moderate | Low | Low | Low | Low | Moderate |
| Sanchis, USA, 2008 (148) | Low | Moderate | Moderate | Moderate | Low | Moderate | Moderate |

*Supplementary Table 8.1 - Risk of bias of studies examining type two diabetes mellitus. This was conducted using the QUIPs tool (25). Domain 1 – Study Participation, domain 2 – Study Attrition, domain 3 – Prognostic factor measurement, domain 4 – outcome measurement, domain 5– study confounding, domain 6 – Statistical analysis and reporting*

| Study | Domain 1 | Domain 2 | Domain 3 | Domain 4 | Domain 5 | Domain 6 | Overall |
|---|---|---|---|---|---|---|---|
| Elder, Australia, 2006 (149) | Low | Moderate | Low | NA | Low | Low | Moderate |
| Diercks, USA, 2002 (150) | Low | Low | Low | NA | Low | Low | Low |
| Chandra, USA, 2002 (151) | Low | Low | Low | NA | Low | Low | Low |
| Tan, Australia, 2013 (126) | Low | Low | Low | NA | Low | Low | Low |

*Supplementary Table 8.2 - Risk of bias of Dyslipidaemia studies. This was conducted using the QUIPs tool. Domain 1 – Study Participation, domain .22 – Study Attrition, domain 3 – Prognostic factor measurement, domain 4 – outcome measurement, domain 5 – study confounding, domain 6 – statistical analysis and reporting, N/A denotes not applicable*

| Study Author | Domain 1 | Domain 2 | Domain 3 | Domain 4 | Domain 5 | Domain 6 | Overall |
|---|---|---|---|---|---|---|---|
| Sanchis (148) | Low | Moderate | Moderate | Moderate | Low | Moderate | Moderate |
| Chaikriangkrai (152) | Low | High | Moderate | Low | Low | Moderate | High |

*Supplementary Table 8.3 Risk of bias of chronic kidney disease prognostication studies. This was conducted using the QUIPs tool (25). Domain 1 – Study Participation, domain 2 – Study Attrition, domain 3 – Prognostic factor measurement, domain 4 – outcome measurement, domain 5 – study confounding, domain 6 – Statistical analysis and reporting*

| | **Cohort 1** | **Cohort 2** | **Cohort 3** |
|---|---|---|---|
| Heart Rate (bpm) | | | |
| Mean | 84.1 | 84.0 | 83.0 |
| SD | 17.1 | 17.7 | 16.9 |
| Missing (n) | 19.1% (1154) | 12.1% (734) | 9.1% (625) |
| | | | |
| Systolic blood pressure (mmHg) | | | |
| Mean | 144.0 | 134.6 | 136.6 |
| SD | 18.9 | 19.5 | 22.0 |
| Missing (n) | 19.1% (1158) | 12.1% (737) | 9.1% (623) |
| | | | |
| Diastolic blood pressure (mmHg) | | | |
| Mean | 80.1 | 79.0 | 82.8 |
| SD | 14.1 | 13.4 | 14.8 |
| Missing (n) | 19.2% (1160) | 12.0% (733) | 9.2% (627) |
| | | | |
| Oxygen Saturations (%) | | | |
| Mean | 98.1 | 98.1 | 97.7 |
| SD | 2.0 | 1.9 | 2.2 |
| Missing (n) | 28.6% (1732) | 12.2% (740) | 9.1% (622) |
| | | | |
| Glasgow Coma Scale | | | |
| Median | 15 | 15 | 15 |
| Range | 12-15 | 15 | 15 |
| Missing (n) | 26.5% (1604) | 12.0% (732) | 9.1% (623) |
| | | | |
| Temperature (degrees Celsius) | | | |
| Mean | 36.4 | 36.5 | 36.5 |
| SD | 0.6 | 0.6 | 0.6 |
| Missing (n) | 19.3% (1170) | 12.2% (745) | 9.1% (623) |
| | | | |
| Index of Multiple Deprivation | | | |
| Mean | 41.9 | 41.2 | 41.1 |
| SD | 17.3 | 17.9 | 17.4 |
| Missing (n) | 57.3% (3470) | 64.3% (3915) | 71.0% (4836) |
| | | | |
| Haemoglobin (g/L) | | | |
| Mean | 13.6 | 13.7 | 13.6 |
| SD | 1.81 | 1.8 | 18.1 |
| Missing (n) | 50.44% (3054) | 30.1% (1834) | 27.8% (1900) |
| | | | |
| White Blood Cell Count ($10^9$/L) | | | |
| Mean | 8.4 | 8.2 | 8.5 |
| SD | 3.3 | 3.0 | 3.7 |
| Missing (n) | 50.2% (3042) | 29.8% (1815) | 27.6% (1886) |
| | | | |
| Estimated Glomerular Filtration Rate (ml/min/1.73$m^2$) | | | |
| Mean | 69.8 | 71.4 | 71.1 |
| SD | 17.7 | 15.9 | 16.0 |

| | | | |
|---|---|---|---|
| Missing (n) | 79.9% (4838) | 61.5% (3744) | 59.0% (4042) |
| | | | |
| Rural Urban Index | | | |
| Median | 5 | 5 | 5 |
| Range | 2-9 | 1-8 | 1-9 |
| Missing (n) | 57.1% (3459) | 64.3% (3912) | 70.6% (4832) |
| | | | |

*Supplementary Table 8.4 Additional demographic characteristics of cohorts*

| Variable | Skewness | Histogram/QQ Plot | Transformation | New skewness |
|---|---|---|---|---|
| Age | 0.44 | Right Skew | Nil | NA |
| Heart Rate | 0.87 | Right Skew | Logarithmic | -0.02 |
| SBP | 0.61 | Right Skew | Logarithmic | 0.09 |
| DBP | 0.15 | Nil | Nil | NA |
| Temperature | 0.44 | Right Skew | Nil | NA |
| Respiratory Rate | 2.92 | Right Skew | Logarithmic | -0.02 |
| EWS | 1.52 | Right Skew | Square root | 0.04 |
| $SaO_2$ | -4.38 | Left Skew | Exponential | 0.93 |
| WBC | 2.1 | Left Skew | Logarithmic | 0.20 |
| Haemoglobin | -0.47 | Right Skew | Nil | NA |
| IMD | -0.04 | Nil | Nil | NA |
| eGFR | -1.26 | Left Skew | Cube | -0.20 |
| cTnT | 18.67 | Right Skew | Logarithmic | 6.29 |
| hs-cTnT | 15.33 | Right Skew | Logarithmic | 1.74 |
| TMACS | 3.64 | Right Skew | Logarithmic | 1.21 |

*Supplementary Table 8.5 Variable transformation summary*

| | | Restricted Cubic Spline | | |
|---|---|---|---|---|
| Variable | Univariate | 3 knots | 4 knots | 5 knots |
| Age | | | | |
| AIC | 10371.04 | 10166.92 | 10160.77 | 10162.59 |
| BIC | 10375.48 | 10180.19 | 10174.09 | 10180.35 |
| Respiratory Rate | | | | |
| AIC | 8289.32 | 8290.83 | 8292.71 | 8292.71 |
| BIC | 8293.52 | 8299.23 | 8305.30 | 8305.30 |
| Temperature | | | | |
| AIC | 8267.96 | 8268.24 | 8270.11 | 8271.91 |
| BIC | 8272.16 | 8276.63 | 8282.69 | 8288.69 |
| WBC | | | | |
| AIC | 6212.85 | 6197.58 | 6198.75 | 6200.29 |
| BIC | 6216.82 | 6205.52 | 6210.65 | 6216.16 |
| Heart Rate | | | | |
| AIC | 8272.49 | 8273.72 | 8275.78 | 8277.53 |
| BIC | 8276.69 | 8282.11 | 8286.88 | 8294.32 |
| SBP | | | | |
| AIC | 8279.56 | 8272.95 | 8274.29 | 8274.44 |
| BIC | 8283.76 | 8286.95 | 8286.88 | 8291.22 |
| IMD | | | | |
| AIC | 9584.89 | 9584.10 | 9586.12 | 9587.69 |
| BIC | 9589.32 | 9595.81 | 9599.41 | 9605.41 |
| eGFR | | | | |
| AIC | 3124.64 | 3124.30 | 3126.17 | 3126.61 |
| BIC | 3128.04 | 3131.12 | 3136.39 | 3140.24 |
| cTnT | | | | |
| AIC | 2671.11 | 2671.10 | 2671.67 | 2673.21 |
| BIC | 2674.37 | 2677.63 | 2681.45 | 2686.26 |
| hs-cTnT | | | | |
| AIC | 4322.74 | 4330.90 | 4328.00 | 4314.96 |
| BIC | 4336.39 | 4338.20 | 4335.31 | 4325.92 |
| TMACS | | | | |
| AIC | 2377.77 | 2374.87 | 2355.46 | 1357.84 |
| BIC | 2380.89 | 2381.12 | 2364.83 | 2370.34 |

*Supplementary Table 8.6 Restricted Cubic Spline comparison.*

| | Age | RUI | IMD | eGFR | WBC | Hb | cTnT | Temp | EWS | GCS | RR | SaO2 | SBP | DBP | HR | P- time |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Cor** | -0.02 | -0.03 | 0.03 | 0.18 | 0.03 | 0.08 | -0.07 | 0.16 | -0.06 | 0.16 | -0.20 | 0.21 | 0.03 | -0.12 | -0.03 |
| Age | **(p-value)** | (0.43) | (0.16) | (0.40) | (<0.01) | (0.19) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (0.05) | (<0.01) | (0.03) |
| | -0.02 | **Cor** | -0.14 | 0.05 | 0.02 | 0.06 | 0.21 | 0.02 | 0.01 | 0.01 | -0.03 | -0.03 | 0.02 | 0.03 | -0.01 | 0.02 |
| RUI | (0.43) | **(p-value)** | (<0.01) | (0.21) | (0.27) | (0.01) | (<0.01) | (0.34) | (0.69) | (0.69) | (0.23) | (0.21) | (0.43) | (0.15) | (0.65) | (0.34) |
| | -0.03 | -0.14 | **Cor** | -0.04 | 0.04 | -0.01 | -0.07 | 0.04 | 0.05 | -0.01 | 0.08 | 0.01 | 0.03 | 0.00 | 0.09 | -0.02 |
| IMD | (0.16) | (<0.01) | **(p-value)** | (0.33) | (0.12) | (0.62) | (0.10) | (0.05) | (0.02) | (0.85) | (<0.01) | (0.61) | (0.23) | (0.86) | (<0.01) | (0.28) |
| | 0.03 | 0.05 | -0.04 | **Cor** | -0.13 | 0.62 | -0.18 | 0.01 | -0.16 | -0.01 | -0.21 | 0.07 | 0.02 | 0.15 | 0.02 | 0.02 |
| eGFR | (0.40) | (0.21) | (0.33) | **(p-value)** | (<0.01) | (<0.01) | (<0.01) | (0.73) | (<0.01) | (0.81) | (<0.01) | (0.03) | (0.57) | (<0.01) | (0.40) | (0.27) |
| | 0.18 | 0.02 | 0.04 | -0.13 | **Cor** | 0.04 | 0.07 | 0.15 | 0.16 | -0.02 | 0.18 | -0.08 | -0.05 | -0.07 | 0.18 | -0.06 |
| WBC | (<0.01) | (0.27) | (0.12) | (<0.01) | **(p-value)** | (0.04) | (0.02) | (<0.01) | (<0.01) | (0.37) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (0.04) |
| | 0.03 | 0.06 | -0.01 | 0.62 | 0.04 | **Cor** | -0.05 | -0.06 | -0.05 | 0.04 | -0.08 | -0.12 | 0.08 | 0.22 | 0.03 | -0.02 |
| Hb | (0.19) | (0.01) | (0.62) | (<0.01) | (0.04) | **(p-value)** | (0.13) | (<0.01) | (0.01) | (0.03) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (0.19) | (0.25) |
| | 0.08 | 0.21 | -0.07 | -0.18 | 0.07 | -0.05 | **Cor** | -0.06 | 0.04 | 0.02 | 0.02 | 0.12 | -0.02 | -0.03 | 0.05 | -0.06 |
| cTnT | (<0.01) | (<0.01) | (0.10) | (<0.01) | (0.02) | (0.13) | **(p-value)** | (0.08) | (0.19( | (0.55) | (0.50) | (<0.01) | (0.60) | (0.33) | (0.16) | (0.04) |
| | -0.07 | 0.02 | 0.04 | 0.01 | 0.15 | -0.06 | -0.06 | **Cor** | -0.18 | -0.00 | 0.08 | -0.05 | 0.03 | -0.02 | 0.24 | -0.06 |
| Temp | (<0.01) | (0.34) | (0.05) | (0.73) | (<0.01) | (<0.01) | (0.08) | **(p-value)** | (<0.01) | (0.76) | (<0.01) | (<0.01) | (0.04) | (0.21) | (<0.01) | (<0.01) |
| | 0.16 | 0.01 | 0.05 | -0.16 | 0.16 | -0.05 | 0.04 | -0.18 | **Cor** | -0.01 | 0.53 | -0.06 | 0.00 | 0.01 | 0.29 | -0.06 |
| EWS | (<0.01) | (0.69) | (0.02) | (<0.01) | (<0.01) | (0.01) | (0.19( | (<0.01) | **(p-value)** | (0.36) | (<0.01) | (<0.01) | (0.94) | (0.29) | (<0.01) | (<0.01) |
| | -0.06 | 0.01 | -0.01 | -0.01 | -0.02 | 0.04 | 0.02 | -0.00 | -0.01 | **Cor** | 0.03 | -0.06 | -0.01 | 0.01 | 0.03 | 0.01 |
| GCS | (<0.01) | (0.69) | (0.85) | (0.81) | (0.37) | (0.03) | (0.55) | (0.76) | (0.36) | **(p-value)** | (0.03) | (0.00) | (0.09) | (0.68) | (0.02) | (0.39) |
| | 0.16 | -0.03 | 0.08 | -0.21 | 0.18 | -0.08 | 0.02 | 0.08 | 0.53 | 0.03 | **Cor** | -0.02 | 0.02 | -0.01 | 0.17 | 0.00 |
| RR | (<0.01) | (0.23) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (0.50) | (<0.01) | (<0.01) | (0.03) | **(p-value)** | (0.21) | (0.09) | (0.56) | (<0.01) | (0.87) |
| | -0.20 | -0.03 | 0.01 | 0.07 | -0.08 | -0.12 | 0.12 | -0.05 | -0.06 | -0.06 | -0.02 | **Cor** | -0.04 | -0.01 | -0.04 | -0.02 |
| SaO2 | (<0.01) | (0.21) | (0.61) | (0.03) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | (0.00) | (0.21) | **(p-value)** | (<0.01) | (0.56) | (<0.01) | (0.29) |
| | 0.21 | 0.02 | 0.03 | 0.02 | -0.05 | 0.08 | -0.02 | 0.03 | 0.00 | -0.01 | 0.02 | -0.04 | **Cor** | 0.60 | 0.07 | 0.02 |
| SBP | (<0.01) | (0.43) | (0.23) | (0.57) | (<0.01) | (<0.01) | (0.60) | (0.04) | (0.94) | (0.09) | (0.09) | (<0.01) | **(p-value)** | (<0.01) | (<0.01) | (0.16) |
| | 0.03 | 0.03 | 0.00 | 0.15 | -0.07 | 0.22 | -0.03 | -0.02 | 0.01 | 0.01 | -0.01 | -0.03 | 0.60 | **Cor** | -0.04 | 0.01 |
| DBP | (0.05) | (0.15) | (0.86) | (<0.01) | (<0.01) | (<0.01) | (0.33) | (0.21) | (0.29) | (0.68) | (0.56) | (0.03) | (<0.01) | **(p-value)** | (<0.01) | (0.52) |
| | -0.12 | -0.01 | 0.09 | 0.02 | 0.18 | 0.03 | 0.05 | 0.24 | 0.29 | 0.03 | 0.17 | -0.04 | 0.07 | -0.04 | **Cor** | 0.02 |
| HR | (<0.01) | (0.65) | (<0.01) | (0.40) | (<0.01) | (0.19) | (0.16) | (<0.01) | (<0.01) | (0.02) | (<0.01) | (<0.01) | (<0.01) | (<0.01) | **(p-value)** | (0.16) |
| | -0.03 | 0.02 | -0.02 | -0.03 | 0.02 | -0.02 | -0.06 | -0.06 | -0.06 | 0.01 | 0.00 | -0.02 | 0.02 | 0.01 | 0.02 | **Cor** |
| P-time | (0.03) | (0.34) | (0.28) | (0.35) | (0.27) | (0.25) | (0.04) | (<0.01) | (<0.01) | (0.39) | (0.87) | (0.29) | (0.16) | (0.52) | (0.16) | **(p-value)** |

*Supplementary Table 8.7 Co-linearity check*

| | Adjusted Hazard Ratio (using all-inclusive model) | | |
|---|---|---|---|
| **Variable** | **Cohort 1** | **Cohort 2** | **Cohort 3** |
| Gender – Female | 0.56 (0.47 - 0.67)* | 0.52 (0.42 -0.64)* | 0.50 (0.39 - 0.63)* |
| Ethnicity | | | |
| Asian | 1.25 (1.03 – 1.51)* | 1.23 (1.00 – 1.51)* | 0.90 (0.71 - 1.15) |
| Black | 0.84 (0.62 – 1.14) | 0.51 (0.34 – 0.77)* | 0.61 (0.40 – 0.94)* |
| Mixed | 0.70 (0.43 – 1.14) | 0.74 (0.44 – 1.25) | 0.27 (0.12 – 0.60)* |
| Other | 0.64 (0.41 – 0.98)* | 0.72 (0.46 – 1.13) | 0.49 (0.23 -1.06) |
| WBC - summarised | All significant | All significant | Nonsignificant |
| Heart rate | 0.68 (0.45 – 1.00)† | 0.79 (0.50 – 1.25) | 0.51 (0.30 – 0.86)* |
| Time | | | |
| Morning - reference | - | - | - |
| Afternoon | 0.92 (0.74 – 1.14) | 0.89 (0.70 – 1.13) | 1.38 (1.02 - 1.87)* |
| Evening | 1.29 (1.00 – 1.65)* | 1.11 (0.84 – 1.46) | 1.58 (1.10 - 2.25)* |
| OOH | 1.20 (0.97 – 1.49) | 1.12 (0.87 – 1.46) | 1.86 (1.36 – 2.54)* |
| Haemoglobin | 1.04 (0.95 – 1.12) | 0.89 (0.84 – 0.94)* | 1.00 (0.99 – 1.00) |
| SBP – summarised | Nonsignificant | All significant | Nonsignificant |

*Supplementary Table 8.8 - Hazard ratios of variables that were statistically significant in at least one cohort (excluding conventional troponin T, high sensitivity cardiac troponin T and TMACS variables).*

| | AIC | $R^2$ | Concordance (c-statistic) |
|---|---|---|---|
| Cohort 1 | | | |
| Stepwise selection | 10292.50 | 0.14 | 0.79 (0.77 – 0.80) |
| All variables | 10308.86 | 0.14 | 0.79 (0.77 - 0.81) |
| | | | |
| Cohort 2 | | | |
| Stepwise selection | 7651.56 | 0.13 | 0.81 (0.80 - 0.83) |
| All variables | 7647.04 | 0.13 | 0.81 (0.79 - 0.82) |
| | | | |
| Cohort 3 | | | |
| Stepwise selection | 5806.08 | 0.14 | 0.84 (0.83 -0.86) |
| All variables | 5823.12 | 0.14 | 0.85 (0.83 – 0.86) |

*Supplementary Table 8.9 - Summary statistics for the underlying adjustment models by cohort year.*

| Risk factor | Data source | EM Staff | non-EM staff | Patient facing materials |
|---|---|---|---|---|
| Blood pressure | Patient record | Novel heart disease practitioner | GP | Leaflet |
| Smoking | Receptionist | EM nurse | Pharmacist | Video advert |
| Cholesterol | Triage nurse | EM clinician | Lifestyle guide | Email |
| Diabetes | Treating nurse | Ring fenced discharging EM clinician | Specialty clinic | Letter given at discharge |
| BMI | Clinician | | | Infographic |
| Lifestyle | Scavenged blood samples | | | Ambulatory BP monitor |
| Prediction model output | | | | Letter sent after visit |
| Chronic kidney disease | | | | Text messages |
| | | | | Website |
| | | | | Medication |

*Supplementary Table 8.10 Logic model of potential care pathway inputs*

| Professional facing materials | Discharge outcomes | Clinical responsibility | Care outcomes |
|---|---|---|---|
| Discharge summary | Self-referral | EM clinician | Patient awareness of risk factor |
| Bespoke letter | Referral to GP or primary care practitioner | GP | Non-pharmacological management of risk factor |
| Pure data transfer only | Referral to other specialty | Specialty | Pharmacological management of risk factor |
| Verbal handover | Educational material | Patient | Patient engagement in preventative services |
| Infographic | CVD prevention practitioners | | Clinician awareness of patient's risk factor |
| | Initiation of non-pharmacological intervention | | Reduction in cardiovascular risk |
| | Initiation of pharmacological intervention | | Holistic EM care |
| | Results streamed to GP | | |
| | Results streamed to another provider | | |
| | Referred to new EM clinic | | |
| | Counselling in EM | | |
| | Repeated discussion in EM | | |

*Supplementary Table 8.11 Logic model of potential care pathway outputs*

| **OPCS/ICD-10 code** | **Definition** (172,240) |
|---|---|
| I21 | “Acute Myocardial Infarction (AMI)” |
| I22 | “Subsequent Myocardial Infarction” |
| I23 | “Certain current complications following AMI“ |
| I46 | “Cardiac arrest“ |
| R96 | “Other sudden death, cause unknown“ |
| R99 | “Other ill-defined and unspecified causes of mortality“ |
| K40 | “Saphenous vein graft replacement of coronary artery“ |
| K41 | “Other autograft replacement of coronary artery“ |
| K42 | “Allograft replacement of coronary artery“ |
| K43 | “Prosthetic replacement of coronary artery“ |
| K44 | “Other replacement of coronary artery“ |
| K45 | “Connection of thoracic artery to coronary artery“ |
| K46 | “Other bypass of coronary artery“ |
| K47 | “Repair of coronary artery“ |
| K48 | “Other open operations on coronary artery“ |
| K49 | “Transluminal balloon angioplasty of coronary artery“ |
| K50 | “Other therapeutic transluminal operations on coronary artery“ |
| K63 | “Contrast radiology of heart“ (including coronary arteriography) |
| K75 | “Percutaneous transluminal balloon angioplasty and insertion of stent into coronary artery“ |

*Supplementary Table 8.12 OPCS/ICD-10 codes and definitions*

| **Day of index event** | |
|---|---|
| Sun | 1508 (11.4%) |
| Mon | 2201 (17.4%) |
| Tue | 2064 (15.6%) |
| Wed | 1864 (14.1%) |
| Thu | 2032 (15.4%) |
| Fri | 1924 (14.6%) |
| Sat | 1614 (12.2%) |
| | |
| **Month of index event** | |
| Jan | 1190 (9.0%) |
| Feb | 1174 (8.9%) |
| Mar | 1059 (8.0%) |
| Apr | 972 (7.4%) |
| May | 938 (7.1%) |
| Jun | 1001 (7.6%) |
| Jul | 1155 (8.7%) |
| Aug | 1259 (9.5%) |
| Sep | 1233 (9.3%) |
| Oct | 1272 (9.6%) |
| Nov | 887 (6.7%) |
| Dec | 1067 (8.1%) |

*Supplementary Table 8.13 TMACS entry day and month frequency*

| | |
|---|---|
| **Troponin values** | |
| **MRI** | |
| Mean | 16.3 |
| Range | 2-5270 |
| Standard Deviation | 16.35233 |
| Missing | 5 |
| | |
| **Troponin values** | |
| **Blackburn** | |
| Mean | 60.75 |
| Range | 2-11000 |
| Standard Deviation | 22.660 |
| | |
| **Burnley** | |
| Mean | 0.094 |
| Range | 0.009 – 23.860 |
| Standard Deviation | 0.882 |
| | |
| **Urea** | |
| Mean | 5.76 |
| Range | 0.70-75.10 |
| Standard Deviation | 3.14 |
| Missing | 2766 |
| | |
| **Creatinine** | |
| Mean | 88.51 |
| Range | 24-1545 |
| Standard Deviation | 66.31 |
| Missing | 2767 |
| | |
| **Estimated glomerular filtration rate** | |
| Mean | 84.00 |
| Range | 2.51-159.30 |
| Standard Deviation | 24.27 |
| Missing | 2793 |
| | |
| **Physiological Parameters** | |
| **Temperature** | |
| Mean | 36.42 |
| Range | 33.30-41.90 |
| Standard Deviation | 0.56 |
| Missing | 4065 |
| | |
| **Respiratory rate** | |
| Mean | 17.16 |
| Range | 6 – 126 |
| Standard Deviation | 2.80 |
| Missing | 4065 |
| | |
| **Heart Rate** | |

| | |
|---|---|
| Mean | 79.58 |
| Range | 10.00-206.00 |
| Standard Deviation | 16.50 |
| Missing | 4065 |
| | |
| **Oxygen Saturations** | |
| Mean | 97.48 |
| Range | 70.00-100.00 |
| Standard Deviation | 1.96 |
| Missing | 4065 |
| | |
| **Glasgow Coma Scale** | |
| Median | 15 |
| Missing | 4065 |
| | |
| **Alert-Voice-Pain-Unresponsive scale** | |
| Mode | Alert |
| Missing | 4065 |

*Supplementary Table 8.14 – Additional TMACS cohort demographics*

| **Variable** | **Co-efficient** | **P-value** |
|---|---|---|
| Intercept | $-2.419e^{01}$ | 1.81e-15 |
| Age Spline term 1 | $3.440e^{-01}$ | 5.72e-06 |
| Age Spline term 2 | $-6.797e^{-01}$ | 0.00354 |
| Age Spline term 3 | 1.778 | 0.07158 |
| Age Spline term 4 | -1.100 | 0.38034 |
| Kidney function | $3.025e^{-04}$ | $< 2e^{-16}$ |
| Gender - Female | $-3.556e^{-01}$ | 0.02988 |
| Ethnicity - Asian | $9.714e^{-01}$ | $1.09e^{-08}$ |
| Ethnicity – Black | -1.859 | $2.38e^{-05}$ |
| Ethnicity – Mixed | $5.595e^{-01}$ | 0.08760 |
| Ethnicity - Other | -2.143 | $6.95e^{-05}$ |

*Supplementary Table 8.15 Extended TMACS model variable coefficients and their respective p-values.*

| | Extension | | |
|---|---|---|---|
| | **Derivation** | **Adjusted for optimism** | **Status Quo** |
| **AUC** | | | |
| Age Q1 | 0.882 (0.823 – 0.941) | 0.878 | 0.92 (0.88-0.95) |
| Age Q2 | 0.852 (0.816 – 0.889) | 0.852 | 0.89 (0.86-0.91) |
| Age Q3 | 0.815 (0.778 – 0.852) | 0.814 | 0.86 (0.83-0.89) |
| Age Q4 | 0.795 (0.758 -0.832) | 0.793 | 0.85 (0.82-0.88) |
| **CITL** | | | |
| Age Q1 | 0.603 (0.087 – 1.120) | 0.652 | -4.71 (-5.42 - -3.99) |
| Age Q2 | -0.187 (-0.480 – 0.105) | -0.180 | -3.35 (-3.73 - -2.96) |
| Age Q3 | 0.301 (-0.233 – 0.262) | 0.012 | -3.21 (-3.51 - -2.92) |
| Age Q4 | 0.015 (-0.272 – 0.301) | -0.010 | -4.83 (-5.19 - -4.48) |

*Supplementary Table 8.16 Extended TMACS summary statistics by age quartile. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Extension | | |
|---|---|---|---|
| | **Derivation** | **Adjusted for optimism** | **Status Quo** |
| **AUC** | | | |
| IMD Q1 | 0.809 (0.766 – 0.852) | 0.810 | 0.89 (0.86-0.92) |
| IMD Q2 | 0.817 (0.772 – 0.862) | 0.814 | 0.85 (0.81-0.88) |
| IMD Q3 | 0.838 (0.788 – 0.887) | 0.838 | 0.89 (0.86-0.92) |
| IMD Q4 | 0.846 (0.804 – 0.888) | 0.844 | 0.87 (0.83-0.91) |
| IMD Q5 | 0.767 (0.713 – 0.820) | 0.767 | 0.84 (0.80-0.88) |
| **CITL** | | | |
| IMD Q1 | 0.740 (0.438 – 1.042) | 0.741 | -3.44 (-3.85--3.04) |
| IMD Q2 | -0.101 (-0.453 – 0.252) | -0.106 | -3.94 (-4.38--3.50) |
| IMD Q3 | 0.153 (-0.197 – 0.504) | 0.126 | -2.97 (-3.36--2.58) |
| IMD Q4 | -0.033 (-0.364 – 0.340) | -0.019 | -3.18 (-3.570--2.80) |
| IMD Q5 | 0.548 (0.227 – 0.851) | 0.548 | -3.30 (-3.70--2.91) |

*Supplementary Table 8.17 Extended TMACS summary statistics by index of multiple deprivation quintile (IMD). AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Extension | | |
|---|---|---|---|
| | **Derivation** | **Adjusted for optimism** | **Status Quo** |
| **AUC** | | | |
| Stage 1 | 0.905 (0.878 – 0.931) | 0.903 | 0.91 (0.88-0.93) |
| Stage 2 | 0.848 (0.815 – 0.881) | 0.848 | 0.87 (0.84-0.90) |
| Stage 3 | 0.812 (0.759 -0.866) | 0.811 | 0.81 (0.75-0.86) |
| Stage 4 | 0.868 (0.763 – 0.974) | 0.877 | 0.89 (0.80-0.98) |
| Stage 5 | 0.826 (0.657 -0.996) | 0.823 | 0.78 (0.54-1.00) |
| **CITL** | | | |
| Stage 1 | -0.399 (-0.657 - -0.118) | -0.378 | -2.08 (-2.36 - -1.80) |
| Stage 2 | 0.682 (0.440 -0.924) | 0.679 | -2.32 (-2.57 - -2.08) |
| Stage 3 | 0.724 (0.310 – 01139) | 0.763 | -3.63 (-4.05 - -3.20) |
| Stage 4 | -2.523(-4.141 - -0.910) | -2.488 | -7.08 (-8.59 - -5.56) |
| Stage 5 | -10.355 (-12.968 - -8.396) | -10.214 | -14.96 (-17.28 - -12.65) |

*Supplementary Table 8.18 Extended TMACS summary statistics by kidney function categorised by chronic kidney disease stage. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Extension | | |
|---|---|---|---|
| | **Derivation** | **Adjusted for optimism** | **Status Quo** |
| **AUC** | | | |
| Asian | 0.832 (0.796 -0.869) | 0.832 | 0.87 (0.84 – 0.90) |
| Black | 0.887 (0.804 -0.970) | 0.886 | 0.90 (0.83 – 0.97) |
| Mixed | 0.830 (0.744 – 0.917) | 0.830 | 0.92 (0.87 – 0.97) |
| Other | 0.855 (0.757 – 0.954) | 0.855 | 0.92 (0.86 -0.97) |
| White | 0.810 (0.781 – 0.839) | 0.809 | 0.87 (0.85 – 0.89) |
| **CITL** | | | |
| Asian | 0.047 (-0.205 – 0.300) | 0.041 | -2.52 (-2.80 - -2.24) |
| Black | 0.184 (-0.670 – 1.028) | 0.178 | -5.98 (-6.86 - -5.10) |
| Mixed | 0.050 (-0.603 – 0.623) | 0.044 | -3.71 (-4.60 - -2.81) |
| Other | 0.454 (-0.763 – 1.473) | 0.448 | -3.86 (-4.91 - -2.82) |
| White | 0.052 (-0.158 – 0.261) | 0.045 | -4.03 (-4.28 - -3.77) |

*Supplementary Table 8.19 Extended TMACS summary statistics by ethnicity. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| Quarter | CITL | CITL lower 95% CI | CITL upper 95% CI | AUC | AUC lower CI | AUC upper CI |
|---|---|---|---|---|---|---|
| 1 | -2.15956 | -3.44933 | -0.8698 | 0.925358 | 0.851837 | 0.998879 |
| 2 | -2.02779 | -2.6831 | -1.37248 | 0.931566 | 0.898485 | 0.964647 |
| 3 | -3.78715 | -4.94238 | -2.63192 | 0.945646 | 0.918345 | 0.972947 |
| 4 | -0.9672 | -1.53247 | -0.40193 | 0.897424 | 0.845633 | 0.949216 |
| 5 | -2.27358 | -3.06929 | -1.47788 | 0.92056 | 0.868488 | 0.972633 |
| 6 | -1.06512 | -1.74511 | -0.38513 | 0.870215 | 0.797435 | 0.942994 |
| 7 | -2.7947 | -3.94655 | -1.64285 | 0.920052 | 0.841137 | 0.998966 |
| 8 | -1.08672 | -1.89031 | -0.28313 | 0.887939 | 0.804551 | 0.971328 |
| 9 | -1.46692 | -2.37315 | -0.56069 | 0.813458 | 0.717758 | 0.909159 |
| 10 | -0.80734 | -1.52014 | -0.09453 | 0.778386 | 0.665618 | 0.891153 |
| 11 | -1.83748 | -2.5205 | -1.15447 | 0.891795 | 0.841768 | 0.941822 |
| 12 | -3.4177 | -4.45226 | -2.38314 | 0.913092 | 0.857295 | 0.968888 |
| 13 | -1.62887 | -2.47441 | -0.78333 | 0.846833 | 0.760009 | 0.933656 |
| 14 | -0.53342 | -1.12628 | 0.059438 | 0.895466 | 0.855758 | 0.935173 |
| 15 | -0.38555 | -1.13519 | 0.364098 | 0.897611 | 0.83635 | 0.958871 |
| 16 | -0.95971 | -1.58094 | -0.33849 | 0.873165 | 0.826689 | 0.919642 |
| 17 | -0.14914 | -0.78774 | 0.489457 | 0.825706 | 0.738871 | 0.91254 |
| 18 | -1.19835 | -2.01506 | -0.38164 | 0.891063 | 0.826987 | 0.955138 |
| 19 | 0.142026 | -1.0538 | 1.337848 | 0.931217 | 0.87047 | 0.991964 |

*Supplementary Table 8.20 Model descriptive statistics by quarter for dynamically updated TMACS. CITL- calibration in the large, 95% CI – 95% confidence interval, AUC- area under the curve*

| | **Dynamic Updated TMACS** | **Status Quo** |
|---|---|---|
| **AUC** | | |
| Age Q1 | 0.89 (0.83 - 0.95) | 0.92 (0.88-0.95) |
| Age Q2 | 0.88 (0.85 - 0.92) | 0.89 (0.86-0.91) |
| Age Q3 | 0.85 (0.82 - 0.89) | 0.86 (0.83-0.89) |
| Age Q4 | 0.84 (0.80 - 0.87) | 0.85 (0.82-0.88) |
| **CITL** | | |
| Age Q1 | -1.14 (-1.67 - -0.6) | -4.71 (-5.42 - -3.99) |
| Age Q2 | -0.93 (-1.29 - -0.56) | -3.35 (-3.73 - -2.96) |
| Age Q3 | -0.81 (-1.09 - -0.53) | -3.21 (-3.51 - -2.92) |
| Age Q4 | -2.43 (-2.77 - -2.1) | -4.83 (-5.19 - -4.48) |

*Supplementary Table 8.21 Dynamically updated TMACS summary statistics by age quartile (Manchester and Blackburn). AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Dynamic Updated TMACS | Status Quo |
|---|---|---|
| **AUC** | | |
| IMD Q1 | 0.87 (0.83 - 0.91) | 0.89 (0.86-0.92) |
| IMD Q2 | 0.84 (0.8 - 0.88) | 0.85 (0.81-0.88) |
| IMD Q3 | 0.89 (0.85 - 0.92) | 0.89 (0.86-0.92) |
| IMD Q4 | 0.88 (0.83 - 0.92) | 0.87 (0.83-0.91) |
| IMD Q5 | 0.81 (0.76 - 0.86) | 0.84 (0.80-0.88) |
| **CITL** | | |
| IMD Q1 | -0.90 (-1.26 - -0.54) | -3.44 (-3.85--3.04) |
| IMD Q2 | -1.42 (-1.81 - -1.03) | -3.94 (-4.38--3.50) |
| IMD Q3 | -1.13 (-1.54 - -0.72) | -2.97 (-3.36--2.58) |
| IMD Q4 | -1.06 (-1.47 - -0.66) | -3.18 (-3.570--2.80) |
| IMD Q5 | -0.89 (-1.26 - -0.52) | -3.30 (-3.70--2.91) |

*Supplementary Table 8.22 Dynamically updated TMACS summary statistics by index of multiple deprivation quintile (IMD) – Manchester and Blackburn only. AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Dynamic Updated TMACS | Status Quo |
|---|---|---|
| **AUC** | | |
| Male | 0.85 (0.83 - 0.88) | 0.87 (0.85-0.89) |
| 1Female | 0.88 (0.85 - 0.91) | 0.89 (0.87-0.92) |
| **CITL** | | |
| Male | -1.11 (-1.33 - -0.89) | -3.58 (-3.81--3.35) |
| Female | -1.22 (-1.51 - -0.93) | -4.25 (-4.58--3.92) |

*Supplementary Table 8.23 Dynamically updated TMACS summary statistics by gender (Manchester and Blackburn). AUC - area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | Dynamic Updated TMACS | Status Quo |
|---|---|---|
| **AUC** | | |
| Asian | 0.87 (0.83 - 0.9) | 0.87 (0.84 – 0.90) |
| Black | 0.9 (0.84 - 0.97) | 0.90 (0.83 – 0.97) |
| Mixed | 0.87 (0.8 - 0.95) | 0.92 (0.87 – 0.97) |
| Other | 0.91 (0.84 - 0.97) | 0.92 (0.86 -0.97) |
| White | 0.86 (0.83 - 0.88) | 0.87 (0.85 – 0.89) |
| **CITL** | | |
| Asian | -0.3 (-0.57 - -0.03) | -2.52 (-2.80 - -2.24) |
| Black | -3.66 (-4.52 - -2.79) | -5.98 (-6.86 - -5.10) |
| Mixed | -0.87 (-1.61 - -0.13) | -3.71 (-4.60 - -2.81) |
| Other | -1.81 (-2.91 - -0.7) | -3.86 (-4.91 - -2.82) |
| White | -1.72 (-1.97 - -1.47) | -4.03 (-4.28 - -3.77) |

*Supplementary Table 8.24 Dynamically updated TMACS summary statistics by ethnicity (Manchester and Blackburn). AUC - area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

| | **Dynamic Updated TMACS** | **Status Quo** |
|---|---|---|
| **AUC** | | |
| Stage 1 | 0.89 (0.86 - 0.92) | 0.91 (0.88-0.93) |
| Stage 2 | 0.85 (0.82 - 0.88) | 0.87 (0.84-0.90) |
| Stage 3 | 0.81 (0.75 - 0.87) | 0.81 (0.75-0.86) |
| Stage 4 | 0.85 (0.72 - 0.99) | 0.89 (0.80-0.98) |
| Stage 5 | 0.77 (0.57 - 0.98) | 0.78 (0.54-1.00) |
| **CITL** | | |
| Stage 1 | -0.29 (-0.56 - -0.03) | -2.08 (-2.36 - -1.80) |
| Stage 2 | -0.50 (-0.73 - -0.27) | -2.32 (-2.57 - -2.08) |
| Stage 3 | -1.72 (-2.11 - -1.32) | -3.63 (-4.05 - -3.20) |
| Stage 4 | -5.53 (-7.13 - -3.93) | -7.08 (-8.59 - -5.56) |
| Stage 5 | -13.26 (-15.49 - -11.03) | -14.96 (-17.28 - -12.65) |

*Supplementary Table 8.25 Dynamically updated TMACS summary statistics by kidney function categorised by chronic kidney disease stage (Manchester and Blackburn). AUC -area under the curve, CITL – calibration in the large, 95% CI – 95% confidence interval.*

## 8.3 Supplementary notes

### Chapter 2

*Medline Search strategy*

Population

1. exp Emergency Service, Hospital/ or exp Emergency Medical Services/ or emergency service.mp.

2. emergency medicine.mp. or exp Emergency Medicine/

3. acute care.mp.

4. (accident and emergency).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

5. A&E.mp.

6. 1 OR 2 OR 3 OR 4 OR 5

Outcome

1. cardiovascular diseases.mp. or exp Cardiovascular Diseases/

2. exp DIAGNOSTIC TECHNIQUES, CARDIOVASCULAR/ or exp MODELS, CARDIOVASCULAR/ or cardiovascular.mp.

3. acute coronary syndromes.mp. or exp Acute Coronary Syndrome/

4. exp Myocardial Infarction/ or acute myocardial infarction.mp.

5. non-st elevation myocardial infarction.mp. or exp Non-ST Elevated Myocardial Infarction/

6. exp Angina, Unstable/ or st elevation myocardial infarction.mp. or exp ST Elevation Myocardial Infarction/

7. exp Myocardial Revascularization/ or exp Coronary Artery Disease/ or exp Coronary Disease/ or coronary revascularisation.mp. or exp Coronary Artery Bypass/

8. death.mp. or exp DEATH/ or exp DEATH, SUDDEN, CARDIAC/ or exp DEATH, SUDDEN/

9. stroke.mp. or STROKE, LACUNAR/ or exp STROKE/

10. cerebral vascular accident.mp.

11. cerebrovascular accident.mp. or Stroke/

12. Cerebrovascular disease.mp. or exp Cerebrovascular Disorders/

13. CVA.mp.

14. exp Ischemic Attack, Transient/ or TIA.mp.

15. Brain Ischemia/ or Transient ischaemic attack.mp.

16. 1 OR 2 OR 3 OR 4 OR 5 .... OR 15

Prognostic Factors

**a) Blood pressure**

1. systolic blood pressure.mp.

2. exp HYPERTENSION, RENAL/ or exp ESSENTIAL HYPERTENSION/ or exp HYPERTENSION, MALIGNANT/ or exp HYPERTENSION/ or exp HYPERTENSION, RENOVASCULAR/ or hypertension.mp. or WHITE COAT HYPERTENSION/ or exp Blood Pressure/

3. hyperten*.mp.

4. 1 OR 2 OR 3

**b) Dyslipidaemia**

1. exp LIPIDS/ or lipid.mp.

2. exp Dyslipidemias/ or dyslipidaemia.mp. or Hyperlipidemias/

3. exp Hypercholesterolemia/ or hypercholestraemia.mp.

4. exp Metabolic Syndrome/ or exp Triglycerides/ or exp Hypertriglyceridemia/ or hypertriglyceridaemia.mp. or exp Dyslipidemias/ or exp Hyperlipidemias/

5. 1 OR 2 OR 3 OR 4

**c) Diabetes mellitus**

1. exp DIABETES MELLITUS, EXPERIMENTAL/ or exp DIABETES MELLITUS, TYPE 2/ or diabetes.mp. or exp DIABETES COMPLICATIONS/ or exp DIABETES MELLITUS, TYPE 1/ or exp LATENT AUTOIMMUNE DIABETES IN ADULTS/ or exp DIABETES MELLITUS/

2. exp Diabetes Mellitus, Type 2/ or exp Diabetes Mellitus, Type 1/ or T1DM.mp.

3. T2DM.mp. or exp Insulin Resistance/

1 OR 2 OR 3

**d) Chronic kidney disease**

1. chronic kidney disease.mp. or exp Renal Insufficiency, Chronic/

2. exp Kidney Failure, Chronic/ or exp Renal Insufficiency, Chronic/ or CKD.mp. or exp Glomerular Filtration Rate/ or exp Kidney Diseases/

3. exp Renal Dialysis/ or exp Kidney/ or renal disease.mp.

4. 1 OR 2 OR 3

Limited to:

limit 20 to (English language and humans and all adults (19 plus years) and (adaptive clinical trial or clinical study or clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or comparative study or controlled clinical trial or equivalence trial or evaluation studies or meta analysis or multicenter study or observational study or pragmatic clinical trial or randomized controlled trial or systematic reviews or validation studies) and last 20 years)

*Embase Search strategy*

Population

1. emergency service.mp. or exp emergency health service/

2. emergency department.mp. or exp emergency ward/

3. emergency medicine.mp. or exp emergency medicine/

4. (accident and emergency).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]

5. ER.mp. or exp emergency care/

1 OR 2 OR 3 OR 4 OR 5 OR 6

Outcome

1. cardiovascular disease.mp. or exp cardiovascular disease/

2. acute coronary syndromes.mp. or exp acute coronary syndrome/

3. myocardial infarction.mp. or exp heart infarction/

4. acute myocardial infarction.mp. or exp acute heart infarction/

5. non-st elevation myocardial infarction.mp. or exp non ST segment elevation myocardial infarction/

6. NSTEMI.mp.

7. unstable angina.mp. or exp unstable angina pectoris/

8. st elevation myocardial infarction.mp. or exp ST segment elevation myocardial infarction/

9. STEMI.mp.

10. myocardial revascularisation.mp. or exp heart muscle revascularization/

11. coronary artery disease.mp. or exp coronary artery disease/

12. coronary disease.mp.

13. coronary revascularisation.mp.

14. coronary artery bypass.mp. or exp coronary artery bypass graft/

15. exp death/ or exp sudden cardiac death/ or death.mp. or exp "cause of death"/ or exp sudden death/ or brain death/ or exp heart death/

16. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15

Prognostic Factors

a) **blood pressure**

1. exp essential hypertension/ or exp hypertension/ or hyperten*.mp.

2. exp malignant hypertension/ or exp masked hypertension/ or exp essential hypertension/ or exp orthostatic hypertension/ or exp diabetic hypertension/ or exp resistant hypertension/ or exp hypertension/ or exp systolic hypertension/ or exp borderline hypertension/ or hypertension.mp. or exp renovascular hypertension/ or exp hereditary hypertension/ or exp white coat hypertension/

3. blood pressure.mp. or exp blood pressure/

4. 1 OR 2 OR 3

**b) lipid profile**

1. exp lipid storage/ or lipid.mp. or exp lipid metabolism/ or exp lipid/ or exp lipid blood level/ or exp lipid homeostasis/ or exp lipid liver level/

2. dyslipidaemia.mp. or exp dyslipidemia/

3. exp hyperlipidemia/ or hyperlipidemia.mp.

4. exp hypercholesterolemia/

5. metabolic syndrome.mp. or exp metabolic syndrome X/

6. Hypertriglyceridemia.mp. or exp hypertriglyceridemia/

1 OR 2 OR 3 OR 4 OR 5 OR 6

**c) diabetes mellitus**

1. diabetes mellitus.mp. or exp diabetes mellitus/

2. type 2 diabetes mellitus.mp. or exp non insulin dependent diabetes mellitus/

3. T2DM.mp.

4. insulin resistance.mp. or exp insulin resistance/

5. 1 OR 2 OR 3 OR 4 OR 5

**d) Chronic kidney disease**

1. chronic kidney disease.mp. or exp chronic kidney failure/

2. renal insufficiency.mp. or exp kidney failure/

3. exp chronic kidney failure/ or kidney/

4. exp chronic kidney disease/ or exp chronic kidney failure/ or CKD.mp.

5. renal disease.mp.

Limits

29. limit 28 to (human and english language and (evidence based medicine or consensus development or meta analysis or outcomes research or "systematic review") and yr="2000 - 2020"

and (article or article in press or conference abstract or conference paper or "conference review" or "review"))

## Chapter 3

### *Variable Transformation – additional detail*

The aforementioned variables were carried forward into the study. Their distribution and skew was checked by the Joanes and Gill method (181). The distribution was then visually inspected with histograms and quantile – quantile plots. Any variable with a skewness of > 0.5 or <-0.5 considered for transformation along with those with a non-normal distribution. For those variables selected, multiple transformations were trialled, and the final was selected based on the skewness metric. Based on this method 10/15 continuous variables underwent transformation, the most common being logarithmic (Supplementary Table 8.5). The resultant transformation was then checked again in a quantile-quantile plot to visually assess the improvement and a new skewness metric was calculated (Supplementary Figure 8.3). Presentation time was categorised into morning (0800 - 1159), afternoon (1200 – 1659), evening (1700 – 2059) and out of hours (2200 – 0759) to enable interpretation of the resultant function.

I then checked the variables for linearity with the outcome, using Martingale residual plots and outcome-predictor plots. RCS transformations were trialling different univariate cox proportional hazard models and comparing AIC and BIC. A univariate cox proportional hazard models were derived for the original predictor, the predictor transformed into a restricted cubic spline with 3,4 and 5 knots. Each model was assessed using AIC and BIC metrics, given the increasing sample size requirements for increasing number of knots a substantial increase in model performance was deemed to be needed to use another knot. Of the 11 predictors suspected of being non-linear, 6 were found to have better prognostic performance with an RCS leaving 5 untransformed.

I also examined for opportunities to conduct variable reduction. I found systolic blood pressure (SBP) and diastolic blood pressure (DBP) to be colinear by the study's definition. I opted for the exclusion of DBP on the grounds of data reduction given the magnitude of the collinearity and that SBP had fewer missing data. The candidate predictors included a clustered predictor already, the early warning score (EWS). This incorporated physiological predictors into one metric. As such, the physiological variables already included in the EWS were removed from the predictor pool and only physiological predictors with previously demonstrated long term CVD prognostic ability were retained (heart rate and SBP).

After the variables had been transformed, and variable clustering was considered collinearity was examined. For continuous variables Pearson's correlation coefficient was calculated and colinear plots were produced to provide a visual inspection (*Supplementary Table 8.7* & *Supplementary Figure 8.5*). Substantial collinearity was defined as a correlation co-efficient of > 0.2 or <-0.2 with a significant p-value, biological plausibility and confirmation on visual inspection. Collinearity was identified for seven pairs of variables. These were haemoglobin with eGFR, cTnT with RUI, hs-cTnT with age, hs-cTnT with eGFR, TMACS with age, and finally TMACS with eGFR. These colinear interaction terms were included to improve the performance of the adjustment multivariable models.

I then checked the variables for linearity with the outcome, using Martingale residual plots and outcome-predictor plots. Systolic blood pressure demonstrated a parabolic relationship with the outcome and the Martingale residual plot confirms this with a non-straight line (*Supplementary Figure 8.4*). In addition to SBP non-linearity was suspected in age, respiratory rate, temperature, white blood cell count, heart rate, IMD, eGFR, cTnT, hs-cTnT, and TMACs variables. Restricted cubic splines were assessed to identify if they yielded better prognostic performance. Univariate cox proportional hazard models were derived for the original predictor, the predictor transformed into a restricted cubic spline with 3,4 and 5 knots. Each model was assessed using AIC and BIC metrics, given the increasing sample size requirements for increasing number of knots a substantial increase in model

performance was deemed to be needed to use another knot. Of the 11 predictors suspected of being non-linear, 5 were found to have better prognostic performance without an RCS and were not transformed any further. Of the remaining 6 WBC, SBP and RR had the best prognostic performance when transformed in a 3 knot RCS. The predictor age had the best prognostic performance when in 4 knot RCS and hs-cTnT/TMACS had the best performance in a 5 knot RCS.

I also detected co-linearity between TMACS and hs-cTnT which met the study criteria, as TMACS already incorporated hs-cTnT within it I removed it in cohort 3 only. The interaction between the predictors age and SBP was found to have a correlation coefficient of -0.21, a p-value of <0.01, had biological plausibility and was confirmed on visual inspection (*Supplementary Figure 8.5*) This was also the case for haemoglobin with eGFR, cTnT with RUI, hs-cTnT with age, hs-cTnT with eGFR, TMACS with age, and finally TMACS with eGFR. These colinear interaction terms were included to improve the performance of the adjustment multivariable models.

### *Variation of hazard ratios across cohorts*

I considered the benefit of pooling all three cohorts to better understand the prognostic characteristics of the baseline variables. Prior to this I examined the variation in the hazard ratio per predictor overtime to see if the proportional hazards assumption was maintained between cohorts. I did not include the cardiac diagnostic variables as they were not available across cohorts.

Only female gender was found to be consistently statistically significant (*Supplementary Table 8.8*). Whilst the categorical variable ethnicity was consistently included by the stepwise selection mode no level was persistently statistically significant. Asian and Black ethnicities were significant across two consecutive cohorts. This may indicate a change in the association of ethnicity with cardiovascular disease over time between the three cohorts. If a multivariable model was to be derived pooling a wider time fame this variation should be considered however this is beyond the

scope of this thesis. 3.6.3 Adjustment models - Multivariable cox proportional hazard models – additional detail

I have described the models that enabled the adjustment of the hazard ratios. These are not primarily intended to be deployable clinical prediction models, instead they are presented to offer perspective for future research. A model derivation study could understand what might be expected from work building on this prognostic factor study. Across the multivariable models for each cohort the $R^2$ value was found to be 0.14, 0.13 and 0.14 (*Supplementary Table 8.9*). This was significantly below the $R^2$ in QRisk 3 used for the sample size calculation (19). If work was to be conducted in the future deriving a multivariable model for use as CPM, then these values should be used as the base case.

The AIC and concordance summary statistics of the multivariable models improve as the diagnostic technology innovated between cohorts from cTnT (0.79 95% CI 0.77 – 0.81) to TMACS (0.85 95% CI 0.83 – 0.86). These statistics and underlying models have not been validated, as this was not the primary purpose. As such these summary statistics should be interpreted with caution.

The calibration was assessed with calibration plots (Supplementary Figure 8.6, Supplementary Figure 8.8, and Supplementary Figure 8.7). The baseline cohort model underestimated the risk when the predicted survival was greater than 0.85 and overestimated when the predicted risk was greater than 0.85 (Supplementary Figure 8.6). The highest three deciles were markedly mis-calibrated.

The calibration of the hs-cTnT model (cohort 2), appeared improved with only the last two deciles being markedly mis-calibrated rather than 3 in cohort 1. The calibration of the TMACS model was again visually improved. Only one decile demonstrated significant risk underestimation. This must be taken in the context of a shorter follow up time for the TMACS cohort. These results would signal that the calibration as well as the discrimination of the multivariable models improved with the diagnostic innovations.

The code for the adjustment models is as follows:

Function of stepwise selected model

```
function(age_imp5 = 43,sex_imp5 = "Male",logwbc = 2.0412203,ethn_imp5 = "W
hite",logtrop_mi = -4.6051702,loghr = 4.4067192,time_imp5 = "Afternoon") {
age_imp5 <- as.numeric(age_imp5);-6.6319705+0.1514914* age_imp5-1.2403479e
-05*pmax(age_imp5-20,0)^3-5.1814036e-05*pmax(age_imp5-35,0)^3+0.0001091631
6*pmax(age_imp5-51,0)^3-4.4945646e-05*pmax(age_imp5-78,0)^3-0.56409371*(se
x_imp5=="Female")+0.98521565* logwbc-1.8002758*pmax(logwbc-1.6292405,0)^3+
3.4118622*pmax(logwbc-2.0412203,0)^3-1.6115864*pmax(logwbc-2.501436,0)^3+0
.22646909*(ethn_imp5=="Asian")-0.15278649*(ethn_imp5=="Black")-0.35802643*
(ethn_imp5=="Mixed")-0.45808773*(ethn_imp5=="Other")-0.19983302*logtrop_mi
-0.3840946*loghr-0.078043276*(time_imp5=="Afternoon")+0.25850546*(time_imp
5=="Evening")+0.18791949*(time_imp5=="OOH") }
```

Function of stepwise selected model with categorised cTnT

```
function(cat_oldtrop = "Missing",age_imp5 = 43,sex_imp5 = "Male",logwbc =
2.0412203,ethn_imp5 = "White",loghr = 4.4067192,time_imp5 = "Afternoon") {
age_imp5 <- as.numeric(age_imp5);-4.2798331-0.31208245*(cat_oldtrop=="Miss
ing")-0.55297394*(cat_oldtrop=="High")+0.13375357* age_imp5+2.7942204e-06*
pmax(age_imp5-20,0)^3-8.2492236e-05*pmax(age_imp5-35,0)^3+0.00012537412*pm
ax(age_imp5-51,0)^3-4.5676109e-05*pmax(age_imp5-78,0)^3-0.56454257*(sex_im
p5=="Female")+0.78967752* logwbc-1.491918*pmax(logwbc-1.6292405,0)^3+2.877
2212*pmax(logwbc-2.0412203,0)^3-1.3853032*pmax(logwbc-2.4849066,0)^3+0.227
6588*(ethn_imp5=="Asian")-0.1022863*(ethn_imp5=="Black")-0.27773461*(ethn_
imp5=="Mixed")-0.49729042*(ethn_imp5=="Other")-0.46303019*loghr-0.01938328
1*(time_imp5=="Afternoon")+0.19950745*(time_imp5=="Evening")+0.21757445*(t
ime_imp5=="OOH") }
<environment: 0x000000002771c248>
```

Function for all-inclusive model

```
function(hb_egfr = 1065.6,age_sbp = 1.33,tnt_rui = 0.05,logtrop_mi = -4.60
51702,cubegfr_mi = 493039,loghr = 4.4067192,logwbc = 2.0412203,sqrtews = 1
,age_imp5 = 43,logsbp = 4.8828019,imd_imp5 = 42.28,hb_imp5 = 13.8,sex_imp5
= "Male",ethn_imp5 = "White",rui_imp5 = 5,time_imp5 = "Afternoon") {age_im
p5 <- as.numeric(age_imp5);logsbp <- as.numeric(logsbp);-6.6430144-0.00065
700265*hb_egfr+0.01322783*age_sbp-0.43644671*tnt_rui-0.12402813*logtrop_mi
+5.7451681e-07*cubegfr_mi-0.39272516*loghr+0.99503274* logwbc-1.8293413*pm
ax(logwbc-1.6292405,0)^3+3.4669469*pmax(logwbc-2.0412203,0)^3-1.6376055*pm
ax(logwbc-2.501436,0)^3+0.0085741369*sqrtews+0.15016176* age_imp5-1.090364
1e-05*pmax(age_imp5-20,0)^3-5.5884899e-05*pmax(age_imp5-35,0)^3+0.00011242
451*pmax(age_imp5-51,0)^3-4.5635972e-05*pmax(age_imp5-78,0)^3+0.031450622*
logsbp-1.7777433*pmax(logsbp-4.7095302,0)^3+3.621653*pmax(logsbp-4.8828019
,0)^3-1.8439097*pmax(logsbp-5.049856,0)^3+0.0013172555*imd_imp5+0.03503185
2*hb_imp5-0.5748073*(sex_imp5=="Female")+0.22351636*(ethn_imp5=="Asian")-0
.17152245*(ethn_imp5=="Black")-0.3570141*(ethn_imp5=="Mixed")-0.45184814*(
ethn_imp5=="Other")+0.036779024*rui_imp5-0.085733948*(time_imp5=="Afternoo
n")+0.25210837*(time_imp5=="Evening")+0.17820774*(time_imp5=="OOH") }
<environment: 0x00000000380cd6f8>
```

# Chapter 4

## *Wave 1 Topic Guide*

Thank you for attending today and taking the time out of your schedule. The session ought not to last more than one hour.

Before starting I would just like to remind you to avoid identifying yourself whilst we record the audio of this session.

Today I would like to explore what we can do for patients who attend the emergency department with chest pain but have a heart attack excluded. We have been told by previous patient groups that they are often left feeling dissatisfied at this abrupt end of care, many are left wanting more information particularly about long term risk.

I would like to start with a fictional case; Imagine a patient has presented to the emergency department with chest pain. Serial blood tests and an ECG are conducted looking for evidence of a heart attack. After five hours all the results are returned, and a heart attack is excluded.

**Once the patient is told that this wasn't a heart attack, what more could be done for them?**

**What other information might be useful at that stage?**

**What about long term heart health (cardiovascular disease)?**

**A common risk factor for long term heart health is high blood pressure, if this patient had a high blood pressure what should we do for them in the ED?**

-High cholesterol - smoking, kidney disease - diabetes

**What would be the best way to communicate this?**

-leaflets -apps -emails – text alerts

**There are medications available to treat some of these risk factors, how could we guide patients towards them?**

**Who should we approach with this extra information?**

**Do you think it should be only given to patients with chest pain?**

### *Wave 2 Topic Guide*

Thank you for your time today, this will be about 30 minutes.

Please refrain from saying your name to protect your privacy. We can remove identifiable sections of the interview later.

Patients told us that they were interested in their long-term heart health when attending for acute chest pain.

We have interviewed stakeholders and formulated prototype care pathways that uses information from the Emergency Department to inform long term heart health.

First, I would like to get some feedback on the themes we found from first interviews:

- Who is responsible?
- Primary secondary care interface
- Acute condition diagnosis
- Reductive EM care
- Efficiency of preventative care in EM

There are three prototype care pathways (1) best pathway (2) slim line pathway (3) worst case pathway.

*Explain through pathway with print outs*

*Invite feedback against Peter's et al's framework*

Peters, David H., Taghreed Adam, Olakunle Alonge, Irene Akua Agyepong, and Nhan Tran. "Implementation research: what it is and how to do it." *Bmj* 347 (2013).

Chapter 5

*CPM data validation procedure*

Due to the presence of duplicate entries by clinicians in the MRI TMACS CPM database, the last entry within the admission was considered to the valid and final entry (prime entry). There is the possibility that the selected prime entry, whilst the last data inputted, was not the version used by the clinician, if so then the correct entry may be in the discarded duplicates. To assess this, we will seek to validate the data analysis, by examining the clinical co-variates (troponin, visibly sweating, systolic blood pressure <100mmhg etc).

## Aim

In patient episodes that have multiple TMACS entries we will seek to identify if an incorrect entry has been selected as the prime version.

We will seek to validate this by firstly detecting discrepancies between other sources of the clinical inputs and the TMACS database, and secondly attempting to identify if another TMACS entry matches those validated outputs.

## Outcome definition

Incorrect entry selection – when the current prime TMACS entry has incorrect clinical inputs and there is a correct TMACS entry present.

If all the TMACS entries are incorrect then this definition will not have been met, this will instead be a clinician error.

Clinician error – when there are no TMACS entries with valid clinical inputs

## Phase 1 - automated

This will be an automated phase using databases external to TMACS to validate the clinical inputs.

We will automatically proceed to phase 2

## Phase 2 – case note review

We will randomly select 100 patient episodes that have TMACS duplicates. These episodes will undergo case note review (as per the case note review process below), to validate which duplicate is the final clinical version.

The case note review will be blinded to the TMACS clinical data points.

If we detect an error rate of more than 5% we will review all duplicate cases (phase 3)

## Phase 3

We will review all case notes, as per the case note review flow chart

*Case note review process*

*n.b. we have assumed that in the absence of contradictory data the clinical input is a match*

(1) Does the troponin match from the automated review?

Yes – continue to question 2

No – check duplicates – if match restart process - if no match mark as clinician error

(2) Does the blood pressure match from the automated review?
Yes – continue to question 3

No – check duplicates – if match restart process - if no match mark as clinician error

(3) Is there a scanned ECG in the notes?
Yes – go to Q3a

No – got to Q3b

(3a) Does the scanned ECG match the prime TMACS clinical input for evidence of ischaemia on the ECG?

Yes – continue to question 4

No – check duplicates – if match restart process – if no match mark as clinician error

(3b) Does a written statement confirming or refuting the presence of ischaemia match the TMACS clinical inputs?

Yes – continue to question 4

No - check duplicates – if match restart process – if no match mark as clinician error

No statement – continue to question 4

(4) Is there a statement regarding the radiation of pain in the clinical notes?
Yes – go to Q4a
No – go to Q5

(4a) Does a written statement regarding the radiation of pain match the prime TMACS clinical input?
Yes – go to Q5

No - check duplicates – if match restart process – if no match mark as

clinician error

(5) Is there a statement regarding crescendo or worsening angina in the notes?
Yes – go to Q5a
No – got to Q6

(5a) Does the statement in the clinical notes regarding angina match the prime TMACS clinical input?
Yes – go to Q6
No - check duplicates – if match restart process – if no match mark as clinician error

(6) Is there a statement regarding sweating in the notes?
Yes – go to Q6a
No – go to Q7

(6a) Does the statement in the clinical notes regarding angina match the prime TMACS clinical input?
Yes – got to Q7
No - check duplicates – if match restart process – if no match mark as clinician error

(7) Is there a statement regarding vomiting in the notes
Yes - go to Q7a
No – go to Q8

(7a) Does the statement in the clinical notes regarding vomiting match the prime TMACS clinical input?
Yes – go to Q8
No - check duplicates – if match restart process – if no match mark as clinician error

(8) Is the now validated TMACS entry the original prime?
Yes – PASS
No – incorrect entry selection criteria met

# References


1. Hospital Admitted Patient Care Activity, 2016-17 - NHS Digital [Internet]. [cited 2017 Dec 11]. Available from: https://digital.nhs.uk/catalogue/PUB30098

2. Chest pain of recent onset: assessment and diagnosis | Guidance and guidelines | NICE [Internet]. [cited 2017 Dec 11]. Available from: https://www.nice.org.uk/guidance/cg95

3. Myocardial infarction (acute): Early rule out using high-sensitivity troponin tests (Elecsys Troponin T high-sensitive, ARCHITECT STAT High Sensitive Troponin-I and AccuTnI+3 assays) | Guidance and guidelines | NICE [Internet]. [cited 2017 Dec 11]. Available from: https://www.nice.org.uk/guidance/dg15

4. Chapman AR, Lee KK, McAllister DA, Cullen L, Greenslade JH, Parsonage W, et al. Association of High-Sensitivity Cardiac Troponin I Concentration With Cardiac Outcomes in Patients With Suspected Acute Coronary Syndrome. JAMA. 2017 Nov 21;318(19):1913–24.

5. Body R, Carlton E, Sperrin M, Lewis PS, Burrows G, Carley S, et al. Troponin-only Manchester Acute Coronary Syndromes (T-MACS) decision aid: single biomarker re-derivation and external validation in three cohorts. Emergency Medicine Journal. 2016 Aug 26;emermed-2016-205983.

6. Backus BE, Six AJ, Kelder JC, Bosschaert MAR, Mast EG, Mosterd A, et al. A prospective validation of the HEART score for chest pain patients at the emergency department. International Journal of Cardiology. 2013 Oct;168(3):2153–8.

7. Than MP, Pickering JW, Aldous SJ, Cullen L, Frampton CMA, Peacock WF, et al. Effectiveness of EDACS Versus ADAPT Accelerated Diagnostic Pathways for Chest Pain: A Pragmatic Randomized Controlled Trial Embedded Within Practice. Ann Emerg Med. 2016;68(1):93-102.e1.

8. Reynard C, Jenkins D, Martin GP, Kontopantelis E, Body R. Is your clinical prediction model past its sell by date? Emergency Medicine Journal. 2022;

9. Body R, Boachie C, McConnachie A, Carley S, Berg PVD, Lecky FE. Feasibility of the Manchester Acute Coronary Syndromes (MACS) decision rule to safely reduce unnecessary hospital admissions: a pilot randomised controlled trial. Emerg Med J. 2017 Sep 1;34(9):586–92.

10. MacKichan F, Brangan E, Wye L, Checkland K, Lasserson D, Huntley A, et al. Why do patients seek primary medical care in emergency departments? An ethnographic exploration of access to general practice. BMJ open. 2017;7(4):e013816.

11. Lowe RA, Localio AR, Schwarz DF, Williams S, Tuton LW, Maroney S, et al. Association between primary care practice characteristics and emergency department use in a Medicaid managed care organization. Medical care. 2005;792–800.

12. Body R, Carley S, Wibberley C, McDowell G, Ferguson J, Mackway-Jones K. The value of symptoms and signs in the emergent diagnosis of acute coronary syndromes. Resuscitation. 2010;81(3):281–6.

13. Kostis JB. The importance of managing hypertension and dyslipidemia to decrease cardiovascular disease. Cardiovascular drugs and therapy. 2007;21(4):297–309.

14. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HAW, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. The Lancet. 2004;364(9435):685–96.

15. Dahlöf B, Sever PS, Poulter NR, Wedel H, Beevers DG, Caulfield M, et al. Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendroflumethiazide as required, in the Anglo-Scandinavian Cardiac Outcomes Trial-Blood Pressure Lowering Arm (ASCOT-BPLA): a multicentre randomised controlled trial. The Lancet. 2005;366(9489):895–906.

16. British Heart Foundation. CVD Statistics—BHF UK Factsheet. 2020;

17. Wilkins E, Wilson L, Wickramasinghe K, Bhatnagar P, Leal J, Luengo-Fernandez R, et al. European cardiovascular disease statistics 2017. 2017;

18. Heart and Circulatory Diseases Statistics 2018. [Internet]. British Heart Foundation, University of Birmingham and the Office for National Statistics; [cited 2018 Jul 11]. Available from: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics/heart-statistics-publications/cardiovascular-disease-statistics-2018

19. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. bmj. 2017;357.

20. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics. 2018 Aug 13;

21. Fox CS, Matsushita K, Woodward M, Bilo HJ, Chalmers J, Heerspink HJL, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. The Lancet. 2012 Nov;380(9854):1662–73.

22. Hobbs FR. Cardiovascular disease: different strategies for primary and secondary prevention? Heart. 2004;90(10):1217–23.

23. Chaturvedi N. Ethnic differences in cardiovascular disease. Heart. 2003;89(6):681–6.

24. Hajar R. Risk factors for coronary artery disease: historical perspectives. Heart views: the official journal of the Gulf Heart Association. 2017;18(3):109.

25. Alves AJ, Viana JL, Cavalcante SL, Oliveira NL, Duarte JA, Mota J, et al. Physical activity in primary and secondary prevention of cardiovascular disease: Overview updated. World journal of cardiology. 2016;8(10):575.

26. Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J. 2013;34(38):2949–3003.

27. Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al. Fourth universal definition of myocardial infarction (2018). Journal of the American College of Cardiology. 2018;72(18):2231–64.

28. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). European heart journal. 2017;39(2):119–77.

29. Poldervaart JM, Reitsma JB, Backus BE, Koffijberg H, Veldkamp RF, ten Haaf ME, et al. Effect of Using the HEART Score in Patients With Chest Pain in the Emergency Department: A Stepped-Wedge, Cluster Randomized Trial. Annals of Internal Medicine. 2017 May 16;166(10):689.

30. Carlton EW, Ingram J, Taylor H, Glynn J, Kandiyali R, Campbell S, et al. Limit of detection of troponin discharge strategy versus usual care: randomised controlled trial. Heart. 2020;106(20):1586–94.

31. Anand A, Lee KK, Chapman AR, Ferry AV, Adamson PD, Strachan FE, et al. High-sensitivity cardiac troponin on presentation to rule out myocardial infarction: a stepped-wedge cluster randomized controlled trial. Circulation. 2021;143(23):2214–24.

32. Takeda S, Yamashita A, Maeda K, Maéda Y. Structure of the core domain of human cardiac troponin in the Ca2+-saturated form. Nature. 2003 Jul 1;424(6944):35–41.

33. Cummins B, Auckland ML, Cummins P. Cardiac-specific troponin-l radioimmunoassay in the diagnosis of acute myocardial infarction. American heart journal. 1987;113(6):1333–44.

34. Chapelle JP. Cardiac troponin I and troponin T: recent players in the field of myocardial markers. 1999;

35. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. The clinical biochemist reviews. 2008;29(Suppl 1):S49.

36. Aakre KM, Saenger AK, Body R, Collinson P, Hammarsten O, Jaffe AS, et al. Analytical Considerations in Deriving 99th Percentile Upper Reference Limits for High-Sensitivity Cardiac Troponin Assays: Educational Recommendations from the IFCC Committee on Clinical Application of Cardiac Bio-Markers. Clinical Chemistry. 2022 Aug 3;68(8):1022–30.

37. Apple FS, Jaffe AS, Collinson P, Mockel M, Ordonez-Llanos J, Lindahl B, et al. IFCC educational materials on selected analytical and clinical applications of high sensitivity cardiac troponin assays. Clinical Biochemistry. 2015 Mar 1;48(4):201–3.

38. High-sensitivity troponin tests for the early rule out of NSTEMI [Internet]. National Institute for Health and Care Excellence; 2020 Aug [cited 2022 Sep 19]. Report No.: DG40. Available from: https://www.nice.org.uk/guidance/dg40/chapter/1-Recommendations

39. Zhelev Z, Hyde C, Youngman E, Rogers M, Fleming S, Slade T, et al. Diagnostic accuracy of single baseline measurement of Elecsys Troponin T high-sensitive assay for diagnosis of acute myocardial infarction in emergency department: systematic review and meta-analysis. BMJ. 2015;350(8900488, bmj, 101090866):h15.

40. Bularga A, Lee KK, Stewart S, Ferry AV, Chapman AR, Marshall L, et al. High-sensitivity troponin and the application of risk stratification thresholds in patients with suspected acute coronary syndrome. Circulation. 2019;140(19):1557–68.

41. Chiang CH, Chiang CH, Lee GH, Gi WT, Wu YK, Huang SS, et al. Safety and efficacy of the European Society of Cardiology 0/1-hour algorithm for diagnosis of myocardial infarction: systematic review and meta-analysis. Heart. 2020;106(13):985–91.

42. Marcusson J, Nord M, Dong HJ, Lyth J. Clinically useful prediction of hospital admissions in an older population. BMC geriatrics. 2020;20(1):1–9.

43. Hodgson LE, Dimitrov BD, Roderick PJ, Venn R, Forni LG. Predicting AKI in emergency admissions: an external validation study of the acute kidney injury prediction score (APS). BMJ Open. 2017;7(3):e013511.

44. van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. Journal of clinical epidemiology. 2021;132:142–5.

45. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Annals of internal medicine. 2015;162(1):W1–73.

46. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux P, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. Jama. 2017;318(14):1377–84.

47. Wessler BS, Nelson J, Park JG, McGinnes H, Gulati G, Brazil R, et al. External validations of cardiovascular clinical prediction models: a large-scale review of the literature. Circulation: Cardiovascular Quality and Outcomes. 2021;14(8):e007858.

48. Body R, McDowell G, Carley S, Mackway-Jones K. Do risk factors for chronic coronary heart disease help diagnose acute myocardial infarction in the Emergency Department? Resuscitation. 2008;79(1):41–5.

49. Greenslade JH, Nayer R, Parsonage W, Doig S, Young J, Pickering JW, et al. Validating the Manchester Acute Coronary Syndromes (MACS) and Troponin-only Manchester Acute Coronary Syndromes (T-MACS) rules for the prediction of acute myocardial infarction in patients presenting to the emergency department with chest pain. Emergency Medicine Journal. 2017 Aug;34(8):517–23.

50. Ruangsomboon O, Thirawattanasoot N, Chakorn T, Limsuwat C, Monsomboon A, Praphruetkit N, et al. The utility of the 1-hour high-sensitivity cardiac troponin T algorithm compared with and combined with five early rule-out scores in high-acuity chest pain emergency patients. International journal of cardiology. 2020;

51. Steiro OT, Tjora HL, Langørgen J, Bjørneklett R, Nygård OK, Skadberg Ø, et al. Clinical risk scores identify more patients at risk for cardiovascular events within 30 days as compared to standard ACS risk criteria: the WESTCOR study. European Heart Journal Acute Cardiovascular Care. 2020;

52. Body R, Boachie C, McConnachie A, Carley S, Van Den Berg P, Lecky FE. Feasibility of the Manchester Acute Coronary Syndromes (MACS) decision rule to safely reduce unnecessary hospital admissions: a pilot randomised controlled trial. Emergency Medicine Journal. 2017;34(9):586–92.

53. Body R, Morris N, Reynard C, Collinson PO. Comparison of four decision aids for the early diagnosis of acute coronary syndromes in the emergency department. Emergency Medicine Journal. 2020;37(1):8–13.

54. Dawber TR, Meadors GF, Moore Jr FE. Epidemiological approaches to heart disease: the Framingham Study. American Journal of Public Health and the Nations Health. 1951;41(3):279–86.

55. Kannel WB, Dawber TR, Kagan A, Revotskie N, STOKES III J. Factors of risk in the development of coronary heart disease—six-year follow-up experience: the Framingham Study. Annals of internal medicine. 1961;55(1):33–50.

56. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97(18):1837–47.

57. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. Bmj. 2007;335(7611):136.

58. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. Bmj. 2008;336(7659):1475–82.

59. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ [Internet]. 2017 May 23;357. Available from: http://www.bmj.com/content/357/bmj.j2099.abstract

60. Rabar S, Harker M, O'Flynn N, Wierzbicki AS. Lipid modification and cardiovascular risk assessment for the primary and secondary prevention of cardiovascular disease: summary of updated NICE guidance. Bmj. 2014;349.

61. National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification. CG18. 2014;

62. Hickey GL, Grant SW, Murphy GJ, Bhabra M, Pagano D, McAllister K, et al. Dynamic trends in cardiac surgery: why the logistic EuroSCORE is no longer suitable for contemporary cardiac surgery and implications for future risk models. Eur J Cardiothorac Surg. 2013 Jun 1;43(6):1146–52.

63. Davis SE, Lasko TA, Chen G, Siew ED, Matheny ME. Calibration drift in regression and machine learning models for acute kidney injury. Journal of the American Medical Informatics Association. 2017 Nov 1;24(6):1052–61.

64. Davis SE, Lasko TA, Chen G, Matheny ME. Calibration Drift Among Regression and Machine Learning Models for Hospital Mortality. AMIA Annu Symp Proc. 2018 Apr 16;2017:625–34.

65. Jenkins DA, Martin GP, Sperrin M, Riley RD, Debray TP, Collins GS, et al. Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? Diagnostic and Prognostic Research. 2021;5(1):1–7.

66. Su TL, Jaki T, Hickey GL, Buchan I, Sperrin M. A review of statistical updating methods for clinical prediction models. Stat Methods Med Res. 2016 Jul 26;0962280215626466.

67. Martin GP, Mamas MA, Peek N, Buchan I, Sperrin M. Clinical prediction in defined populations: a simulation study investigating when and how to aggregate existing models. BMC Medical Research Methodology. 2017;17:1.

68. Debray TPA, Koffijberg H, Nieboer D, Vergouwe Y, Steyerberg EW, Moons KGM. Meta-analysis and aggregation of multiple published prediction models. Statist Med. 2014 Jun 30;33(14):2341–62.

69. Debray TPA, Koffijberg H, Vergouwe Y, Moons KGM, Steyerberg EW. Aggregating published prediction models with individual participant data: a comparison of different approaches. Stat Med. 2012 Oct 15;31(23):2697–712.

70. Raftery AE, Kárný M, Ettler P. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. Technometrics. 2010;52(1):52–66.

71. McCormick TH, Raftery AE, Madigan D, Burd RS. Dynamic logistic regression and dynamic model averaging for binary classification. Biometrics. 2012;68(1):23–30.

72. Siregar S, Nieboer D, Vergouwe Y, Versteegh MIM, Noyez L, Vonk ABA, et al. Improved Prediction by Dynamic Modeling. Circ Cardiovasc Qual Outcomes. 2016 Mar 1;9(2):171.

73. Strobl AN, Vickers AJ, Calster BV, Steyerberg E, Leach RJ, Thompson IM, et al. Improving patient prostate cancer risk assessment: Moving from static, globally-applied to dynamic, practice-specific risk calculators. Journal of biomedical informatics, Journal of biomedical informatics. 2015 Aug 1;56:87–93.

74. Batalden P. Getting more health from healthcare: quality improvement must acknowledge patient coproduction—an essay by Paul Batalden. BMJ. 2018 Sep 6;k3617.

75. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Quality & Safety. 2016 Jul;25(7):509–17.

76. Clarke D, Jones F, Harris R, Robert G. What outcomes are associated with developing and implementing co-produced interventions in acute healthcare settings? A rapid evidence synthesis. BMJ Open. 2017 Jul;7(7):e014650.

77. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. bmj. 2021;374.

78. Cowling TE, Cecil EV, Soljak MA, Lee JT, Millett C, Majeed A, et al. Access to Primary Care and Visits to Emergency Departments in England: A Cross-Sectional, Population-Based Study. PLOS ONE. 2013 Jun 12;8(6):e66699.

79. Cowling TE, Harris MJ, Watt HC, Gibbons DC, Majeed A. Access to general practice and visits to accident and emergency departments in England: cross-sectional analysis of a national patient survey. British Journal of General Practice. 2014;64(624):e434–9.

80. Tanner-Smith EE, Lipsey MW. Brief Alcohol Interventions for Adolescents and Young Adults: A Systematic Review and Meta-Analysis. Journal of Substance Abuse Treatment. 2015 Apr 1;51:1–18.

81. Fleming J, Meredith C, Henry J. Detection of hypertension in the emergency department. Emergency Medicine Journal. 2005 Sep 1;22(9):636–40.

82. Lee S, You C youn, Kim J, Jo YH, Ro YS, Kang SH, et al. Long-term cardiovascular risk of hypertensive events in emergency department: A population-based 10-year follow-up study. PLOS ONE. 2018 Feb 15;13(2):e0191738.

83. Tanabe P, Persell SD, Adams JG, McCormick JC, Martinovich Z, Baker DW. Increased Blood Pressure in the Emergency Department: Pain, Anxiety, or Undiagnosed Hypertension? Annals of Emergency Medicine. 2008 Mar;51(3):221–9.

84. Ferry AV, Strachan FE, Stewart SD, Marshall L, Lee KK, Anand A, et al. Exploring Patient Experience of Chest Pain Before and After Implementation of an Early Rule-Out Pathway for Myocardial Infarction: A Qualitative Study. Annals of Emergency Medicine. 2020;

85. Katz DA, Graber M, Birrer E, Lounsbury P, Baldwin A, Hillis SL, et al. Health Beliefs toward Cardiovascular Risk Reduction in Patients Admitted to Chest Pain Observation Units. Academic Emergency Medicine. 2009 May;16(5):379–87.

86. Katz DA, Graber M, Lounsbury P, Vander Weg MW, Phillips EK, Clair C, et al. Multiple Risk Factor Counseling to Promote Heart-healthy Lifestyles in the Chest Pain Observation Unit: Pilot Randomized Controlled Trial. Acad Emerg Med. 2017;24(8):968–82.

87. Collins K, Gough S, Clancy M. Screening for hypertension in the emergency department. Emergency Medicine Journal. 2008 Apr 1;25(4):196–9.

88. Julliard K, Orvieto C, Win A, Sadler S, Al-Haddad B, Simmons B, et al. Feasibility of Referral of Patients with Elevated Blood Pressure from the Emergency Department. J Community Health. 2012 Feb 1;37(1):159–64.

89. Lee J. Should emergency departments really be screening for hypertension? Emergency medicine journal. 2005;22(9):640–640.

90. Murray MJ, LeBlanc CH. Clinic Follow-up From the Emergency Department: Do Patients Show Up? Annals of Emergency Medicine. 1996 Jan 1;27(1):56–8.

91. Schwalm JD, McKee M, Huffman MD, Yusuf S. Resource Effective Strategies to Prevent and Treat Cardiovascular Disease. Circulation. 2016 Feb 23;133(8):742–55.

92. Collins GS, Altman DG. An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. Bmj. 2010;340:c2442.

93. Reynard C, van den Berg P, Oliver G, Naguib MP, Sammut-Powell C, McMillan B, et al. The prognostic value of emergency department measured hypertension: A systematic review and meta-analysis. Academic Emergency Medicine. n/a(n/a).

94. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Systematic Reviews. 2012 Feb 9;1(1):2.

95. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P, others. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine. 2009;6(7):e1000097.

96. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol. 2018 Jan 10;18.

97. National Institute for Clinical Excellence. Hypertension in adults: diagnosis and management NICE guideline [NG136]. London NICE. 2019;

98. Williams B, Mancia G, Spiering W, Rosei EA, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of HypertensionThe Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension. Journal of hypertension. 2018;36(10):1953–2041.

99. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med. 1999 Mar 16;130(6):461–70.

100. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009 May 5;150(9):604–12.

101. WHO [Internet]. [cited 2019 Jun 26]. WHO | Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. Available from: https://www.who.int/diabetes/publications/diagnosis_diabetes2011/en/

102. World Health Organization. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. 2006;

103. Association AD. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2019. Diabetes Care. 2019 Jan 1;42(Supplement 1):S13–28.

104. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013 Feb 19;158(4):280–6.

105. Hamza TH, van Houwelingen HC, Stijnen T. The binomial distribution of meta-analysis was preferred to model within-study variability. Journal of clinical epidemiology. 2008;61(1):41–51.

106. Team RC. R: A language and environment for statistical computing. 2013;

107. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of statistical software. 2010;36(3):1–48.

108. Wickham H. ggplot2: elegant graphics for data analysis. Springer; 2016.

109. Schwarzer G. meta: An R package for meta-analysis. R news. 2007;7(3):40–5.

110. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine. 2002;21(11):1539–58.

111. Mason S, Kuczawski M, Teare MD, Stevenson M, Goodacre S, Ramlakhan S, et al. AHEAD Study: an observational study of the management of anticoagulated patients who suffer head injury. BMJ Open. 2017;7(1):e014324.

112. Niska RW. Blood pressure measurements at emergency department visits by adults: United States, 2007-2008. NCHS Data Brief. 2011 Aug;(72):1–8.

113. Gleason-Comstock J, Streater A, Ager J, Goodman A, Brody A, Kivell L, et al. Patient education and follow-up as an intervention for hypertensive patients discharged from an emergency department: a randomized control trial study protocol. BMC Emerg Med. 2015 Dec 21;15:38.

114. Achieving Blood Pressure Control Through Enhanced Discharge - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Jun 27]. Available from: https://clinicaltrials.gov/ct2/show/NCT02069015

115. Targeting of UnControlled Hypertension in Emergency Department - Full Text View - ClinicalTrials.gov [Internet]. [cited 2019 Jun 27]. Available from: https://clinicaltrials.gov/ct2/show/NCT03749499

116. McNaughton CD, Self WH, Zhu Y, Janke AT, Storrow AB, Levy P. Incidence of Hypertension-Related Emergency Department Visits in the United States, 2006 to 2012. The American Journal of Cardiology. 2015 Dec;116(11):1717–23.

117. Tanabe P, Steinmann R, Kippenhan M, Stehman C, Beach C. Undiagnosed hypertension in the ED setting—an unrecognized opportunity by emergency nurses. Journal of Emergency Nursing. 2004;30(3):225–9.

118. Karras DJ, Ufberg JW, Heilpern KL, Cienki JJ, Chiang WK, Wald MM, et al. Elevated blood pressure in urban emergency department patients. Academic emergency medicine. 2005;12(9):835–43.

119. Tilman K, DeLashaw M, Lowe S, Springer S, Hundley S, Counselman FL. Recognizing asymptomatic elevated blood pressure in ED patients: how good (bad) are we? The American journal of emergency medicine. 2007;25(3):313–7.

120. Baumann BM, Abate NL, Cowan RM, Chansky ME, Rosa K, Boudreaux ED. Characteristics and referral of emergency department patients with elevated blood pressure. Academic Emergency Medicine. 2007;14(9):779–84.

121. Svenson JE, Repplinger M. Hypertension in the ED: still an unrecognized problem. The American Journal of Emergency Medicine. 2008 Oct;26(8):913–7.

122. Umscheid CA, Maguire MG, Pines JM, Everett WW, Baren JM, Townsend RR, et al. Untreated hypertension and the emergency department: a chance to intervene? Academic Emergency Medicine. 2008;15(6):529–36.

123. Baumann BM, Cline DM, Cienki JJ, Egging D, Lehrmann JF, Tanabe P. Provider Self-Report and Practice: Reassessment and Referral of Emergency Department Patients With Elevated Blood Pressure. American Journal of Hypertension. 2009 Jun 1;22(6):604–10.

124. Shah T, Aronow WS, Peterson SJ, Goldwag D. Diagnosis, Treatment, and Referral of Hypertension or Prehypertension in an Emergency Department After an Educational Program: Preliminary Results: Hypertension and Management. The Journal of Clinical Hypertension. 2011 Jun;13(6):413–5.

125. Arhami Dolatabadi A, Motamedi M, Hatamabadi H, Alimohamadi H. Prevalence of Undiagnosed Hypertension in the Emergency Department. Trauma Monthly. 2014 Jan 25;19(1).

126. Tan N, Taylor DM. Feasibility and outcomes of screening for cardiovascular risk factors in the emergency department. Emergency Medicine Australasia. 2013 Apr 1;25(2):175–81.

127. Goldberg EM, Wilson T, Saucier C, Brody AM, Levy PD, Eaton CB, et al. Achieving the BpTRUth: emergency department hypertension screening and the Centers for Medicare & Medicaid Services quality measure. Journal of the American Society of Hypertension. 2017 May;11(5):290–4.

128. Oras P, Häbel H, Skoglund PH, Svensson P. Elevated blood pressure in the emergency department: a risk factor for incident cardiovascular disease. Hypertension. 2020;75(1):229–36.

129. Souffront K, Gestal C, Melkus GD, Richardson L. Recognition of Asymptomatic Hypertension in an Urban Emergency Department: Where Are We Now? Advanced emergency nursing journal. 2016 Dec;38(4):320–6.

130. Cienki JJ, Guerrera AD, Steed NR, Kubo EN, Baumann BM. Impact of an Electronic Medical Record System on Emergency Department Discharge Instructions for Patients With Hypertension. Postgraduate Medicine. 2013 Sep;125(5):59–66.

131. Adhikari S, Mathiasen R, Lander L. Elevated blood pressure in the emergency department: lack of adherence to clinical practice guidelines. Blood Pressure Monitoring. 2016 Feb;21(1):54–8.

132. Backer HD, Decker L, Ackerson L. Reproducibility of increased blood pressure during an emergency department or urgent care visit. Annals of emergency medicine. 2003;41(4):507–12.

133. Shiber-Ofer S, Shohat Z, Grossman A. Elevated Diastolic, But Not Systolic, Blood Pressure Measured in the Emergency Department Predicts Future Development of Hypertension in Normotensive Individuals. The Journal of Clinical Hypertension. 2015 May;17(5):359–63.

134. Cline DM, Selvan V, Ferrario CM, Ayala C, Lekiachvili A, Mensah G. Emergency department versus clinic follow-up blood pressures. Academic Emergency Medicine. 2006;13(5 Supplement 1):S188.

135. 蔡振興, 董春香, 黃麗儀. Nurse-led hypertension referral system in an emergency department for asymptomatic elevated blood pressure. Hong Kong Med J. 2012;18(3):201–6.

136. Meurer WJ, Dome M, Brown D, Delemos D, Oska S, Gorom V, et al. Feasibility of Emergency Department–initiated, Mobile Health Blood Pressure Intervention: An Exploratory, Randomized Clinical Trial. Academic Emergency Medicine. 2019;26(5):517–27.

137. Masood S, Austin PC, Atzema CL. A population-based analysis of outcomes in patients with a primary diagnosis of hypertension in the emergency department. Annals of emergency medicine. 2016;68(3):258–67.

138. Kontopantelis E, Springate DA, Reeves D. A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses. PloS one. 2013;8(7):e69930.

139. Silverman RA, Pahk R, Carbone M, Wells E, Mitzner R, Burris K, et al. The relationship of plasma glucose and HbA1c levels among emergency department patients with no prior history of diabetes mellitus. Academic emergency medicine. 2006;13(7):722–6.

140. Silverman RA, Thakker U, Ellman T, Wong I, Smith K, Ito K, et al. Hemoglobin A1c as a Screen for Previously Undiagnosed Prediabetes and Diabetes in an Acute-Care Setting. Diabetes Care. 2011 Sep 1;34(9):1908–12.

141. George P, Valabhji J, Dawood M, Henry J. Screening for type 2 diabetes in the accident and emergency department. Diabetic medicine. 2005;22(12):1766–9.

142. Charfen MA, Ipp E, Kaji AH, Saleh T, Qazi MF, Lewis RJ. Detection of Undiagnosed Diabetes and Prediabetic States in High-risk Emergency Department Patients. Academic Emergency Medicine. 2009 May 1;16(5):394–402.

143. Ginde AA, Cagliero E, Nathan DM, Camargo Jr CA. Point-of-care glucose and hemoglobin A1c in emergency department patients without known diabetes: implications for opportunistic screening. Academic Emergency Medicine. 2008;15(12):1241–7.

144. Hewat N, McD Taylor D, MacDonald E. Pilot study of random finger prick glucose testing as a screening tool for type 2 diabetes mellitus in the emergency department. Emergency Medicine Journal. 2009 Oct 1;26(10):732–3.

145. Jelinek GA, Weiland TJ, Moore G, Tan G, Maslin M, Bowman K, et al. Screening for type 2 diabetes with random finger-prick glucose and bedside HbA1c in an Australian emergency department: T2D screening in the ED. Emergency Medicine Australasia. 2010 Oct;22(5):427–34.

146. Hng TM, Hor A, Ravi S, Feng X, Lin J, Astell-Burt T, et al. Diabetes case finding in the emergency department, using HbA1c: an opportunity to improve diabetes detection, prevention, and care. BMJ Open Diabetes Res Care. 2016 Jun 3;4(1).

147. Farkouh ME, Aneja A, Reeder GS, Smars PA, Bansilal S, Lennon RJ, et al. Clinical risk stratification in the emergency department predicts long-term cardiovascular outcomes in a population-based cohort presenting with acute chest pain: primary results of the Olmsted county chest pain study. Medicine. 2009;88(5).

148. Sanchis J, Bodí V, Núñez J, Bosch X, Loma-Osorio P, Mainar L, et al. Limitations of clinical history for evaluation of patients with acute chest pain, non-diagnostic electrocardiogram, and normal troponin. Am J Cardiol. 2008 Mar 1;101(5):613–7.

149. Elder CV, Kerr D, Davey RX, Kelly AM. Potential health promotion benefits of lipid testing for all patients presenting with chest pain to an emergency department. Emergency Medicine Journal. 2006 Jan 1;23(1):23–6.

150. Diercks DB, Kirk JD, Turnipseed SD, Gershoff L, Amsterdam EA. Cholesterol Screening in an ED-Based Chest Pain Unit. The American Journal of Emergency Medicine. 2002 Oct 1;20(6):510–2.

151. Chandra A, Compton S, Sochor M, Puri S, Zalenski RJ. Untreated Hypercholesterolemia in an Emergency Department Chest Pain Observation Unit Population. Academic Emergency Medicine. 2002;9(7):699–702.

152. Chaikriangkrai K, Nabi F, Mahmarian JJ, Chang SM. Additive prognostic value of coronary artery calcium score and renal function in patients with acute chest pain without known coronary artery disease: up to 5-year follow-up. The International Journal of Cardiovascular Imaging. 2015 Dec;31(8):1619–26.

153. Altman DG, Royston P. The cost of dichotomising continuous variables. Bmj. 2006;332(7549):1080.

154. Bernstein SL, Haukoos JS. Public health, prevention, and emergency medicine: a critical juxtaposition. Academic emergency medicine: official journal of the Society for Academic Emergency Medicine. 2008;15(2):190–3.

155. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med. 2009 May;36(5):452–7.

156. Lawson PJ, Flocke SA. Teachable moments for health behavior change: a concept analysis. Patient education and counseling. 2009;76(1):25–30.

157. Wolf SJ, Lo B, Shih RD, Smith MD, Fesmire FM. Clinical policy: critical issues in the evaluation and management of adult patients in the emergency department with asymptomatic elevated blood pressure. Annals of emergency medicine. 2013;62(1):59–68.

158. Kavsak PA, Worster A, Ma J, Shortt C, Clayton N, Sherbino J, et al. High-Sensitivity Cardiac Troponin Risk Cutoffs for Acute Cardiac Outcomes at Emergency Department Presentation. Can J Cardiol. 2017;33(7):898–903.

159. Etaher A, Nguyen TL, Saad YM, Frost S, Ferguson I, Juergens CP, et al. Mortality at 5 Years Among Very Elderly Patients Undergoing High Sensitivity Troponin T Testing for Suspected Acute Coronary Syndromes. Heart, Lung and Circulation. 2020;29(11):1696–703.

160. Heinze G, Wallisch C, Dunkler D. Variable selection–a review and recommendations for the practicing statistician. Biometrical journal. 2018;60(3):431–49.

161. Tibshirani R. The lasso method for variable selection in the cox model. Statistics in medicine. 1997 Feb 28;16(4).

162. Petersson S, Sehlstedt K, Herbertsson A. Variable selection techniques for the Cox proportional hazards model: A comparative study [Internet]. University of Gothenburg; 2018 Feb [cited 2023 Feb 9]. Available from: https://core.ac.uk/download/pdf/152600668.pdf

163. Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine. 1996;15(4):361–87.

164. Dessai S, Patil V. Testing and interpreting assumptions of COX regression analysis. Cancer Research, Statistics, and Treatment. 2019;2(1):108–11.

165. Gauthier J, Wu Q, Gooley T. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. Bone marrow transplantation. 2020;55(4):675–80.

166. Keele L. Proportionally difficult: testing for nonproportional hazards in Cox models. Political Analysis. 2010;18(2):189–205.

167. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81(3):515–26.

168. Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. Statistics in medicine. 1995;14(15):1707–23.

169. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj. 2009;338.

170. Kontopantelis E, White IR, Sperrin M, Buchan I. Outcome-sensitive multiple imputation: a simulation study. BMC medical research methodology. 2017;17(1):1–13.

171. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005;1130–9.

172. NHS Digital. National Clinical Coding Standards: Opcs-4 (2017), Accurate Data for Quality Information. TSO; 2017.

173. Riley RD, Ensor J, Snell KI, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. Bmj. 2020;368.

174. Riley RD, Van Calster B, Collins GS. A note on estimating the Cox-Snell R2 from a reported C statistic (AUROC) to inform sample size calculations for developing a prediction model with a binary outcome. Statistics in Medicine. 2021 Feb 20;40(4):859–64.

175. RStudio T. RStudio: integrated development for R. Rstudio Team, PBC, Boston, MA URL http://www rstudio com. 2020;

176. R Core Team (2013). R: A language and environment for statistical computing [Internet]. . Vienna, Austria.: R Foundation for Statistical Computing,; Available from: URL http://www.R-project.org/.

177. Harrell Jr FE, Harrell Jr MFE. Package 'hmisc'. CRAN2018. 2019;2019:235–6.

178. Potter JM, Simpson AJ, Kerrigan J, Southcott E, Salib MM, Koerbin G, et al. The relationship of plasma creatinine (as eGFR) and high-sensitivity cardiac troponin and NT-proBNP concentrations in a hospital and community outpatient population. Clinical Biochemistry. 2017;50(15):813–5.

179. Therneau T, Lumley T. R survival package. R Core Team. 2013;

180. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239–41.

181. Joanes DN, Gill CA. Comparing measures of sample skewness and kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician). 1998;47(1):183–9.

182. Woodward M, Brindle P, Tunstall-Pedoe H. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart. 2007;93(2):172–6.

183. Hippisley-Cox J. Update for QRISK2 and Release of Open Source Software. London: BMJ. 2010;

184. Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D. Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019—results from a systematic review and meta-analysis. Critical Care. 2020;24(1):1–9.

185. National General Practice Profiles [Internet]. Public Health England; [cited 2022 Apr 1]. Available from: https://fingertips.phe.org.uk/profile/general-practice/data#page/8

186. Conroy R, Pyörälä K, Fitzgerald A el, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. European heart journal. 2003;24(11):987–1003.

187. Kavsak PA, Cerasuolo JO, Ko DT, Perez R, Seow H, Ma J, et al. High-Sensitivity Cardiac Troponin T Testing and Cardiovascular Outcomes at 30 Days and 1 Year in Patients Discharged Home from the Emergency Department with Chest Pain. The Journal of Applied Laboratory Medicine. 2020;5(4):821–4.

188. Stewart J, Lu J, Goudie A, Bennamoun M, Sprivulis P, Sanfillipo F, et al. Applications of machine learning to undifferentiated chest pain in the emergency department: A systematic review. PloS one. 2021;16(8):e0252612.

189. Khand A, Frost F, Grainger R, Fisher M, Chew P, Mullen L, et al. Identification of high-risk non-ST elevation myocardial infarction at presentation to emergency department. A prospective observational cohort study in North West England. BMJ open. 2020;10(6):e030128.

190. Hinton J, Mariathas M, Gabara L, Allan R, Nicholas Z, Kwok CS, et al. Relation of high-sensitivity troponin to 1 year mortality in 20,000 consecutive hospital patients undergoing a blood test for any reason. The American Journal of Cardiology. 2021;158:124–31.

191. Su TL, Jaki T, Hickey GL, Buchan I, Sperrin M. A review of statistical updating methods for clinical prediction models. Statistical methods in medical research. 2018;27(1):185–97.

192. Vogel JA, Rising KL, Jones J, Bowden ML, Ginde AA, Havranek EP. Reasons patients choose the emergency department over primary care: a qualitative metasynthesis. Journal of General Internal Medicine. 2019;34:2610–9.

193. Yang SC, Zwar N, Vagholkar S, Dennis S, Redmond H. Factors influencing general practice follow-up attendances of patients with complex medical problems after hospitalization. Family practice. 2010;27(1):62–8.

194. Niska RW. Blood pressure measurements at emergency department visits by adults: United States, 2007-2008. US Department of Health and Human Services, Centers for Disease Control and …; 2011.

195. Atzema CL, Yu B, Schull MJ, Jackevicius CA, Ivers NM, Lee DS, et al. Physician follow-up and long-term use of evidence-based medication for patients with hypertension who were discharged from an emergency department: a prospective cohort study. Canadian Medical Association Open Access Journal. 2018;6(2):E151–61.

196. Fox NJ. Post-positivism. The SAGE encyclopedia of qualitative research methods. 2008;2:659–64.

197. Adding Theoretical Grounding to Grounded Theory: Toward Multi-Grounded Theory - Göran Goldkuhl, Stefan Cronholm, 2010 [Internet]. [cited 2019 Mar 2]. Available from: https://journals.sagepub.com/doi/full/10.1177/160940691000900205

198. Smith L, Elwick H, Mhizha-Murira JR, Topcu G, Bale C, Evangelou N, et al. Developing a clinical pathway to identify and manage cognitive problems in Multiple Sclerosis: Qualitative findings

from patients, family members, charity volunteers, clinicians and healthcare commissioners. Multiple sclerosis and related disorders. 2021;49:102563.

199. Bradbury K, Steele M, Corbett T, Geraghty AW, Krusche A, Heber E, et al. Developing a digital intervention for cancer survivors: an evidence-, theory-and person-based approach. NPJ digital medicine. 2019;2(1):85.

200. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. Bmj. 2013;347.

201. EBCD: experience-based co-design toolkit: the point of care Foundation, 2016. [Internet]. The Point of Care Foundation; [cited 2023 Jul 18]. Available from: : https://www. pointofcarefoundation. org. uk/ resource/experience-based-co-design-ebcd-toolkit/step-bystep-guide/1-experience-based-co-design/

202. Green T, Bonner A, Teleni L, Bradford N, Purtell L, Douglas C, et al. Use and reporting of experience-based codesign studies in the healthcare setting: a systematic review. BMJ quality & safety. 2020;29(1):64–76.

203. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ open. 2019;9(8):e029954.

204. Rotter T, Kinsman L, James E, Machotta A, Steyerberg EW. The quality of the evidence base for clinical pathway effectiveness: room for improvement in the design of evaluation trials. BMC medical research methodology. 2012;12:1–8.

205. Gore M. Odette, Seliger Stephen L., deFilippi Christopher R., Nambi Vijay, Christenson Robert H., Hashim Ibrahim A., et al. Age- and Sex-Dependent Upper Reference Limits for the High-Sensitivity Cardiac Troponin T Assay. Journal of the American College of Cardiology. 2014 Apr 15;63(14):1441–8.

206. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ. 2016 Jun 22;353:i3140.

207. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P, for the CHD Risk Prediction Group. Validation of the Framingham Coronary Heart Disease Prediction Scores: Results of a Multiple Ethnic Groups Investigation. JAMA. 2001 Jul 11;286(2):180–7.

208. Cubbin C, Sundquist K, Ahlen H, Johansson SE, WInkleby M, Sunquist J. Neighborhood deprivation and cardiovascular disease risk factors: Protective and harmful effects. Scandinavian Journal of Public Health. 2006;34(3):228–37.

209. Ramsay SE, Morris RW, Whincup PH, Subramanian SV, Papacosta AO, Lennon LT, et al. The influence of neighbourhood-level socioeconomic deprivation on cardiovascular disease mortality in older age: longitudinal multilevel analyses from a cohort of older British men. J Epidemiol Community Health. 2015 Dec 1;69(12):1224.

210. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clinical chemistry. 2008;54(1):17–23.

211. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European heart journal. 2014;35(29):1925–31.

212. Van Calster B, McLernon DJ, Van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative Patrick Bossuyt Gary S. Collins Petra Macaskill David J. McLernon Karel GM Moons Ewout W. Steyerberg Ben Van Calster Maarten van Smeden Andrew J. Vickers. Calibration: the Achilles heel of predictive analytics. BMC medicine. 2019;17:1–7.

213. Stevens RJ, Poppe KK. Validation of clinical prediction models: what does the "calibration slope" really measure? Journal of clinical epidemiology. 2020;118:93–9.

214. Rufibach K. Use of Brier score to assess binary predictions. Journal of clinical epidemiology. 2010;63(8):938–9.

215. Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. Journal of the American Medical Informatics Association. 2020;27(4):621–33.

216. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. Journal of the American Medical Informatics Association. 2013;20(1):144–51.

217. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;837–45.

218. Koukkunen H, Pyörälä K, Halinen MO. Low-risk patients with chest pain and without evidence of myocardial infarction may be safely discharged from emergency department. European Heart Journal. 2004 Feb 1;25(4):329–34.

219. Levin A, Stevens PE, Bilous RW, Coresh J, De Francisco AL, De Jong PE, et al. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney international supplements. 2013;3(1):1–150.

220. Body R, Carlton E, Sperrin M, Lewis PS, Burrows G, Carley S, et al. Troponin-only Manchester Acute Coronary Syndromes (T-MACS) decision aid: single biomarker re-derivation and external validation in three cohorts. Emerg Med J. 2017;34(6):349–56.

221. Va Den Berg P, Burrows G, Lewis P, Carley S, Body R. Validation of the (Troponin-only) Manchester ACS decision aid with a contemporary cardiac troponin I assay. Am J Emerg Med. 2018;36(4):602–7.

222. Body R, Almashali M, Morris N, Moss P, Jarman H, Appelboam A, et al. Diagnostic accuracy of the T-MACS decision aid with a contemporary point-of-care troponin assay. Heart. 2019;105(10):768–74.

223. Alghamdi A, Reynard C, Morris N, Moss P, Jarman H, Hardy E, et al. Diagnostic accuracy of the Troponin-only Manchester Acute Coronary Syndromes (T-MACS) decision aid with a point-of-care cardiac troponin assay. Emergency Medicine Journal. 2020;37(4):223–8.

224. Than M, Herbert M, Flaws D, Cullen L, Hess E, Hollander JE, et al. What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the Emergency Department?: a clinical survey. International journal of cardiology. 2013;166(3):752–4.

225. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Statistics in medicine. 2014;33(3):517–35.

226. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. Journal of clinical epidemiology. 2016;74:167–76.

227. Nguyen MT, Kung WM, Lehman S, Chew DP. High-sensitivity troponin in chronic kidney disease: Considerations in myocardial infarction and beyond. Reviews in Cardiovascular Medicine. 2020;21(2):191–203.

228. Herzog CA, Asinger RW, Berger AK, Charytan DM, Díez J, Hart RG, et al. Cardiovascular disease in chronic kidney disease. A clinical update from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney international. 2011;80(6):572–86.

229. Hickey GL, Grant SW, Murphy GJ, Bhabra M, Pagano D, McAllister K, et al. Dynamic trends in cardiac surgery: why the logistic EuroSCORE is no longer suitable for contemporary cardiac surgery and implications for future risk models. Eur J Cardiothorac Surg. 2013 Jun;43(6):1146–52.

230. Genders TS, Steyerberg EW, Alkadhi H, Leschka S, Desbiolles L, Nieman K, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. European heart journal. 2011;32(11):1316–30.

231. Body R, Carlton E, Sperrin M, Lewis PS, Burrows G, Carley S, et al. Troponin-only Manchester Acute Coronary Syndromes (T-MACS) decision aid: single biomarker re-derivation and external validation in three cohorts. Vol. 34. 2017. p. 349.

232. Riley RD, Snell KI, Ensor J, Burke DL, Harrell Jr FE, Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II-binary and time-to-event outcomes. Statistics in medicine. 2019;38(7):1276–96.

233. Jenkins D. Dynamic modelling methods for clinical prediction model updating and monitoring: identification and comparison of dynamic models for healthcare use. [Manchester]: University of Manchester; 2022.

234. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. Euroscore ii. European journal of cardio-thoracic surgery. 2012;41(4):734–45.

235. Hickey GL, Grant SW, Caiado C, Kendall S, Dunning J, Poullis M, et al. Dynamic prediction modeling approaches for cardiac surgery. Circulation: Cardiovascular Quality and Outcomes. 2013;6(6):649–58.

236. Jenkins DA, Sperrin M, Martin GP, Peek N. Dynamic models to predict health outcomes: current status and methodological challenges. Diagnostic and Prognostic Research. 2018 Dec 18;2(1):23.

237. Coles B, Teece L, Weston C, de Belder MA, Oliver-Williams C, Welch CA, et al. Case-ascertainment of acute myocardial infarction hospitalizations in cancer patients: a cohort study

using English linked electronic health data. European Heart Journal-Quality of Care and Clinical Outcomes. 2022;8(1):86–95.

238. National Data Opt-out, May 2023 [Internet]. NHS Digital; 2023 Jun [cited 2023 Jul 21]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/may-23

239. King KM, Khan NA, Quan H. Ethnic variation in acute myocardial infarction presentation and access to care. The American journal of cardiology. 2009;103(10):1368–73.

240. International Statistical Classification of Diseases and related health problems: Alphabetical index. World Health Organisation; 2022. Report No.: Volume 3.