Advancing the use of routinely collected health data in observational research to study comparative treatment effects: Two natural experiments in UK primary and secondary care

PATRICK BRIAN BIDULKA

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy of the University of London

July 2024

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology & Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by the National Institute for Health and Care Research (NIHR), Kidney Research UK, & The Health Foundation

Research group affiliation: Electronic Health Records Group

DECLARATION

I, Patrick Brian Bidulka, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Observational studies using routinely collected health data can inform health technology assessments (HTA) and change clinical practice. However, when observational studies are used to investigate comparative treatment effects, a major concern is confounding bias. Careful study design is required to reduce this risk of bias and generate evidence that is useful for national and individual-level clinical decision-making.

In this thesis, I aim to advance the use of routinely collected health data to assess comparative effectiveness. I use the target trial emulation (TTE) framework and instrumental variable analyses to reduce the risk of confounding in two case studies in England focused on the relative effectiveness of (1) sulfonylureas (SU), dipeptidyl peptidase-4 inhibitors (DPP4i), and sodium-glucose co-transporter 2 inhibitors (SGLT2i), all added to metformin, as second-line oral antidiabetic treatments for people with type 2 diabetes mellitus, and (2) invasive versus conservative cardiac management for acute myocardial infarction (AMI) among people with kidney impairment.

In Case Study 1, I find evidence that SGLT2i are better than the alternatives at improving important clinical measures and reducing the hazards of cardiovascular and kidney outcomes. I exploit exogenous variation in second-line antidiabetic prescribing at the general practice group-level as a preference-based instrument. I also find evidence of treatment heterogeneity for DPP4i versus SU on mean change in haemoglobin A1c.

In Case Study 2, I illustrate potential biases when using secondary care datasets to study AMI treatment in people with kidney impairment. I also demonstrate variation in alternative AMI treatment at the cardiology-centre level which could be used in future natural experiments.

I conclude that careful study design and analyses helped reduce the risk of bias in these observational studies using routinely collected health data. I demonstrate challenges and opportunities for using these data to inform clinical practice and HTA, and to improve outcomes for patients.

ACKNOWLEDGEMENTS

I owe many people thanks for helping me complete this thesis.

Firstly, I thank my PhD supervisors, Dorothea Nitsch, Richard Grieve, and Helen Strongman.

Dorothea and Richard have provided unwavering support as my PhD supervisors and line managers at LSHTM. I am greatly appreciative of their leadership and the opportunities they have facilitated for me to explore my interests and develop as a well-rounded academic over the past several years. I am lucky to have them as mentors.

Helen has been a steadfast colleague, running buddy, mentor, and PhD supervisor. I am grateful for her support throughout my LSHTM employment and PhD registration.

I am grateful to have had the opportunity to work on multidisciplinary teams of researchers from the PERMIT and the QECKD/NACARAI studies. In particular, I would like to thank David Lugo-Palacios, Stephen O'Neill, Orlagh Carroll, Anirban Basu, Amanda Adler, Kamlesh Khunti, Liam Smeeth, Richard Silverwood, Paul Charlton, David Adlam, Clive Weston, Mark de Belder, John Deanfield, and Jemima Scott who have all spent time and energy coaching me to produce better research and to be a better colleague.

I also thank Charlotte Warren-Gash, Elizabeth 'Fizz' Williamson, and Rachael Williams, who provided valuable input throughout my PhD registration to improve this thesis.

I am extremely grateful to Ian Douglas, for seeing my potential as my MSc supervisor and continuing to guide me as I navigate my early-career research years; to Laurie Tomlinson, for mentoring me in my first job post-MSc; and to my department heads, David Leon and Sanjay Kinra, for giving me valuable advice throughout my time at LSHTM.

There are many people in the LSHTM Electronic Health Records group, past and present, who have directly or indirectly helped me complete this thesis. In particular, Angel Wong, Anna Schultze, John Tazare, Kate Mansfield, Krishnan Bhaskaran, Sinéad Langan, Stephen Evans, Kevin Wing, Ruth Costello, Julian Matthewman, Anne Suffel, Helen McDonald, Rutendo Muzambi, Harriett Forbes, Emily Herrett, Penny Bloore, Paris Baptiste, Sophie Eastwood, Sophie Hutton, and Rohini Mathur.

I was fortunate to be hired for my first research gig by Jackie Bosch, who has taught me important lessons in humility, kindness, and hospitality over the past seven years.

And finally, I am a product of the friends and family I am so fortunate to have in my life. I could not have achieved this milestone without the support of my McGill and London friends, CPFs, grandparents, aunts, uncles, cousins, Andrew and Ellen, and of course, my wonderful parents.

GLOSSARY

2SLS	Two-stage least squares
2SRI	Two-stage residual inclusion
ACEI/ARB	Angiotensin converting enzyme inhibitor/angiotensin receptor blocker
ACS	Acute coronary syndrome
AKI	Acute kidney injury
AMI	Acute myocardial infarction
ANCOVA	Analysis of covariance
ATE	Average treatment effect
ATT	Average treatment effect in the treated
BHF	British Heart Foundation
BMI	Body-mass index
CATE	Conditional average treatment effect
CCG	Clinical Commissioning Group
CI	Confidence interval
CKD	Chronic kidney disease
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular disease
DAG	Directed acyclic graph
DBP	Diastolic blood pressure
DPP4i	Dipeptidyl peptidase-4 inhibitors
eGFR	estimated glomerular filtration rate
EHR	Electronic Health Record
GLP1-RA	Glucagon-like peptide-1 receptor agonists
GP	General practice
HbA1c	Haemoglobin A1c
HES	Hospital Episode Statistics
HQIP	Healthcare Quality Improvement Partnership
HTA	Health technology assessments
ICD-10	International Classification of Diseases – 10^{th} Edition
ICS	Integrated Care System
IMD	Index of Multiple Deprivation
IPTW	Inverse probability of treatment weighting
ISAC	Independent Scientific Advisory Committee
ITT	Intention-to-treat
IV	Instrumental variable
KRUK	Kidney Research UK
LATE	Local average treatment effect
LIV	Local instrumental variable

LSHTM	London School of Hygiene & Tropical Medicine
MACE	Major adverse cardiovascular event
MAKE	Major adverse kidney event
MINAP	Myocardial Ischaemia National Audit Project
MHRA	Medicines and Healthcare products Regulatory Agency
MTE	Marginal treatment effect
NACARAI	Improving acute cardiac care of patients with renal disease through linkage of national audits in the UK: the National Cardiac and Renal Audit Initiative project
NAPCI	National Audit of Percutaneous Coronary Interventions
NCKDA	National Chronic Kidney Disease Audit
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NHS	National Health Service
NRS	Non-randomised studies
NSTEMI	non-ST-elevated myocardial infarction
OLS	Ordinary least squares
ONS	Office of National Statistics
PCI	Percutaneous coronary intervention
PERMIT	Personalised Medicine for Intensification of Treatment study
PH	Proportional hazards
PPI	Patient and public involvement
QECKD	Quality and Equity of Care in Kidney Disease study
RCT	Randomised controlled trials
RWD	Real-world data
SAP	Statistical analysis plan
SBP	Systolic blood pressure
SCr	Serum creatinine
SD	Standard deviation
SGLT2i	Sodium-glucose co-transporter 2 inhibitors
STEMI	ST-elevated myocardial infarction
SU	Sulfonylureas
T2DM	Type 2 diabetes mellitus
TTE	Target trial emulation
TTP	Tendency to prescribe
UK	United Kingdom
US	United States

Table of Contents

Table of Tables

Table of Figures

CHAPTER 1. INTRODUCTION

Overview

In this chapter, I provide background information on routinely collected health data and its relevance to health technology assessment (HTA) and clinical practice, particularly in the United Kingdom (UK) healthcare system. I then summarise causal inference in pharmacoepidemiology and how routinely collected health data can be used to generate evidence on the comparative effectiveness of treatments in the presence of clinical uncertainty. In particular, I concentrate on the target trial emulation (TTE) framework combined with an instrumental variable (IV) analysis as a strategy for causal inference. I then describe two examples of clinical uncertainty: (i) second-line oral antidiabetic treatments for people with type 2 diabetes mellitus (T2DM), and (ii) alternative acute myocardial infarction (AMI) treatments among people with kidney impairment. Finally, I outline the rationale, aims, objectives, and structure of this thesis.

1.1. Routinely collected health data and its relevance to HTA in the UK

Routinely collected health data, sometimes referred to as 'real-world data' (RWD), are increasingly used globally to generate evidence to inform HTA and regulatory decisions.[1, 2] Availability of these data has increased substantially in many countries,[3] with the United Kingdom (UK) being a world-leader in organising and utilising these data for health research.[1, 4-7] A major source of health data in the UK is the National Health Service (NHS), funded through general taxation and free at the point of use to the general UK population 'from cradle to grave'.[8] The NHS healthcare ecosystem includes primary and secondary care, which generate vast amounts of routinely collected health data that are used widely in health research.[9]

1.1.1. NHS primary and secondary care

Referred to as the primary care system, general practitioners manage a range of services for most diseases and conditions. General practitioners act as the gatekeepers to specialised healthcare provided in hospitals and outpatient clinics. Almost everyone living in the UK is registered with a general practice (GP).[10] Between 2013 to 2022, GPs were grouped into Clinical Commissioning Groups (CCGs), which commissioned health services at the local level for the populations they served. Approximately 60% of the NHS budget was managed by CCGs.[11] In 2022, the NHS reorganised CCGs[12] into Integrated Care Systems (ICSs) in response to the challenges of the ageing population increasingly living with multiple long-term conditions.[13] Each of the 42 ICSs has one Integrated Care Board which, like CCGs, commission health services tailored to the local population with the goal of joining up care between primary and secondary care to reduce health inequalities and improve outcomes.[12]

Secondary care services evolved separately from GPs and are delivered by specialist clinicians in hospitals or specialised clinics.[8] These services, organised in an organ-specific delivery model, have been slow to adapt to the changing epidemiological profiles of the general population, which is older and more commonly living with more than one long-term condition (i.e., multimorbidity).[13, 14]

These healthcare services generate large health datasets. Linking these disparate datasets is key to leverage these data for high-quality health research useful to healthcare providers and HTA.

1.1.2. UK Electronic Health Records (EHRs) and clinical audits used for research

UK primary care EHR data were first collected by a commercial company named VAMP (Value Added Medical Products) Research Bank starting in 1987.[15] The goal was to create a representative research database including four million patients in UK primary care for post-marketing studies for

pharmaceutical companies.[15] Since 1994, these primary care data have been managed by the UK government's Department of Health, and evolved into the General Practice Research Database, and now, since 2012, the Clinical Practice Research Datalink (CPRD).[16] Primary care EHR in the UK are supported by only four major software providers, two of which contribute to the CPRD (Vision and EMIS).[17] Because GPs are usually the first point of contact for patients in the NHS other than emergency visits to the hospital, these data are a rich source of information related to individuals' sociodemographic characteristics, clinical diagnoses, laboratory tests and results, and prescription data.[16, 18]

UK secondary care EHR data developed in parallel. These EHR hospital data were first collected in 1989.[19] Today, NHS England collates the Hospital Episode Statistics (HES) dataset which are used to manage and plan services and re-imburse hospitals in England.[19] HES Admitted Patient Care data contain demographic information (e.g., age and sex) and diagnoses translated from hospital letters by human coders using the International Classification of Diseases 10th Edition (ICD-10) codes. Some specialities in certain hospital trusts are still paper-based for aspects of care such as in-patient drug prescribing, meaning there is no national database in England for hospital prescriptions.[19]

Clinical audits are funded separately from primary and secondary care services by organisations like NHS England or the Healthcare Quality Improvement Partnership (HQIP).[20, 21] These audits are driven by clinical specialities who seek to ensure quality care, and tend to focus on a particular disease area. Audits use standardised electronic capture tools to retrospectively assess GP, hospital, or out-patient health records to understand whether recommended treatment guidelines are followed by the healthcare provider. These data are collected and used to improve aspects of clinical care.[20, 21]

Certain linkages between primary and secondary care are routine. These linkages are based on the unique NHS number (patient identifier) and incorporate several other variables to increase the accurateness and quality of the linkage.[22] Health services planning, auditors, and researchers benefit from these linkages as it more completely captures the health status of the patient and the health and social care the patient receives.[23]

Linking health data in observational and randomised studies alike is important to improve the efficiency and quality of research. These data are increasingly used in large-scale observational studies,[4, 5, 24, 25] and more recently, randomised controlled trials (RCTs).[26, 27]

1.1.3. Challenges in generating evidence from routinely collected health data for HTA and clinical practice

RCTs are considered the 'gold-standard' study design to evaluate interventions, including pharmacological treatments. This is because the randomisation of treatments across study participants removes any differences in baseline characteristics between comparison groups on average except due to chance. However, RCTs can be difficult to generalise to the target populations of interest. Treatment effects estimated in RCTs may not necessarily be consistent across people excluded from trials because of differences in the setting and characteristics of people included in the trial versus in routine care. This creates challenges for HTA agencies, who must make decisions with incomplete information on the effectiveness of treatments evaluated in RCTs. Observational studies using routinely collected health data can address these evidence gaps to complement RCT evidence and inform clinical practice.

The National Institute for Health and Care Excellence (NICE), an HTA agency for England and Wales, recognises the value of 'real-world evidence' generated using routinely collected health data.[1] NICE already uses evidence generated from RWD to describe health conditions, processes of care, and patient experiences, create and populate economic models, and characterise health inequalities.[1] NICE also recognises the potential of using routinely collected health data to estimate comparative treatment effects where RCT data are not available. However, these studies face major challenges in minimising biases to generate high-quality evidence, since these data are usually not collected specifically for research purposes and treatments are not randomly allocated.[1] Significant advancements in causal inference has helped to address these important limitations of observational health research.

1.2. Causal inference in pharmacoepidemiological studies

1.2.1. Causal inference frameworks

Causal inference is a broad concept which includes many 'frameworks' to approach the difficult task of attributing a causal effect of an exposure on an outcome.[28] In 1965, Bradford Hill proposed a list of 'viewpoints' to guide epidemiological analyses seeking to elucidate causal relationships between an exposure and outcome.[29, 30] This list included: (i) strength of association, (ii) consistency, (iii) specificity, (iv) temporality, (v) biological gradient, (vi) plausibility, (vii) coherence, (viii) experiment, and (ix) analogy.[29] These criterion are useful in designing and hypothesising whether an observed association

between a treatment and outcome is causal. But more nuanced and rigorous frameworks are needed to estimate treatment effects in routinely collected health data.

Since the acceptance that smoking is a cause of a lung cancer and the advent of RCTs, the potential outcomes framework for causal inference has undergone substantial development[31] as one branch in a diverse suite of causal inference frameworks, focused on interventions which can be manipulated by humans (e.g., a pharmacological treatment).[28] This framework is popular in pharmacoepidemiology, "the study of interactions between drugs and human populations, investigating, in real conditions of life, benefits, risks and use of drugs."[32] In this scientific field, researchers are often interested in elucidating the causal association between a pharmacological agent and an outcome.

Under the potential outcomes framework, the probability of the outcome occurring changes in the presence or absence of the exposure/treatment.[33-35] However, the fundamental problem of causal inference is that we cannot observe both potential outcomes (or the counterfactual outcome) – we can only observe the outcome which actually occurs under a particular exposure. This induces a missing data problem, where we are missing a potential outcome with which to draw our causal inference.[36]

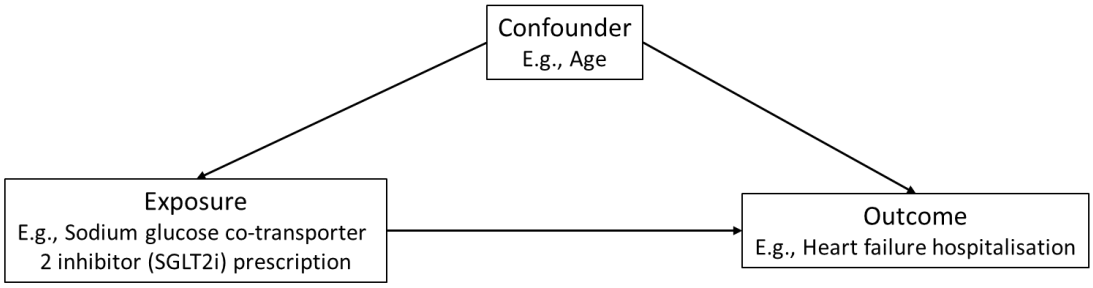
Because of this missing data problem, it is impossible to measure individual treatment effects. However, we can estimate various causal estimands in epidemiological analyses, including the average treatment effect (ATE). The ATE is the difference in the average potential outcome when everyone is exposed and the average potential outcome when no one is exposed.[37] This and other causal estimands, such as the average treatment effect in the treated, the average treatment effect in the untreated, and conditional average treatment effects (CATEs) are useful to clinicians and policymakers to understand which treatment is best for particular people.[37]

1.2.2. Causal diagrams, confounding, and colliders

Causal diagrams, such as directed acyclic graphs (DAGs), are important tools in causal inference when estimating causal treatment effects. DAGs are useful in pharmacoepidemiological studies to explicitly illustrate the causal framework and assumptions underlying the relationship between an exposure and outcome,[38] as well as other variables which will impact the analysis strategy such as confounders, colliders, mediators, and instruments.

Confounding in particular is an important bias that must be addressed in causal inference studies. A confounder is a third variable which directly influences treatment receipt, is a cause of the outcome, and does not lie on the causal pathway between the treatment and outcome (**Figure 1.1**).

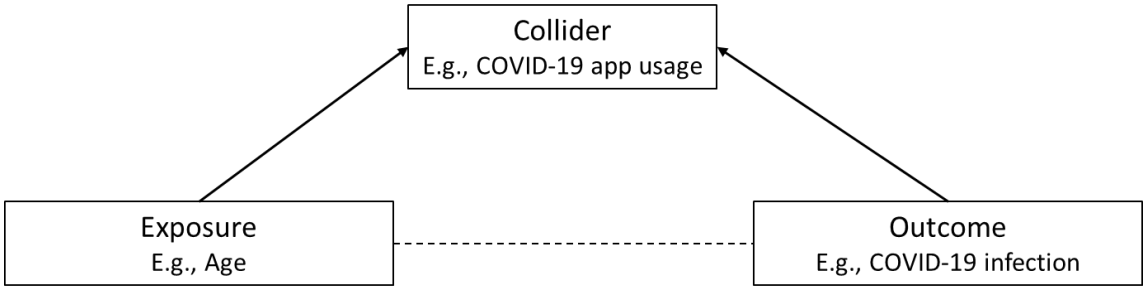
Figure 1.1. Causal diagram illustrating confounding with an example relevant to this thesis.



Differences between the treated and untreated groups in relation to a confounder will induce bias in the estimated treatment effect. In the example presented in **Figure 1.1.**, people who are prescribed sodium-glucose co-transporter 2 inhibitors (SGLT2i), a newer antidiabetic treatment, might be younger than people prescribed alternatives.[39] These younger people will have a lower risk of heart failure hospitalisation compared with older people. Thus, when we estimate the effect of SGLT2i prescribing on heart failure hospitalisation, we would likely exaggerate any potential benefit of SGLT2i on heart failure hospitalisation due to confounding by indication.

A collider is a variable which is directly influenced by two variables: the exposure/treatment and the outcome (**Figure 1.2.**).

Figure 1.2. Causal diagram illustrating collider bias described in a previous study.[40]



Conditioning on a collider, either by adjustment in a multivariable model or by the selection of a study population, may induce a spurious, non-causal association between the exposure/treatment and the outcome. In the example presented in **Figure 1.2.**, being younger increases the probability of using the COVID-19 app (the collider). Separately, being infected with COVID-19 increases the probability of using the COVID-19 app (the collider). If the study population is selected conditional on the collider,

then a spurious association will be observed between younger age and increased risk of COVID-19 infection.[40]

1.2.3. Main assumptions in causal inference

Several assumptions are implicit when applying these causal frameworks to estimate causal treatment effects. These assumptions rely on careful justification as they cannot be fully tested.[37] I briefly summarise these assumptions in **Table 1.1.**

Table 1.1. Causal inference assumptions and brief description.[37]

Causal inference assumption	Brief description
Exchangeability	No confounding.[41]
Non-interference	No spill-over effects – outcome in one individual is influenced by their own treatment and not the treatment of others.[35, 42]
Consistency	Exposure is defined with enough specificity that different variants of the exposure will not have different effects on the outcome.[43]
Positivity	There must be variation in treatment in all covariate strata (i.e., being treated or untreated is possible at every combination of covariates).[44]

The exchangeability assumption requires that there are no confounders biasing the treatment effect.[41] Exchangeability is usually achieved through randomisation in RCTs. In observational studies, exchangeability is more challenging to achieve, and usually relies on the assumption of no unmeasured confounding after analyses have accounted for measured confounders.[37]

The non-interference assumption requires that the potential outcomes of one person do not depend on the treatment status of another person.[35, 42] This assumption can be violated for treatments like vaccines, where the vaccination status of one individual can impact outcomes in another.[37]

The consistency assumption requires that the treatment is defined well-enough such that a change in the level of the same treatment does not change the potential outcome.[43] This assumption can be violated where different dosages of a drug lead to different outcomes, in which case an average treatment effect across dosages is measured.[37] The non-interference and consistency assumptions are referred to together as the stable unit treatment value assumption.

Finally, the positivity assumption requires that each level of the treatment is possible across all individuals and across all strata of covariates in an analysis.[44] This assumption can be violated by

chance or systematically when, for instance, a treatment is contraindicated in a particular group of people included in a pharmacoepidemiological study.[37]

These assumptions are fundamental to causal inference and must be considered when evaluating treatment effects. However, these assumptions are not exhaustive in pharmacoepidemiological studies seeking to make causal inferences. Researchers are most often forced to make additional assumptions depending on the study design, methods, and analyses selected.

1.2.4. Causal inference in RCTs

As described in section 1.1.3., an RCT is considered the ideal design to infer causal treatment effects because it removes confounding through the random allocation of the treatment-of-interest.[45] This random allocation ensures comparability/exchangeability of treatment groups, meaning they have similar distributions of both observed and unobserved confounders on average. This random allocation also ensures the positivity assumption is met so long as the study population is sufficiently large, since the random treatment assignment is possible across all subgroups included in the study. Other strengths of the RCT design, namely double (or triple) blinding of the treatment allocation to (i) the participant, (ii) the managing clinical team, and (iii) the analysts, ensures that the randomisation itself does not influence the outcome except via the treatment.

RCTs generally estimate the intention-to-treat (ITT) effect.[46] This represents the average effect of being allocated to the treatment group, rather than estimating the effect of actually taking the assigned treatment. This analysis strategy preserves the exchangeability of treatment groups, thereby minimising the risk of confounders biasing the causal estimand. If compliance to the randomised treatment is high, the ITT effect should be similar to the true ATE.[47]

Despite these benefits, RCTs suffer from several limitations and challenges. These studies are expensive, require long follow-up of study participants,[48] may use different comparators which are not standard of care in the NHS, and may include study populations which are not representative of the general population being treated in the NHS.[1] Generalisability of study results to those excluded, such as people with advanced chronic illness, the elderly, and people living with advanced frailty or more than one chronic condition, is therefore challenging.

These limitations introduce clinical uncertainty for clinicians, healthcare commissioners, and policymakers when making healthcare decisions for diverse patient populations with complex healthcare needs.[1] The increasing availability of observational data can be useful in generating evidence to address evidence gaps unanswered by RCT data.

1.2.5. Causal inference in observational pharmacoepidemiological studies using routinely collected health data

Observational studies using routinely collected health data to study comparative treatment effects are increasingly being used to complement randomised evidence.[1] However, these studies almost always suffer from important biases, such as unmeasured confounding and collider bias. These biases must be considered in the design and interpretation of observational studies seeking to make causal inferences.[45]

In 2022, NICE published a 'living' guideline titled the "NICE real-world evidence framework".[1] This framework "aims to improve the quality of real-world evidence informing [NICE] guidance" and highlights best practices in planning, conducting, and reporting studies using real-world data.[1] This framework outlines principles which should be followed when conducting observational studies using routinely collected health data, which include (i) "ensuring data is of good provenance, relevant and of sufficient quality to answer the research question", (ii) "generate evidence in a transparent way and with integrity from study planning through to study conduct and reporting", and (iii) "use analytical methods that minimise the risk of bias and characterise uncertainty".[1] This guidance is a useful tool that researchers can use to generate high-quality observational evidence which can potentially inform guideline development and clinical practice. I will organise my commentary on the conduct of observational studies using routinely collected health data according to these NICE principles.[1]

(i) Ensuring data is of good provenance, relevant and of sufficient quality to answer the research question

An important first step in designing a pharmacoepidemiological study is illustrating the causal question using a casual diagram, such as a DAG. This is important to clearly outline the casual assumptions about the relationship between the treatment, the outcome, and other measured and unmeasured variables (e.g., confounders, colliders) which must be considered in the study design. Next the researcher must identify suitable data sources to define the study population, treatment, outcome, and covariates, with consideration of potential limitations that may lead to bias (e.g., misclassification (i.e., information bias), selection bias, confounding).

(ii) Generate evidence in a transparent way and with integrity from study planning through to study conduct and reporting

To increase the credibility and rigour of the observational research using routinely collected health data, researchers should aim for transparency in the conduct and reporting of analyses. This can be achieved by publishing protocols and statistical analysis plans (SAP) which pre-specify analyses and provide justification for any amendments.

(iii) Use analytical methods that minimise the risk of bias and characterise uncertainty

When specifying the study design, researchers must choose between several methods and analysis strategies (or often a combination) to infer causal treatment effects from routinely collected health data. There are many study design strategies and analyses which can reduce the risk of confounding when applied to observational data to study causal exposure or treatment effects. These include prospective cohorts (e.g., occupational cohorts),[49] traditional multivariable regression, propensity score analyses,[50, 51] and natural or quasi-experiments, such as studies using regression discontinuity,[52] interrupted time series,[53] difference in differences,[54] and instrumental variable (IV) analyses,[55] including mendelian randomisation.[56]

The NICE real-world evidence framework emphasises the strength of using a target trial emulation (TTE) strategy to generate evidence to inform HTA and clinical guidelines.[1] This is a framework for causal inference, which must be combined with other study design and analysis strategies to reduce the risk of common biases in pharmacoepidemiological research, including confounding bias.

The TTE framework

While the concept of considering the ideal RCT when designing an observational study is not new,[57, 58] the TTE framework was only first proposed in 2016.[59] This framework involves a two-step process:[59, 60]

1. The researcher clearly specifies the causal research question. Following this, the researcher can then specify the ideal trial design and how it will be adapted to suit the observational data in which the study will take place. The ideal trial design usually focuses on the following key aspects of a trial: eligibility criteria, treatment strategy and assignment, follow-up including the specification of time 0 (baseline) and the end of follow-up, the outcomes of interest, causal contrasts, and the analysis plan.
2. The researcher then applies this target trial design in the observational data. Successful application of the TTE framework can minimise common biases in observational research like immortal time bias. [59-62] However, the TTE framework must be combined with other analysis strategies for causal inference to minimise the risk of confounding bias.

In **Table 1.2.**, I present a non-exhaustive list of methods and analysis strategies which can be combined with the TTE framework to study causal treatment effects in routinely collected health data.

Table 1.2. Summary of common methods and analysis strategies which can be combined with the target trial emulation (TTE) framework to study causal treatment effects using routinely collected health data.

Method/Analysis Strategy	Brief description	Confounding addressed
Negative control outcome	A secondary outcome is selected that shares the same confounding structure as the exposure and outcome-of-interest but has no known causal relationship with the exposure of interest. The association between the exposure-of-interest and the negative-control outcome is measured using the same methods and analysis as the primary outcome. If a null effect is found for the association between the exposure and negative-control outcome, it supports the strategy to minimise the risk of confounding bias (and the assumption of no unmeasured confounding) in the main comparative effectiveness analysis.	Measured and, to some extent, unmeasured confounding.
Traditional multivariable regression	Measured confounders are included as covariates in a multivariable model. This assumes no unmeasured confounding.	Measured confounding only.
Propensity scores[50, 51]	A propensity score is estimated, describing the conditional probability of treatment given a set of measured covariates. The propensity score can then be used to improve the exchangeability of the treatment groups on measured confounders by inverse probability of treatment weighting (IPTW), matching, stratification, or regression adjustment. This method assumes no unmeasured confounding.	Measured confounding only.
Instrumental variable (IV) analysis[55]	An 'instrument' independent of the confounding framework of the exposure and outcome under study is used to estimate the causal treatment effect. This is akin to an RCT, where the 'instrument' is the random allocation of treatment. The more likely the IV assumptions are met, the more likely that this analysis will minimise confounding bias in the treatment effect estimation.	Measured and unmeasured confounding.
Epidemiological triangulation[63]	Causal treatment effects from different studies subject to different biases or which make different assumptions are compared. Consistency in the treatment effect across these different sources will support the causal conclusions.	It depends on the types of methods and analyses used in the triangulated studies.

The methods and analysis strategies presented in **Table 1.2.** are complementary. A pharmacoepidemiological study should consider combinations of these methods within the TTE framework to understand biases or violations of key assumptions which could undermine the conclusions of a single study, as well as triangulating results across different studies which apply different strategies subject to alternative forms of bias. For example, pharmacoepidemiological studies often use traditional multivariable regression or a propensity score analysis to estimate causal treatment effects. These analyses assume no unmeasured confounding, either by imperfectly measured variables (e.g., smoking status measured by self-report, body-mass index (BMI) using outdated measurements) and unmeasured variables (e.g., diet, frailty, caregiver support). This assumption may not always be appropriate, particularly in routinely collected health data. These data are not collected for research purposes and will exclude pertinent confounder information which likely bias any estimated treatment effect.

Thus, addressing questions of comparative treatment effectiveness using an IV analysis similarly applied within the TTE framework can provide valuable evidence on the comparative effectiveness of treatments which does not assume no unmeasured confounding. This type of analysis is increasingly used to create 'natural experiments' in epidemiological analyses to estimate causal effects.[55, 56, 64]

1.3. IV analyses in pharmacoepidemiology

An IV analysis is a useful tool that can take advantage of a natural experiment to infer causal effects. In the context of pharmacoepidemiology, an IV analysis is possible where there is some degree of exogenous variation in treatment uptake across the study population of interest. This exogenous variation can be exploited to estimate treatment effects independent of measured and unmeasured confounders.

In other words, the IV analysis seeks to find a randomised experiment within observational data, by (i) defining a valid instrument; (ii) using this instrument to extract variation in the treatment received that is independent of confounders (exogenous); and (iii) applying this unbiased variation to estimate the causal effect of the exposure/treatment.[55, 65] Popular in econometrics, this analysis is gaining increasing interest in epidemiology, (e.g., mendelian randomisation studies[66, 67]) as well as in other pharmacoepidemiology and health services research.[55, 68, 69]

The major challenge of IV analyses is selecting an appropriate instrument which meets four key assumptions listed in **Table 1.3.**[55, 64] Of these four assumptions, only the relevance assumption can be formally tested. Researchers cannot use the same routinely collected health data as in the IV analysis to test the other assumptions.[70] Instead, researchers usually rely on careful rationale and justification

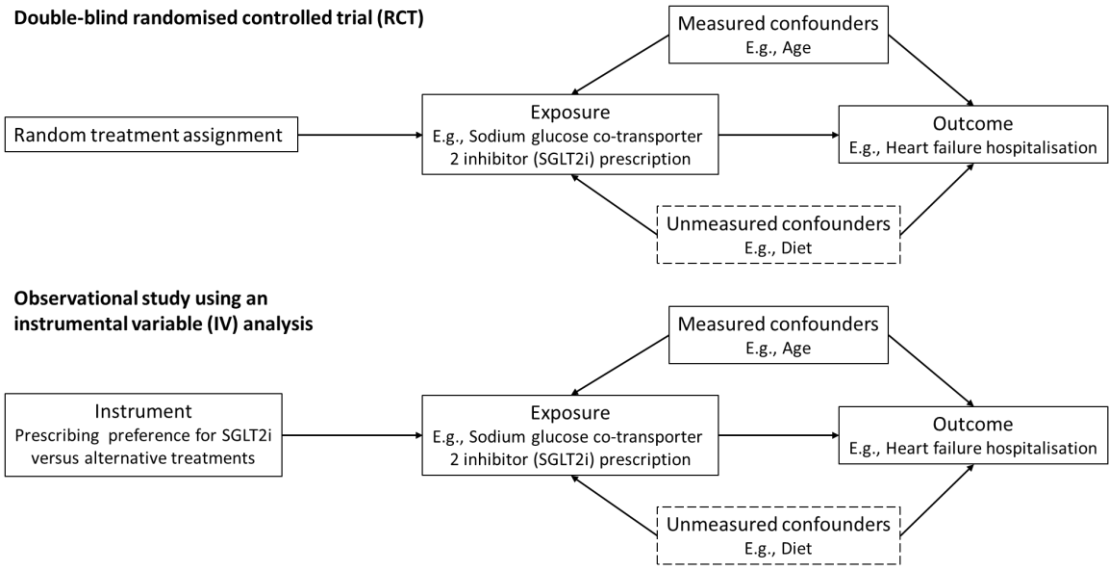
on how the instrument can meet the other key assumptions when using this analysis to estimate causal treatment effects.

Table 1.3. Key assumptions in an instrumental variable (IV) analysis.

Assumption	Details
1. The relevance assumption	The instrument must cause a change in the treatment received.
2. The exchangeability assumption	The instrument must be independent of unmeasured confounders.
3. The exclusion restriction	The instrument only affects the outcome via the treatment.
4. The monotonicity assumption	An increase in the level of the instrument always results in a higher or equal level of treatment assignment (i.e., no 'defiers').

An IV analysis is akin to an RCT (**Figure 1.3.**). In this example, we are interested in the causal effect of SGLT2i (the treatment) on heart failure hospitalisation (the outcome). Here, study participants are randomly allocated to either SGLT2i treatment or the control group (often placebo, added to standard therapy). The primary analysis of this hypothetical RCT would then estimate the ITT effect – the effect of being randomly allocated, but not necessarily taking, SGLT2i versus placebo. This random allocation can be thought of as the 'instrument'.

Figure 1.3. Directed acyclic graph (DAG) comparing a hypothetical randomised controlled trial (RCT) with a hypothetical instrumental variable (IV) analysis.



Here, the relevance assumption is met, since being randomly assigned to SGLT2i would strongly predict taking SGLT2i. The exchangeability assumption is also met, since the random allocation will be independent of confounders, measured or unmeasured, except due to chance. The exclusion restriction assumption would be met in a double (or triple) blinded RCT, where the study participant, the healthcare team, and the analysts are 'blinded' (i.e., unaware of) to the actual treatment allocation. This blinding ensures that the randomisation procedure (the 'instrument') would only affect the outcome via the treatment actually taken. Finally, the monotonicity assumption in an RCT is also usually met through double-blinding the treatment allocation. This assumption would be violated if a subgroup of people randomised to the treatment 'defied' this randomisation by dropping out of the study or taking the alternative treatment to what was randomly allocated. I describe this assumption in more detail in thesis section 2.4.4.

An IV analysis in an observational study shares a similar structure to the RCT (**Figure 1.3.**) – but defining an instrument that meets the necessary assumptions is challenging. Several different types of instruments can be used in observational pharmacoepidemiology and health services research (**Table 1.4**).

Table 1.4. Examples of pharmaco-epidemiological and health services research exposures/treatments of interest and their corresponding instrument in IV analyses

Exposure/treatment of interest	Instrument	Citation
COX-2 selective nonsteroidal anti-inflammatory drugs (NSAID)	Physician/provider preference for treatment[71]	Davies et al, 2013[68]
Neonatal intensive care unit	Distance/time to travel to treatment centre	Lorch et al, 2012[72]
Prompt admission to critical care ward	Critical care bed occupancy	Harris et al, 2018[69]

The example in **Figure 1.3.** uses a preference-based instrument, where the prescribing preference of the healthcare provider is used as the instrument. This type of instrument seeks to use differences in the prescribing preferences at the level of an individual prescriber (e.g., the general practitioner) or a group of prescribers (e.g., a group of GPs) to create a natural experiment. In this thesis, I use a preference-based instrument to study comparative treatment effects.

In the example in **Figure 1.3.**, the preference for SGLT2i versus the alternative treatment is defined as the instrument. Because this is a latent variable which cannot be directly measured, we would need to measure this preference indirectly. A common method is to use the prescribing history of the prescriber, measured as the proportion of people prescribed drug A (in this case, SGLT2i) versus drug B (in this case, for example, dipeptidyl peptidase-4 inhibitors (DPP4i), an alternative antidiabetic treatment with similar indications for blood glucose lowering[73]). We could ensure this preference strongly predicts the treatment actually received to meet the relevance assumption of an IV analysis. We could then observe whether the preference-based instrument is associated with the measured confounders (e.g., age) to explore the exchangeability assumption of an IV analysis. However, we cannot evaluate if the preference for prescribing SGLT2i versus DPP4i is associated with unmeasured confounders (e.g., diet). We also cannot know if the preference for prescribing SGLT2i versus DPP4i acts on heart failure hospitalisation via a causal path independent of the treatment actually prescribed to meet the exclusion restriction assumption of an IV analysis. And finally, we cannot observe counterfactual treatment prescribing across different prescribers with different prescribing preferences using routinely collected health data to confirm the monotonicity assumption of an IV analysis.

Clearly, there are challenges to selecting an appropriate instrument in an observational study which meets the key assumptions of an IV analysis. But with careful scrutiny and appropriate epidemiological triangulation, observational studies using an IV analysis within the TTE framework have the potential to generate high-quality observational evidence useful to HTA and policymakers.[1] This evidence can be

most useful in circumstances where a lack of RCT evidence for a particular treatment decision leads to substantial exogenous variation in treatment.

1.4. Evidence gaps in clinical treatment decisions: two examples in United Kingdom (UK) primary and secondary care

1.4.1. T2DM second-line oral antidiabetic treatment prescribed in primary care

T2DM is a metabolic disorder caused by insulin resistance and insufficient insulin secretion which leads to persistently high levels of blood glucose. T2DM is one of the most common chronic diseases in the UK and can result in many long term complications including cardiovascular disease (CVD), kidney disease, neuropathy, eye disease, and death.[74] People with T2DM generally manage their condition with their GP in primary care.[74]

NICE guidelines recommend metformin monotherapy as first-line oral antidiabetic treatment after diet and lifestyle modifications.[73] Prior to 2022, these guidelines recommended a choice of three main second-line treatment options for treatment intensification if a patient's blood glucose (haemoglobin A1c, HbA1c) is not adequately controlled by metformin monotherapy: (1) sulfonylureas (SU); (2) DPP4i; and (3) SGLT2i, all in combination with metformin.[73] Due to the lack of direct RCT comparisons of these three second-line treatments, NICE guidelines did not clearly recommend a particular treatment.

More recently, placebo-controlled CVD-safety trials demonstrated that the newer SGLT2i drug class has cardio- and kidney-protective effects among people with and without T2DM.[75-79] Updated NICE guidelines, published in 2022, recommend SGLT2i for people with chronic heart failure or established CVD in addition to metformin as first-line treatment.[73] However, up to this point, with the lack of clear recommendations on which drug should be given to which patients, second-line treatment decision-making would likely have been strongly influenced by patient preference (among well-informed patients), GP awareness of drug benefits for particular patients, GP personal preference, or directives from the management of groups of GPs which were known as CCGs prior to 2022 (see section 1.1.1.). An example where directives from the CCG may have led to exogenous treatment variation in second-line antidiabetic treatment prescribed includes the recommendation from the London South East CCG that SU should be preferably prescribed as second-line antidiabetic treatment added to metformin monotherapy.[80] These decisions could be made by the CCG for a variety of reasons, including budgets, since SGLT2i are under patent and are therefore more expensive treatments compared with SU and DPP4i.[81]

Previous research has observed substantial variation in second-line oral antidiabetic treatment prescribing at the CCG-level across England.[82] Further research to understand potential sources of this variation, and how this variation could be applied in a natural experiment using an IV analysis, would be useful to inform guidelines and health services.

1.4.2. AMI treatment strategies in people with reduced kidney function in secondary care

AMI is an acute manifestation of coronary heart disease where the blood supply to the heart is disrupted, leading to insufficient oxygen supply to the cardiac muscles.[83] NICE guidelines recommend two main early management treatment strategies for AMI which are provided in secondary care: (i) an invasive cardiac strategy, where cardiac imaging (angiography) and follow-up percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) is provided, or (ii) conservative management, where ticagrelor with aspirin, clopidogrel with aspirin, or aspirin alone can may be offered to non-invasively manage the AMI.[84]

Among those hospitalised with ST-elevated myocardial infarction (STEMI) – usually a complete blockage of a coronary blood vessel[85] – NICE guidelines recommend that almost all individuals should be treated with an invasive cardiac strategy to restore oxygen supply to the heart.[84] Among those hospitalised with non-ST elevated myocardial infarction (NSTEMI) – caused by several factors such as arterial narrowing or a partially blocked cardiac blood vessel[85] – NICE again recommends an invasive management strategy if predicted mortality is intermediate or high risk.[84] However, these guidelines also advise caution in providing invasive treatment strategies to people at higher risks of complications due to bleeding or comorbidities.[84]

Contrast-induced acute kidney injury (AKI) during an angiography[86, 87] is considered a risk for complications from invasive cardiac management, particularly for people with chronic kidney disease (CKD) and kidney impairment, who represent approximately 40% of those people hospitalised for AMI.[88-90] Because RCTs which demonstrated the benefits of invasive cardiac management have largely excluded people with CKD,[91] the balance of the benefits versus risks of invasive versus conservative cardiac management is unclear in this vulnerable patient population. This likely contributes to observed disparities in invasive cardiac management for AMI, particularly for NSTEMI, among people with CKD.[89, 92-94]

This evidence gap in RCT data has been addressed by several observational studies. These studies found evidence to suggest people with CKD might benefit from invasive cardiac management strategies.[92-95] However, these studies apply observational methods such as multivariable regression and propensity scores which are prone to bias from unmeasured confounding. In the persistent

absence of RCT data, additional observational evidence to triangulate these findings is needed to strengthen the observational evidence base and support AMI guideline development.

Because people with CKD are usually older, with multiple long-term conditions,[96] they are particularly unsuited to the organ-specific healthcare delivery structure of NHS secondary care services. Data capture in secondary care EHRs may therefore be different according to CKD status – patients with competing long-term illnesses may not be managed by a cardiologist and are therefore not captured by cardiac-specific audits. Furthermore, not all hospitals in England provide cardiac interventions, and may transfer patients between hospitals. It is therefore important to consider the suitability of different secondary care data sources to define the study population admitted to hospital for AMI with kidney impairment before any comparative effectiveness work. Once the suitability of the secondary care data is confirmed, these data could be used to study the comparative effectiveness of invasive AMI interventions among people with CKD in a natural experiment using IV analyses.

1.5. Rationale, aims, and objectives

1.5.1. Thesis rationale

While RCT evidence has driven major improvements in treatments and outcomes for T2DM and AMI, two diseases which are a major burden in the UK population and to the NHS,[73, 84, 97] there are outstanding evidence gaps due to (1) the lack of head-to-head comparisons of the three most commonly prescribed second-line oral antidiabetic treatment options and (2) the exclusion of people with CKD in RCTs investigating alternative AMI treatments.[91] Using routinely collected health data and advanced quantitative methods to study these unanswered clinical questions is important in the absence of RCT evidence.

1.5.2. Thesis aim and objectives

Aim

This thesis aims to advance the use of routinely collected health data to study the comparative effectiveness of treatments. To address this aim, I include two case studies set in England which address important areas of clinical uncertainty. These case studies are as follows:

Objectives

Case Study 1: Alternative second-line oral antidiabetic treatments among people with type 2 diabetes mellitus (T2DM) in primary care

1A: To examine inequalities in second-line antidiabetic treatment prescribing according to sociodemographic characteristics which are likely to be important potential confounders when studying the comparative effectiveness of these drugs (Chapter 3).

1B: To inform the selection of potential instruments for natural experiments comparing alternative second-line antidiabetic treatments by investigating treatment variation at the clinical commissioning group (CCG) level (Chapters 3 to 5).

1C: To design and conduct an IV analysis to estimate the relative effectiveness of alternative second-line antidiabetic treatments with respect to outcomes important to patients, healthcare providers, and policymakers (Chapter 4 to 5).

1D: To investigate heterogeneous treatment effects across the target population of people with T2DM in English primary care (Chapter 6).

Case Study 2: Alternative acute myocardial infarction (AMI) treatments among people with kidney impairment in secondary care

2A: To investigate potential biases in defining a study population of people hospitalised for AMI with reduced kidney function in English secondary care using primary and secondary care data sources (Chapter 7).

2B: To examine inequalities in AMI treatment and outcomes by kidney function (Chapter 8).

2C: To explore variation in AMI treatment strategies at the individual and cardiology centre-level to inform the selection of a preference-based instrument for future comparative effectiveness studies using an IV analysis (Chapter 9).

1.6. Thesis structure

In this chapter, I presented the introduction to this thesis, including background information describing routinely collected health data from the UK and its relevance to HTA, causal inference in pharmacoepidemiology including the TTE framework and IV analyses, and important evidence gaps in treatment decision-making in T2DM and AMI care. I then concluded with the rationale, aim, and objectives for this thesis.

In Chapter 2, I present an overview of the general methods used across the two case studies included in this thesis. This methods chapter will include details about the research projects which funded this work, a summary of the approach I used to study the comparative effectiveness of treatments using routinely collected health data, the data sources for each case study, and details of the statistical models I used in the research papers included in this thesis.

In Chapters 3 to 6, I present 4 research papers (2 published,[98, 99] 1 *in press*, and 1 in submission at a peer-reviewed journal) all focused on alternative second-line oral antidiabetic treatments for people with T2DM in English primary care (Case Study 1).

In Chapters 7 to 9, I present 3 research papers (2 published[100, 101] and 1 draft manuscript) which focus on alternative AMI treatment strategies among people with kidney impairment in English secondary care (Case Study 2).

I begin each of these chapters with an overview of the research paper, followed by the research paper in its published, accepted, or draft form. A cover sheet accompanies each research paper, signed by the senior author, on which I detail my involvement and contributions to the paper and its publication status. I conclude each of these chapters with a brief discussion summarising why the research paper is relevant to my thesis. Key tables and figures from the supplementary materials of each paper are included within the 'Relevance to my thesis' discussions following each research paper. Further supplementary materials referenced in the main text are provided in the thesis appendix.

Finally, in Chapter 10, I include a discussion of the thesis, which summarises the main findings of each case study, the original scientific contributions I have made in this thesis, the limitations of my research, and reflections on my personal learning and development over the course of my PhD registration period. In the final section of the discussion, I conclude this thesis.

References are presented in the JAMA style and are listed at the end of each chapter.

1.7. References

1. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

2. Real-World Evidence. Web. U.S. Food & Drug Administration (FDA). Accessed 4 April, 2024. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

3. Dang A. Real-World Evidence: A Primer. *Pharmaceut Med*. Jan 2023;37(1):25-36. doi:10.1007/s40290-022-00456-6

4. Hulme WJ, Williamson EJ, Green ACA, et al. Comparative effectiveness of ChAdOx1 versus BNT162b2 covid-19 vaccines in health and social care workers in England: cohort study using OpenSAFELY. *BMJ (Clinical research ed)*. 2022;378:e068946. doi:10.1136/bmj-2021-068946

5. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020/08/01 2020;584(7821):430-436. doi:10.1038/s41586-020-2521-4

6. Ghosh RE, Crellin E, Beatty S, Donegan K, Myles P, Williams R. How Clinical Practice Research Datalink data are used to support pharmacovigilance. *Ther Adv Drug Saf*. 2019;10:2042098619854010. doi:10.1177/2042098619854010

7. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ (Clinical research ed)*. Apr 7 2021;373:n826. doi:10.1136/bmj.n826

8. Greengross PG, Ken; Collini, Elizabeth. *The history and development of the UK National Health Service: 1948-1999*. 1999. https://assets.publishing.service.gov.uk/media/57a08d91e5274a31e000192c/The-history-and-development-of-the-UK-NHS.pdf

9. The healthcare ecosystem. Web. NHS England. Accessed 4 April, 2024. https://digital.nhs.uk/developer/guides-and-documentation/introduction-to-healthcare-technology/the-healthcare-ecosystem

10. Patients Registered at a GP Practice. Web. NHS Digital. Accessed 3 March, 2022. https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice

11. Clinical commissioning groups (CCGs). Web. NHS England. Accessed 3 March, 2022. https://www.england.nhs.uk/commissioning/who-commissions-nhs-services/ccgs/

12. What are integrated care systems? Web. NHS England. Accessed 4 April, 2024. https://www.england.nhs.uk/integratedcare/what-is-integrated-care/

13. Whitty CJM, MacEwen C, Goddard A, et al. Rising to the challenge of multimorbidity. *BMJ (Clinical research ed)*. 2020;368:l6964. doi:10.1136/bmj.l6964

14. Kingston A, Robinson L, Booth H, Knapp M, Jagger C, project M. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. *Age and ageing*. 2018;47(3):374-380.

15. Walley T, Mantgani A. The UK General Practice Research Database. *The Lancet*. 1997/10/11/ 1997;350(9084):1097-1099. doi:https://doi.org/10.1016/S0140-6736(97)04248-7

16. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. Jun 2015;44(3):827-36. doi:10.1093/ije/dyv098

17. GP IT Futures systems and services. Web. NHS Digital. Updated 30 September 2021. Accessed 2 March, 2022. https://digital.nhs.uk/services/gp-it-futures-systems

18. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology*. 2019;48(6):1740-1740g. doi:10.1093/ije/dyz034

19. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *International journal of epidemiology*. 2017;46(4):1093-1093i. doi:10.1093/ije/dyx015

20. Clinical audit. Web. NHS England. Accessed 4 April, 2024. https://www.england.nhs.uk/clinaudit/

21. Clinical audits and registries. Web. NHS England. Accessed 4 April, 2024. https://digital.nhs.uk/data-and-information/clinical-audits-and-registries

22. Ray D, Roebuck C, Smith O. *Delivering linked datasets to support health and care delivery and research*. 2018. https://digital.nhs.uk/services/data-access-request-service-dars/linked-datasets-supporting-health-and-care-delivery-and-research

23. Mourby MJ, Doidge J, Jones KH, et al. Health Data Linkage for UK Public Interest Research: Key Obstacles and Solutions. *Int J Popul Data Sci*. Apr 2 2019;4(1):1093. doi:10.23889/ijpds.v4i1.1093

24. Mathur R, Rentsch CT, Morton CE, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet (London, England)*. May 8 2021;397(10286):1711-1724. doi:10.1016/s0140-6736(21)00634-6

25. Bhaskaran K, Dos-Santos-Silva I, Leon DA, Douglas IJ, Smeeth L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. *Lancet Diabetes Endocrinol*. Dec 2018;6(12):944-953. doi:10.1016/s2213-8587(18)30288-2

26. Sydes MR, Barbachano Y, Bowman L, et al. Realising the full potential of data-enabled trials in the UK: a call for action. *BMJ Open*. Jun 16 2021;11(6):e043906. doi:10.1136/bmjopen-2020-043906

27. Dexamethasone in Hospitalized Patients with Covid-19. *New England Journal of Medicine*. 2021;384(8):693-704. doi:doi:10.1056/NEJMoa2021436

28. Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol*. Dec 1 2016;45(6):1776-1786. doi:10.1093/ije/dyv341

29. Hill AB. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc R Soc Med*. May 1965;58(5):295-300.

30. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol*. 2015;12:14. doi:10.1186/s12982-015-0037-4

31. VanderWeele TJ. Commentary: On Causes, Causal Inference, and Potential Outcomes. *Int J Epidemiol*. Dec 1 2016;45(6):1809-1816. doi:10.1093/ije/dyw230

32. Montastruc JL, Benevent J, Montastruc F, et al. What is pharmacoepidemiology? Definition, methods, interest and clinical applications. *Therapie*. Apr 2019;74(2):169-174. doi:10.1016/j.therap.2018.08.001

33. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*. 1974;66(5):688-701. doi:10.1037/h0037350

34. Splawa-Neyman J, Dabrowska DM, Speed TP. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*. 1990;5(4):465-472.

35. Rubin DB. Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association*. 2005/03/01 2005;100(469):322-331. doi:10.1198/016214504000001880

36. Holland PW. Statistics and Causal Inference. *Journal of the American Statistical Association*. 1986/12/01 1986;81(396):945-960. doi:10.1080/01621459.1986.10478354

37. Erik I, Peter C, Jim L, John L, Anna P, Srinivasa Vittal K. Causal inference and effect estimation using observational data. *Journal of Epidemiology and Community Health*. 2022;76(11):960. doi:10.1136/jech-2022-219267

38. Greenland S, Pearl J, Robins JM. Causal Diagrams for Epidemiologic Research. *Epidemiology (Cambridge, Mass)*. 1999;10(1)

39. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

40. Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun*. Nov 12 2020;11(1):5749. doi:10.1038/s41467-020-19478-2

41. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health*. Jul 2006;60(7):578-86. doi:10.1136/jech.2004.029496

42. VanderWeele TJ, Hernán MA. Causal Inference Under Multiple Versions of Treatment. *J Causal Inference*. May 1 2013;1(1):1-20. doi:10.1515/jci-2012-0002

43. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology (Cambridge, Mass)*. Nov 2009;20(6):880-3. doi:10.1097/EDE.0b013e3181bd5638

44. Petersen ML, Porter KE, Gruber S, Wang Y, van der Laan MJ. Diagnosing and responding to violations in the positivity assumption. *Stat Methods Med Res*. Feb 2012;21(1):31-54. doi:10.1177/0962280210386207

45. Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *The Lancet*. 2001/02/03/ 2001;357(9253):373-380. doi:https://doi.org/10.1016/S0140-6736(00)03651-5

46. *ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials*. 2020. 17 Februrary. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf

47. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology (Cambridge, Mass)*. Jul 2006;17(4):360-72. doi:10.1097/01.ede.0000222409.00878.37

48. Reith C, Landray M, Devereaux PJ, et al. Randomized Clinical Trials — Removing Unnecessary Obstacles. *New England Journal of Medicine*. 2013/09/12 2013;369(11):1061-1065. doi:10.1056/NEJMsb1300760

49. Pearce N, Vandenbroucke JP, Lawlor DA. Causal Inference in Environmental Epidemiology: Old and New Approaches. *Epidemiology (Cambridge, Mass)*. May 2019;30(3):311-316. doi:10.1097/ede.0000000000000987

50. Stürmer T, Wyss R, Glynn RJ, Brookhart MA. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *Journal of internal medicine*. Jun 2014;275(6):570-80. doi:10.1111/joim.12197

51. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ (Clinical research ed)*. 2019;367:l5657. doi:10.1136/bmj.l5657

52. Venkataramani AS, Bor J, Jena AB. Regression discontinuity designs in healthcare research. *BMJ (Clinical research ed)*. 2016;352:i1216. doi:10.1136/bmj.i1216

53. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol*. Feb 1 2017;46(1):348-355. doi:10.1093/ije/dyw098

54. Wing C, Simon K, Bello-Gomez RA. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annu Rev Public Health*. Apr 1 2018;39:453-469. doi:10.1146/annurev-publhealth-040617-013507

55. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. Jun 15 2014;33(13):2297-340. doi:10.1002/sim.6128

56. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nature Reviews Methods Primers*. 2022/02/10 2022;2(1):6. doi:10.1038/s43586-021-00092-5

57. Dorn HF. Philosophy of inferences from retrospective studies. *Am J Public Health Nations Health*. Jun 1953;43(6 Pt 1):677-83. doi:10.2105/ajph.43.6_pt_1.677

58. Pearce N, Vandenbroucke JP. Are Target Trial Emulations the Gold Standard for Observational Studies? *Epidemiology (Cambridge, Mass)*. Sep 1 2023;34(5):614-618. doi:10.1097/ede.0000000000001636

59. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

60. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

61. Fu EL. Target Trial Emulation to Improve Causal Inference from Observational Data: What, Why, and How? *Journal of the American Society of Nephrology*. 2023;34(8)

62. Yadav K, Lewis RJ. Immortal Time Bias in Observational Studies. *Jama*. 2021;325(7):686-687. doi:10.1001/jama.2020.9151

63. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. Dec 1 2016;45(6):1866-1886. doi:10.1093/ije/dyw314

64. Keele L, Small D. Instrumental variables: Don't throw the baby out with the bathwater. *Health Serv Res*. 2019;54(3):543-546. doi:10.1111/1475-6773.13130

65. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. *Emerging Themes in Epidemiology*. 2018/01/22 2018;15(1):1. doi:10.1186/s12982-018-0069-7

66. Walker VM, Davey Smith G, Davies NM, Martin RM. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. *International Journal of Epidemiology*. 2017;46(6):2078-2089. doi:10.1093/ije/dyx207

67. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafò MR. Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates.

Economics & Human Biology. 2014/03/01/ 2014;13:99-106. doi:https://doi.org/10.1016/j.ehb.2013.12.002

68. Davies NM, Smith GD, Windmeijer F, Martin RM. COX-2 selective nonsteroidal anti-inflammatory drugs and risk of gastrointestinal tract complications and myocardial infarction: an instrumental variable analysis. *Epidemiology (Cambridge, Mass)*. May 2013;24(3):352-62. doi:10.1097/EDE.0b013e318289e024

69. Harris S, Singer M, Sanderson C, Grieve R, Harrison D, Rowan K. Impact on mortality of prompt admission to critical care for deteriorating ward patients: an instrumental variable analysis using critical care bed strain. *Intensive Care Medicine*. 2018/05/01 2018;44(5):606-615. doi:10.1007/s00134-018-5148-2

70. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and evaluation of the monotonicity condition for preference-based instruments. *Epidemiology (Cambridge, Mass)*. May 2015;26(3):414-20. doi:10.1097/ede.0000000000000279

71. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *The international journal of biostatistics*. 2007;3(1):Article 14. doi:10.2202/1557-4679.1072

72. Lorch SA, Baiocchi M, Ahlberg CE, Small DS. The differential impact of delivery hospital on the outcomes of premature infants. *Pediatrics*. Aug 2012;130(2):270-8. doi:10.1542/peds.2011-2820

73. NICE guideline [NG28]: Type 2 diabetes in adults: management. Web. NICE. Accessed 3 March, 2022. https://www.nice.org.uk/guidance/ng28/chapter/Recommendations#reviewing-drug-treatments

74. Diabetes - type 2. Web. National Institute for Health and Care Excellence. Accessed 5 April, 2024. https://cks.nice.org.uk/topics/diabetes-type-2/

75. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *New England Journal of Medicine*. 2019/06/13 2019;380(24):2295-2306. doi:10.1056/NEJMoa1811744

76. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *New England Journal of Medicine*. 2015/11/26 2015;373(22):2117-2128. doi:10.1056/NEJMoa1504720

77. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *New England Journal of Medicine*. 2017/08/17 2017;377(7):644-657. doi:10.1056/NEJMoa1611925

78. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine*. 2019/01/24 2018;380(4):347-357. doi:10.1056/NEJMoa1812389

79. Empagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2022;388(2):117-127. doi:10.1056/NEJMoa2204233

80. Type 2 Diabetes Mellitus in Adults: A guide for Southwark General Practice. South East London Integrated Care System NHS. Accessed 25 March, 2024. https://selondonccg.nhs.uk/wp-content/uploads/2021/11/CE-Diabetes-Guide-2021-v-4.3-1.pdf

81. Taylor SI. The High Cost of Diabetes Drugs: Disparate Impact on the Most Vulnerable Patients. *Diabetes care*. Oct 2020;43(10):2330-2332. doi:10.2337/dci20-0039

82. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

83. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *European heart journal*. Aug 25 2018;doi:10.1093/eurheartj/ehy462

84. NICE guideline [NG185]: Acute coronary syndrome. Web. NICE. Accessed March 10, 2022. https://www.nice.org.uk/guidance/ng185

85. Bhatt DL, Lopes RD, Harrington RA. Diagnosis and Treatment of Acute Coronary Syndromes: A Review. *Jama*. Feb 15 2022;327(7):662-675. doi:10.1001/jama.2022.0358

86. Barrett BJ, Parfrey PS. Clinical practice. Preventing nephropathy induced by contrast medium. *The New England journal of medicine*. Jan 26 2006;354(4):379-86. doi:10.1056/NEJMcp050801

87. Rear R, Bell RM, Hausenloy DJ. Contrast-induced nephropathy following angiography and cardiac interventions. *Heart (British Cardiac Society)*. 2016;102(8):638. doi:10.1136/heartjnl-2014-306962

88. Santopinto JJ, Fox KA, Goldberg RJ, et al. Creatinine clearance and adverse hospital outcomes in patients with acute coronary syndromes: findings from the global registry of acute coronary events (GRACE). *Heart (British Cardiac Society)*. Sep 2003;89(9):1003-8. doi:10.1136/heart.89.9.1003

89. Sofia Sederholm L, Joakim A, Karolina S, Mats F, Eva S. Prevalence and prognostic impact of chronic kidney disease in STEMI from a gender perspective: data from the SWEDEHEART register, a large Swedish prospective cohort. *BMJ Open*. 2015;5(6):e008188. doi:10.1136/bmjopen-2015-008188

90. Hanna EB, Chen AY, Roe MT, Saucedo JF. Characteristics and in-hospital outcomes of patients presenting with non-ST-segment elevation myocardial infarction found to have significant coronary

artery disease on coronary angiography and managed medically: stratification according to renal function. *Am Heart J*. Jul 2012;164(1):52-7.e1. doi:10.1016/j.ahj.2012.04.009

91. Konstantinidis I, Nadkarni GN, Yacoub R, et al. Representation of Patients With Kidney Disease in Trials of Cardiovascular Interventions: An Updated Systematic Review. *JAMA internal medicine*. Jan 2016;176(1):121-4. doi:10.1001/jamainternmed.2015.6102

92. Fox CS, Muntner P, Chen AY, et al. Use of evidence-based therapies in short-term outcomes of ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction in patients with chronic kidney disease: a report from the National Cardiovascular Data Acute Coronary Treatment and Intervention Outcomes Network registry. *Circulation*. 2010/01// 2010;121(3):357-365. doi:10.1161/circulationaha.109.865352

93. Bhatia S, Arora S, Bhatia SM, et al. Non-ST-Segment-Elevation Myocardial Infarction Among Patients With Chronic Kidney Disease: A Propensity Score-Matched Comparison of Percutaneous Coronary Intervention Versus Conservative Management. *J Am Heart Assoc*. Mar 10 2018;7(6)doi:10.1161/jaha.117.007920

94. Shaw C, Nitsch D, Steenkamp R, et al. Inpatient Coronary Angiography and Revascularisation following Non-ST-Elevation Acute Coronary Syndrome in Patients with Renal Impairment: A Cohort Study Using the Myocardial Ischaemia National Audit Project. *PLOS ONE*. 2014;9(6):e99925. doi:10.1371/journal.pone.0099925

95. Shaw C, Nitsch D, Lee J, Fogarty D, Sharpe CC. Impact of an Early Invasive Strategy versus Conservative Strategy for Unstable Angina and Non-ST Elevation Acute Coronary Syndrome in Patients with Chronic Kidney Disease: A Systematic Review. *PLOS ONE*. 2016;11(5):e0153478. doi:10.1371/journal.pone.0153478

96. Tonelli M, Wiebe N, Guthrie B, et al. Comorbidity as a driver of adverse outcomes in people with chronic kidney disease. *Kidney Int*. Oct 2015;88(4):859-66. doi:10.1038/ki.2015.228

97. Hex N, Bartlett C, Wright D, Taylor M, Varley D. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic medicine : a journal of the British Diabetic Association*. Jul 2012;29(7):855-62. doi:10.1111/j.1464-5491.2012.03698.x

98. Bidulka P, Mathur R, Lugo-Palacios DG, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes, Obesity and Metabolism*. 2023;25(1):282-292. doi:https://doi.org/10.1111/dom.14874

99. Bidulka P, O'Neill S, Basu A, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open*. 2021;11(9):e046912. doi:10.1136/bmjopen-2020-046912
100. Bidulka P, Scott J, Taylor DM, et al. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. *BMJ Open*. 2022;12(3):e057909. doi:10.1136/bmjopen-2021-057909
101. Scott J, Bidulka P, Taylor DM, et al. Management and outcomes of myocardial infarction in people with impaired kidney function in England. *BMC Nephrology*. 2023/11/02 2023;24(1):325. doi:10.1186/s12882-023-03377-x

CHAPTER 2. METHODS

Overview

In this chapter, I provide an overview of the methods used in this thesis. This overview complements the individual methods sections in the research papers included in Chapters 3 to 9, which provide more specific details about study populations, exposures, covariates, outcomes, and analysis plans.

This chapter begins with section 2.1., which describes the research projects that employed me throughout my PhD registration. In sections 2.2. and 2.3, I outline the general study designs and data sources for the research included in this thesis, respectively. In section 2.4., I then describe the main statistical methods and models I used in this thesis. In section 2.5., I summarise the LSHTM and external ethics approvals covering the analyses included in this thesis. Finally, in section 2.6., I conclude this methods chapter.

2.1. Description of research projects funding my PhD research and my roles on the project teams

2.1.1. The Personalised Medicine for Intensification of Treatment (PERMIT): the case for type 2 diabetes mellitus study

The PERMIT study[1] is funded by the National Institute for Health and Care Research (NIHR). The aim of this study is to evaluate the comparative effectiveness of second-line oral antidiabetic treatments and predict longer term outcomes using a microsimulation model developed in US data. This project funded the research included in Case Study 1 of this thesis (Chapters 3 to 6) and part of my salary during the PhD registration period.

I was the first research analyst hired by Professor Richard Grieve at the start of the project. I was responsible for operationalising the successful grant application into an actionable protocol. I led all ethics and data applications to the LSHTM and to the Independent Scientific Advisory Committee (ISAC) at the Medicines and Healthcare products Regulatory Agency (MHRA) at the start of the project and all subsequent amendments to the approved data application.[2] I was solely responsible for extracting and creating the study cohort and led the reporting of protocols and analyses to co-investigators at quarterly grant holders' meetings throughout my PhD. I act as a 'central node' for the PERMIT project as the internal team has expanded, being the lead author on the published protocol paper (Chapter 4) and the Statistical Analysis Plan (SAP)[3] for the research paper presented in Chapter 5. I co-led analyses, manuscript drafting, and reporting to project stakeholders. I was also the lead representative of the PERMIT team at one of two patient and public involvement (PPI) panels.

2.1.2. The Quality and Equity of Care in Kidney Disease: the promise of big data (QECKD) study

The QECKD study was a Kidney Research UK (KRUK) funded study which aimed to describe inequalities in AMI treatment among people with CKD in England. This project funded part of the research in Case Study 2 of this thesis and part of my salary during the PhD registration period.

I was hired by Professor Dorothea Nitsch as the only research analyst funded by the project. I was responsible for operationalising the successful grant application into a protocol. I led all ethics and data applications to the LSHTM, the steering committees for the National Chronic Kidney Disease Audit (NCKDA), and the Myocardial Ischaemia National Audit Project (MINAP). I was solely responsible for extracting and creating the study cohort and statistical analyses. I led the reporting of protocols and analyses to co-investigators at grant holders' meetings and co-led drafting of the resultant peer-reviewed publications presented in Chapters 7 to 8.

2.1.3. Improving acute cardiac care of patients with renal disease through linkage of national audits in the UK: the National Cardiac and Renal Audit Initiative (NACARAI) project

The NACARAI project is a Health Foundation funded project[4] which aims to investigate the comparative effectiveness of alternative AMI treatments among people with kidney impairment using variation in AMI treatment across hospitals in England. This project funded part of the research in Case Study 2 of this thesis and part of my salary during the PhD.

My research from the QECKD project (section 2.1.2.) inspired this Health Foundation funded project. I had input to the content of the grant application, including the study design, and was the named analyst in the successful funding application. I lead all ethics and data applications to the LSHTM and externally, including to the Health Quality Improvement Partnership and the British Heart Foundation (BHF) Data Centre (for future work described in the thesis discussion section 10.5.2.). I lead all analyses funded by this work and included in Chapters 7 to 9.

2.2. Description of my approach to study the comparative effectiveness of treatments

I used a multi-stage approach to study the comparative effectiveness of treatment in English primary and secondary care. This approach aligns with guidance from the NICE real-world evidence framework published in 2022, which recommends the core principles of ensuring the suitability of the data to answer the study question, generating evidence transparently, and using analytical methods to reduce the risk of bias[5] (see also section 1.2.5.).

The steps I used in my approach to study the comparative effectiveness of treatments in routinely collected health data are as follows:

1. Ensure the main data source is suitable to define the study population of interest.
2. Understand important patient-level characteristics that influence treatment receipt and represent inequalities in treatment.
3. Examine variation in treatment receipt to inform the selection of a preference-based instrument.
4. Design a comparative effectiveness analysis with input from a multidisciplinary team, including patients.
5. Conduct a comparative effectiveness analysis to estimate the comparative effectiveness of treatments.

In **Figure 2.1.**, I offer a visual guide mapping how the research papers in this thesis (Chapters 3 to 9) address each step outlined above. **Table 2.1.** accompanies this figure to summarise the details of the research papers included in this thesis. In the following sections, I will describe in greater detail the approach outlined in **Figure 2.1.** as it relates to the two case studies.

Figure 2.1. Illustration of the approach I used to plan and conduct two comparative effectiveness studies using an instrumental variable analysis in routinely collected health data and how each step maps to the thesis objectives and research papers included in this thesis.

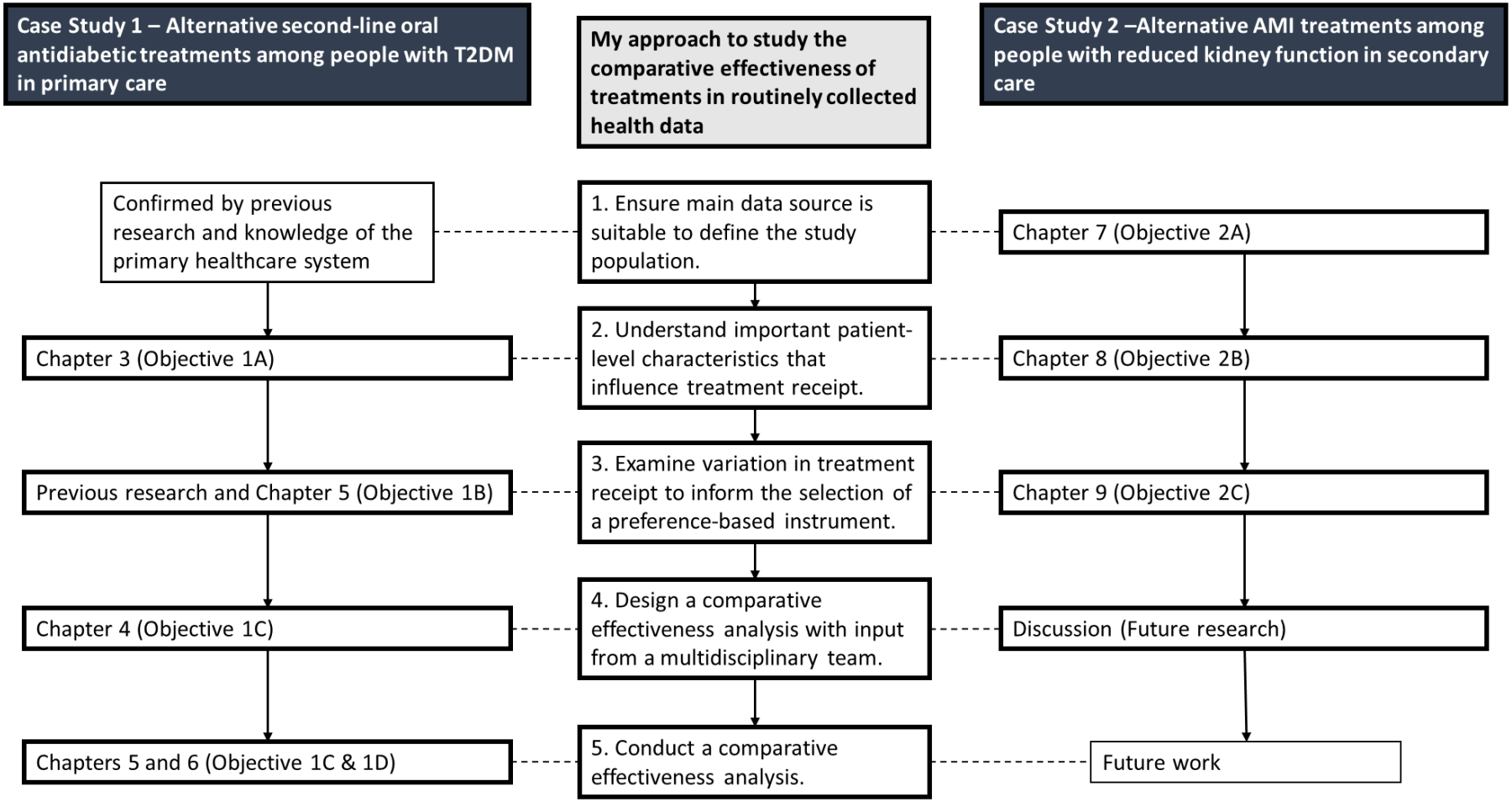


Table 2.1. Key aspects of the study designs for included research papers in Chapters 3 to 9 of this thesis.

Thesis chapter no.	Research paper name	Aim of the research paper	Key aspects of the study design	Study population	Main exposure/independent variable of interest	Main outcome(s) of interest	Modelling strategy
Case study 1: Relative effectiveness of alternative second-line oral antidiabetic treatments among people with type 2 diabetes mellitus (T2DM) in primary care							
3	Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment in people with type 2 diabetes in England: a cross-sectional study.	To describe the association between sociodemographic person-level characteristics and alternative second-line antidiabetic treatment prescribing in England primary care.	Cross-sectional study.	People with T2DM who initiate second-line oral antidiabetic treatment.	Ethnicity and deprivation quintile.	Initiation of second-line oral antidiabetic treatment with one of SU, DPP4i, or SGLT2i added to metformin monotherapy.	Multivariable multinomial mixed effect logistic regression
4	Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study.	To detail the protocol for a cohort study using an IV analysis to determine the comparative effectiveness of alternative second-line oral antidiabetic treatments.	Protocol paper.	People with T2DM who initiate second-line oral antidiabetic treatment.	-	-	-
5	Comparative effectiveness of alternative second-	To determine the comparative effectiveness of the	Cohort study using a target	People with T2DM who initiate second-	Initiation of second-line oral antidiabetic treatment with one of	Metabolic and clinical measures	Non-parametric survival analysis (plotting

Thesis chapter no.	Research paper name	Aim of the research paper	Key aspects of the study design	Study population	Main exposure/independent variable of interest	Main outcome(s) of interest	Modelling strategy
	line oral antidiabetic treatments on metabolic, kidney, and cardiovascular outcomes amongst people with type 2 diabetes mellitus: a cohort study using routinely collected health data.	most commonly prescribed alternative second-line oral antidiabetic treatments in England on important outcomes for people with T2DM.	trial design and an IV analysis.	line oral antidiabetic treatment.	SU, DPP4i, or SGLT2i added to metformin monotherapy.	(mean change in HbA1c, BMI, eGFR, and blood pressure). Cardiovascular and kidney outcomes (MACE, heart failure hospitalisation, MAKE). All-cause mortality.	cumulative failure curves for survival outcomes). IV analysis: 2SRI (primary analysis) with linear regression or Cox PH regression in the second stage model. 2SLS (secondary analysis) with linear second stage model. Traditional multivariable analysis: Multivariable linear and Cox PH regression. Other: Propensity score-IPTW-RA
6	Going beyond Randomised Controlled Trials to assess treatment effect	To emulate a published RCT comparing DPP4i and SU added to metformin as second-	Cohort study using a target trial emulation	People with T2DM who initiate second-line oral antidiabetic treatment.	Initiation of second-line oral antidiabetic treatment with one of SU or DPP4i added to	Mean change in HbA1c.	IV analysis: LIV (primary analysis) with linear regression in the second stage model.

Thesis chapter no.	Research paper name	Aim of the research paper	Key aspects of the study design	Study population	Main exposure/independent variable of interest	Main outcome(s) of interest	Modelling strategy
	heterogeneity across target populations	line antidiabetic treatment in observational data and extend the results to those who would have been excluded from the RCT.	design and an IV analysis.		metformin monotherapy.		Other: Propensity score-IPTW-RA
Case study 2: Relative effectiveness of alternative AMI treatments among people with reduced kidney function in secondary care							
7	Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study.	To understand the differential capture of AMI hospitalisations according to individual-level kidney function across two secondary care datasets.	Cohort study.	People at-risk of or with kidney impairment/CKD included in the NCKDA who are hospitalised for AMI.	Kidney impairment (defined as having the most recent eGFR from primary care being <60mL/min/1.73m^2).	AMI case ascertainment in the MINAP dataset only, the HES dataset only, or both MINAP & HES.	Traditional multivariable regression: Multivariable multinomial logistic regression. Multivariable logistic regression.
8	Management and outcomes of myocardial infarction in people with impaired kidney function in England.	To describe the association between kidney function and alternative AMI treatments.	Cohort study.	People at-risk of or with kidney impairment/chronic kidney disease (CKD) included in the NCKDA who are hospitalised for AMI.	Level of kidney impairment (defined as eGFR stage 1-2, stage 3a, stage 3b, and stage 4-5).	Receipt of invasive (angiography and/or PCI) versus conservative cardiac management strategy.	Traditional multivariable regression: Multivariable logistic regression. Multivariable Cox PH regression.

Thesis chapter no.	Research paper name	Aim of the research paper	Key aspects of the study design	Study population	Main exposure/independent variable of interest	Main outcome(s) of interest	Modelling strategy
9	Acute myocardial infarction treatment variation and inequalities by kidney function: a cross-sectional study using the Myocardial Ischaemia National Audit Project (MINAP).	To describe centre-level and individual-level variation in AMI treatment by level of kidney impairment in the MINAP dataset.	Cross-sectional study.	People hospitalised for AMI and included in the MINAP audit.	Level of kidney impairment (defined as eGFR stage 1-2, stage 3a, stage 3b, and stage 4-5).	Receipt of invasive (angiography and/or PCI) versus conservative cardiac management strategy.	Multivariable logistic regression.

2SLS: two-stage least squares; 2SRI: two-stage residual inclusion; AMI: acute myocardial infarction; CKD: chronic kidney disease; eGFR: estimated glomerular filtration rate; HbA1c: haemoglobin A1c; IPTW-RA: inverse probability of treatment weighting with regression adjustment; IV: instrumental variable; LIV: local instrumental variable; MACE: major adverse cardiovascular event; MAKE: major adverse kidney event; MINAP: Myocardial Ischaemia National Audit Project; NCKDA: National Chronic Kidney Disease Audit; PCI: percutaneous coronary intervention; PH: proportional hazards; T2DM: type 2 diabetes mellitus

2.2.1. Case study 1: Relative effectiveness of alternative second-line oral antidiabetic treatments among people with T2DM in primary care

In this case study, I used a combination of cross-sectional and cohort study designs. The cohort studies included in this thesis applied an IV analysis[6] within the target trial emulation framework[7, 8] (see also sections 1.2.5. and 1.3.).

Step 1: Ensure main data source is suitable to define the study population

I relied on previously published studies to confirm the suitability of the primary care data to be used to define the study population and treatments (**Figure 2.1.**). People with T2DM in England are managed mainly in primary care, where the majority of antidiabetic prescribing occurs.[9, 10] Several validation studies have investigated the validity of primary care data to capture T2DM diagnoses in the general UK population.[11, 12] These studies suggest that T2DM case ascertainment is unreliable in secondary care and primary care data should be prioritised. Further, case ascertainment in primary care is sensitive to diagnosis codes used to define a study population with prevalent T2DM. Incorporating antidiabetic prescriptions is recommended to improve T2DM case ascertainment.[11]

Previous studies investigating comparative antidiabetic treatment effects in the UK follow these recommendations and ascertain T2DM cases in primary care data using a combination of diagnosis codes (e.g., Read[13] and/or Snomed-CT[14] codes) and recorded antidiabetic prescriptions.[15-18] I follow similar methodologies to define a cohort of people living with T2DM and initiating second-line oral antidiabetic treatment following metformin monotherapy.

Step 2: Understand important patient-level characteristics that influence treatment receipt and represent inequalities in treatment

In Chapter 3, I present a research paper published in *Diabetes, Obesity & Metabolism*[19] which used a cross-sectional study design to understand important socio-demographic characteristics which are associated with second-line antidiabetic treatment prescribing.

Step 3: Examine variation in treatment receipt to inform the selection of a preference-based instrument

Precedent research by Wilkinson et al (2018) demonstrated substantial variation in second-line antidiabetic treatment prescribing across groups of GPs in England between 2000-2017.[10] These data

were used to inform the design of the comparative effectiveness analyses using variation in antidiabetic prescribing at the CCG-level to derive a preference-based IV. In the research paper presented in Chapter 5, I update the findings by Wilkinson et al (2018)[10] to demonstrate persistent variation in second-line oral antidiabetic treatment prescribing at the CCG-level during the study period of interest (2014 to 2021).

Step 4: Design a comparative effectiveness analysis with input from a multidisciplinary team

In Chapter 4, I present a protocol paper published in *BMJ Open*[20] which outlines the protocol for a comparative effectiveness IV analysis comparing the three most commonly prescribed second-line oral antidiabetic treatments in England.[10] Following the protocol paper, I include a table from the subsequently published statistical analysis plan[3] which summarises the changes I made to the protocol[20] post-publication which were necessary due to statistical or logistical challenges. This protocol and SAP were designed and co-authored by a multidisciplinary group of clinicians, statisticians, health economists, policymakers, and patients.

Step 5: Conduct a comparative effectiveness analysis

In Chapter 5, I present a paper *in press* at *The British Medical Journal (BMJ)* which used a cohort study design to estimate the comparative effectiveness of the three alternative second-line oral antidiabetic treatments of interest in this thesis. I used the TTE framework combined with an IV analysis following guidance from the NICE real-world evidence framework to reduce the impact of biases, including that from measured and unmeasured confounding, in real-world studies of comparative effects.[5] In this study, I defined the ideal RCT which could answer the causal question of which alternative antidiabetic treatment was better in terms of specific outcomes relevant to T2DM patients.[7, 8] I then translated this ideal RCT into a feasible observational study using routinely collected health data from England. I could only emulate an ideal RCT, since there was no published RCT which included a direct comparison of all three alternative antidiabetic treatments of interest in this study. This research paper also describes the CCG-level variation in alternative second-line antidiabetic prescribing to confirm this variation exists in my contemporary cohort (Step 3 in **Figure 2.1.**).

Finally, in Chapter 6, I present a manuscript *in submission* which similarly used a cohort study designed using an IV analysis applied within the TTE framework. This study specifically aimed to describe heterogenous treatment effects with respect to mean change in HbA1c from baseline to 1-year follow-up across the study population initiating second line antidiabetic treatment.

Heterogenous treatment effects are "non-random, explainable variability in the direction and magnitude of treatment effects for individuals within a population."[21] In epidemiology, heterogenous treatment effects are often described as effect modification or interactions between the treatment and some other variable. Treatment heterogeneity due to observed variables is classified as overt heterogeneity and is most often investigated using subgroup analyses and likelihood ratio tests for interaction effects.[21] However, it is likely that treatment heterogeneity is also explained by other unobserved variables, which can create problems when estimating treatment effects in routinely collected health data. Treatment heterogeneity due to observed and unobserved variables is classified as essential heterogeneity. While it is not possible to investigate the interaction effect of a specific unmeasured variable with a treatment, one can consistently estimate the ATE and CATEs in a study population, accounting for essential heterogeneity, using a particular type of IV analysis, called the local IV (LIV)[22] (further details in section 2.4.4.).

The LIV can only consider two-way treatment comparisons at the time of writing this thesis. Therefore, in this study which aims to investigate heterogenous treatment effects, I narrowed the focus of this research paper to two of the three second-line oral antidiabetic treatments of interest in this thesis (SU and DPP4i, both added to metformin monotherapy). I focused on these two treatments because there are many trials comparing these two treatments directly as opposed to SGLT2i trials which are mainly placebo controlled.[23] In a literature review described in Chapter 6, I was able to identify a suitable trial[24] which I could feasibly emulate in the study data to (a) understand if results from the LIV analysis agreed with results from the RCT when applying similar inclusion/exclusion criteria to define a trial eligible subpopulation within the observational cohort, and (b) to transport the IV analysis to the trial ineligible subpopulation and the overall target population and investigate heterogeneous treatment effects.

2.2.2. Case study 2: Relative effectiveness of alternative AMI treatments among people with reduced kidney function in secondary care

In this case study, I used a combination of cohort studies and a cross-sectional study to follow my approach to design and conduct a comparative effectiveness analysis using routinely collected health data (**Figure 2.1.**).

Step 1: Ensure main data source is suitable to define the study population

I could not rely on previously published work for this first step as I did for Case Study 1. Since AMI are primarily managed in secondary care, the study population needed to be defined in this setting. As described in the thesis introduction (section 1.1.), secondary care data are more fragmented than primary care data. Previous work found that AMI case ascertainment was incomplete using individual secondary care datasets,[25] but no study to my knowledge investigated potential differential AMI case ascertainment according to kidney disease status.

Thus, in Chapter 7, I present a research paper published in *BMJ Open*[26] which used a cohort study design to describe AMI case ascertainment among people with reduced kidney function in two secondary care data sources. This paper also investigated the agreement between eGFR recorded in primary care (considered the best estimate of baseline kidney function in routinely collected health data)[27] and eGFR recorded in secondary care to understand if baseline kidney impairment could be reliably defined in secondary care data.

Step 2: Understand important patient-level characteristics that influence treatment receipt and represent inequalities in treatment

In Chapter 8, I present a research paper published in *BMC Nephrology*[28] which used the same cohort as in Chapter 7 to describe disparities in AMI treatment and outcomes according to individuals' level of kidney impairment.

Step 3: Examine variation in treatment receipt to inform the selection of a preference-based instrument

In Chapter 9, I present a manuscript *in preparation* which uses a cross-sectional study design to describe AMI treatment variation at the cardiology centre-level across England and how this is influenced by kidney impairment. These analyses inform the definition of a preference-based instrument, namely the tendency for invasive management among people with impaired kidney function hospitalised for AMI.

Steps 4 to 5: Design and conduct a comparative effectiveness analysis with input from a multidisciplinary team

In the thesis discussion (section 10.5.2.), I discuss the design and planned comparative effectiveness analysis, similarly applying an IV analysis within the TTE framework as in Case Study 1, to estimate the comparative effectiveness of alternative AMI treatments among people with kidney impairment.

These analyses were not completed within the PhD registration period since prerequisite research on the suitability of the secondary care data sources was required before undertaking these advanced analyses in the routine data (Chapters 7 to 9).

2.2.3. Summary of the approach used to study comparative effectiveness of treatments in this thesis

In this section, I summarised the approach I used to study the comparative effectiveness of treatments in two case studies set in English primary and secondary care. This approach followed the core principles and guidance from NICE on how to generate real-world evidence from routinely collected health data.[5] In the next section, I will describe the routinely collected health data sources I used to conduct these case studies.

2.3. Data resources

I used several pseudonymised routinely collected health datasets. These datasets come from EHRs recorded by the GP or hospital, disease-specific national audits, small-area level deprivation data from the UK government, and death data from the Office of National Statistics (ONS). In **Table 2.2**., I detail the data sources used in the case studies included in this thesis.

Table 2.2. Summary of data sources used in each case study of this thesis.

Data source	**Case study 1** (Relative effectiveness of alternative antidiabetics among people with T2DM)	**Case study 2** (Relative effectiveness of alternative AMI treatments among people with impaired kidney function)
Primary care	CPRD	NCKDA
Secondary care	HES	HES, MINAP
Deprivation status	IMD	IMD
Death information	ONS	ONS

AMI: Acute myocardial infarction; CPRD: Clinical Practice Research Datalink; HES: Hospital Episode Statistics; IMD: Index of Multiple Deprivation; MINAP: Myocardial Infarction National Audit Project; NCKDA: National Chronic Kidney Disease Audit; ONS: Office of National Statistics

2.3.1. Primary care data

Clinical Practice Research Datalink (CPRD)

The CPRD is made up of two large population-based datasets: CPRD Gold and Aurum. CPRD Gold data[29] are collected from GPs who use Vision software and have agreed to contribute data. CPRD Gold user-base is decreasing over time. CPRD Aurum data[30] are collected from GPs who use EMIS Health software and have similarly agreed to contribute data. CPRD Aurum user-base is growing over time. Together, these datasets, referred to hereafter as CPRD, contain primary care data collected from approximately 20% of the UK population and are broadly representative of the UK population in terms of age and sex.[29, 30] These CPRD data include diagnoses (recorded using Read codes[13] (CPRD Gold) or Snomed codes[14] (CPRD Aurum)), prescriptions, laboratory test results, and demographic and lifestyle factors (e.g., age, sex, height, weight, smoking status, and alcohol intake).

National Chronic Kidney Disease Audit (NCKDA)

The NCKDA was a clinical audit commissioned by the HQIP. This audit aimed to understand and improve routine clinical care for people at risk of or living with CKD.[31, 32] The NCKDA included 1,005 GPs, which covers approximately 75% of the Welsh patient population and 10% of the English patient population. People at risk of CKD (**Table 2.3.**) or with CKD were identified and included in the audit using clinical diagnosis codes and laboratory test results. Only data relevant to CKD were collected in two main extracts (2014-2016) from the GP historical records, including relevant diagnoses, prescriptions, laboratory test results, and lifestyle factors.

Table 2.3. Risk factors for CKD which fulfilled inclusion criteria for the NCKDA[31]

Risk factors for CKD
• Relevant cardiovascular disease • Diabetes mellitus • Hypertension • Connective tissue disorders • Prostatic disease • Kidney stones • Previous AKI • Family history of CKD • Previous prescriptions for kidney-damaging medications such as lithium or calcineurin inhibitors

AKI: acute kidney injury; CKD: chronic kidney disease

2.3.2. Secondary care

Hospital Episode Statistics (HES)

HES Admitted Patient Care data, hereafter referred to as HES data, includes all NHS-funded in-patient hospitalisations in England, collated by NHS England.[33] Hospitalisations, referred to as 'spells', are split into 'episodes', which designate the time spent under the management of a particular consultant during the hospitalisation. Each episode has up to 20 diagnoses, in order of relevance to the hospitalisation, coded using International Classification of Diseases 10th edition (ICD-10). HES data also include dates, methods of admission and discharge, and demographic data (e.g., ethnicity).

Myocardial Ischaemia National Audit Project (MINAP)

MINAP is an on-going cardiovascular audit which aims to describe and improve the patient pathway from acute coronary syndrome (ACS) hospitalisation to discharge.[34, 35] MINAP aims to capture all people hospitalised with ACS (including type 1 AMI and unstable angina) across all hospitals in England, Wales, and Northern Ireland. These data include patient demographics, comorbidities, smoking status, diagnostic tests, serum creatinine within 24 hours of hospitalisation (used to calculate estimated glomerular filtration rate (eGFR), a key measure of kidney function),[36, 37] medications prescribed at discharge, procedures and treatments received in-hospital, and discharge or death information.

2.3.3. Death data

The Office of National Statistics (ONS) mortality data includes a death date and cause of death (ICD-10) for all deaths registered in England and Wales. These data are routinely linked to health data for research purposes.[38]

2.3.4. Deprivation data

The Index of Multiple Deprivation (IMD) is used in England to rank neighbourhoods from 1 to 32,844 in terms of relative deprivation by combining seven indices at the Lower-level Super Output Area or neighbourhood level.[39] Individuals are then assigned a relative measure of deprivation at the small area level based on their post code of residence. These rankings are commonly split into ordered

quintiles for research and have high completeness in routinely collected primary care data like the CPRD.[40]

2.3.5. Summary of data sources used in both case studies of this thesis

In Case Study 1: alternative second-line oral antidiabetic treatments in people with T2DM, I use CPRD data linked with HES, IMD, and ONS data (**Table 2.2.**). These data are widely used in pharmacoepidemiological analyses and are suitable to study the comparative effectiveness of alternative antidiabetic treatments.

In Case Study 2: alternative AMI treatments in people with kidney impairment, I use NCKDA audit data linked with HES, IMD, ONS, and MINAP data (**Table 2.2.**). I used the NCKDA data instead of alternative primary care data like the CPRD because linkage to MINAP and other audit data such as the UK Renal Registry[41] are not yet routine and would require substantial data governance and time to enable. Permissions to link MINAP data to the NCKDA primary care audit were in place and feasible within my PhD registration period. In the thesis discussion (section 10.5.2.), I discuss the datasets I will use in future research to optimise the linked data sources available for this case study.

While data governance and linkages are complex and lengthy processes, in part to protect sensitive patient data, using linked data from across the spectrum of routinely collected health data is important to enhance the quality of any pharmacoepidemiological analysis.[42, 43] The linked datasets used in both case studies offer valuable information to understand and describe aspects of patient care and relevant health-related factors which must be considered in pharmacoepidemiological studies. In this thesis, these linkages are particularly important because of the complex health profiles and, consequently, interactions with the NHS the people included in both case studies experience.

These complex health profiles and interactions with the health service captured in routinely collected health data make any observational study prone to bias, particularly confounding. In the next section, I describe the general statistical models I applied in these linked data sources to minimise the risk of confounding when I report treatment inequalities and comparative effectiveness.

2.4. Statistical models used in this thesis

In this section, I describe the statistical models I use in the research papers presented in Chapters 3 to 9 to model the relationship between dependent (outcomes) and independent variables (exposures/treatments and measured confounders). These models include generalised linear models (GLMs) (linear regression, logistic regression), Cox proportional hazards (PH) regression, the two-stage

IV models (two-stage least squares (2SLS), two-stage residual inclusion (2SRI), and local IV (LIV)), and the propensity score with inverse probability of treatment weighting with regression adjustment (IPTW-RA). These details supplement and extend on those provided in the methods sections of the individual research papers and are provided here for completeness.

When applied in pharmacoepidemiological studies to investigate the causal effect of an exposure on an outcome, the multivariable linear and Cox PH regression models, as well as the propensity score IPTW-RA, estimate the ATE assuming no unmeasured confounding (**Table 2.4.**). This assumption is unlikely to hold when using routinely collected health data for causal inference in pharmacoepidemiology, since many important confounders are unmeasured or imperfectly measured, like diet, exercise, and frailty.

The IV models make alternative assumptions that do not include the assumption of no unmeasured confounding (section 1.3. and Appendix B.1.). I will review these assumptions and how I investigate their plausibility in my research in section 2.4.4.

2.4.1. Multivariable linear regression to model continuous outcomes

In Case Study 1, I investigated continuous outcomes such as the change in HbA1c and body-mass index (BMI). To model the change in these continuous outcomes, I used multivariable linear regression which used ordinary least squares (OLS) estimation to model the assumed linear relationship between the treatments of interest and the outcome. Other independent variables included in the model were potential confounders. This model can be expressed as:

$$Y = \beta_0 + \beta_1 D_1 + \boldsymbol{\beta}_o \boldsymbol{X}_o + \varepsilon$$

Where Y is the dependent variable (outcome), β_0 is the intercept, representing the value of Y when all independent variables are equal to zero, D_1 is the treatment variable, β_1 is the coefficient for the treatment variable, D_1, $\boldsymbol{\beta}_o$ is a vector of coefficients $(\beta_2, \beta_3, \dots, \beta_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_2, X_3, \dots, X_p)$, where p represents the total number of observed confounders, and ε is the residual error term, representing the difference between the observed and predicted values of Y.

The OLS model will estimate coefficients for the independent variables of interest which minimise the sum of squared differences between the observed values of the dependent variable and the predicted

values from the linear equation. This model relies on several assumptions, namely (i) there is a linear relationship between the independent variables and the dependent variable, (ii) homoscedasticity, meaning the variance of the error term is constant across all values of the independent variables, (iii) independence, meaning all observations are independent of each other, and (iv) the distribution of the errors follows a normal distribution.[44]

Change scores using multivariable linear regression

In the research papers presented in Chapters 5 and 6, I investigate the change in clinical measures (e.g., HbA1c) from baseline to pre-specified points in follow-up (e.g., 1-year), which are modelled using multivariable linear regression models. The 'change score', that is, the difference in the clinical measure between baseline and 1-year follow-up, serves as the dependent outcome variable. When modelling change scores as an outcome, there is debate about whether to include or exclude the baseline outcome measure as an independent variable in the outcome regression model – this decision can impact the type of causal estimand one estimates.[45-48]

Glymour et al (2006) illustrate the bias which can be induced by measurement error and regression to the mean in an applied example of the effect of education on changes in cognitive function in older age.[49] Tennant et al (2022) stress that the choice to adjust for the baseline outcome measure is dependent on whether this measure is a confounder or mediator of the causal effect of interest.[45]

In Chapters 5 and 6, where I investigate the causal effect of antidiabetic treatments on the change scores for clinical measures such as HbA1c, I chose to adjust for the baseline clinical measure in an analysis of covariance (ANCOVA), as this was measured prior to or on the same day as the 'index date' (i.e., the date of second-line antidiabetic treatment prescription). Thus, this variable likely acts as a confounder and not a mediator of the causal relationship of interest. I discuss the potential limitations of this decision in the thesis discussion (section 10.4.6.).

2.4.2. Multivariable logistic and multinomial logistic regression to model binary or categorical outcomes

In addition to continuous outcomes, I investigate binary outcomes such as AMI treatment with an invasive versus conservative cardiac strategy. To model binary outcomes, I use multivariable logistic regression. This model can be expressed as:

$$logit(P(Y = 1 \mid D_1, \boldsymbol{X}_o)) = \beta_0 + \beta_1 D_1 + \boldsymbol{\beta}_o \boldsymbol{X}_o$$

Where $logit(Y)$ is the natural logarithm of the odds of the outcome, β_0 is the intercept, representing the log odds when all other coefficients are set to zero, β_1 is the coefficient for the treatment variable, D_1, representing the odds ratio (OR) on the log-scale, and $\boldsymbol{\beta}_o$ is a vector of coefficients $(\beta_2, \beta_3, \dots, \beta_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_2, X_3, \dots, X_p)$, where p represents the total number of observed confounders.

I also use multivariable multinomial logistic regression. This is used when the dependent variable has more than two categories. I use this, for example, when I am modelling as the dependent variable second-line oral antidiabetic treatment prescribing (SU, DPP4i, or SGLT2i added to metformin) or AMI case ascertainment (AMI captured in HES & MINAP, MINAP only, or HES only). The formula is like the multivariable logistic regression formula but extends to accommodate more than two categories of the dependent variable. This can be expressed as:

$$\log\left(\frac{P(Y = k \mid D_1, \boldsymbol{X}_o)}{P(Y = 1 \mid D_1, \boldsymbol{X}_o)}\right) = \beta_{k0} + \beta_{k1} D_1 + \boldsymbol{\beta}_{ko} \boldsymbol{X}_o$$

Where Y is the categorical dependent variable with K categories, where K>2, k is the level of the categorical dependent variable Y, β_{k0} is the intercept belonging to the category k of the dependent variable Y, D_1 is the treatment variable, β_{k1} is the coefficient for the treatment variable, D_1, belonging to category k, representing the OR on the log-scale, and $\boldsymbol{\beta}_{ko}$ is a vector of coefficients $(\beta_{k2}, \beta_{k3}, \dots, \beta_{kp})$ for the vector of observed confounders, $\boldsymbol{X}_{ko} = (X_{k2}, X_{k3}, \dots, X_{kp})$ for category k, where p represents the total number of observed confounders.

Misinterpretations of the OR

The OR is easily estimated using logistic regression and commonly reported in the medical literature for binary outcomes.[50] However, where the outcome is common, interpretations of the OR will be exaggerated when compared to the probability scale. To better understand this issue, I present the general formula for the odds:

$$Odds = \frac{p}{1 - p}$$

Where p represents the probability of the outcome.

When the probability of the outcome is rare (i.e., <~10%) the odds will be similar to the risk. When the outcome is common (>~10%) the odds will be considerably greater than the risk. When taking the ratio of the odds, this ratio will be exaggerated for common outcomes. Thus, the sign of the OR is much easier to interpret compared with the magnitude.[50]

For example, the probability of an outcome among people exposed is 0.4 and among those unexposed is 0.7. The risk ratio could be calculated as 0.4/0.7 = 0.57. The odds ratio could be calculated as (0.4/1-0.4) / (0.7/1-0.7) = 0.29. The interpretation of this OR, that the exposure leads to 71% lower odds of the outcome compared with the unexposed, is not easily interpretable.[50]

In the case of a common outcome, the predicted probabilities of the outcome with respect to the exposure derived from the multivariable logistic regression model is a useful method to illustrate the magnitude of effect on the probability scale (0-1).[51] These differences in the adjusted predicted probabilities can be more easily interpreted in the clinical context. Thus, where I use OR in this thesis, I also report the adjusted predicted probabilities to better understand and interpret the association between the exposures and outcomes of interest.

2.4.3. Multivariable Cox Proportional Hazards regression to model time-to-event outcomes

In both case studies, I investigate time-to-event or survival outcomes such as time-to-heart failure hospitalisation or time-to-all-cause mortality. Time-to-event outcomes are often described with the hazard function, which represents the instantaneous probability of an event conditional on surviving event-free up to a certain time.[52] The hazard functions between treatment groups can be compared and expressed as a hazard ratio (HR) using multivariable Cox proportional hazards (PH) regression.[52] This model can be expressed as:

$$h(t \mid D_1, \boldsymbol{X}) = h_0(t) \times \exp(\beta_1 D_1 + \boldsymbol{\beta}_o \boldsymbol{X}_o)$$

Where $h(t|D_1, \boldsymbol{X})$ is the hazard function at time t for an individual with independent variables D_1 and $\boldsymbol{X}$, $h_0(t)$ is the baseline hazard function, representing the hazard for an individual with all independent variables equal to zero, exp is an exponential function denoted as $\exp(x) = e^x$ with e being the base of the natural logarithm, β_1 is the log HR associated with the treatment variable, D_1, and $\boldsymbol{\beta}_o$ is a vector of coefficients $(\beta_2, \beta_3, \dots, \beta_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_2, X_3, \dots, X_p)$, where p represents the total number of observed confounders.

The Cox PH regression model[52] and the HR are widely used in reporting treatment effects in RCTs and observational studies when analysing time-to-event outcomes.[53] This model parameterises the independent variables but does not parameterise the baseline hazard function, allowing for flexibility in the modelling of the time-to-event data. However, the Cox model assumes proportional hazards, meaning that the hazards for the exposed/treated and unexposed/untreated groups are proportional across the entire follow-up period. This assumption can be evaluated by examining non-parametric plots, by exploring whether the hazard ratio varies over time, and/or by exploring whether the Schoenfeld residuals change over time.[53] Stensrud et al (2020) argue this assumption is unlikely to hold in most studies comparing treatments since treatment effects are rarely constant over time.[53] In response, Sjölander et al (2024) argue that these tests are still informative and should not be disregarded.[54] In this thesis, I focused on examining and testing the Schoenfeld residuals to verify the PH assumption.

2.4.4. IV two-stage regression

I use an IV analysis in my primary analyses to estimate the comparative effectiveness of alternative second-line oral antidiabetic treatments in Chapters 5 and 6. More specifically, in Chapter 5, I use 2SLS and 2SRI models to estimate the LATE and ATE estimands, respectively. In Chapter 6, I use an LIV model to estimate the ATE accounting for essential heterogeneity. In the thesis discussion (section 10.5.2.), I briefly discuss my future research which will use these IV models to study the comparative effectiveness of alternative AMI treatments among people with kidney impairment. **Table 2.4.** summarises the key features of the alternative IV analyses I use in this thesis. These features will be expanded upon in the text where I describe each IV analysis in greater detail.

Table 2.4. Key features of the IV analyses used in this thesis and a comparison with traditional multivariable regression models.

	Analysis		
	2SLS	**2SRI**	**LIV**
Chapter where the type of IV analysis is used (primary or secondary analysis)	5 (alternative analysis)	5 (primary analysis)	6 (primary analysis)

	Analysis		
	2SLS	**2SRI**	**LIV**
Estimand which can be estimated	ATE (assuming homogeneous treatment effects) LATE (assuming monotonicty)	ATE, ATT, ATU, CATE (assuming monotonicity)	ATE, ATT, ATU, CATE (assuming monotonicity)
Treatment heterogeneity	Treatment effect estimates can account for overt heterogeneity.	Treatment effect estimates can account for overt heterogeneity.	Treatment effect estimates account for essential heterogeneity.[55]
Instrument requirements	Continuous or binary IV.	Continuous or binary IV.	Continuous IV.
Outcome model type	Linear.	Linear or non-linear.	Linear or non-linear.
Model specification	Assumes both models are correct.	Assumes both models are correct.	Assumes both models are correct.
Consideration for 3-way treatment comparisons	Can consider 3-way treatment comparisons.	Can consider 3-way treatment comparisons.	Cannot consider 3-way treatment comparisons.

2SLS: two-stage least squares; 2SRI: two-stage residual inclusion; ATE: average treatment effect; ATT: average treatment effect in the treated; ATU: average treatment effect in the untreated; CATE: conditional average treatment effect; IV: instrumental variable; LATE: local average treatment effect; LIV: local instrumental variable

Defining the instrument

In this thesis, I use a preference-based continuous instrument.[56] As I described in the introduction (1.3.), prescribing preference is a latent variable. I use a proxy, the 'tendency to prescribe (TTP)' a certain treatment, which can be calculated by choosing a look-back period from the index date for each individual (e.g., 1-year) and calculating the proportion of people prescribed each drug treatment of interest in that period.

In Case Study 1, I used the TTP at the CCG-level in the 1-year prior to the index date for the first second-line antidiabetic treatment prescription. I made the decision to use the tendency to prescribe at the CCG level rather than the individual GP level because many prescribing policies are made by

commissioners of the health service at the group-level. These policies are likely to contribute to the exogenous treatment variation across these groups of GPs which can be used to conduct an IV analysis (section 1.4.1.).

Evaluating the IV assumptions using statistical methods

The relevance assumption, that the instrument strongly predicts the treatment, can be formally tested using the partial F-statistic from the first stage IV regression model.[6] This F-statistic measures the strength of association between the instrument and the treatment actually prescribed. The general rule of thumb is that this F-statistic should be >10 to adequately demonstrate that an IV meets the relevance assumption. Recent work, however, suggests this F-statistic should be >100.[6, 57] I report the F-statistic describing the strength of association between the instrument and the treatment in both case studies to test the first IV assumption.

The exchangeability assumption, that the instrument is not associated with confounders, can only be partially evaluated by comparing levels of the instrument across levels of the measured confounders. To do this falsification test in my thesis, I standardise each covariate by dividing the value of the covariate by its standard deviation (SD) across the sample population of interest. I then plot the mean standardised values of the covariates by deciles of the instrument (the TTP) to observe any associations between these variables. Any association between a measured confounder and the instrument will be accounted for in the multivariable first stage model. But these observed associations can draw suspicion that the instrument is associated with unmeasured confounders, thus violating the exchangeability assumption. I offer further discussion about the plausibility of this assumption in the research papers (Chapters 5 and 6) and the thesis discussion (section 10.4.1.).

The exclusion restriction, that the instrument is a cause of the outcome only via the treatment, cannot be formally evaluated. I use DAGs to illustrate this assumption and offer further discussion about its plausibility in the discussion section of the research paper included in Chapter 5 and the thesis discussion (section 10.4.1.).

Finally, a more nuanced fourth assumption must also be considered in an IV analysis. One must assume either (i) treatment homogeneity, that is the treatment has a homogenous effect across the study population, or (ii) in the presence of treatment heterogeneity, the monotonicity assumption, that is the instrument has a consistent effect on the treatment, meaning that there are no 'defiers' in the study population (Appendix B.2.).

The monotonicity assumption can be more formally expressed as:

If $Z' > Z$ then $D_{Z'} \geq D_Z$ with probability 1

Where Z' and Z are alternative levels of the instrument and $D_{Z'}$ and D_Z are the corresponding alternative levels of the treatment.

In other words, this means that an increase in the level of the instrument, Z, always results in a higher or equal level of treatment assignment, D_Z. This assumption is difficult to verify since we cannot observe the counterfactual treatment decisions for the same patient across different prescribers.[58] Swanson et al (2015) proposed a survey study design of prescribers to understand their treatment plans for hypothetical patients and their prescribing preferences, with the ultimate aim of understanding counterfactual treatments each hypothetical patient would receive across all prescribers in the survey.[58] However, this is difficult to implement in practice, and the hypothetical patients will not be representative of all patients in an IV analysis. I offer a brief commentary about my assumption of monotonicity in the research papers presented in Chapter 5 and the thesis discussion (section 10.4.1.).

After carefully defining a suitable instrument, the IV can then be used in comparative effectiveness analyses to estimate treatment effects. The choice of IV model depends on the estimand of interest, the outcome and instrument type (continuous or binary), and accommodations for treatment heterogeneity (**Table 2.4.**). I will first describe the 2SLS model, which is commonly used in the economic literature and which I use in an alternative analysis in Chapter 5.

2SLS model

The 2SLS model regresses the instrument and other observed confounders on the treatment variable in the first stage model. The treatment prediction from the first stage model then replaces the observed treatment variable in the second stage model. The parameter for the treatment prediction in this second stage model is assumed to be independent of the error term.[59]

The 2SLS model is a particular case of the two-stage predictor substitution model that can only be used with a continuous outcome and a continuous or binary instrument.[60] The 2SLS model only estimates the ATE when treatment effects are homogenous across all individuals. Otherwise, when allowing for treatment heterogeneity, the 2SLS model estimates the local average treatment effect (LATE), which is the ATE among the 'compliers' (Appendix B.2.),[59] i.e., those who would be prescribed treatment A when the prescribing clinician prefers treatment A and would also be prescribed treatment B when a

different prescribing clinician prefers treatment B.[58] This group is difficult to define and thus the LATE is difficult to generalise, making it less useful for policy makers and clinical decision-making.[22]

2SRI model

The 2SRI model[61] is a particular type of IV model which measures the ATE, allowing for heterogenous treatment effects according to observed confounders under the monotonicity assumption.[58] Like the 2SLS model, the treatment effect is measured in two stages.

Stage 1 model

In the first stage, the treatment is regressed on the instrument using a probit or logit model. In this thesis, I used a probit model in the first stage regression. This model calculates the probability of treatment given the instrument and the independent variables included in the probit model, i.e., the propensity score. The first stage model can be expressed as:

$$P(D_1 = 1 \mid Z) = \Phi(\alpha_0 + \alpha_1 Z + \boldsymbol{\alpha}_o \boldsymbol{X}_o)$$

Where D_1 is the treatment variable, Z is the instrumental variable, $\boldsymbol{\alpha_o}$ is a vector of coefficients $(\alpha_2, \alpha_3, \dots, \alpha_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_2, X_3, \dots, X_p)$, where p represents the total number of observed confounders, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

The residuals, that is the difference between the observed treatment and the predicted probabilities of treatment, are then calculated and used in the second stage regression model as additional independent variables. In this thesis, I used the generalised residuals[62] as they were demonstrated to produce the least biased estimates of the ATE in a simulation study which compared alternative residual forms in a 2SRI model.[60] These were calculated using the score option in the predict command in Stata.

Stage 2 model

In the second stage, the generalised residuals[62] are included as independent variables along with the observed treatment and any other independent variables (i.e., observed confounders). This is unlike

2SLS, where the treatment is replaced by the predicted treatment from the first-stage model. The type of regression model used in the second stage will depend on the type of outcome under study.

For continuous outcomes, a linear model can be used. This second stage model can be expressed as:

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 \hat{\varepsilon}_{2SRI} + \boldsymbol{\beta}_o \boldsymbol{X}_o + u$$

Where Y is the dependent variable (outcome), β_0 is the intercept, representing the value of Y when all independent variables are equal to zero, D_1 is the treatment variable, β_1 is the coefficient for the treatment variable, D_1, β_2 is the coefficient for the residuals from the first stage model, $\hat{\varepsilon}_{2SRI}$, $\boldsymbol{\beta}_o$ is a vector of coefficients $(\beta_3, \beta_4, \dots, \beta_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_3, X_4, \dots, X_p)$, where p represents the total number of observed confounders, and u is the residual error term between the observed and predicted values of Y, which we assume here to be independent of the treatment, D_1, in this second stage IV model.

The counterfactual outcomes can then be predicted from this second stage model, with the ATE being the mean difference between these counterfactual predicted outcomes under the alternative treatments.

For time-to-event outcomes, a Cox PH regression model can be used. The second stage model can be expressed as:

$$h(t \mid D_1, \hat{\varepsilon}_{2SRI}, \boldsymbol{X_0}) = h_0(t) \times \exp(\beta_1 D_1 + \beta_2 \hat{\varepsilon}_{2SRI} + \boldsymbol{\beta}_o \boldsymbol{X}_o)$$

Where $h(t|D_1, \hat{\varepsilon}_{2SRI}, \boldsymbol{X})$ is the hazard function at time t for an individual with independent variables D_1, $\hat{\varepsilon}_{2SRI}$, and $\boldsymbol{X_0}$, $h_0(t)$ is the baseline hazard function, representing the hazard for an individual with all independent variables equal to zero, exp is an exponential function denoted as $\exp(x) = e^x$ with e being the base of the natural logarithm, β_1 is the log HR associated with the treatment variable, D_1, β_2 is the coefficient for the residuals from the first stage model, $\hat{\varepsilon}_{2SRI}$, and $\boldsymbol{\beta}_o$ is a vector of coefficients $(\beta_2, \beta_3, \dots, \beta_p)$ for the vector of observed confounders, $\boldsymbol{X}_o = (X_2, X_3, \dots, X_p)$, where p represents the total number of observed confounders.

I interpreted the HR from this second stage model as the ATE. To account for uncertainty in estimating the propensity score and residuals, bootstrapping is needed to estimate standard errors to generate confidence intervals (CI).[55, 60, 63] Non-parametric bootstrapping with replacement was used to generate

t-based 95% CI for the treatment effects generated by the IV models. Further details are provided in the supplementary methods of the research paper in Chapter 5 (Appendix E1.1. to E.4.).

Including the first stage model residuals in the second stage model will minimise the association between the confounders, measured and unmeasured, and the treatment. The more likely the IV assumptions are met, the more likely the model will minimise the risk of bias due to confounding, measured and unmeasured.[61]

The key limitation of the 2SRI model is that the procedure is sensitive to misspecification of the first stage model. This can lead to errors in the residuals estimated from the first stage model, which can then bias the results from the second stage model.[60, 61] Previous work demonstrated that using the generalised residuals resulted in the lowest risk of bias, hence why I chose to use this form of the residuals in this study.[60]

LIV model

The LIV model[64, 65] is an alternative two-stage IV model which requires a continuous instrument and is suitable for continuous or binary outcomes (**Table 2.4.**). Like the 2SRI model, the LIV model can estimate the ATE; however, the model has particular advantages over the 2SRI model, in that it can account for essential heterogeneity; that is heterogenous treatment effects from observed and unobserved variables (e.g., frailty).[64]

Here I provide a short overview of the LIV method. Additional details are provided in the research paper in Chapter 6. Briefly, the LIV works by identifying marginal treatment effects (MTEs) for patients at the 'margins of treatment choice'.[64, 66] These MTEs are estimated in situations of clinical equipoise, where there is balance between the observed characteristics (including the IV) which encourage the treatment and unobserved characteristics that discourage the treatment.[57] In this case, a small change in the instrument is sufficient to 'nudge' these hypothetical patients into the treatment group without changing the distribution of the underlying characteristics, both observed and unobserved. The differences in the outcome for these people with marginally different levels of the instrument, but different treatment allocations, identifies a series of MTEs for hypothetical individuals at the margins of treatment choice. These MTEs can be aggregated to calculate the ATE or CATEs for particular subgroups, accounting for essential heterogeneity.[22] Basu (2014, 2015) extended this approach to assign individuals a weighted MTE that predicts an individual-level treatment effect, called the person-centred treatment effect, which are useful in describing heterogeneity across individuals in the study population.[55, 67]

The first stage model is similar to the other IV methods: the propensity score for the treatment is calculated by regressing the treatment variable on the independent variables and the continuous instrument using a probit or logit model. The second stage regresses the outcome on the observed confounders and a function of the propensity score for treatment, which includes interactions with observed confounders. This outcome model is then differentiated by the function of the propensity score for treatment. The personalised treatment effects for each individual are then estimated using numerical integration. Again, the standard errors to calculate confidence intervals can be calculated using non-parametric bootstrapping.[55, 67]

The LIV method was used in the primary analysis to study heterogenous treatment effects of SU and DPP4i as second-line oral antidiabetic treatments. I could not use this model in Chapter 5 because the theory on how to apply this IV model in a three-way treatment comparison setting did not exist at the time of writing this thesis (**Table 2.4.**).

2.4.5. Propensity score with inverse probability of treatment weighting with weighted regression adjustment (IPTW-RA)

I use a propensity score model with IPTW-RA as an alternative analytical method to estimate the ATE under the assumption of no unmeasured confounding in Chapters 5 (in response to reviewer comments) and 6 (pre-specified).

Like the IV analyses, this is a two-stage regression model. The propensity score is calculated in the first stage model using a probit or logit function. The propensity score can then be used in several ways to adjust for measured confounders, including IPTW.[68, 69] With IPTW, the inverse of the probability of being treated (the propensity score) is calculated for each individual to create a 'weight'. Using this weight, a pseudo-population is created in which the confounder distributions in the treated and untreated groups are the same as in the original total population.[70] This allows for estimation of the ATE which is unbiased by confounding, assuming that there are no unmeasured confounders and that both models are specified correctly.

The IPTW-RA follows this approach but adds the doubly robust estimator. By adding the doubly robust estimator (IPTW-RA), the assumption that both models are specified correctly is loosened so that only one of the two models must be specified correctly (see Funk et al (2011) – Appendix 1).[71]

2.4.6. Summary of regression models used in this thesis

I use a variety of regression models in this thesis to estimate inequalities in treatment and causal treatment effects. These models make different assumptions which must be carefully considered, particularly when interpreting model outputs as causal treatment effects. The traditional multivariable regression models (sections 2.4.1. to 2.4.3.) and the propensity score IPTW-RA model (section 2.4.5.) all assume no unmeasured confounding and no essential heterogeneity when estimating the ATE. These assumptions are unlikely to hold in routinely collected health data.

The IV analyses presented in section 2.4.4. do not assume no unmeasured confounding. Further, the LIV model accounts for essential heterogeneity when estimating the ATE or CATEs in subgroups of the study population. These advantages over the models which assume no unmeasured confounding are balanced by the alternative assumptions the IV analyses must make. Thus, these analyses require careful scrutiny in their design and application to minimise the risk of bias in pharmacoepidemiological analyses.

2.5 Ethics

Before applying these methods to study the comparative effectiveness of treatments, I obtained ethics approvals for all research included in this thesis from internal and external committees. These ethics approvals were obtained from the Independent Scientific Advisory Committee of the Medicines and Health Regulatory Agency, the NCKDA steering committee, and LSHTM (**Appendix A**).

2.6. Chapter summary

In this chapter, I summarised the overarching methods used in this PhD thesis, particularly the detail not included in the research papers included in Chapters 3 to 9. These methods included the description of the research projects funding my salary and this research, the general approach I used to study the comparative effectiveness of treatments in both case studies, the data sources, and details about the statistical models I used to measure associations and causal treatment effects.

In Chapters 3 to 6 (case study 1: alternative second-line oral antidiabetic treatments among people with T2DM in primary care) and chapters 7 to 9 (alternative AMI treatments among people with kidney impairment in secondary care), I present research papers which apply the methods I have outlined here to accomplish the aims and objectives of this thesis.

2.7. References

1. PERMIT: Investigating the effectiveness of different treatments for Type 2 Diabetes Mellitus, to help choose the right treatments for patients. Web. London School of Hygiene & Tropical Medicine. Accessed May 19, 2023. https://www.lshtm.ac.uk/research/centres-projects-groups/permit#welcome

2. PERsonalised Medicine for Intensification of Treatment (PERMIT): the case of type 2 diabetes mellitus. Web. Medicines & Healthcare products Regulatory Agency: CPRD. Accessed 23 March, 2024. https://www.cprd.com/approved-studies/personalised-medicine-intensification-treatment-permit-case-type-2-diabetes

3. Bidulka P, Lugo-Palacios DG, Carroll O, O'Neill S, Grieve R. Statistical Analysis Plan (SAP): PERsonalised Medicine for Intensification of Treatment (PERMIT) study, Version 1.0, July 2023. Web. Accessed 28 November, 2023. https://www.lshtm.ac.uk/media/72276

4. Improving acute cardiac care of patients with renal disease through linkage of national audits in the UK: London School of Hygiene & Tropical Medicine. Web. The Health Foundation. Accessed 23 March, 2024. https://www.health.org.uk/funding-and-partnerships/programmes/improving-acute-cardiac-care-of-patients-with-renal-disease

5. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

6. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. Jun 15 2014;33(13):2297-340. doi:10.1002/sim.6128

7. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

8. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

9. Diabetes primary and community care. Web. Diabetes UK. Accessed 23 March, 2024. https://www.diabetes.org.uk/for-professionals/improving-care/good-practice/primary-and-community-care#:~:text=With%20rising%20demand%20for%20services,care%20for%20people%20with%20diabetes.

10. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

11. Tate AR, Sheena D, Simon G, Natalia B, Rachael W, Tim W. Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect incidence estimates? Cross-

sectional study using the CPRD database. *BMJ Open*. 2017;7(1):e012905. doi:10.1136/bmjopen-2016-012905

12. Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes in UK Biobank. *PLoS One*. 2016;11(9):e0162388. doi:10.1371/journal.pone.0162388

13. Read codes. Web. NHS England. Accessed 7 April, 2024. https://digital.nhs.uk/services/terminology-and-classifications/read-codes

14. Snomed CT. Web. NHS England. Accessed 7 April, 2024. https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct

15. Wilkinson S, Williamson E, Pokrajac A, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. *Diabetes, Obesity and Metabolism*. 2020;22(5):847-856. doi:https://doi.org/10.1111/dom.13970

16. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

17. Eastwood SV, Mathur R, Sattar N, Smeeth L, Bhaskaran K, Chaturvedi N. Ethnic differences in guideline-indicated statin initiation for people with type 2 diabetes in UK primary care, 2006–2019: A cohort study. *PLoS medicine*. 2021;18(6):e1003672. doi:10.1371/journal.pmed.1003672

18. Mathur R, Farmer RE, Eastwood SV, Chaturvedi N, Douglas I, Smeeth L. Ethnic disparities in initiation and intensification of diabetes treatment in adults with type 2 diabetes in the UK, 1990–2017: A cohort study. *PLoS medicine*. 2020;17(5):e1003106. doi:10.1371/journal.pmed.1003106

19. Bidulka P, Mathur R, Lugo-Palacios DG, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes, Obesity and Metabolism*. 2023;25(1):282-292. doi:https://doi.org/10.1111/dom.14874

20. Bidulka P, O'Neill S, Basu A, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open*. 2021;11(9):e046912. doi:10.1136/bmjopen-2020-046912

21. Varadhan R, Seeger JD. *Developing a protocol for Observational Comparative Effectiveness Research: A User's Guide*. Agency for Healthcare Research and Quality (US); 2013.

22. Moler-Zapata S, Grieve R, Lugo-Palacios D, et al. Local Instrumental Variable Methods to Address Confounding and Heterogeneity when Using Electronic Health Records: An Application to

Emergency Surgery. *Medical decision making : an international journal of the Society for Medical Decision Making*. Nov 2022;42(8):1010-1026. doi:10.1177/0272989x221100799

23. McGuire DK, Shih WJ, Cosentino F, et al. Association of SGLT2 Inhibitors With Cardiovascular and Kidney Outcomes in Patients With Type 2 Diabetes: A Meta-analysis. *JAMA Cardiol*. Feb 1 2021;6(2):148-158. doi:10.1001/jamacardio.2020.4511

24. Nauck MA, Meininger G, Sheng D, Terranella L, Stein PP. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes, obesity & metabolism*. Mar 2007;9(2):194-205. doi:10.1111/j.1463-1326.2006.00704.x

25. Herrett E, Shah AD, Boggon R, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ : British Medical Journal*. 2013;346:f2350. doi:10.1136/bmj.f2350

26. Bidulka P, Scott J, Taylor DM, et al. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. *BMJ Open*. 2022;12(3):e057909. doi:10.1136/bmjopen-2021-057909

27. Iwagami M, Tomlinson LA, Mansfield KE, et al. Validity of estimated prevalence of decreased kidney function and renal replacement therapy from primary care electronic health records compared with national survey and registry data in the United Kingdom. *Nephrol Dial Transplant*. Apr 1 2017;32(suppl_2):ii142-ii150. doi:10.1093/ndt/gfw318

28. Scott J, Bidulka P, Taylor DM, et al. Management and outcomes of myocardial infarction in people with impaired kidney function in England. *BMC Nephrology*. 2023/11/02 2023;24(1):325. doi:10.1186/s12882-023-03377-x

29. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. Jun 2015;44(3):827-36. doi:10.1093/ije/dyv098

30. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology*. 2019;48(6):1740-1740g. doi:10.1093/ije/dyz034

31. Nitsch DC, Ben; Hull, Sally; Wheeler, David. *First National CKD Audit Report 2017*. 2017. https://www.lshtm.ac.uk/files/ckd_audit_report.pdf

32. Mudie KC, Faye; Caplin, Ben; Wheeler, David; Hull, Sally; Nitsch, Dorothea. *Second National CKD Audit Report 2017*. 2017. https://www.lshtm.ac.uk/media/9951

33. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *International journal of epidemiology*. 2017;46(4):1093-1093i. doi:10.1093/ije/dyx015

34. Herrett E, Smeeth L, Walker L, Weston C, on behalf of the MAG. The Myocardial Ischaemia National Audit Project (MINAP). *Heart (British Cardiac Society)*. 2010;96(16):1264. doi:10.1136/hrt.2009.192328

35. Myocardial Ischaemia/MINAP (Heart Attack audit). Web. The National Institute for Cardiovascular Outcomes Research. Accessed 18 May, 2021. https://www.nicor.org.uk/national-cardiac-audit-programme/myocardial-ischaemia-minap-heart-attack-audit/

36. Levey AS, Stevens LA, Schmid CH, et al. A New Equation to Estimate Glomerular Filtration Rate. *Annals of Internal Medicine*. 2009/05/05 2009;150(9):604-612. doi:10.7326/0003-4819-150-9-200905050-00006

37. Levey AS, Coresh J, Greene T, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med*. Aug 15 2006;145(4):247-54. doi:10.7326/0003-4819-145-4-200608150-00004

38. Linked HES-ONS mortality data. Web. NHS Digital. Accessed 4 March, 2022. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data#ons-mortality-data

39. Penney B. *The English Indices of Deprivation 2019 (IoD2019)*. 2019. Accessed 4 March 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf

40. Preveina M, Mia H, Stuart F, et al. Completeness and representativeness of small area socioeconomic data linked with the UK Clinical Practice Research Datalink (CPRD). *Journal of Epidemiology and Community Health*. 2022;76(10):880. doi:10.1136/jech-2022-219200

41. UK Renal Registry (UKRR): Home. Web. UKRR. Accessed 12, 2023. https://ukkidney.org/about-us/who-we-are/uk-renal-registry

42. Mourby MJ, Doidge J, Jones KH, et al. Health Data Linkage for UK Public Interest Research: Key Obstacles and Solutions. *Int J Popul Data Sci*. Apr 2 2019;4(1):1093. doi:10.23889/ijpds.v4i1.1093

43. Grieve R, Yang Y, Abbott S, et al. The importance of investing in data, models, experiments, team science, and public trust to help policymakers prepare for the next pandemic. *PLOS Global Public Health*. 2023;3(11):e0002601. doi:10.1371/journal.pgph.0002601

44. Casson RJ, Farmer LD. Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & Experimental Ophthalmology*. 2014;42(6):590-596. doi:https://doi.org/10.1111/ceo.12358

45. Tennant PWG, Arnold KF, Ellison GTH, Gilthorpe MS. Analyses of 'change scores' do not estimate causal effects in observational data. *Int J Epidemiol*. Oct 13 2022;51(5):1604-1615. doi:10.1093/ije/dyab050

46. Glymour MM. Commentary: Modelling change in a causal framework. *Int J Epidemiol*. Oct 13 2022;51(5):1615-1621. doi:10.1093/ije/dyac151

47. Lord FM. A paradox in the interpretation of group comparisons. *Psychol Bull*. Nov 1967;68(5):304-5. doi:10.1037/h0025105

48. Pearl J. Lord's Paradox Revisited – (Oh Lord! Kumbaya!). *Journal of Causal Inference*. 2016;4(2)doi:doi:10.1515/jci-2016-0021

49. Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM. When is baseline adjustment useful in analyses of change? An example with education and cognitive change. *Am J Epidemiol*. Aug 1 2005;162(3):267-78. doi:10.1093/aje/kwi187

50. Norton EC, Dowd BE, Maciejewski ML. Odds Ratios—Current Best Practice and Use. *Jama*. 2018;320(1):84-85. doi:10.1001/jama.2018.6971

51. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *International Journal of Epidemiology*. 2014;43(3):962-970. doi:10.1093/ije/dyu029

52. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34(2):187-220.

53. Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? *Jama*. 2020;323(14):1401-1402. doi:10.1001/jama.2020.1267

54. Sjölander A, Dickman P. Why test for proportional hazards – or any other model assumptions? *Am J Epidemiol*. 2024;doi:10.1093/aje/kwae002

55. Basu A. ESTIMATING PERSON-CENTERED TREATMENT (PeT) EFFECTS USING INSTRUMENTAL VARIABLES: AN APPLICATION TO EVALUATING PROSTATE CANCER TREATMENTS. *J Appl Econ (Chichester Engl)*. June/July 2014;29(4):671-691. doi:10.1002/jae.2343

56. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *The international journal of biostatistics*. 2007;3(1):Article 14. doi:10.2202/1557-4679.1072

57. Moler-Zapata S, Grieve R, Basu A, O'Neill S. How does a local instrumental variable method perform across settings with instruments of differing strengths? A simulation study and an evaluation of emergency surgery. *Health Econ*. Jun 11 2023;doi:10.1002/hec.4719

58. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and evaluation of the monotonicity condition for preference-based instruments. *Epidemiology (Cambridge, Mass)*. May 2015;26(3):414-20. doi:10.1097/ede.0000000000000279

59. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*. 1996;91(434):444-455. doi:10.2307/2291629

60. Basu A, Coe NB, Chapman CG. 2SLS versus 2SRI: Appropriate methods for rare outcomes and/or rare exposures. *Health Econ*. Jun 2018;27(6):937-955. doi:10.1002/hec.3647

61. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*. 2008/05/01/ 2008;27(3):531-543. doi:https://doi.org/10.1016/j.jhealeco.2007.09.009

62. Gourieroux C, Monfort A, Renault E, Trognon A. Generalised residuals. *Journal of Econometrics*. 1987/01/01/ 1987;34(1):5-32. doi:https://doi.org/10.1016/0304-4076(87)90065-0

63. Mooney CZ, Duval RD. *Bootstrapping: A nonparametric Approach to Statistical Inference*. SAGE Publications, Inc.; 1993.

64. Heckman JJ, Vytlacil EJ. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proc Natl Acad Sci U S A*. Apr 13 1999;96(8):4730-4. doi:10.1073/pnas.96.8.4730

65. Heckman JJ, Vytlacil E. Policy-Relevant Treatment Effects. *The American Economic Review*. 2001;91(2):107-111.

66. Björklund A, Moffitt R. The Estimation of Wage Gains and Welfare Gains in Self-Selection Models. *The Review of Economics and Statistics*. 1987;69(1):42-49. doi:10.2307/1937899

67. Basu A. Person-centred treatment (PeT) effects: Individualized treatment effects using instrumental variables. *The Stata Journal*. 2015;15(2):397-410.

68. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ (Clinical research ed)*. 2019;367:l5657. doi:10.1136/bmj.l5657

69. Ali MS, Prieto-Alhambra D, Lopes LC, et al. Propensity Score Methods in Health Technology Assessment: Principles, Extended Applications, and Recent Advances. *Front Pharmacol*. 2019;10:973. doi:10.3389/fphar.2019.00973

70. Stürmer T, Rothman KJ, Glynn RJ. Insights into different results from different causal contrasts in the presence of effect-measure modification. *Pharmacoepidemiol Drug Saf*. Oct 2006;15(10):698-709. doi:10.1002/pds.1231

71. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly Robust Estimation of Causal Effects. *Am J Epidemiol*. 2011;173(7):761-767. doi:10.1093/aje/kwq439

CHAPTER 3. RESEARCH PAPER – ETHNIC AND SOCIOECONOMIC DISPARITIES IN INITIATION OF SECOND-LINE ANTIDIABETIC TREATMENT IN PEOPLE WITH TYPE 2 DIABETES IN ENGLAND: A CROSS SECTIONAL STUDY.

OVERVIEW

In this chapter, I include a research paper published in *Diabetes, Obesity & Metabolism* for which I am the first author. I presented this work as an oral presentation at the International Conference in Pharmacoepidemiology (ICPE) 2022 in Copenhagen, Denmark, and as an invited speaker at Imperial College London, McMaster University, and internally at LSHTM.

In this study, I investigate the association between important sociodemographic characteristics, namely ethnicity and deprivation status, and prescription for alternative second-line oral antidiabetic treatments using linked English primary (CPRD) and secondary care (HES) data. Following the research paper, I include a brief discussion of the relevance of this paper to my thesis. This discussion includes a key table from the supplementary materials of the published paper which is particularly important in the interpretation of the study in the context of this thesis.

3.1. Published research paper

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study comparative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Diabetes, Obesity, & Metabolism		
When was the work published?	22 September 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I am first and correpsonding author of this paper. I led the design of this study with input from my co-authors and supervisor, Richard Grieve. I led the data management, including the extraction and creation of the study dataset from the raw CPRD files. I led the data analysis, with support from my co-author, Stephen O'Neill. I led the literature review and the interpretation of all study results. I wrote the first draft of the manuscript. Co-authors reviewed multiple drafts before sending to the journal for publication. I led the peer review responses with input from my co-authors.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Richard Grieve
Date	07 April 2024

Received: 13 July 2022 | Revised: 6 September 2022 | Accepted: 18 September 2022

DOI: 10.1111/dom.14874

ORIGINAL ARTICLE

WILEY

Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study

Patrick Bidulka MSc[1] | Rohini Mathur PhD[1] | David G. Lugo-Palacios PhD[2] | Stephen O'Neill PhD[2] | Anirban Basu PhD[3] | Richard J. Silverwood PhD[4] | Paul Charlton MA[5] | Andrew Briggs DPhil[2] | Liam Smeeth MBChB[1] | Amanda I. Adler MD[6] | Ian J. Douglas PhD[1] | Kamlesh Khunti MD[7] | Richard Grieve PhD[2]

[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK

[2]Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK

[3]The Comparative Health Outcomes, Policy & Economics (CHOICE) Institute, University of Washington School of Pharmacy, Seattle, Washington

[4]Centre for Longitudinal Studies, UCL Social Research Institute, University College London, London, UK

[5]Patient Research Champion Team, National Institute for Health Research, Twickenham, UK

[6]Diabetes Trials Unit, The Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Headington, UK

[7]Diabetes Research Centre, University of Leicester, Leicester, UK

Correspondence
Patrick Bidulka, Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK.
Email: patrick.bidulka1@lshtm.ac.uk

Funding information
National Institute for Health and Care Research (NIHR), Grant/Award Number: NIHR128490

Abstract

Aims: To assess any disparities in the initiation of second-line antidiabetic treatments prescribed among people with type 2 diabetes mellitus (T2DM) in England according to ethnicity and social deprivation level.

Materials and methods: This cross-sectional study used linked primary (Clinical Practice Research Datalink) and secondary care data (Hospital Episode Statistics), and the Index of Multiple Deprivation (IMD). We included people aged 18 years or older with T2DM who intensified to second-line oral antidiabetic medication between 2014 and 2020 to investigate disparities in second-line antidiabetic treatment prescribing (one of sulphonylureas [SUs], dipeptidyl peptidase-4 [DPP-4] inhibitors, or sodium-glucose cotransporter-2 [SGLT2] inhibitors, in combination with metformin) by ethnicity (White, South Asian, Black, mixed/other) and deprivation level (IMD quintiles). We report prescriptions of the alternative treatments by ethnicity and deprivation level according to predicted percentages derived from multivariable, multinomial logistic regression.

Results: Among 36 023 people, 85% were White, 10% South Asian, 4% Black and 1% mixed/other. After adjustment, the predicted percentages for SGLT2 inhibitor prescribing by ethnicity were 21% (95% confidence interval [CI] 19–23%), 20% (95% CI 18–22%), 19% (95% CI 16–22%) and 17% (95% CI 14–21%) among people with White, South Asian, Black, and mixed/other ethnicity, respectively. After adjustment, the predicted percentages for SGLT2 inhibitor prescribing by deprivation were 22% (95% CI 20–25%) and 19% (95% CI 17–21%) for the least deprived and the most deprived quintile, respectively. When stratifying by prevalent cardiovascular disease (CVD) status, we found lower predicted percentages of people with prevalent CVD

prescribed SGLT2 inhibitors compared with people without prevalent CVD across all ethnicity groups and all levels of social deprivation.

Conclusions: Among people with T2DM, there were no substantial differences by ethnicity or deprivation level in the percentage prescribed either SGLT2 inhibitors, DPP-4 inhibitors or SUs as second-line antidiabetic treatment.

KEYWORDS
ethnicity, oral antidiabetics, pharmacoepidemiology, socioeconomic deprivation, type 2 diabetes

1 | INTRODUCTION

Most healthcare systems report inequities in disease incidence, healthcare delivery and outcomes according to people's socioeconomic status and ethnicity.[1,2] For countries with single-payer systems such as England, national recommendations from agencies like the National Institute for Health and Care Excellence (NICE) encourage access to effective and cost-effective interventions to maximize clinical benefit while also reducing health inequalities.[3-5] Nonetheless in countries such as England, inequities in using healthcare interventions according to people's socioeconomic characteristics persist for diseases such as cardiovascular disease (CVD),[6] chronic kidney disease (CKD),[7] and type 2 diabetes mellitus (T2DM).[8]

There are inequities in T2DM prevalence and outcomes according to ethnicity[8] and deprivation.[9] People of Black and South Asian ethnicity, and people with lower income or lower educational attainment have a higher prevalence of T2DM, worse blood glucose control and earlier onset of macro- and microvascular complications compared with people of ethnicities other than Black and South Asian, higher incomes or higher educational attainment.[8-13] Ethnic minorities also tend to experience delays in T2DM treatment intensification when clinically indicated (therapeutic inertia),[12] which may contribute to worse outcomes compared with White people.[14-18] Other ethnic and socioeconomic inequities in T2DM treatment that could impact clinical outcomes, such as the type of second-line antidiabetic treatment prescribed at treatment intensification from metformin monotherapy, are less well understood. Hence, we chose to examine the potential disparities in second-line antidiabetic treatment prescribing by ethnicity and deprivation status.

For people with T2DM whose glycated haemoglobin (HbA1c) levels are poorly controlled, an important choice is which second-line oral antidiabetic therapy to prescribe in addition to metformin.[19] Between 2015 and 2021, NICE technology appraisals and clinical guidelines recommended that, for most people with T2DM, several second-line oral treatment options should be available, including sodium-glucose cotransporter 2 (SGLT2) inhibitors, dipeptidyl peptidase-4 (DPP-4) inhibitors or the lower-cost option of sulphonylureas (SUs).[19,20] Updated NICE guidelines (2022) recommend SGLT2 inhibitors for individuals at high risk of or with prevalent CVD but that, for other eligible patients, any of these three treatments may be suitable.[21] The decision to allow local discretion in the choice of second-line treatment may reflect the uncertainty over comparative effectiveness and cost-effectiveness of these three treatment choices, which is partly related to the lack of randomized controlled trials (RCTs) providing head-to-head comparisons of these three oral antidiabetic drugs. In contrast, international diabetes guidance and consensus reports recommend SGLT2 inhibitors for people with established atherosclerotic CVD, heart failure and CKD,[22] irrespective of the additional costs of SGLT2 inhibitors compared to SUs, drawing on evidence from placebo-controlled RCTs showing improved CVD and kidney disease outcomes when prescribing SGLT2 inhibitors.

Previous research has found wide variation in clinical practice in the United Kingdom (UK) in the choice of second-line oral antidiabetic treatment.[23,24] However, no previous study has considered whether disparities exist in prescriptions of this second-line treatment according to ethnicity and socioeconomic status. We aimed to assess whether ethnic minorities and people with higher deprivation status had a lower probability of being prescribed SGLT2 inhibitors compared with DPP-4 inhibitors or the lower-cost SUs, both overall and by prevalent CVD status.

2 | METHODS

2.1 | Study design and setting

We conducted a cross-sectional study, nested within the Personalized Medicine for Intensification of Treatment in people with T2DM (PERMIT) cohort study,[25] to investigate disparities in second-line antidiabetic treatment prescribed to people with T2DM by ethnicity and by deprivation status. Data sources included the Clinical Practice Research Datalink (CPRD) Gold and Aurum datasets (primary care), Hospital Episode Statistics (HES; secondary care), the Index of Multiple Deprivation (IMD), and the Office of National Statistics (ONS) death data.

The CPRD is a large, population-based dataset covering approximately 20% of the UK population and is representative in terms of age, sex and ethnicity.[26,27] These data include clinical diagnoses, laboratory test results, and prescribing information recorded in primary care. Linkage of CPRD data to HES data is available for approximately 80% of people in the CPRD registered at general practices in England. HES data include diagnoses and demographic information related to all NHS-funded hospitalizations.[28] The IMD is commonly used in epidemiological research as a proxy for socioeconomic status in England.

It ranks individuals according to deprivation status based on their postcode, and is usually reported in quintiles (1 being the least deprived, 5 being most deprived).[29] ONS mortality data include information on all deaths registered in England and Wales.[30]

2.2 | Patient and public involvement

One patient and public (PP) representative (P.C.) was involved in this study's design, provided feedback on this manuscript, and is a co-author. The PERMIT study protocol describes PP contributions to the study design.[25] PP representatives will assist with drafting lay summaries, which we will share on the study website (https://www.lshtm.ac.uk/research/centres-projects-groups/permit) and at study workshops with a wider group of multi-ethnic PP representatives. We will work with the Centre for Ethnic Health Research, led by co-author K.K., to make culturally adapted lay summaries.

2.3 | Study population

We included people aged 18 years or older with a T2DM diagnosis, in whom incident second-line oral antidiabetic treatment was prescribed for the first time between January 1, 2014 and March 31, 2020 after first-line antidiabetic treatment with metformin monotherapy. We used the complete historical general practice (GP) electronic health record to ensure this was the first time each person had a record of being prescribed an SU, a DPP-4 inhibitor or an SGLT2 inhibitor. The second-line therapy, an SU, a DPP-4 inhibitor or an SGLT2 inhibitor, had to have been added on to metformin, and had not replaced it. These three treatments constituted approximately 99% of the second-line treatments prescribed, therefore, other second-line antidiabetic treatments were excluded from this study.[23,31] Eligible people had to have had a prescription for metformin monotherapy within 60 days prior to the first prescription for second-line treatment to ensure they were continuous users of metformin monotherapy prior to intensification. Also, to ensure the second-line treatments were an addition to, rather than a switch from, metformin, the individuals were required to have been prescribed metformin on the same day or within 60 days after the first prescription for the second-line antidiabetic treatment.

We excluded women with a record of pregnancy within 12 months prior to second-line treatment initiation since antidiabetic prescribing guidelines are different for this population.[19] We also excluded people whose last recorded estimated glomerular filtration rate (eGFR) was less than 30 ml/min/1.73 m^2 since metformin is contraindicated in this group, and SGLT2 inhibitors are not recommended for this group in the United Kingdom for the purpose of lowering blood glucose.[19,32]

2.4 | Definitions of ethnicity and deprivation

We defined ethnicity according to clinical and demographic codes recorded within the CPRD or linked HES data prior to or on the same day as the first-ever prescription date for one of the three second-line antidiabetic treatments of interest, that is, the index date. Ethnicity was grouped into 16 categories in primary care and 11 categories in secondary care, which we further re-grouped as the following: (1) White, (2) South Asian, (3) Black, and (4) Mixed/other (Table S1). We considered this re-grouping necessary to ensure sufficient sample sizes within each ethnic group, and to follow precedent studies using the same data sources,[12,33,34] as well as the ethnic groupings used in the 2011 England and Wales census.[35] If the two sources for ethnicity provided different categorizations, then we used ethnicity as defined in the CPRD since these data have been shown to be more reliable than HES inpatient data.[33] If no ethnicity data were available within the CRPD, we categorized ethnicity using HES data. If ethnicity was not recorded in either source, we considered ethnicity as missing and the individual was excluded from the complete case analyses.

We used the small area IMD to define deprivation. The IMD combines seven indices which capture dimensions of deprivation at the Lower-Layer Super Output Area or neighbourhood level, and ranks each neighbourhood from 1 to 32 844.[29] Neighbourhood rankings were divided into quintiles and used to compare relative levels of deprivation among people in this study living in different neighbourhoods in England.

We also considered how the proportion of patients receiving the alternative second-line treatments may differ according to calendar time, recognizing that the dissemination and awareness of the safety and efficacy of SGLT2 inhibitors for patients with T2DM increased over the time period, with the publication of important RCT results.[36-38] We considered this hypothesis in grouping calendar time into years 2014, 2015 to 2016, 2017 to 2018, and 2019 to 2020.

2.5 | Covariates

We adjusted for several additional variables, derived from data captured before or on the same day as the index date. These were sex, age, duration of time on metformin monotherapy, number of patients registered at the individual's general practice, geographic region, co-prescriptions for renin-angiotensin system inhibitors and/or statins, history of proteinuria, history of hypoglycaemia, clinical measures (body mass index [BMI] and HbA1c), smoking status, alcohol intake, and comorbidities at the time of second-line antidiabetic treatment initiation. Comorbidities included CKD stage (no known CKD, stages 1, 2, 3a, and 3b, assigned using the latest recorded eGFR), cancer (any), blindness, congestive heart failure, previous myocardial infarction (MI), unstable angina, previous stroke, other ischaemic heart disease, and uncontrolled hypertension based on the most recent blood pressure measures recorded in primary care. We defined prevalent CVD as a composite of heart failure, ischaemic heart disease, unstable angina, previous myocardial infarction, or previous stroke.

2.6 | Treatment prescribed

Our dependent variable of interest was incident second-line oral antidiabetic treatment prescribed (SUs, DPP-4 inhibitors, or SGLT2 inhibitors, in addition to metformin), defined using CPRD prescribing data.

2.7 | Analysis

We described baseline characteristics of the study population stratified by ethnicity and IMD. We then built mixed-effect multivariable, multinomial logistic regression models which compared the odds of initiating SGLT2 inhibitors and DPP-4 inhibitors versus SUs (reference outcome), as well as, in a separate model, SGLT2 inhibitors versus DPP-4 inhibitors (reference outcome), first adjusting for just age and sex. In the final adjusted model, we adjusted for all covariates, as well as mutual adjustment for ethnicity and deprivation (fixed effects) and clustering at the Clinical Commissioning Group (CCG) level (random effect). Because odds ratios can be misleading, particularly when the outcome is common,[39] we calculated and plotted predicted percentages from the adjusted model using recycled predictions.[40] These percentages refer to people prescribed each second-line antidiabetic treatment stratified by ethnicity, and separately by deprivation, while still adjusting for all measured covariates, and accounting for clustering at the CCG level. We obtained *P* values from Wald tests comparing the predicted percentage of being prescribed one of SUs, DPP-4 inhibitors or SGLT2 inhibitors by each non-White ethnic group versus White ethnic group and by deprivation Quintiles 2 to 5 versus deprivation Quintile 1. We also performed joint tests to test whether predicted percentages for each ethnic group or for each deprivation quintile were equal for each second-line antidiabetic treatment. These percentages and *P* values were used to support our final conclusions on disparities in second-line antidiabetic treatment prescribing by ethnicity or by deprivation.[39] We then stratified the adjusted predicted percentages by prevalent CVD status at baseline to determine if there were differences in prescribing by ethnicity or by deprivation quintile according to prevalent CVD status.

In the secondary analyses, we compared the change in odds ratios between the fixed-effect model (model including ethnicity, deprivation, and all covariates) and the mixed-effect model (the fixed effect model plus accounting for CCG clustering as a random effect). We also compared the final adjusted mixed-effect model with and without adjustment for deprivation to observe any changes in ethnic disparities in second-line antidiabetic treatment prescribed when adding this variable to the multivariable model, since deprivation could be a mediator between ethnicity and second-line antidiabetic treatment prescribed. Because awareness of the cardio- and kidney-protective effects of SGLT2 inhibitors versus placebo have increased over time, we considered year of second-line oral antidiabetic treatment initiation as the independent variable of interest, repeating the main analysis to observe any differences in second-line antidiabetic treatment prescribed over time, overall and stratified by prevalent CVD status. Finally, we investigated whether there were interactions between (1) ethnicity and IMD, (2) ethnicity and calendar time, and (3) deprivation and calendar time, informed by the results of likelihood ratio tests on the final adjusted multinomial models, and by joint tests on whether predicted percentages for interaction terms were equal for each of the three second-line treatments prescribed.

Data management and analyses were performed using Stata 17.

3 | RESULTS

3.1 | Baseline characteristics

The study population included 36 023 people with complete data on all variables of interest who initiated second-line oral antidiabetic treatment during the study period with linked secondary care data (Figure 1). Eighty-four percent of the cohort were White, 10% were South Asian, 4% were Black, and 1% were Mixed/other ethnicity. We excluded 6150 people with missing data for at least one variable, including 348 with missing ethnicity data and 20 with missing IMD data (Table S4).

Overall, 41% of the cohort was female, with a mean age of 59 years (Table 1). People of White ethnicity were more likely to be male (60%) and older (mean age 60.1 years) compared with people of South Asian ethnicity (53% male, mean age 52.6 years), Black ethnicity (49% male, mean age 55.2 years), and people of Mixed/other ethnicity (58% male, mean age 55.4 years). People of South Asian and Black ethnicities were over-represented in the lowest IMD quintile (34% and 46%, respectively). Recorded CVD prevalence was 24% overall and was lower in people of Black (15%) compared to White ethnicity (23%). Mean BMI was highest in people of White ethnicity (33.8 kg/m^2) and lowest in people with South Asian ethnicity (30.2 kg/m^2).

After stratifying by IMD quintile, we found that people in the most deprived quintile were over-represented (25%) and people in the least deprived quintile were under-represented (16%; Table S2). People in the most deprived quintile were younger (mean age 56.6 years) compared with the least deprived quintile (mean age 61.8 years). The most deprived quintile included a higher proportion of South Asian and Black people (13% and 7%, respectively) compared with the least deprived quintile (5% and 1%, respectively).

People with missing covariate information were similar according to sex, age, IMD and prevalent CVD versus those with fully observed covariate information (Tables S3,S4).

3.2 | Ethnicity and second-line antidiabetic treatment choice

In people of Black ethnicity, the most common second-line treatment during the study period was SUs (593 [46%]), whereas for all other ethnic groups DPP-4 inhibitors was the most commonly prescribed second-line treatment (13 398 [43%], 1530 [44%], and 249 [45%], among people of White, South Asian, and Mixed/other ethnic groups, respectively; Table 2). SGLT2 inhibitors were the least common second-line treatment across all ethnic groups, ranging from 14% prescribed, among people of Black ethnicity, to 21%, among people of White ethnicity (Table 2).

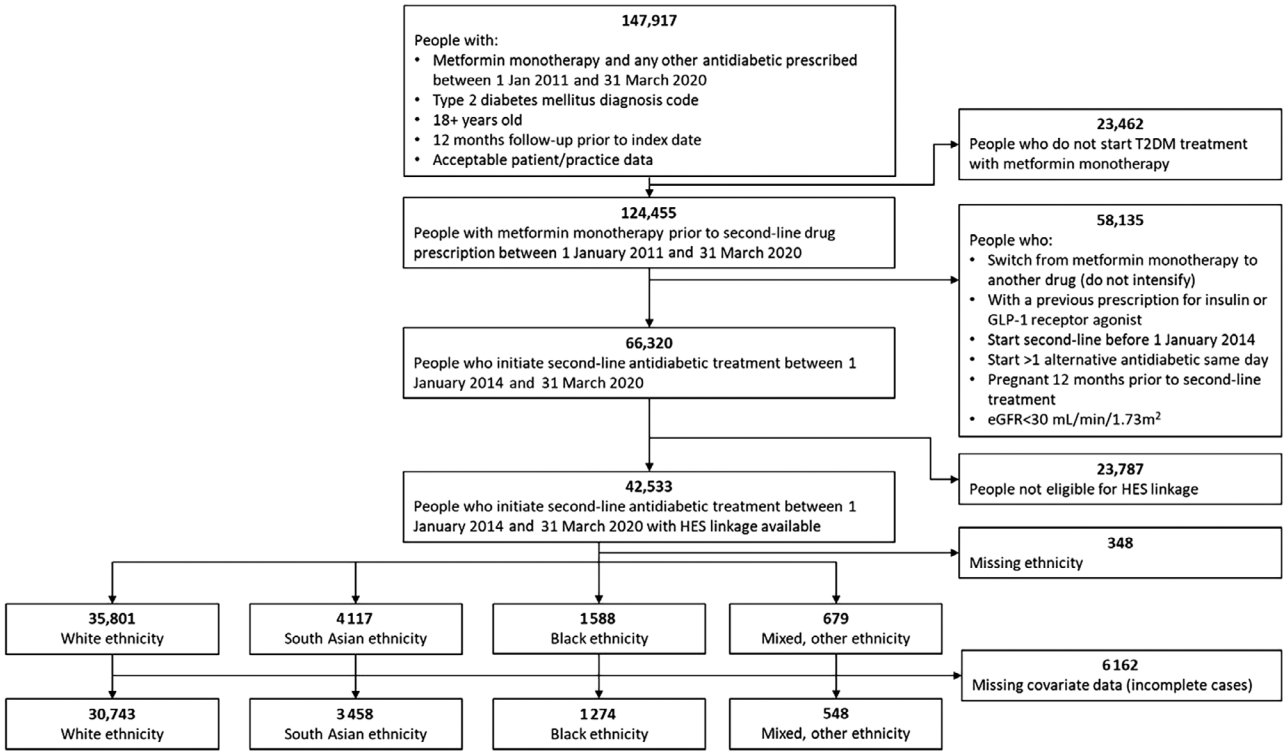


FIGURE 1 Flow diagram illustrating selection of study population diagnosed with type 2 diabetes mellitus (T2DM) and initiating second-line antidiabetic treatment. eGFR, estimated glomerular filtration rate; HES, Hospital Episodes Statistics.

There was some evidence that adjusted predicted percentages for being prescribed SGLT2 inhibitors were greatest for White people (21% [95% CI 19–23%]) compared with South Asian people (20% [95% CI 18–22%]), Black people (19% [95% CI 16–22%]) and Mixed/other people (17% [95% CI 14–21%]; $P = 0.003$ [Figure 2, Table S5]). There was no evidence of differences in adjusted predicted percentages for being prescribed DPP-4 inhibitors or SUs according to ethnicity (Figure 2, Table S5). The results from the multinomial, multivariable logistic regression model used to calculate these adjusted predicted percentages are described in Table S6.

3.3 | Social deprivation and second-line antidiabetic treatment choice

The crude proportion of people prescribed each second-line antidiabetic treatment option across deprivation quintiles are presented in Table 2.

There was some evidence of a small difference in adjusted predicted percentages of people prescribed SGLT2 inhibitors according to deprivation: 19% (95% CI 17–21%) were prescribed SGLT2 inhibitors in the most deprived quintile, and 22% (95% CI 20–24%) were prescribed SGLT2 inhibitors in the least deprived quintile ($P < 0.001$; Figure 2, Table S5). Conversely, there was some evidence that people in the most deprived quintile had a small increase in the adjusted predicted percent of being prescribed DPP-4 inhibitors 44% (95% CI 42–47%) versus 42% (95% CI 39–45%) of people in the least deprived quintile ($P = 0.04$). There was no evidence of any differences in the adjusted predicted percentages of people prescribed SUs according to deprivation ($P = 0.26$).

3.4 | Second-line antidiabetic treatment prescribed among people with prevalent CVD

When stratifying by prevalent CVD status ($n = 8466$ with prevalent CVD, $n = 27\,557$ without prevalent CVD), adjusted predicted percentages showed no substantial differences in SU prescribing across ethnicities and across deprivation quintiles. However, adjusted predicted percentages showed evidence of a slightly higher proportion of people prescribed DPP-4 inhibitors with prevalent CVD versus no CVD across all ethnicities. Conversely, adjusted predicted percentages showed evidence of slightly less SGLT2 inhibitor prescribing among people with prevalent CVD versus no CVD across all ethnicities (Table S7).

3.5 | Secondary analyses

Results were similar in the fixed-effect and mixed-effect models (Table S6), and we did not observe any substantial mediation by deprivation level on the association between ethnicity and second-line treatment prescribed (Table S8).

TABLE 1 Baseline characteristics of study population (at time of second-line treatment initiation)

	Total 36 023 (100%)	White 30 743 (85%)	South Asian 3458 (10%)	Black 1274 (4%)	Mixed/Other 548 (1%)
Female, n (%)	14 643 (41)	12 149 (40)	1616 (47)	646 (51)	232 (42)
Age at baseline, n (%)					
18-49 years	7734 (21)	5711 (19)	1450 (42)	398 (31)	175 (32)
50–59 years	11 128 (31)	9396 (31)	1059 (31)	483 (38)	190 (35)
60-69 years	9622 (27)	8613 (28)	662 (19)	236 (19)	111 (20)
70+ years	7539 (21)	7023 (23)	287 (8)	157 (12)	72 (13)
Year of second-line treatment initiation, n (%)					
2014	4092 (11)	3584 (12)	318 (9)	139 (11)	51 (9)
2015-2016	10 910 (30)	9467 (31)	915 (26)	364 (29)	164 (30)
2017-2018	12 221 (34)	10 379 (34)	1186 (34)	459 (36)	197 (36)
2019-2020	8800 (24)	7313 (24)	1039 (30)	312 (24)	136 (25)
Years on first-line (metformin monotherapy), n (%)					
<0.5	6676 (19)	5710 (19)	539 (16)	320 (25)	107 (20)
0.5-0.99	3654 (10)	3183 (10)	288 (8)	123 (10)	60 (11)
≥1	25 693 (71)	21 850 (71)	2631 (76)	831 (65)	381 (70)
Median (IQR) no. of patients registered at the person's GP	10 295 (6981-14254)	10 471 (7184-14275)	8793 (5213-12861)	10 357 (6484-15253)	10 674 (7377-15196)
Hospitalization (any) within 1 year prior to second-line treatment initiation, n (%)	10 216 (28)	8818 (29)	873 (25)	361 (28)	164 (30)
IMD quintile, n (%)					
1 (least deprived)	5739 (16)	5322 (17)	271 (8)	45 (4)	101 (18)
2	6484 (18)	5922 (19)	419 (12)	77 (6)	66 (12)
3	6915 (19)	6075 (20)	601 (17)	153 (12)	86 (16)
4	8020 (22)	6473 (21)	995 (29)	416 (33)	136 (25)
5 (most deprived)	8865 (25)	6951 (23)	1172 (34)	583 (46)	159 (29)
HbA1c at baseline, n (%)		9.1	8	9	9.1
<53 mmol/mol (7%)	1318 (4)	1070 (3)	139 (4)	84 (7)	25 (5)
53-74 mmol/mol	19 443 (54)	16 548 (54)	2057 (59)	541 (42)	297 (54)
75+ mmol/mol (9%)	15 262 (42)	13 125 (43)	1262 (36)	649 (51)	226 (41)
Uncontrolled hypertension, based on last recorded blood pressure, n (%)					
Normotensive	9749 (27)	8073 (26)	1155 (33)	351 (28)	170 (31)
Hypertensive	26 274 (73)	22 670 (74)	2303 (67)	923 (72)	378 (69)
Mean (SD) BMI, kg/m^2	33.4 (7.0)	33.8 (7.1)	30.2 (5.8)	32.2 (6.9)	31.3 (6.8)
Smoking status, n (%)					
Non-smoker	7371 (20)	5691 (19)	1145 (33)	389 (31)	146 (27)
Current smoker	9874 (27)	8608 (28)	804 (23)	317 (25)	145 (26)
Ex-smoker	18 778 (52)	16 444 (53)	1509 (44)	568 (45)	257 (47)
Alcohol status, n (%)					
Non-drinker	3846 (11)	2276 (7)	1199 (35)	239 (19)	132 (24)
Current drinker	22 082 (61)	20 136 (65)	1112 (32)	583 (46)	251 (46)
Ex-drinker	10 095 (28)	8331 (27)	1147 (33)	452 (35)	165 (30)
Co-prescriptions, n (%)					
RAS inhibitors	17 949 (50)	15 700 (51)	1450 (42)	535 (42)	264 (48)
Statins	24 907 (69)	21 472 (70)	2376 (69)	713 (56)	346 (63)
Cancer (any)	4048 (11)	3794 (12)	129 (4)	91 (7)	34 (6)

(Continues)

TABLE 1 (Continued)

	Total 36 023 (100%)	White 30 743 (85%)	South Asian 3458 (10%)	Black 1274 (4%)	Mixed/Other 548 (1%)
Macrovascular comorbidities, n (%)					
CVD composite[a]	8466 (24)	7589 (25)	600 (17)	191 (15)	86 (16)
Amputation	283 (1)	270 (1)	8 (0)	<5(0)	<5 (0)
Heart failure	2110 (6)	1936 (6)	109 (3)	51 (4)	14 (3)
Myocardial infarction	2521 (7)	2288 (7)	174 (5)	37 (3)	22 (4)
Stroke	1640 (5)	1470 (5)	103 (3)	54 (4)	13 (2)
Ischaemic heart disease	6823 (19)	6117 (20)	495 (14)	134 (11)	77 (14)
Unstable angina	1175 (3)	1049 (3)	83 (2)	29 (2)	14 (3)
Microvascular comorbidities, n (%)					
eGFR at baseline category (mL/min/1.73 m^2)					
No known CKD (eGFR missing)	675 (2)	612 (2)	31 (1)	11 (1)	21 (4)
90+ (Stage 1)	21 391 (59)	17 676 (57)	2694 (78)	640 (50)	381 (70)
60-89 (Stage 2)	11 913 (33)	10 608 (35)	648 (19)	539 (42)	118 (22)
45-59 (Stage 3a)	1585 (4)	1432 (5)	65 (2)	66 (5)	22 (4)
30-44 (Stage 3b)	459 (1)	415 (1)	20 (1)	18 (1)	6 (1)
Blindness	486 (1)	432 (1)	33 (1)	13 (1)	8 (1)
Hypoglycaemia	320 (1)	272 (1)	23 (1)	16 (1)	9 (2)
Proteinuria	2586 (7)	2202 (7)	260 (8)	80 (6)	44 (8)

[a]CVD composite: heart failure, ischaemic heart disease, myocardial infarction, stroke, unstable angina.
Abbreviations: BMI, body mass index; CPRD, Clinical Practice Research Datalink; CVD, cardiovascular; DPP-4, dipeptidyl peptidase-4; eGFR, estimated glomerular filtration rate; GP, general practice; HES, Hospital Episode Statistics; IMD, Index of Multiple Deprivation; IQR, interquartile range; RAS, renin-angiotensin system; SGLT2, sodium-glucose cotransporter 2; SD, standard deviation; SU, sulphonylureas; UK, United Kingdom.

TABLE 2 Crude proportions of second-line treatment prescribed by ethnicity or by deprivation

	Second-line antidiabetic treatment prescribed		
Variable of interest	**Metformin-SUs**	**Metformin-DPP-4 inhibitors**	**Metformin-SGLT2 inhibitors**
Ethnicity, *n* (row %)			
White	11 584 (37)	13 398 (43)	6455 (21)
South Asian	1316 (37)	1530 (44)	667 (19)
Black	593 (46)	524 (40)	179 (14)
Mixed/other	216 (39)	249 (45)	93 (17)
IMD quintile, *n* (row %)			
1 (least deprived)	2155 (37)	2483 (42)	1264 (21)
2	2432 (37)	2833 (43)	1388 (21)
3	2654 (37)	2972 (42)	1476 (21)
4	3112 (38)	3975 (44)	1628 (20)
5 (most deprived)	3356 (37)	3975 (44)	1638 (18)

Abbreviations: DPP-4, dipeptidyl peptidase-4; IMD, Index of Multiple Deprivation; SGLT2, sodium-glucose cotransporter 2; SU, sulphonylurea.

Second-line antidiabetic treatments prescribed changed over time (Tables S9–S11, Figure S1). Briefly, SU prescribing decreased substantially over time (adjusted predicted percent 60% [95% CI 57–62%] in 2014 versus 23% [95% CI 21–24%] in 2019–2020), while prescribing of DPP-4 inhibitors and SGLT2 inhibitors increased over the same time periods (DPP-4 inhibitors: 34% [95% CI 31–36%] in 2014 versus 43% [95% CI 40–46%] in 2019–2020; and SGLT2 inhibitors: 7% [95% CI 5–8%] in 2014 versus 34% [95% CI 32–37%] in 2019–2020; Table S10). Among people with and without prevalent CVD, those with prevalent CVD had consistently lower probabilities of being

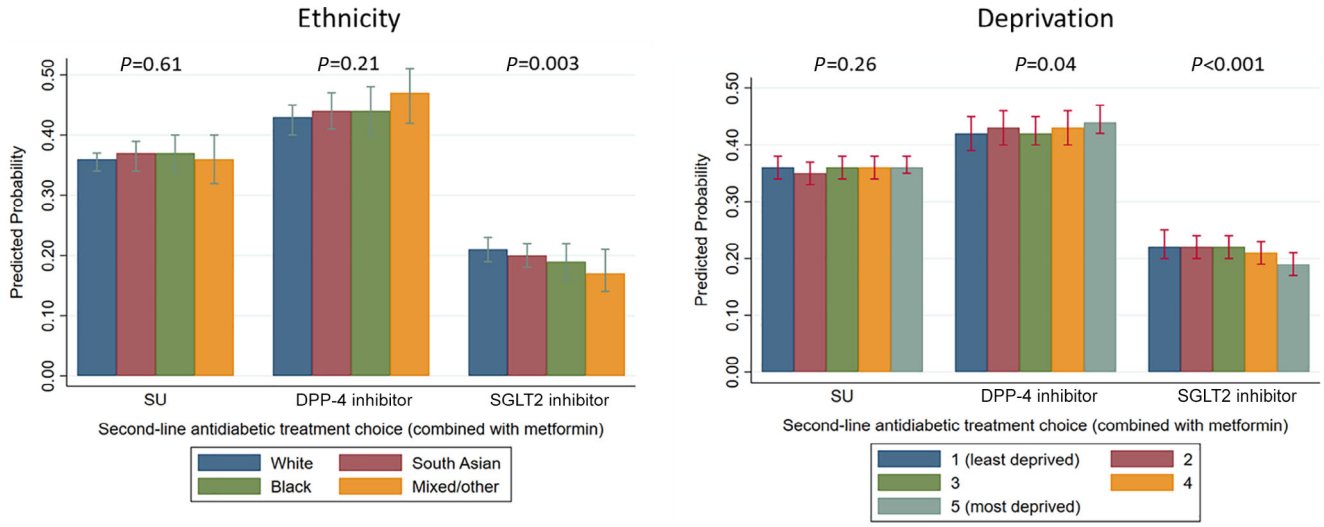


FIGURE 2 Adjusted predicted percentages of second-line antidiabetic treatment prescribed, according to ethnicity or deprivation. *P* values are generated from joint tests comparing the adjusted predicted percentages for being prescribed a particular second-line antidiabetic drug (a sulphonylurea [SU], a dipeptidyl peptidase-4 [DPP-4] inhibitor or a sodium-glucose cotransporter-2 [SGLT2] inhibitor) across ethnic groups or across deprivation levels. Predicted percentages are mutually adjusted for deprivation (ethnicity estimates) and ethnicity (deprivation estimates), as well as the number of patients registered at the patient's general practice, years on first-line treatment category, age category, sex, last glycated haemoglobin value prior to second-line treatment initiation category, body mass index, prevalent heart failure, ischaemic heart disease, myocardial infarction, stroke, unstable angina, renin-angiotensin system inhibitors and/or statin co-prescription, chronic kidney disease category, blood pressure category, history of proteinuria, blindness, cancer (any), hospitalization (any) in past year, smoking status, alcohol status, region, all as fixed effects, and Clinical Commissioning Group-clustering as a random effect.

prescribed SGLT2 inhibitors compared with those without prevalent CVD across time periods (Table S10).

There was no evidence of an interaction effect between ethnicity and IMD on second-line antidiabetic treatment prescribed on the odds scale ($P = 0.45$). On the adjusted predicted percent scale, there was some evidence that the percent prescribed SGLT2 inhibitors among people of White ethnicity and Mixed/other ethnicity decreased with increasing deprivation ($P < 0.001$ and $P = 0.04$). There was no evidence of an interaction between calendar time and ethnicity ($P = 0.11$), nor between calendar time and deprivation quintile ($P = 0.66$).

4 | DISCUSSION

We found statistically significant, but small absolute differences in SU, DPP-4 inhibitor and SGLT2 inhibitor prescribing as second-line antidiabetic treatment, in combination with metformin, according to ethnicity and deprivation level, after accounting for several covariates and clustering at the CCG level in England. There was some evidence that across ethnic groups and levels of deprivation, people with prevalent CVD had a lower probability of being prescribed SGLT2 inhibitors compared with those without prevalent CVD.

It is reassuring that we did not observe substantial ethnic differences in second-line antidiabetic treatment prescribing, as previous research has described many other ethnic disparities related to T2DM. In the UK, ethnic minorities with T2DM had longer delays in intensification to second-line treatment than White people with T2DM, and experienced greater treatment inertia following identification of uncontrolled HbA1c.[12] In the United States, which does not have a universal healthcare system, ethnic minorities are less likely to be prescribed SGLT2 inhibitors at any time after T2DM diagnosis compared to White people, even after adjusting for deprivation level.[41]

This cross-sectional study used one of the largest primary care datasets in the world.[26,27] While our results suggest statistical evidence of a lower percentage of ethnic minorities and people from more deprived areas in England being prescribed SGLT2 inhibitors compared with people of White ethnicity, these differences were not substantial and unlikely to represent major disparities in T2DM care. Factors such as willingness to try newer treatments on the part of both the healthcare team and the patient are unmeasured, and could have contributed to the small differences in second-line antidiabetic treatments prescribed that we observed in this study.

It is, however, concerning that people with prevalent CVD had a lower probability of receiving SGLT2 inhibitors versus those without, since trials comparing SGLT2 inhibitors versus placebo have shown substantial improvements in diabetic-related outcomes among those with atherosclerotic CVD, heart failure, and kidney disease.[38,42,43] However, national and international guidance/guidelines recommending SGLT2 inhibitors among those with prevalent CVD were only updated towards the end of our study period.[22,44] We hope future research shows increased SGLT2 inhibitor prescribing in those with prevalent CVD after 2020.

This study has some limitations. We did not include people with missing documentation of ethnicity; however, only 348 people (0.8%) in the total sample had missing ethnicity data, which limited our ability to include these people as a separate group in our analyses. Further, some exposure misclassification may exist in our study, since deprivation was measured at the small area/neighbourhood level and patient-level deprivation status may differ from this measure. There is also likely to be residual confounding according to other clinical characteristics such as history of alcohol misuse, pancreatic disease, urinary tract infections, mycotic urinary infections, and unobserved factors such as prescriber characteristics and patient frailty status. Our data came from electronic health records, which are not designed primarily for research and thus some degree of misclassification is expected for covariates, particularly those which are not necessarily recorded on the same day as second-line treatment initiation (eg, HbA1c, BMI, eGFR, blood pressure). Finally, we were limited to prescribing data from primary care to define second-line antidiabetic treatment choice. We were unable to use dispensing data from pharmacies since these data are not available for linkage to CPRD data, nor were we able to determine if treatment initiation occurred during an inpatient stay (secondary care), where prescription of newer drug classes may be relatively more likely. We adjusted for any hospitalization in the past year to try and account for this. However, even if a prescription was initiated or recommended by specialist care, primary care would probably continue prescriptions of these therapies.

In conclusion, we found statistically significant, but small differences in second-line oral antidiabetic treatment prescribing by ethnicity and social deprivation status in England. These differences are unlikely to be clinically important. We consider it encouraging that, after accounting for various clinical characteristics and variation at the CCG level, ethnic minorities and people from more deprived backgrounds did not have substantially lower probabilities of being prescribed SGLT2 inhibitors compared with DPP-4 inhibitors and SUs. Future work should investigate other factors at the individual and local CCG level which may drive treatment choice to understand how these treatments are used in routine care, and to highlight the need for future research to directly evaluate the comparative effectiveness of these three second-line antidiabetic treatment choices to optimize oral antidiabetic treatment prescribing.

AUTHOR CONTRIBUTIONS

Patrick Bidulka, David Lugo-Palacios, Stephen O'Neill, Anirban Basu, Kamlesh Khunti and Richard Grieve conceived and designed the study. Patrick Bidulka and Stephen O'Neill conducted the data management and analyses. Patrick Bidulka wrote the first draft of the manuscript. All authors (Patrick Bidulka, Rohini Mathur, David Lugo-Palacios, Stephen O'Neill, Anirban Basu, Richard Silverwood, Paul Charlton, Andrew Briggs, Liam Smeeth, Amanda Adler, Ian Douglas, Kamlesh Khunti and Richard Grieve) reviewed and commented on the manuscript, and approved the final version for submission.

ACKNOWLEDGMENTS

This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data are provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the authors alone. The study was approved by the Independent Scientific Advisory Committee (approval number: 20_064). Kamlesh Khunti is supported by the National Institute for Health Research (NIHR; grant number NIHR128490) Applied Research Collaboration East Midlands (ARC EM) and the NIHR Leicester Biomedical Research Centre (BRC). Rohini Mathur is supported by a Sir Henry Wellcome Postdoctoral Fellowship (201375/Z/16/Z). Patrick Bidulka and Richard Grieve are joint guarantors and take full responsibility for this work as a whole, including the study design, access to data, and the decision to submit and publish this manuscript. This work was supported by National Institute for Health and Care Research (NIHR) grant number NIHR128490.

CONFLICT OF INTERESTS

All authors have completed an ICJME form. Patrick Bidulka, Stephen O'Neill, Anirban Basu, Richard Silverwood and Liam Smeeth have nothing to declare. Kamlesh Khunti has acted as a consultant, speaker or received grants for investigator-initiated studies for Astra Zeneca, Novartis, Novo Nordisk, Sanofi-Aventis, Lilly and Merck Sharp & Dohme, Boehringer Ingelheim, Bayer, Berlin-Chemie AG / Menarini Group, Janssen and Napp. Rohini Mathur has received consulting fees from AMGEN. Paul Charlton sat on an NIHR HTA Commissioning Committee member until September 2021. Anirban Basu is an economic advisor on the DiRECT trial, with ongoing responsibility for economic analysis during the long-term follow-up phase, and has also acted as consultant to GlaxoSmithKline, Merck, Novo Nordisk and Boehringer Ingelheim in relation to their diabetes products. Amanda Adler receives salary from the NIHR BRC via the Oxford Centre for Diabetes, Endocrinology and Metabolism. Ian Douglas holds an unrestricted research grant from GSK and holds shares in GSK. Richard Grieve sits on the NIHR commissioning committee.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/dom.14874.

DATA AVAILABILITY STATEMENT

Due to data-sharing restrictions, we cannot share the data used in this study directly. However, researchers may apply to use CPRD data linked with other health datasets. Please see the CPRD website for further instruction https://cprd.com/. Codelists to create exposure, outcome, and covariates will be published on LSHTM DataCompass https://datacompass.lshtm.ac.uk/.

ETHICS APPROVAL

This research was approved by the London School of Hygiene & Tropical Medicine ethics committee (reference 21 395) and the Independent Scientific Advisory Committee (reference 20_064).

TRANSPARENCY STATEMENT

Patrick Bidulka, as corresponding author, confirms that the manuscript is an honest, accurate, and transparent account of the study being reported and that no important aspects of the study have been omitted, and that any discrepancies from the study as planned have been explained.

ORCID

Patrick Bidulka https://orcid.org/0000-0001-7644-2030
Kamlesh Khunti https://orcid.org/0000-0003-2343-7099

REFERENCES

1. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep*. 2014;129(Suppl 2):19-31. doi:10.1177/00333549141291S206
2. Mackenbach JP, Stirbu I, Roskam A-JR, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008 2008; 358(23):2468-2481. doi:10.1056/NEJMsa0707519
3. NICE: 20 years of evidence-based decision making. Web. National Institute for Health and Care Excellence (NICE). Accessed September 10, 2021. https://indepth.nice.org.uk/20-years-of-NICE/index.html
4. NHS body 'to end postcode prescribing'. Web. BBC News. Accessed September 10, 2021. http://news.bbc.co.uk/1/hi/health/271522.stm
5. Our principles: The principles that guide the development of NICE guidance and standards. Web. National Institute for Health and Care Excellence. https://www.nice.org.uk/about/who-we-are/our-principles
6. Scarborough PB, Bhatnagar P, Kaur A, Smolina K, Wickramasinghe K, Rayner M. *Ethnic differences in Cardiovascular Disease 2010*. 2010. https://www.bhf.org.uk/informationsupport/publications/statistics/ethnic-differences-in-cardiovascular-disease-2010#
7. Caskey F, Gavin D. *Kidney Health Inequalities in the UK: an agenda for change*. 2018. https://kidneyresearchuk.org/wp-content/uploads/2019/09/Health_Inequalities_lay_report_FINAL_WEB_20190311.pdf
8. Health Survey for England 2004: The Health of Minority Ethnic Groups. 2005. https://files.digital.nhs.uk/publicationimport/pub01xxx/pub01209/heal-surv-hea-eth-min-hea-tab-eng-2004-rep.pdf
9. Ng Fat L. Health survey for England 2019 adults. *Health*. 2020;9-13: 12-13. https://files.digital.nhs.uk/23/6B5DEA/HSE19-Adult-health-rep.pdf
10. Mathur R, Bhaskaran K, Edwards E, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004–2014. *BMJ Open*. 2017;7(2):e014444. doi:10.1136/bmjopen-2016-014444
11. Tillin T, Hughes AD, Mayet J, et al. The relationship between metabolic risk factors and incident cardiovascular disease in Europeans, South Asians, and African Caribbeans: SABRE (Southall and Brent revisited)—a prospective population-based study. *J Am Coll Cardiol* 2013;61(17):1777-1786. doi:https://doi.org/10.1016/j.jacc.2012.12.046
12. Mathur R, Farmer RE, Eastwood SV, Chaturvedi N, Douglas I, Smeeth L. Ethnic disparities in initiation and intensification of diabetes treatment in adults with type 2 diabetes in the UK, 1990–2017: a cohort study. *PLoS Med*. 2020;17(5):e1003106. doi:10.1371/journal.pmed.1003106
13. Bachmann MO, Eachus J, Hopper CD, et al. Socio-economic inequalities in diabetes complications, control, attitudes and health service use: a cross-sectional study. *Diabet Med*. 2003;20(11):921-929. https://doi.org/10.1046/j.1464-5491.2003.01050.x
14. Khunti K, Gomes MB, Pocock S, et al. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: a systematic review. *Diabetes Obes Metab*. 2018;20(2):427-437. 10.1111/dom.13088
15. Khunti K, Seidu S. Therapeutic inertia and the legacy of dysglycemia on the microvascular and macrovascular complications of diabetes. *Diabetes Care*. 2019;42(3):349-351. doi:10.2337/dci18-0030
16. Khunti K, Wolden ML, Thorsted BL, Andersen M, Davies MJ. Clinical inertia in people with Type 2 diabetes. *Diabetes Care*. 2013;36(11): 3411-3417. doi:10.2337/dc13-0331
17. Khunti S, Khunti K, Seidu S. Therapeutic inertia in type 2 diabetes: prevalence, causes, consequences and methods to overcome inertia. *Ther Adv Endocrinol Metab*. 2019;10:2042018819844694. doi:10.1177/2042018819844694
18. Schernthaner G, Shehadeh N, Ametov AS, et al. Worldwide inertia to the use of cardiorenal protective glucose-lowering drugs (SGLT2i and GLP-1 RA) in high-risk patients with type 2 diabetes. *Cardiovasc Diabetol*. 2020;19(1):185. doi:10.1186/s12933-020-01154-w
19. NG28: Type 2 diabetes in adults: management. NICE. https://www.nice.org.uk/guidance/ng28
20. Canagliflozin in combination therapy for treating type 2 diabetes: Technology appraisal guidance TA315. Web. NICE. Updated June 25, 2014. https://www.nice.org.uk/guidance/ta315/chapter/2-The-technology
21. NICE guideline [NG28]: Type 2 diabetes in adults: management. Web. NICE. Accessed March 3, 2022, 2022. https://www.nice.org.uk/guidance/ng28/chapter/Recommendations#reviewing-drug-treatments
22. Buse JB, Wexler DJ, Tsapas A, et al. Update to: management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of diabetes (EASD). *Diabet care*. 2019, 2020;43(2):487-493. doi:10.2337/dci19-0066
23. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768
24. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142
25. Bidulka P, O'Neill S, Basu A, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open*. 2021;11:e046912.
26. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol* Jun 2015;44(3):827-36. doi:10.1093/ije/dyv098
27. Wolf A, Dedman D, Campbell J, et al. Data resource profile: clinical practice research datalink (CPRD) aurum. *Int J Epidemiol*. 2019;48(6): 1740-1740g. doi:10.1093/ije/dyz034
28. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data resource profile: hospital episode statistics admitted patient care (HES APC). *Int J Epidemiol*. 2017;46(4):1093-1093i. doi:10.1093/ije/dyx015
29. The English Indices of Deprivation 2015 (GOV.UK) (2015).
30. Linked HES-ONS mortality data. Web. NHS Digital. Accessed March 4, 2022. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data#ons-mortality-data
31. Wilding J, Godec T, Khunti K, et al. Changes in HbA1c and weight, and treatment persistence, over the 18 months following initiation of second-line therapy in patients with type 2 diabetes: results from the United Kingdom clinical practice research datalink. *BMC Med*. 2018; 16(1):116. doi:10.1186/s12916-018-1085-8
32. TA390: Canagliflozing, dapagliflozin and empagliflozin as monotherapies for treating type 2 diabetes. NICE.
33. Mathur R, Bhaskaran K, Chaturvedi N, et al. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health*. 2014;36(4):684-692. doi:10.1093/pubmed/fdt116

34. Eastwood SV, Mathur R, Sattar N, Smeeth L, Bhaskaran K, Chaturvedi N. Ethnic differences in guideline-indicated statin initiation for people with type 2 diabetes in UK primary care, 2006–2019: a cohort study. *PLoS Med*. 2021;18(6):e1003672. doi:10.1371/journal.pmed.1003672
35. 2011 Census Variables and Classifications: 2014. 2014. Accessed May 3, 2022. https://www.ons.gov.uk/census/2011census/2011censusdata/2011censususerguide/variablesandclassifications
36. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med*. 2017;377(7):644-657. doi:10.1056/NEJMoa1611925
37. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2019;380(4):347-357. doi:10.1056/NEJMoa1812389
38. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373(22):2117-2128. doi:10.1056/NEJMoa1504720
39. Norton EC, Dowd BE, Maciejewski ML. Odds ratios—current best practice and use. *JAMA*. 2018;320(1):84-85. doi:10.1001/jama.2018.6971
40. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol*. 2014;43(3):962-970. doi:10.1093/ije/dyu029
41. Eberly LA, Yang L, Eneanya ND, et al. Association of Race/ethnicity, gender, and socioeconomic status with sodium-glucose cotransporter 2 inhibitor use among patients with diabetes in the US. *JAMA Netw Open*. 2021;4(4):e216139. doi:10.1001/jamanetworkopen.2021.6139
42. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. *N Engl J Med*. 2019;380(24):2295-2306. doi:10.1056/NEJMoa1811744
43. McMurray JJV, Solomon SD, Inzucchi SE, et al. Dapagliflozin in patients with heart failure and reduced ejection fraction. *N Engl J Med*. 2019;381(21):1995-2008. doi:10.1056/NEJMoa1911303
44. Guideline update: Type 2 diabetes in adults: management; Draft for consultation, September 2021. Web. National Institute of Health and Care Excellence. Accessed September 9, 2021. https://www.nice.org.uk/guidance/GID-NG10246/documents/draft-guideline

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bidulka P, Mathur R, Lugo-Palacios DG, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes Obes Metab*. 2023;25(1):282-292. doi:10.1111/dom.14874

3.2. Relevance to my thesis

This research paper addresses thesis objective 1A: to examine inequalities in second-line antidiabetic treatment prescribing according to sociodemographic patient characteristics. In this study, I found statistically significant but small disparities in second-line antidiabetic treatment prescribing by ethnicity and deprivation (Appendix C.1.).

In this study, I aimed to investigate second-line oral antidiabetic treatment prescribed and its association with an individual's ethnicity and deprivation status, based on the area where individuals live. Previous work has described substantial variation in second-line antidiabetic prescribing at the CCG level, which is unrelated to measured patient characteristics.[1] There are many factors which could drive this variation in prescribing at the CCG level, including budgetary decisions, engagement with the pharmaceutical industry, and hospital and specialist influence.[2] Thus, using a statistical model to acknowledge this clustering is necessary when calculating point estimates and standard errors for the associations of interest.[3,4]

There are several approaches to account for this clustering; namely, a mixed effect model with CCG modelled as a random effect, generalised estimating equations (GEE),[5] and using robust standard errors.[6] In this study, I chose to use a mixed model, including CCG as a random effect, since simulation work has shown that these can provide unbiased, statistically efficient estimates across a multitude of settings.[3,4] Using a mixed effect or hierarchical model, rather than GEE, allowed me to estimate the individual-level association between second-line antidiabetic treatment prescribing and sociodemographic characteristics, rather than averaging at the population-level.

NICE guidance recognises the usefulness of using routinely collected health data to study inequalities in healthcare services and treatment.[7] The lack of any large differences in the predicted probability of being prescribed alternative second-line antidiabetic treatments by ethnicity and deprivation status was reassuring in the context of other stark ethnic and socioeconomic inequalities in health services and treatment described in the UK[8,9] and many countries globally.[10] Previous work has shown that people of ethnic minorities more commonly experience treatment inertia for indicated antidiabetic treatments[11] and statins[12] compared to people of white ethnicity in England. In addition, people of lower SES are less likely to be prescribed newer antidiabetic treatments (e.g., SGTL2i) compared with people of higher SES in the USA.[13] These sociodemographic disparities in treatment were less strong in my own research, after accounting for measured confounders and clustering by CCG.

This research paper also studied inequalities in crude proportion and the adjusted predicted probability of being prescribed SGLT2i compared with DPP4i and SU over the follow-up period (**Table 3.1.**). I found that people with prevalent CVD at second-line antidiabetic treatment initiation had a

consistently lower predicted probability of being prescribed SGLT2i compared with those without prevalent CVD. This inequality in treatment was concerning, since published placebo controlled RCTs published over the course of the study period suggested that SGLT2i were efficacious at reducing the risk of CVD events like heart failure hospitalisation. This finding was in line with other evidence from the UK suggesting a lower probability of people with prevalent CVD were being prescribed SGLT2i.[14-16] In the research paper, I suggest future work should monitor if this disparity decreases over time, particularly as new evidence accumulates that this drug class is beneficial for the subgroup of people with T2DM that have prevalent CVD.

Table 3.1. Adjusted predicted percentages of second-line treatment prescribed by year of second-line treatment initiation, overall and stratified by prevalent CVD status (Supplementary table 10 in the supplementary materials of the published paper).

Outcome	Year of second-line initiation	Overall, adjusted predicted probability (95% CI)	No prevalent CVD, Adjusted predicted probability (95% CI)	Prevalent CVD, Adjusted predicted probability (95% CI)
SU	2014	0.60 (0.57-0.62)	0.60 (0.58-0.62)	0.60 (0.57-0.63)
	2015-16	0.45 (0.43-0.47)	0.45 (0.42-0.47)	0.45 (0.43-0.48)
	2017-18	0.29 (0.27-0.31)	0.29 (0.27-0.31)	0.30 (0.28-0.32)
	2019-20	0.23 (0.21-0.24)	0.22 (0.21-0.24)	0.24 (0.22-0.26)
DPP4i	2014	0.34 (0.31-0.36)	0.33 (0.30-0.36)	0.35 (0.32-0.38)
	2015-16	0.41 (0.39-0.44)	0.41 (0.38-0.43)	0.44 (0.41-0.47)
	2017-18	0.48 (0.45-0.51)	0.47 (0.44-0.49)	0.52 (0.49-0.54)
	2019-20	0.43 (0.40-0.46)	0.41 (0.38-0.44)	0.48 (0.45-0.51)
SGLT2i	2014	0.07 (0.05-0.08)	0.07 (0.06-0.08)	0.05 (0.04-0.06)
	2015-16	0.14 (0.12-0.16)	0.45 (0.13-0.17)	0.11 (0.09-0.12)
	2017-18	0.23 (0.21-0.25)	0.24 (0.22-0.27)	0.18 (0.16-0.21)
	2019-20	0.34 (0.32-0.37)	0.36 (0.33-0.39)	0.28 (0.26-0.31)

CI: confidence interval; CVD: cardiovascular disease; DPP4i: dipeptidyl peptidase-4 inhibitors; SGLT2i: sodium-glucose co-transporter 2 inhibitors; SU: sulfonylureas

The research paper presented in this chapter builds on previous work which considered the association between many independent variables, including ethnicity and deprivation status, and second-line oral antidiabetic prescribing in English primary care using CPRD data.[15] In this updated cohort study including more recent years of follow-up (2014-19), I chose to focus on the association between

ethnicity and deprivation with second-line antidiabetic treatment prescribing because of the heightened interest in healthcare inequalities from policymakers, healthcare professionals, and patients.[17]

The results from this paper, as well as results from previous work, clinical expertise from the PERMIT project's clinical experts, and DAGs (see Chapters 4 to 5) were important to define potential measured confounders of the causal effect of SU, DPP4i, and SGLT2i as second-line antidiabetic treatments on outcomes investigated in Chapters 5 to 6 of this thesis. These potential confounders are described in greater detail in the protocol paper presented in the next chapter.

3.3. References

1. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

2. Scoggins A, Tiessen J, Ling T, Rabinovich L. *Prescribing in primary care: Understanding what shapes GP's prescribing choices and how might these be changed*. 2007. https://www.rand.org/content/dam/rand/pubs/technical_reports/2007/RAND_TR443.pdf

3. Gomes M, Ng ES, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. *Medical decision making : an international journal of the Society for Medical Decision Making*. Mar-Apr 2012;32(2):350-61. doi:10.1177/0272989x11418372

4. Gomes M, Grieve R, Nixon R, Ng ES-W, Carpenter J, Thompson SG. METHODS FOR COVARIATE ADJUSTMENT IN COST-EFFECTIVENESS ANALYSIS THAT USE CLUSTER RANDOMISED TRIALS. *Health Economics*. 2012;21(9):1101-1118. doi:https://doi.org/10.1002/hec.2812

5. Liang K-Y, Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*. 1986;73(1):13-22. doi:10.2307/2336267

6. Rogers W. Regression standard errors in clustered samples. *Stata Technical Bulletin*. 1994;3(13)

7. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

8. Williams E, Buck D, Babalola G, Maguire D. What are health inequalities. Web. The King's Fund. Accessed 29 March, 2024. https://www.kingsfund.org.uk/insight-and-analysis/long-reads/what-are-health-inequalities#inequalities-in-healthy-life-expectancy

9. What are healthcare inequalities. Web. NHS England. Accessed 29 March, 2024. https://www.england.nhs.uk/about/equality/equality-hub/national-healthcare-inequalities-improvement-programme/what-are-healthcare-inequalities/

10. *Monitoring health inequality: An essential step for achieving health equity*. 2015. https://iris.who.int/bitstream/handle/10665/164530/WHO_HIS_HSI_2015.1_eng.pdf?sequence=1

11. Mathur R, Farmer RE, Eastwood SV, Chaturvedi N, Douglas I, Smeeth L. Ethnic disparities in initiation and intensification of diabetes treatment in adults with type 2 diabetes in the UK, 1990–2017: A cohort study. *PLoS medicine*. 2020;17(5):e1003106. doi:10.1371/journal.pmed.1003106

12. Eastwood SV, Mathur R, Sattar N, Smeeth L, Bhaskaran K, Chaturvedi N. Ethnic differences in guideline-indicated statin initiation for people with type 2 diabetes in UK primary care, 2006–2019: A cohort study. *PLoS medicine*. 2021;18(6):e1003672. doi:10.1371/journal.pmed.1003672

13. Eberly LA, Yang L, Eneanya ND, et al. Association of Race/Ethnicity, Gender, and Socioeconomic Status With Sodium-Glucose Cotransporter 2 Inhibitor Use Among Patients With Diabetes in the US. *JAMA Network Open*. 2021;4(4):e216139-e216139. doi:10.1001/jamanetworkopen.2021.6139

14. Wilkinson S, Williamson E, Pokrajac A, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. *Diabetes, Obesity and Metabolism*. 2020;22(5):847-856. doi:https://doi.org/10.1111/dom.13970

15. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

16. Farmer RE, Beard I, Raza SI, et al. Prescribing in Type 2 Diabetes Patients With and Without Cardiovascular Disease History: A Descriptive Analysis in the UK CPRD. *Clinical Therapeutics*. 2021;43(2):320-335. doi:10.1016/j.clinthera.2020.12.015

17. NICE and health inequalities. National Institute for Health and Care Excellence. Accessed 3 April, 2024. https://www.nice.org.uk/about/what-we-do/nice-and-health-inequalities

CHAPTER 4. PROTOCOL PAPER – PROTOCOL FOR AN OBSERVATIONAL COHORT STUDY INVESTIGATING PERSONALISED MEDICINE FOR INTENSIFICATION OF TREATMENT IN PEOPLE WITH TYPE 2 DIABETES MELLITUS: THE PERMIT STUDY.

OVERVIEW

In this chapter, I present a protocol paper published in *BMJ Open* for which I am the first author. This protocol specifies the clinical and epidemiological justification and methods for a comparative effectiveness analysis of alternative second-line oral antidiabetic treatments in a cohort of people with T2DM in English primary care. In this protocol, I introduce the rationale and evidence gap to be addressed with this research, the data sources to be used, the methods, including the IV analysis, and patient and public involvement in the research project.

This protocol was amended post-publication which is summarised immediately following the published paper in **Table 4.1.** This table is copied from the statistical analysis plan (SAP) published on the PERMIT study website. The SAP provides additional detail, updates, and amendments to the protocol and the research presented in Chapter 5. These amendments were necessary due to challenges with the study design, the statistical analysis, and logistical problems (computing power required for the analyses) while conducting the research.

Following the protocol paper and amendment summary, I include a brief discussion of this paper's relevance to my thesis.

4.1. Published research paper

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMJ Open		
When was the work published?	31 August 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	This is a protocol paper, which describes the analyses undertaken in case study 1 of this thesis (comparative effects of second-line antidiabetic treatments for people with type 2 diabetes mellitus). I lead the design of this study in the lead up to registering for my PhD.		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I led the design of this protocol paper with my supervisor, Richard Grieve. I wrote the first draft of the manuscript. Co-authors reviewed multiple drafts before sending to the journal for publication. I led the peer review responses to the journal for publication.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Richard Grieve
Date	7 April 2024

BMJ Open

Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study

Patrick Bidulka ,[1] Stephen O'Neill,[2] Anirban Basu,[3] Samantha Wilkinson,[4] Richard J Silverwood ,[5] Paul Charlton,[6] Andrew Briggs,[2] Amanda I Adler,[7] Kamlesh Khunti,[8] Laurie A Tomlinson,[1] Liam Smeeth,[1] Ian J Douglas,[1] Richard Grieve[2]

To cite: Bidulka P, O'Neill S, Basu A, *et al*. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open* 2021;**11**:e046912. doi:10.1136/bmjopen-2020-046912

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-046912).

Received 12 November 2020
Accepted 31 August 2021

For numbered affiliations see end of article.

Correspondence to
Mr Patrick Bidulka;
patrick.bidulka1@lshtm.ac.uk

ABSTRACT

Introduction For people with type 2 diabetes mellitus (T2DM) who require an antidiabetic drug as an add-on to metformin, there is controversy about whether newer drug classes such as dipeptidyl peptidase-4 inhibitors (DPP4i) or sodium-glucose co-transporter-2 inhibitors (SGLT2i) reduce the risk of long-term complications compared with sulfonylureas (SU). There is widespread variation across National Health Service Clinical Commissioning Groups (CCGs) in drug choice for second-line treatment in part because National Institute for Health and Care Excellence guidelines do not specify a single preferred drug class, either overall or within specific patient subgroups. This study will evaluate the relative effectiveness of the three most common second-line treatments in the UK (SU, DPP4i and SGLT2i as add-ons to metformin) and help target treatments according to individual risk profiles.

Methods and analysis The study includes people with T2DM prescribed one of the second-line treatments-of-interest between 2014 and 2020 within the UK Clinical Practice Research Datalink linked with Hospital Episode Statistics and Office of National Statistics. We will use an instrumental variable (IV) method to estimate short-term and long-term relative effectiveness of second-line treatments according to individuals' risk profiles. This method minimises bias from unmeasured confounders by exploiting the natural variation in second-line prescribing across CCGs as an IV for the choice of prescribed treatment. The primary outcome to assess short-term effectiveness will be change in haemoglobin A1c (%) 12 months after treatment initiation. Outcome measures to assess longer-term effectiveness (maximum ~6 years) will include microvascular and macrovascular complications, all-cause mortality and hospital admissions during follow-up.

Ethics and dissemination This study was approved by the Independent Scientific Advisory Committee (20-064) and the London School of Hygiene & Tropical Medicine Research Ethics Committee (21395). Results, codelists and other analysis code will be made available to patients, clinicians, policy-makers and researchers.

Strengths and limitations of this study

- ► This large representative study of UK clinical practice will describe variation in second-line antidiabetic treatment for people with type 2 diabetes mellitus.
- ► The instrumental variable (IV) design will minimise bias due to confounding by indication and provide person-level estimates of second-line treatment effectiveness.
- ► The IV design relies on assumptions which can only be partly tested using the data available.
- ► We will not consider less commonly used injectable second-line antidiabetic treatments in the UK, namely glucagon-like peptide-1 receptor agonists and insulin.

INTRODUCTION

Around 3.5 million people in the UK have been diagnosed with type 2 diabetes mellitus (T2DM) accounting for ~10% of National Health Service (NHS) expenditure.[1] This proportion is predicted to rise to ~17% by 2035–2036.[2] T2DM is a progressive disease which requires careful management of blood glucose and diabetes-associated complications.[1] The National Institute for Health and Care Excellence (NICE) recommends metformin as the first-line antidiabetic treatment in people with T2DM.[1] In many cases, people with T2DM need further treatment in addition to metformin monotherapy to maintain sufficient glycaemic control.

NICE guidance recommends several drug classes as add-ons to metformin for first-stage intensification, hereafter referred to as second-line treatment. These include sulfonylureas (SU), pioglitazone, dipeptidyl peptidase-4 inhibitors (DPP4i) or sodium-glucose

co-transporter-2 inhibitors (SGLT2i).[1 3] NICE guidance recommends considering individual clinical circumstances when selecting T2DM drug treatment. For instance, SGLT2i are recommended as second-line treatment if the person is at high risk of hypoglycaemia or when SU are not tolerated or are contraindicated.[1 3] However, these guidelines do not specify a single preferred drug class, either overall or within specific groups sharing clinical characteristics.[1] Research using a representative sample of the UK primary care population up to 2017 showed that SU, DPP4i and SGLT2i, each in combination with metformin, are the most commonly prescribed second-line treatments.[4]

There is wide variation in the proportion of people prescribed these drugs in addition to metformin across NHS Clinical Commissioning Groups (CCGs), who commission local NHS services, suggesting clinician preference may influence treatment choice.[4] In particular, the variation in SGLT2i prescribing suggests that some clinicians may prescribe these drugs even for those patients who are not considered at high risk of hypoglycaemia and therefore eligible for SU.

Similar to NICE guidance, an international consensus statement published in 2018 did not specify a single preferred drug class for second-line treatment, but recommended that choice is 'personalised' to individual characteristics and risk profiles. This statement was updated in 2019 in light of new evidence supporting SGLT2i or glucagon-like peptide-1 receptor agonist (GLP1RA) use after metformin for those with atherosclerotic cardiovascular disease (CVD), or those at high CVD risk.[5] However, regulators, clinicians and patients remain uncertain about how best to tailor second-line antidiabetic treatment based on individual characteristics.

Meta-analyses have reported that compared with other antidiabetic treatments, second-generation or third-generation SU are not associated with higher risk of death or cardiovascular (CV) events.[6 7] A recent CV outcome trial reported that the safety profile of glimepiride (SU drug class) was similar to linagliptin (DPP4i drug class).[8] Several placebo-controlled trials reported that SGLT2i reduced major CV events in people with T2DM.[9–12] While head-to-head randomised controlled trials (RCTs) can provide unbiased estimates of relative effectiveness, the range and number of participants included in head-to-head RCTs of alternative second-line treatments are insufficient to provide reliable estimates of long-term effectiveness according to individual-level risk profiles.[13–15]

Observational studies comparing outcomes of alternative second-line drug regimens have reported that SU, DPP4i and SGLT2i combined with metformin are all associated with haemoglobin A1c (HbA1c) reductions compared with metformin alone[16 17]; however, some reported that SU are associated with higher risk of CV events.[7 18] These observational studies did not recognise that people who receive SU may have been a more severe case mix according to unmeasured prognostic variables (eg, frailty) meaning results are likely biased due to confounding by indication.[19]

Aims and objectives

This study aims to investigate the relative effectiveness of SU, DPP4i or SGLT2i in combination with metformin as second-line antidiabetic drug treatments on key T2DM outcomes, and how treatment decisions should be tailored to an individual's risk factor profile to maximise clinical benefit. We will use advanced quantitative methods to minimise the impact of confounding by indication and allow for heterogeneity according to patient characteristics.

The study's objectives are to: (1) Describe baseline characteristics and treatment patterns overall, and by clinically important subgroups, for SU, DPP4i and SGLT2i in combination with metformin as second-line T2DM treatment; (2) Estimate the relative short-term (12 month) effectiveness of SU, DPP4i or SGLT2i combined with metformin on levels of HbA1c, overall and according to individual risk-factor profiles and (3) Estimate the long-term (maximum ~6 years) effectiveness of SU, DPP4i or SGLT2i combined with metformin on incident micro-vascular and macro-vascular complications, overall and according to individual risk-factor profiles.

METHODS AND ANALYSIS

Data resources

We will identify the study population using the UK Clinical Practice Research Datalink (CPRD),[20 21] a pseudonymised primary care database which includes detailed demographic/lifestyle data, clinical diagnoses and measurements, primary care prescriptions, referrals and laboratory test results for approximately 20% of the UK population. Both the CPRD Gold and Aurum datasets will be used to identify people eligible for inclusion, providing a representative population of people with T2DM eligible for the second-line treatments of interest.[20 21]

Linkage to Hospital Episode Statistics (HES) is available for approximately 70% of English practices and will be used to gather secondary care data for the study population. HES Admitted Patient Care data includes complete in-patient admissions data to all NHS hospitals in England.[22] These secondary care data include admission and discharge dates, diagnoses and other descriptive information (eg, ethnicity). Linkages will also be made to the Office of National Statistics to obtain mortality data and the Index of Multiple Deprivation (IMD) as a person-level proxy of socioeconomic status (SES).

Sample selection/study population

The study population will include people registered with a CPRD-contributing practice, aged 18 years or older, diagnosed with T2DM who intensify antidiabetic treatment from metformin-monotherapy to a combination of metformin and SU, DPP4i or SGLT2i (second-line treatment) between 2014 and 2020 (figure 1).

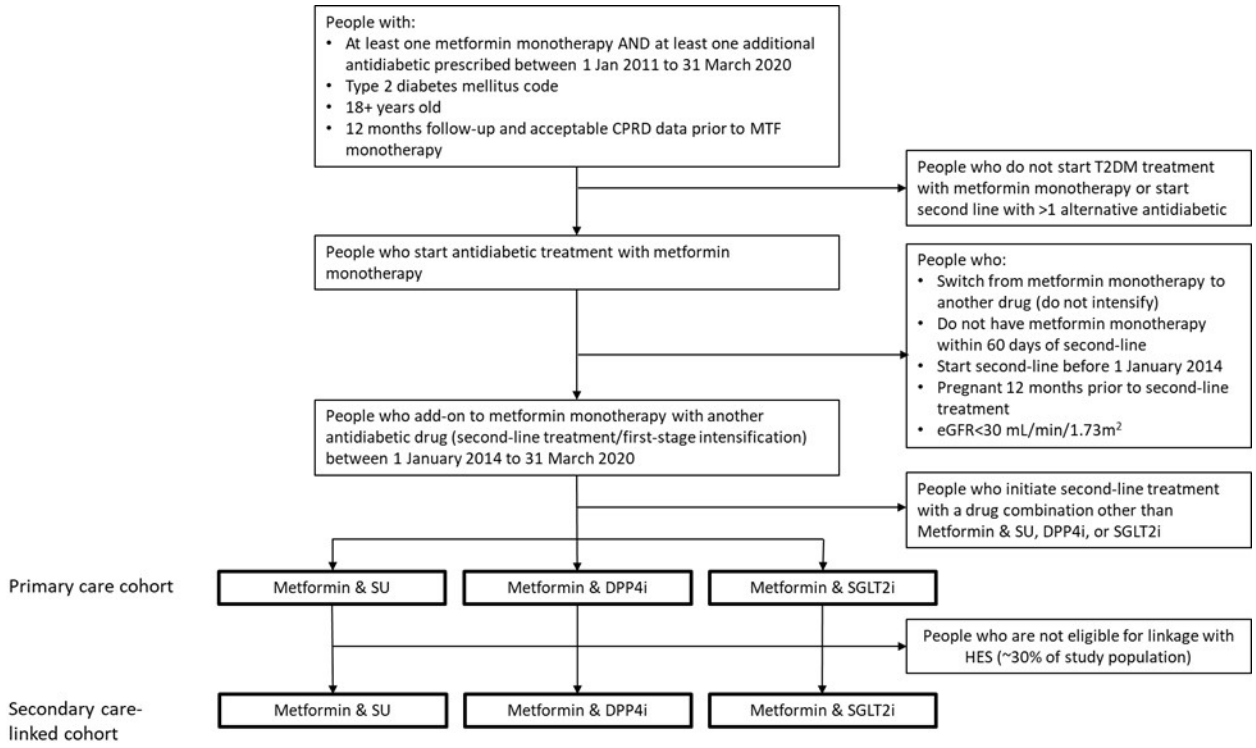


Figure 1 Flow diagram illustrating the identification of the study cohort of people with type 2 diabetes mellitus (T2DM) who initiate second-line antidiabetic treatment with metformin and one of sulfonylurea (SU), dipeptidyl peptidase four inhibitor (DPP4i) or sodium-glucose cotransporter two inhibitors (SGLT2i). CPRD, Clinical Practice Research Datalink; eGFR, estimated glomerular filtration rate; HES, Hospital Episode Statistics; MTF; metformin.

We will identify people within CPRD with at least one prescription for metformin monotherapy and one other antidiabetic medication in primary care between 1 January 2011 and 31 March 2020, registered at a general practice (GP) contributing research-quality data at the prescription date, and registered with their GP for at least 1 year prior to the first metformin or other antidiabetic prescription, to ensure that we are studying new users. The study population will be limited to people with a T2DM primary care code, on or before the antidiabetic index date to exclude those prescribed antidiabetic medications for other indications (eg, polycystic ovarian syndrome or pre-diabetes). Using the individual's entire prescribing history with their registered GP, we will include only people who initiate antidiabetic treatment with metformin, and intensify metformin-monotherapy with a first-time prescription for one of the three second-line antidiabetic drug treatments of interest after 1 January 2014. We chose this date for the evaluation of these three treatments as prior to this only a small minority of people in the UK were prescribed SGLT2i.[4] People who intensify with two or more drug classes on the same date, who discontinue metformin monotherapy prior to a prescription for SU, DPP4i or SGLT2i, or who are prescribed a different drug class as second-line treatment (eg, thiazolidinedione (TZD), insulin, GLP1RA) will be excluded. In addition, we will exclude people with estimated glomerular filtration rate (eGFR) below 30 mL/min/1.73m^2 since SGLT2i are contraindicated for this group.[23] We will also exclude women who were pregnant in the 12 months prior to second-line treatment initiation since prescribing guidelines recommend different treatments for pregnant and breastfeeding women.[24]

Exposures

Exposure groups will include people prescribed SU, DPP4i or SGLT2i as an add-on to metformin. We will not consider other less commonly prescribed T2DM intensification treatments namely TZD, insulin and GLP1-RA since these treatments combined account for less than 10% of second-line therapy regimens in the UK.[4]

The first prescription date for second-line treatment will be considered baseline. To reduce misclassification of people who switch treatments rather than add-on to metformin, we require an additional prescription for metformin on the same date or within 60 days after the first prescription for the second-line drug prescription (SU, DPP4i or SGLT2i). This follows precedent research which used the same definition for second-line antidiabetic treatment in the same database.[4 25] The study will take an intention-to-treat approach, where each person will contribute to the original exposure group to which they were assigned, irrespective of which treatments they may be prescribed subsequently. People will remain exposed until the date the data are censored by death, the patient leaving the GP practice, the GP practice stops contributing to CPRD, or 31 July 2020. We will use the prescription duration recorded in primary care, plus a

60-day grace-period to account for stock-piling medicines, to mark the end of a prescription. Where these data are missing, we will impute the length of prescription with the mean duration of prescription at the practice level, plus the 60-day grace-period, to mark the end of a prescription. Data on subsequent anti-diabetic treatments (third line, fourth line, etc) will be described in those who discontinue second-line treatment.

The study requires information on each person's adherence to antidiabetic treatments. First, to provide a 'baseline' measure of adherence, we require a measure of each person's adherence to metformin monotherapy in the year prior to second-line treatment. We will then assess whether baseline adherence modifies the relative effectiveness of the alternative second-line treatments. Second, we will calculate treatment adherence during the follow-up to help interpret the estimates of the relative effectiveness of the alternative second-line treatments. For both measures of adherence, we will calculate defined daily dose (DDD) from the number of tablets and dosage instructions prescribed versus the duration of the period in question.

Covariates

We will use primary care demographic data, diagnosis codes (Read or SNOMED for CPRD Gold and Aurum, respectively), and laboratory test results recorded prior to second-line treatment initiation, to define our main list of potential confounders. These include age, sex, IMD, time on first-line antidiabetic treatment (as a proxy for diabetes duration), GP size, relevant coprescriptions prescribed within 60 days of second-line treatment initiation (renin-angiotensin system inhibitors, statins), history of proteinuria and comorbidities at baseline (myocardial infarction (MI), unstable angina, stroke, ischaemic heart disease, hypoglycaemia, congestive heart failure (CHF), chronic kidney disease (CKD), end-stage renal disease (ESRD), cancer (any), advanced eye disease and lower extremity amputation). CKD status will be defined using serum creatinine test results to derive eGFR, using cut-points defined by the Kidney Disease Improving Global Outcomes guidelines for CKD, but without requiring two measures 3 months apart.[26] We will also identify HbA1c, systolic blood pressure (SBP), diastolic blood pressure (DBP), eGFR, body weight and body mass index (BMI)[27] using values recorded in the 180 days period before the second-line antidiabetic treatment initiation date in the primary care record. Time between baseline clinical measures and second-line treatment initiation will also be included as covariates. We will follow previous observational research,[28] in undertaking secondary analyses that include additional potential confounders that are defined in primary care records, but for which we anticipate relatively high levels of missing data, namely: ethnicity, high-density lipoprotein (HDL), low-density lipoprotein (LDL), total cholesterol, triglycerides, smoking and alcohol status. In the HES-linked cohort, we will use International Classification of Diseases 10th Revision (ICD-10) diagnosis codes recorded as part of previous hospitalisations in conjunction with primary care data to define comorbidities. We will also use ethnicity recorded in HES for people whose ethnicity is missing within the primary care data. Codelists for all covariates defined in primary and secondary care will be published alongside study results.

Outcomes

The primary outcome for objective 2 (short-term relative effectiveness) will be absolute change in HbA1c% at 12 months follow-up. This change in HbA1c% will be quantified by contrasting follow-up versus baseline laboratory test data recorded in CPRD for each exposure. Secondary outcomes that will also be reported at 12 months after baseline include HDL, LDL, total cholesterol, triglycerides, SBP, DBP, eGFR, body weight and BMI. In defining the 12-month follow-up measurement, the available measure that is closest in time to the 12 months from baseline will be used, recognising that within the pilot data, a median interval in HbA1c measurement of 5 months was observed. Patients without the relevant measurement between 9 and 15 months will be designated as having 'missing 12-month data' (see missing data section). We will also report change in HbA1c at 6–18, 24–30 and 36 months follow-up, again using the closest HbA1c measure in the 3 months before and after the follow-up time point of interest.

Outcomes for long-term relative effectiveness (objective 3) will include macrovascular and microvascular conditions such as CV outcomes (MI, CHF, unstable angina, stroke), renal outcomes (nephropathy, ESRD, 40% decline in eGFR from baseline[29]) and lower limb amputation. Additional outcomes will include hypoglycaemia, time-to-cessation of second-line treatment or treatment switching, adherence calculated according to DDD, all-cause mortality and number of hospital admissions (any reason).

The assessment of long-term outcomes will use the maximum available follow-up. The investigation of microvascular and macrovascular complications including any hospitalisations will require HES data linked to CPRD, and so the patients in the CPRD cohort who cannot be linked to HES (an expected 30%–40%) will be excluded from this aspect of the evaluation.[20 21] Hospital admissions, including microvascular and macrovascular complications, will be identified in HES using ICD-10 diagnosis codes. Clinical diagnoses in primary care coded using Read (CPRD Gold) or SNOMED (CPRD Aurum) codes will be used in addition to secondary care data to identify outcome events. eGFR will be calculated using serum creatinine recorded in primary care as an input in the CKD-EPI formula.[30] We will define nephropathy as new-onset albuminuria or eGFR <60 mL/min/1.73m^2 in people with eGFR ≥60 mL/min/1.73m^2 and no raised albumin to creatinine ratio within 2 years of second-line treatment initiation. A 40% decline in eGFR will be defined as an eGFR measure ≤40% of baseline eGFR.

ESRD will be identified by primary care coding for ESRD and/or renal replacement therapy (RRT) by the GP.

Analytical approach

Objective 1: Describing UK treatment patterns for second-line T2DM treatment, and summarising the results of relevant published RCTs to contextualise the study findings

We will describe trends in prescribing for T2DM second-line treatment for the duration of the study period across the UK and between CCGs. This analysis will update previous research which described the same second-line treatment use in the UK from 2000 to 2017, and will employ similar methods.[4] These descriptive statistics will inform the assessment of the validity of the assumptions that underlie the overall study design. Baseline characteristics listed in the covariates section will also be described for this cohort, overall and stratified by exposure group. We will also conduct a literature review to summarise published RCTs which describe the relative effectiveness of alternative second-line antidiabetic treatments of interest to this study. This will help contextualise the results of this observational study (cf. objectives 2 and 3). We will consider reasons for any possible differences between this observational study compared with published RCTs, including residual confounding and differences in the study populations.

Objectives 2 and 3: Instrumental variable (IV) design to estimate relative treatment effectiveness overall and by subgroup

Studies which apply traditional risk adjustment approaches with little information on case severity may provide biased estimates of treatment effectiveness. We will therefore use an IV design[31 32] to estimate treatment effectiveness in the presence of residual confounding. The IV for second-line drug treatment in this study will be each CCG's prescribing history, recognising that the choice of second-line treatment may involve the hospital diabetologist, the GP, other healthcare professionals, and the individual. We will define 'CCG prescribing history' as the proportion of people prescribed each second-line treatment in the CCG for the last complete calendar year prior to the treatment intensification currently under consideration. This IV encourages receipt of the treatment but does not have a direct effect on outcomes except through the treatment prescribed (figure 2). Using CCG prescribing history as the IV follows pharmacoepidemiological research[32] that uses provider preference as an instrument for treatment prescribed.

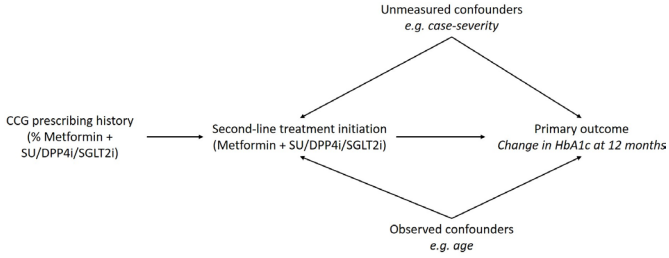


Figure 2 Instrumental variable design to be applied in this study comparing three options for second-line antidiabetic treatment. CCG, Clinical Commissioning Group; DPP4i, dipeptidyl peptidase-4 inhibitors; HbA1c, haemoglobin A1c; SGLT2i, sodium-glucose cotransporter-2 inhibitor; SU, sulfonylureas.

In our pilot CPRD data,[4] the proportions of people prescribed each second-line treatment regimen varied widely. For example, in 2014 the ranges across CCGs were 5%–100% (SU), 0%–90% (DPP4i) and 0%–35% (SGLT2i).[4] These proportions have changed over time but similar people received different second-line treatment regimens simply according to CCG prescribing preference or time period.[4]

This study's design will exploit this wide variation in the choice of second-line treatment. We will use this IV to estimate the relative effectiveness of alternative second-line treatments while minimising bias from unobserved confounding. We will use a 'local IV estimator'[33] to allow for heterogeneity according to unobserved characteristics (eg, lifestyle choices) as well as observed characteristics (eg, baseline HbA1c) when reporting the relative effectiveness of the alternative second-line treatments according to individual risk factor profiles.

IV assumptions

The validity of our IV design relies on three key assumptions: the IV must (1) strongly predict the treatment prescribed; (2) be independent of baseline unmeasured covariates; and (3) only affect the outcome through the treatment prescribed.[31] The IV design will lead to bias if the prescribing history of the CCG has a direct effect on the outcome. We carefully assessed whether the CCG's prescribing history met the criteria for an IV. Our pilot data showed it was strongly associated with the second-line treatment regimen prescribed (assumption 1). We also found that prescribing history balanced the observed covariates (assumption 2, figure 3). We are unable to assess empirically whether clinicians' prescribing history is independent of unmeasured confounders; however, it is likely that participants will attend their local GP without considering their prescribing history, and unlikely that the CCGs prescribing history would have a direct effect on outcomes (assumption 3). For example, it is unlikely that simply because a CCG shows a preference for prescribing SU the participants' outcomes would be better (or worse) regardless of the treatment actually prescribed. We will reassess each assumption using the full study dataset and undertake sensitivity analyses to test these assumptions.

Power considerations

Power calculations were conducted prior to accessing study data. Clinically meaningful between-treatment difference in HbA1c from baseline is considered to be 0.3 percentage points (eg, from 8.0% to 7.7%) by the European Medicines Agency[34] and 0.5 percentage points by NICE.[35] We based our power calculations on these numbers and assuming an SD of 2.4.[36] We follow methodological recommendations for power calculations with

These demonstrate that clinician choice of different first-stage intensification therapies show remarkably little relationship with clinical risk factors, strongly suggesting the IV approach will produce balanced and valid comparison groups.

(A) Preference for DPP4i prescription (prescriptions of DPP4i excluding current participant / total prescriptions) by year and CCG.

(B) Preference for SGLT2i prescription (prescriptions of SGLT2i excluding current participant / total prescriptions) by year and CCG.

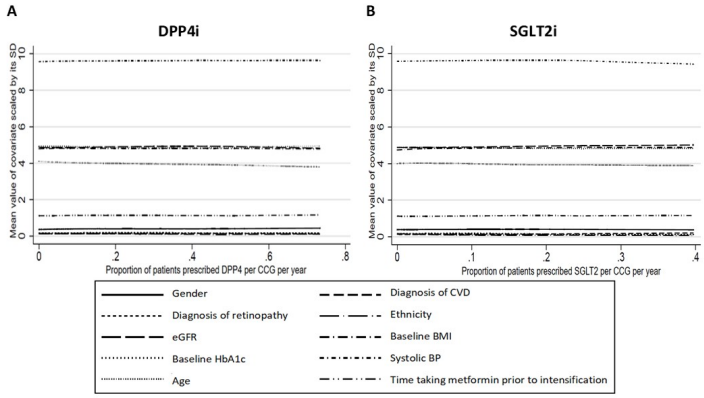


Figure 3 Covariate balance across levels of CCG prescribing history (this study's IV) (2014–2017). BMI, body mass index; BP, blood pressure; CCG, Clinical Commissioning Group; CVD, cardiovascular disease; DPP4i, dipeptidyl peptidase-4 inhibitors; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; IV, instrumental variable; SGLT2i, sodium-glucose cotransporter-2 inhibitor; SU, sulfonylureas.

IV designs and consider that the proportion of people who actually receive the treatment predicted by the IV is 80%, but also consider scenarios where the IV is weaker (70% compliance) and stronger (90% compliance).[31] We require 80% power at the 5% (two-sided) level of statistical significance, with a Bonferroni correction to account for multiple comparisons to get a familywise error rate of 5%. Table 1 shows the requisite sample sizes of the two treatment groups projected to have the fewest participants (SU and SGLT2i). The study will include approximately 25 700 participants (SU=6000, DPP4i=13 000, SGLT2i=6700) based on an initial feasibility count, which will be more than sufficient for detecting whether clinically significant differences in the primary endpoint are statistically significant.

Planned analyses

We will examine the relevant trends in prescribing between 2014 and 2020 by CCG, and by year, and summarise baseline covariates using data collected prior to the index date for second-line treatment.

We will provide personalised estimates of treatment effectiveness using the local IV (LIV) approach[33 37] to predict the counterfactual outcomes that each person would experience if they were prescribed each second-line treatment. We will use probit regression models[38] to estimate the propensity to receive each treatment according to observed characteristics, and CCG preference for each second-line regimen (the IV). We will estimate the relationship of each outcome with observed characteristics, and the propensity for each second-line treatment using generalised linear models (GLMs)[39] for continuous and count outcomes. For time-to-event outcomes (eg, time to second-line treatment cessation, time to each microvascular or macrovascular complication), we will recognise that the period of observation may differ across individuals due to censoring. We will describe each endpoint by plotting Kaplan-Meier curves, and estimate each treatment effect using discrete-time hazard models.[40] SEs will be calculated with non-parametric bootstrapping, and will account for clustering of individuals within practices.

These models will be used to estimate the relative effect of prescription of SGLT2i vs SU, DPP4i versus SU and SGLT2i vs DPP4i for the primary and secondary outcomes. Person-level treatment effects will be calculated as the difference

Table 1 Required sample size (N) for the IV design according to instrument strength (level of compliance) and magnitude of effect size at 80% power and 5% (two-sided) level of statistical significance

	Level of compliance (IV strength)					
	70%		80%		90%	
Effect size: between-treatment difference in mean HbA1c reduction baseline to 12 months	SU	SGLT2i	SU	SGLT2i	SU	SGLT2i
0.3	4556	1952	3488	1495	2756	1181
0.4	2563	1098	1962	841	1550	664
0.5	1640	703	1256	538	992	425

DPP4i, dipeptidyl peptidase-4 inhibitors; IV, instrumental variable; SGLT2i, sodium-glucose cotransporter-2 inhibitor; SU, sulfonylureas.

in predicted outcomes following prescription of the alternative drugs. These person-level treatment effects will be aggregated to report the relative effectiveness of the treatments prescribed overall, and by prespecified subgroups. These prespecified subgroups will include: people with and without CV comorbidities overall and by subtype of CVD, people with baseline eGFR ≥60 mL/min/1.73 m^2 vs baseline eGFR <60 mL/min/1.73 m^2, age groups, sex, ethnicity, BMI (based on WHO categorical definition[41]), adherence to metformin and baseline HbA1c levels. We will consider finer eGFR subgroupings, age categories and HbA1c levels based on descriptive statistics (objective 1) prior to any relative effectiveness analyses (objectives 2–3). Any additional subgroups will be informed by descriptive statistics of each covariate, and the advice of a panel of healthcare professionals, building on those identified in a literature review. The clinical panellists will include diabetologists, GPs and practice nurses involved in care for people with T2DM.

Missing data

In our primary analysis, we will use a complete-case approach based on the main potential confounders listed in the covariates section. We will conduct secondary analyses using complete cases for the full list of potential confounders, including those expected to have a high proportion of missing data (see covariates section), which we do not expect to be missing at random. Because we cannot assume covariate measurements are missing at random and the IV model is computationally intensive, we will not use multiple imputation.

We will adopt two main approaches based on the type of missingness for outcome data: (1) linear interpolation using values recorded during follow-up, and (2) inverse probability weighting (IPW) to those people lost to follow-up with no subsequent outcome measure. We will use linear interpolation for values that are intermittently missing during follow-up, for example, if HBA1c at 12 months is required, but the available measures are at 3-month and 17-month follow-up, which fall outside the requisite time window (9–15 months). This method was used in precedent diabetes research with observational data.[42] For those settings, were the patient is lost to follow-up and there is therefore no subsequent HbA1c measure available, we will use IPW, reweighting the data for those with available observations to represent the group lost to follow-up, assuming therefore that the HbA1c data are missing at random.[42 43]

Sensitivity analyses

We will conduct sensitivity analyses falling under three broad categories: (1) Modified study population inclusion and exclusion criteria to evaluate the validity of our IV assumptions in subgroups where there is arguably less equipoise in the choice of second-line treatment[44]; (2) Comparing the larger primary care unlinked cohort with the primary care population linked to secondary care data (HES) and (3) Evaluating the robustness of our statistical methods.

Under the first category of sensitivity analyses, we will exclude people with contraindications for SU (eg, liver disease) who are prescribed SGLT2i, as this prescribing may not be due to CCG preference. We will also expand the eGFR exclusion to all those with eGFR <60 mL/min/1.73m^2 (vs eGFR <30 mL/min/1.73m^2 in the main analysis). In addition, we will include people who are censored or die during the first 60 days after a prescription for SU, DPP4i or SGLT2i without a prescription for metformin in the same time period to consider the impact of potentially misclassifying these people as switching from metformin monotherapy instead of adding on to metformin monotherapy. Under the second category of sensitivity analyses, we will repeat the analyses for objectives 1 and 2 limited to the HES-linked subpopulation from CPRD who are eligible for the long-term outcomes (objective 3). Third, we will assess the robustness of findings to alternative statistical models for the LIV approach, outcome regressions and alternative approaches to handling missing data.[45]

Patient and public involvement

Two PP representatives were consulted when designing this study prior to obtaining funding. One has close family experience of type 1 diabetes as a carer and the other was recently diagnosed with T2DM. Both PP representatives have discussed the study with local patients and obtained future workshop interest. Our PP representatives reinforced that the study design, outcomes and interpretation should recognise the importance of personalising treatment choice according to the individual's experience, and according to their age, weight, ethnicity and more general lifestyle choices. The PP representatives have supported plans for two study workshops that will inform the translation of results to patients and the public. The PP representatives have emphasised the importance of developing accessible preworkshop information to help participants prepare. The PP representatives will help inform the way the study presents and communicates results so they are accessible to patients and the general public.

Strength and limitations

This study will exploit the natural variation in prescribing patterns for second-line antidiabetic treatment across CCGs within similar groups of people by using an IV study design. This design minimises potential biases resulting from unmeasured confounders, a major limitation in observational research. The large and representative sample from UK primary care will improve the generalisability of this study's results and allow for stratification on prespecified baseline risk factors, helping patients and their providers choose treatments based on personal risk profiles to maximise clinical benefit. While the IV relies on three major assumptions which may limit the validity of our estimates, we will evaluate the strength of our assumptions in sensitivity analyses.

A potential limitation is that the required natural variation in prescribing may not exist for those people who are prescribed SGLT2i as second-line treatment because they do not tolerate or have a contraindication for SU, as per current[6]

NICE guidelines.[3] We will investigate this potential source of bias by undertaking a sensitivity analysis excluding those in the SGLT2i exposure group with contraindications for SU. In addition, our study may be susceptible to non-differential outcome misclassification, as we are unable to link our data to additional audit datasets with more detailed outcome information (eg, laboratory tests) such as the Myocardial Ischaemia National Audit Project (MINAP).[46] However, a previous study shows that the majority of MINAP acute MI events in the general England and Wales populations are also recorded in CPRD and HES.[47]

Future work

The results of this study will be used in future research which aims to predict long term outcomes and associated costs to the NHS beyond this study's maximum follow-up. To do this, we will adapt a diabetes microsimulation model developed using observational data from the United States Veterans' Affairs database[48] to the UK setting. We will use a personalised approach to second-line treatment by using the estimates of relative effectiveness within the subgroup analyses in this study. We will publish the analysis plan for this work separately.

Ethics and dissemination

Ethics

This study will be based in part on data from the CPRD obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data are provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this protocol are those of the authors alone. The study was approved by the Independent Scientific Advisory Committee (approval number 20-064) and the London School of Hygiene & Tropical Medicine Research Ethics Committee (reference number 21395). GPs have opted-in to contributing data to the CPRD, while individuals registered at these GPs may opt-out. Individual-level consent was not necessary since these data are deidentified.

Dissemination/outputs

This study's outputs will be designed in collaboration with our expert advisory panel and PP representatives to help ensure this study can inform future clinical guidelines and care for people with T2DM. Results will be published open-access in peer-reviewed journals and presented at scientific conferences. Additional emphasis will be placed on the implementation of advanced quantitative methods in this study, which will provide general guidance for future studies on how the overall approach of combining these methods with routinely available electronic health data can provide insights to inform person-level care. We will provide recommendations via the Academic Health Sciences Networks to commissioners and T2DM care providers on how to target second-line antidiabetic treatment to individuals and patient groups.

Data visualisations of key results and lay summaries will also be published on this website as a resource to be shared with key stakeholders and as accessible information for the general public.

Author affiliations
[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[2]Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK
[3]The Comparative Health Outcomes, Policy & Economics (CHOICE) Institute, University of Washington School of Pharmacy, Seattle, Washington, USA
[4]Personalized Healthcare Data Science, Roche Products Limited, Welwyn Garden City, UK
[5]Centre for Longitudinal Studies, University College London, London, UK
[6]Patient Research Champion Team, National Institute for Health Research, Twickenham, UK
[7]Diabetes Trials Unit, The Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK
[8]Diabetes Research Centre, University of Leicester, Leicester, UK

Acknowledgements KK is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration East Midlands (ARC EM) and the NIHR Leicester Biomedical Research Centre (BRC).

Contributors PB and RG drafted the manuscript. SO'N, AB, SW, RJS, PC, AB, AIA, KK, LAT, LS and IJD assisted in the revision of the draft paper. Each of the authors has approved the final version of the manuscript for submission.

Funding This work was supported by National Institutes of Health Research (NIHR) grant number NIHR128490.

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NIHR, NHS or the Department of Health and Social Care.

Competing interests PB, SO'N, AB, RJS, PC, LAT and LS have nothing to declare. SW is employed by Roche and holds stock in Roche. AB is an economic advisor on DiRECT trial with ongoing responsibility for economic analysis during long-term follow-up phase, and has also acted as consultant to GlaxoSmithKline, Merck, Novo Nordisk and Boehringer Ingelheim in relation to their diabetes products. AIA receives salary from the National Institute for Health Research (NIHR) via the University of Oxford and Addenbrooke's Hospital, and also chairs an NICE technology appraisal committee, is a member of Diabetes UK, and manages people whose salaries are partially funded by completed trials involving sitagliptin and exenatide and an ongoing trial of empagliflozin. KK has acted as a consultant, speaker or received grants for investigator-initiated studies for Novartis, Novo Nordisk, Sanofi-Aventis, Lilly and Merck Sharp & Dohme, Boehringer Ingelheim, Bayer, Berlin-Chemie AG/Menarini Group, Janssen, and Napp. IJD holds an unrestricted research grant from GSK and holds shares in GSK. RG sits on the NIHR commissioning committee.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Patrick Bidulka http://orcid.org/0000-0001-7644-2030
Richard J Silverwood http://orcid.org/0000-0002-2744-1194

REFERENCES

1 NICE. NG28: type 2 diabetes in adults: management, 2015. Available: https://www.nice.org.uk/guidance/ng28

2 Hex N, Bartlett C, Wright D, *et al*. Estimating the current and future costs of type 1 and type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabet Med* 2012;29:855–62.
3 NICE. *TA390: canagliflozing, dapagliflozin and empagliflozin as monotherapies for treating type 2 diabetes*. London: NICE, 2016.
4 Wilkinson S, Douglas I, Stirnadel-Farrant H, *et al*. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000-2017. *BMJ Open* 2018;8:e022768.
5 Buse JB, Wexler DJ, Tsapas A, *et al*. 2019 update to: management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American diabetes association (ADA) and the European association for the study of diabetes (EASD). *Diabetes Care* 2020;43:487–93.
6 Varvaki Rados D, Catani Pinto L, Reck Remonti L, *et al*. The association between sulfonylurea use and all-cause and cardiovascular mortality: a meta-analysis with trial sequential analysis of randomized clinical trials. *PLoS Med* 2016;13:e1001992.
7 Khunti K, Chatterjee S, Gerstein HC, *et al*. Do sulphonylureas still have a place in clinical practice? *Lancet Diabetes Endocrinol* 2018;6:821–32.
8 Rosenstock J, Kahn SE, Johansen OE, *et al*. Effect of linagliptin vs glimepiride on major adverse cardiovascular outcomes in patients with type 2 diabetes: the Carolina randomized clinical trial. *JAMA* 2019;322:1155–66.
9 Neal B, Perkovic V, Mahaffey KW, *et al*. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med Overseas Ed* 2017;377:644–57.
10 Zinman B, Wanner C, Lachin JM, *et al*. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med Overseas Ed* 2015;373:2117–28.
11 Wiviott SD, Raz I, Bonaca MP, *et al*. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med Overseas Ed* 2019;380:347–57.
12 Zelniker TA, Wiviott SD, Raz I, *et al*. Sglt2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. *Lancet* 2019;393:31–9.
13 McGovern A, Feher M, Munro N, *et al*. Sodium-Glucose co-transporter 2 (SGLT2) inhibitor: comparing trial data and real-world use. *Diabetes Ther* 2017;8:365–76.
14 Rothwell PM. Treating individuals 2. subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 2005;365:176–86.
15 Mannucci E, Naletto L, Vaccaro G, *et al*. Efficacy and safety of glucose-lowering agents in patients with type 2 diabetes: a network meta-analysis of randomized, active comparator-controlled trials. *Nutr Metab Cardiovasc Dis* 2021;31:1027–34.
16 Khunti K, Godec TR, Medina J, *et al*. Patterns of glycaemic control in patients with type 2 diabetes mellitus initiating second-line therapy after metformin monotherapy: Retrospective data for 10 256 individuals from the United Kingdom and Germany. *Diabetes Obes Metab* 2018;20:389–99.
17 Wilding J, Fernando K, Milne N, *et al*. Sglt2 inhibitors in type 2 diabetes management: key evidence and implications for clinical practice. *Diabetes Ther* 2018;9:1757–73.
18 O'Brien MJ, Karam SL, Wallia A, *et al*. Association of second-line antidiabetic medications with cardiovascular events among insured adults with type 2 diabetes. *JAMA Netw Open* 2018;1:e186125.
19 Powell WR, Christiansen CL, Miller DR. Meta-analysis of sulfonylurea therapy on long-term risk of mortality and cardiovascular events compared to other oral glucose-lowering treatments. *Diabetes Ther* 2018;9:1431–40.
20 Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: clinical practice research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
21 Wolf A, Dedman D, Campbell J. Data resource profile: clinical practice research Datalink (CPRD) aurum. *Int J Epidemiol* 2019;48:1740.
22 Herbert A, Wijlaars L, Zylbersztejn A. Data resource profile: Hospital episode statistics admitted patient care (Hes APC). *Int J Epidemiol* 2017;46:1093.
23 NICE. Canagliflozin in combination therapy for treating type 2 diabetes: technology appraisal guidance TA315 [Web], 2014. Available: https://www.nice.org.uk/guidance/ta315/chapter/2-The-technology
24 National Institute for Health and Care Excellence. Diabetes, pregnancy and breast-feeding [Web], 2017. Available: https://bnf.nice.org.uk/treatment-summary/diabetes-pregnancy-and-breast-feeding.html
25 Wilkinson S, Douglas IJ, Williamson E, *et al*. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol* 2018;10:1639–48.
26 Stevens PE, Levin A, Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med* 2013;158:825–30.
27 Bhaskaran K, Forbes HJ, Douglas I, *et al*. Representativeness and optimal use of body mass index (BMI) in the UK clinical practice research Datalink (CPRD). *BMJ Open* 2013;3:e003389.
28 Williamson EJ, Walker AJ, Bhaskaran K, *et al*. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430–6.
29 Levey AS, Inker LA, Matsushita K, *et al*. Gfr decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National kidney Foundation and the US food and drug administration. *Am J Kidney Dis* 2014;64:821–35.
30 Levey AS, Stevens LA, Schmid CH, *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–12.
31 Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med* 2014;33:2297–340.
32 Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostat* 2007;3:Article 14.
33 Basu A. Estimating PERSON-CENTERED treatment (PET) effects using instrumental variables: an application to evaluating prostate cancer treatments. *J Appl Econ* 2014;29:671–91.
34 European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment or prevention of diabetes mellitus [Web], 2018. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-clinical-investigation-medicinal-products-treatment-prevention-diabetes-mellitus_en.pdf [Accessed 29 Jan 2018].
35 National Insititute for Health and Care Excellence. Type 2 diabetes mellitus: medicines optimisation priorities, 2019. Available: https://www.nice.org.uk/advice/ktt12/chapter/Evidence-context
36 Farmer AJ, Rodgers LR, Lonergan M, *et al*. Adherence to oral glucose-lowering therapies and associations with 1-year HbA1c: a retrospective cohort analysis in a large primary care database. *Diabetes Care* 2016;39:258–63.
37 Basu A, Gore JL. Are elderly patients with clinically localized prostate cancer Overtreated? exploring heterogeneity in survival effects. *Med Care* 2015;53:79–86.
38 Bliss CI. The method of PROBITS. *Science* 1934;79:38.
39 Nelder JA, Wedderburn RWM. Generalized linear models. *J R Stat Soc Ser A* 1972;135:370–84.
40 Allison PD. Discrete-Time methods for the analysis of event histories. *Sociol Methodol* 1982;13:61–98.
41 World Health Organisation. Body mass index - BMI, 2020. Available: https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi
42 Basu A, Sohn M-W, Bartle B, *et al*. Development and validation of the real-world progression in diabetes (RAPIDS) model. *Medical Decision Making* 2019;39:137–51.
43 Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res* 2013;22:278–95.
44 Small DS, Rosenbaum PR. War and wages. *J Am Stat Assoc* 2008;103:924–33.
45 Welch CA, Petersen I, Bartlett JW, *et al*. Evaluation of two-fold fully conditional specification multiple imputation for longitudinal electronic health record data. *Stat Med* 2014;33:3725–37.
46 Herrett E, Smeeth L, Walker L, *et al*. The myocardial ischaemia national audit project (MINAP). *Heart* 2010;96:1264–7.
47 Herrett E, Shah AD, Boggon R, *et al*. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ* 2013;346:f2350.
48 Basu A, Sohn M-W, Bartle B, *et al*. Development and validation of the real-world progression in diabetes (rapids) model. *Med Decis Making* 2019;39:137–51.

4.2. Amendments to the published protocol

Amendments to this protocol were pre-specified prior to conducting the main IV analyses to estimate the comparative effectiveness of the antidiabetic treatments of interest (Chapter 5). I outline these amendments and their justifications in **Table 4.1.**, which is copied from the published SAP.[1] Major amendments to the protocol included (i) changing the type of IV analysis from LIV to 2SRI due to methodological issues in using the LIV for a three-way treatment comparison, (ii) using MI to impute missing baseline and continuous outcome data (e.g., HbA1c during follow-up), and (iii) extending the follow-up period. The SAP can be accessed here: https://www.lshtm.ac.uk/media/72276.

Table 4.1. Amendments to the published protocol for the PERMIT study comparative effectiveness analysis of alternative second-line oral antidiabetic treatments.

The original text from the protocol is written in column 2, and the amended text is written in column 3. The pieces of text that are bolded are the specific aspects of the original protocol which were amended.

Amendment number	**Original plan (as per Bidulka et al. 2021[2])**	**Amended plan**
1	Study population: The study population will include people registered with **a CPRD-contributing practice,** aged 18 years or older, diagnosed with T2DM who intensify antidiabetic treatment from metformin-monotherapy to a combination of metformin and SU, DPP4i or SGLT2i (second-line treatment) between **2014 and 2020.**	Study population: The study population will include people registered with **a CPRD-contributing practice with HES/ONS/IMD linkage available,** aged 18 years or older, diagnosed with T2DM who intensify antidiabetic treatment from metformin-monotherapy to a combination of metformin and SU, DPP4i or SGLT2i (second-line treatment) between **1 January 2015 and 31 December 2020.** Justification: The time to event outcomes require HES/ONS/IMD linkage. Further, additional data from HES (e.g., ethnicity data, comorbidity data), ONS (gold-standard death date in England and Wales), and the IMD (small-area deprivation) are important to reduce data missingness and misclassification. When designing the study, we anticipated that only 70% of the CPRD population in England would be eligible for HES-linkage based on data resource profiles for CPRD.[3, 4] However, in our updated dataset the proportion eligible for linkage is 91%. Thus, we feel that the balance between maximising study power for all outcomes versus the benefits in analysing and

		interpreting results for one clearly defined cohort is in favour of latter. In addition, we have changed the study time period in which people can enter the study. We require one year of historical prescribing data from each person's index date to define the instrument in our instrumental variable analysis – thus we must exclude people who initiate second-line antidiabetic treatment in 2014, since prior to this year SGLT2i were not widely prescribed. We also extended the dataset to allow for the inclusion of people who intensify second-line antidiabetic treatment up until 31 December 2020.
2	End of follow-up: People will remain exposed until the date the data are censored by death, the patient leaving the GP practice, the GP practice stops contributing to CPRD, **or 31 July 2020.**	End of follow-up: People will remain exposed until the date the data are censored by death, the patient leaving the GP practice, the GP practice stops contributing to CPRD, **or 31 December 2021 (for outcomes defined in primary care only) and 31 March 2021 (for outcomes defined in secondary care (HES) or ONS data).** Justification: The dataset was extended to allow for follow-up until 31 December 2021 for outcomes defined in primary care data. For outcomes defined in secondary care data, we must end follow-up at 31 March 2021 as this was the maximum amount of follow-up available at the time of study data extraction.
3	Covariates: We will also identify HbA1c, systolic blood pressure (SBP), diastolic blood pressure (DBP), eGFR, body weight, and body-	Covariates: We will also identify HbA1c, systolic blood pressure (SBP), diastolic blood pressure (DBP), eGFR, body weight, and body-

	mass index (BMI) using values recorded **in the 180 days period before the second-line antidiabetic treatment initiation date** in the primary care record.	mass index (BMI) using values recorded in the 180 days (HbA1c) and 540 days (SBP, DBP, eGFR, BMI) period before the second-line antidiabetic treatment initiation date in the primary care record. Justification: We follow precedent research that used a 540 day window pre-baseline to define these clinical measures at baseline.[5] However, we still require HbA1c to have been measured within 180 days as we expect older values of HbA1c to be unrepresentative of the patient's HbA1c status at baseline.
4	Primary outcome: The primary outcome for objective 2 (short-term relative effectiveness) will be absolute change in HbA1c**%** at 12 months follow-up.	Primary outcome: The primary outcome for objective 2 (short-term relative effectiveness) will be absolute change in HbA1c (**mmol/mol**) at 12 months follow-up. Justification: UK is aligning with Europe and reporting HbA1c using International Federation of Clinical Chemistry (IFCC) units (mmol/mol) rather than Diabetes Control and Complications Trial (DCCT) unit (%).
5	Secondary outcomes: We will also report change in HbA1c at 6–18, 24–30 and 36months follow-up, again using the closest HbA1c measure in the 3 months before and after the follow-up time point of interest.	Secondary outcomes: We will also report change in HbA1c**, eGFR, SBP, and BMI** at **6-, 24-, 36-, 48-, and 60-months follow-up,** again using the closest outcome measure in the 3 months before and after the follow-up time point of interest.

		Justification: Our dataset was updated to include a maximum follow-up time of 7 years. We therefore increased the duration of follow-up time points of interest, and simplified to be at yearly intervals (other than the first 6-month time-point). While we planned to investigate eGFR, BMI, and SBP at 12 months follow-up, we also plan to investigate these outcomes at every other follow-up time point of interest as per HbA1c.
6	Secondary outcomes: Outcomes for long-term relative effectiveness (objective 3) will include macrovascular and microvascular conditions **such as CV outcomes** (MI, CHF, unstable angina, stroke)**, renal outcomes (**nephropathy, ESRD, 40% decline in eGFR from baseline**)** and lower limb amputation.	Secondary outcomes: Outcomes for long-term relative effectiveness (objective 3) will include macrovascular and microvascular conditions **such as CV outcomes (3-point major adverse cardiovascular event (MACE), a composite outcome of myocardial infarction, stroke, and all-cause mortality),** MI, CHF, unstable angina, stroke**), renal outcomes (a composite kidney outcome (40% decline in eGFR from baseline, end-stage kidney/renal disease (ESKD), and all-cause mortality), as well as** nephropathy, ESRD, 40% decline in eGFR from baseline**)** and lower limb amputation. Justification: We seek to emulate trials which compare these second-line antidiabetic drugs, which often report MACE and a composite kidney outcome.[6, 7] The components of these outcomes were already specified in the protocol, and pre-specified adding the composite end-points before conducting analyses.
7	Secondary outcomes: A 40% decline in eGFR will be defined as an eGFR measure **≤40% of baseline eGFR.**	Secondary outcomes: A 40% decline in eGFR will be defined as an eGFR measure **≤60% of baseline eGFR.**

		Justification: This was a mistake in the original protocol – the 40% decline should represent an eGFR measure that is ≤60% of baseline eGFR.
8	Analytical approach (Objective 1): We will describe trends in prescribing for T2DM second-line treatment for the duration of the study period across the UK and between CCGs. This analysis will update previous research which described the same second-line treatment use in the UK from 2000 to 2017, and will employ similar methods.	Analytical approach (Objective 1): We will describe trends in prescribing for T2DM second-line treatment for the duration of the study period across the UK and between CCGs, **particularly with respect to clinically important factors predicting which type of second-line antidiabetic treatment people are prescribed, such as ethnicity and deprivation.** This analysis will update previous research which described the same second-line treatment use in the UK from 2000 to 2017, and will employ similar methods. Justification: Understanding factors which predict prescribing for particular second-line antidiabetic treatments is helpful to design the instrumental variable analysis. Previous work by Wilkinson et al (2018)[8] described factors associated with choice of second-line antidiabetic treatment. We build off this work in understanding whether there are sociodemographic disparities in which type of second-line antidiabetic treatment is prescribed.
9	Planned analyses – the IV:	Planned analyses – the IV:

	We will provide **personalised estimates of treatment effectiveness using the local IV (LIV) approach** to predict the counterfactual outcomes that each person would experience if they were prescribed each second-line treatment.	Due to challenges in developing the methodology to compare three rather than two treatments using the local IV approach, we will instead use the two-stage residual inclusions (2SRI) model to conduct this analysis. This approach also enables treatment effectiveness to be reported for the overall populations and subpopulations of prime interest.
10	Missing data: **In our primary analysis, we will use a complete-case approach based on the main potential confounders listed in the covariates section. We will conduct secondary analyses using complete cases for the full list of potential confounders, including those expected to have a high proportion of missing data (see covariates section), which we do not expect to be missing at random. Because we cannot assume covariate measurements are missing at random and the IV model is computationally intensive, we will not use multiple imputation.** We will adopt two main approaches based on the type of missingness for outcome data: (1) linear interpolation using values recorded during follow-up, and (2) inverse probability weighting (IPW) to those people lost to follow-up with no subsequent outcome measure.	Missing data: We now propose using multiple imputation (MI) to handle missing values in covariates and intermittent missingness in the continuous outcomes as it will impute unobserved values with plausible substitutes based on the distribution of the observed data. We will handle loss to follow-up with inverse probability weighting. Justification. The initial data descriptions highlighted the non-linear trajectory of the continuous outcomes, it is best to utilise all information when imputing the outcome values. MI has the advantage of accounting for uncertainty in the imputed value while also incorporating observed relationships between the variable being imputed and other variables in the dataset. For loss to follow-up, previous work has shown the problems of using MI for imputing for timepoints beyond the observed data, and that IPW that only relies on baseline values to reweight the observations is more appropriate for handling this problem.

4.3. Relevance to my thesis

This protocol paper and amendments detailed in the SAP clearly outline the comparative effectiveness analysis using an IV for Case Study 1 of this thesis (Chapter 5). Publishing this protocol paper was an important pre-requisite to clearly specifying the justification and methodology of the cohort study aiming to estimate the causal effects of alternative antidiabetic treatments, and follows the NICE real-world evidence framework which emphasises the need to "generate evidence in a transparent way and with integrity from study planning through to study conduct and reporting".[9] A previously published paper by my colleagues, Wilkinson et al (2018) described variation at the CCG-level in second-line antidiabetic treatment prescribing.[10] This evidence was used to justify the IV analysis using the TTP at the CCG-level as the instrument. Thus, this protocol addressed steps 3 and 4 (**Figure 2.1.**) of my approach to implementing a comparative effectiveness analysis using routinely collected health data.

This protocol paper does not clearly articulate that this cohort study followed the 'target trial emulation' framework.[11] The term 'target trial emulation' was coined by Hernán and Robins in 2016 shortly before the start of my PhD studies. While I did not incorporate this framework explicitly in this protocol paper at the time of writing in 2020, the protocol clearly articulates key aspects of the formal framework including clear definitions of the study population, the treatment assignment and contrasts of interest, time 0 (i.e., the start of follow-up), and the statistical analyses used to estimate causal treatment effects while minimising the risk of bias. In the subsequently published SAP[1] and comparative effectiveness results paper (Chapter 5), I explicitly apply the target trial framework[11] by comparing the ideal RCT design to the observational design of the comparative effectiveness work for Case Study 1 of this thesis.

This study pre-specification, a requirement for most RCTs, is an important step to designing and conducting rigorous observational studies useful to policymakers and healthcare providers to improve services and outcomes for patients.[9] Tools and frameworks such as the Strengthening the Reporting of Observational studies in Epidemiology (STROBE),[12] reporting of studies conducted using observational routinely collected health data for pharmacoepidemiology (RECORD-PE),[13] and the target trial emulation framework,[11, 14] formalise and bolster the strength of observational study designs, with the latter aimed at minimising biases when estimating causal treatment effects. Using these tools and frameworks instil confidence in the integrity and validity of results from observational studies which use complex study designs and analyses with the goal of generating useful evidence to complement that from RCTs to inform HTA and clinical practice.

4.4. References

1. Bidulka P, Lugo-Palacios DG, Carroll O, O'Neill S, Grieve R. Statistical Analysis Plan (SAP): PERsonalised Medicine for Intensification of Treatment (PERMIT) study, Version 1.0, July 2023. Web. Accessed 28 November, 2023. https://www.lshtm.ac.uk/media/72276

2. Bidulka P, O'Neill S, Basu A, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open*. 2021;11(9):e046912. doi:10.1136/bmjopen-2020-046912

3. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. Jun 2015;44(3):827-36. doi:10.1093/ije/dyv098

4. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology*. 2019;48(6):1740-1740g. doi:10.1093/ije/dyz034

5. Wilkinson S, Williamson E, Pokrajac A, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. *Diabetes, Obesity and Metabolism*. 2020;22(5):847-856. doi:https://doi.org/10.1111/dom.13970

6. Glycemia Reduction in Type 2 Diabetes — Microvascular and Cardiovascular Outcomes. *New England Journal of Medicine*. 2022/09/22 2022;387(12):1075-1088. doi:10.1056/NEJMoa2200436

7. Empagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2022;388(2):117-127. doi:10.1056/NEJMoa2204233

8. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

9. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

10. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

11. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

12. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS medicine*. Oct 16 2007;4(10):e297. doi:10.1371/journal.pmed.0040297

13. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ (Clinical research ed)*. 2018;363:k3532. doi:10.1136/bmj.k3532

14. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

CHAPTER 5. RESEARCH PAPER – COMPARATIVE EFFECTIVENESS OF ALTERNATIVE SECOND-LINE ORAL ANTIDIABETIC TREATMENTS ON METABOLIC, KIDNEY, AND CARDIOVASCULAR OUTCOMES AMONGST PEOPLE WITH TYPE 2 DIABETES MELLITUS: A COHORT STUDY USING ROUTINELY COLLECTED HEALTH DATA.

OVERVIEW

In this chapter, I include the accepted manuscript for a research paper *in press* at the *BMJ* (accepted March 2024) for which I am joint first author. I presented this work at the Health Data Research (HDR)-UK 2024 annual conference, as well as at invited talks at Imperial College London and internally at LSHTM.

In this research paper, I present the results from a cohort study applying the target trial framework and an IV analysis to compare alternative second-line oral antidiabetic treatments among a general population of people with T2DM in English primary care. The design of this cohort study was presented in Chapter 4. This study uses an updated cohort of people included in the analyses in Chapter 3 of this thesis, applying the same inclusion and exclusion criteria and treatment definitions but with an extended study period (2014-2021).

This study demonstrates how variation in second-line antidiabetic treatment prescribing across CCGs in England can be used as a preference-based instrument to estimate comparative treatment effects for important T2DM-related outcomes in a general population of people with T2DM in primary care. Following the research paper, I include a brief discussion of this paper's relevance to my thesis. This discussion includes key tables from the supplementary materials of the published paper which are particularly important in the interpretation of the study in the context of this thesis.

5.1. Accepted research paper

The main tables and figures are provided immediately following the references for this accepted manuscript. I added the prefix '5'.X for each main figure and table to indicate this is the 5th chapter of this thesis.

Select supplementary materials important to this thesis, including the supplementary methods, are provided in the Appendix F, referenced in section 5.2.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ
Please list the paper's authors in the intended authorship order:	Patrick Bidulka*, David G Lugo-Palacios*, Orlagh Carroll, Stephen O'Neill, Amanda I Adler, Anirban Basu, Richard J Silverwood, Jonathan W Bartlett, Dorothea Nitsch, Paul Charlton, Andrew H Briggs, Liam Smeeth, Ian J Douglas,

	Kamlesh Khunti, Richard Grieve
Stage of publication	**In press**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I am joint first author and the corresponding author of this paper. I led the design of the study (lead author on the protocol paper and the statistical analysis plan) along with the wider PERMIT study team. I led the data management, including the extraction and creation of the study dataset from the raw CPRD files. I co-led the main analyses with DLP, SON, and OC, including the descriptive analyses, traditional multivariable regression analyses, and IV analyses. OC led the multiple imputation. DLP and SON led the implementation of the LASSO approach for covariate selection and the propensity score-IPTW-RA alternative analyses. I supported DLP who led the drawing of the DAGs. I led the literature review and the interpretaion of all study results. I wrote the first draft of the manuscript with support from my supervisor, Richard Grieve. SON and OC led the writing of the supplementary methods. Co-authors reviewed multiple drafts before sending to the journal for publication. I led the peer review responses with input from my co-authors. I share lead authorship with DLP.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Richard Grieve
Date	07 March 2024

Comparative effectiveness of alternative second-line oral antidiabetic treatments on metabolic, kidney, and cardiovascular outcomes amongst people with type 2 diabetes mellitus: cohort study using routinely collected health data

Order	Affiliation	Author name	Job title	Email
1*	1	Patrick Bidulka	Research Fellow	Patrick.bidulka1@lshtm.ac.uk
2*	2	David G Lugo-Palacios	Assistant Professor	David.Lugo-Palacios@lshtm.ac.uk
3	2	Orlagh Carroll	Research Fellow	Orlagh.Carroll@lshtm.ac.uk
4	2	Stephen O'Neill	Associate Professor	Stephen.ONeill@lshtm.ac.uk
5	3	Amanda I Adler	Professor	amanda.adler@dtu.ox.ac.uk
6	4	Anirban Basu	Professor	basua@uw.edu
7	5	Richard J Silverwood	Associate Professor	r.silverwood@ucl.ac.uk
8	6	Jonathan W. Bartlett	Professor	Jonathan.Bartlett1@lshtm.ac.uk
9	1	Dorothea Nitsch	Professor	Dorothea.nitsch@lshtm.ac.uk
10	7	Paul Charlton	Patient Research Champion	charlton808@btinternet.com
11	2	Andrew H Briggs	Professor	Andrew.Briggs@lshtm.ac.uk
12	1	Liam Smeeth	Professor	Liam.Smeeth@lshtm.ac.uk
13	1	Ian J Douglas	Professor	Ian.DOUGLAS@lshtm.ac.uk
14	8	Kamlesh Khunti	Professor	kk22@leicester.ac.uk
15	2	Richard Grieve	Professor	Richard.Grieve@lshtm.ac.uk

*Contributed equally

Affiliations

(1) Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London UK WC1E 7HT
(2) Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, Keppel Street, London UK WC1E 7HT
(3) Diabetes Trials Unit, The Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, OCDEM Building Churchill Hospital, Old Road, Headington UK OX3 7LJ
(4) The Comparative Health Outcomes, Policy & Economics (CHOICE) Institute, University of Washington School of Pharmacy, Seattle USA 98195
(5) Centre for Longitudinal Studies, UCL Social Research Institute, University College London, 20 Bedford Way, London UK WC1H 0AL
(6) Department of Medical Statistics, London School of Hygiene & Tropical Medicine, Keppel Street, London UK WC1E 7HT
(7) Patient Research Champion Team, National Institute for Health Research, 15 Church Street, Twickenham UK TW1 3NL
(8) Diabetes Research Centre, University of Leicester, Gwendolen Road, Leicester UK LE5 4PW

Word count: 6,423 (main text), 376 (abstract). Tables: 1 Figures: 5 References: 89

Corresponding author:

Patrick Bidulka

patrick.bidulka1@lshtm.ac.uk

ORCID number: 0000-0001-7644-2030

Twitter handle: @pbidulka

ABSTRACT

Objective

To compare the effectiveness of sulfonylureas (SU), dipeptidyl peptidase-4 inhibitors (DPP4i), and sodium-glucose co-transporter 2 inhibitors (SGLT2i) added to metformin for people with type 2 diabetes mellitus (T2DM) who require second-line treatment in routine clinical practice.

Design

Cohort study emulating a comparative effectiveness trial (target trial) with an instrumental variable analysis to reduce the risk of both measured and unmeasured confounding.

Setting

Linked primary care, hospital, and death data in England between 2015-2021.

Participants

75,739 adults with T2DM who start second-line oral antidiabetic treatment.

Treatments of interest

SU or DPP4i or SGLT2i, added to metformin.

Main outcome measures

Primary outcome: change in haemoglobin A1c (HbA1c) between baseline and 1-year follow-up. Secondary outcomes: changes in body mass index (BMI), systolic blood pressure (SBP), and estimated glomerular filtration rate (eGFR) at 1- and 2-years (also HbA1c), and time to ≥40% decline in eGFR, major adverse kidney event (MAKE), heart failure hospitalisation, major adverse cardiovascular event (MACE), and all-cause mortality.

Results

75,739 people initiated second-line oral antidiabetic treatment with SU (34%), DPP4i (46%), or SGLT2i (20%). SGLT2i were more effective than either DPP4i or SU in reducing mean HbA1c levels between baseline and 1-year. After the instrumental variable analysis, the mean (95% confidence interval (CI)) differences in HbA1c change between baseline and 1-year were: -2.5 mmol/mol (-3.7 to -1.3) for SGLT2i versus SU, and -3.2 mmol/mol (-4.6 to -1.8) for SGLT2i versus DPP4i. SGLT2i was more effective in reducing BMI and SBP compared to either SU or DPP4i.

For some secondary endpoints there was no evidence that SGLT2i were more effective. For example the hazard ratio (HR) for MACE was 0.99 (95% CI 0.60 to 1.63) versus SU, and 0.91 (0.50 to 1.66) versus

DPP4i. SGLT2i had reduced hazards of heart failure hospitalisation compared with DPP4i (HR 0.32, 0.12 to 0.90), and SU (0.46, 0.19 to 1.05). The HR for a ≥40% decline in eGFR indicated a protective effect versus SU (0.42, 0.22 to 0.82), with uncertainty in the estimated HR versus DPP4i (0.64, 0.29 to 1.43).

Conclusions

SGLT2i were more effective than SU or DPP4i in lowering mean HbA1c, BMI and SBP, and reducing the hazards of heart failure hospitalisation (versus DPP4i) and kidney disease progression (versus SU), with no evidence of differences in other clinical endpoints.

SUMMARY BOX

Section 1: What is already known on this topic

- Placebo-controlled randomised trials have demonstrated that SGLT2i are cardio- and kidney-protective among people with T2DM.
- For people with T2DM who have cardiovascular disease or at high-risk of cardiovascular disease, guidance and guidelines recommend that SGLT2i are added to metformin as second-line oral antidiabetic treatment; however, for the broader T2DM population without these indications whose glycaemic control is inadequate following metformin monotherapy, current guidelines recommend that either SU, DPP4i or SGLT2i are added to metformin.
- The comparative effectiveness of these three alternative second-line oral antidiabetic treatments have not been assessed directly in randomised controlled trials, and evidence from observational studies is prone to confounding by indication.

Section 2: What this study adds

- For a broad population of people with T2DM, SGLT2i were more effective than SU or DPP4i in lowering mean HbA1c, BMI and SBP, and reducing the hazards of heart failure hospitalisation (versus DPP4i) and kidney disease progression (versus SU).
- A target trial design was combined with an instrumental variable analysis to help the study reduce the risk of bias from confounding, and to supplement previous studies in providing useful evidence that applies directly to routine clinical practice.

ABBREVIATIONS

(Listed alphabetically)

BMI – body-mass index

CCG – clinical commissioning group

CI – confidence intervals

CKD – chronic kidney disease

CPRD – Clinical Practice Research Datalink

CVD – cardiovascular disease

DAG – directed acyclic graph

DBP – diastolic blood pressure

DPP4i – dipeptidyl peptidase-4 inhibitors

eGFR – estimated glomerular filtration rate

ESKD – end-stage kidney disease

GLP1-RA glucagon-like peptide-1 receptor agonists

GP – general practice

GRADE – Glycemia Reduction Approaches In Diabetes: A Comparative Effectiveness Study

HbA1c – haemoglobin A1c

HES – Hospital Episode Statistics

HR – hazard ratio

HTA – health technology assessment

IHD – ischaemic heart disease

IMD – index of multiple deprivation

IPTW – inverse probability of treatment weighting

IQR – interquartile range

IV – instrumental variable

LASSO – least absolute shrinkage and selection operator

MACE – 3-point major adverse cardiovascular event

MAKE – major adverse kidney event

MI – myocardial infarction

MICE – multiple imputation with chained equations

NHS – National Health Service

NICE – National Institute of Health and Care Excellence

OLS – ordinary least squares

ONS – Office of National Statistics

PERMIT – PERsonalised Medicine for Intensification of Treatment

PPI – patient and public involvement

RASi – renin-angiotensin system inhibitors

RCT – randomised controlled trials

SAP – statistical analysis plan

SBP – systolic blood pressure

SGLT2i – sodium-glucose co-transporter 2 inhibitors

STEMI – ST-elevated myocardial infarction

SU – sulfonylureas

TTP – tendency to prescribe

T2DM – type 2 diabetes mellitus

UK – United Kingdom

US – United States

2SRI – two-stage residual inclusion

2SLS – two-stage least squares

INTRODUCTION

About 463 million people worldwide (9.3%) have type 2 diabetes mellitus (T2DM).(1) In most cases, T2DM is a progressive disease, with risks of multiple complications including cardiovascular disease (CVD) and chronic kidney disease (CKD).(2) Interventions that improve T2DM biomarkers such as glycated haemoglobin A1c (HbA1c), blood pressure, and lipids can reduce the risk of these complications.(3-6) International clinical guidance and guidelines recommend additional drugs (second-line therapy) if glycaemic control is inadequate following metformin monotherapy.(7-9) A recent study that considered second-line treatments for people with T2DM across 38 countries reported the most prevalent second-line treatments were three oral drug treatments: dipeptidyl peptidase-4 inhibitors (DPP4i) (48.3%), sulphonylureas (SU) (40.9%), and sodium-glucose co-transporter-2 inhibitors (SGLT2i) (8.3%).(10)

Of these oral treatments, DPP4i and SGLT2i are newer and more costly classes of drugs.(11) In England, SGLT2i are recommended second-line treatments in preference to other drug classes for some people with T2DM, that is those with pre-existing CVD, those at high risk of CVD, or those with kidney disease.(7) However, for most people with T2DM, evidence on the comparative effectiveness of the alternative drugs classes, in particular according to HbA1c reduction, is insufficient to recommend a particular second-line treatment.(7) An international consensus statement(9) and guidelines from the National Institute of Health and Care Excellence (NICE)(7) therefore leaves the choice of second-line treatment for most people with T2DM to clinician and patient choice; this has led to wide variation across groups of primary care providers in England in the proportion of people prescribed each drug class.(12) Current NICE (2022) guidelines stipulate that other antidiabetic treatments, such as insulin-based therapy and glucagon-like peptide-1 receptor agonists (GLP1-RA), are only recommended if HbA1c is not controlled *after* second-line oral antidiabetic treatment.(7) Hence, in many countries, including England, the proportion of people with T2DM who are prescribed GLP1-RA as second-line antidiabetic treatment is low.(10, 12, 13)

Most randomised controlled trials (RCT) that have assessed the effectiveness and safety of SGLT2i or DPP4i have randomised groups to an active intervention versus a placebo comparator.(14-26) Hence, while these placebo-controlled RCTs have reported fewer CVD and kidney events following allocation to SGLT2i for people with and without T2DM, these results are difficult to apply to routine clinical practice, where the relevant populations and comparators are different.(16-24) Of the RCTs that included an active comparator, some have compared DPP4i to SU(27-30) or SGLT2i to SU,(31) but none of these trials have compared all three of these drug classes. Thus, the comparative effectiveness of SGLT2i versus alternative second-line oral antidiabetic treatments on outcomes important to patients,

particularly HbA1c reduction, remains unclear. Results from previous observational studies which compared these treatments(32-34) are at risk of bias due to residual (unmeasured) confounding. While a recent observational study(35) emulated some of the results of the GRADE (Glycemia Reduction Approaches In Diabetes: A Comparative Effectiveness Study) randomised trial,(29, 36, 37) neither the trial nor the observational study considered SGLT2i, which limits the applicability of the results to routine clinical practice.

Recent advances in real-world data combined with developments in quantitative methods offer important opportunities for generating evidence on comparative treatment effectiveness with direct relevance to clinical practice.(35) In this study, we illustrate the potential and challenges of using real-world data from the Clinical Practice Research Datalink (CPRD). We emulate the design of a hypothetical pragmatic RCT by contrasting all three antidiabetic drug classes (SU, DPP4i, and SGLT2i) of interest for the broad population of people with T2DM who, according to current NICE guidelines, are eligible for any of these second-line treatments. We consider intermediate metabolic outcomes, in particular HbA1c, but also kidney- and cardiovascular-related complications. We help reduce the risk of unmeasured confounding by using prescriber variation as an instrumental variable (IV) in our analysis in estimating treatment effectiveness from routine data.(38, 39) Our study complements a recent target trial emulation that assessed the comparative effectiveness of alternative second-line treatments using data from the Department of United States (US) Veterans Affairs,(40) but which underrepresents females (less than 10%), and in the main analyses assumed that there was no unmeasured confounding.

The aim of this paper is to assess the comparative effectiveness of the three most common second-line antidiabetic treatments prescribed in the United Kingdom (UK) according to metabolic and other clinical measures (changes from baseline in HbA1c, estimated glomerular filtration rate (eGFR), body-mass index (BMI), and systolic blood pressure (SBP)), and adverse clinical endpoints (kidney and cardiovascular outcomes, and death).

METHODS

Study design

This study, the PERsonalised Medicine for Intensification of Treatment (PERMIT) study, was designed according to the target trial framework;(41) pre-specified in both the published versions of the study protocol(42) and the statistical analysis plan (SAP).(43) Details of how each of the standpoints of the target trial were emulated are reported in **Supplementary table 1**. In brief, a target trial is a hypothetical RCT for assessing comparative effectiveness from observational data which requires the definition of the main elements of a trial's protocol, including eligibility criteria, the respective treatment strategies, definition of time 'zero', and an analysis plan.(41)

We applied target trial principles to primary care data from the CPRD to identify people with T2DM who had similar prognosis prior to initiating any of the three second-line antidiabetic drug treatments under comparison. CPRD covers approximately 20% of the UK population registered at general practices (GPs), and includes longitudinal information on primary care diagnoses, prescriptions, demographic information, and laboratory test results.(44, 45) Linkage from CPRD to Hospital Episode Statistics (HES) in-patient data was available for approximately 90% of participating practices in England. We accessed information from the HES admitted patient care database on diagnoses, procedures, socio-demographic characteristics, admission and discharge dates.(46) We used CPRD-HES linked data to ascertain cardiovascular and kidney outcomes, as this has been shown to improve capture of these events, and reduce risks of misclassification, rather than relying on a single data source.(47, 48) Information on each person's vital status was available from linkage to the Office for National Statistics (ONS) death record.(49, 50)

Study population

We defined the study population according to eligibility criteria, which had to be met prior to 'time zero' ('baseline') and is analogous to the time of randomisation in an RCT. Time zero was defined by the date of the first prescription for any of the three oral second-line treatments that were added to metformin (**Supplementary table 1**). We followed precedent research by including people diagnosed with T2DM aged 18 years or older(33, 51) registered with a GP in England who intensified treatment from first- to second-line oral antidiabetic treatments between 1 January 2015 to 31 December 2020 with a first-ever prescription of SU, DPP4i or SGLT2i, added to metformin. Those eligible had at least one prescription for metformin monotherapy within 60 days *prior* to the first prescription for second-line treatment, to ensure their use of metformin monotherapy was continuous prior to intensification.

We excluded women with a record of pregnancy within 12 months prior to second-line treatment initiation and people whose last recorded eGFR was less than 30mL/min/1.73m^2, since prescribing guidelines recommend different treatments for these groups. We also excluded people whose GPs had not consented to the required linkage of HES data. We followed precedent research in excluding those who were not prescribed metformin on the same day, or within 60 days *after* initiating second-line treatment,(33) as it is unlikely that their treatment with metformin continued. Detailed inclusion/exclusion criteria are presented in **Supplementary tables 1-2**.

Treatments under comparison

We compared DPP4i versus SU, SGLT2i versus SU, and SGLT2i versus DPP4i as second-line oral antidiabetic treatments added to metformin. We extracted information on the prescribed duration of each second-line treatment, of metformin, and of any subsequent antidiabetic therapy.

The study took an 'intention-to-treat' approach so that each person contributed to the treatment group to which they were assigned at baseline until the end of the follow-up period (**Supplementary table 1**), irrespective of the extent to which they adhered to the second-line antidiabetic treatment prescribed. We defined the end of follow-up as the earliest of: the date the GP stopped contributing to CPRD, the date the person left the GP, the date of death, or the last date of available data (31 December 2021 for continuous outcomes or 31 March 2021 for time-to-event outcomes (see also Statistical Analysis section). We described the duration on second-line and third-line treatments by comparison group.

Covariates

We have previously described the covariates in detail, (11, 42) and these are summarised in **Supplementary table 3**. Briefly, we defined patient sociodemographic characteristics (age, sex, ethnicity, index of multiple deprivation (IMD)), time since T2DM diagnosis, year of second-line antidiabetic treatment initiation, National Health Service (NHS) region (East of England, London, Midlands, North East and Yorkshire, North West, South East, and South West),(52) number of patients registered with the participants' GP, smoking and alcohol status, relevant co-prescriptions (renin-angiotensin system inhibitors (RASi) or statins) issued within 60 days prior to baseline, hospitalisation (any) in the previous year, and comorbidities recorded at baseline (previous myocardial infarction (MI), unstable angina, stroke, ischaemic heart disease (IHD), hypoglycaemia, heart failure, history of cancer (any), history of proteinuria, advanced eye disease, lower extremity amputation, CKD). We also defined

HbA1c, SBP, diastolic blood pressure (DBP), eGFR, and BMI(53) using the most recent measures recorded in primary care.

For the primary endpoint, HbA1c, we only considered the most recent measure within 180 days prior to baseline in line with NICE guidance which recommends that HbA1C is measured every 6 months.(7) For SBP, DBP, and eGFR, we followed previous research in considering the most recent measure within 540 days prior to baseline(33) (**Supplementary table 3**). Any values recorded in advance of these time windows were considered out-dated and were not used to define baseline characteristics. For BMI we followed a previously published algorithm in using the most recent measure available, which for the vast majority of cases was within 6-months (see results).(53)

Outcomes

The primary outcome was the absolute change in HbA1c (mmol/mol) between baseline and 1-year following each second-line treatment prescription (HbA1c at 1-year – HbA1c at baseline). Treatment groups were compared according to the mean change in HbA1c. We used the measurement closest in time to the 1-year follow-up timepoint and allowed for measures within ±90 days, otherwise the measure was designated as missing (see Statistical analysis section).

Secondary outcomes included: change in HbA1c at 2-years, and change in BMI, SBP, and eGFR all at 1- and 2-years.(33) We also reported the time to the following first events before 2-years follow-up: (a) a 40% decline in eGFR from baseline, which could be a marker for the rarer end-stage kidney disease (ESKD) outcome,(54) (b) a major adverse kidney event (MAKE), a composite outcome for the earliest of a decline in eGFR from baseline of 40%, end-stage kidney disease (ESKD), and all-cause mortality,(55) (c) heart failure hospitalisation, (d) 3-point major adverse cardiovascular event (MACE), a composite outcome for the earliest of myocardial infarction, stroke, and CVD death, and (e) all-cause mortality. We also reported time to myocardial infarction and stroke individually. We could not report time to ESKD and CVD-specific mortality due to the low number of events. Individuals were followed until they experienced the event of interest, died or CPRD-HES data were no longer available (31 March 2021). For these time-to-event measures, we considered outcomes over 2-years in the base case, as it was anticipated that at later timepoints a high proportion would have censored or missing data (see also Alternative analysis). Details on all outcome definitions, including data sources, are described in **Supplementary table 4**.

Statistical analysis

An IV analysis was chosen to help reduce the risk of confounding due to unobserved baseline measures, such as information on diet and exercise prior to initiation of second-line treatment (for details see **Supplementary Methods, Supplementary table 1, Supplementary figures 1A-B**).(38) The IV was the primary care providers' tendency to prescribe (TTP) the three classes of second-line treatment. In England most primary care clinicians work within a group, and over the study's timeframe this was defined as a clinical commissioning group (CCG), who informed health funding decisions for their respective geographic region. Some CCGs recommended that SGLT2i were too costly which means that GPs within that CCGs region had a strong disincentive to prescribe these drugs. So, defining CCGs rather than individual GPs as the unit for the IV reflected decision-making, and was strongly associated with the choice of second-line treatment.(11, 12)

We also found wide variation across CCGs in the proportion of people prescribed each of these three classes of second-line treatment (**Figure 5.1.**)**.** This 'natural variation' implied that people of similar prognosis at baseline received a different second-line treatment simply according to their CCG. We defined TTP as the proportion of eligible people prescribed each second-line treatment within the 12 months preceding the specific baseline (time zero) for each person. A valid instrument must meet four main conditions (see also the Direct Acyclic Graph (DAGs) in **Supplementary figures 1A-B**).(38) First, the instrument must predict the treatment prescribed, which can be formally assessed.(56) Here, we assessed the relevance of the CCGs TTP with a weak instrument test that is robust to heteroscedasticity and clustering by NHS region (see Results). Recent work has suggested that to meet the requirement that the instrument is of sufficient strength, the F-statistic summarising the relationship between the IV and the treatment received must exceed 100.(38, 57) Second, the instrument must be independent of covariates that predict the outcomes of interest, which can be partially evaluated. We assessed the extent to which observed prognostic covariates differed across levels of the instrument (see **Supplementary figures 2A-C**). Third, the instrument must have an effect on the outcomes only through the treatment received, which cannot be evaluated empirically (see Discussion). Large imbalances in measured covariates across levels of the TTPs would raise concerns about the second and third IV assumptions. We followed our pre-specified protocol(42), the SAP,(43) and were guided by the DAGs (**Supplementary figures 1A-B**), in choosing to adjust for measured contextual and temporal confounders in the second-stage (outcome) regression. By including these contextual covariates in the second-stage regression, we were able to make weaker assumptions, that the TTP was independent of the outcome, and only had an effect on the outcome via the treatment received, *after* adjusting for any differences in region, GP practice size, and time period (see Results and Supplement). Fourth, the IV assumes monotonicity, which implies that there are 'no defiers' so that as the levels of the IV change

this should have the same direction of effect on the treatment prescribed across similar individuals. However, this assumption cannot be verified.(58) Indeed, in our study, we cannot observe the same treatment choice for a particular individual according to their attendance at two CCGs with difference levels of prescribing preference for SGLT2i (vs DPP4i or SU). For the population this assumption applies that the average treatment choice must increases or decrease monotonically with the level of the IV.(59) Hence, it is plausible to assume that if a group of patients who attended a CCG with a moderate preference for prescribing SGLT2i were prescribed this drug class, then a similar group of patients who attended a CCG with a stronger preference for prescribing SGLT2i would not be prescribed DPP4i or an SU.(59)

The IV approach taken was the two-stage residual inclusion (2SRI) method,(60) which enabled us to assess comparative effectiveness across the full study populations of interest, that is, to report average treatment effects, while reducing the risk of bias from unmeasured confounding. The first stage models estimated the probabilities that each person was prescribed each treatment given their baseline covariates and their CCGs TTP for that treatment.(61) The second-stage outcome models then included generalised residuals from the first-stage (propensity score) models. The outcome models were estimated by ordinary least squares (OLS) for continuous outcomes (e.g. 1-year HbA1c), and Cox proportional hazards models for time-to-event outcomes with an individual frailty.[32] Models for both stages included all measured baseline covariates, with polynomials and covariate interactions selected via a post-double selection approach using Least Absolute Shrinkage and Selection Operator (LASSO) regression(62-64) (**Supplementary methods table S1**). The purpose of including person-level covariates in the second-stage (outcome regression) was to gain precision in estimating the relative treatment effects.

Some data were missing for outcomes (metabolic and other clinical measures) and baseline covariates (ethnicity, IMD, HbA1c, SBP, DBP, BMI, eGFR, smoking and alcohol status) because the patient did not have these measures recorded by the GP at all, or within the requisite time window for a specific timepoint. The percentages of missing values for HbA1c were 33.7% (1 year) and 36.4% (2 years), BMI 44.7% (1 year) and 47.8% (2 years), SBP 33.6% (1 year) and 37.2% (2 years) and eGFR 37.4% (1 year) and 40.0% (2 years). For some people a measurement that was not available at a particular timepoint (e.g., 2 years) was available at other timepoints (e.g., 1 year and 3 years) (**Supplementary methods table S2).** It was also possible that at any particular timepoint one measure (e.g. BMI) was not available whereas (e.g. HbA1c SBP, and eGFR) were.

We chose to handle all missing baseline and longitudinal outcome data by Multiple Imputation(65) with Chained Equations (MICE).(66) This approach assumed data were 'missing at random'. The

imputation of each longitudinal outcome at a given timepoint used all relevant information including measurements of the same outcome at other time points. This use of auxiliary information can help the study recover more accurate estimates of the unknown outcome values.(67) This also ensured our study population was comparable at each time point. Partially-observed covariates and outcomes(67, 68) were multiply imputed by predictive mean matching with 10 donors,(69) producing five imputed datasets. The number of imputations was driven by the need to balance computational time with improved inference from increasing the number of imputations (**Supplementary methods, page 12** for further details). The imputation models developed for each covariate were congenial with the form of outcome(70) (continuous or time-to-event). For the time-to-event endpoints, it was assumed there were no missing data. All imputation models were stratified by second-line treatment (DPP4i, SGLT2i, SU) and by whether the individual died, or was censored prior to the relevant study end-date (see **Supplementary methods**).

We reported differences between the comparison groups according to absolute change in outcomes between baseline and follow-up for continuous measures, and according to time-to-event measures. We reported results overall and according to whether or not patients had CVD (at least one of previous MI, previous stroke, heart failure, IHD, or unstable angina) recorded prior to initiation of second-line treatment. To recognise statistical uncertainty in the estimates of treatment effects, the data were bootstrapped 500 times, stratified by CCG, treatment group, and death and censoring status, to maintain the structure of the original sample across replicates. Within each bootstrap resample MICE was implemented(71, 72) with Rubin's first rule(65) applied across the five imputed datasets to obtain overall treatment effects for each bootstrap sample, which were then used to estimate variances and calculate t-based bootstrap confidence intervals (CI). The imputation procedure and time-to-event analyses were performed with MICE and the survival package(73, 74) in R 4.2.2 respectively(75), and the analysis of the clinical measures in Stata 17.(76)

Alternative analyses

We undertook alternative analyses to check the impact of making different statistical assumptions on our results. Firstly, we applied complete-case analysis rather than MICE (base case) to examine whether the results were robust when alternative approaches were applied to handle the missing data. Secondly, we applied two-stage least squares (2SLS) (continuous outcomes), multivariable linear regression (continuous outcomes) and Cox regression analysis (time-to-event), adjusting for all measured baseline covariates, to assess the sensitivity of our approach to confounding adjustment. Thirdly, we extended the follow-up period to 5-years rather than 2-years. Fourthly, in additional

analyses that were not pre-specified, we further checked the impact of applying approaches that, like multivariable regression, assumed 'no unmeasured confounding', but can be less sensitive to the form of outcome regression model. We applied two approaches based on propensity scores - inverse probability of treatment weighting (IPTW),(77) and IPTW-weighted regression adjustment,(78) with unstablised and stabilised weights.(79) We also used asymmetric trimming to understand the impact, if any, of large weights in the IPTW- regression adjustment analysis.(80, 81) IPTW-weighted regression adjustment has the so-called 'double-robustness property' in that subject to the assumption of no unobserved confounding it can still provide consistent estimates provided *either* the propensity score or regression model is correctly specified.(78, 82) The multivariable regression analyses and the IPTW analyses all estimate the average treatment effects as in the base case. We undertook all the alternative analyses on the complete cases only.

Patient and public involvement (PPI)

PPI advisors, including a co-author on this paper (PC), helped inform the design and proposed analysis, including the choice of outcome measures. We will reconvene a PPI workshop to discuss the study findings and co-produce a lay summary that will be available on the PERMIT study website.(83)

RESULTS

Study population and baseline characteristics

We included 75,739 people with T2DM who initiated second-line antidiabetic treatment with SU, DPP4i, or SGLT2i and met all eligibility criteria (**Figure 5.2.**). Of these, 25,693 (34%) were prescribed SU, 34,464 (46%) DPP4i, and 15,582 (20%) SGLT2i, in addition to metformin. The frequencies of each drug prescribed within each drug class are reported in **Supplementary table 5**. The most common drugs prescribed within each drug class were gliclazide (SU), sitagliptin (DPP4i), and empagliflozin (SGLT2i). The mean age of people prescribed SGLT2i (56 years, SD 11) was lower than those prescribed DPP4i (62, SD 12) or SU (60, SD 13) (**Table 1**). Baseline mean HbA1c was higher for people prescribed SU (81 mmol/mol, SD 22) versus DPP4i (72 mmol/mol, SD 16) and SGLT2i (75 mmol/mol, SD 17), and a lower proportion of people prescribed SGLT2i had comorbidities; for example, 17.2% of those prescribed SGLT2i had pre-existing CVD, compared to 22.8% of those prescribed SU and 23.5% prescribed DPP4i. The proportion of people prescribed SGLT2i increased from 7.3% in 2015 to 24.9%

in 2020. The median time between recorded BMI and the index date was 19 days (interquartile range (IQR) 0-140).

Within two-years follow-up, the median (IQR) time prescribed second-line antidiabetic treatment was lower for the SU group (248 days, IQR 67 to 671) compared with DPP4i (345 days, IQR 96 to 730) and SGLT2i (328 days, IQR 84 to 730). The proportions who switched to a third-line treatment within two years of the index date were: 59% (SU), 52% (DPP4i) and 53% (SGLT2i), with metformin monotherapy the most common third-line treatment for all three comparison groups (**Supplementary table 6**). In each comparison group, the proportions of people whose third-line treatment was triple-therapy were: 25% (SU), 32% (DPP4i) and 22% (SGLT2i).

Empirical assessment of instrumental variable assumptions

The TTP, met a major requirement for being a valid IV, in that it was strongly associated with the second-line treatment prescribed (assumption 1) with accompanying F-statistics of 1,902 for DPP4i and 1,935 for SGLT2i, which indicated that the instrumental variable was of sufficient strength (F>100).(38, 57) The measured potential confounders were balanced across levels of the TTP (assumption 2), aside from time period, which was included within the covariate adjustment of the IV analysis (see **Supplementary figure 2A-C** and Discussion).

Intermediate metabolic and other clinical measures

The crude change in mean HbA1c from baseline to 1-year follow-up among people with observed follow-up measures was greatest for people prescribed SU (-18 mmol/mol) compared with DPP4i (-10 mmol/mol) and SGLT2i (-14 mmol/mol, **Figure 5.3., Supplementary figure 3**). Of those people not censored by 1-year follow-up (N=72,066), 32.1% were missing HbA1c at this time point (**Supplementary methods table 2**). Although levels of missing data were higher for those timepoints that occurred after the onset of the COVID-19 pandemic, the levels of missing data remained similar across the comparison groups (see **Supplementary table 7**).

The crude changes in mean BMI and SBP from baseline were small across all time points (**Figure 5.3., Supplementary figure 3**). The crude change in mean eGFR from baseline to 1-year follow-up was similar across the three second-line treatments of interest (-2 mL/min/1.73m^2), with smaller decreases in mean eGFR across subsequent follow-up periods amongst people prescribed SGLT2i rather than SU or DPP4i (**Figure 5.3., Supplementary figure 3**).

Figure 5.4. presents the results after addressing confounding (instrumental variable analysis) and missing data (MICE) and applies to the full study population. There was strong evidence that SGLT2i were more effective in reducing HbA1c between baseline and 1-year follow-up, with a mean (95% CI) reduction of -2.5 mmol/mol (95% CI -3.7 to -1.3) versus SU, and of -3.2 mmol/mol (95% CI -4.6 to -1.8) versus DPP4i (**Figure 5.4., Supplementary table 8**). After addressing confounding and missing data, SGLT2i were more effective in improving BMI and SBP (**Figure 5.4.**). People prescribed SGLT2i had a greater reduction in BMI between baseline and 1-year with a mean difference of -1.55 kg/m^2 (95% CI -1.72 to -1.37) versus SU, and -0.85 kg/m^2 (95% CI -1.03 to -0.66) versus DPP4i. For SBP the mean (95% CI) difference was -2.07 mm Hg (95% CI -3.10 to -1.04) versus SU, and -1.76 mm Hg (95% CI -2.99 to -0.53) versus DPP4i, with these improvements maintained at 2-years follow-up. SGLT2i slowed the decline in eGFR at 2-years follow-up compared to SU (mean difference of 1.39 mL/min/1.73m^2, 95% CI 0.49 to 2.30), but not versus DPP4i (mean difference of -0.04 mL/min/1.73m2, 95% CI -1.09 to 1.01).

Kidney, cardiovascular, and mortality outcomes

People prescribed SGLT2i had lower crude rates of all adverse kidney, cardiovascular, and mortality events compared to those prescribed SU and DPP4i (**Supplementary table 9, Supplementary figures 4-9**). After accounting for confounding and missing data, we found that over 2-years follow-up (base case), SGLT2i were more effective in preventing a ≥40% decline in eGFR from baseline versus SU (hazard ratio (HR) 0.42, 95% CI 0.22 to 0.82), but the estimated HR versus DPP4i were highly uncertain (HR 0.64, 95% CI 0.29 to 1.43) (**Figure 5.5.**). The rates of heart failure hospitalisation were lower following SGLT2i versus SU (HR 0.46, 95% CI 0.19 to 1.05) and DPP4i (HR 0.32, 95% CI 0.12 to 0.90). For the other endpoints, there was no evidence of a difference in the comparative effectiveness of the second-line antidiabetic treatments (**Figure 5.5., Supplementary table 10)**. There was no evidence that having CVD prior to starting second-line treatment modified the relative effectiveness of these three treatments (**Supplementary tables 11-12**).

Alternative analyses

The findings from the complete case analyses were similar to those from applying multiple imputation to address the problem of missing data (**Supplementary tables 10 and 13**). The results were also similar if the risk of confounding was addressed with 2SLS, an alternative instrumental variable approach. (**Supplementary table 14**). The regression analyses which assumed that there were no unmeasured confounders, reported that the benefits of SGLT2i were greater than for the base case,

and more precisely estimated (**Supplementary tables 10 and 15**). When the study time frame was extended to five years, the gains following initial receipt of SGLT2i were maintained, although by this timepoint few people had complete follow-up information or were still prescribed the same second-line treatment (**Supplementary tables 6, 8, 10, 13-15**). The results were similar to the main analyses if IPTW or IPTW-weighted regression adjustment were used to address observed confounding (**Supplementary tables 16-20, Supplementary figures 10-11**).

DISCUSSION

Principal findings

In this comparative effectiveness study, we found that second-line treatment for people with T2DM was more effective with SGLT2i than with SU or DPP4i in reducing mean HbA1c, BMI, and SBP after reducing the risk of confounding with an instrumental variable analysis. SGLT2i were also more effective at reducing the hazards of heart failure hospitalisation (versus DPP4i) and ≥40% decline in eGFR (versus SU). We did not find strong evidence for other meaningful differences for the other study endpoints over the two-year study period.

In any study that aims to assess comparative effectiveness from routine data, a major concern is bias from confounding, in particular due to unmeasured prognostic differences between the comparison groups. This risk of bias can never be eliminated. However, a crucial advantage of our study design is that it followed recommended methods of target trial emulation in pre-specifying the population eligibility criteria, time 'zero', treatment comparisons, outcomes, and analyses.(41, 84, 85) Our main analysis used an IV to further reduce the risk of residual confounding. We are therefore able to provide useful evidence about the comparative effectiveness of these three treatments as they were prescribed in routine clinical practice for a diverse population of people with T2DM.

The aim of the PERMIT study was to assess the relative effectiveness of the three most common second-line treatments for an unselected population in routine clinical practice. By contrast, published RCTs have aimed to demonstrate the *safely and efficacy* of one of these drug classes *compared to placebo* in *selected populations.* For the comparison of SGLT2i versus SU, published RCT do not include general populations of people with T2DM who meet national guideline's eligibility criteria for these three second-line treatments (**Supplementary table 21**). (7) It is therefore challenging to interpret any comparison of the results of the PERMIT study with those of the published RCTs.

In **Supplementary tables 21-22**, we describe the results of the PERMIT study alongside those of the corresponding RCTs for common endpoints such as heart failure hospitalisation, MACE, MAKE and all-

cause death. We find that the point estimates for the PERMIT target trial emulation fall within the estimated 95% CI of the corresponding treatment effect reported in the RCTs, i.e. they meet previously defined criteria for 'agreement'(85) (**Supplementary table 22**). This concordance also applies to the few published RCTs, including the GRADE trial,(29, 36) that have compared two active treatments, DPP4is and SUs for general populations of people with T2DM. Unlike the GRADE trial, the PERMIT study did not exclude people with HbA1c outside the range 6.8-8.5%. A previous target trial emulated the GRADE trial in applying strict eligibility criteria, but unlike our study was unable to investigate MACE, heart failure, and all-cause mortality due to low event rates from a small study population. Our larger study found protective effects of SGLT2i for heart failure compared to DPP4i, similar to meta-analyses of RCTs(86, 87) and observational studies.(32) However, even with this relatively large sample, the number of people followed over the full follow-up period was insufficient to detect other clinically important differences for outcomes such as MAKE, and to investigate ESKD and CVD-specific mortality individually.

In our alternative analysis, we made the common assumption assuming no unmeasured confounding, and found that after adjusting for all measured confounders, SGLT2i were associated with greater improvement in all endpoints, including all-cause mortality. However, people prescribed SGLT2i had fewer comorbidities, and were likely to be healthier according to baseline characteristics that were not measured. A previous study that considered uptake of SGLT2i as second-line antidiabetic treatment also reported that compared to people who received SU or DPP4i, those who received SGLT2i were healthier, and at lower risk of all cause death.(34) For an endpoint such as all-cause death, it is particularly challenging to capture all the potential confounders from routine data sources (**Supplementary figure 1B**). In particular, for this endpoint important potential confounders include the individual's overall health, diet, exercise, and lifestyle prior to second-line treatment. If an instrumental variable is valid, it reduces the risk of bias from these unmeasured confounders, whereas approaches such as regression do not. Hence, the finding from the regression analysis that within the two-year follow-up period SGLT2i were associated with reduced hazards of all-cause mortality compared to SU or DPP4i could reflect these unmeasured baseline differences, i.e., residual confounding.

Strengths and weaknesses of the study and comparison with other studies

In this study, we directly compared the three most commonly prescribed second-line antidiabetic drug treatments using a large, linked dataset which is representative of the UK primary care population in terms of age and sex.(44, 45) Our direct comparison of SU, DPP4i, and SGLT2i is in contrast to previous

trials(16-18, 20, 26, 29, 36) and meta-analyses(86, 87) which did not include an active second-line treatment as a comparator. We did not restrict the study population to those with baseline HbA1c in a particular range as many RCTs have done previously.(16-19) This study therefore includes people with a broader range of glycaemic control, which is reflective of the UK primary care population with T2DM.

We add to the evidence reported in previous observational studies,(33-35, 40, 88) which make direct comparisons between antidiabetic treatments, by using an instrumental variable analysis as the main analysis to reduce the risk of confounding from both measured and unmeasured baseline confounders, and provide evidence on comparative effectiveness for those three drug classes that are most commonly prescribed in publicly funded health systems for a general population of people with T2DM. We investigated intermediate metabolic and other clinical measures, but also adverse kidney and cardiovascular events which are important to patients. The benefits we observe of SGLT2i improving HbA1c, BMI, and SBP, and reducing the risks of heart failure hospitalisation (versus DPP4i) and ≥40% decline in eGFR (versus SU), are indicative of a causal mechanism that has some biological plausibility.

Our DAGs provided a framework for the analysis that recognised second-line treatment with SGLT2i in routine practice could improve any of the intermediate clinical endpoints listed, which may in turn lead to reduced risks of subsequent events. In particular, the pharmacological action of SGLT2i, namely reducing blood pressure and cardiac pre- and after-load via diuretic mechanisms,(89) would imply protective effects on heart failure hospitalisation, and kidney endpoints; however, this would not necessarily translate to immediate protective effects during an ST-elevated myocardial infarction (STEMI)/acute coronary plaque rupture.

We acknowledge limitations in our study. We did not consider GLP1-RA since, in the UK, this class was rarely prescribed as a second-line antidiabetic treatment during the study period,(12, 13) and is still not recommended as a second-line treatment for people with T2DM.(12) However, GLP1-RA prescribing is increasing in the US, and warrants further study as the number of people prescribed these drugs increases in routinely collected data. Our instrumental variable analysis relies on three major assumptions. While we were able to empirically demonstrate that the instrument strongly predicts treatment receipt (assumption 1), we could only partially evaluate whether the instrument is balanced across confounders (assumption 2). We adjusted for measured confounders in the second stage of the regression model to account for any residual imbalances across levels of the instrument in particular with regard to time period and contextual measures such as region and GP size. However, if assumption 2 is not met then unmeasured confounders would be imbalanced across levels of the instrument leading to biased estimates. We must also assume that the instrument, the TTP, does not directly impact outcomes except via the treatment prescribed (assumption 3). We cannot test this

assumption, and it is possible it is violated if, for example, after adjusting for region, and practice size, there are CCGs with a higher TTP preference for SGLT2i who also deliver higher quality of care which impacts outcomes independent of the prescribing history.

The PERMIT study used routine data, and the requisite outcome data were not available for all those included. For continuous measures, the proportion of people with missing values at the one-year timepoint ranged from 36.4% (HbA1c) to 47.8% (BMI). In the main analysis we addressed these missing data for all the continuous outcomes along with any missing covariate information with multiple imputation, and undertook complete case analysis as alternative analyses. The results from these alternative approaches that make different underlying assumptions about why the data were missing gave similar results. For the time-to-event endpoints, we used linked primary and secondary care and ONS death datasets to ascertain cardiovascular, kidney, and mortality outcomes to improve the capture of events, rather than relying on a single source. However, a limitation shared with other target trial emulations using routine data, is that we do not know if data on events pertaining to kidney disease or CVD is 'missing'. People may experience an event that is diagnosed and recorded in outpatient clinics that is not recorded in the linked primary-secondary care data. For major events such as MI or stroke, levels of under-recording in the linked data are likely to be small and similar across the comparison groups, and lead to reduced statistical power rather than bias in the estimates of relative effectiveness.

While the study did consider endpoints up to five-years after initiation of second-line treatment, by this timepoint levels of missing data were high (from 46.9% HbA1C to 59.4% BMI), and after two years the majority of people will have stopped their second-line treatment. Hence, while we have reported results for the pre-specified five-year endpoint, given the levels of missing data, appropriate caution should be exercised when interpreting these results.

Policy implications

This study provides evidence that SGLT2i offer clinically important benefits when provided in routine clinical practice compared with common alternative oral antidiabetic medicines that are added to metformin for people with T2DM. These findings apply to a wide range of people with T2DM, and therefore complement the evidence available from RCTs,(16-24) and previous studies that have emulated trials.(35, 40) In recent updated guidelines, NICE and other health technology assessment (HTA) agencies have published guidance and guidelines which are neutral about the use of SGLT2i versus DPP4i versus SU as second-line treatments, unless people have pre-existing CVD including heart failure, are at high risk of CVD, or have kidney disease. For these subgroups, SGLT2i are recommended,

in addition to metformin. Our study reported similar advantages for SGLT2i (versus SU and DPP4i) as second-line treatments for people who did not have pre-existing CVD as well as for those who did. Future guidelines could draw from this study and related evidence to also recommend SGLT2i for those without CVD, including those at relatively low risk of subsequent CVD.

Further research is needed to understand the long-term effectiveness and cost-effectiveness of increasing the use of SGLT2i for people with T2DM. Future research can use the information from this study to predict whether SGLT2i can lead to sufficient improvement in long-term outcomes, for example, from reduced incidence and costs of complications such as retinopathy, amputation, or ESKD, to justify any additional costs. Further research is also required to assess the comparative effectiveness of GLP1-RA with the three alternative second-line oral antidiabetic treatments among people with T2DM, and to assess how best to personalise the order in which these treatments are prescribed.

Conclusions

We found that for a broad population of people with T2DM, compared to DPP4i or SU, SGLT2i are more effective second-line treatments in routine clinical practice in improving HbA1c, BMI, and SBP. SGLT2i were also more effective at reducing the hazards of heart failure hospitalisation (versus DPP4i) and ≥40% decline in eGFR (versus SU). We did not find evidence for differences in the other study endpoints over the two-year study period.

CONTRIBUTORS

PB, DGLP, OC, SON, ABA, KK, AIA, and RG were responsible for the conception and design of the study. All authors critically appraised and contributed to the study design and analysis plan. PB requested and extracted the data. PB, DGLP, OC, and SON performed the data management and analysis. PB, DGLP, OC, SON, and RG wrote the manuscript with contributions from all authors. All authors critically revised and approved the manuscript. PB and RG are the guarantors. PB and DGLP share first authorship. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

FUNDING

This work was funded by the National Institute for Health and Care Research (NIHR) grant number NIHR128490. The funder had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication. A CC BY license is required for publication.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: PB, DLP, OC, SON, RJS, LS, and RG have no disclosures. AIA receives salary from the NIHR BRC via the Oxford Centre for Diabetes, Endocrinology and Metabolism. AB is an economic advisor on the DiRECT trial, with ongoing responsibility for economic analysis during the long-term follow-up phase. He serves as a consultant to Salutis Consulting LLC on projects not related to diabetes. JWB has acted as a consultant to AstraZeneca; Bayer; Novartis; Roche. DN is the UK Kidney Association Director of Informatics Research; she previously was involved in two GSK funded studies in Sub-Saharan Africa unrelated to this work. PC sat on an NIHR HTA Commissioning Committee member until September 2021. AHB has acted as consultant to the following pharmaceutical companies within the past three years: Astra Zeneca; Boehringer Ingelheim; Daiichi Sankyo; Eli Lilly; Gilead; Idorsia; Novo Nordisk; Rhythm; Roche and Sanofi. IJD has unrestricted research grants from and share in GSK, and a research grant from AstraZeneca. KK has acted as a consultant, speaker or received grants for investigator-initiated studies for Astra Zeneca, Bayer, Novartis, Novo Nordisk, Sanofi-Aventis, Lilly and Merck Sharp & Dohme, Boehringer Ingelheim, Oramed Pharmaceuticals, Roche and Applied Therapeutics.

PATIENT CONSENT

Individual patient consent was not needed for this study.

ETHICAL APPROVAL

This research was approved by the London School of Hygiene & Tropical Medicine ethics committee (reference 21 395) and the Independent Scientific Advisory Committee (reference 20_064). While

individual consent was not obtained, general practices opt-in to sharing their data for research and individual people can opt-out of sharing their data for research.

DATA SHARING

Due to data-sharing restrictions, we cannot share the data used in this study directly. However, researchers may apply to use CPRD data linked with other health datasets. Please see the CPRD website for further instruction https://cprd.com/. Codelists to create exposure, outcome, and covariates are published on LSHTM DataCompass: https://datacompass.lshtm.ac.uk/id/eprint/3743/

TRANSPARENCY STATEMENT

The lead authors (PB and DGLP) and the manuscript's guarantor (PB and RG) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned have been explained.

DISSEMINATION TO PARTICIPANTS AND RELATED PATIENT AND PUBLIC COMMUNITIES

After publication, the results will be disseminated to policy makers, including those at the National Institute of Care and Clinical Excellence (NICE). We have shared preliminary methodology and results related to this work with two patient panels. With the help of our PPI and clinical colleagues, we will disseminate the study results after publication with patients, the public, and healthcare professionals through press releases, media communications, social media postings, and presentations at scientific conferences.

PROVENANCE AND PEER REVIEW

Provenance and peer review: Not commissioned; externally peer reviewed.

ACKNOWLEDGMENTS

This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data are provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the authors alone. The study was approved by the Independent Scientific Advisory Committee (approval number: 20_064).

We acknowledge and appreciate the support of the PERMIT study advisory team (Ken Paterson, Jim Lewsey, Ewan Pearson, Sarah Finer, Stephen Evans, Bill Huston, and Rahul Mohan), and the management and administrative support provided by Elizabeth Silver, Anna Carnegie and Paula Fry.

KK is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration East Midlands (ARC EM) and the NIHR Leicester Biomedical Research Centre (BRC).

PB and DGLP share first authorship.

REFERENCES

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. Diabetes research and clinical practice. 2019;157:107843.

2. Rodriguez-Gutierrez R, McCoy RG. Measuring What Matters in Diabetes. Jama. 2019;321(19):1865-6.

3. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet (London, England). 1998;352(9131):837-53.

4. Rodriguez-Gutierrez R, Gonzalez-Gonzalez JG, Zuñiga-Hernandez JA, McCoy RG. Benefits and harms of intensive glycemic control in patients with type 2 diabetes. BMJ (Clinical research ed). 2019;367:l5887.

5. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2020. Diabetes care. 2020;43(Suppl 1):S66-s76.

6. TA390: Canagliflozing, dapagliflozin and empagliflozin as monotherapies for treating type 2 diabetes: NICE; 2016 [

7. NICE guideline [NG28]: Type 2 diabetes in adults: management [Web]. NICE; 2022 [Available from: https://www.nice.org.uk/guidance/ng28/chapter/Recommendations#reviewing-drug-treatments.

8. SIGN 154: Pharmacological management of glycaemic control in people with type 2 diabetes: SIGN; 2017 [

9. Davies MJ, Aroda VR, Collins BS, Gabbay RA, Green J, Maruthur NM, et al. Management of Hyperglycemia in Type 2 Diabetes, 2022. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes care. 2022;45(11):2753-86.

10. Khunti K, Charbonnel B, Cooper A, Gomes MB, Ji L, Leigh P, et al. Associations between second-line glucose-lowering combination therapies with metformin and HbA1c, body weight, quality of life, hypoglycaemic events and glucose-lowering treatment intensification: The DISCOVER study. Diabetes, Obesity and Metabolism. 2021;23(8):1823-33.

11. Bidulka P, Mathur R, Lugo-Palacios DG, O'Neill S, Basu A, Silverwood RJ, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. Diabetes, Obesity and Metabolism. 2023;25(1):282-92.

12. Wilkinson S, Douglas I, Stirnadel-Farrant H, Fogarty D, Pokrajac A, Smeeth L, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. BMJ Open. 2018;8(7):e022768.

13. Farmer RE, Beard I, Raza SI, Gollop ND, Patel N, Tebboth A, et al. Prescribing in Type 2 Diabetes Patients With and Without Cardiovascular Disease History: A Descriptive Analysis in the UK CPRD. Clinical Therapeutics. 2021;43(2):320-35.

14. Jennifer RD, Catherine AG, Eugene C, Carlo AM, Kris A-B, Karissa J, et al. Comparative safety of the sodium glucose co-transporter 2 (SGLT2) inhibitors: a systematic review and meta-analysis. BMJ Open. 2019;9(1):e022577.

15. Baigent C, Emberson J, Haynes R, Herrington WG, Judge P, Landray MJ, et al. Impact of diabetes on the effects of sodium glucose co-transporter-2 inhibitors on kidney outcomes: collaborative meta-analysis of large placebo-controlled trials. The Lancet.

16. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. New England Journal of Medicine. 2015;373(22):2117-28.

17. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. New England Journal of Medicine. 2017;377(7):644-57.

18. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine. 2018;380(4):347-57.

19. Cannon CP, Pratley R, Dagogo-Jack S, Mancuso J, Huyck S, Masiukiewicz U, et al. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. New England Journal of Medicine. 2020;383(15):1425-35.

20. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJL, Charytan DM, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. New England Journal of Medicine. 2019;380(24):2295-306.

21. Empagliflozin in Patients with Chronic Kidney Disease. New England Journal of Medicine. 2022;388(2):117-27.

22. McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, et al. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. New England Journal of Medicine. 2019;381(21):1995-2008.

23. Heerspink HJL, Stefánsson BV, Correa-Rotter R, Chertow GM, Greene T, Hou F-F, et al. Dapagliflozin in Patients with Chronic Kidney Disease. New England Journal of Medicine. 2020;383(15):1436-46.

24. Packer M, Anker SD, Butler J, Filippatos G, Pocock SJ, Carson P, et al. Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure. New England Journal of Medicine. 2020;383(15):1413-24.

25. Rosenstock J, Perkovic V, Johansen OE, Cooper ME, Kahn SE, Marx N, et al. Effect of Linagliptin vs Placebo on Major Cardiovascular Events in Adults With Type 2 Diabetes and High Cardiovascular and Renal Risk: The CARMELINA Randomized Clinical Trial. Jama. 2019;321(1):69-79.

26. Green JB, Bethel MA, Armstrong PW, Buse JB, Engel SS, Garg J, et al. Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine. 2015;373(3):232-42.

27. Ferrannini E, Fonseca V, Zinman B, Matthews D, Ahren B, Byiers S, et al. Fifty-two-week efficacy and safety of vildagliptin vs. glimepiride in patients with type 2 diabetes mellitus inadequately controlled on metformin monotherapy. Diabetes Obes Metab. 2009;11(2):157-66.

28. Nauck MA, Meininger G, Sheng D, Terranella L, Stein PP, Sitagliptin Study G. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. Diabetes, obesity & metabolism. 2007;9(2):194-205.

29. Glycemia Reduction in Type 2 Diabetes — Glycemic Outcomes. New England Journal of Medicine. 2022;387(12):1063-74.

30. Rosenstock J, Kahn SE, Johansen OE, Zinman B, Espeland MA, Woerle HJ, et al. Effect of Linagliptin vs Glimepiride on Major Adverse Cardiovascular Outcomes in Patients With Type 2 Diabetes: The CAROLINA Randomized Clinical Trial. Jama. 2019;322(12):1155-66.

31. Leiter LA, Yoon KH, Arias P, Langslet G, Xie J, Balis DA, et al. Canagliflozin provides durable glycemic improvements and body weight reduction over 104 weeks versus glimepiride in patients with type 2 diabetes on metformin: a randomized, double-blind, phase 3 study. Diabetes care. 2015;38(3):355-64.

32. D'Andrea E, Wexler DJ, Kim SC, Paik JM, Alt E, Patorno E. Comparing Effectiveness and Safety of SGLT2 Inhibitors vs DPP-4 Inhibitors in Patients With Type 2 Diabetes and Varying Baseline HbA1c Levels. JAMA internal medicine. 2023;183(3):242-54.

33. Wilkinson S, Williamson E, Pokrajac A, Fogarty D, Stirnadel-Farrant H, Smeeth L, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. Diabetes, Obesity and Metabolism. 2020;22(5):847-56.

34. Khunti K, Kosiborod M, Kim DJ, Kohsaka S, Lam CSP, Goh SY, et al. Cardiovascular outcomes with sodium-glucose cotransporter-2 inhibitors vs other glucose-lowering drugs in 13 countries across three continents: analysis of CVD-REAL data. Cardiovasc Diabetol. 2021;20(1):159.

35. Deng Y, Polley EC, Wallach JD, Dhruva SS, Herrin J, Quinto K, et al. Emulating the GRADE trial using real world data: retrospective comparative effectiveness study. BMJ (Clinical research ed). 2022;379:e070717.

36. Glycemia Reduction in Type 2 Diabetes — Microvascular and Cardiovascular Outcomes. New England Journal of Medicine. 2022;387(12):1075-88.

37. Wexler DJ, de Boer IH, Ghosh A, Younes N, Bebu I, Inzucchi SE, et al. Comparative Effects of Glucose-Lowering Medications on Kidney Outcomes in Type 2 Diabetes: The GRADE Randomized Clinical Trial. JAMA internal medicine. 2023.

38. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med. 2014;33(13):2297-340.

39. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. The international journal of biostatistics. 2007;3(1):Article 14.

40. Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. The Lancet Diabetes & Endocrinology. 2023;11(9):644-56.

41. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.

42. Bidulka P, O'Neill S, Basu A, Wilkinson S, Silverwood RJ, Charlton P, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. BMJ Open. 2021;11(9):e046912.

43. Bidulka P, Lugo-Palacios DG, Carroll O, O'Neill S, Grieve R. Statistical Analysis Plan (SAP): PERsonalised Medicine for Intensification of Treatment (PERMIT) study, Version 1.0, July 2023 [Web]. 2023 [Available from: https://www.lshtm.ac.uk/media/72276.

44. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.

45. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. International Journal of Epidemiology. 2019;48(6):1740-g.

46. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). International journal of epidemiology. 2017;46(4):1093-i.

47. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. BMJ : British Medical Journal. 2013;346:f2350.

48. Bidulka P, Scott J, Taylor DM, Udayaraj U, Caskey F, Teece L, et al. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. BMJ Open. 2022;12(3):e057909.

49. Linked HES-ONS mortality data [Web]. NHS Digital; 2020 [Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data#ons-mortality-data.

50. Penney B. The English Indices of Deprivation 2019 (IoD2019). 2019.

51. Wilkinson S, Douglas IJ, Williamson E, Stirnadel-Farrant HA, Fogarty D, Pokrajac A, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. Clin Epidemiol. 2018;10:1639-48.

52. NHS England - About Us - Regional Teams [Web]. NHS England; [Available from: https://www.england.nhs.uk/about/regional-area-teams/.

53. Bhaskaran K, Forbes HJ, Douglas I, Leon DA, Smeeth L. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). BMJ Open. 2013;3(9):e003389.

54. Levey AS, Inker LA, Matsushita K, Greene T, Willis K, Lewis E, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. American journal of kidney diseases : the official journal of the National Kidney Foundation. 2014;64(6):821-35.

55. Levin A, Agarwal R, Herrington WG, Heerspink HL, Mann JFE, Shahinfar S, et al. International consensus definitions of clinical trial outcomes for kidney failure: 2020. Kidney Int. 2020;98(4):849-59.

56. Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments. Econometrica. 1997;65(3):557-86.

57. Moler-Zapata S, Grieve R, Basu A, O'Neill S. How does a local instrumental variable method perform across settings with instruments of differing strengths? A simulation study and an evaluation of emergency surgery. Health Econ. 2023.

58. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and evaluation of the monotonicity condition for preference-based instruments. Epidemiology (Cambridge, Mass). 2015;26(3):414-20.

59. Vytlacil E. Independence, Monotonicity, and Latent Index Models: An Equivalence Result. Econometrica. 2002;70(1):331-41.

60. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. Journal of Health Economics. 2008;27(3):531-43.

61. Basu A, Coe NB, Chapman CG. 2SLS versus 2SRI: Appropriate methods for rare outcomes and/or rare exposures. Health Econ. 2018;27(6):937-55.

62. Frank lE, Friedman JH. A Statistical View of Some Chemometrics Regression Tools. Technometrics. 1993;35(2):109-35.

63. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996;58(1):267-88.

64. Belloni A, Chen D, Chernozhukov V, Hansen C. Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. Econometrica. 2012;80(6):2369-429.

65. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Hoboken, NJ: John Wiley & Sons, Inc.; 1987.

66. van Buuren S, Oudshoorn, C.G.M. Multivariate Imputation by Chained Equations: MICE V1.0 User's manual. TNO Report PG/VGZ/00.038. Leiden, Netherlands; 2000.

67. Powney M, Williamson P, Kirkham J, Kolamunnage-Dona R. A review of the handling of missing longitudinal outcome data in clinical trials. Trials. 2014;15:237.

68. Lee KJ, Roberts G, Doyle LW, Anderson PJ, Carlin JB. Multiple imputation for missing data in a longitudinal cohort study: a tutorial based on a detailed case study involving imputation of missing outcome data. International Journal of Social Research Methodology. 2016;19(5):575-91.

69. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Medical Research Methodology. 2014;14(1):75.

70. White IR, Royston P. Imputing missing covariate values for the Cox model. Stat Med. 2009;28(15):1982-98.

71. Bartlett JW, Hughes RA. Bootstrap inference for multiple imputation under uncongeniality and misspecification. Statistical Methods in Medical Research. 2020;29(12):3533-46.

72. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. Stat Med. 2018;37(14):2252-66.

73. Therneau T. A Package for Survival Analysis in R: R package version 3.4-0 2022 [Available from: https://CRAN.R-project.org/package=survival.

74. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011;45(3):1 - 67.

75. Team RC. R: A language and environment for statistical computing [Web]. Vienna, Austria: R Foundation for Statistical Computing; 2022 [Available from: https://www.R-project.org/.

76. StataCorp. Stata Statistical Software: Release 17 College Station, Texas: StataCorp LLC; 2021 [

77. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. BMJ (Clinical research ed). 2019;367:l5657.

78. Wooldridge JM. Inverse probability weighted estimation for general missing data problems. Journal of Econometrics. 2007;141(2):1281-301.

79. Xu S, Ross C, Raebel MA, Shetterly S, Blanchette C, Smith D. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2010;13(2):273-7.

80. Stürmer T, Webster-Clark M, Lund JL, Wyss R, Ellis AR, Lunt M, et al. Propensity Score Weighting and Trimming Strategies for Reducing Variance and Bias of Treatment Effect Estimates: A Simulation Study. Am J Epidemiol. 2021;190(8):1659-70.

81. Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. Am J Epidemiol. 2010;172(7):843-54.

82. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly Robust Estimation of Causal Effects. Am J Epidemiol. 2011;173(7):761-7.

83. PERMIT: Investigating the effectiveness of different treatments for Type 2 Diabetes Mellitus, to help choose the right treatments for patients [Web]. London School of Hygiene & Tropical Medicine; 2023 [Available from: https://www.lshtm.ac.uk/research/centres-projects-groups/permit#welcome.

84. Hansford HJ, Cashin AG, Jones MD, Swanson SA, Islam N, Douglas SRG, et al. Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials: A Systematic Review. JAMA Netw Open. 2023;6(9):e2336023.

85. Wang SV, Schneeweiss S, Initiative R-D. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. Jama. 2023;329(16):1376-85.

86. Type 2 diabetes [B] Pharmacological therapies with cardiovascular and other benefits in people with type 2 diabetes, NICE Guideline NG28 [Web]. National Institute for Health and Care Excellence (NICE); 2022 [Available from: https://www.nice.org.uk/guidance/ng28/evidence/b-pharmacological-therapies-with-cardiovascular-and-other-benefits-in-people-with-type-2-diabetes-pdf-10956473392.

87. Sim R, Chong CW, Loganadan NK, Fong AYY, Navaravong L, Hussein Z, et al. Comparative effectiveness of cardiovascular, renal and safety outcomes of second-line antidiabetic drugs use in people with type 2 diabetes: A systematic review and network meta-analysis of randomised controlled trials. Diabetic Medicine. 2022;39(3):e14780.

88. Cavender MA, Norhammar A, Birkeland KI, Jørgensen ME, Wilding JP, Khunti K, et al. SGLT-2 Inhibitors and Cardiovascular Risk: An Analysis of CVD-REAL. J Am Coll Cardiol. 2018;71(22):2497-506.

89. Ni L, Yuan C, Chen G, Zhang C, Wu X. SGLT2i: beyond the glucose-lowering effect. Cardiovascular Diabetology. 2020;19(1):98.

Table 5.1. Baseline characteristics of the primary-secondary care linked study population, stratified by the second-line antidiabetic treatment prescribed. n (column %) unless specified otherwise.

	SU	DPP4i	SGLT2i
N (row %)	**25,693 (34)**	**34,464 (46)**	**15,582 (20)**
Female	9,840 (38.3)	13,456 (39.0)	5,996 (38.5)
Age in years, mean (SD)	60 (13)	62 (12)	56 (11)
Ethnicity			
White	19,665 (76.5)	27,308 (79.2)	12,543 (80.5)
South Asian	3,522 (13.7)	4,616 (13.4)	1,961 (12.6)
Black	1,625 (6.3)	1,451 (4.2)	542 (3.5)
Mixed/Other	534 (2.1)	612 (1.8)	231 (1.5)
Missing	347 (1.4)	477 (1.4)	305 (2.0)
Index of Multiple Deprivation quintile			
1 (least deprived)	3,619 (14.1)	5,161 (15.0)	2,604 (16.7)
2	4,504 (17.5)	6,175 (17.9)	2,793 (17.9)
3	4,955 (19.3)	6,642 (19.3)	2,953 (19.0)
4	6,152 (23.9)	7,677 (22.3)	3,408 (21.9)
5 (most deprived)	6,449 (25.1)	8,785 (25.5)	3,815 (24.5)
Missing	14 (0.1)	24 (0.1)	9 (0.1)
Year of second-line antidiabetic treatment initiation			
2015	6,996 (27.2)	4,958 (14.4)	1,145 (7.3)
2016	5,221 (20.3)	6,057 (17.6)	1,525 (9.8)
2017	4,260 (16.6)	6,309 (18.3)	2,178 (14.0)
2018	3,562 (13.9)	6,771 (19.6)	2,912 (18.7)
2019	3,112 (12.1)	5,995 (17.4)	3,939 (25.3)
2020	2,542 (9.9)	4,374 (12.7)	3,883 (24.9)
Years on first-line treatment (metformin monotherapy), median (IQR)	2.9 (1.1-5.4)	3.6 (1.7-6.3)	2.8 (1.3-5.2)
General Practice size, mean number of patients registered, median (IQR)	9690 (6250-13628)	9971 (6538-13795)	10143 (6896-13881)
Last HbA1c value (mmol/mol) recorded prior to index date, mean (SD)	81 (22)	72 (16)	75 (17)
Last HbA1c value (%) recorded prior to index date, mean (SD)	9.1 (2.1)	8.2 (1.5)	8.5 (1.6)
Last HbA1c value (mmol/mol) recorded prior to index date			
<53	713 (2.8)	1,053 (3.1)	515 (3.3)
53-74	10,818 (42.1)	21,870 (63.5)	8,410 (54.0)
75+	12,579 (49.0)	10,398 (30.2)	6,134 (39.4)
Missing	1,583 (6.2)	1,143 (3.3)	523 (3.4)
Last systolic blood pressure measure (mm Hg) recorded prior to index date, mean (SD)	132 (14)	132 (14)	133 (14)

	SU	DPP4i	SGLT2i
N (row %)	**25,693 (34)**	**34,464 (46)**	**15,582 (20)**
Last diastolic blood pressure measure (mm Hg) recorded prior to index date, mean (SD)	78 (9)	77 (9)	80 (9)
Hypertensive, based on last recorded blood pressure measure			
Normotensive	7,123 (27.7)	9,424 (27.3)	3,664 (23.5)
Hypertensive	18,525 (72.1)	25,002 (72.5)	11,906 (76.4)
Missing	45 (0.2)	38 (0.1)	12 (0.1)
BMI at index date (kg/m^2), mean (SD)	31.5 (6.6)	32.2 (6.5)	35.1 (7.0)
BMI at index date (kg/m^2)			
Under/normal weight	2,718 (10.6)	2,782 (8.1)	394 (2.5)
Overweight	8,110 (31.6)	10,180 (29.5)	2,867 (18.4)
Obese	14,702 (57.2)	21,375 (62.0)	12,283 (78.8)
Missing	163 (0.6)	127 (0.4)	38 (0.2)
Last recorded eGFR (mL/min/1.73m^2) prior to index date, mean (SD)	91 (19)	88 (19)	97 (15)
eGFR category (mL/min/1.73m^2)			
Stage 1-2 (eGFR≥60)	23,282 (90.6)	30,823 (89.4)	15,186 (97.5)
Stage 3a-3b (eGFR 30-59)	1,770 (6.9)	3,199 (9.3)	161 (1.0)
Missing	641 (2.5)	442 (1.3)	235 (1.5)
Comorbidities			
Prevalent CVD	5,858 (22.8)	8,108 (23.5)	2,680 (17.2)
Previous amputation	227 (0.9)	265 (0.8)	76 (0.5)
Heart failure	1,457 (5.7)	2,007 (5.8)	598 (3.8)
Previous myocardial infarction	1,644 (6.4)	2,226 (6.5)	842 (5.4)
Previous stroke	1,378 (5.4)	1,678 (4.9)	512 (3.3)
Ischaemic heart disease	4,572 (17.8)	6,538 (19.0)	2,175 (14.0)
Unstable angina	777 (3.0)	1,099 (3.2)	362 (2.3)
History of cancer	4,254 (16.6)	5,397 (15.7)	1,447 (9.3)
Blindness	425 (1.7)	527 (1.5)	140 (0.9)
Previous hypoglycaemia	260 (1.0)	302 (0.9)	129 (0.8)
Proteinuria	3,658 (14.2)	4,679 (13.6)	1,585 (10.2)
Co-prescriptions			
RASi	12,584 (49.0)	18,911 (54.9)	8,108 (52.0)
Statin	17,729 (69.0)	25,690 (74.5)	10,838 (69.6)
Smoking status			
Non-smoker	5,720 (22.3)	7,455 (21.6)	3,562 (22.9)
Ex-smoker	12,640 (49.2)	18,009 (52.3)	7,865 (50.5)
Current smoker	7,327 (28.5)	8,992 (26.1)	4,154 (26.7)
Missing	6 (0.0)	8 (0.0)	1 (0.0)
Alcohol intake			
Non-drinker	3,192 (12.4)	3,716 (10.8)	1,630 (10.5)
Ex-drinker	7,248 (28.2)	10,009 (29.0)	4,179 (26.8)
Current drinker	14,899 (58.0)	20,367 (59.1)	9,582 (61.5)
Missing	354 (1.4)	372 (1.1)	191 (1.2)

BMI: body-mass index; DPP4: dipeptidyl peptidase 4 inhibitors; eGFR: estimated glomerular filtration rate; IQR: interquartile range; RASi: renin-angiotensin system inhibitors; SD: standard deviation

Figure 5.1. Stacked bar chart illustrating the variation in second-line antidiabetic treatment prescribed among people included in the study at the clinical commissioning group (CCG) level in England, 2014-2020

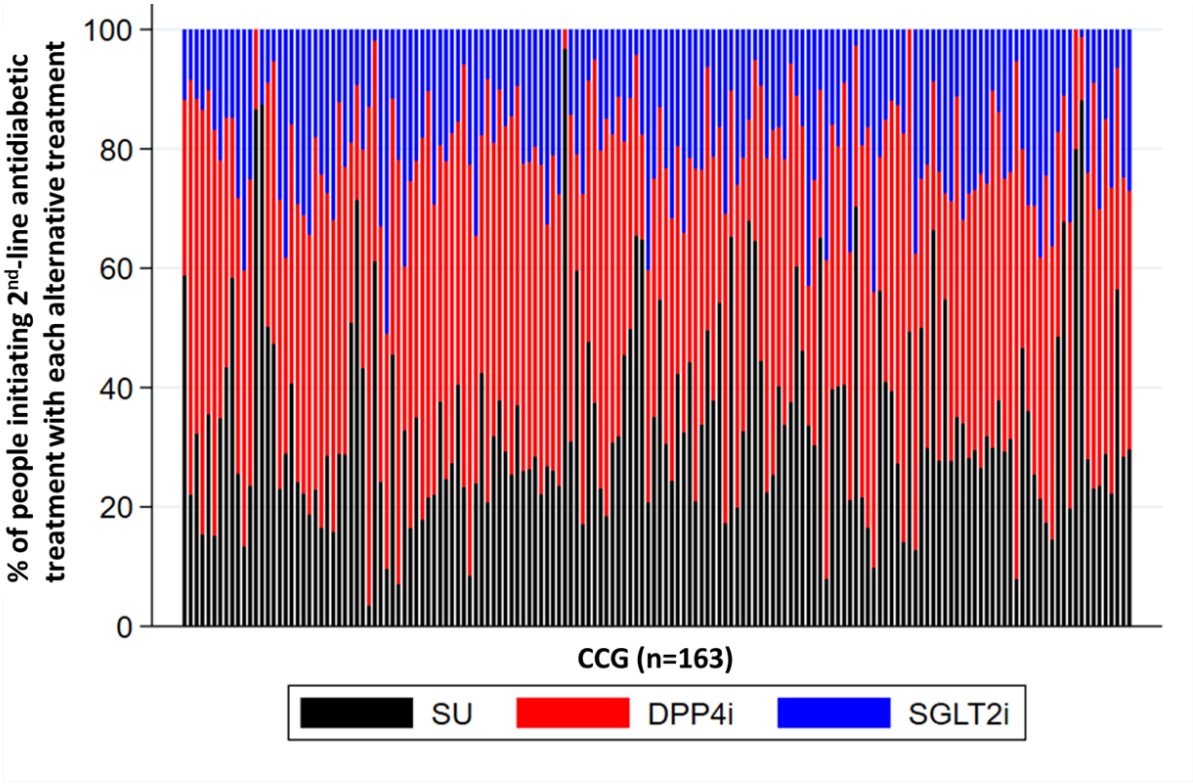


Figure 5.2. Flow diagram illustrating the study population inclusion and exclusion criteria

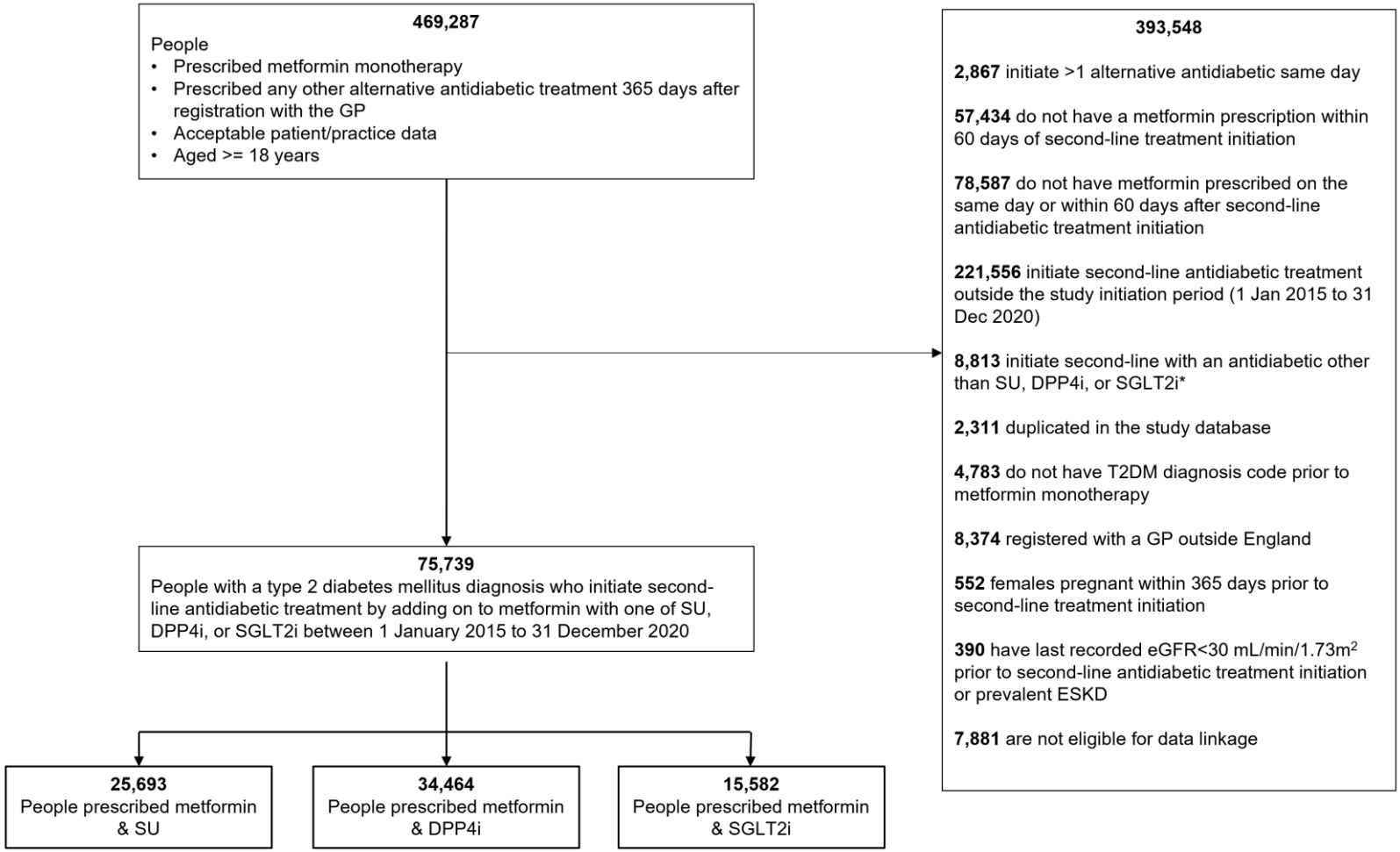


Figure 5.3. Mean haemoglobin A1c (HbA1c, mmol/mol), estimated glomerular filtration rate (eGFR, mL/min/1.73m2), body-mass index (BMI, kg/m2), and systolic blood pressure (systolic BP, mm Hg) at each follow-up time point of interest, stratified by treatment group

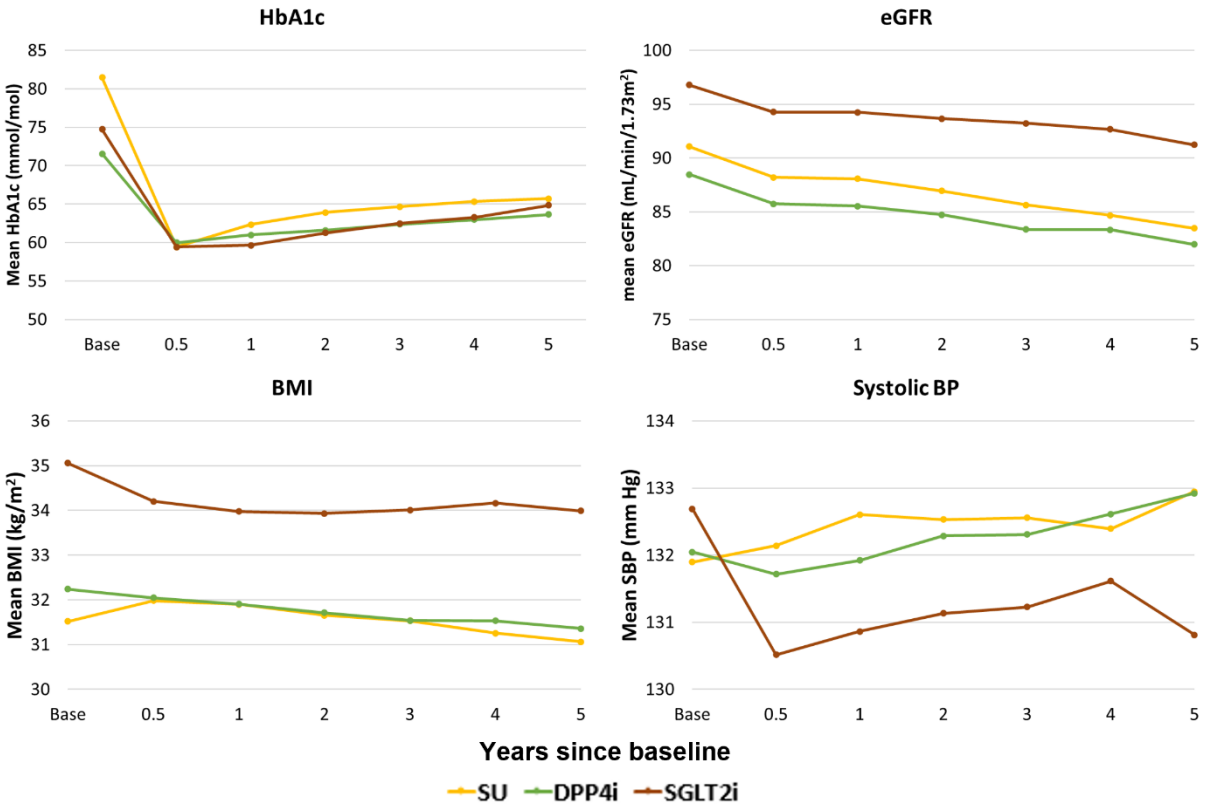


Figure 5.4. Forest plot showing differences in the change between baseline and 1- or 2- year follow-up in continuous clinical measures for (i) SGLT2i (A) compared to SU (B), (ii) SGLT2i (A) compared to DPP4i (B), and (iii) DPP4i (A) compared to SU (B).

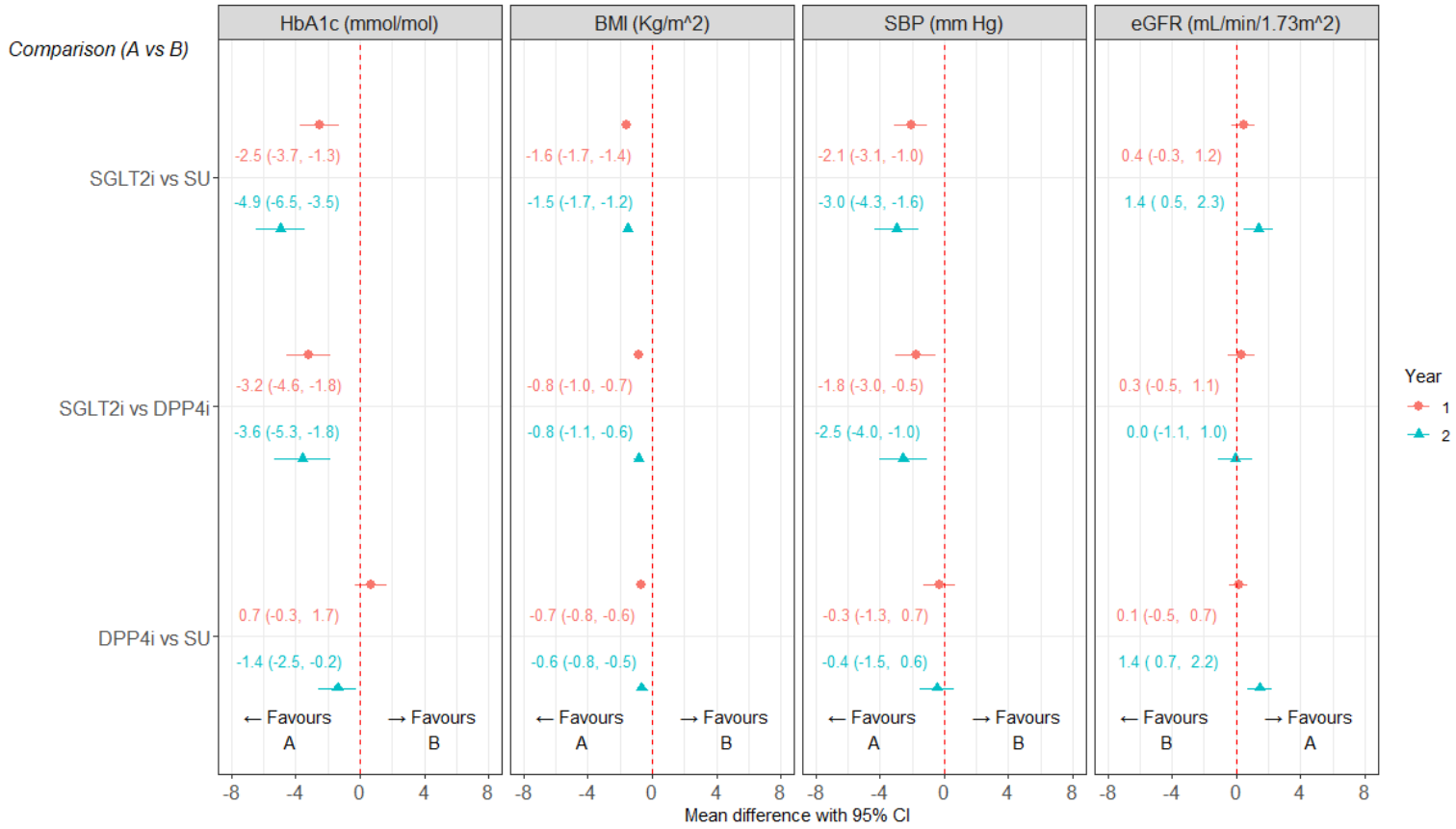
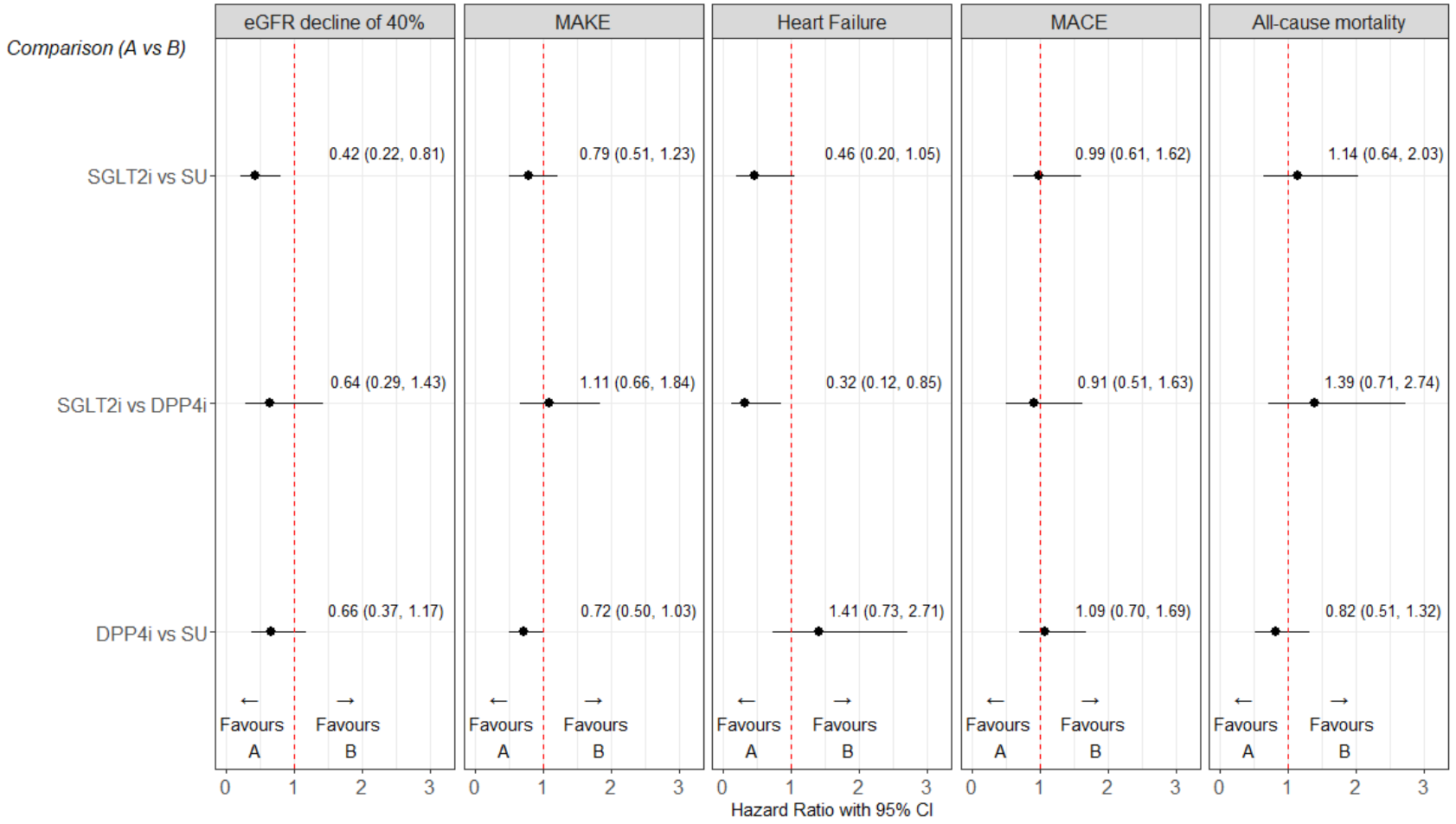


Figure 5.5. Forest plot showing adjusted hazard ratios for instrumental variable survival analysis for CVD and kidney outcomes when comparing (i) SGLT2i to SU, (ii) SGLT2i to DPP4i, and (iii) DPP4i to SU.



5.2. Relevance to my thesis

This research paper addressed thesis objectives 1B to C: to describe variation in second-line oral antidiabetic treatment prescribing across CCGs in England and use this variation as an instrument in an IV analysis to compare these alternative treatments on outcomes important to patients, healthcare providers, and policymakers alike (see also steps step 3 to 5 of **Figure 2.1.**).

In this research, I confirm the variation in second-line antidiabetic treatment prescribing across CCGs in England which was previously described by Wilkinson et al (2018) in a less contemporary cohort.[1] I used this variation in an IV analysis, which can reduce the risk of confounding provided several assumptions are met. These assumptions are carefully described in the paper, and I provide evidence that suggests they are plausible in this setting using DAGs (Appendices E.1. to E.2.) and covariate balance plots according to the instrument (Appendices E.3. to E.5.). Supplementary information about the methods (including model selection and multiple imputation) are described in Appendices E.6.

I chose a 12-month look-back window to define the TTP (the instrument) in this chapter's analyses. This was shown in unpublished pilot work using the cohort defined by Wilkinson et al (2018) to strongly predict the treatment actually prescribed,[1-3] suggesting it would be a strong instrument. I pre-specified the 12-month previous prescribing history in the study protocol, with the input of a multidisciplinary panel of GPs, diabetologists, epidemiologists, and statisticians, as an appropriate look-back window that would capture the TTP of each CCG included in the study, with updates over the study time period as preferences for the second-line antidiabetic treatments changed over the course of the 5-year study period.

Pharmacoepidemiological studies using an IV analysis can include sensitivity analyses to consider different look-back windows when defining the preference-based instrument.[4] Often, this is to consider whether a different look-back window would strengthen the instrument in predicting the treatment, thus reducing the risk of bias in the IV analysis. Because (i) the IV used in Case Study 1 of this thesis was strong based on the partial F-statistics in the first-stage regression models and (ii) the IV analyses, which included multiple imputation and bootstrapping, were computationally intense, I chose not to include a sensitivity analysis considering other look-back windows (e.g., 6-months). Instead, I conducted several alternative analyses which helped evaluate the IV assumptions, including the 2SLS IV model instead of the 2SRI IV model, a complete case analysis, a propensity score with IPTW-RA, and traditional multivariable regression. In future work including a sensitivity analysis to compare different look-back windows for the preference-based IV would be helpful, particularly where the instrument is weaker.

In the research paper presented in this chapter, I found that SGLT2i were more effective than both DPP4i and SU (all added to metformin monotherapy) at reducing mean HbA1c, BMI, and SBP (Appendix E.7.), and reducing the hazards of heart failure hospitalisation (versus DPP4i) and kidney disease progression (versus SU) (Appendix E.8.).

This research paper followed the protocol described in Chapter 4 and applied the TTE framework[5,6] to emulate the ideal RCT which could investigate the comparative effectiveness of these three alternative antidiabetic treatments. I included a table in the supplementary materials of this paper, which I have copied into this chapter (**Table 5.2.**), to explicitly report the key criteria for the ideal RCT and the observational study using an IV analysis presented here. In this study, we could not emulate a published RCT, since there is no published RCT at the time of writing this thesis, to my knowledge, that directly compares SU, DPP4i, and SGLT2i added to metformin monotherapy as second-line oral antidiabetic treatment.

Table 5.2. Summary of the target trial emulation in the research paper presented in Chapter 5, copied from the supplementary materials of the research paper (Supplementary table 1: Summary of target trial emulation)

	Target trial	**Emulation**
Eligibility criteria	Inclusion criteria: • People aged ≥18 years with a T2DM diagnosis. • Initiate first-line oral antidiabetic treatment with metformin monotherapy. • Initiate second-line oral antidiabetic treatment with one of SU, DPP4i or SGLT2i added on to metformin. Exclusion criteria: • Women with a record of pregnancy within 12 months prior to second-line treatment initiation. • People whose last recorded eGFR<30mL/min/1.73m^2. • People whose primary care data cannot be linked to secondary care data (essential for outcome definitions).	Inclusion criteria: • People aged ≥18 years with a T2DM diagnosis. • Initiate first-line oral antidiabetic treatment with metformin monotherapy. • Initiate second-line oral antidiabetic treatment with one of SU, DPP4i or SGLT2i added on to metformin. • At least one metformin prescription within 60 days prior to second-line treatment initiation. • At least one metformin prescription on the same day or within 60 days post-second-line treatment initiation. Exclusion criteria: • Women with a record of pregnancy within 12 months prior to second-line treatment initiation. • People whose last recorded eGFR<30mL/min/1.73m^2. • People whose primary care data cannot be linked to secondary care data (essential for outcome definitions).
Treatment assignment	Participants randomly assigned to add one of SU, DPP4i, or SGLT2i to metformin monotherapy.	We used the tendency to prescribe DPP4i or SGLT2i versus SU as the instrumental variable for receipt of these alternative second-line oral antidiabetic treatments. The instrumental variable analysis aimed to reduce the risk of confounding (thus mimicking randomisation in the target trial) (see details in Main Text (Methods pages 14-15, Supplementary figures 1A-B).
Treatment initiation	Initiation of one of SU, DPP4i or SGLT2i, all added to metformin, at randomisation.	We used GP prescriptions from the CPRD for one of SU, DPP4i, or SGLT2i, added to metformin monotherapy. The day of first prescription for SU, DPP4i, or SGLT2i served as the index date. All participants must also have had a prescription for metformin on the same day or within 60 days post index date to ensure that participants are adding on to metformin monotherapy rather than stopping metformin when switching to SU or DPP4i, or SGLT2i.[2,3]

	Target trial	Emulation
Treatment strategy	The duration of second-line treatment, and then all subsequent treatments, including reversion to monotherapy, or further intensification with additional oral treatments of insulin was determined over follow-up. Participants may change their treatment through the course of the study. Changes may be captured using additional GP prescribing data.	The duration of second-line treatment was extracted from prescription data. Information was collected on whether participants changed their treatment during study follow-up, and the form of treatment and duration of any subsequent treatment during the follow-up period. All continuous courses of treatment were defined using the duration field of the CPRD prescribing data. A grace period of 60 days was added to the end of each prescription to allow for delays in filling new prescriptions for a continuous course of treatment. (See also causal contrasts).
Follow-up	Follow-up starts at treatment initiation. Participants are followed until 31 December 2021. Death/outcome of interest are censoring events.	Follow-up started at treatment initiation. Participants were followed until the outcome date, or 31 December 2021 (continuous outcomes defined in primary care, e.g., HbA1c) or 31 March 2021 (time-to-event outcomes defined in primary or secondary care, e.g., MACE). Linked hospital data were only available up to 31 March 2021. Death/outcome of interest are censoring events.
Outcomes	**Primary outcome:** Change in HbA1c (mmol/mol) at 1 year follow-up. **Secondary outcomes:** Change in BMI, systolic blood pressure, eGFR at 0.5, 1, 2, 3, 4, 5 years follow-up, and change in HbA1c at 0.5, 2, 3, 4, 5 years follow-up. 40% decline in eGFR from baseline. Major adverse kidney event (MAKE): composite of 40% decline in eGFR from baseline, end-stage kidney disease, or all-cause death. Heart failure hospitalisation. Major adverse cardiovascular event (MACE): stroke, myocardial infarction, or cardiovascular-specific death. All-cause death.	**Primary outcome:** Change in HbA1c (mmol/mol) at 1 year follow-up. **Secondary outcomes:** Change in BMI, systolic blood pressure, eGFR at 0.5, 1, 2, 3, 4, 5 years follow-up, and change in HbA1c at 0.5, 2, 3, 4, 5 years follow-up. 40% decline in eGFR from baseline. Major adverse kidney event (MAKE): composite of 40% decline in eGFR from baseline, end-stage kidney disease, or all-cause death. Heart failure hospitalisation. Major adverse cardiovascular event (MACE): stroke, myocardial infarction, or cardiovascular-specific death. All-cause death.

	Target trial	Emulation
Causal contrasts of interest	Intention-to-treat average treatment effect.	Intention-to-treat average treatment effect.
Analysis plan to estimate causal contrasts of interest	Multivariable survival analysis adjusting for any chance imbalances in the treatment groups. Average treatment effect estimated as change scores with 95% confidence intervals (mean change in outcome from baseline) for continuous outcomes and as hazard ratios with 95% confidence intervals for time-to-event outcomes.	Applied 2SRI model. In the first stage we estimated propensity score models to estimate probabilities that each person was prescribed each treatment based on their baseline covariates and their clinical commissioning group's tendency to prescribe that treatment. The second stage outcome models included the generalised residuals from the first stage models in an ordinary least squares (OLS) regression model (continuous outcomes) or Cox proportional hazards model (time-to-event outcomes) with an individual frailty. Models in both stages will include all measured baseline covariates, with additional polynomials and covariate interactions selected by Least Absolute Shrinkage and Selection Operator (LASSO) regression. The rationale for including contextual variables in the outcome regression model was to enable the IV approach to make more plausible assumptions (see text and supplement). Alternative analyses included multivariable regression analysis, adjusting for measured confounders. Post-hoc, we conducted an inverse probability of treatment (IPTW) analysis, and an IPTW-weighted regression (doubly robust), as a further alternative analysis. Average treatment effects were reported as change scores (mean change in outcome from baseline) or hazard ratios with 95% confidence intervals.

As a triangulation exercise to support the causal inferences made in this work, I took two main approaches to consider the impact of potential biases introduced by the primary IV analysis which impact the conclusions of this work.

First, I made internal comparisons of the primary IV analysis with alternative analyses that used the 2SLS IV model (Appendix E.9.), traditional multivariable OLS and Cox PH regression (Appendices E.8. and E.10.), and the propensity score IPTW-RA (Appendix E.11.). These alternative analyses were largely consistent with the main analysis.

Second, I triangulated the primary IV analysis results to those from key published RCTs relevant to this study which covered some of the comparisons and endpoints of interest. These key RCTs focused on evaluating the cardiovascular safety of SGLT2i but also include studies which directly compared DPP4i and SU (**Tables 5.3.** and **5.4.**). There was strong evidence that SGLT2i versus placebo were efficacious at reducing the risk of a variety of adverse cardiovascular outcomes, particularly heart failure hospitalisation,[7-12] as well as preventing adverse kidney outcomes[12-15] in these trials. However, as I illustrate in **Table 5.3.**, most of these studies were conducted in select patient populations and were placebo-controlled, with some exceptions.[16-19] Although none of these exceptions compared all three second-line antidiabetic treatments investigated in this thesis and most commonly used in the UK[1,20] and globally.[21]

I demonstrate similar findings to these placebo controlled RCTs, although the benefits of SGLT2i for CVD and kidney endpoints were more modest in my observational study. I hypothesise that these differences could be due to (i) chance, since the IV analysis only uses the exogenous variation in treatment assignment which increases the standard errors of the estimated treatment effects, (ii) measurement error, since we may not have perfectly measured the outcomes investigated in this study, (iii) violations in the IV assumptions which may have biased our results, or (iv) a true diminution of the benefits of SGLT2i when comparing against an active comparator versus placebo. I will expand on these potential limitations in the thesis discussion (section 10.4.).

Table 5.3. Main features of the study populations and comparison groups for the PERMIT study and relevant randomised controlled trials (RCTs) that included a randomisation to either SGLT2i or DPP4i (Supplementary table 21 in the Supplementary materials of the accepted manuscript)

Study	**Year**	**Key study eligibility criteria**	**'Active' treatment**	**'Comparator'**	**Antidiabetic treatment prior to randomisation, % of study population[1]**				
					Metformin	**SU**	**DPP4i**	**Insulin**	**GLP1-RA**
PERMIT[22]	2023	T2DM, general 2nd line initiators, eGFR>30mL/min/1.73m^2, no antidiabetic treatment prior to randomisation except metformin	SGLT2i or DPP4i	DPP4i or SU	100	0	0	0	0
EMPA-REG[7]	2015	T2DM, established CVD, HbA1c 7-9%, eGFR>30mL/min/1.73m^2, BMI<45kg/m^2	SGLT2i (empagliflozin)	Placebo	74	42	11	48	3
CANVAS-R[8]	2017	T2DM, at high-CVD risk, HbA1c 7.0-10.5%	SGLT2i (canagliflozin)	Placebo	77	43	12	50	4
DECLARE-TIMI 58[9]	2019	T2DM, at high-CVD risk, HbA1c 6.5-12, creatinine clearance ≥60mL/min	SGLT2i (dapagliflozin)	Placebo	82	43	17	41	4
CAROLINA[18]	2019	T2DM, at high-CVD risk, HbA1c 6.5-8.5%	DPP4i (linagliptin)	SU (glimepiride)	84	29	Unknown	Unknown	Unknown
ERTUGLIFLOZIN CVOT[10]	2020	T2DM, established CVD, HbA1c 7.0-10.5%, eGFR>30mL/min/1.73m^2	SGLT2i (ertugliflozin)	Placebo	77	41	11	48	3
CREDENCE[13]	2019	T2DM, CKD, HbA1c 6.5-12.0%	SGLT2i (canagliflozin)	Placebo	58	29	17	66	4
EMPA-Kidney[15]	2023	People with (96.6%) or without (3.4%) T2DM, eGFR 45-90mL/min/1.73m^2	SGLT2i (empagliflozin)	Placebo	10	9	13	25	5
GRADE[16,17]	2023	T2DM, excluded if major CVD in past year, HbA1c 6.8-8.5%, treated with MET alone	DPP4i (sitagliptin)	SU (glimepiride)	100	0	0	0	0
DAPA-HF[11]	2019	People with (41.8%) or without (58.2%) T2DM, with heart failure, eGFR>30mL/min/1.73m^2	SGLT2i (dapagliflozin)	Placebo	51	23	16	27	1

Study	Year	Key study eligibility criteria	'Active' treatment	'Comparator'	Antidiabetic treatment prior to randomisation, % of study population[1]				
					Metformin	SU	DPP4i	Insulin	GLP1-RA
DAPA-CKD[14]	2020	People with (67%) or without (33%) T2DM, eGFR 25-75mL/min/1.73m^2, ACEI/ARB	SGLT2i (dapagliflozin)	Placebo	Unknown	Unknown	Unknown	Unknown	Unknown
EMPEROR-Reduced[12]	2020	People with (49.8%) or without (50.2%) T2DM, heart failure, BMI<45kg/m^2	SGLT2i (empagliflozin)	Placebo	Unknown	Unknown	Unknown	Unknown	Unknown

ACEI/ARB: angiotensin converting enzyme inhibitor/angiotensin receptor blocker: BMI: body-mass index; CKD: chronic kidney disease; CVD: cardiovascular disease; DPP4i: dipeptidyl peptidase-4 inhibitor; eGFR: estimated glomerular filtration rate; GLP1-RA: glucagon-like peptide-1 receptor agonist; HbA1c: haemoglobin A1c; MACE-3: 3-point major adverse cardiovascular event; MAKE: major adverse kidney event; SGLT2i: sodium-glucose co-transporter 2 inhibitor; SU: sulfonylurea; T2DM: type 2 diabetes mellitus

[1]Note that for some studies people received more than one antidiabetic treatment prior to randomisation. For these studies the sum of the percentages can exceed 100%.

Table 5.4. Results from main analysis of the PERMIT study and those from relevant RCTs for common endpoints (Supplementary table 22 in the Supplementary materials of the accepted manuscript)

			Outcome, All reported with hazard ratios (HR) (95% confidence intervals (CI))				
Study	**Year**	**Simplified summary of study population**	**Heart failure hospitalisation**	**MACE-3**	**MAKE**	**40% decline in eGFR**	**All-cause mortality**
Studies comparing DPP4i versus SU							
PERMIT (DPP4i v SU)	2023	T2DM	1.41 (0.73, 2.71)	1.09 (0.70, 1.69)	0.72 (0.50, 1.03)	0.66 (0.37, 1.17)	0.82 (0.51, 1.32)
CAROLINA[18] (DPP4i v SU)	2019	T2DM, CVD	1.21 (0.92, 1.59)	0.98 (0.84, 1.14)	N/A	N/A	0.91 (0.78, 1.06)
GRADE[17] (DPP4i v SU)	2023	T2DM	0.99 (0.60, 1.64)	1.16 (0.82, 1.64)	0.93 (0.75, 1.18)	N/A	0.93 (0.61, 1.45)
Studies comparing SGLT2i versus placebo or active comparator							
PERMIT (SGLT2i v SU)	2023	T2DM	0.46 (0.20, 1.05)	0.99 (0.61, 1.62)	0.79 (0.51, 1.23)	0.42 (0.22, 0.81)	1.14 (0.64, 2.03)
PERMIT (SGLT2i v DPP4i)	2023	T2DM	0.32 (0.12, 0.85)	0.91 (0.51, 1.63)	1.11 (0.66, 1.84)	0.64 (0.29, 0.81)	1.39 (0.71, 2.74)
EMPA-REG[7] (SGLT2i v placebo)	2015	T2DM, CVD	0.65 (0.50, 0.85)	0.86 (0.74, 0.99)	N/A	N/A	0.68 (0.57, 0.82)
CANVAS-R[8] (SGLT2i v placebo)	2017	T2DM, CVD	0.67 (0.52, 0.87)	0.86 (0.75, 0.97)	0.60 (0.47, 0.77)	N/A	0.87 (0.74, 1.01)
DECLARE-TIMI 58[9] (SGLT2i v placebo)	2019	T2DM, CVD	0.76 (0.61, 0.88)	0.93 (0.84, 1.03)	0.53 (0.43, 0.66)	N/A	0.93 (0.82, 1.04)
ERTUGLIFLOZIN CVOT[10] (SGLT2i v placebo)	2020	T2DM, CVD	0.70 (0.54, 0.90)	0.97 (0.85, 1.11)	0.81 (0.63, 1.04)	N/A	0.93 (0.80, 1.08)
CREDENCE[13] (SGLT2i v placebo)	2019	T2DM, CKD	0.61 (0.47, 0.80)	0.80 (0.67, 0.95)	0.70 (0.59, 0.82)	0.60 (0.48, 0.76)	0.83 (0.68, 1.02)
EMPA-Kidney[15] (SGLT2i v placebo)	2023	T2DM, CKD	0.84 (0.67, 1.07)	N/A	0.72 (0.64, 0.82)	0.70 (0.61, 0.81)	0.87 (0.67, 1.07)
DAPA-HF[11] (SGLT2i v placebo)	2019	CVD	0.70 (0.59, 0.83)	N/A	0.71 (0.44, 1.16)	N/A	0.83 (0.71, 0.97)
DAPA-CKD[14]	2020	CKD	0.71 (0.55, 0.92)	N/A	0.61 (0.51, 0.72)	0.53 (0.42, 0.67)	0.69 (0.53, 0.88)

			Outcome, All reported with hazard ratios (HR) (95% confidence intervals (CI))				
Study	**Year**	**Simplified summary of study population**	**Heart failure hospitalisation**	**MACE-3**	**MAKE**	**40% decline in eGFR**	**All-cause mortality**
(SGLT2i v placebo)							
EMPEROR-Reduced[12] (SGLT2i v placebo)	2020	CVD	0.69 (0.59, 0.81)	N/A	0.50 (0.32, 0.77)	N/A	0.92 (0.77, 1.10)

CKD: chronic kidney disease; CVD: cardiovascular disease; DPP4i: dipeptidyl peptidase-4 inhibitor; eGFR: estimated glomerular filtration rate; MACE-3: 3-point major adverse cardiovascular event; MAKE: major adverse kidney event; N/A: not applicable (in this case, the study did not include this outcome); SGLT2i: sodium-glucose co-transporter 2 inhibitor; SU: sulfonylurea; T2DM: type 2 diabetes mellitus

In this analysis, I also stratified the results by prevalent CVD status at baseline to investigate treatment heterogeneity by this important subgroup. I did not find evidence of any differences in the treatment effect, although this could have been due to the lack of statistical power for this subgroup analysis.

As I describe in the methods overview of this thesis (section 2.4.4.), the 2SRI IV analysis used in this research paper can only account for overt heterogeneity. It is possible that the ATEs estimated using the 2SRI IV model are biased by essential heterogeneity. In the next chapter, I take a different IV approach to account for essential heterogeneity in a two-way treatment comparison (DPP4i versus SU) again using the TTE framework.

5.3. References

1. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

2. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

3. Wilkinson S, Williamson E, Pokrajac A, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. *Diabetes, Obesity and Metabolism*. 2020;22(5):847-856. doi:https://doi.org/10.1111/dom.13970

4. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *The international journal of biostatistics*. 2007;3(1):Article 14. doi:10.2202/1557-4679.1072

5. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

6. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

7. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *New England Journal of Medicine*. 2015/11/26 2015;373(22):2117-2128. doi:10.1056/NEJMoa1504720

8. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *New England Journal of Medicine*. 2017/08/17 2017;377(7):644-657. doi:10.1056/NEJMoa1611925

9. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine*. 2019/01/24 2018;380(4):347-357. doi:10.1056/NEJMoa1812389

10. Cannon CP, Pratley R, Dagogo-Jack S, et al. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. *New England Journal of Medicine*. 2020;383(15):1425-1435. doi:10.1056/NEJMoa2004967

11. McMurray JJV, Solomon SD, Inzucchi SE, et al. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. *New England Journal of Medicine*. 2019/11/21 2019;381(21):1995-2008. doi:10.1056/NEJMoa1911303

12. Packer M, Anker SD, Butler J, et al. Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure. *New England Journal of Medicine*. 2020;383(15):1413-1424. doi:10.1056/NEJMoa2022190

13. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *New England Journal of Medicine*. 2019/06/13 2019;380(24):2295-2306. doi:10.1056/NEJMoa1811744

14. Heerspink HJL, Stefánsson BV, Correa-Rotter R, et al. Dapagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2020;383(15):1436-1446. doi:10.1056/NEJMoa2024816

15. Empagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2022;388(2):117-127. doi:10.1056/NEJMoa2204233

16. Glycemia Reduction in Type 2 Diabetes — Glycemic Outcomes. *New England Journal of Medicine*. 2022/09/22 2022;387(12):1063-1074. doi:10.1056/NEJMoa2200433

17. Glycemia Reduction in Type 2 Diabetes — Microvascular and Cardiovascular Outcomes. *New England Journal of Medicine*. 2022/09/22 2022;387(12):1075-1088. doi:10.1056/NEJMoa2200436

18. Rosenstock J, Kahn SE, Johansen OE, et al. Effect of Linagliptin vs Glimepiride on Major Adverse Cardiovascular Outcomes in Patients With Type 2 Diabetes: The CAROLINA Randomized Clinical Trial. *Jama*. Sep 19 2019;322(12):1155-66. doi:10.1001/jama.2019.13772

19. Nauck MA, Meininger G, Sheng D, Terranella L, Stein PP. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes, obesity & metabolism*. Mar 2007;9(2):194-205. doi:10.1111/j.1463-1326.2006.00704.x

20. NICE guideline [NG28]: Type 2 diabetes in adults: management. Web. NICE. Accessed 10 June, 2024. https://www.nice.org.uk/guidance/ng28/chapter/Recommendations#reviewing-drug-treatments

21. Khunti K, Kosiborod M, Kim DJ, et al. Cardiovascular outcomes with sodium-glucose cotransporter-2 inhibitors vs other glucose-lowering drugs in 13 countries across three continents: analysis of CVD-REAL data. *Cardiovasc Diabetol*. Jul 31 2021;20(1):159. doi:10.1186/s12933-021-01345-z

22. Bidulka P, O'Neill S, Basu A, et al. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open*. 2021;11(9):e046912. doi:10.1136/bmjopen-2020-046912

CHAPTER 6. RESEARCH PAPER – GOING BEYOND RANDOMISED CONTROLLED TRIALS TO ASSESS TREATMENT EFFECT HETEROGENEITY ACROSS TARGET POPULATIONS.

OVERVIEW

In this chapter, I include a submitted research paper for which I am second author. This manuscript details a target trial emulation combined with an IV analysis (LIV) to compare DPP4i and SU as second-line oral antidiabetic treatments. This manuscript explores treatment heterogeneity according to observed variables (overt heterogeneity) and unobserved variables (essential heterogeneity). As part of this examination of heterogeneity I emulated the published RCT by Nauck et al, 2007 in the 'trial eligible' subpopulation of the observational cohort. I then examine heterogeneity across the 'trial ineligible' subpopulation of the observational cohort and consider the implications more generally for transporting results from trial eligible to trial ineligible populations. I explore essential heterogeneity, as well as overt heterogeneity according to the estimated change in HbA1c at 1-year follow-up. I pre-specify subgroups to consider whether there are differences in heterogeneity between the 'trial eligible' and 'trial ineligible' subpopulations. I offer conclusions on how LIV methods can be useful to investigate heterogenous treatment effects in populations excluded from RCTs.

Following the research paper, I include a brief discussion of this research paper's relevance to my thesis. This discussion includes key tables and figures from the supplementary materials of the published paper which aid in the interpretation of the study in the context of this thesis.

6.1. Submitted research paper

The main tables and figures are provided immediately following the references for this submitted manuscript. I added the prefix '6'.X for each main figure and table to indicate this is the 6th chapter of this thesis.

Select supplementary materials important to this thesis are provided in the Appendix F, referenced in section 6.2.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Health Economics
Please list the paper's authors in the intended authorship order:	David G Lugo-Palacios, Patrick Bidulka, Stephen O'Neill, Orlagh Carroll, Anirban Basu, Amanda I Adler, Karla Diaz-Ordaz, Andrew H Briggs, Richard Grieve

Stage of publication	**Submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I am second author of this paper. I supported DLP (first author) in the study design. I led the literature review to select the trial to use for the target trial emulation. I led the data management, including the extraction and creation of the study dataset from the raw CPRD files and applying the inclusion/exclusion criteria of the trial to the cohort. I supported DLP in the analyses and interpretation of the study results. DLP wrote the first draft of the manuscript, which I edited before sharing with co-authors for comments. I will contribute to the peer review responses.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Richard Grieve
Date	07 April 2024

Going beyond Randomised Controlled Trials to assess treatment effect heterogeneity across target populations

Abstract

Methods have been developed for transporting evidence from Randomised Controlled Trials (RCTs) to target populations. However, these approaches allow only for differences in characteristics observed in the RCT and real-world data (overt heterogeneity). These approaches do not recognise heterogeneity of treatment effects (HTE) according to unmeasured characteristics (essential heterogeneity).

We use a target trial design and apply a local instrumental variable (LIV) approach to electronic health records (EHR) and examine both forms of heterogeneity in assessing the comparative effectiveness of two second-line treatments for type 2 diabetes mellitus. We first estimate individualised estimates of HTE across the entire target population defined by applying eligibility criteria from national guidelines (n=13,240) within an overall target trial framework. We define a subpopulation who meet a published RCT's eligibility criteria ('RCT-eligible', n=6,497), and a subpopulation who do not ('RCT-ineligible', n=6,743). We compare average treatment effects for pre-specified subgroups within the RCT-eligible subpopulation, the RCT-ineligible subpopulation, and within the overall target population. We find differences across these subpopulations in the magnitude of subgroup-level treatment effects, but that the direction of estimated effects is stable. Our results highlight that LIV methods can provide useful evidence about treatment effect heterogeneity including for those subpopulations excluded from RCTs.

Keywords: Transportability, heterogeneous treatment effects, Electronic Health Records, Target Trial Emulation, Instrumental Variables, Diabetes Mellitus

Word count: 6,924

Title: Going beyond Randomised Controlled Trials to assess treatment effect heterogeneity across target populations

Short title: Assessing heterogeneity beyond RCTs

Authors: David G. Lugo-Palacios*, Patrick Bidulka, Stephen O'Neill, Orlagh Carroll, Anirban Basu, Amanda Adler, Karla DíazOrdaz, Andrew Briggs, Richard Grieve

Order	Affiliation	Author name	Email
1	1	David G Lugo-Palacios*	David.Lugo-Palacios@lshtm.ac.uk
2	2	Patrick Bidulka	Patrick.bidulka1@lshtm.ac.uk
3	1	Stephen O'Neill	Stephen.ONeill@lshtm.ac.uk
4	1	Orlagh Carroll	Orlagh.Carroll@lshtm.ac.uk
5	3	Anirban Basu	basua@uw.edu
6	4	Amanda I Adler	amanda.adler@dtu.ox.ac.uk
7	5	Karla Díaz-Ordaz	karla.diaz-ordaz@ucl.ac.uk
8	1	Andrew H Briggs	Andrew.Briggs@lshtm.ac.uk
9	1	Richard Grieve	Richard.Grieve@lshtm.ac.uk

***** Corresponding author

Affiliations

(1) Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London UK WC1H 9SH
(2) Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London UK WC1E 7HT
(3) The Comparative Health Outcomes, Policy & Economics (CHOICE) Institute, University of Washington School of Pharmacy, Seattle USA 98195
(4) Diabetes Trials Unit, The Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, OCDEM Building Churchill Hospital, Old Road, Headington UK OX3 7LJ
(5) Department of Statistical Science, University College London, Gower Street, London UK WC1E 6BT

Funding: This work was funded by the National Institute for Health and Care Research (NIHR) grant number NIHR128490. The funder had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Acknowledgements: This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency (MHRA). The data are provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the authors alone. We gratefully acknowledge the valuable input from Silvia Moler-Zapata and the other PERMIT co-applicants (Ian Douglas, Richard Silverwood, Paul Charlton, Liam Smeeth and Kamlesh Khunti), as well as the support of the PERMIT study advisory team (Ken Paterson, Jim Lewsey, Ewan Pearson, Sarah Finer, Stephen Evans, Bill Huston, and Rahul Mohan), and the management and administrative support provided by Elizabeth Silver, Anna Carnegie and Paula Fry. We also thank David McAllister and Elaine Butterly for sharing early findings from their systematic review on RCTs assessing the comparative effectiveness of alternative second-line treatments.

Conflict of interest statement: PB, OC, SON, AIA, and KDO have no disclosures. DLP served as external consultant for Blutitude Consultores on a project not related to diabetes. AB reports Personal Fees from Salutis Consulting LLC outside the scope of the submitted work, AHB reports grants from NIHR during the conduct of the study and personal fees outside the submitted work from Boehringer Ingelheim, Novartis, Takeda, Idorsia, Rhythm, Daiichi-Sankyo, Bayer, GSK, and Gilead. RG reports grants from NIHR during the conduct of the study.

Introduction

Health policy-makers and reimbursement agencies require evidence of comparative effectiveness for target populations and subpopulations relevant to the decision context. Randomised Controlled Trials (RCTs) are the recommended source of evidence for estimating treatment effects **(NICE, 2013, ICER, 2020)**. However, the selection of trial centres and participants leads to differences in the characteristics of trial populations versus those eligible for the same interventions in routine practice, i.e. the 'target populations'. RCT eligibility criteria, especially for trials designed to assess safety and efficacy rather than comparative effectiveness, may exclude subpopulations of interest to national and local decision-makers, with these exclusions partly captured by observed measures **(Dahabreh and Hernán, 2019, Elliott et al., 2023, Stuart et al., 2010, Hartman et al., 2015, Gheorghe et al., 2013)**. Statistical methods have been developed to 'transport' estimates of average treatment effects (ATE) or conditional average treatment effects (CATEs) from RCTs to target populations and subpopulations defined by routine data **(Allcott and Mullainathan, 2012, Dahabreh and Hernán, 2019, Elliott et al., 2023, Stuart et al., 2010, Hartman et al., 2015, Degtiar and Rose, 2023)**. However, these methods only account for differences between settings in observed characteristics (overt heterogeneity). RCT participation may also reflect characteristics that are *unmeasured* in both the RCT or routine data, such as frailty levels, attitude to risk, preferences, or contextual factors such as quality of care. If unmeasured characteristics associated with RCT participation also modify the relative effectiveness of the treatment alternatives, i.e. there is essential heterogeneity, then this leads to *study selection bias and* related concerns about the transportability of the results to the decision context **(Hartman et al., 2015, Heckman et al., 2006, Robertson et al., 2023).**

Non-randomised studies (NRS) can provide complementary evidence on comparative effectiveness in subpopulations excluded from RCTs and may be less prone to sample selection bias **(Imai et al., 2008)**. The improved quality and availability of data from electronic health records (EHRs) offer new opportunities for NRS to provide estimates of treatment effects including for subgroups who do not meet explicit measured RCT eligibility criteria. A further advantage is that an NRS may include the treatment comparators of direct interest to the decision-context and incorporate treatment protocols and adherence levels from routine practice rather than those driven by the RCT design. However, the major challenge when using NRS to assess comparative effectiveness is that the treatments groups of interest are selected according to measured and unmeasured prognostic measures which leads to *treatment selection bias* (confounding). A recommended NRS design to reduce the risk of confounding and other sources of bias is 'target trial emulation' **(Hernán and Robins, 2016, Hernán et al., 2016, Wang et al., 2023, NICE, 2021, Gomes et al., 2022, Moler-Zapata et al., 2023b)**. The target trial framework requires the analyst to conceptualise the NRS as if it were an RCT, and make explicit decisions at the design stage to reduce the risk of bias, for example by defining the study populations, treatment regimes, and analytical methods. The target trial

framework has helped some NRS replicate treatment effect estimates from RCTs **(Hernán and Robins, 2016, Wang et al., 2023)**, and the recent update to the NICE methods guide advocates its use when RCT evidence on comparative effectiveness is unavailable or is judged insufficient **(NICE, 2021)**. However, if the target trial applies the same eligibility criteria as an RCT, it may face some of the same concerns about sample selection bias, especially as the stated eligibility criteria (e.g. age) are likely to be correlated with measures (e.g. frailty levels) unobserved in the data used by the target trial. A promising approach is to design a target trial to assess HTE across the full target populations and subpopulations of decision-making relevance, including subpopulations who would have been ineligible for the published RCT(s). For such a target trial design to provide evidence on comparative effectiveness of direct relevance to decision-makers, it is essential to reduce the risk of unobserved confounding, but also of essential heterogeneity across the subpopulations of interest.

Instrumental Variable (IV) methods, such as two-stage least squares (2SLS), can reduce the risk of unobserved as well as observed confounding. However in settings with essential heterogeneity, 2SLS no longer identifies policy-relevant estimands, such as the ATE, even if the instrument is strong and valid **(Heckman et al., 2006)**. Local Instrumental Variable (LIV) approaches can provide consistent estimates of the ATE and CATEs **(Heckman and Vytlacil, 2005, Cornelissen et al., 2016).** LIV methods draw on choice theory to identify 'marginal treatment effects' (MTEs) for those at the 'margin of treatment choice' for whom the level of a continuous IV balances observed and unobserved characteristics **(Björklund and Moffitt, 1987, Heckman and Vytlacil, 1999)**. For these patients at the 'marginal choice' a small change in the level of the valid, continuous IV changes the treatment decision, without altering the distribution of the underlying risk factors. Therefore, we can identify MTEs for individuals who *comply* with the change in treatment that is due to a small change in the level of the IV, by comparing mean outcomes between two groups of similar patients who are only separated by a small change in the IV **(Heckman and Vytlacil, 1999)**. Hence, given observed covariates, and common support, MTEs can be estimated along the continuum of the IV, and aggregated to provide CATEs and ATEs **(Heckman and Vytlacil, 2001, Heckman and Vytlacil, 2005, Heckman and Vytlacil, 1999)**.

The theoretical properties of these LIV methods in settings with essential heterogeneity have been established by **Basu et al., 2007, Heckman et al., 2006, Angrist and Fernández-Val, 2013** *inter alia.* A recent simulation study showed that given an IV of sufficient strength, an LIV method can provide consistent estimates of treatment effects in settings with overt and essential heterogeneity **(Moler-Zapata et al., 2023a)**. LIV methods have been applied across a range of settings including cardiovascular and bariatric surgery, universal child care programs and transfers to intensive care units **(Basu et al., 2018, Cornelissen et al., 2018, Grieve et al., 2019, Reynolds et al., 2021)**.

In this paper, we use a LIV method to examine HTE across an entire target population. We recognise that there may be essential heterogeneity for subpopulations in the target population who were ineligible for the RCT. We illustrate our approach for addressing this challenge within the running example of the PERsonalised Medicine for Intensification of Treatment (PERMIT) study which evaluates alternative second-line drug treatments for people with type 2 diabetes mellitus (T2DM) treated with metformin **(Bidulka et al., 2021)**. Here, we compare two second-line treatments, dipeptidyl peptidase-4 inhibitors (DPP4i) and sulfonylureas (SU) as 'add on' therapies to metformin (for details see **(Bidulka et al., 2021)**). The primary endpoint is improvement in glycaemic control, measured by the change in haemoglobin A1c (HbA_{1c}) between baseline and 1-year follow-up. Published RCTs report divergent results, with some reporting that SUs lead to greater improvements in HbA_{1c} than DPP4i, and others that DPP4i lead to greater improvements **(Marathur et al., 2016, GRADE Study Research Group, 2022, Rosenstock et al., 2019)**. A general problem is that the RCTs have explicitly excluded people with poor glycaemic control who are likely to differ in their responses to the treatments compared to people with better glycaemic control at baseline. Consequently, treatment recommendations based on RCT evidence may be suboptimal for subpopulations excluded from these studies. Faced with insufficient evidence from RCTs including the full target population of interest, NICE clinical guidelines have recommended allowing for individual's risk factors and circumstances when choosing second-line treatments between options that include DPP4is and SUs, but there is little evidence on HTE to help decision-makers transport the findings from RCTs to target populations of interest. Hence, the PERMIT study exemplifies the common situation where decision-makers require further evidence on HTE including subgroups excluded from an RCT.

The paper proceeds as follows. In Section 2, we offer an overview of the main features of the PERMIT study. In Section 3, we outline the LIV approach for handling confounding and examining heterogeneity. In section 4, we define target trial protocols for the overall target population of interest, and subpopulations who do and do not meet 'RCT eligibility criteria'. In Section 5, we report the empirical results. In Section 6, we discuss the results in the context of the extant literature for examining HTE when transporting results from RCTs to the routine practice setting, and outline future research directions.

2. Overview of the PERMIT study

The PERMIT study aims to assess the comparative effectiveness of alternative second-line pharmacological treatments for people with T2DM in England who meet national eligibility criteria for these treatments, and to provide evidence to help personalise the choice of second-line treatment according to individual-level characteristics. In this paper, we evaluate DPP4i versus SU which were the most prevalent classes of second-line treatments for T2DM in the UK between 2011 and 2015 **(Wilkinson et al., 2018)**. We used

Clinical Practice Research Datalink (CPRD) data from England on clinical and demographic characteristics, clinical diagnoses, laboratory test results, prescribing, and outcome information recorded in primary care, with further information on resource use and outcomes from linkage to hospital episodes statistics (HES) data **(Herbert et al., 2017, Herrett et al., 2015, Wolf et al., 2019).** Full details of the PERMIT study have been published elsewhere **(Bidulka et al., 2024, Bidulka et al., 2023, Bidulka et al., 2021)**. In brief, we undertook a target trial emulation to define the target populations and treatment comparisons of interest from the CPRD data. The study addressed the potential concerns about confounding with a continuous preference-based instrumental variable (IV) **(Baiocchi et al., 2014)**. The IV was the clinical commissioning groups (CCG)'s tendency to prescribe (TTP) the alternative classes of second-line treatment. Over the study time-frame CCGs informed health funding decisions about whether DPP4is or SUs were prescribed within their respective geographic region **(Bidulka et al., 2023, Wilkinson et al., 2018).**

We exploited the wide variation across CCGs in the proportion of people prescribed DPP4is versus SUs (see supplement **Figure 6.1.**)**.** This 'natural variation' implied that people of similar baseline prognosis received a different second-line treatment simply according to their CCG. We defined TTP as the proportion of eligible people prescribed each second-line treatment within the 12 months preceding the specific baseline (time zero) for each person. A valid instrument must meet four main conditions (see also the Direct Acyclic Graph (DAGs) in **Figure S1** of the online supporting material) **(Baiocchi et al., 2014)**. First, the instrument must predict the treatment prescribed **(Staiger and Stock, 1997).** Here, we assessed the relevance of the CCGs TTP with a weak instrument test that is robust to heteroscedasticity and clustering by NHS region, and reported F-statistics of around 1,115, compared to a benchmark of 100 **(Baiocchi et al., 2014, Moler-Zapata et al., 2023a).** Second, the instrument must be independent of unmeasured covariates that predict the outcomes of interest, which can be partially evaluated through its relationship with measured covariates. We found that levels of observed prognostic covariates were similar across levels of the instrument (see online supporting material). Third, the instrument must have an effect on the outcomes only through the treatment received. We adjusted for contextual and temporal confounders, and made the weaker assumption, that the TTP was independent of the outcome, and only had an effect on the outcome via the treatment received, *after* adjusting for any differences in region, GP practice size, and time period (see online supplement). Fourth, we assume that the average treatment choice must increase or decrease monotonically with the level of the IV **(Vytlacil, 2002)**. Here, it is plausible to assume that if a group of patients who attended a CCG with a moderate preference for prescribing DPP4is were prescribed this drug class, then a similar group of patients who attended a CCG with a stronger preference for prescribing DPP4is would also be prescribed this drug class.

The PERMIT study previously used the 2-stage residual inclusion (2SRI) to report ATEs for the overall target population of interest **(Bidulka et al., 2024)**. However, an outstanding concern is to explore HTE, that may arise according to baseline risk factors that are observed (e.g. HbA_{1C}) as well as those that are not readily observed (e.g. patient preferences). We now formally state the IV assumptions and define the LIV approach for estimating policy relevant estimands of interest, namely ATE and CATEs.

3. Methods

3.1. Instrumental variables methods

Following the Neyman-Rubin potential outcomes framework **(Rubin, 1974, Neyman, 1990)**, let Y_D denote the observed outcome, D_Z the choice of treatment, Z the IV observed for each individual (Y_D, D_Z, Z). Let $Y_1 = \mu_1(X_O, X_U, \vartheta)$ and $Y_0 = \mu_0(X_O, X_U, \vartheta)$ denote the individual's potential outcomes under each treatment, where X_O and X_U are vectors of measured and unmeasured confounders, and ϑ captures the remaining unobserved random variables. We assume exogeneity of the covariates (A1), so that the treatment assignment is the only source of endogeneity, such that $(X_O, X_U) \perp \vartheta$ and $X_O \perp X_U$.

3.1.1. Identification assumptions

We follow Abadie (2003) and Tan (2006) in making the following assumptions which are the conditional version of the assumptions outlined by Angrist et al., (1993) for the local average treatment effect (LATE):

(A2)	Unconfoundedness of Z	$(Y_{d_z}, D_z) \perp Z \mid X_O$
(A3)	Exclusion restriction	$Y_{d_z} = Y_d$ with probability 1
(A4)	Relevance	$0 < P(Z = z) < 1$
(A5)	Monotonicity	If $z' > z$ then $D_{z'} \geq D_z$ with probability 1
(A6)	Stable Unit Treatment Value Assumption	$D = D_Z$ and $Y = Y_D$

Assumption (A2) requires that within levels of X_O Z is as good as randomly assigned. Assumption (A3) rules out that Z has a direct effect on the outcome other than through D_z. Assumptions (A2) and (A3) ensure that the only effect of the Z on the outcome is through D_z. Assumption (A4) ensures that Z and D_z are correlated conditional on X_O. Assumption (A5) requires that an increase in Z always results in a higher or equal level of treatment assignment, and this is needed to point-identify our estimand of choice. Assumption (A6) requires that one individual's potential outcomes (Y_D) and treatments (D_z) are not influenced by other individuals' levels of Z (i.e., no interference), nor by how the instrument or treatment is delivered (i.e., no different versions of Z or D_z).

3.1.2. Estimands

(Angrist et al., 1996, Imbens and Angrist, 1994) show that following the above assumptions, the LATE can be defined as $\Delta^{LATE}(x_o, z, z') = E[Y_1 - Y_0|X_O = x_o, D_z < D_{z'}]$ and is identified by the IV estimand:

$$\frac{E[Y|X_O = x_o,\ Z = z'] - E[Y|X_O = x_o,\ Z = z]}{E[D|X_O = x_o,\ Z = z'] - E[D|X_O = x_o,\ Z = z]}$$

(Vytlacil, 2002, Tan, 2006) showed that the independence (A2 and A3) and monotonicity assumptions (A5) within the LATE framework are equivalent to those imposed by a non-parametric selection model, where treatment assignment depends on whether a latent index ($\mu_D(X_O, Z)$) crosses a particular threshold (X_{U_D}):

$$D_z = 1\{\mu_D(X_O, Z) \geq X_{U_D}\}$$

where X_{U_D} is a random variable that captures X_U and all other factors influencing treatment assignment but not the outcomes. We follow **(Heckman and Vytlacil, 1999, Heckman and Vytlacil, 2001)**, in rewriting this equation as D_z=1$\{P(X_O, Z) > V\}$, where $V = F_{X_{U_D}}[X_{U_D}|X_O = x_O, Z = z]$ with $V \perp (Z, X_O)$ and $P(x_O, z) = F_{X_{U_D}|x_O,z}[\mu_D(X_O, Z)]$ is the propensity for treatment, and F represents a cumulative distribution function. Therefore, for any arbitrary distribution of X_{U_D} conditional on X_O and Z, by definition $V \sim Uniform[0,1]$ conditional on X_O and Z. Then, the MTE can be defined as, $\Delta^{MTE}(x_O, p) = E(Y_1 - Y_0|X_O = x_O, V = v)$ and **(Heckman and Vytlacil, 1999, Heckman and Vytlacil, 2001)** showed that, under the standard IV assumptions, it can be identified by:

$$\frac{\partial E_\vartheta(Y|X_O = x_o, Z = z)}{\partial p} = E_\vartheta[(Y_1 - Y_0)|X_O = x_o, V = v]$$

(Heckman et al., 2006) showed that MTEs can be aggregated to obtain estimates of the ATE. **(Basu, 2014)** showed that MTEs can be used to derive personalised treatment (PeT) effects for each individual that recognise the plausible range of values that V may take for each patient, compatible with the levels of their observed covariates, the IV and their observed treatment assignment (see next section). The crucial insight underlying this approach is that given the observed covariates and the level of the IV, the treatment assignment status observed provides some information on X_{U_D}. For patients in the treatment group ($D_z = 1$), the propensity to choose treatment based on X_O and Z must outweigh the propensity to choose the comparator strategy based on X_{U_D}, i.e. $P(x_O, z) > v$. For patients in the comparator strategy ($D_z = 0$), the converse is true. The PeT effect for an individual is therefore obtained by averaging the MTEs corresponding to that individual's level of X_O and Z over levels of unobserved variables that are compatible with the individual's assigned treatment. Hence, $\Delta^{PeT}(x_O, p, D) = E(Y_1 - Y_0|X_O = x_O, P(z, x_O) > v)$ for

individuals with $D_z = 1$ and $\Delta^{PeT}(x_O, p, D) = E(Y_1 - Y_0 | X_O = x_O, P(z, x_O) < v)$ for individuals with $D_z = 0$.

All required treatment effect estimands, including ATE and CATEs, can be derived by appropriately aggregating the PeT effects since these are defined at the individual level (see next section)

3.2 Local Instrumental Variable (LIV) estimator: estimating PeT effects

Basu, 2014, Basu, 2015 provides a detailed description for using LIV methods to estimate PeT effects. Briefly, the analyst must first estimate the propensity for treatment $p(x_O, z)$, based on observed covariates and the instrument. Next, the outcome Y is regressed on X_O and a function of $\hat{p}(x_O, z)$ including interactions with X_O. The approach outlined in **Basu, 2014** involves differentiating the outcome model $g(Y)$ by $\hat{p}(x_O, z)$ to obtain MTE estimates. Next, PeT effects for each individual can be obtained by performing numerical integration, with MTE $(\partial \hat{g}(Y)/\partial \hat{p})$ evaluated by replacing $\hat{p}$ using 1,000 random draws of $u \sim unif(\min(\hat{p}(x_O, z)), \max(\hat{p}(x_O, z)))$. Then, the corresponding treatment assignment at each value of u is given by $D^* = \Phi^{-1}\{\hat{p}(x_O, z)\} + \Phi^{-1}(1 - u)$. PeT effects can be computed by averaging $\partial \hat{g}(Y)/\partial \hat{p}$ over values of u consistent with the observed treatment decision ($D^* > 0$ if $D = 1$; or over values of $D^* \leq 0$ if $D = 0$). Finally, an estimate of the ATE for the population can be obtained by averaging PeT effects over all of the observations, and the CATE for the subgroups of interest by aggregating over the appropriate strata of X_O. Standard errors can be computed using nonparametric bootstrap methods **(Basu, 2015)**.

4. Target trial design

4.1 Overview

The aim of the target trial was to emulate a hypothetical RCT for estimating the ATEs and CATEs of second-line treatment with DPP4i versus SUs in routine clinical practice in England. The LIV approach was applied to estimate individual-level treatment effects which were aggregated across the entire target population to report ATEs, and for a small number of pre-specified subgroups to report CATEs. We identified subpopulations from within the target population who met a published RCT's eligibility criteria ('RCT eligible') and a subpopulation who did not ('RCT ineligible). We compared the ATE for the 'RCT eligible' to the RCT with the same eligibility criteria (the 'RCT benchmark'). We then compared CATEs for the overall target population, and the 'RCT eligible' and 'RCT ineligible' subpopulations. The CATEs were according to pre-specified subgroups for whom there were prior hypotheses of HTE for DPP4i vs SUs. These subgroups were: age group (younger than 50 years old, 50-59, 60-69. 70-78, 79 or over) **(Khunti et al., 2018, Nauck et al., 2007)**, ethnicity (white, Asian (South Asian), black/mixed/other) **(Gan et al., 2020)**, baseline levels of

HbA_{1c} measured in mmol/mol (<44; 44 to <64; 64 to <75; 75 to < 88; ≥ 88) **(Nauck et al., 2007, Canivell et al., 2019)** and BMI defined according to WHO categories **(Khunti et al., 2018)**.

4.2 Eligibility criteria for identifying the overall target population

We used the eligibility criteria stipulated in NICE clinical guidelines for T2DM to define the overall target population **(NICE, 2022)**. Individuals had to initiate second-line antidiabetic treatment with either SU or DPP4i in addition to metformin between 1st January 2011 and 31st December 2015 (for specific inclusion criteria see **Table S1** in the online supporting material). These eligibility criteria ensured the target population was of direct relevance to the decision problem, and that there was general equipoise in the choice of SU and DPP4i, as either second-line treatment was an option in the target population irrespective of their baseline characteristics.

4.3 Treatment strategies and 'day zero'

The treatment strategies were SU or DPP4i for second-line treatment with each drug class prescribed as an addition to metformin monotherapy. The definition of the comparators was broad to allow for any drug within either drug class, and included the specific drugs defined by the randomised groups in published RCTs **(GRADE Study Research Group, 2022, Rosenstock et al., 2019, Nauck et al., 2007)**. *Day zero* was analogous to the time of randomisation and was when an individual met all eligibility criteria, had a first prescription for either SUs or DPP4is, and therefore started follow-up.

4.4 Covariates

We pre-specified baseline covariates for consideration in the LIV and alternative analyses and to define pre-specified subgroups for estimating CATEs. From the CPRD-HES lined data we defined patient sociodemographic characteristics, for example, age, sex, ethnicity (see **Table S2** for full list and definitions). We also extracted information on HbA_{1c}, systolic blood pressure (SBP), diastolic blood pressure (DBP), estimated glomerular filtration rate (eGFR), and body mass index (BMI) using the most recent measures recorded in primary care. For HbA_{1c} we only considered the most recent measure within 180 days prior to baseline, and for SBP, DBP, and eGFR, we used the most recent measure within 540 days prior to baseline **(Wilkinson et al., 2018)**.

4.5 Outcome and causal contrast of interest

The primary outcome was the change in HbA_{1c} between day zero and week 52 reported in mmol/mol. We reported the difference in the change as ΔDPP4i - ΔSU, such that a negative difference meant that DPP4is were better (reduced HbA_{1c} more), and a positive difference that SUs were better. A clinically meaningful between-treatment difference in HbA_{1c} was defined as 3.3 mmol/mol (0.3%) **(European Medicines Agency, 2018)**. We used the HbA_{1c} measurement closest in time to the 1-year follow-up timepoint and allowed for measures within ±90 days, otherwise the measure was designated as missing (see Statistical Analysis section).

The causal contrast of interest was the 'intention-to-treat' (ITT) effect of DPP4i vs SU and so the full sample who met the eligibility criteria were considered in the analysis including those who switched to 'third-line' treatments, or did not adhere to their 'assigned' treatment within the 12-month follow-up (see also statistical analysis, missing data).

4.6 Selection of RCT benchmark

We reviewed the literature to identify published RCTs that evaluated second-line antidiabetic drugs as add-on treatments to metformin for people with T2DM. Our selection criteria for emulating and benchmarking purposes were that the RCT: had to randomise between SUs and DPP4i, and report essential details required for emulation, including the eligibility criteria, treatment strategies and primary outcome (see **supporting note S1**). We selected the RCT by **(Nauck et al., 2007)** as it was the only trial that met all the inclusion criteria. The Nauck et al trial was a multinational, parallel-group RCT with a non-inferiority design that compared the efficacy of DPP4is versus SUs in patients with T2DM and inadequate glycaemic control following metformin monotherapy. The primary analysis of this RCT assessed whether DPP4is were non-inferior versus SUs with regard to the change in HbA_{1c} from baseline to Week 52 according to a non-inferiority margin of 3.3 mmol/mol (0.3%). The RCT exemplified key challenges in the use of RCTs designed to evaluate efficacy and safety for decision-making, in that the study excluded patients aged over 78 years, and those with very poor glycaemic control which was defined by HbA_{1c} > 87 (10%). The study made a limited attempt to consider HTE, the only pre-specified subgroup analysis was according to baseline HbA_{1C} and the authors did not report CATEs with confidence intervals.

We applied the additional inclusion criteria stipulated by Nauck et al to the CPRD data to define a 'RCT eligible' subpopulation, which included people aged 18-78 years who had baseline HbA_{1c} of 44 - 87 mmol/mol, (6.5%-10%) at study entry (day zero). We defined an 'RCT ineligible' subpopulation as the remaining subsample from the target population who did not meet the inclusion criteria for the Nauck et al RCT.

4.7 Statistical analysis: estimating ATEs and CATEs for the overall target population and for the 'RCT-eligible' and 'RCT ineligible' subpopulations

We applied LIV to estimate the effects of DPP4i versus SU for each person in the entire target population. These individual effects were then aggregated to report ATE and CATEs for the overall target population, and for the 'RCT eligible' and 'RCT ineligible' subpopulations. Within the LIV estimation, the first stage models estimated the probability that each person was prescribed DDP4i given their baseline covariates and their CCG's TTP **(Basu et al., 2018**). The second-stage outcome models then included the predicted probabilities from the first-stage (propensity score) models, covariates and their interactions. Probit regression models were used to estimate the initial propensity score (first stage), while GLMs were applied to the outcome data, with the most appropriate family (gaussian) and link function (identity) chosen according to root mean squared error, with Hosmer-Lemeshow and Pregibon tests also used to check model fit and appropriateness **(Hosmer and Lemeshow, 2000, Pregibon, 1980)**. In addition to including main covariate effects in the models for both stages, we also considered the quadratic forms of both age and baseline HbA_{1c}. We also considered the interaction of baseline HbA_{1c} with age, sex, and baseline BMI. We considered interactions of the IV (first-stage models) or for the treatment indicator (second-stage models) with baseline HbA_{1c}, eGFR, BMI, SBP and age. To select the final set of interactions, we applied the rigorous Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm for the first and second stage models **(Frank and Friedman, 1993, Belloni et al., 2012, Tibshirani, 2018).** Following the approach detailed in **(Bidulka et al., 2024),** the final model specification included those variables selected by the rigorous LASSO in at least one stage.

For the *RCT-eligible population* we compared the ATE to the corresponding estimand from the RCT with the binary agreement metric proposed by **Franklin et al., 2021** which designates *estimate agreement* if the ATE from the target trial is within the 95% CI of the RCT estimate.

4.7.1 Missing Data

Some measurements were missing for the HbA_{1C} outcome (31.8%) and baseline covariates (e.g. ethnicity, baseline HbA_{1c}). In previous work we have showed that the estimated levels of HbA_{1C} were similar with a complete case analysis, which implies that the missingness mechanism is independent of the outcome given the covariates included in the analytical models, versus multiple imputation (MI) which assumes the data are missing at random (MAR) **(Bidulka et al., 2024)**. Hence, in this paper, we adopt the simpler approach of complete case analyses (See flow diagram in **Figure S2**).

4.7.2 Alternative analyses

We conducted alternative analyses that avoided assuming that the IV was valid, but assumed instead that there was no unobserved confounding (selection on observables) and no essential heterogeneity. Using the same covariates as described in Sections 4.4 and 4.7, we applied inverse probability of treatment weighting with regression adjustment (IPTW-RA) **(Wooldridge, 2007)**, which has the so-called 'double-robustness property' in that subject to the 'no unobserved confounding' assumption it can still provide consistent estimates provided *either* the propensity score or the outcome regression model is correctly specified **(Funk et al., 2011)**.

We followed the same principle as for the main analysis by using recycled predictions **(StataCorp, 2023, Basu and Rathouz, 2005)** to estimate individual level effects by predicting potential outcomes ($\hat{Y}^{DPP4i}, \hat{Y}^{SU}$) for each person with each treatment, and then calculated individual-level differences in these predictions $\widehat{\tau_i} = \hat{Y}^{DPP4i} - \hat{Y}^{SU}$, representing individualised treatment effect estimates. It is worth noting that IPTW-RA with recycled predictions is equivalent to g-computation regression adjustment incorporating inverse probability weighting **(Smith et al., 2022)**. We then aggregated these individualised treatment effect estimates to each pre-specified subgroup, and the overall populations and subpopulations of interest to obtain estimates of the CATEs and ATE.

We undertook all these alternative analyses on the same samples of complete cases as for the main analyses.

5. Results

We identified 13,240 people from 162 CCGs in the CPRD data who met the inclusion criteria for the target population. **Table 6.1.** compares the baseline characteristics for those who had DPP4is versus SUs. For baseline measures such as age, gender and ethnicity there were only small differences between the comparison groups, but for HbA_{1C} and BMI there were important differences, those with high baseline HbA_{1C} (≥ 88 mmol/mol) were more likely to receive a SU, and those in obesity class 2 or 3 were more likely to have a DPP4i. **Figure 6.1.** shows the variation in the estimated individualised treatment effects across the overall target population.

Table 6.2. presents the LIV individualised treatment effect estimates aggregated according to the overall population (ATE) and to pre-specified subgroups (CATE). Except for those patients in the highest baseline HbA_{1C} category, the estimated ATE and CATEs were small, and less than the thresholds for clinical or statistical significance. **Table 6.2.** reports evidence of HTE according to baseline HbA_{1C} category. For the stratum with high baseline HbA_{1C} (>=88 mmol/mol) the mean difference in HbA_{1C} in favour of DPP4i versus SU was large (-5.3) albeit with CIs that included zero (-12.8 to 2.2).

Table 6.3. presents the baseline characteristics from the Nauck et al. RCT versus the subpopulations from the CPRD data who met the RCT eligibility criteria (RCT-eligible), and for those who did not meet these criteria (RCT-ineligible) as well as for the overall target populations. By definition, the 'RCT-eligible population' excluded the stratum with the older age group and higher baseline HbA_{1c}. The RCT-eligible sample were representative of the broader target population according to characteristics not used to define eligibility criteria such as BMI, gender and ethnicity. However, mean baseline HbA_{1C} was much lower in the Nauck et al RCT and the corresponding 'RCT-eligible' subpopulation, than for the 'RCT-ineligible' population.

For the target trials there were some differences between the treatment groups in observed baseline characteristics. For the 'RCT-eligible' subpopulation, most strata were balanced, but there were between-treatment group differences in the proportion of patients in obesity class 2 and 3, and in baseline HbA_{1c} >=75 to 87. As anticipated, baseline imbalances were more pronounced for the 'RCT-ineligible' subpopulation, with large between-treatment differences in the proportions with HbA1C ≥ 88 and in the proportions in Obesity Classes 2 and 3. These observed measures may be correlated with unobserved characteristics such as dietary patterns, levels of exercise and adherence to previous medications, which increases the risk of unobserved confounding and also essential heterogeneity in the 'RCT ineligible' subpopulation.

Figure 6.2. compares the ATE from the Nauck et al RCT to the appropriate ATE from the LIV method, which comprise of individual-level treatment effects aggregated to the 'RCT-eligible' subpopulation. Neither of the estimated ATEs were of 'clinical' or 'statistical' significance. The point estimates for the 'RCT-eligible' population were within the 95% CIs of those for the RCT, indicating 'estimate agreement'. As the RCT eligibility criteria meant that there were important differences in baseline characteristics that were anticipated to modify the treatment effect, the ATEs for the 'RCT-ineligible' and the 'target' populations were not directly comparable to those for the Nauck et al RCT, but are reported here for completeness. The ATE estimate for the 'RCT-ineligible' population was more uncertain than for the 'RCT-eligible' population which reflected greater variation in the individualised treatment effect estimates across the broader patient group, which included those who failed the eligibility criteria according to observed levels of baseline HbA_{1C} and age (see **Figure 6.3.** and corresponding figures for these two baseline measures in the supplement).

The alternative analysis assuming no unobserved confounding provided estimates of ATEs that were in the opposite direction to those from the LIV, but for the 'RCT-eligible' population still met the criteria for 'estimate agreement' with those from Nauck et al. The alternative analyses like the LIV reported more uncertainty in the estimates for the 'RCT-ineligible' versus 'RCT-eligible' subpopulations. While the

estimated ATEs from IPTW-RA are statistically different from zero, none of them were of a magnitude that met the criteria for clinical significance (See **Figure S5** in the Supplement).

Figure 6.4. reports the CATE estimates from the LIV method for the 'RCT-eligible' and 'RCT-ineligible' subpopulations as well as for the target population. The results show some evidence of HTE in particular according to baseline HbA_{1C}, age group, and BMI (additional subgroups available in Supplementary Figures S6-S8). The estimated CATEs were in similar directions for the 'RCT-eligible' and 'RCT-ineligible' subpopulations and therefore for the target populations. For some subgroups, the magnitude of the estimated CATEs within the 'RCT-ineligible' subpopulation were somewhat different to those for the 'RCT-eligible' population, and the estimated CATEs differed between the subgroups excluded from the RCT versus those included. In particular, for people with high levels of baseline HbA1C (>=88 mmol/mol) the estimated improvement in HbA_{1C} following DPP4i versus SU was of clinical significance, albeit estimated with high levels of uncertainty.

6. Discussion

Target populations for decision-making may differ from those eligible for RCTs according to baseline characteristics that may modify the relative effectiveness of health care interventions. We use a target trial emulation with an LIV method to enable us to fully examine treatment effect heterogeneity including effect modification according to levels of unobserved covariates (essential heterogeneity) across subpopulations eligible and ineligible for a published RCT. We consider the approach within a case study evaluating the effectiveness of two alternative second-line treatments for people with T2DM. We applied the LIV method to estimate individualised treatment effects, that we then aggregated to report ATEs and CATEs across the full target population, defined by a national clinical guideline, and for the 'RCT-eligible' and 'RCT-ineligible' subpopulations. The estimated ATEs for the 'RCT-eligible' population are similar to those from a published RCT. The estimated CATEs are in the same direction for the subpopulations included versus excluded from the RCT, but differ in magnitude. The variation in the estimated individual treatment effects is greater across the broader sample of people who do not meet the RCT inclusion criteria than for those who do.

This paper contributes to three related areas of research: the transportability of results from RCT eligible populations to target populations for decisions, essential heterogeneity, and emulating target trials. First, previous work has developed methods for transporting estimates of ATEs and CATEs from RCTs to a target population, with recent expansions including the use of flexible machine learning methods **(Allcott and Mullainathan, 2012, Dahabreh and Hernán, 2019, Elliott et al., 2023, Stuart et al., 2010, Hartman et al., 2015, Degtiar and Rose, 2023)**. This extant literature has made the common crucial assumption, that there is no essential heterogeneity, which can also be expressed as no trial selection according to unobserved

variables that modify the relative treatment effect. The plausibility of this assumption will depend on the setting, and will relate to issues around the RCT design and the availability of common baseline measures between the RCT and target population. An RCT with a more pragmatic design may impose less restrictive eligibility criteria and be less prone to select participants according to unobserved characteristics that are likely to modify the treatment effect. Also, an RCT nested within a data source such as a disease registry that collects a common set of baseline variables including all potential effect modifies is better placed to transport the RCT result. More generally, given the non-random selection of RCT participants, those included in RCTs are likely to differ according to measures that are not fully observed in the RCT(s) or observational data, including those correlated with the explicit inclusion criteria. For example, in our case study, people with poor baseline glycaemic control were excluded which is correlated with diet and previous adherence to medication, neither of which were observed in the EHR or RCT data **(Zaccardi et al., 2020, Nauck et al., 2007)**. Similarly, older patients were excluded which is correlated with frailty, which was also not observed. Hence, an RCT finding, in this case of no significant differences in clinical effectiveness between the alternative treatments, may not transport to the target population and subpopulations. The approach taken provides new evidence, for policy relevant subgroups excluded from the RCT.

Second, the paper contributes to the literature on assessing essential heterogeneity. Previous work has developed conceptual frameworks **(Heckman and Vytlacil, 1999, Heckman and Vytlacil, 2001, Heckman and Vytlacil, 2005, Cornelissen et al., 2016)** for understanding essential heterogeneity in general settings, shown how under standard IV assumptions and with a continuous IV the requisite marginal treatment effects can be estimated **(Heckman and Vytlacil, 1999, Heckman and Vytlacil, 2001)**, and how these can be aggregated to estimate policy-relevant estimands including the ATE and CATE **(Heckman and Vytlacil, 2001, Basu, 2014, Basu, 2015)**. The LIV approach taken in this paper has been applied across a diverse range of settings, including to assess the effects of interventions in education as well as health care **(Basu, 2014, Basu and Gore, 2015, Basu et al., 2007, Basu et al., 2014, Grieve et al., 2019)**. A recent simulation study showed that in the presence of both overt and essential heterogeneity, the LIV approach taken in this paper can report consistent effect estimates of ATEs and CATEs provided the IV is sufficiently strong (F statistic>100), and the sample size sufficiently large (>5000) **(Moler-Zapata et al., 2023a).** This paper adds to this literature by using target trial emulation to assess the performance of the approach when an ATE from a RCT is available for a subpopulation, and to explore essential heterogeneity across an entire target population including subpopulations who met RCT eligibility criteria, and those who did not. The approach taken has wider application to settings where the likely presence of essential heterogeneity raises challenges for the transportability of RCT(s) results to the target populations of interest.

Third, the paper contributes to the literature emulating target trials for decision-making. Previous studies have applied aspects of the approach we take, in using a target trial design, firstly to emulate a published RCT, and secondly to estimate ATEs for a population excluded from the RCT. However, these previous target trials have only applied analytical methods that assume 'no unobserved confounding'. In our study we found that for the 'RCT-eligible subpopulation' a method that assumes 'no observed confounding' (IPTW-RA) provides similar ATE estimates to those from the LIV approach and to the RCT benchmark, and also similar CATE estimates to the LIV approach. By contrast for the RCT-ineligible population, there were wide baseline imbalances according to observed potential confounders, such as age, which were likely correlated with unobserved confounders such as frailty. Hence, this is a more challenging setting for approaches that assume no unobserved confounding and no essential heterogeneity as it is likely these assumptions are implausible, and this may explain why the estimated CATEs from the IPTW-RA approach differed to those from the LIV method. Moreover, there was less variation in the estimated individualised treatment effects when they were calculated using the IPTW-RA approach versus the LIV. Hence, while the LIV estimates of the CATEs and ATEs are less precise than those for the IPTW-RA this may partly reflect the appropriate capture of essential heterogeneity. Future target trial approaches should consider LIV approaches that can address both confounding and heterogeneity, and avoid relying on methods that assume no unobserved confounding and no essential heterogeneity, especially when a valid instrument is available and interest lies in broader populations for which there is no RCT benchmark.

This paper is subject to some limitations. First, we explored HTE within a single clinical scenario which cannot cover all the features that arise in practice when attempting to transport comparative effectiveness evidence from RCT eligible populations to target populations. In other settings, where an RCT requiring informed consent is nested within the target population, this may imply somewhat different selection mechanisms, and imply further challenges when transporting the trial results to the target population. While the potential importance of deploying approaches that consider essential heterogeneity remain, an extra problem is that the form of treatment may differ between the specific protocols required for the RCT versus those used in practice. Second, the comparators of interest in the target population may not be included in any particular RCT. In response to this concern a network meta-analyses may include RCTs with the relevant comparators. Here, the challenge of transporting results from the RCT setting to the target population is somewhat different as the RCTs may well have different eligibility criteria, and so the approach taken here would need to be extended to consider HTE across the target population including subpopulations who do meet and do not meet the eligibility criteria for different RCTs within the network. Third, this paper considered a single endpoint, but the approach could be applied to multiple endpoints recognising that HTE may differ across the different outcomes. Fourth, in our example, we had a strong IV and moderately large sample size, but the LIV estimates that incorporated essential heterogeneity were

still somewhat imprecise. Stein-like approaches that combine efficient but inconsistent estimators (e.g. OLS) with consistent but inefficient estimators (e.g. 2SLS/LIV) to improve precision at the expense of a somewhat higher risk of bias may warrant consideration for estimation of HTE (**Hansen, 2017)**. This may be particularly useful for studies focussing on smaller subgroups or that have weaker instruments.

Future research is required to explore HTE in populations explicitly excluded from RCTs, across a broader array of settings including those where the RCT is nested within the EHR data to allow this aspect of selection to be formally studied. It would also be helpful to consider settings with multiple outcomes, only some of which may be available in the RCT with others in the EHR data, and in settings where the issue is in transporting findings from network meta-analyses of RCTs to a target population. In settings, where more than two treatment comparators are of interest, further development of the requisite LIV methods are required as current approaches identify the marginal treatment effect of a treatment versus the next best alternative **(Heckman et al., 2008)**, but do not readily provide CATE estimates for specific treatment comparisons.

References

ALLCOTT, H. & MULLAINATHAN, S. 2012. External validity and partner selection bias. *National Bureau of Economic Research*.

ANGRIST, J. D. & FERNÁNDEZ-VAL, I. 2013. ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework. *In:* ACEMOGLU, D., ARELLANO, M. & DEKEL, E. (eds.) *Advances in Economics and Econometrics: Tenth World Congress: Volume 3: Econometrics.* Cambridge: Cambridge University Press.

ANGRIST, J. D., IMBENS, G. W. & RUBIN, D. B. 1996. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association,* 91**,** 444-455.

BAIOCCHI, M., CHENG, J. & SMALL, D. S. 2014. Instrumental variable methods for causal inference. *Statistics in Medicine,* 33**,** 2297-2340.

BASU, A. 2014. ESTIMATING PERSON-CENTERED TREATMENT (PeT) EFFECTS USING INSTRUMENTAL VARIABLES: AN APPLICATION TO EVALUATING PROSTATE CANCER TREATMENTS. *Journal of Applied Econometrics,* 29**,** 671-691.

BASU, A. 2015. Person-centered Treatment (PeT) Effects: Individualized Treatment Effects Using Instrumental Variables. *The Stata Journal,* 15**,** 397-410.

BASU, A., COE, N. B. & CHAPMAN, C. G. 2018. 2SLS versus 2SRI: Appropriate methods for rare outcomes and/or rare exposures. *Health Economics,* 27**,** 937-955.

BASU, A. & GORE, J. L. 2015. Are Elderly Patients With Clinically Localized Prostate Cancer Overtreated? Exploring Heterogeneity in Survival Effects. *Medical Care,* 53**,** 79-86.

BASU, A., HECKMAN, J., NAVARRO-LOZANO, S. & URZUA, S. 2007. Use of instrumental variables in the presence of heterogeneity and self-selection: an application to treatments of breast cancer patients. *Health Economics,* 16**,** 1133-1157.

BASU, A., JONES, A. & ROSA DIAS, P. 2014. The roles of cognitive and non-cognitive skills in moderating the effects of mixed-ability schools on long-term health. *National Bureau of Economic Research*.

BASU, A. & RATHOUZ, P. J. 2005. Estimating marginal and incremental effects on health outcomes using flexible link and variance function models. *Biostatistics,* 6**,** 93-109.

BELLONI, A., CHEN, D., CHERNOZHUKOV, V. & HANSEN, C. 2012. Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica,* 80**,** 2369-2429.

BIDULKA, P., LUGO-PALACIOS, D. G., CARROLL, O., O'NEILL, S., ADLER, A. I., BASU, A., SILVERWOOD, R., BARTLETT, J. W., NITSCH, D., CHARLTON, P., BRIGGS, A., SMEETH, L., DOUGLAS, I. J., KHUNTI, K. & GRIEVE, R. 2024. Comparative effectiveness of alternative second-line oral antidiabetic treatments on metabolic, kidney, and cardiovascular outcomes amongst people with type 2 diabetes mellitus: a cohort study using routinely collected health data. *BMJ,* In Press.

BIDULKA, P., MATHUR, R., LUGO-PALACIOS, D. G., O'NEILL, S., BASU, A., SILVERWOOD, R. J., CHARLTON, P., BRIGGS, A., SMEETH, L., ADLER, A. I., DOUGLAS, I. J., KHUNTI, K. & GRIEVE, R. 2023. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes Obes Metab,* 25**,** 282-292.

BIDULKA, P., O'NEILL, S., BASU, A., WILKINSON, S., SILVERWOOD, R. J., CHARLTON, P., BRIGGS, A., ADLER, A. I., KHUNTI, K., TOMLINSON, L. A., SMEETH, L., DOUGLAS, I. J. & GRIEVE, R. 2021. Protocol for an observational cohort study investigating personalised medicine for intensification of treatment in people with type 2 diabetes mellitus: the PERMIT study. *BMJ Open,* 11**,** e046912.

BJÖRKLUND, A. & MOFFITT, R. 1987. The Estimation of Wage Gains and Welfare Gains in Self-Selection Models. *The Review of Economics and Statistics,* 69**,** 42-49.

CANIVELL, S., MATA-CASES, M., REAL, J., FRANCH-NADAL, J., VLACHO, B., KHUNTI, K., GRATACÒS, M. & MAURICIO, D. 2019. Glycaemic control after treatment intensification in patients with type 2 diabetes uncontrolled on two or more non-insulin antidiabetic drugs in a real-world setting. *Diabetes, Obesity and Metabolism,* 21**,** 1373-1380.

CORNELISSEN, T., DUSTMANN, C., RAUTE, A. & SCHÖNBERG, U. 2016. From LATE to MTE: Alternative methods for the evaluation of policy interventions. *Labour Economics,* 41**,** 47-60.

CORNELISSEN, T., DUSTMANN, C., RAUTE, A. & SCHÖNBERG, U. 2018. Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy,* 126**,** 2356-2409.

DAHABREH, I. J. & HERNÁN, M. A. 2019. Extending inferences from a randomized trial to a target population. *European Journal of Epidemiology,* 34**,** 719-722.

DEGTIAR, I. & ROSE, S. 2023. A Review of Generalizability and Transportability. *Annual Review of Statistics and Its Application,* 10**,** 501-524.

ELLIOTT, M. R., CARROLL, O., GRIEVE, R. & CARPENTER, J. 2023. Improving transportability of randomized controlled trial inference using robust prediction methods. *Statistical Methods in Medical Research,* 32**,** 2365-2385.

EUROPEAN MEDICINES AGENCY. 2018. *Guideline on clinical investigation of medicinal products in the treatment or prevention of diabetes mellitus* [Online]. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-or-prevention-diabetes-mellitus-revision-1_en.pdf [Accessed 18 March 2024].

FRANK, L. E. & FRIEDMAN, J. H. 1993. A Statistical View of Some Chemometrics Regression Tools. *Technometrics,* 35**,** 109-135.

FRANKLIN, J. M., PATORNO, E., DESAI, R. J., GLYNN, R. J., MARTIN, D., QUINTO, K., PAWAR, A., BESSETTE, L. G., LEE, H., GARRY, E. M., GAUTAM, N. & SCHNEEWEISS, S. 2021. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies. *Circulation,* 143**,** 1002-1013.

FUNK, M. J., WESTREICH, D., WIESEN, C., STÜRMER, T., BROOKHART, M. A. & DAVIDIAN, M. 2011. Doubly Robust Estimation of Causal Effects. *American Journal of Epidemiology,* 173**,** 761-767.

GAN, S., DAWED, A. Y., DONNELLY, L. A., NAIR, A. T. N., PALMER, C. N. A., MOHAN, V. & PEARSON, E. R. 2020. Efficacy of Modern Diabetes Treatments DPP-4i, SGLT-2i, and GLP-1RA in White and Asian Patients With Diabetes: A Systematic Review and Meta-analysis of Randomized Controlled Trials. *Diabetes Care,* 43**,** 1948-1957.

GHEORGHE, A., ROBERTS, T. E., IVES, J. C., FLETCHER, B. R. & CALVERT, M. 2013. Centre Selection for Clinical Trials and the Generalisability of Results: A Mixed Methods Study. *PLOS ONE,* 8**,** e56560.

GOMES, M., LATIMER, N., SOARES, M., DIAS, S., BAIO, G., FREEMANTLE, N., DAWOUD, D., WAILOO, A. & GRIEVE, R. 2022. Target Trial Emulation for Transparent and Robust Estimation of Treatment Effects for Health Technology Assessment Using Real-World Data: Opportunities and Challenges. *PharmacoEconomics,* 40**,** 577-586.

GRADE STUDY RESEARCH GROUP 2022. Glycemia Reduction in Type 2 Diabetes — Glycemic Outcomes. *New England Journal of Medicine,* 387**,** 1063-1074.

GRIEVE, R., O'NEILL, S., BASU, A., KEELE, L., ROWAN, K. M. & HARRIS, S. 2019. Analysis of Benefit of Intensive Care Unit Transfer for Deteriorating Ward Patients: A Patient-Centered Approach to Clinical Evaluation. *JAMA Network Open,* 2**,** e187704-e187704.

HANSEN, B. E. 2017. Stein-like 2SLS estimator. *Econometric Reviews,* 36**,** 840-852.

HARTMAN, E., GRIEVE, R., RAMSAHAI, R. & SEKHON, J. S. 2015. From Sample Average Treatment Effect to Population Average Treatment Effect on the Treated: Combining Experimental with Observational Studies to Estimate Population Treatment Effects. *Journal of the Royal Statistical Society Series A: Statistics in Society,* 178**,** 757-778.

HECKMAN, J., URZUA, S. & VYTLACIL, E. 2006. Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics,* 88**,** 389-432.

HECKMAN, J. & VYTLACIL, E. 1999. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences,* 96**,** 4730-4734.

HECKMAN, J. & VYTLACIL, E. 2001. Policy-Relevant Treatment Effects. *American Economic Review,* 91**,** 107-111.

HECKMAN, J. & VYTLACIL, E. 2005. Structural Equations, Treatment Effects, and Econometric Policy Evaluation1. *Econometrica,* 73**,** 669-738.

HECKMAN, J. J., URZUA, S. & VYTLACIL, E. 2008. Instrumental Variables in Models with Multiple Outcomes: the General Unordered Case. *Annales d'Économie et de Statistique*, 151-174.
HERBERT, A., WIJLAARS, L., ZYLBERSZTEJN, A., CROMWELL, D. & HARDELID, P. 2017. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *International Journal of Epidemiology,* 46, 1093-1093i.
HERNÁN, M. A. & ROBINS, J. M. 2016. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology,* 183, 758-764.
HERNÁN, M. A., SAUER, B. C., HERNÁNDEZ-DÍAZ, S., PLATT, R. & SHRIER, I. 2016. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology,* 79, 70-75.
HERRETT, E., GALLAGHER, A. M., BHASKARAN, K., FORBES, H., MATHUR, R., VAN STAA, T. & SMEETH, L. 2015. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology,* 44, 827-836.
HOSMER, D. W. & LEMESHOW, S. 2000. *Applied Logistic Regression* New York, John Wiley & Sons.
ICER 2020. 2020-2023 Value Assessment Framework. *In:* ICER (ed.). Boston, MA: ICER.
IMAI, K., KING, G. & STUART, E. A. 2008. Misunderstandings Between Experimentalists and Observationalists about Causal Inference. *Journal of the Royal Statistical Society Series A: Statistics in Society,* 171, 481-502.
IMBENS, G. W. & ANGRIST, J. D. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica,* 62, 467-475.
KHUNTI, K., GODEC, T. R., MEDINA, J., GARCIA-ALVAREZ, L., HILLER, J., GOMES, M. B., CID-RUZAFA, J., CHARBONNEL, B., FENICI, P., HAMMAR, N., HASHIGAMI, K., KOSIBOROD, M., NICOLUCCI, A., SHESTAKOVA, M. V., JI, L. & POCOCK, S. 2018. Patterns of glycaemic control in patients with type 2 diabetes mellitus initiating second-line therapy after metformin monotherapy: Retrospective data for 10 256 individuals from the United Kingdom and Germany. *Diabetes, Obesity and Metabolism,* 20, 389-399.
MARATHUR, N. M., TSENG, E., HUTFLESS, S., WILSON, L. M., SUAREZ-CUERVO, C., BERGER, Z., CHU, Y., IYOHA, E., SEGAL, J. B. & BOLEN, S. 2016. Diabetes Medications as Monotherapy or Metformin-Based Combination Therapy for Type 2 Diabetes. A Systematic Review and Meta-analysis. *Annals of Internal Medicine,* 164, 740-751.
MOLER-ZAPATA, S., GRIEVE, R., BASU, A. & O'NEILL, S. 2023a. How does a local instrumental variable method perform across settings with instruments of differing strengths? A simulation study and an evaluation of emergency surgery. *Health Economics,* 32, 2113-2126.
MOLER-ZAPATA, S., HUTCHINGS, A., O'NEILL, S., SILVERWOOD, R. J. & GRIEVE, R. 2023b. Emulating Target Trials With Real-World Data to Inform Health Technology Assessment: Findings and Lessons From an Application to Emergency Surgery. *Value in Health,* 26, 1164-1174.
NAUCK, M. A., MEININGER, G., SHENG, D., TERRANELLA, L. & STEIN, P. P. 2007. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes Obes Metab,* 9, 194-205.
NEYMAN, J. 1990. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science,* 5, 465-472.
NICE 2013. Guide to the methods of technology appraisal. *In:* NICE (ed.). London: NICE.
NICE. 2021. *Methods, processes and topic selection for health technology evaluation: proposals for change. Appendix 1: Real World Evidence framework* [Online]. Available: https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/nice-guidance/chte-methods-and-processes-consultation/appendix-real-world-evidence-framework.docx [Accessed 25 November 2021].
NICE. 2022. *NG28: Type 2 diabetes in adults: management* [Online]. Available: https://www.nice.org.uk/guidance/ng28 [Accessed 22 December 2023].

PREGIBON, D. 1980. Goodness of Link Tests for Generalized Linear Models. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 29**,** 15-14.

REYNOLDS, K., BARTON, L., BASU, A., FISCHER, F., ARTERBURN, D., BARTHOLD, D., COURCOULAS, A., CRAWFORD, C., KIM, B., FEDORKA, P., MUN, E., MURALI, S., ZANE, R. & COLEMAN, K. 2021. Comparative Effectiveness of Gastric Bypass and Vertical Sleeve Gastrectomy for Hypertension Remission and Relapse: The ENGAGE CVD Study. *Hypertension,* 78**,** 1116-1125.

ROBERTSON, S. E., STEINGRIMSSON, J. A. & DAHABREH, I. J. 2023. Regression-based estimation of heterogeneous treatment effects when extending inferences from a randomized trial to a target population. *European Journal of Epidemiology,* 38**,** 123-133.

ROSENSTOCK, J., KAHN, S. E., JOHANSEN, O. E., ZINMAN, B., ESPELAND, M. A., WOERLE, H. J., PFARR, E., KELLER, A., MATTHEUS, M., BAANSTRA, D., MEINICKE, T., GEORGE, J. T., VON EYNATTEN, M., MCGUIRE, D. K., MARX, N. & INVESTIGATORS, F. T. C. 2019. Effect of Linagliptin vs Glimepiride on Major Adverse Cardiovascular Outcomes in Patients With Type 2 Diabetes: The CAROLINA Randomized Clinical Trial. *JAMA,* 322**,** 1155-1166.

RUBIN, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology,* 66**,** 688-701.

SMITH, M. J., MANSOURNIA, M. A., MARINGE, C., ZIVICH, P. N., COLE, S. R., LEYRAT, C., BELOT, A., RACHET, B. & LUQUE-FERNANDEZ, M. A. 2022. Introduction to computational causal inference using reproducible Stata, R, and Python code: A tutorial. *Statistics in Medicine,* 41**,** 407-432.

STAIGER, D. & STOCK, J. H. 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica,* 65**,** 557-586.

STATACORP 2023. *Stata: Release 18. Statistical Software. STATA causal inference and treatment-effects estimation reference manual,* College Station, TX, StataCorp.

STUART, E. A., COLE, S. R., BRADSHAW, C. P. & LEAF, P. J. 2010. The Use of Propensity Scores to Assess the Generalizability of Results from Randomized Trials. *Journal of the Royal Statistical Society Series A: Statistics in Society,* 174**,** 369-386.

TAN, Z. 2006. Regression and Weighting Methods for Causal Inference Using Instrumental Variables. *Journal of the American Statistical Association,* 101**,** 1607-1618.

TIBSHIRANI, R. 2018. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological),* 58**,** 267-288.

VYTLACIL, E. 2002. Independence, Monotonicity, and Latent Index Models: An Equivalence Result. *Econometrica,* 70**,** 331-341.

WANG, S. V., SCHNEEWEISS, S. & RCT-DUPLICATE INITIATIVE 2023. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA,* 329**,** 1376-1385.

WILKINSON, S., DOUGLAS, I., STIRNADEL-FARRANT, H., FOGARTY, D., POKRAJAC, A., SMEETH, L. & TOMLINSON, L. 2018. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open,* 8**,** e022768.

WOLF, A., DEDMAN, D., CAMPBELL, J., BOOTH, H., LUNN, D., CHAPMAN, J. & MYLES, P. 2019. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology,* 48**,** 1740-1740g.

WOOLDRIDGE, J. M. 2007. Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics,* 141**,** 1281-1301.

ZACCARDI, F., JACQUOT, E., CORTESE, V., TYRER, F., SEIDU, S., DAVIES, M. J. & KHUNTI, K. 2020. Comparative effectiveness of gliclazide modified release versus sitagliptin as second-line treatment after metformin monotherapy in patients with uncontrolled type 2 diabetes. *Diabetes, Obesity and Metabolism,* n/a.

Table 6.1. Baseline measures for the treatment groups in the Target Population

Variable	Target Population (N=13,240)	
	SU	DPP4i
	(n=8,289 - 62.6%)	(n=4,951– 37.4%)
Age: Mean (SD)	61.0 (11.7)	60.9 (11.5)
Age group years – N (%)		
Younger than 50	1,465 (17.7%)	854 (17.3%)
50-59	2,151 (26.0%)	1,360 (27.5%)
60-69	2,588 (31.2%)	1,543 (31.2%)
70-78	1,584 (19.1%)	919 (18.6%)
79 or older	501 (6.0%)	275 (5.6%)
Female N (%)	3,290 (39.7%)	1,977 (39.9%)
BMI: kg/m^2 – Mean (SD)	32.0 (6.2)	33.4 (6.4)
Under/normal weight – BMI (15-24.9)	805 (9.7%)	315 (6.4%)
Overweight – BMI (25-29.9)	2,687 (32.4%)	1,314 (26.5%)
Obese (class 1) – BMI (30-34.9)	2,582 (31.2%)	1,630 (32.9%)
Obese (class 2 & 3) –BMI (35 and higher)	2,215 (26.7%)	1,592 (34.2%)
Baseline HbA_{1c}: mmol/mol – Mean (SD)	76.0 (19.1)	70.4 (14.8)
HbA1c distribution at baseline N (%)		
HbA1c < 44 mmol/mol	21 (0.3%)	11 (0.2%)
HbA_{1c} ≥ 44 to 63 mmol/mol	2,299 (27.7%)	1,891 (38.2%)
HbA_{1c} ≥64 to 74 mmol/mol	2,562 (30.9%)	1,648 (33.3%)
HbA_{1c} ≥75 to 87 mmol/mol	1,567 (18.9%)	810 (16.4%)
HbA_{1c} ≥88 mmol/mol	1,840 (22.2%)	591 (11.9%)
Ethnicity (%)		
White	7,112 (85.8%)	4,426 (89.4%)
South Asian	819 (9.9%)	371 (7.5%)
Black / Mixed / Other	358 (4.3%)	154 (3.1%)
Duration of diabetes: years – Mean (SD)	5.3 (4.8)	5.8 (4.8)

Table 6.2. ATE and CATEs for the Target population from the LIV approach

* Difference in the change in HbA_{1c} (mmol/mol) from baseline

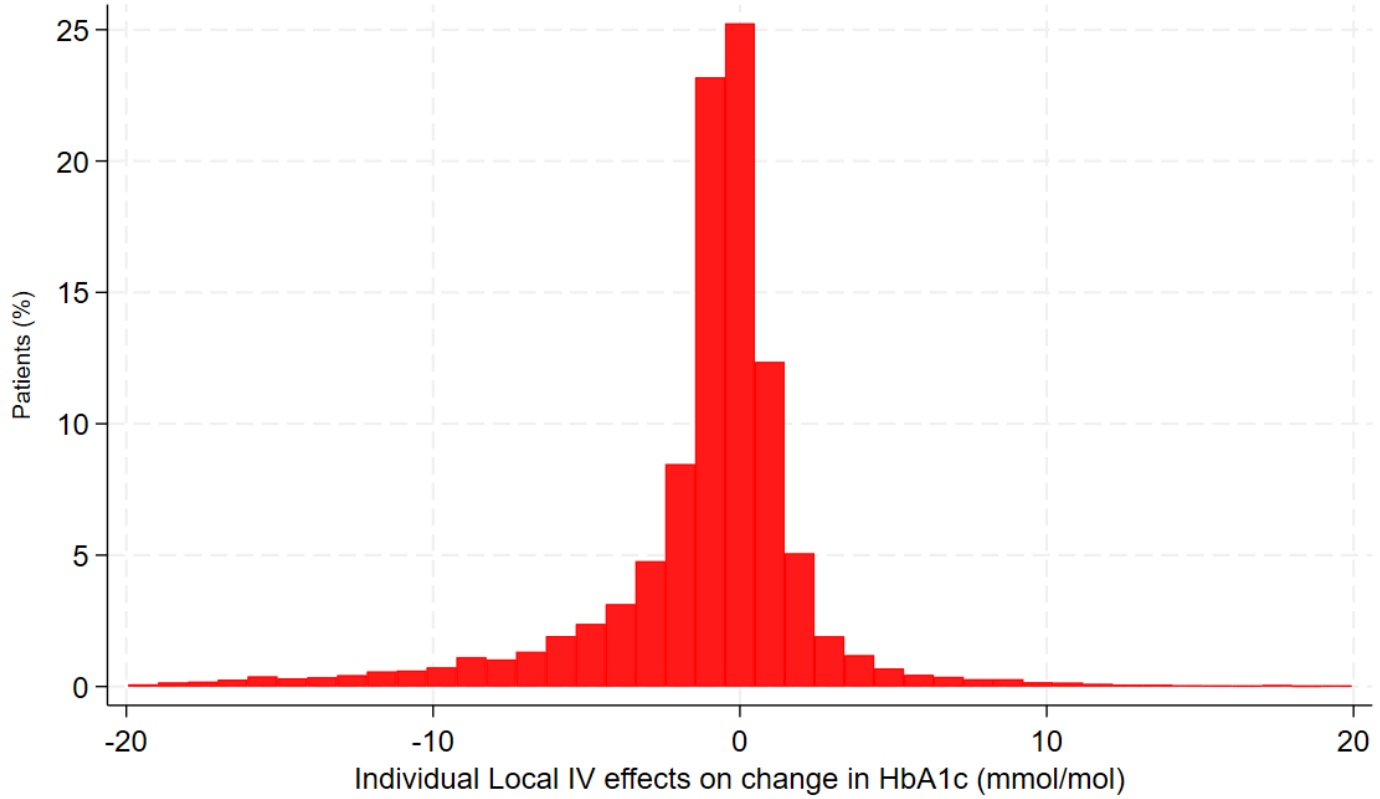
Variable	Patients per subgroup (N)	Treatment effect (DPP4i vs SU)* Point estimate (95% CI)
Overall	13,240	-1.3 (-3.3, 0.8)
Age group years		
Younger than 50	2,319	-1.3 (-5.9, 3.3)
50-59	3,511	-3.0 (-5.6, -0.4)
60-69	4,131	-1.6 (-3.8, 0.6)
70-78	2,503	0.6 (-1.7, 2.8)
79 or older	776	2.7 (-1.3, 6.6)
Female	5,267	-1.2 (-3.1, 0.8)
BMI (kg/m^2)		
Under/normal weight – BMI (15-24.9)	1,120	-0.2 (-3.6, 3.1)
Overweight – BMI (25-29.9)	4,001	-0.5 (-2.9, 1.9)
Obese (class 1) – BMI (30-34.9)	4,212	-1.1 (-3.1, 0.9)
Obese (class 2 & 3) –BMI (35 and higher)	3,907	-2.5 (-5.2, 0.1)
HbA1c distribution at baseline		
HbA1c < 44 mmol/mol	32	-2.2 (-10.4, 6.0)
HbA_{1c} ≥ 44 to 63 mmol/mol	4,190	-0.1 (-1.8, 1.6)
HbA_{1c} ≥64 to 74 mmol/mol	4,210	-0.3 (-2.0, 1.5)
HbA_{1c} ≥75 to 87 mmol/mol	2,377	-1.1 (-3.9, 1.7)
HbA_{1c} ≥88 mmol/mol	2,431	-5.3 (-12.8, 2.2)
Ethnicity		
White	11,538	-1.3 (-3.3, 0.7)
South Asian	1,190	-0.7 (-3.5, 2.1)
Black / Mixed / Other	512	-1.5 (-5.1, 2.0)

Table 6.3. Baseline measures by treatment group for the Benchmark RCT (shaded), and for the 'RCT eligible', 'RCT ineligible and overall target populations from the CPRD data (target trials).

Variable	Benchmark RCT Nauck et al (N=1,172)		Target Trial RCT-eligible Population (N=6,497)		RCT ineligible population (N=6,743)		Target Population (N=13,240)	
	SU (n=588 - 50.2%)	DPP4i (n=584 - 49.8%)	SU (n=3,931 - 60.5%)	DPP4i (n=2,566– 39.5%)	SU (n=4,358 - 64.6%)	DPP4i (n=2,385– 35.4%)	SU (n=8,289 - 62.6%)	DPP4i (n=4,951– 37.4%)
Age: years – Mean (SD)	56.6 (9.8)	56.8 (9.3)	60.8 (10.4)	60.3 (10.2)	61.2 (12.8)	61.6 (12.7)	61.0 (11.7)	60.9 (11.5)
Age subgroups N (%)								
Younger than 50			614 (15.6%)	426 (16.6%)	851 (19.5%)	428 (18.0%)	1,465 (17.7%)	854 (17.3%)
50-59	*	*	1,035 (26.3%)	726 (28.3%)	1,116 (25.6%)	634 (26.6%)	2,151 (26.0%)	1,360 (27.5%)
60-69	*	*	1,385 (35.2%)	883 (34.4%)	1,203 (27.6%)	660 (27.7%)	2,588 (31.2%)	1,543 (31.2%)
70-78	*	*	897 (22.8%)	531 (20.7%)	687 (15.8%)	388 (16.3%)	1,584 (19.1%)	919 (18.6%)
79 or older	*Ineligible*	*Ineligible*	*Ineligible*	*Ineligible*	501 (11.5%)	275 (11.5%)	501 (6.0%)	275 (5.6%)
Female (%)	226 (38.7%)	252 (42.9%)	1,579 (40.2%)	1,012 (39.4%)	1,711 (39.3%)	965 (40.5%)	3,290 (39.7%)	1,977 (39.9%)
BMI: kg/m^2 – Mean (SD)	31.3 (5.2)	31.2 (5.0)	32.2 (6.1)	33.6 (6.3)	31.8 (6.3)	33.1 (6.5)	32.0 (6.2)	33.4 (6.4)
BMI subgroups N (%)								
Under/normal weight	*	*	337 (8.6%)	132 (5.1%)	468 (10.7%)	183 (7.7%)	805 (9.7%)	315 (6.4%)
Overweight	*	*	1,289 (32.8%)	669 (26.1%)	1,398 (32.1%)	645 (27.04%)	2,687 (32.4%)	1,314 (26.5%)
Obese (class 1)	*	*	1,237 (31.5%)	867 (33.8%)	1,345 (30.9%)	763 (32.0%)	2,582 (31.2%)	1,630 (32.9%)
Obese (class 2 & 3)	*	*	1,068 (27.2%)	898 (35.0%)	1,147 (26.3%)	794 (33.3%)	2,215 (26.7%)	1,592 (34.2%)
Baseline HbA_{1c}: mmol/mol – Mean (SD)	58.8 (6.6)	58.4 (5.9)	68.8 (8.3)	67.4 (8.2)	82.5 (23.3)	73.6 (19.1)	76.0 (19.1)	70.4 (14.8)
Baseline HbA1c subgroups N (%)								
HbA1c < 44 mmol/mol	*Ineligible*	*Ineligible*	*Ineligible*	*Ineligible*	21 (0.5%)	11 (0.5%)	21 (0.3%)	11 (0.2%)
HbA_{1c} ≥ 44 to 63 mmol/mol	381 (65.5%)	375 (64.0%)	1,225 (31.1%)	993 (38.7%)	1,074 (24.6%)	898 (37.7%)	2,299 (27.7%)	1,891 (38.2%)
HbA_{1c} ≥64 to 74 mmol/mol	141 (24.2%)	151 (25.8%)	1,674 (42.6%)	1,053 (41.0%)	888 (20.4%)	595 (25.0%)	2,562 (30.9%)	1,648 (33.3%)
HbA_{1c} ≥75 to 87 mmol/mol	60 (10.3%)	60 (10.2%)	1,032 (26.3%)	520 (20.3%)	535 (12.3%)	290 (12.2%)	1,567 (18.9%)	810 (16.4%)
HbA_{1c} ≥88 mmol/mol	*Ineligible*	*Ineligible*	*Ineligible*	*Ineligible*	1,840 (42.2%)	591 (24.8%)	1,840 (22.2%)	591 (11.9%)
Ethnicity N (%)								
White	74.3%	73.5%	3,421 (87.0%)	2,315 (90.2%)	3,691 (84.7%)	2,111 (88.5%)	7,112 (85.8%)	4,426 (89.4%)
South Asian	8.4%	8.5%	379 (9.6%)	185 (7.2%)	440 (10.1%)	186 (7.8%)	819 (9.9%)	371 (7.5%)
Black / Mixed / Other	17.3%	18.0%	131 (3.3%)	66 (2.6%)	227 (5.2%)	88 (3.7%)	358 (4.3%)	154 (3.1%)

Variable	Benchmark RCT Nauck et al (N=1,172)		Target Trial					
			RCT-eligible Population (N=6,497)		RCT ineligible population (N=6,743)		Target Population (N=13,240)	
	SU (n=588 - 50.2%)	DPP4i (n=584 - 49.8%)	SU (n=3,931 - 60.5%)	DPP4i (n=2,566– 39.5%)	SU (n=4,358 - 64.6%)	DPP4i (n=2,385– 35.4%)	SU (n=8,289 - 62.6%)	DPP4i (n=4,951– 37.4%)
Duration of diabetes: Years – Mean (SD)	6.5 (6.1)	6.2 (5.4)	5.5 (4.7)	5.9 (5.0)	5.2 (4.9)	5.6 (4.7)	5.3 (4.8)	5.8 (4.8)

Figure 6.1. Distribution of estimated individual treatment effects reported as expected difference (DPP4i-SUs) in change in HbA_{1C} (mmol/mol) between baseline and 1 year for the target population from the LIV approach†



† For presentational purposes, 135 individual effects (1% of target population) outside the range (-20,20) mmol/mol were excluded.

Figure 6.2. ATEs from Nauck et al. (2007), and for the corresponding 'RCT eligible' subpopulation, 'RCT eligible' and overall target populations from the target trial using the LIV approach. Average Treatment effects (ATEs) reported as difference (DPP4i-SUs) in change in HbA_{1C} (mmol/mol) between baseline and 1 year.

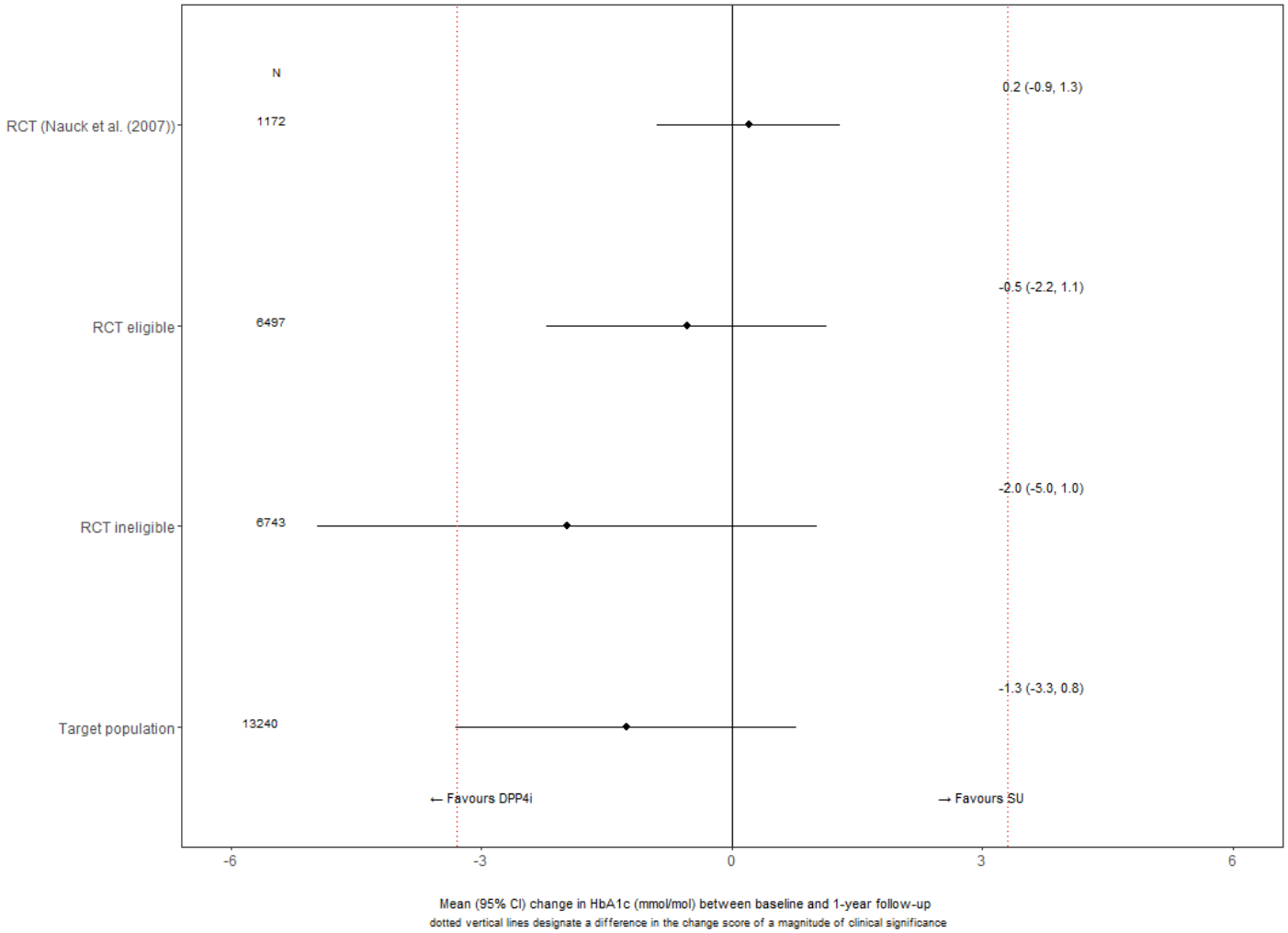
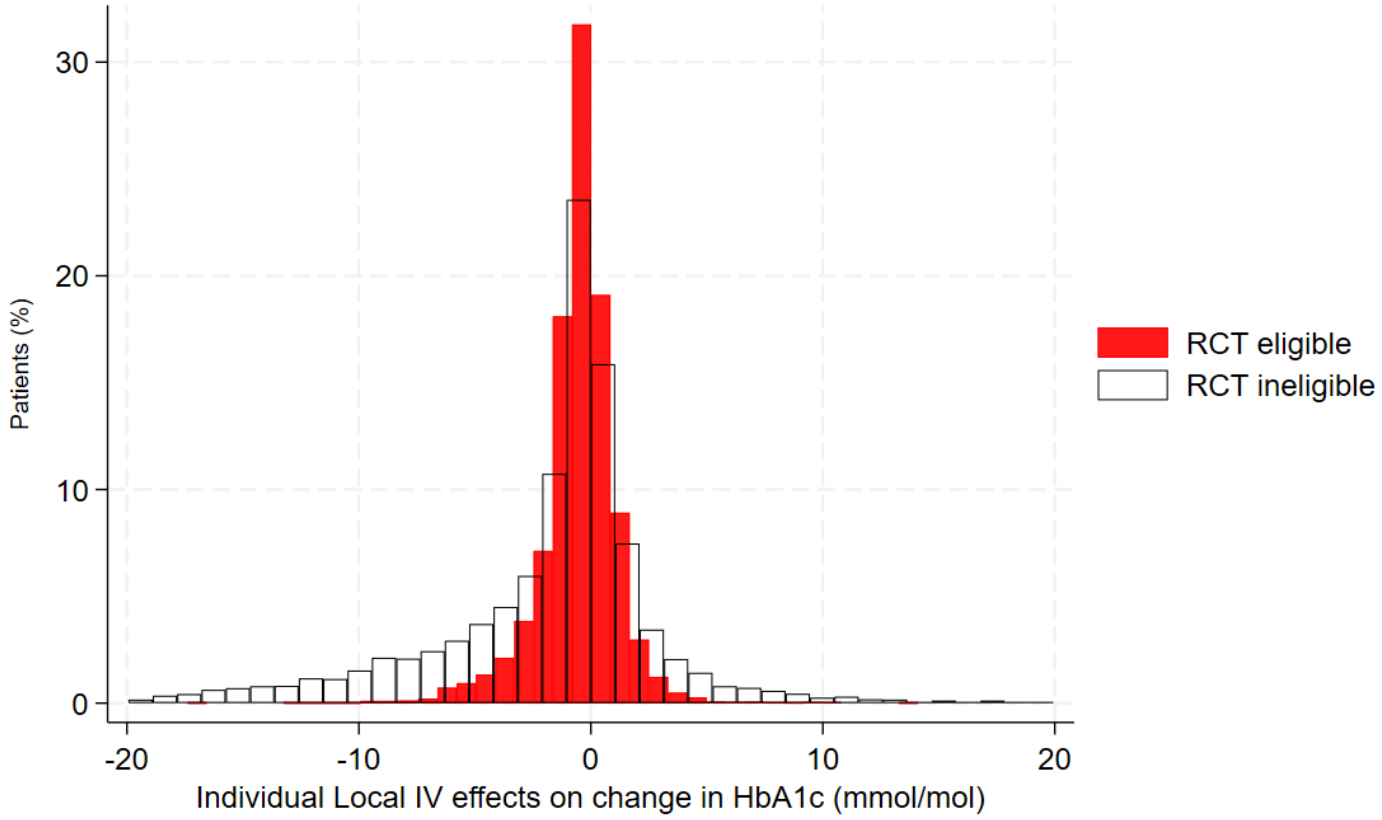
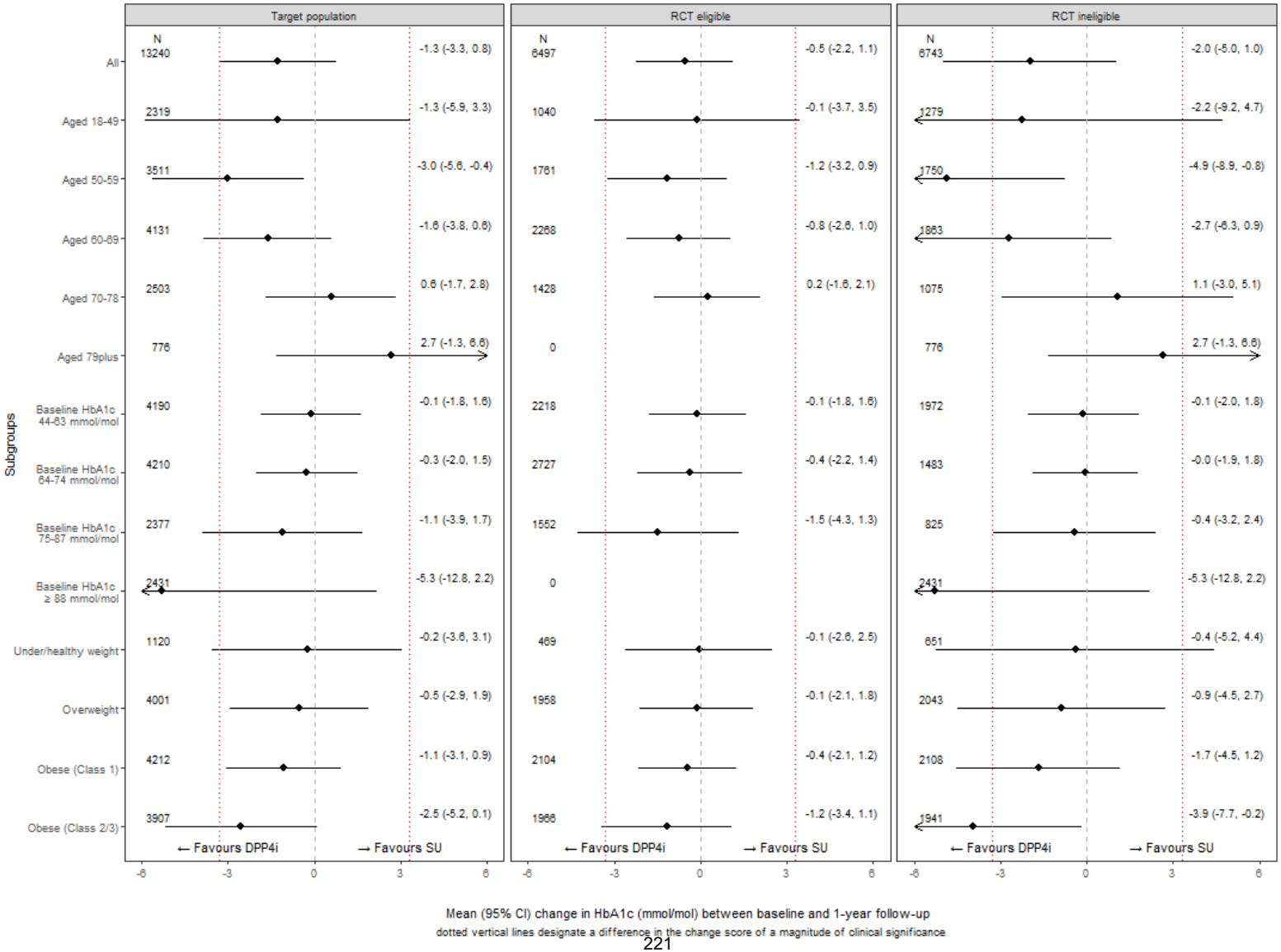


Figure 6.3. Distribution of estimated individual treatment effects reported as expected difference (DPP4i-SUs) in change in HbA_{1c} (mmol/mol) between baseline and 1 year, for the 'RCT eligible' versus 'RCT ineligible' subpopulations. †



† For presentational purposes, 135 individual effects (all from the RCT ineligible population and accounting for 1% of target population) outside the range (-20,20) mmol/mol were excluded.

Figure 6.4. Conditional Average Treatment Effects (CATEs) for the 'RCT eligible' subpopulation, 'RCT eligible' and overall target populations from the target trial using the LIV approach for age and baseline HbA_{1c} subgroups. CATEs reported as difference (DPP4i-SUs) in change in HbA_{1C} (mmol/mol) between baseline and 1 year.



6.2. Relevance to my thesis

This research paper addressed thesis objective 1D: to investigate heterogenous treatment effects across the target population of people with T2DM in English primary care. NICE guidance on generating real-world evidence relevant for HTA emphasises the potential for NRS to transport trial results to those populations excluded from RCTs to improve clinical decision-making.[1] NICE also highlights the potential for NRS to investigate heterogenous treatment effects, which is often difficult in RCTs due to the large sample sizes required for this type of investigation.[1] Transporting trial results to those subpopulations who do not meet trial inclusion criteria is further complicated by potential differences in the underlying distribution of treatment effect modifiers, which can be unmeasured or imperfectly measured in routinely collected health data (e.g., diet and frailty, respectively).

Thus, this chapter uses an LIV analysis[2-4] to account for this essential heterogeneity when estimating the comparative effectiveness of DPP4i and SU: two of the three most commonly prescribed second-line oral antidiabetic treatments globally.[5, 6] The LIV can estimate individual-level treatment effects, which can be aggregated to the subgroup or overall study population level to estimate CATEs and the ATE.[2] Combined with this LIV approach, I again used the target trial emulation framework;[7, 8] however, in this chapter, unlike the previous, I was able to identify a suitable *published* RCT to emulate.[9]

I led the review of trials comparing DPP4i and SU to select a suitable trial(s) for this TTE. I used a list of trials provided by Dr Elaine Butterly at the University of Glasgow, as part of her on-going systematic review and meta-analysis that is unpublished at the time of writing this thesis. From this list of 35 published trials, I used the screening process outlined in **Table 6.4.** to select the trial(s) which were suitable for the target trial emulation.

Table 6.4. Description of the trial review screening process to identify suitable trials for this target trial emulation (Table S3a in the supplementary materials of the submitted manuscript)

Screening step	**Exclusion Details**
Screen 1	Review trial registration (e.g., ClinicalTrials.gov) and exclude trials which: • Are not phase 3 trials. • Are not double-blind (e.g., open-label). • Are not in the general type 2 diabetes mellitus (T2DM) population (e.g., include only elderly patients, patients with chronic kidney disease). • Do not report HbA1c at 52-weeks (1-year) as an outcome.
Screen 2	Review trial registration and published peer-reviewed papers and exclude trials which: • Analyse results with a per-protocol analysis as the primary analysis.

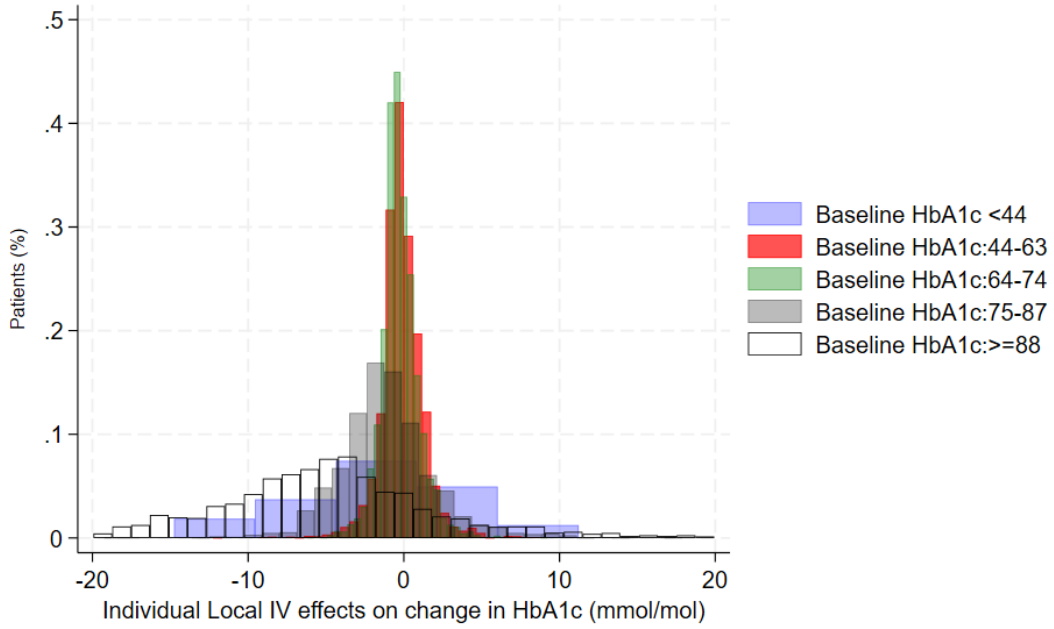
	• Impute missing outcome data (HbA1c at 1-year) without rigorous methodology (e.g., last observation carried forward). • Exclude based on criteria difficult to define in the Clinical Practice Research Datalink (CPRD) (e.g., liver laboratory test results, family history of disease, clinical judgements). • Exclude people with a history of cancer. • Only report changes in HbA1c at 1-year using a figure (i.e., cannot extract precise outcome measures). • Peer-reviewed article not available online.

CPRD: Clinical Practice Research Datalink; HbA1c: haemoglobin A1c

Details of the RCT screening process are provided in Appendix F.1. to F.2. I selected one trial by Nauck et al (2007) which was suitable for this study.[9] After defining the study population (Appendix F.3.) and summarising the target population's baseline characteristics, I applied the RCT inclusion and exclusion criteria as best as I could emulate in the routinely collected health data to define the 'trial eligible' and 'trial ineligible' subpopulations.

The LIV analysis was then used to estimate the change in the mean difference in HbA1c from baseline to 1-year follow-up for DPP4i versus SU, reducing the risk of confounding under the IV assumptions and accounting for essential heterogeneity in the target population. The individual treatment effects were aggregated at the subpopulation levels to estimate the ATE in the 'trial eligible' and 'trial ineligible' subpopulations, as well as in the overall target population. Comparisons to the ATE from the RCT showed agreement for the 'trial eligible' subpopulation. For the 'trial ineligible' population the estimates of the ATE and CATE were more uncertain. Substantial variation in individual level treatment effects was reported in the target population, particularly among the subgroup of people with HbA1c ≥ 88mmol/mol at baseline (who were excluded from the RCT) (**Figure 6.5**). This variation indicates the challenges in generalising the ATEs estimated in trials to the target population in the presence of heterogeneity.

Figure 6.5. Distribution of estimated individual treatment effects reported as expected difference (DPP4i-SUs) in change in HbA_{1C} (mmol/mol) between baseline and 1 year for the target population from the LIV approach across pre-specified subgroups of HbA1c at baseline (Figure S4 from the supplementary materials of the submitted paper).†



† For presentation purposes, 135 individual effects (1% of target population) outside the range (-20,20) were excluded.

This application of the LIV analysis follows previous examples where unmeasured confounding and treatment heterogeneity according to observed and unobserved variables are considered major sources of potential bias.[2, 3, 10] Following on from these studies, I have applied the LIV analysis within the target trial emulation framework in a study of antidiabetic treatment for people with T2DM, an area of great clinical interest in the UK and globally. Specifically, this chapter makes three important contributions to this thesis:

1. The agreement between the ATE measured in the RCT and the ATE measured in the trial eligible subpopulation offers no evidence to suggest that there are violations of the major IV assumptions in Chapters 5 to 6.
2. The agreement also suggests that the IV is minimising residual confounding in the NRS in Chapters 5 to 6.

3. An additional example of how the TTE framework combined with an IV analysis can be useful in generating evidence from routinely collected health data that is relevant in policymaking decisions.

In the thesis discussion (section 10.5.2.), I discuss areas of future work where the LIV method can be applied to generate useful evidence for HTA and clinical practice.

This is the final chapter in Case Study 1 of this thesis. In this case study, I was able to address an important area of clinical uncertainty using advanced quantitative methods and analyses to generate high-quality evidence from non-randomised data. There are limitations to the component studies of this case study which I describe in each research paper's discussion section, as well as in the overall thesis discussion (section 10.4.). I have conducted several alternative analyses to partly address these limitations.

In the next case study, I consider a different area of clinical uncertainty – alternative AMI treatment among people with reduced kidney function. This case study follows a similar approach to designing a comparative effectiveness analysis using routinely collected health data as in Case Study 1. However, as I illustrate in the next chapter, the suitability of the secondary care data sources is more challenging to confirm in the secondary care setting.

6.3. References

1. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

2. Basu A. ESTIMATING PERSON-CENTERED TREATMENT (PeT) EFFECTS USING INSTRUMENTAL VARIABLES: AN APPLICATION TO EVALUATING PROSTATE CANCER TREATMENTS. *J Appl Econ (Chichester Engl)*. June/July 2014;29(4):671-691. doi:10.1002/jae.2343

3. Moler-Zapata S, Grieve R, Lugo-Palacios D, et al. Local Instrumental Variable Methods to Address Confounding and Heterogeneity when Using Electronic Health Records: An Application to Emergency Surgery. *Medical decision making : an international journal of the Society for Medical Decision Making*. Nov 2022;42(8):1010-1026. doi:10.1177/0272989x221100799

4. Heckman JJ, Vytlacil EJ. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proc Natl Acad Sci U S A*. Apr 13 1999;96(8):4730-4. doi:10.1073/pnas.96.8.4730

5. Khunti K, Chatterjee S, Gerstein HC, Zoungas S, Davies MJ. Do sulphonylureas still have a place in clinical practice? *The Lancet Diabetes & Endocrinology*. 2018/10/01/ 2018;6(10):821-832. doi:https://doi.org/10.1016/S2213-8587(18)30025-1

6. Khunti K, Godec TR, Medina J, et al. Patterns of glycaemic control in patients with type 2 diabetes mellitus initiating second-line therapy after metformin monotherapy: Retrospective data for 10 256 individuals from the United Kingdom and Germany. *Diabetes, obesity & metabolism*. Feb 2018;20(2):389-399. doi:10.1111/dom.13083

7. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

8. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

9. Nauck MA, Meininger G, Sheng D, Terranella L, Stein PP. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes, obesity & metabolism*. Mar 2007;9(2):194-205. doi:10.1111/j.1463-1326.2006.00704.x

10. Basu A, Gore JL. Are Elderly Patients With Clinically Localized Prostate Cancer Overtreated? Exploring Heterogeneity in Survival Effects. *Medical care*. Jan 2015;53(1):79-86. doi:10.1097/mlr.0000000000000260

CHAPTER 7. RESEARCH PAPER – IMPACT OF CHRONIC KIDNEY DISEASE ON CASE ASCERTAINMENT FOR HOSPITALISED ACUTE MYOCARDIAL INFARCTION: AN ENGLISH COHORT STUDY.

OVERVIEW

In this chapter, I include a research paper published in *BMJ Open* for which I am joint first author. I presented this work at UK Kidney Week (virtual conference) as a poster presentation with a pre-recorded video presentation in 2021, as well as at invited talks at McMaster University and internally at LSHTM.

This is the first research paper part of Case Study 2: Relative effectiveness of alternative AMI treatments among people with reduced kidney function in secondary care. In this study, I explore biases in the data sources to be used in this case study. First, I investigate AMI case ascertainment across MINAP and HES, two secondary care datasets, and how this case ascertainment is associated with kidney impairment among individuals hospitalised for AMI. In addition, I investigate the agreement between kidney function estimated from primary care versus secondary care data.

Following the research paper, I include a brief discussion of this paper's relevance to my thesis. This discussion includes key tables and figures from the supplementary materials of the published paper, as well as two unpublished figures which aid in the interpretation of the study in the context of this thesis.

7.1. Published research paper

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMJ Open		
When was the work published?	23 February 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I am joint first author and corresponding author of this paper. I co-led the study design with JS (joint first author), DA, and DN. I was solely responsible for data extraction, data management, and data analysis. I led the interpretation of results with support from my co-authors. I co-wrote the first draft of the manuscript with JS. Co-authors reviewed multiple drafts before sending to the journal for publication. I led the peer review responses with input from my co-authors.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Dorothea Nitsch
Date	4/4/24

BMJ Open

Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study

Patrick Bidulka ,[1] Jemima Scott,[2,3] Dominic M Taylor,[2,3] Udaya Udayaraj,[4,5] Fergus Caskey,[2,3] Lucy Teece,[6] Michael Sweeting,[6] John Deanfield,[7,8] Mark de Belder,[7] Spiros Denaxas,[9,10] Clive Weston,[11] David Adlam,[12] Dorothea Nitsch[1]

To cite: Bidulka P, Scott J, Taylor DM, *et al*. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. *BMJ Open* 2022;**12**:e057909. doi:10.1136/bmjopen-2021-057909

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-057909).

PB and JS contributed equally. DA and DN contributed equally.

Received 04 October 2021
Accepted 23 February 2022

For numbered affiliations see end of article.

Correspondence to
Patrick Bidulka;
patrick.bidulka1@lshtm.ac.uk

ABSTRACT

Objectives Acute myocardial infarction (AMI) case ascertainment improves for the UK general population using linked health data sets. Because care pathways for people with chronic kidney disease (CKD) change based on disease severity, AMI case ascertainment for these people may differ compared with the general population. We aimed to determine the association between CKD severity and AMI case ascertainment in two secondary care data sets, and the agreement in estimated glomerular filtration rate (eGFR) between the same data sets.

Methods We used a cohort study design. Primary care records for people with CKD or risk factors for CKD, identified using the National CKD Audit (2015–2017), were linked to the Myocardial Ischaemia National Audit Project (MINAP, 2007–2017) and Hospital Episode Statistics (HES, 2007–2017) secondary care registries. People with an AMI recorded in either MINAP, HES or both were included in the study cohort. CKD status was defined using eGFR, derived from the most recent serum creatinine value recorded in primary care. Moderate–severe CKD was defined as eGFR <60 mL/min/1.73 m^2, and mild CKD or at risk of CKD was defined as eGFR ≥60 mL/min/1.73 m^2 or eGFR missing. CKD stages were grouped as (1) At risk of CKD and Stages 1–2 (eGFR missing or ≥60 mL/min/1.73 m^2), (2) Stage 3a (eGFR 45–59 mL/min/1.73 m^2), (3) Stage 3b (eGFR 30–44 mL/min/1.73 m^2) and (4) Stages 4–5 (eGFR <30 mL/min/1.73 m^2).

Results We identified 6748 AMIs: 23% were recorded in both MINAP and HES, 66% in HES only and 11% in MINAP only. Compared with people at risk of CKD or with mild CKD, AMIs in people with moderate–severe CKD were more likely to be recorded in both MINAP and HES (42% vs 11%, respectively), or MINAP only (22% vs 5%), and less likely to be recorded in HES only (36% vs 84%). People with AMIs recorded in HES only or MINAP only had increased odds of death during hospitalisation compared with those recorded in both (adjusted OR 1.61, 95% CI 1.32 to 1.96 and OR 1.60, 95% CI 1.26 to 2.04, respectively). Agreement between eGFR at AMI admission (MINAP) and in primary care was poor (kappa (K) 0.42, SE 0.012).

Conclusions AMI case ascertainment is incomplete in both MINAP and HES, and is associated with CKD severity.

Strengths and limitations of this study

- ► Our study includes a large sample size of 6748 acute myocardial infarction (AMI) events.
- ► We have assessed the completeness of AMI hospitalisations recorded in two healthcare data sets widely used in observational research in England.
- ► We evaluated, for the first time, the validity of using serum creatinine recorded in secondary care at the time of an AMI to estimate pre-AMI chronic kidney disease (CKD) stage.
- ► Generalisability to the general population is limited as the National Chronic Kidney Disease Audit only included people with CKD and/or risk factors for CKD.

INTRODUCTION

Prognosis following acute myocardial infarction (AMI) has improved considerably over the past 50 years such that 85% of individuals now live longer than 1 year post-AMI.[1] Improved survival is the result of advances in AMI management, driven by evidence from large-scale randomised controlled trials (RCTs).[2–6] Of those admitted to hospital with AMI, 30%–40% have chronic kidney disease (CKD)[7]: a sustained reduction in kidney function associated with poor outcomes.[8 9] Among those with dialysis-dependent CKD only 40% will survive their first year post-AMI.[10] These inferior outcomes may result from higher prevalence of comorbidity,[2] calcific coronary artery disease[3] and the pro-inflammatory effects of uraemia.[4]

Most major RCTs investigating AMI interventions excluded patients with advanced CKD.[11] However, current AMI guidelines

from Europe and the USA apply the results of these RCTs to those with or without CKD.[5–7] Clinicians' unease with the dearth of evidence may explain diversion from these AMI guidelines when treating people with CKD.[2 10 11] In the absence of specific RCTs in CKD populations, well-conducted observational analyses can contribute significantly to our understanding and improved management of AMI.

In the UK, data on AMI treatment and outcomes is collected in unlinked, disease-specific registries or in broad registration databases. While there are known differences in the reliability and validity of AMI case ascertainment using these resources in the general population,[12] it is unclear to what extent these differences persist in people with underlying CKD. Multimorbidity and differences in admission and treatment pathways in people with CKD may influence AMI case recording. Reliably identifying which patients with AMI have CKD using AMI audit data is also difficult; previous studies used admission serum creatinine (SCr) as a proxy for pre-admission CKD stage.[13–15] This unvalidated method risks misclassifying people as having CKD because of the co-incidence of acute kidney injury (AKI) and AMI.[16]

In this study we linked records from the National Chronic Kidney Disease Audit (NCKDA) to the Myocardial Ischaemia National Audit Project (MINAP) and Hospital Episode Statistics (HES) to determine the reliability of these data sources to investigate cardiovascular disease comorbidity and outcomes in people with or at risk of CKD in England. Our objectives were to: (1) Compare case ascertainment of AMI hospitalisations in secondary care data sets (MINAP and HES); (2) determine if MINAP and/or HES case ascertainment defines populations of patients with CKD with different risks of death during and after AMI; and (3) compare CKD stage classification using admission SCr recorded in secondary care (MINAP) versus primary care (NCKDA).

METHODS

Data sources

Data from all sources were restricted to patients treated in England. People with or at risk of CKD were identified using primary care data from the NCKDA.[17 18] The NCKDA aimed to optimise the identification and management of people with CKD and/or risk factors for CKD in primary care, and included 10% of English General Practices (GP).[17 18] NCKDA data were collected between 2015 and 2016 in two main cross-sectional data extracts for people with either blood or urine laboratory results indicating CKD and/or risk factors for CKD (prevalent hypertension, diabetes mellitus, cardiovascular disease, connective tissue disorders, kidney stones, prostatic disease, family history of kidney disease, previous AKI and users of kidney-damaging medications such as lithium or calcineurin inhibitors).[17] People without an estimated glomerular filtration rate (eGFR) recorded in primary care were included in the NCKDA only if they were at risk of CKD.

The population identified from NCKDA was linked with Office of National Statistics (ONS) data, January 1998 to September 2019, as well as secondary care data from HES Admitted Patient Care (APC) and MINAP, April 2007 to April 2017. HES APC includes hospital admission data for National Health Service-treated patients in England, including admission and discharge dates, and diagnoses recorded using International Classification of Diseases 10th Edition (ICD-10) codes.[19] MINAP is an ongoing AMI audit in England, Wales and Northern Ireland which was designed to optimise the care of patients with type one AMI by evaluating the patient pathway from hospital admission with AMI[20] to discharge. MINAP includes admission and discharge dates, treatments and comorbidities.[21]

Study design

Cohort study.

Study participants

We included people in the NCKDA registered with a GP in England, with one or more AMI hospitalisation recorded in HES or MINAP. People in each NCKDA extract must have been alive according to GP records at the time of that extract. We therefore included people with an AMI hospitalisation recorded in MINAP or HES only after the date of their GP's final NCKDA extract. People with an AMI hospitalisation that started prior to the extract date and ended after the extract date were added to the cohort, since they were at risk of death (n=183). In addition, people with an ONS death date indicating death during an AMI hospitalisation that occurred within 90 days prior to the NCKDA extract date were included (n=96), since they were likely misclassified as alive at the time of the extract because of delays in updating the death date in the GP systems. People with an ONS death date earlier than 90 days prior to the extract were excluded (n=4).

Exposures

Our main exposure variable was moderate to severe CKD (eGFR <60 mL/min/1.73 m^2, CKD stages 3–5), defined using the most recent eGFR recorded in primary care (NCKDA data) prior to the AMI hospitalisation. People with no eGFR recorded in primary care or an eGFR ≥60 mL/min/1.73 m^2 were categorised as at risk of CKD or having mild CKD, respectively. We assumed people with no eGFR recorded in primary care did not have moderate to severe CKD since these people are much less likely to have CKD than those with eGFR recorded.[22] eGFR was calculated using primary care SCr measures and the revised Modification of Diet in Renal Disease (MDRD) equation.[23 24]

Our secondary exposure was CKD stage, defined by the Kidney Disease Improving Global Outcomes CKD staging, based on a single eGFR record without the requirement for two measures 3 months apart.[25] We combined some CKD stages due to low numbers of AMI cases: (1) At

risk of CKD and Stages 1–2 (eGFR missing or ≥60 mL/min/1.73 m^2), (2) Stage 3a (eGFR 45–59 mL/min/1.73 m^2), (3) Stage 3b (eGFR 30–44 mL/min/1.73 m^2) and (4) Stages 4–5 (eGFR <30 mL/min/1.73 m^2).

We used the latest eGFR recorded prior to the AMI hospitalisation to categorise people with a history of kidney transplant into the primary and secondary exposure groups. We categorised people with a history of dialysis prior to the AMI hospitalisation as moderate to severe CKD for the main exposure and CKD stages 4–5 for the secondary exposure, even if the latest eGFR did not agree.

As the use of a single SCr test at the time of AMI hospitalisation to determine CKD stage has not previously been validated, we have used the term 'eGFR stage' in place of CKD stage to refer to the eGFR level calculated from this test.

Outcomes

Primary outcome

The primary outcome was AMI case ascertainment, defined as the data set(s) in which the AMI hospitalisation was recorded. We defined an AMI as being recorded in both HES and MINAP if an AMI hospitalisation in HES was within 30 days of an AMI hospitalisation in MINAP. Where multiple HES AMI hospitalisations fell within 30 days of a MINAP AMI hospitalisation, the HES AMI hospitalisation closest in time to the MINAP AMI admission was selected as the single matched event. AMI hospitalisations without a match were categorised as HES or MINAP only. Study participants could contribute multiple AMI hospitalisations.

We defined an AMI in HES data using ICD-10 codes I.21, I.22 or I.23 in the primary admission diagnosis field (first diagnostic position in the first episode of an admission).[26] We categorised AMI subtype (ST-segment elevation myocardial infarction (STEMI) and non-ST-segment elevation myocardial infarction (NSTEMI)) using the UK Biobank coding definitions[27] (online supplemental table 1). We defined AMI in MINAP using a previously developed algorithm which uses the discharge diagnosis, ECG results and the presence of elevated cardiac markers to identify acute coronary syndrome events and subtypes (online supplemental table 2). We excluded MINAP hospitalisations classified as unstable angina or other from the analysis.[12]

Secondary outcomes

We investigated in-hospital mortality during each person's first AMI hospitalisation within the study period. In those who survived and were discharged from their first AMI hospitalisation, we also investigated post-discharge mortality using the ONS death date (up to 15 September 2019). Variables in HES, MINAP and ONS used to define death are described in online supplemental table 3. People were considered to have died during AMI hospitalisation if any of these variables indicated in-hospital death, or the ONS death date fell on or between the admission and discharge dates. We used the earliest of the HES or MINAP admission dates and the latest of the HES or MINAP discharge dates to define these dates for AMI hospitalisations recorded in both data sets.

We investigated the agreement between CKD stage derived from the most recent primary care SCr test (NCKDA data) and eGFR stage derived from the secondary care SCr test conducted within 24 hours of AMI hospitalisation (MINAP data). We used the same methods to determine eGFR stage in MINAP data as we did for NCKDA data.

Covariates

We described age at AMI admission (mean and SD as well as age category in years: 18–49, 50–64, 65–79, 80+), sex, ethnicity (white or other), index of multiple deprivation quintiles (IMD, as a proxy for socioeconomic status) and relevant comorbidities including angina, cerebrovascular disease, chronic obstruction pulmonary disease (COPD), diabetes mellitus (type 1 and 2), heart failure, hypertension, previous myocardial infarction and peripheral vascular disease. We also described dialysis and transplant status, and smoking status. Data sources for each key covariate are described in online supplemental table 4.

Data analysis

Objective 1—AMI case ascertainment

We summarised key covariates by CKD status. We used Venn diagrams to describe AMI case ascertainment overall and stratified by CKD status (at risk of CKD or mild CKD vs moderate to severe CKD). We used multinomial, multivariable logistic regression to quantify the association between CKD stage and AMI case ascertainment. We used the 'HES and MINAP' category as the base outcome and reported crude and adjusted relative risk ratios (RRR) and 95% CIs, using 'At risk of CKD and Stages 1–2' as the reference exposure category. We adjusted for sex, age category, ethnicity, IMD quintile, previous AMI, heart failure, COPD, diabetes mellitus and clustering by participant (using cluster-robust standard errors (SEs)). We used a complete case analysis since we could not assume that missing values for ethnicity were missing at random. In a secondary analysis, we stratified these regressions by AMI subtype (STEMI and NSTEMI).

Objective 2—risk of death

We used multivariable logistic regression to calculate the odds of death in hospital during each person's first AMI hospitalisation in people with AMI recorded in MINAP only or HES only, relative to MINAP and HES. After confirming the proportional hazards assumption with a Schoenfeld Residuals test on the full multivariable model (p=0.35), we used multivariable Cox regression to estimate HRs for death during total follow-up in those who survived their first AMI hospitalisation with AMI recorded in MINAP only or HES only, relative to MINAP and HES.

Objective 3—agreement between eGFR in primary and secondary care

Finally, to assess the validity of using MINAP-recorded eGFR at AMI admission as a proxy for pre-admission CKD status, we compared eGFR and its corresponding eGFR stage within 24 hours of AMI admission (MINAP data) to the most recent eGFR and its corresponding CKD stage in primary care (NCKDA data). In this analysis, we excluded people with eGFR measures greater than 120 mL/min/1.73 m^2 in either NCKDA or MINAP as these are unlikely to be true values. We drew a Bland-Altman plot to describe differences in the distribution of eGFR measures in primary and secondary care[28] and calculated the per cent agreement and kappa agreement statistics between CKD and eGFR stage derived using primary and secondary care eGFRs, respectively. Secondary analyses re-calculated agreements and kappa statistics restricting to people with stages 3a–5 (eGFR <60 mL/min/1.73 m^2) in primary care and grouped by time between the most recent primary care eGFR measure and the AMI hospitalisation (0–5, 6–11, 12–23, 24–36 month gaps).

Sensitivity analyses

We repeated the main analyses for AMI events occurring prior to the study start date (the latest NCKDA extract). People who experienced AMI hospitalisation before the study start were survivors, since only people alive at the time of the NCKDA were included in the study.

In addition, we re-drew the Venn diagrams after including HES AMI hospitalisations recorded in both the first and second diagnostic positions of the first episode to include AMIs recorded as co-primary diagnoses.[26] We also repeated the matching process between MINAP and HES AMI hospitalisations after combining all HES AMIs within 30 days of each other into a single HES hospitalisation, since it is likely that some AMI events in our data set have multiple HES hospitalisations recorded if, for example, a patient is transferred between hospitals for treatment. Furthermore, we repeated our multivariable analyses after excluding people with a history of dialysis. Finally, to understand why people may have an AMI hospitalisation recorded in MINAP but not HES, we searched for non-AMI HES hospitalisations within 30 days of the AMI recorded in MINAP and described the ICD-10 diagnoses in the first episode of the first diagnostic position.

Missing data

We did a complete case analysis when building our multivariable models. People with missing ethnicity (~1%) and IMD data (<1%) were excluded prior to building our unadjusted, partially adjusted and adjusted multinomial models.

We used discharge dates to help re-categorise people who were in-hospital at the time of the NCKDA extract into the cohort, as well as to determine death in hospital and the start of follow-up in those who survived their first AMI hospitalisation. Discharge date was missing in 19% and 1% of the MINAP and HES data sets, respectively. We assumed these dates were missing at random and used the median length of admissions in those without missing admission and discharge dates (5 and 4 days in MINAP and HES, respectively) to impute the missing discharge dates.

Patient and public involvement

The Kidney Care UK patient organisation (https://www.kidneycareuk.org/) supported the research questions, grant applications and the related record linkage application for section 251 permissions critical to the development of the NCKDA. Patient members of the UK Renal Registry Patient Council (https://renal.org/patients/patient-council) reviewed the study results. Their feedback supported a further planned record linkage of renal and cardiac data to look at patient outcomes.

RESULTS

Study population and baseline characteristics

From 1 702 345 people in England included in the NCKDA, we identified 6042 (0.4%) people with or at risk of CKD who experienced 6748 AMIs between the final NCKDA extract and 1 April 2017 (online supplemental figure 1). Baseline characteristics stratified by CKD stage are described in table 1. People with moderate to severe CKD accounted for 38% of AMI hospitalisations (2,575). Average age at the time of AMI was 73 years (SD 13). People with moderate to severe CKD were older on average than people with mild CKD or at risk of CKD. Most people were white (92%) and men (61%). The most prevalent comorbidities were hypertension (61%) and diabetes mellitus (35%).

AMI recording in HES and MINAP

Overall, 23% of AMI hospitalisations were captured by both MINAP and HES data sets (1552 AMI hospitalisations) (figure 1). There was no substantial change in AMI case ascertainment over time (online supplemental figure 2). In people with moderate to severe CKD, 42% of all AMI hospitalisations were captured by both MINAP and HES (1092 AMI hospitalisations). In people with mild CKD or at risk of CKD, 11% of all AMI hospitalisations were captured by MINAP and HES (460 AMI hospitalisations) (figure 1).

Relative association between CKD stage and AMI recording

Crude and adjusted RRRs and 95% CIs describing the association between CKD stage and AMI case ascertainment are presented in table 2. After adjusting for key covariates, we observed weak evidence of an increased likelihood of AMI recorded in MINAP only, compared with MINAP and HES, in people with CKD stages 4–5 versus the at risk of CKD/stages 1–2 group (RRR 1.34, 95% CI 0.97 to 1.85). Furthermore, compared with the at-risk of CKD/stages 1–2 group, people with CKD stages 3a, 3b and 4–5 were less likely to have an AMI hospitalisation recorded in HES only versus MINAP

Table 1 Baseline characteristics by CKD stage for all AMI events captured after the study start. n (column %) unless specified otherwise.

CKD status (main exposure)	At risk of CKD or mild CKD	Moderate to severe CKD			
CKD stage (secondary exposure)	At risk of CKD and 1–2	3a	3b	4–5	Total
Unique individuals	3751	1210	732	349	6042
Total number of AMI events, N	4173	1353	825	397	6748
Age at AMI event, years, mean (SD)	70 (13)	79 (10)	82 (9)	79 (12)	73 (13)
Age category at AMI event, years					
18–50	299 (7)	20 (1)	<5	10 (3)	330 (5)
50–64	1163 (28)	91 (7)	44 (5)	38 (10)	1336 (20)
65–79	1655 (40)	531 (39)	251 (30)	122 (31)	2559 (38)
80+	1056 (25)	711 (53)	529 (64)	227 (57)	2523 (37)
Female	1430 (34)	638 (47)	416 (50)	168 (42)	2652 (39)
Ethnicity					
White	3807 (91)	1263 (93)	773 (94)	361 (91)	6204 (92)
Other	323 (8)	73 (5)	44 (5)	34 (9)	474 (7)
Missing	43 (1)	17 (1)	8 (1)	<5	70 (1)
IMD quintile					
1 (least deprived)	732 (18)	252 (19)	152 (18)	68 (17)	1204 (18)
2	843 (20)	306 (23)	176 (21)	85 (21)	1410 (21)
3	934 (22)	331 (24)	193 (23)	79 (20)	1537 (23)
4	951 (23)	255 (19)	185 (22)	96 (24)	1487 (22)
5 (most deprived)	690 (17)	206 (15)	114 (14)	69 (17)	1079 (16)
Missing	23 (1)	<5	5 (1)	0 (0)	31 (0)
Dialysis in primary care					
Peritoneal dialysis	0 (0)	0 (0)	0 (0)	15 (4)	15 (0)
Haemodialysis	0 (0)	0 (0)	0 (0)	24 (6)	24 (0)
Renal dialysis, unspecified	0 (0)	0 (0)	0 (0)	7 (2)	7 (0)
Kidney transplant	5 (0)	0 (0)	<5	17 (4)	26 (0)
Comorbidities					
Angina	959 (23)	399 (29)	275 (33)	155 (39)	1788 (26)
Cerebrovascular disease	390 (9)	178 (13)	139 (17)	81 (20)	788 (12)
COPD	514 (12)	209 (15)	168 (20)	58 (15)	949 (14)
Diabetes mellitus	1293 (31)	465 (34)	356 (43)	233 (59)	2347 (35)
Heart failure	400 (10)	234 (17)	211 (26)	123 (31)	968 (14)
Hypertension	2333 (56)	884 (65)	583 (71)	322 (81)	4122 (61)
Myocardial infarction	1050 (25)	430 (32)	274 (33)	163 (41)	1917 (28)
Peripheral vascular disease	229 (5)	108 (8)	74 (9)	47 (12)	458 (7)
Smoking status					
Non-smoker	1953 (47)	566 (42)	306 (37)	151 (38)	2976 (44)
Ever-smoker	2018 (48)	530 (39)	318 (39)	140 (35)	3006 (45)
Missing	202 (5)	257 (19)	201 (24)	106 (27)	766 (11)

AMI, acute myocardial infarction; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; IMD, Index of Multiple Deprivation.;

and HES. We did not observe any differences in the likelihood of recording of AMI hospitalisation when stratifying by AMI subtype (online supplemental table 5).

Mortality during AMI hospitalisation and post-discharge

Of those with a first AMI recorded in both HES and MINAP, 209 people (15%) died during the AMI hospitalisation, compared with 151 (23%) with a first AMI

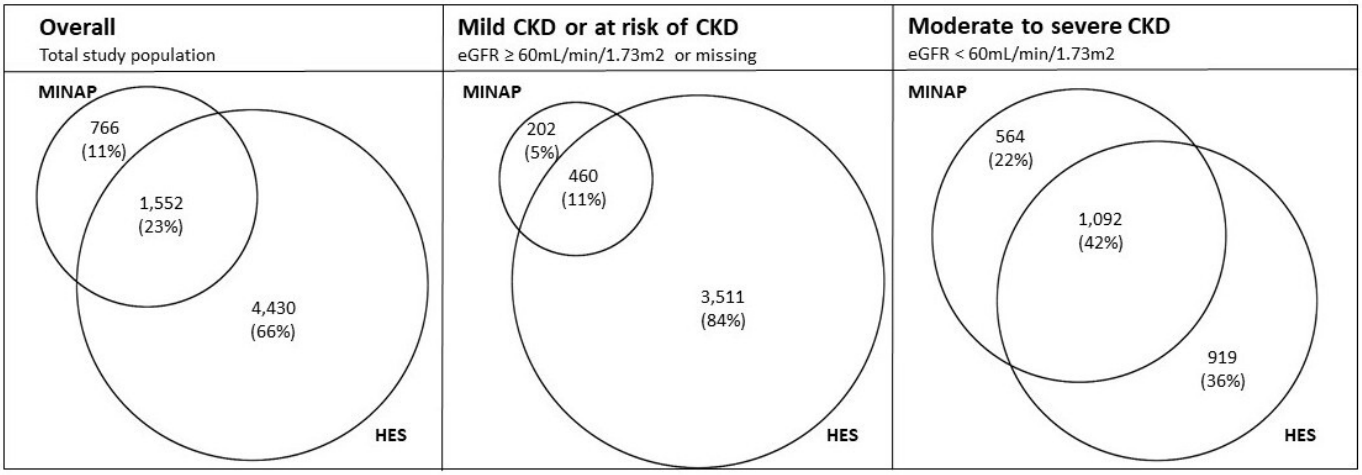


Figure 1 Venn diagrams illustrating acute myocardial infarction (AMI) recording in MINAP and HES secondary care data sets. Venn diagrams presented overall, and stratified by CKD status (at-risk of or mild CKD, eGFR ≥60 mL/min/1.73 m^2 or moderate to severe CKD, eGFR <60 mL/min/1.73 m^2). Circle areas are proportional to the number of AMI events in each data set.
CKD, chronic kidney disease, eGFR, estimated glomerular filtration rate, HES, hospital episode statistics, MINAP, Myocardial Ischaemia National Audit Project.

recorded in MINAP only and 579 (15%) recorded in HES only (table 3). After adjusting for key covariates, people with AMI recorded in MINAP only and HES only had increased odds of in-hospital death compared with people with AMI recorded in both MINAP and HES (OR 1.60, 95% CI 1.26 to 2.04 and OR 1.61, 95% CI 1.32 to 1.96, respectively).

Mean follow-up among people who survived a first AMI hospitalisation was 2.4 years. The rate of death per 100 person-years during complete follow-up was 18.0 (95% CI 16.4 to 19.7) for AMI recorded in MINAP and HES, 23.3 (95% CI 20.6 to 26.5) for AMI recorded in MINAP only and 10.3 (95% CI 9.61 to 11.0) for AMI recorded in HES only (table 3). After adjusting for key covariates, there was no evidence of a difference in death during follow-up based on which data set(s) captured the first AMI hospitalisation.

Agreement between eGFR derived from secondary care versus primary care data

Of the AMI hospitalisations recorded in MINAP, 2240 (97%) had SCr recorded within 24 hours of AMI admission

Table 2 Multinomial logistic regression comparing the RRR of AMI recording across HES and MINAP according to CKD stage. The comparator outcome is people with AMI recorded in both HES and MINAP databases.

AMI recording (outcome, compared with people with AMI recorded in MINAP and HES)	CKD stage (exposure)	Number of AMI admissions, n=	Unadjusted* RR (95% CI)	Partially adjusted† RRR (95% CI)	Adjusted‡ RRR (95% CI)
MINAP only (N=742)	At risk of CKD/stages 1–2	196	1	1	1
	Stage 3a	245	1.07 (0.85 to 1.34)	0.98 (0.77 to 1.25)	0.98 (0.77 to 1.25)
	Stage 3b	197	1.17 (0.92 to 1.49)	1.04 (0.80 to 1.36)	1.03 (0.79 to 1.34)
	Stages 4–5	104	1.50 (1.11 to 2.03)	1.38 (1.01 to 1.90)	1.34 (0.97 to 1.85)
HES only (N=4367)	At risk of CKD/stages 1–2	3456	1	1	1
	Stage 3a	557	0.14 (0.12 to 0.16)	0.14 (0.12 to 0.17)	0.14 (0.12 to 0.16)
	Stage 3b	224	0.08 (0.06 to 0.09)	0.08 (0.07 to 0.10)	0.08 (0.06 to 0.10)
	Stages 4–5	130	0.11 (0.08 to 0.14)	0.11 (0.09 to 0.15)	0.12 (0.09 to 0.16)

*Complete cases for adjusted model.
†Adjusted for sex, age at AMI admission, ethnicity (white, other), IMD quintile, clustering by participant.
‡Additionally adjusted for previous AMI, heart failure, COPD, diabetes mellitus.
AMI, acute myocardial infarction; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; HES, Hospital Episode Statistics; IMD, Index of Multiple Deprivation; MINAP, Myocardial Ischaemia National Audit Project; RRR, relative risk ratios.

Table 3 Death during and after first AMI hospitalisation in total study population at risk of or with CKD

Death during first AMI hospitalisation (N=5919)*	Number who died, n (%)	–	Unadjusted OR (95% CI)	Adjusted† OR (95% CI)
MINAP and HES	209 (15)	–	1	1
MINAP only	151 (23)	–	1.67 (1.32 to 2.11)	1.60 (1.26 to 2.04)
HES only	579 (15)	–	0.98 (0.82 to 1.16)	1.61 (1.32 to 1.96)
Death during complete follow-up in those who survive first AMI hospitalisation (N=5009)*	**Number who died during follow-up, n**	**Rate per 100 person-years (95% CI)**	**Unadjusted HR (95% CI)**	**Adjusted† HR (95% CI)**
MINAP and HES	456	18.0 (16.4 to 19.7)	1	1
MINAP only	237	23.3 (20.6 to 26.5)	1.27 (1.08 to 1.48)	1.12 (0.96 to 1.31)
HES only	847	10.3 (9.61 to 11.0)	0.59 (0.53 to 0.67)	1.07 (0.94 to 1.21)

*Complete cases for adjusted model.
†Adjusted for sex, age at AMI admission, ethnicity (white, other), IMD quintile, previous AMI, heart failure, COPD, diabetes mellitus.
AMI, acute myocardial infarction; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; HES, Hospital Episode Statistics; IMD, Index of Multiple Deprivation; MINAP, Myocardial Ischaemia National Audit Project; OR, odds ratio.

(online supplemental table 6). Median eGFR at time of admission was 47.6 mL/min/1.73 m^2 (IQR 33.5–61.6). The Bland-Altman plot comparing the primary care eGFR and the secondary care eGFR indicated a negligible mean difference but wide variation (mean difference 3.35 mL/min/1.73 m^2, 95% CI −23.4 to 30.1) (online supplemental figure 3).

The per cent agreements and kappa statistics between eGFR stage derived from MINAP eGFR at AMI admission and the CKD stage derived from NCKDA using primary care data are shown in table 4. Overall, there was 57.2% agreement in staging (kappa statistic (K) 0.42 (SE 0.012)). When restricting to people with NCKDA-derived CKD stages 3–5, the % agreement and K indicated worse agreement (table 4). However, when looking at agreement in categorising people as having moderate to severe CKD (stages 3–5) versus mild CKD (stages 1–2), agreement improved (82.1% agreement, K 0.55 (SE 0.021)).

When stratifying by months between the primary and secondary care eGFR measures, we observed the best agreement in staging within a 0–5 month gap between the primary and secondary care eGFR measures: 61.0%, K 0.48 (SE 0.03) (table 4). Agreement was worse when the time between eGFR measures increased.

Sensitivity analyses

AMI case ascertainment in MINAP and HES was similar in AMI hospitalisations recorded prior to the study start (sensitivity analysis) compared with after the study start (main analysis) (online supplemental tables 7-8, figure 4). There were also no major differences in agreement between CKD staging derived in primary versus secondary care when investigating AMI hospitalisations prior to the study start (online supplemental table 9).

After expanding the AMI definition in HES to include any hospitalisations with AMI coded in the second diagnostic position as well as the first, the proportion of AMI hospitalisations captured in both HES and MINAP decreased slightly (online supplemental figure 5). After combining HES AMI admissions within 30 days of each other for the same person, we observed a 1% increase in the proportion of AMI hospitalisations recorded in both MINAP and HES (online supplemental figure 5). Results were also similar when excluding people with a history of dialysis (online supplemental tables 10-11).

Table 4 Agreement between primary care-derived CKD stage (NCKDA) and secondary care-derived eGFR stage (MINAP)

	% agreement	Kappa statistic (SE)
Overall*	57.2	0.42 (0.012)
CKD stages 3a, 3b, 4–5†	53.2	0.34 (0.015)
CKD stages 1–2, 3a–5‡	82.1	0.55 (0.021)
Overall,* by time from NCKDA SCr test (primary care) to MINAP SCr test (at AMI secondary care admission)		
0–5 months	61.0	0.48 (0.03)
6–11 months	56.7	0.42 (0.02)
12–23 months	55.9	0.40 (0.02)
24–36 months	56.8	0.41 (0.04)

*Overall agreement when grouping as (1) Stages 1–2 (eGFR 60–120 mL/min/1.73 m^2), (2) Stage 3a (eGFR 45–59 mL/min/1.73 m^2), (3) Stage 3b (eGFR 30–44 mL/min/1.73 m^2) and (4) Stages 4–5 (eGFR 0–30 mL/min/1.73 m^2).
†Agreement when restricting to people with CKD stages 3a–5, grouped as (1) Stage 3a (eGFR 45–59 mL/min/1.73 m^2), (2) Stage 3b (eGFR 30–44 mL/min/1.73 m^2) and (3) Stages 4–5 (eGFR 0–30 mL/min/1.73 m^2).
‡Agreement when grouping as (1) Stages 1–2 (eGFR 60–120 mL/min/1.73 m^2) and (2) Stages 3a–5 (eGFR 0–59 mL/min/1.73 m^2).
AMI, acute myocardial infarction; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; MINAP, Myocardial Ischaemia National Audit Project; NCKDA, National Chronic Kidney Disease Audit; SCr, serum creatinine; SE, Standard error.

Finally, the 10 most common diagnoses in HES matching with the MINAP only AMI hospitalisations from the main analysis are shown in online supplemental table 12. Eighty-eight per cent of unmatched MINAP AMIs had a non-AMI HES hospitalisation within 30 days. These were mainly CVD or respiratory infection-related ICD-10 diagnoses.

DISCUSSION

We compared recording of AMI hospitalisations for people with CKD between two large secondary healthcare data sets in England. In a cohort of 6042 people, we found that both HES and MINAP missed a significant proportion of AMI hospitalisations. CKD stage influenced likelihood of AMI recording by data set: AMI hospitalisations in people with moderate to severe CKD were more likely to be recorded in MINAP compared with people at risk of CKD or with mild CKD. We found an association between AMI hospitalisation recording by data set and in-hospital mortality. There was marked variation between eGFR at AMI admission and preceding eGFR measurements in primary care, but no obvious systematic bias in terms of over/underestimation of eGFR at AMI admission.

Our results agree with previous research demonstrating incomplete capture of AMI events by individual healthcare data sets in the overall English population and extend them to a population with CKD. Herrett *et al*[12] showed 46% agreement when restricting to MINAP and HES recorded AMI hospitalisations, which is close to the 42% agreement we found in people with moderate to severe CKD. A smaller single-centre study by Torabi *et al*[29] found 32% agreement between MINAP and the hospital information department (responsible for HES coding).

In contrast to both studies, we found significantly worse agreement in case ascertainment for AMI hospitalisations between MINAP and HES for people at risk of CKD or with mild CKD. Torabi *et al*[29] collected data on renal function but found likelihood of AMI case recording in MINAP to reduce with advancing CKD stage. Differences in results between this and our study could be ascribed to changes in management of patients and/or event recording over time, differences in the populations studied and local practice in the single centre analysed by Torabi *et al*[29]

The high prevalence of CKD risk factors in people at risk of or with mild CKD could put them at greater risk of type 2 AMI; a mismatch of myocardial oxygen supply and demand in the absence of the 'classical' coronary artery plaque rupture with thrombosis reflective of type 1 AMI.[20] People with type 2 AMI are typically older, with a greater burden of comorbidities than those with type 1 AMI, and have poor outcomes.[30] HES is likely to include more type 2 AMI than MINAP as clinical coders for the latter are asked to select type 1 AMI only.

People with AMI recorded in both MINAP and HES had lower in-hospital mortality compared with those with AMI recorded in either MINAP or HES only. Our findings agree with Herrett *et al*[12]; patients with AMI recorded in only one source had a higher mortality than those with events recorded in more than one source. Higher in-hospital mortality in the MINAP only cases is likely to reflect the referral of severe and complex AMI cases to cardiology, including a higher STEMI to NSTEMI ratio.

Across all levels of eGFR, we found significant variation between eGFR stage derived from SCr taken within 24 hours of AMI admission (recorded in MINAP) and that derived from SCr in primary care, which is in line with reported variability of eGFR in validation studies.[23 24] As expected with known limitations of using MDRD eGFR to estimate kidney function for GFRs above 60 mL/min/1.73 m^2, binary classification between individuals with CKD stages 3–5 and those with stages 1–2 is more reliable than classification by CKD stage. These findings suggest that although previous research[13–15] using SCr at AMI admission recorded in MINAP as a proxy for baseline CKD stage may result in misclassification, it is unlikely to have resulted in a systematic bias in either overestimation or underestimation of CKD stage, despite our initial hypothesis that there would be systematic underestimation of kidney function due to the substantially increased risk of AKI during an AMI hospitalisation.[31] Differences between SCr recorded in primary care and SCr recorded in MINAP may reflect progression of CKD, differential use of medication that affects the renin–angiotension–aldosterone system, AKI at the time of serum sampling (although changes in SCr are unlikely to show within 24 hours of AMI onset), or variation around the mean.

Limitations

The NCKDA only included people with CKD and/or risk factors for CKD; therefore, we cannot generalise our results to people without risk factors for CKD. We may have incorrectly misclassified people who have no documented tests for CKD in primary care as having risk factors for CKD only; however, previous work has shown this group of people are much less likely to have CKD than those who do have CKD tests recorded in primary care.[22] Furthermore, we included people with at least one reduced kidney function test as potentially having CKD since not every patient undergoes regular CKD testing in our routine clinical data sets. Defining CKD using one eGFR measurement will have led to some misclassification. However, as people with CKD have very high cardiovascular risk and because of the infrequent SCr measurement in primary care, applying the chronicity criterion would have led to a selected cohort of people who did not develop a myocardial infarction until the second measurement had been done. Our results are likely impacted by residual confounding, since we were limited in the number of relevant comorbidities we could include in our multivariable models. Finally, AMI misclassification in HES data may have occurred due to the structure and level of detail available in this data set. For example, we may have missed AMI cases by including only those recorded in the first diagnostic position of the first episode of an HES admission; however, our sensitivity

analysis which included AMI hospitalisations recorded in the first or second diagnostic position showed similar results. In addition, unlike MINAP data, HES data do not include ECG results and troponin levels, which we could have used to reduce potential misclassification. Inclusion of the first diagnostic position of later episodes was undertaken in a similar study investigating AMI case ascertainment in people with malignancy, with little improvement in agreement between data sets.[32]

Future research

This study demonstrates how AMI case ascertainment in England can be improved by using linked healthcare data sets. Further research investigating cardiovascular and kidney disease incidence, prevalence and outcomes should follow this approach. Other countries with similarly rich, yet fragmented healthcare data sets would benefit from applying similar methods to evaluate the validity and completeness of cardiovascular and kidney disease capture in similar data. Optimising data quality in healthcare data sets and simplifying the process of data linkage would facilitate high-quality observational research to inform the design of future RCTs and provide estimated treatment effects where RCT data are lacking.

CONCLUSION

The use of linked healthcare data sets should be prioritised in observational research investigating multimorbidity.

Author affiliations
[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
[2]Population Health Sciences, University of Bristol, Bristol, UK
[3]Richard Bright Renal Service, North Bristol NHS Trust, Southmead Hospital, Bristol, UK
[4]Oxford Kidney Unit, Churchill Hospital, Oxford, UK
[5]Nuffield Department of Medicine, University of Oxford, Oxford, UK
[6]Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, UK
[7]National Institute for Cardiovascular Outcomes Research (NICOR), Barts Health NHS Trust, London, UK
[8]Institute of Cardiovascular Sciences, University College London, London, UK
[9]Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK
[10]Health Data Research UK, London, UK
[11]Glangwili General Hospital, Carmarthen, UK
[12]Department of Cardiovascular Sciences, University of Leicester and NIHR Leicester Biomedical Research Centre, Leicester, UK

Twitter Lucy Teece @LucyTeece

Acknowledgements MINAP is commissioned, as part of the 'National Cardiac Audit Project' by the Healthcare Quality Improvement Partnership. This work uses data that has been provided by patients and collected by the NHS as part of their care and support. We acknowledge and thank UK cardiovascular, renal and primary care physicians as well as hospital and community audit and coding teams whose diligent data collection has provided the data for these analyses. In particular, we would like to thank James Chal, Anil Gunesh, Andrew Harrison, and Akosua Donkor from NICOR; and Kathryn Griffith, Matthew Harker, Yvonne Silove, Tasneem Hoosain, Nick Wilson, Ronnie Moodley, Richard Fluck, Chris Gush, David Wheeler, Liam Smeeth, Ron Cullen, Andy Syme, Richard Gunn, Paul Wright, Hugh Gallagher, Sion Edwards, Fiona Loud, Nick Palmer, Richard Fluck, Anita Sharma, Kate Cheema, Andy Syme, Sally Hull, Ben Caplin, Lois Kim, and Faye Cleary from the NCKDA steering, clinical reference and statistical groups.

Contributors JS, DT, DN, DA, FC and UU conceived the study. JS, DT, DN, DA, FC, UU and PB designed the study. PB and DN had access to and analysed the data. PB and JS drafted the manuscript. PB, JS, DT, UU, FC, LT, MS, JD, MdB, SD, CW, DA and DN (all authors) critically appraised the results and edited the manuscript. All authors approved the final version of the manuscript. The lead authors confirm all authors meet ICJME authorship criteria, and no one who meets ICJME criteria have been excluded. PB and DN are the guarantors of this study, and accept full responsibility for the work and conduct of the study, had access to the data, and controlled the decision to publish.

Funding This work was supported by Kidney Research UK (grant number IN_008_20180304) and the Health Foundation (grant number 1725841). JS is a Doctoral Research Fellow funded by the NIHR (NIHR 300906). DA is funded by a joint research grant from the British Heart Foundation (SP/16/5/32415) and Cancer Research UK (C53325/A21134).

Competing interests All authors have completed an ICJME form. PB, JS, DT, FC, LT, MdB and SD have nothing to declare. UU declares a grant from the Health Foundation to undertake quality improvement unrelated to this work. MS declares funding from Cancer Research UK (C53325/A21134) and British Heart Foundation (SP/16/5/32415) for research activities related to this work looking at acute myocardial infarction ascertainment in a cancer population. JS declares Doctoral research funding from the NIHR for related research looking at acute myocardial infarction care for people with chronic kidney disease. JD declares grants from British Heart Foundation paid to his institution unrelated to this work. JD also declares consulting fees from Novo Nordisk, and honoraria from Amgen, Boehringer Ingelheim, Merck, Pfizer, Aegerion, Novartis, Sanofi, Takeda, Novo Nordisk and Bayer, unrelated to this work. JD is also member of a study steering committee with Novo Nordisk. CW declares he is clinical lead of the Myocardial Ischaemia National Audit Project. DA reports research funding and in-kind support from AstraZeneca for unrelated research and educational funding from Abbott Vascular to support a clinical research fellow doing unrelated research. DA has conducted consultancy for General Electric to support general research funds. DN is the UK Kidney Association Director of Informatics Research. DN is also on the steering group for two GlaxoSmithKline funded studies that investigate kidney function in children and adults in sub-Saharan Africa.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study was approved by the London School of Hygiene & Tropical Medicine LEO ethics committee (ID 16988) and the National Chronic Kidney Disease Audit (NCKDA) Steering Committee. Audits used in this study are covered by Section 251 approvals (NHS Act 2006) which allows data to be collected without individual patient consent for medical research when it is not possible to use anonymised information and when seeking consent is not practical.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Due to data sharing agreements, we are not able to share these data. Researchers interested in using these data should consult the websites for the NCKDA (https://www.lshtm.ac.uk/research/centres-projects-groups/ckdaudit#welcome) and MINAP (https://www.nicor.org.uk/national-cardiac-audit-programme/myocardial-ischaemia-minap-heart-attack-audit/).

ORCID iD
Patrick Bidulka http://orcid.org/0000-0001-7644-2030

REFERENCES

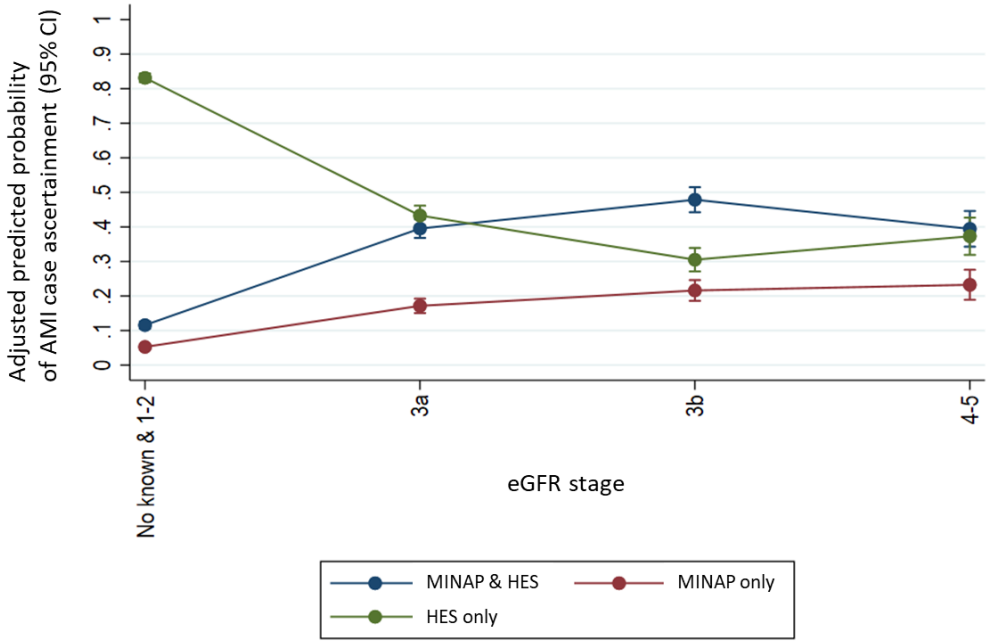
1 Foundation BH. Heart statistics, 2020. Available: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics [Accessed 28 Feb 2020].
2 Pitt B, Remme W, Zannad F, *et al*. Eplerenone, a selective aldosterone blocker, in patients with left ventricular dysfunction after myocardial infarction. *N Engl J Med* 2003;348:1309–21.
3 Januzzi JL, Cannon CP, DiBattiste PM, *et al*. Effects of renal insufficiency on early invasive management in patients with acute coronary syndromes (the TACTICS-TIMI 18 trial). *Am J Cardiol* 2002;90:1246–9.
4 Bhatt DL, Roe MT, Peterson ED, *et al*. Utilization of early invasive management strategies for high-risk patients with non-ST-segment elevation acute coronary syndromes: results from the crusade quality improvement initiative. *JAMA* 2004;292:2096–104.
5 Kedhi E, Fabris E, van der Ent M, *et al*. Six months versus 12 months dual antiplatelet therapy after drug-eluting stent implantation in ST-elevation myocardial infarction (DAPT-STEMI): randomised, multicentre, non-inferiority trial. *BMJ* 2018;363:k3793.
6 Serruys PW, Kogame N, Katagiri Y, *et al*. Clinical outcomes of state-of-the-art percutaneous coronary revascularisation in patients with three-vessel disease: two-year follow-up of the SYNTAX II study. *EuroIntervention* 2019;15:e244–52.
7 Wright RS, Reeder GS, Herzog CA, *et al*. Acute myocardial infarction and renal dysfunction: a high-risk combination. *Ann Intern Med* 2002;137:563–70.
8 Santopinto JJ, Fox KAA, Goldberg RJ, *et al*. Creatinine clearance and adverse hospital outcomes in patients with acute coronary syndromes: findings from the global registry of acute coronary events (grace). *Heart* 2003;89:1003–8.
9 Fox CS, Muntner P, Chen AY, *et al*. Use of evidence-based therapies in short-term outcomes of ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction in patients with chronic kidney disease: a report from the National cardiovascular data acute coronary treatment and intervention outcomes network registry. *Circulation* 2010;121:357–65.
10 Herzog CA, Ma JZ, Collins AJ. Poor long-term survival after acute myocardial infarction among patients on long-term dialysis. *N Engl J Med* 1998;339:799–805.
11 Konstantinidis I, Nadkarni GN, Yacoub R, *et al*. Representation of patients with kidney disease in trials of cardiovascular interventions: an updated systematic review. *JAMA Intern Med* 2016;176:121–4.
12 Herrett E, Shah AD, Boggon R, *et al*. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ* 2013;346:f2350.
13 Gupta T, Paul N, Kolte D, *et al*. Association of chronic renal insufficiency with in-hospital outcomes after percutaneous coronary intervention. *J Am Heart Assoc* 2015;4:e002069.
14 Rozenbaum Z, Benchetrit S, Minha S, *et al*. The effect of admission renal function on the treatment and outcome of patients with acute coronary syndrome. *Cardiorenal Med* 2017;7:169–78.
15 Shaw C, Nitsch D, Steenkamp R, *et al*. Inpatient coronary angiography and revascularisation following non-ST-elevation acute coronary syndrome in patients with renal impairment: a cohort study using the myocardial ischaemia national audit project. *PLoS One* 2014;9:e99925.
16 Grams ME, Astor BC, Bash LD, *et al*. Albuminuria and estimated glomerular filtration rate independently associate with acute kidney injury. *J Am Soc Nephrol* 2010;21:1757–64.
17 Wheeler D. *National chronic kidney disease audit - national report part 1*, 2017. https://www.lshtm.ac.uk/files/ckd_audit_report.pdf
18 Wheeler D, Hull S. *National chronic kidney disease audit: national report Part 2*, 2017.
19 Hospital episode statistics, 2020. Available: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics
20 Thygesen K, Alpert JS, Jaffe AS, *et al*. Fourth universal definition of myocardial infarction (2018). *Circulation* 2018;138:e618-e651.
21 Myocardial Ischaemia/MINAP (heart attack audit). Available: https://www.nicor.org.uk/national-cardiac-audit-programme/myocardial-ischaemia-minap-heart-attack-audit/
22 Iwagami M, Tomlinson LA, Mansfield KE, *et al*. Validity of estimated prevalence of decreased kidney function and renal replacement therapy from primary care electronic health records compared with national survey and registry data in the United Kingdom. *Nephrol Dial Transplant* 2017;32:ii142–50.
23 Levey AS, Coresh J, Greene T, *et al*. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med* 2006;145:247–54.
24 Levey AS, Stevens LA, Schmid CH, *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–12.
25 Stevens PE, Levin A, Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med* 2013;158:825–30.
26 Herbert A, Wijlaars L, Zylbersztejn A, *et al*. Data resource profile: Hospital episode statistics admitted patient care (hES APC). *Int J Epidemiol* 2017;46:1093–1093i.
27 Schnier C, Sudlow C. Definitions of Acute Myocardial Infarction and Main Myocardial Infarction Pathological Types - UK Biobank Phase 1 Outcomes Adjudication. Available: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf [Accessed 12 Jan 2021].
28 Giavarina D. Understanding Bland Altman analysis. *Biochem Med* 2015;25:141–51.
29 Torabi A, Cleland JGF, Sherwi N, *et al*. Influence of case definition on incidence and outcome of acute coronary syndromes. *Open Heart* 2016;3:e000487.
30 Baron T, Hambraeus K, Sundström J, *et al*. Type 2 myocardial infarction in clinical practice. *Heart* 2015;101:101–6.
31 Pickering JW, Blunt IRH, Than MP. Acute kidney injury and mortality prognosis in acute coronary syndrome patients: a meta-analysis. *Nephrology* 2018;23:237–46.
32 Coles B, Teece L, Weston C. Case-ascertainment of acute myocardial infarction hospitalizations in cancer patients: a cohort study using English linked electronic health data. *Eur Heart J Qual Care Clin Outcomes* 2021;8.

7.2. Relevance to my thesis

This research paper addresses objective 2A: to investigate potential selection and misclassification biases in defining a study population hospitalised for AMI with impaired kidney function in routine secondary care data (see also step 1 in **Figure 2.1**.). In this research, I highlight that AMI case ascertainment is incomplete using unlinked MINAP and HES data; however, people with moderate to severe kidney impairment are more likely to have AMI captured in both datasets compared with people with mild kidney impairment. Further, there is considerable variation between eGFR recorded within 24 hours of AMI hospitalisation in MINAP data versus eGFR recorded in primary care data; however, agreement between the two datasets when defining eGFR<60mL/min/1.73m^2 (i.e., moderate to severe kidney impairment) was good. In Appendices G.1. to G.3., I describe the codes and algorithms I used to define AMI, ACS subtypes, and covariates in the NCKDA-HES-MINAP linked data.

In Case Study 2 of this thesis, I aim to study the comparative effectiveness of alternative AMI treatments among people with impaired kidney function. The fragmentation of secondary care in England and the routinely collected data this generates creates challenges in defining the study population for this case study. In this published research paper, I show that linked secondary care data are important to minimise potential selection bias in future comparative effectiveness work. I used ORs and 95% CIs to summarise the differential AMI case ascertainment in MINAP and HES in the published paper, at the time of writing being less conscious of the issue of exaggerated ORs when describing a common outcome[1] (in this case, AMI case ascertainment in MINAP & HES, MINAP only, or HES only). For completeness of this thesis, I have added an unpublished figure describing the adjusted predicted probabilities of AMI case ascertainment by eGFR stage derived from the same multivariable multinomial logistic regression model for the main analysis (**Figure 7.1.**). This figure more intuitively demonstrates that the predicted probability of having an AMI captured in HES only is very high (83%) compared with AMI captured in MINAP only (5%) and in MINAP & HES (12%) for people with eGFR stages 1-2. Differences between the predicted probabilities of the three levels of AMI case ascertainment become less pronounced for people with eGFR stages 3a, 3b, and 4-5. In the discussion of the research paper, I hypothesise that this differential case ascertainment is because people with mild CKD may have a higher probability of experiencing a type 2 AMI while MINAP aims to include people with type 1 AMIs.[2, 3] These type 2 AMIs may be captured in HES data but not MINAP data, accounting for some of the increased probability of AMI being captured in HES only among people with mild CKD. I was unable to differentiate between type 1 and type 2 AMIs using HES and MINAP data to explore this hypothesis, and relied on the clinical input of my co-authors in formulating this theory.

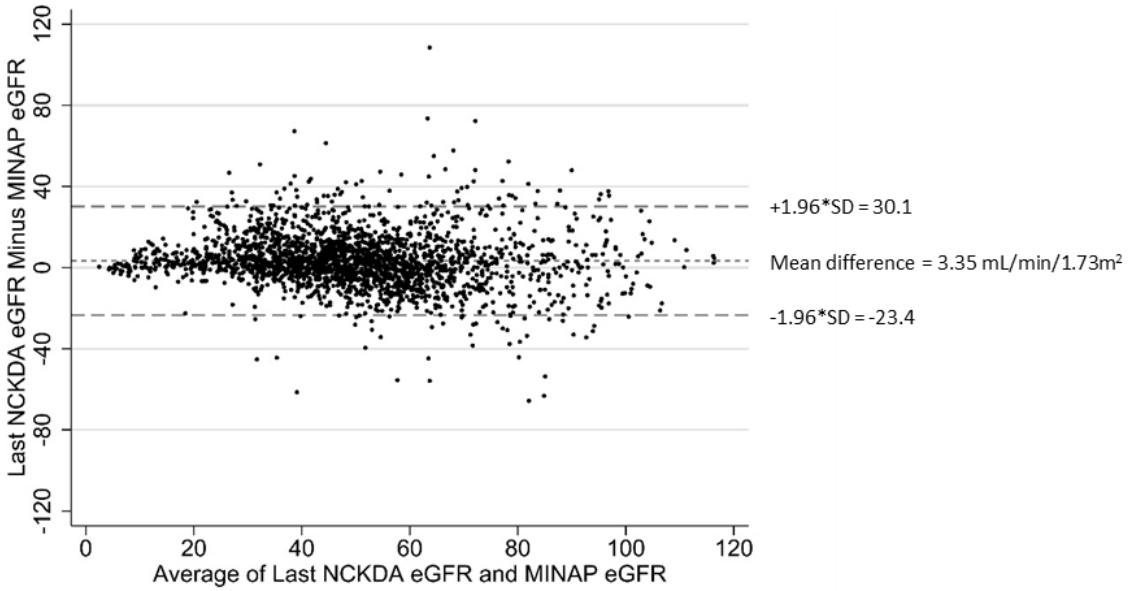
Figure 7.1. Adjusted predicted probabilities describing the association between eGFR stage and AMI case ascertainment (figure not included in the published *BMJ Open* paper presented in Chapter 7)



CI: confidence interval; eGFR: estimated glomerular filtration rate; HES: Hospital Episode Statistics; MINAP: Myocardial Ischaemia National Audit Project

In this research paper, I also highlight the differences in eGFR and eGFR staging between MINAP (secondary care) and the NCKDA (primary care). My hypothesis was that eGFR would be systematically underestimated in secondary care data due to acute decreases in kidney function in consequence of the AMI event. However, as shown in the Bland-Altman plot which I have copied here from the supplementary materials of the published paper (**Figure 7.2.**), there was no systematic under- or over-estimation in the secondary care versus primary care. I posit in the discussion of the main text this was because any acute decrease in eGFR was unlikely to be reflected within 24 hours of the AMI hospitalisation, which is the time window of entry for this data point in the MINAP dataset. When evaluating the agreement between eGFR stage derived from these eGFR measures in MINAP versus NCKDA data, I found that agreement was moderate to good when simplifying the categorisation to two categories (mild kidney impairment and moderate to severe kidney impairment). However, eGFR, like all other non-coded laboratory results, are not available in HES data. Thus, using MINAP eGFR to define a study population of people with moderate to severe kidney impairment will still suffer from substantial missingness in AMI hospitalisations, which is likely to impact any comparative effectiveness work via collider bias, a specific example of selection bias.

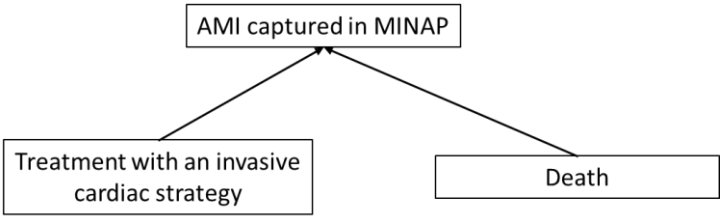
Figure 7.2. Bland-Altman plot comparing the mean eGFR (NCKDA eGFR and MINAP eGFR, x-axis) and the difference between the NCKDA eGFR and the MINAP eGFR (y-axis) (Supplementary figure 3 in the published BMJ Open research paper)



eGFR: estimated glomerular filtration rate, NCKDA: National Chronic Kidney Disease Audit, MINAP: Myocardial Ischaemia National Audit Project; SD: standard deviation

A collider is a variable which is caused by two variables. Conditioning the selection of a study population on a collider may induce a spurious association between the exposure/treatment and the outcome, a form of selection bias described as 'collider bias'.[4] In this case study, the collider would be AMI case ascertainment, i.e., the database in which the AMI hospitalisation is captured. If (i) alternative AMI treatment influences a person's AMI to be captured (or not) in MINAP versus HES data, and (ii) the outcome of interest, for example, death also influences a person's AMI to be captured (or not) in MINAP versus HES data, then selecting a study population conditioned on being captured in MINAP data is likely to result in a biased treatment effect. This collider bias can be illustrated with a DAG (**Figure 7.3.**).

Figure 7.3. Directed Acyclic Graph (DAG) illustrating potential collider bias when conditioning the selection of the study population on AMI capture in MINAP (unpublished)



AMI: acute myocardial infarction; MINAP: Myocardial Ischaemia National Audit Project

Thus, using linked primary and secondary care datasets is important in this case study to estimate causal treatment effects, not only to reduce measurement error and data missingness, but to also minimise the risk of selection bias induced by conditioning the study population on a collider.

In the next chapter, I will investigate the association between patient-level characteristics, namely kidney impairment, and alternative AMI treatment strategies using the same cohort.

7.3. References

1. Norton EC, Dowd BE, Maciejewski ML. Odds Ratios—Current Best Practice and Use. *Jama*. 2018;320(1):84-85. doi:10.1001/jama.2018.6971

2. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *European heart journal*. Aug 25 2018;doi:10.1093/eurheartj/ehy462

3. Herrett E, Smeeth L, Walker L, Weston C, on behalf of the MAG. The Myocardial Ischaemia National Audit Project (MINAP). *Heart (British Cardiac Society)*. 2010;96(16):1264. doi:10.1136/hrt.2009.192328

4. Holmberg MJ, Andersen LW. Collider Bias. *Jama*. 2022;327(13):1282-1283. doi:10.1001/jama.2022.1820

CHAPTER 8. RESEARCH PAPER – MANAGEMENT AND OUTCOMES OF MYOCARDIAL INFARCTION IN PEOPLE WITH IMPAIRED KIDNEY FUNCTION IN ENGLAND.

OVERVIEW

In this chapter, I present a research paper published in *BMC Nephrology* for which I am joint first author. In this study, I use the same cohort as in Chapter 7 to investigate the association between eGFR stage and invasive cardiac treatment strategy for AMI, as well as AMI outcomes like death (in-hospital and post-discharge) and AMI re-admission. These analyses aimed to describe disparities in AMI treatment strategies by kidney function at the individual-level, overall and by AMI subtype (STEMI and NSTEMI). Following the research paper, I include a brief discussion of this paper's relevance to my thesis. This discussion includes a key table from the supplementary materials of the published paper which aids in the interpretation of the study in the context of this thesis.

8.1. Published research paper

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMC Nephrology		
When was the work published?	02/11/2023		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I am joint first author of this paper. I co-led the study design with JS (joint first author), DA, and DN. I was solely responsible for data extraction, data management, and data analysis. I led the interpretation of results with support from my co-authors. I co-wrote the first draft of the manuscript with JS. Co-authors reviewed multiple drafts before sending to the journal for publication. I co-led the responses to peer review comments with JS with input from my co-authors.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Dorothea Nitsch
Date	04/04/2024

BMC Nephrology

RESEARCH Open Access

Management and outcomes of myocardial infarction in people with impaired kidney function in England

Jemima Scott[1,2*†], Patrick Bidulka[3†], Dominic M. Taylor[1,2], Udaya Udayaraj[4,5], Fergus J. Caskey[1,2], Kate Birnie[1], John Deanfield[6,7], Mark de Belder[6], Spiros Denaxas[8,9], Clive Weston[10], David Adlam[11†] and Dorothea Nitsch[3†]

Abstract

Background Acute myocardial infarction (AMI) causes significant mortality and morbidity in people with impaired kidney function. Previous observational research has demonstrated reduced use of invasive management strategies and inferior outcomes in this population. Studies from the USA have suggested that disparities in care have reduced over time. It is unclear whether these findings extend to Europe and the UK.

Methods Linked data from four national healthcare datasets were used to investigate management and outcomes of AMI by estimated glomerular filtration rate (eGFR) category in England. Multivariable logistic and Cox regression models compared management strategies and outcomes by eGFR category among people with kidney impairment hospitalised for AMI between 2015–2017.

Results In a cohort of 5 835 people, we found reduced odds of invasive management in people with eGFR < 60mls/min/1.73m^2 compared with people with eGFR ≥ 60 when hospitalised for non-ST segment elevation MI (NSTEMI). The association between eGFR and odds of invasive management for ST-elevation MI (STEMI) varied depending on the availability of percutaneous coronary intervention. A graded association between mortality and eGFR category was demonstrated both in-hospital and after discharge for all people.

Conclusions In England, patients with reduced eGFR are less likely to receive invasive management compared to those with preserved eGFR. Disparities in care may however be decreasing over time, with the least difference seen in patients with STEMI managed via the primary percutaneous coronary intervention pathway. Reduced eGFR continues to be associated with worse outcomes after AMI.

Keywords CKD, Coronary angiography, Myocardial infarction, Percutaneous coronary intervention, Survival analysis

† Jemima Scott, Patrick Bidulka, David Adlam and Dorothea Nitsch equal contribution/joint-authorship.

*Correspondence:
Jemima Scott
Jemima.scott@bristol.ac.uk
Full list of author information is available at the end of the article

Introduction

The prevalence of moderate to severe chronic kidney disease (CKD) is 11–13% in the general population [1], but 40% in those with acute myocardial infarction (AMI) [2–4]. The increased risk of AMI with CKD results from higher prevalence of traditional cardiovascular risk factors in addition to risk factors specific to CKD pathophysiology [5]. In the United States of America (USA) and Europe, in-hospital and post-discharge mortality following AMI is higher in the CKD population [6, 7]. Progressively worse outcomes are seen with increasing CKD severity [2].

Knowledge of the optimal management of AMI in people with CKD lags behind our understanding within the general population. People with CKD have been excluded from most randomized-controlled trials (RCTs) that have driven forward advances in AMI management over the past 50 years [8]. Evidence in this population is limited to observational data and small subgroup analyses of people with mild to moderate CKD who were included within relevant RCTs [9–12]. The accumulating evidence suggests that people with reduced estimated glomerular filtration rate (eGFR) do benefit from invasive management of AMI, despite increased risk of complications and poorer outcomes compared to those with normal kidney function [13–15]. Current USA and European cardiology guidelines now recommend invasive management for high-risk non-ST-elevation myocardial infarction (NSTEMI) and all ST-elevation myocardial infarction (STEMI) events independent of eGFR [16, 17].

Previous observational research has shown that people with reduced eGFR receive less aggressive AMI management than those with normal kidney function [3, 10, 11, 14, 18]. Despite evidence that disparities in care are falling over time, contemporary studies from the USA continue to demonstrate reduced rates of invasive management in patients with low eGFR [6, 19, 20]. Since the routine introduction of primary percutaneous coronary intervention (PCI) for STEMI in 2009, no studies from Europe or the UK have examined how this may have affected care for patients with reduced eGFR [21]. In England, the latest study relating to AMI in patients with kidney disease used data from 2004–2008 and was limited to NSTEMI [13]. To our knowledge, the association between eGFR and treatment of STEMI has not previously been examined using English data.

Prior research regarding AMI care for people with reduced eGFR in England is limited because granular data on eGFR and AMI are held on distinct healthcare registries. Reliance on a single dataset risks biasing results via misclassification of disease status. To maximise the utility of available data, we used multiple linked healthcare registries to provide an updated description of AMI care and outcomes for people with reduced eGFR in England. Dataset linkage has allowed a) the use of pre-admission rather than in-hospital creatinine readings to define eGFR category, giving a more accurate reflection of baseline kidney function [22], and b) the identification of AMI hospitalisations from more than one database, optimizing sensitivity in identifying these events [22, 23].

Materials and methods

Study design and data sources

This historical cohort study used data from clinical audit and routinely collected health records. Our cohort was defined using primary care data from the National Chronic Kidney Disease Audit (NCKDA) research database, and secondary care data from the Myocardial Ischaemia National Audit Project (MINAP) and Hospital Episode Statistics (HES) Admitted Patient Care (APC).

The NCKDA [24, 25] aimed to audit and improve primary care health services in England and Wales for people with CKD or CKD risk factors (Additional Table 1). The audit included 10% of English General Practices (approximately 1.7 million people with CKD or risk factors for CKD) who invested in audit software and volunteered to participate in the audit, and is now used as a research database to study long-term outcomes of this population [22, 24, 25]. The NCKDA collected complete historical patient-level data for eligible participants from general practices in two main cross-sectional extracts between 2014 and 2016. People who died between extracts, opted-out of data-sharing (person or practice-level), or people who changed GPs were excluded in the second extract. People included in the NCKDA were generally representative of the English population in terms of age and sex [24].

HES Admitted Patient Care data are collected to compensate hospitals for services provided by the NHS. In England, all hospitalisations funded by the NHS (approximately 99%) are captured by HES data [26]. Diagnoses during hospitalisation are recorded using International Classification of Diseases 10th Edition (ICD-10) codes.

MINAP, part of the National Institute of Cardiovascular Outcomes Research (NICOR) audit and research programme, aims to audit all type 1 AMIs admitted to hospitals in England and Wales. Data are collected on patient characteristics, laboratory tests, comorbidities, processes of care, and treatment received during AMI hospitalisation [27, 28].

Office of National Statistics (ONS) data were linked to these primary and secondary care data to determine death dates [29] (Additional Table 2). The Index of Multiple Deprivation (IMD) patient-level data were linked as a proxy for socioeconomic status (SES) [30].

HES and MINAP-linked data were available up to 31 March 2017. Anyone with an AMI hospitalisation between the final NCKDA extract and the end of HES/MINAP linked data were included in the cohort. The ONS linked follow-up death data were available up to 15 September 2019.

Study participants

We included people captured by the NCKDA research database with an AMI hospitalisation recorded in MINAP, HES, or both between 2015–2017, after the final NCKDA cross-sectional extract in which the person appeared [22]. We identified incident AMI hospitalisations and AMI subtypes (STEMI, NSTEMI) in HES using ICD-10 codes (Additional Table 3) recorded in the first diagnostic position of the first episode of the spell, and in MINAP using an algorithm which uses discharge diagnosis, cardiac marker levels, and electrocardiogram results (Additional Table 4). People with CKD risk factors, but no eGFR in the primary care record ($n=118$), were excluded.

Exposures

We calculated the baseline eGFR from the most recent serum creatinine value recorded in primary care prior to the index AMI hospitalisation using the MDRD equation [31]. We defined eGFR categories using the same cut-points KDIGO recommends for the definition of CKD stages: Category 1–2 (eGFR 60-120 mL/min/1.73m^2), 3a (eGFR 45–59), 3b (eGFR 30–44), and 4–5 (eGFR 0–29) [32].

Outcomes

Our primary outcomes were all-cause death during the first AMI hospitalisation recorded during the study period (the index AMI hospitalisation), and all-cause death during follow-up, for those who survived the index AMI hospitalisation. Variables used to define death date are described in Additional Table 5. Secondary outcomes were treatments received during hospitalisation: (1) Angiography and/or percutaneous coronary intervention (PCI), and (2) coronary artery bypass graft (CABG) (Additional Table 6). Other secondary outcomes, among survivors of the index AMI hospitalisation, were AMI re-admission and cardiovascular-specific death post-index AMI discharge.

Covariables

Potential confounding variables available in our dataset were age at AMI hospitalisation (continuous), sex, ethnicity (white, other), IMD quintile, smoking status (non-smoker, ever smoker), receipt of dialysis or kidney transplant, prior AMI, and comorbidities including chronic obstructive pulmonary disease (COPD), type 2 diabetes mellitus (T2DM), heart failure, unstable angina, cerebrovascular disease, hypertension, and peripheral vascular disease. We defined these covariates using a combination of primary and secondary care data (Additional Table 7) [22]. We categorised each hospital centre which contributed patient-level data to this study into two main categories: (1) PCI always available and (2) PCI services not always available (Additional Table 8).

Data analysis

We described baseline characteristics of the study population stratified by eGFR category. We used multivariable logistic regression to estimate the adjusted odds ratios comparing the odds of death during the index AMI hospitalisation (primary outcome) and the odds of invasive management (angiography and/or PCI, coronary artery bypass graft (CABG)) across eGFR categories. We also calculated predicted percentages from the adjusted logistic regression models using recycled predictions, since odds ratios can be misleading when the outcome is common [33]. We looked at these associations in the overall study population and stratified by AMI subtype (STEMI and NSTEMI). We tested for a linear trend in the association between eGFR category and the odds of receiving angiography and/or PCI using a likelihood ratio test.

We used Cox regression to investigate the association between eGFR category and outcomes post-index AMI hospitalisation among survivors, including all-cause mortality (primary outcome), cardiovascular-specific mortality, and AMI re-admission, after confirming the proportional hazard assumption using a global test on the Schoenfeld residuals over time. We first calculated crude rates for each outcome stratified by eGFR stage by dividing the number of outcome events by the total person-time study participants contributed following discharge from the index AMI hospitalisation. We reported these crude rates per 100 person-years. In our multivariable models, we specified a priori to adjust for age (continuous), sex, ethnicity, IMD quintile, COPD, T2DM, heart failure, and prior AMI as we anticipated these to be the most important confounders for this study population.

Secondary/sensitivity analyses

We repeated the main analyses, stratifying by (1) centre type, to understand the impact of PCI availability on the association between eGFR category and the odds of receiving angiography and/or PCI; (2) and relevant comorbidities (prevalent T2DM and heart failure), since it is possible people with these comorbidities experience different management and outcomes compared with those without. We also repeated all main analyses after

excluding people with prior AMI ($n=1,883$) as previous coronary intervention may impact subsequent care.

Missing data

We conducted a complete case analysis, excluding people with missing ethnicity and/or IMD data ($n=107$). Discharge dates were missing in 19% of MINAP and 1% of HES records. We imputed missing discharge dates using the median number of days in-hospital from non-missing records (5 and 4 days in MINAP and HES, respectively) [22].

Patient and public involvement

This study benefited from similar patient and public involvement as described in a related study [22]. The creation and maintenance of the NCKDA research database, including its record linkages and necessary section 251 permissions benefited from the support of the Kidney Care UK patient organisation (https://www.kidneycareuk.org/). Feedback from patient members of the UK Renal Registry Patient Council (https://renal.org/patients/patient-council) supported a further planned record linkage of renal and cardiac data.

Results

Study population and baseline characteristics

A total of 5 835 individuals who were included in the NCKDA and experienced at least one incident AMI hospitalisation captured in HES and/or MINAP were included in this study (Fig. 1). The median time between the most recent eGFR recorded in primary care and the index AMI hospitalisation was 0.97 years (interquartile range (IQR) 0.60 to 1.63) (Additional Fig. 1).

Of the 5 835 people hospitalised for AMI during the study period, 2,260 (39%) had eGFR category 3–5 as their latest primary care record of kidney function (Table 1). People with eGFR category 3b were oldest on average (82 years) and had the highest proportion of females (50%) compared with other eGFR categories. The most prevalent comorbidity was hypertension (60% overall), followed by type two diabetes mellitus (34%), and angina (25%). People with incomplete covariate data ($n=107$) are described in Additional Table 9.

Death during AMI hospitalisation

Overall, 907 people (16%) died during the index AMI hospitalisation. The crude proportion who died was greatest among people in eGFR categories 4–5 (27%) and lowest among people in category 1–2 (11%) (Table 2).

After adjustment, we found that people in eGFR categories 3a, 3b and 4–5 had greater odds of death compared with people in category 1–2 (adjusted OR 1.28 (95% CI 1.06–1.54), 1.51 (95% CI 1.22–1.87), and 1.80 (95% CI 1.37–2.38), respectively) (Table 2). When stratifying by AMI subtype, we found similarly increased odds of death during AMI hospitalisation for people in eGFR categories 3a, 3b and 4–5 compared with category 1–2 among people with NSTEMI and STEMI (although 95% CI overlap the null estimate when comparing eGFR category 3a and 4–5 with category 1–2 for the STEMI subgroup). The predicted percents for death during AMI hospitalisation also showed a higher percent of people dying with eGFR stages 3b and 4–5 compared with eGFR stages 1–2 (Additional Table 10).

Death post-AMI hospitalisation

Among people who survived their first AMI hospitalisation, we observed increasing rates of subsequent death with worsening baseline kidney function, during a mean follow-up of 2.4 years. People in eGFR category 1–2 had a crude rate of death post-AMI hospitalisation of 8.30 per 100 person-years (PY) (95% CI 7.70–8.95), while people in category 4–5 had a crude rate of death

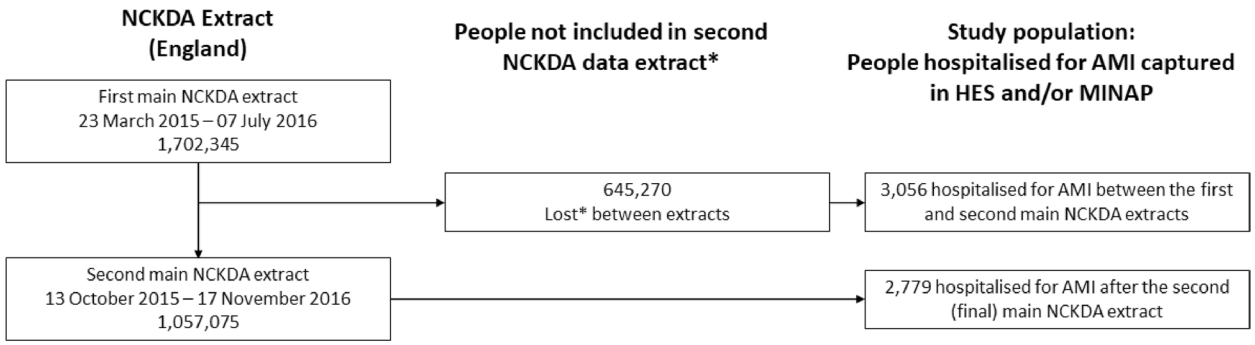


Fig. 1 Flow diagram illustrating the study population hospitalised for AMI derived from the NCKDA research database

Table 1 Baseline characteristics of people included in the NCKDA with an AMI hospitalisation recorded in MINAP, HES, or both

eGFR category	1–2	3a	3b	4–5	Total
	N = 3,574	*N* = 1,193	*N* = 721	*N* = 347	5,835
Age (years) at AMI admission, mean (SD)	70 (13)	79 (10)	82 (9)	79 (12)	74 (13)
Age group at AMI admission, years					
< 50	249 (7)	15 (1)	1 (0)	9 (3)	274 (5)
50–64	976 (27)	83 (7)	36 (5)	35 (10)	1,130 (19)
65–79	1,443 (40)	477 (40)	213 (30)	104 (30)	2,237 (38)
80 +	906 (25)	618 (52)	471 (65)	199 (57)	2,194 (38)
Female	1,226 (34)	564 (47)	361 (50)	142 (41)	2,293 (39)
Ethnicity					
White	3,288 (92)	1,131 (95)	686 (95)	317 (91)	5,422 (93)
Other	286 (8)	62 (5)	35 (5)	30 (9)	413 (7)
IMD quintile					
1 (least deprived)	631 (18)	226 (19)	136 (19)	59 (17)	1,052 (18)
2	736 (21)	274 (23)	154 (21)	68 (20)	1,232 (21)
3	793 (22)	281 (24)	167 (23)	72 (21)	1,313 (23)
4	822 (23)	229 (19)	159 (22)	83 (24)	1,293 (22)
5 (most deprived)	592 (17)	183 (15)	105 (15)	65 (19)	945 (16)
History of dialysis in primary care data					
Peritoneal dialysis	0 (0)	0 (0)	0 (0)	10 (3)	10 (0)
Haemodialysis	0 (0)	0 (0)	0 (0)	21 (6)	21 (0)
Renal dialysis, unspecified	0 (0)	0 (0)	0 (0)	7 (2)	7 (0)
History of kidney transplant in primary care data	2 (0)	0 (0)	3 (0)	15 (4)	20 (0)
Comorbidities					
Angina	770 (22)	332 (28)	228 (32)	134 (39)	1,464 (25)
Cerebrovascular disease	319 (9)	159 (13)	126 (17)	68 (20)	672 (12)
COPD	422 (12)	184 (15)	147 (20)	48 (14)	801 (14)
Type 2 diabetes mellitus	1,086 (30)	408 (34)	307 (43)	205 (59)	2,006 (34)
Heart failure	287 (8)	191 (16)	174 (24)	103 (30)	755 (13)
Hypertension	1,942 (54)	765 (64)	509 (71)	276 (80)	3,492 (60)
History of acute myocardial infarction	650 (18)	314 (26)	205 (28)	128 (37)	1,297 (22)
Peripheral vascular disease	179 (5)	90 (8)	59 (8)	39 (11)	367 (6)
Smoking status					
Non-smoker	1,715 (48)	575 (48)	345 (48)	170 (49)	2,805 (48)
Ever-smoker	1,859 (52)	618 (52)	376 (52)	177 (51)	3,030 (52)

n = (col %) unless specified otherwise

of 54.33 per 100 PY (95% CI 47.04–62.76) (Table 3). There was no evidence that the hazards were not proportional over time ($p = 0.18$).

After adjusting for pre-specified confounders, we observed increased hazards of death among people in eGFR categories 3b and 4–5, compared with people in category 1–2 (adjusted HR 1.40 (95% CI 1.21–1.62) and 2.57 (95% CI 2.16–3.05), respectively). Hazards were similarly greater among people in eGFR categories 3b and 4–5, compared with people in category 1–2 when stratifying by AMI subtype (Table 3).

Processes of care during AMI hospitalisation

Overall, the crude proportion of people who received angiography and/or PCI during their first AMI hospitalisation in the study period ranged from 64% among people in eGFR category 1–2 to 33% among people in category 4–5 (Table 4).

In our adjusted analysis, we observed that people in eGFR categories 3b and 4–5 had lower odds of receiving angiography and/or PCI compared with people in category 1–2 (adjusted OR 0.76 (95% CI 0.63–0.92) and 0.55 (95% CI 0.42–0.71), respectively).

Table 2 Death during the index AMI hospitalisation, overall and stratified by AMI subtype

Subgroup	eGFR category	Deaths, n (row %)	Total number of people	Age and sex adjusted, OR (95% CI)	Adjusted[a], OR (95% CI)
Overall	1–2	399 (11)	3,574	1	1
	3a	231 (19)	1,193	1.30 (1.08–1.56)	1.28 (1.06–1.54)
	3b	183 (25)	721	1.60 (1.30–1.97)	1.51 (1.22–1.87)
	4–5	94 (27)	347	2.00 (1.53–2.62)	1.80 (1.37–2.38)
AMI subtype					
STEMI	1–2	174 (13)	1,339	1	1
	3a	75 (24)	314	1.34 (0.97–1.86)	1.26 (0.91–1.76)
	3b	55 (35)	159	1.81 (1.20–2.71)	1.63 (1.09–2.43)
	4–5	22 (31)	70	1.93 (1.08–3.46)	1.31 (0.73–2.38)
NSTEMI	1–2	225 (10)	2,235	1	1
	3a	156 (18)	879	1.35 (1.07–1.70)	1.33 (1.06–1.68)
	3b	128 (23)	562	1.64 (1.27–2.13)	1.56 (1.20–2.02)
	4–5	72 (26)	277	2.20 (1.58–3.07)	2.07 (1.50–2.86)

[a] Adjusted for age (continuous), sex, IMD quintile, ethnicity (white or other), and history of T2DM, heart failure, COPD, and previous AMI

Table 3 Death post-AMI hospitalisation, among people who survive the index AMI hospitalisation, overall and stratified by AMI subtype

Subgroup	eGFR category	Deaths, *n* =	Rate per 100 person-years	Age and sex adjusted, Hazard Ratio (HR) (95% CI)	Adjusted[a], HR (95% CI)
Overall	1–2	675	8.30 (7.70–8.95)	1	1
	3a	359	16.94 (15.23–18.78)	1.19 (1.04–1.36)	1.10 (0.96–1.26)
	3b	298	30.37 (27.11–34.02)	1.63 (1.41–1.88)	1.40 (1.21–1.62)
	4–5	185	54.33 (47.04–62.76)	3.12 (2.64–3.69)	2.57 (2.16–3.05)
AMI subtype					
STEMI	1–2	187	6.08 (5.26–7.01)	1	1
	3a	71	13.37 (10.59–16.87)	1.18 (0.89–1.56)	1.20 (0.90–1.59)
	3b	50	26.22 (19.88–34.60)	1.52 (1.09–2.12)	1.33 (0.94–1.86)
	4–5	29	45.82 (31.84–65.94)	3.56 (2.38–5.32)	3.47 (2.29–5.26)
NSTEMI	1–2	488	9.65 (8.83–10.55)	1	1
	3a	288	18.13 (16.15–20.35)	1.17 (1.01–1.36)	1.08 (0.93–1.26)
	3b	248	31.37 (27.70–35.53)	1.62 (1.38–1.90)	1.41 (1.20–1.65)
	4–5	156	56.28 (48.10–65.84)	2.99 (2.48–3.60)	2.49 (2.06–3.02)

[a] Adjusted for age (continuous), sex, IMD quintile, ethnicity (white or other), and history of T2DM, heart failure, COPD, and previous AMI

The predicted percents also showed a lower percentage of people receiving angiography and/or PCI with eGFR stages 3b and 4–5 compared with eGFR stages 1–2 (Additional Table 10). When stratifying by AMI subtype, the association persisted among people with NSTEMI. People in eGFR category 4–5 had lower odds of receiving angiography and/or PCI compared with people in category 1–2 (adjusted OR 0.54 (95% CI 0.40–0.73). There was no evidence for a trend in association ($p = 0.32$) (Table 4).

We did not see evidence of an association between eGFR and receiving CABG after adjusting for pre-specified confounders, although our analyses were limited by low numbers of CABG recipients (Table 4).

Other outcomes

The crude rate of CVD-specific death ranged from 3.74 per 100 PY (95% CI 3.34–4.18) in people in eGFR category 1–2 to 23.20 per 100 PY (95% CI 18.61–28.93) in people in category 4–5 among those discharged alive from the index AMI hospitalisation. There was evidence of increased hazards of CVD-specific death among people in eGFR categories 3b and 4–5 versus people in category 1–2 (Additional Table 11).

Table 4 Processes of care (angiography and/or PCI) associated with eGFR at baseline during the index AMI hospitalisation, overall and stratified by AMI subtype

Outcome/Subgroup	eGFR category	N[b] (row %)	Total number of people	Age and sex adjusted, OR (95% CI)	Adjusted[a], OR (95% CI)
Angiography and/or PCI					
Overall	1–2	2,280 (64)	3,574	1	1
	3a	602 (50)	1,193	1.01 (0.87–1.17)	1.08 (0.93–1.26)
	3b	259 (36)	721	0.66 (0.55–0.79)	0.76 (0.63–0.92)
	4–5	117 (33)	347	0.44 (0.34–0.57)	0.55 (0.42–0.71)
AMI subtype					
STEMI	1–2	1,106 (83)	1,339	1	1
	3a	223 (71)	314	0.96 (0.71–1.30)	1.05 (0.76–1.44)
	3b	90 (57)	159	0.68 (0.46–1.00)	0.78 (0.53–1.15)
	4–5	41 (59)	70	0.53 (0.29–0.96)	0.86 (0.47–1.55)
NSTEMI	1–2	1,174 (53)	2,235	1	1
	3a	379 (43)	879	1.15 (0.97–1.36)	1.21 (1.01–1.44)
	3b	169 (30)	562	0.74 (0.60–0.91)	0.82 (0.66–1.02)
	4–5	76 (27)	277	0.50 (0.37–0.67)	0.54 (0.40–0.73)
CABG					
Overall	1–2	83 (2)	3,574	1	1
	3a	33 (3)	1,193	1.68 (1.08–2.59)	1.64 (1.05–2.55)
	3b	9 (1)	721	0.84 (0.41–1.73)	0.77 (0.37–1.60)
	4–5	6 (1)	347	0.99 (0.43–2.32)	0.80 (0.34–1.91)

[a] Adjusted for age (continuous), sex, IMD quintile, ethnicity (white or other), and history of T2DM, heart failure, COPD, and previous AMI

[b] n Is the number of people receiving angiography and/or PCI

Crude rates of AMI re-hospitalisation among people discharged alive after their first AMI hospitalisation in the study period ranged from 22.50 per 100 PY (95% CI 20.15–25.12) among people in eGFR category 3b to 45.04 per 100 PY (95% CI 32.77–61.89) among people in category 4–5. There was no evidence of increased hazards of AMI re-hospitalisation with worsening eGFR after adjustment for potential confounders (Additional Table 11).

Secondary/sensitivity analyses

When stratifying by availability of PCI services, we observed an attenuation of the relative odds of death during the index AMI hospitalisation for people in eGFR categories 3a (adjusted OR 0.84, 95% CI 0.58–1.22) and 3b (adjusted OR 1.23 (95% CI 0.82–1.85)) versus people in category 1–2 in centres where PCI is always available (Additional Table 12). Relative hazards of death post-AMI hospitalisation among people who survived were similarly greater for people with worsening eGFR when stratifying by PCI service availability (Additional Table 13).

The odds of receiving angiography and/or PCI during the index AMI hospitalisation were similarly lower among people in eGFR categories 3b and 4–5 compared with people in category 1–2 both in centres with and without constant PCI availability. When restricting to STEMI hospitalisations, there was no association between eGFR category and odds of receiving angiography and/or PCI in the centres where PCI is always available. However, there were lower odds of people in eGFR categories 3a, 3b and 4–5 receiving angiography and/or PCI compared with people in category 1–2 when restricting to centres where PCI is available sometimes or not at all (Additional Table 14).

We observed no substantial changes to our results when excluding people with a history of AMI (Additional Table 15), nor when stratifying by prevalent T2DM status (Additional Figs. 3 and 4). However, there was no evidence of an association between eGFR category and receipt of angiography and/or PCI among people with recorded prevalent heart failure (Additional Fig. 2).

Discussion

In this analysis of 5 835 AMI hospitalisations from linked primary and secondary care multi-disease registries in England, odds of death both in-hospital and post-discharge were significantly higher in people with a pre-admission eGFR < 60mls/min/1.73m^2, compared to those with an eGFR ≥ 60. We demonstrated a progressive

reduction in the odds of receiving angiography and/or PCI for NSTEMI with reducing eGFR, independent of the availability of PCI services. In contrast, in people with STEMI, we found no association between eGFR category and the odds of invasive management in the population overall. In centres where PCI was not always available however, reduced use of angiography and PCI extended to those with STEMI, suggesting an opportunity to improve outcomes of patients with impaired renal function by better access to specialised centers with primary PCI services.

Reducing eGFR was associated with a progressive increase in the odds of death following all AMI events both within hospital and post-discharge. Inferior mortality outcomes have been reported previously amongst people with kidney disease, with the poorest survival in those with the lowest eGFRs [10, 14, 18]. People with low eGFR have a higher baseline mortality risk prior to AMI, and are more likely to experience complications relating to both the AMI event and its treatment [2, 6, 7]. Increased deaths amongst people with reduced eGFR could be due to residual confounding from severity of comorbidities or unmeasured factors such as frailty. Reduced invasive AMI management has also been suggested to contribute to these worse outcomes [13, 14, 34, 35].

We demonstrated an inverse association between eGFR category and the odds of invasive management after NSTEMI. Reduced use of angiography and revascularisation in people with kidney impairment has been described previously, and may relate to concerns about contrast-induced nephropathy and bleeding risks, or therapeutic nihilism [3, 35, 36]. A study from the USA has however shown narrowing of this treatment gap, with the greatest increase in use of invasive management in those with the worst kidney function [6]. Comparison of our data with that from a study of NSTEMI management in England in 2004–2008 suggests a similar relative increase in the use of angiography in the lowest eGFR categories [13].

We found that reduced eGFR is associated with lower odds of invasive management in people with NSTEMI but not in those with STEMI. Possible explanations for these differences include a) PCI in STEMI is time-critical and clinicians may not have time to review blood results and/or b) clinicians may consider the benefits of PCI in STEMI to outweigh the risks posed to kidney function. It is possible that our small sample size underlies this lack of association. A large study of AMI management and outcomes in the USA demonstrated reduced use of angiography after STEMI in people with CKD in 2007–8, with attenuation of these differences in 2014–15, following the routine introduction of primary PCI [6]. This correlates with our findings of an association between eGFR and the receipt of angiography which is limited to centres that do not always offer PCI. Reduced use of angiography in these centres may reflect reluctance by clinicians to intervene in frail and complex patients, or transfer to centres offering primary PCI. The lack of association between eGFR and odds of invasive management after AMI (any) in people with heart failure may simply reflect poor diagnosis and recording of heart failure amongst patients with kidney disease [37].

There are some limitations to consider. First, although MINAP is designed as an audit of type one AMI, we were unable to exclude type two AMIs from our analyses. These events may occur more frequently in people with low eGFR than without [14]. Similarly, we may have assigned incorrect eGFR categories to patients experiencing AKI either prior to, or at the time of, admission with AMI. Secondly, we were unable to risk stratify our AMI cohort. Reduced eGFR is associated with greater cardiac risk however, so differences in risk are unlikely to explain our findings. Thirdly, the competing risk of death may bias our effect estimates when investigating receipt of invasive management. The number of people dying within the decision-making timeframe are, however, likely to be small. Fourth, residual confounding is likely to affect our results, for example severity of comorbidities and pharmacological management. Fifth, we acknowledge the study population is selected from 10% of GPs in England who self-selected to take part in the NCKDA, and may not be representative of the English population in terms of ethnicity and standard of primary care [24]. Finally, overestimation of baseline kidney function is also possible, since the median time between the most recent serum creatinine test and the index AMI hospitalisation was approximately one year during which time kidney function may have worsened.

This study adds evidence from England to existing international research demonstrating disparities in AMI care and outcomes between those with and without reduced eGFR. Further research is needed to understand why eGFR influences receipt of AMI management and explore whether differences in AMI care represent appropriate risk stratification of people with reduced eGFR, or inequitable access to effective management. Understanding these treatment disparities will enable interventions to be appropriately allocated to optimize care and outcomes for the growing global CKD population.

Abbreviations

AMI	Acute myocardial infarction
CABG	Coronary artery bypass graft
CKD	Chronic kidney disease
COPD	Chronic obstructive pulmonary disease
eGFR	Estimated glomerular filtration rate

HES-APC	Hospital Episode Statistics—Admitted Patient Care
ICD-10	International Statistical Classification of Diseases and Related Health Problems—10th Revision
IMD	Indices of multiple deprivation
KDIGO	Kidney Disease Improving Global Outcomes
MDRD	Modification of Diet in Renal Disease
MINAP	Myocardial Ischaemia National Audit Project
NCKDA	National Chronic Kidney Disease Audit
NICOR	National Institute for Cardiovascular Outcomes Research
NSTEMI	Non ST-elevation myocardial infarction
ONS	Office for National Statistics
PCI	Percutaneous coronary intervention
RCT	Randomised controlled trial
STEMI	ST-elevation myocardial infarction
T2DM	Type two diabetes mellitus

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12882-023-03377-x.

Additional file 1: Additional table 1. Risk factors for CKD which fulfilled inclusion criteria for the NCKDA. **Additional table 2.** Details on linkages between study datasets. **Additional table 3.** ICD-10 codes for AMI identified in HES*. **Additional table 4.** CALIBER definition of AMI subtypes (STEMI, NSTEMI) using MINAP data.*. **Additional table 5.** Variables used to define death in-hospital and post-AMI discharge*. **Additional table 6.** Definitions for processes of AMI care in MINAP and HES datasets. **Additional table 7.** Details on data sources for covariates*. **Additional table 8.** Categorisation of PCI services available at each hospital in England for this study, derived from categorisation of PCI services available at each hospital in England by MINAP researchers. **Additional table 9.** Baseline characteristics of people with incomplete covariate data who are dropped from the complete case analysis (*N*=107). **Additional table 10.** Adjusted predicted percents for dying during the index AMI hospitalisation, and for receiving angiography and/or PCI during the index AMI hospitalisation, stratified by eGFR stage. **Additional table 11.** Other outcomes during the index AMI hospitalisation and post-AMI hospitalisation. **Additional table 12.** Death during the index AMI hospitalisation stratified by centre type. **Additional table 13.** Death post-AMI hospitalisation discharge, among people who survive the first AMI hospitalisation in the study period, stratified by AMI subtype and centre type. **Additional table 14.** Processes of care (angiography and/or PCI) associated with eGFR stage at baseline during the index AMI hospitalisation stratified by centre type in the overall population, and among people with STEMI hospitalisation only. **Additional table 15.** Processes of care (angiography and/or PCI) during the index AMI hospitalisation restricted to people with no previous AMI hospitalisation. **Additional table 16.** MDRD Study equation. **Additional figure 1.** Histogram describing the time between the most recent eGFR recorded in primary care (used to define baseline kidney function) and the index AMI hospitalisation. **Additional figure 2.** Processes of care (angiography and/or PCI) associated with eGFR stage at baseline during the index AMI hospitalisation stratified by prevalent type 2 diabetes mellitus (T2DM) and prevalent heart failure status. **Additional figure 3.** In-hospital death associated with eGFR stage at baseline during the index AMI hospitalisation stratified by prevalent type 2 diabetes mellitus (T2DM) and prevalent heart failure status. **Additional figure 4.** Death post-AMI discharge associated with eGFR stage at baseline during the index AMI hospitalisation stratified by prevalent type 2 diabetes mellitus (T2DM) and prevalent heart failure status.

Acknowledgements

MINAP, as part of the 'National Cardiac Audit Project', was previously commissioned by the Healthcare Quality Improvement Partnership although more recently the National Cardiac Audit Programme has been moved within NHS England. The data used in this study were provided with HQIP's approval. This work uses data that has been provided by patients and collected by the NHS as part of their care and support.

We acknowledge and thank UK cardiovascular, renal, and primary care physicians as well as hospital and community audit and coding teams whose diligent data collection has provided the data for these analyses. In particular, we would like to thank James Chal, Anil Gunesh, Andrew Harrison, and Akosua Donkor from NICOR; and Kathryn Griffith, Matthew Harker, Yvonne Silove, Tasneem Hoosain, Nick Wilson, Ronnie Moodley, Richard Fluck, Chris Gush, David Wheeler, Liam Smeeth, Ron Cullen, Andy Syme, Richard Gunn, Paul Wright, Hugh Gallagher, Sion Edwards, Fiona Loud, Nick Palmer, Richard Fluck, Anita Sharma, Kate Cheema, Andy Syme, Sally Hull, Ben Caplin, Lois Kim, and Faye Cleary from the NCKDA steering, clinical reference, and statistical groups.

Authors' contributions

JS, DMT, DN, DA, FJC, KB and UU conceived the study. JS, DMT, DN, DA, FJC, UU, KB and PB designed the study. PB and DN had access to and analysed the data. PB and JS drafted the manuscript. PB, JS, DMT, UU, FJC, JD, MB, KB, SD, CW, DA, DN (all authors) critically appraised the results and edited the manuscript. All authors approved the final version of the manuscript. The lead authors confirm all authors meet ICJME authorship criteria, and no one who meets ICJME criteria have been excluded.

Funding

This work was supported by Kidney Research UK (grant number IN_008_20180304) and the Health Foundation (grant number 1725841). JS is a Doctoral Research Fellow funded by the NIHR (NIHR 300906). DA is funded by a joint research grant from the British Heart Foundation (SP/16/5/32415) and Cancer Research UK (C53325/A21134).

Availability of data and materials

The data that support the findings of this study are available from the NCKDA and MINAP but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. For further information contact Jemima.Scott@bristol.ac.uk.

Declarations

Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations. This study was approved by the London School of Hygiene & Tropical Medicine LEO ethics committee (ID 16988) and the National Chronic Kidney Disease Audit (NCKDA) Steering Committee. Audits used in this study are covered by Section 251 approvals (NHS Act 2006) which allows data to be collected without individual patient consent for medical research when it is not possible to use anonymised information and when seeking consent is not practical.

Consent for publication

Not applicable.

Competing interests

All authors have completed an ICJME form. PB, DMT, FJC, KB, MDB, and SD have nothing to declare. UU declares a grant from the Health Foundation to undertake quality improvement unrelated to this work. MS declares funding from Cancer Research UK (C53325/A21134) and British Heart Foundation (SP/16/5/32415) for research activities related to this work looking at acute myocardial infarction ascertainment in a cancer population. JS declares Doctoral research funding from the NIHR for related research looking at acute myocardial infarction care for people with chronic kidney disease. JD declares grants from British Heart Foundation paid to his institution unrelated to this work. JD also declares consulting fees from Novo Nordisk, and honoraria from Amgen, Boehringer Ingelheim, Merck, Pfizer, Aegerion, Novartis, Sanofi, Takeda, Novo Nordisk, and Bayer, unrelated to this work. JD is also member of a study steering committee with Novo Nordisk. CW declares he is clinical lead of the Myocardial Ischaemia National Audit Project. DA reports research funding and in-kind support from Astra Zeneca inc. for unrelated research and educational funding from Abbott Vascular inc. to support a clinical research fellow doing unrelated research. DA has conducted consultancy for General Electric inc. to support general research funds. DN is the UK Kidney Association Director of Informatics Research. DN is also on the steering group for two GlaxoSmithKline funded studies that investigate kidney function in children and adults in Sub-Saharan Africa.

Author details
[1]Population Health Sciences, University of Bristol, Bristol, England. [2]Richard Bright Renal Service, North Bristol NHS Trust, Southmead Hospital, Bristol, England. [3]Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London, England. [4]Oxford Kidney Unit, Churchill Hospital, Oxford, England. [5]Nuffield Department of Medicine, University of Oxford, Oxford, England. [6]National Institute for Cardiovascular Outcomes Research (NICOR), NHS Arden & Greater East Midlands Commissioning Support Unit, Leicester, England. [7]Institute of Cardiovascular Sciences, University College London, London, UK. [8]British Heart Foundation, Data Science Centre, London, UK. [9]University College London Hospitals Biomedical Research Centre, London, UK. [10]Glangwili General Hospital, Dolgwili Road, Carmarthen, Wales, UK. [11]Department of Cardiovascular Sciences, University of Leicester, and NIHR Leicester Biomedical Research Centre, Leicester, UK.

Received: 10 May 2023 Accepted: 23 October 2023
Published online: 02 November 2023

References

1. Hill NR, Fatoba ST, Oke JL, Hirst JAO, Callaghan CA, Lasserson DS, Hobbs FDR. Global prevalence of chronic kidney disease - a systematic review and meta-analysis. PLoS One. 2016;11(7):e0158765.
2. Santopinto JJ, Fox KA, Goldberg RJ, Budaj A, Pinero G, Avezum A, et al. Creatinine clearance and adverse hospital outcomes in patients with acute coronary syndromes: findings from the global registry of acute coronary events (GRACE). Heart (British Cardiac Society). 2003;89(9):1003–8.
3. Sederholm Lawesson S, Alfredsson J, Szummer K, Fredrikson M, Swahn E. Prevalence and prognostic impact of chronic kidney disease in STEMI from a gender perspective: data from the SWEDEHEART register, a large Swedish prospective cohort. BMJ Open. 2015;5(6):e008188.
4. Hanna EB, Chen AY, Roe MT, Saucedo JF. Characteristics and in-hospital outcomes of patients presenting with non-ST-segment elevation myocardial infarction found to have significant coronary artery disease on coronary angiography and managed medically: stratification according to renal function. Am Heart J. 2012;164(1):52-7.e1.
5. Major RW, Cheng MRI, Grant RA, Shantikumar S, Xu G, Oozeerally I, et al. Cardiovascular disease risk factors in chronic kidney disease: a systematic review and meta-analysis. PLoS ONE. 2018;13(3):e0192895.
6. Bagai A, Lu D, Lucas J, Goyal A, Herzog CA, Wang TY, et al. Temporal trends in utilization of cardiac therapies and outcomes for myocardial infarction by degree of chronic kidney disease: a report from the NCDR chest pain-MI registry. J Am Heart Assoc. 2018;7(24):e010394.
7. Jakobsson S, Graipe A, Huber D, Bjorklund F, Mooe T. The risk of ischemic stroke after an acute myocardial infarction in patients with decreased renal function. Cerebrovasc Dis. 2014;37(6):460–9.
8. Konstantinidis I, Nadkarni GN, Yacoub R, Saha A, Simoes P, Parikh CR, et al. Representation of patients with kidney disease in trials of cardiovascular interventions: an updated systematic review. JAMA Intern Med. 2016;176(1):121–4.
9. Szummer K, Lundman P, Jacobson SH, Schon S, Lindback J, Stenestrand U, et al. Influence of renal function on the effects of early revascularization in non-ST-elevation myocardial infarction: data from the Swedish Web-System for Enhancement and Development of Evidence-Based Care in Heart Disease Evaluated According to Recommended Therapies (SWEDEHEART). Circulation. 2009;120(10):851–8.
10. Fox CS, Muntner P, Chen AY, Alexander KP, Roe MT, Cannon CP, et al. Use of evidence-based therapies in short-term outcomes of ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction in patients with chronic kidney disease: a report from the national cardiovascular data acute coronary treatment and intervention outcomes network registry. Circulation. 2010;121(3):357–65.
11 Bhatia S, Arora S, Bhatia SM, Al-Hijji M, Reddy YNV, Patel P, et al. Non-ST-segment-elevation myocardial infarction among patients with chronic kidney disease: a propensity score-matched comparison of percutaneous coronary intervention versus conservative management. J Am Heart Assoc. 2018;7(6):e007920.
12. Keeley EC, Kadakia R, Soman S, Borzak S, McCullough PA. Analysis of long-term survival after revascularization in patients with chronic kidney disease presenting with acute coronary syndromes. Am J Cardiol. 2003;92(5):509–14.
13. Shaw C, Nitsch D, Steenkamp R, Junghans C, Shah S, O'Donoghue D, et al. Inpatient coronary angiography and revascularisation following non-ST-elevation acute coronary syndrome in patients with renal impairment: a cohort study using the myocardial ischaemia national audit project. PLoS ONE. 2014;9(6):e99925.
14. Gallacher PJ, Miller-Hodges E, Shah ASV, Farrah TE, Halbesma N, Blackmur JP, et al. High-sensitivity cardiac troponin and the diagnosis of myocardial infarction in patients with kidney impairment. Kidney Int. 2022;102:149–59.
15. Kawsara A, Sulaiman S, Mohamed M, Paul TK, Kashani KB, Boobes K, et al. Treatment effect of percutaneous coronary intervention in dialysis patients with ST-elevation myocardial infarction. Am J Kidney Dis. 2022;79(6):832–40.
16. Lawton JS, Tamis-Holland JE, Bangalore S, Bates ER, Beckie TM, Bischoff JM, et al. 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: executive summary: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation. 2022;145(3):e4–17.
17. Collet JP, Thiele H, Barbato E, Barthélémy O, Bauersachs J, Bhatt DL, et al. 2020 ESC guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. Rev Esp Cardiol (Engl Ed). 2021;74(6):544.
18. Nauta ST, van Domburg RT, Nuis RJ, Akkerhuis M, Deckers JW. Decline in 20-year mortality after myocardial infarction in patients with chronic kidney disease: evolution from the prethrombolysis to the percutaneous coronary intervention era. Kidney Int. 2013;84(2):353–8.
19. Panchal HB, Zheng S, Devani K, White CJ, Leinaar EF, Mukherjee D, et al. Impact of chronic kidney disease on revascularization and outcomes in patients with ST-elevation myocardial infarction. Am J Cardiol. 2021;150:15–23.
20. Khan MZ, Syed M, Osman M, Faisaluddin M, Sulaiman S, Farjo PD, et al. Contemporary trends and outcomes in patients with ST-segment elevation myocardial infarction and end-stage renal disease on dialysis: insight from the national inpatient sample. Cardiovasc Revasc Med. 2020;21(12):1474–81.
21. The DH Vascular Programme Team. Treatment of heart attack national guidance: final report of the National Infarct Angioplasty Project (NIAP). London: Department of Health; 2008.
22. Bidulka P, Scott J, Taylor DM, Udayaraj U, Caskey F, Teece L, et al. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. BMJ Open. 2022;12(3):e057909.
23. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. BMJ. 2013;346:f2350.
24. Nitsch D CB HS, Wheeler D. National chronic kidney disease audit - national report part 1; 2017.
25. Mudie K CF CB, Wheeler D, Hull S, Nitsch D. National chronic kidney disease audit: national report part 2; 2017.
26. Statistics HE. Hospital Episode Statistics (HES); 2022. Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics.
27. NICOR. Myocardial Ischaemia/MINAP (Heart Attack audit); 2022. Available from: https://www.nicor.org.uk/myocardial-ischaemia-minap-heart-attack-audit/.
28. Herrett E, Smeeth L, Walker L, Weston C. The Myocardial Ischaemia National Audit Project (MINAP). Heart. 2010;96(16):1264–7.
29. Statistics OfN. Office for National Statistics, Census 2021. Deaths. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths. Accessed 6 Jan 2022.
30. Gov.uk. English Indices of Deprivation 2019; 2019. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths.
31 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;130(6):461–70.

32. Levey AS, de Jong PE, Coresh J, El Nahas M, Astor BC, Matsushita K, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO controversies conference report. Kidney Int. 2011;80(1):17–28.
33. Norton EC, Dowd BE, Maciejewski ML. Odds ratios-current best practice and use. JAMA. 2018;320(1):84–5.
34. Shaw C, Nitsch D, Lee J, Fogarty D, Sharpe CC. Impact of an early invasive strategy versus conservative strategy for unstable angina and non-ST elevation acute coronary syndrome in patients with chronic kidney disease: a systematic review. PLoS ONE. 2016;11(5):e0153478.
35 Bhatia S, Arora S, Bhatia SM, Al-Hijji M, Reddy YN, Patel P, et al. Non-ST-segment-elevation myocardial infarction among patients with chronic kidney disease: a propensity score-matched comparison of percutaneous coronary intervention versus conservative management. J Am Heart Assoc. 2018;7(6):7920.
36. Patel B, Shah M, Dusaj R, Maynard S, Patel N. Percutaneous coronary intervention and inpatient mortality in patients with advanced chronic kidney disease presenting with acute coronary syndrome. Proc (Baylor Univ Med Cent). 2017;30(4):400–3.
37. McCullough PA, Roberts WC. Influence of chronic renal failure on cardiac structure. J Am Coll Cardiol. 2016;67(10):1183–5.

Publisher's Note

8.2. Relevance to my thesis

This research paper addresses thesis objective 2B: to highlight disparities in AMI treatment by individual-level kidney function (see also step 2 in **Figure 2.1.**). In this research paper, I demonstrate a negative association between the odds of receiving an invasive AMI treatment strategy and worsening kidney function among people from the NCKDA with an AMI hospitalisation recorded in MINAP and/or HES. This negative association was more pronounced for NSTEMI. When converted to the adjusted predicted probability scale, the negative association was less pronounced, but individuals with eGFR stages 3b and 4-5 still had a lower probability of receiving angiography and/or PCI compared with people with eGFR stages 1-2 and 3a (**Table 8.1.**).

Table 8.1. Adjusted predicted percents for dying during the index AMI hospitalisation, and for receiving angiography and/or PCI during the index AMI hospitalisation, stratified by eGFR stage (Additional table 10 in the supplementary materials of the published paper)

Outcome	**eGFR stage**	**Adjusted predicted percent (%) (95% CI)**
Death during index AMI hospitalisation	1-2	13 (12 to 15)
	3a	16 (14 to 18)
	3b	19 (16 to 21)
	4-5	21 (17 to 25)
Angiography and/or PCI during index AMI hospitalisation	1-2	57 (55 to 58)
	3a	59 (56 to 61)
	3b	51 (48 to 55)
	4-5	45 (40 to 50)

AMI: acute myocardial infarction; CI: confidence interval; eGFR: estimated glomerular filtration rate; PCI: percutaneous coronary intervention

There was also a positive association between the odds of death (in-hospital and post-AMI discharge) and worsening kidney function. This analysis relied on traditional multivariable logistic and Cox regression to adjust for measured confounders. In the discussion section of the research paper, I acknowledge the likeliness of residual confounding biasing these associations between AMI treatment and outcomes, particularly since the number of covariates included in the model as potential

confounders was limited due to the relatively small sample size of the cohort (N=5,835). However, the potential collider bias I described in Chapter 7 is reduced because I use both HES and MINAP to define this cohort of people with impaired kidney function hospitalised for AMI. In Appendix H.1., I describe the algorithms and codes used to define cardiac investigation (angiography) and interventions (PCI and CABG) in the study.

In the persistent absence of RCT data including people with impaired kidney function, observational studies using routinely collected data could be useful to clinicians and policymakers to improve practice and guidelines. Previous observational studies have used methods to adjust for confounding such as multivariable regression[1, 2] and propensity scores.[3-5] However, these analytical methods all assume no unmeasured confounding.

Triangulation of these study results by applying alternative advanced quantitative methods which are subject to alternative forms of bias is important to better inform HTA and clinical practice. As in Case Study 1, there is an opportunity to understand whether AMI treatment variation across hospitals in England is a suitable instrument which can be used to design a comparative effectiveness cohort study using an IV analysis to minimise the risk of confounding bias. The sample size of the cohort used in Chapters 7 to 8 was too small to accurately describe this hospital-level variation. In the next chapter, I present pilot work which used a larger copy of MINAP data to describe variation in AMI treatment strategies across cardiology centres in England.

8.3. References

1. Patel B, Shah M, Dusaj R, Maynard S, Patel N. Percutaneous coronary intervention and inpatient mortality in patients with advanced chronic kidney disease presenting with acute coronary syndrome. *Proc (Bayl Univ Med Cent)*. Oct 2017;30(4):400-403. doi:10.1080/08998280.2017.11930205

2. Huang HD, Alam M, Hamzeh I, et al. Patients with severe chronic kidney disease benefit from early revascularization after acute coronary syndrome. *International Journal of Cardiology*. 2013/10/09/ 2013;168(4):3741-3746. doi:https://doi.org/10.1016/j.ijcard.2013.06.013

3. Shaw C, Nitsch D, Steenkamp R, et al. Inpatient Coronary Angiography and Revascularisation following Non-ST-Elevation Acute Coronary Syndrome in Patients with Renal Impairment: A Cohort Study Using the Myocardial Ischaemia National Audit Project. *PLOS ONE*. 2014;9(6):e99925. doi:10.1371/journal.pone.0099925

4. Bhatia S, Arora S, Bhatia SM, et al. Non-ST-Segment-Elevation Myocardial Infarction Among Patients With Chronic Kidney Disease: A Propensity Score-Matched Comparison of Percutaneous Coronary Intervention Versus Conservative Management. *J Am Heart Assoc*. Mar 10 2018;7(6)doi:10.1161/jaha.117.007920

5. Szummer K, Lundman P, Jacobson SH, et al. Influence of renal function on the effects of early revascularization in non-ST-elevation myocardial infarction: data from the Swedish Web-System for Enhancement and Development of Evidence-Based Care in Heart Disease Evaluated According to Recommended Therapies (SWEDEHEART). *Circulation*. Sep 8 2009;120(10):851-8. doi:10.1161/circulationaha.108.838169

CHAPTER 9. RESEARCH PAPER – ACUTE MYOCARDIAL INFARCTION TREATMENT VARIATION AND INEQUALITIES BY KIDNEY FUNCTION: A CROSS SECTIONAL STUDY USING THE MYOCARDIAL ISCHAEMIA NATIONAL AUDIT PROJECT (MINAP).

OVERVIEW

In this chapter, I present a draft manuscript for which I am first author. In this study, I use an unlinked copy of MINAP data, 2014-19, to describe cardiology centre-level and individual-level variation in AMI treatment by level of kidney impairment. These unlinked MINAP data could be impacted by selection bias by conditioning on a collider which I discuss in Chapter 7 and again in the draft manuscript. This research is intended as pilot work for future comparative effectiveness analyses using a TTE framework and an IV analysis in multiple linked data sources. Key tables and figures from the draft supplementary materials are included in the thesis appendix for reference.

Following this manuscript, I include a brief discussion of the paper's relevance to my thesis. This discussion includes additional data which aid in the interpretation of the study in the context of this thesis.

9.1. Draft research paper

The main tables and figures are provided immediately following the references for this draft manuscript. I added the prefix '9'.X for each main figure and table to indicate this is the 9th chapter of this thesis. Select supplementary materials important to this thesis are provided in the Appendix I, referenced within the draft manuscript and also in section 9.2.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	LSH1702213	**Title**	MR
First Name(s)	Patrick Brian		
Surname/Family Name	Bidulka		
Thesis Title	Advancing the use of routinely collected health data in observational research to study relative treatment effects: two natural experiments in UK primary and secondary care.		
Primary Supervisor	Dorothea Nitsch & Richard Grieve		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Nephrology, Dialysis & Transplantation
Please list the paper's authors in the intended authorship order:	Patrick Bidulka, Clive Weston, Mark de Belder, John Deanfield, Rob Konstant-Hambling, Richard Grieve, David Adlam, Dorothea Nitsch

Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I will be first author and corresponding author of this paper. I led the study design with input from my co-authors. I was solely responsible for data extraction, data management, and data analysis. I led the interpretation of results with support from my co-authors. I wrote the first draft of the manuscript. Co-authors have reviewed one draft of the manuscript. I will lead the journal submission and peer review responses.

SECTION E

Student Signature	Patrick Bidulka
Date	01 March 2024

Supervisor Signature	Dorothea Nitsch
Date	4/4/2024

Title

Acute myocardial infarction treatment variation and disparities by kidney function: a cross-sectional study using the Myocardial Ischaemia National Audit Project (MINAP)

Authors

Order	Title	Author name	Affiliation no.	Email
1	Mr	Patrick Bidulka	1	patrick.bidulka1@lshtm.ac.uk
2	Dr	Clive Weston	2	cliveweston@doctors.org.uk
3	Professor	Mark de Belder	3	mark.debelder@nhs.net
4	Professor	John Deanfield	3, 4	j.deanfield@ucl.ac.uk
5	Mr	Rob Konstant-Hambling	5	rob.konstant-hambling@nhs.net
6	Professor	Richard Grieve	6	richard.grieve@lshtm.ac.uk
7	Professor	David Adlam*	7	da134@leicester.ac.uk
8	Professor	Dorothea Nitsch*	1	dorothea.nitsch@lshtm.ac.uk

*Joint-senior authors

Author affiliations

1. Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine
2. Glangwili General Hospital, Dolgwili Road, Carmarthen, Wales
3. National Institute for Cardiovascular Outcomes Research (NICOR), NHS Arden & Greater East Midlands Commissioning Support Unit, Leicester, England
4. Institute of Cardiovascular Sciences, University College London. London, UK
5. NHS England
6. Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine
7. Department of Cardiovascular Sciences, and NIHR Leicester Biomedical Research Centre, Leicester, UK

Word count: 3,340

Tables: 1

Figures: 4

References: 32

ABSTRACT

Background and hypothesis

People with kidney impairment are at increased risk of acute myocardial infarction (AMI). These people have largely been excluded from randomised controlled trials (RCTs) studying AMI treatment, leading to uncertainty as to the benefits versus risks of invasive cardiac treatment strategies in this higher-risk subgroup, particularly for non-ST-elevated myocardial infarction (NSTEMI). We hypothesised that there is substantial variation in AMI treatment across English hospitals, particularly for people hospitalised for NSTEMI and with kidney impairment.

Methods

We used the Myocardial Ischaemia National Audit Project (MINAP), 2014 to 2019, to describe aggregated hospital-level and individual-level AMI treatment variation and disparities according to kidney function. We used multivariable logistic regression and adjusted predicted probabilities to describe the associations between kidney function and AMI treatment (invasive cardiac treatment versus conservative treatment). We also evaluated whether NSTEMI treatment variation is a valid instrument to be used in future natural experiments to evaluate the comparative effectiveness of alternative treatment strategies.

Results

We included 361,259 people with a first hospitalisation for AMI (STEMI or NSTEMI) at 209 hospitals for centre-level analyses, and 292,572 people with complete covariable data at 207 hospitals for individual-level analyses. There was substantial variation in the mean proportion of people with NSTEMI treated with invasive cardiac treatment across centres in England. At the individual-level, people had a lower adjusted predicted probability of being treated with invasive cardiac treatment with worsening eGFR stage, particularly for NSTEMI cases (eGFR stage 2: 76.6% (95% confidence interval (CI) 76.3 to 76.8 versus eGFR stage 5: 44.5% (95% CI 41.2 to 47.5).

Conclusions

There is substantial AMI treatment variation across hospitals in England, particularly among people hospitalised for NSTEMI with impaired kidney function. Future research can explore whether this variation can be exploited in a natural experiment.

KEY LEARNING POINTS

What was known

People with kidney disease are less likely to be treated with invasive cardiac treatment for acute myocardial infarction (AMI), particularly for non-ST-elevated myocardial infarction (NSTEMI), due to some uncertainty in the balance of benefits versus risks for this treatment in this high-risk population. Less is known about AMI treatment variation across English hospitals, particularly according to underlying kidney disease status of treated populations.

This study adds

We have shown substantial variation at the cardiology centre- and individual-levels in AMI treatment across hospitals in England, particularly among people hospitalised with NSTEMI with kidney impairment. The cardiology centre variation is not completely explained by individual-level characteristics.

Potential impact

Our results demonstrate the potential to exploit AMI treatment variation among people with kidney impairment as an instrument in future natural experiments to conduct high-quality comparative effectiveness studies to compare the benefits and risks of invasive cardiac treatment among people with kidney disease.

Keywords

Acute myocardial infarction, kidney disease, audit, health services, treatment variation

INTRODUCTION

People with chronic kidney disease (CKD) are at substantially increased risk of acute myocardial infarction (AMI) compared with the general population.[1, 2] These people experience worse outcomes after AMI,[3] which may in part reflect an older age, the more frequent presence of co-prevalent chronic conditions, such as type 2 diabetes mellitus (T2DM) and heart failure,[4, 5] and adverse effects of treatment, such as bleeding complications.[1] Randomised controlled trials (RCTs) which have driven improvements in AMI treatment and outcomes[6-8] frequently exclude people with kidney impairment,[9] making it challenging to generalise findings from these trials to this high-risk population.

National and international guidelines recommend invasive cardiac treatment strategies, namely angiography, and percutaneous coronary intervention (PCI) if indicated, for almost all ST-elevation myocardial infarction (STEMI) cases. Further, invasive strategies for non-ST-elevation myocardial infarction (NSTEMI) are recommended for those cases judged to be at high risk of mortality.[10-12] Specifically, the National Institute for Health and Care Excellence (NICE) recommend balancing the benefits versus risks of invasive cardiac treatment, particularly for those at elevated risk of complications (e.g., bleeding and acute kidney injury) due to comorbidities such as CKD.[10]

Observational studies and subgroup analyses from RCTs have suggested that people with impaired kidney function may benefit from invasive cardiac treatment despite the risks.[3, 13-18] However, designing and conducting an RCT to specifically study this area of clinical uncertainty is unlikely due to costs. Thus, observational studies which apply advanced quantitative methods to account for biases, including confounding, are needed to reduce the clinical uncertainty for this treatment decision.

The lack of decisive evidence favouring either invasive or conservative cardiac treatment strategies among people with kidney impairment may partly explain both the well-described 'individual-level' association between worsening kidney function and decreased odds of receiving invasive cardiac treatment for NSTEMI and, to some extent, STEMI in England,[15-17] and the overall variation in invasive

versus conservative cardiac treatment strategies, particularly for NSTEMI, observed at the 'centre-level' across hospitals including the general population in England.[19, 20]

This study aimed to describe treatment variation in England among those treated for AMI with impaired kidney function, and suggest how this variation could be used to address the lack of randomised evidence in this high-risk group. Specifically, we aimed to:

1. Describe hospital (centre)-level variation in AMI treatment by aggregate kidney function in England.
2. Describe individual-level variation in AMI treatment by kidney function in England.

METHODS

Study design and data source

We used a cross-sectional study design to investigate AMI treatment variation at centre- and individual-level, according to kidney function, using data from the Myocardial Ischaemia National Audit Project (MINAP) – a prospective national audit programme. MINAP collects detailed information on patient demographics, admission timings and methods, in-patient care including the timeliness of invasive coronary procedures, previous and new drug prescriptions, comorbidity data, and discharge or in-hospital death data.[19, 21]

Study population

We included hospitals in England reporting at least one person hospitalised for AMI (STEMI or NSTEMI) in MINAP between 1 January 2014 to 31 March 2019. We excluded hospitalisations across all centres for unstable angina or any other diagnosis (threatened myocardial infarction, chest pain uncertain cause, myocardial infarction unconfirmed) in aggregated centre-level analyses. Individual-level analyses additionally excluded people with missing covariable data for complete case analyses.

Exposure/independent variable of interest

Our primary exposure of interest was kidney impairment. We measured this using the serum creatinine (SCr), measured within 24 hours of AMI hospitalisation, recorded in MINAP. SCr was converted to eGFR using the CKD-EPI formula,[22] without adjustment for ethnicity. Since we are not applying the chronicity criteria of having two measures <60mL/min/1.73m^2 separated by three months, we cannot define CKD stage. However, we use the same Kidney Disease Improving Global Outcomes guidelines for CKD cut-points[23] to define eGFR stage in this study. These were eGFR stage 1 (≥90mL/min/1.73m^2), stage 2 (60-89), stage 3a (45-59), stage 3b (30-44), stage 4 (15-29), and stage 5 (0-14). Kidney impairment was defined as having eGFR stages 3a-5 or coded chronic renal failure. We categorised people with missing eGFR separately, but assumed they did *not* have kidney impairment in analyses which stratified by kidney function (eGFR stages 1-2, i.e. no kidney impairment, versus eGFR stages 3-5/coded renal failure, i.e. kidney impairment).[24] If a person was transferred between hospitals during an AMI hospitalisation, we used the SCr recorded at the first hospital to which the person was admitted to define eGFR stage, since subsequent SCr measures are likely to be biased upward from baseline kidney function due to acute kidney injury (AKI) either co-incident or resultant from the AMI and its treatment.

Outcomes

Our main outcome of interest was invasive cardiac treatment, defined as early angiography and, if indicated, primary PCI or coronary artery bypass graft (CABG) during the index AMI hospitalisation. Those not recorded as receiving invasive cardiac treatment were considered treated with conservative management. Several MINAP variables were used to define this outcome (thesis Appendix 9.1., and also presented as Supplementary table 1 in the full supplementary materials excluded from this thesis).

Covariables

We considered sex, age, ethnicity (White, Asian, Black, Mixed/Other), comorbidities (previous AMI, angina, hypertension, hypercholesterolaemia, peripheral vascular disease, cerebrovascular disease, COPD, heart failure, T2DM), co-prescriptions (renin-angiotensin system inhibitors (RASi), beta-blocker,

statin), smoking status, and year of AMI hospitalisation as covariables which were to be compared across centres and adjusted for in multivariable models.

We stratified analyses by AMI subtype (STEMI, NSTEMI), defined using a previously developed algorithm for MINAP data (thesis Appendix G.2. and also presented as Supplementary table 2 in the full supplementary materials excluded from this thesis), and by availability of cardiac interventional services at each hospital centre, categorised as (1) PCI available all the time, (2) PCI available sometimes, and (3) PCI available either in exceptional circumstances or never (Supplementary table 3 in the full supplementary materials excluded from this thesis).

Analysis

Objective 1 – Centre-level variation

We described centre-level variation in aggregated individual-level covariables using the median proportions of people with each covariable across hospital centres (reported as median, interquartile range (IQR)). Aggregate centre-level population descriptors were presented overall and by PCI availability at the included centres. People who were transferred between centres during the same AMI event were allocated to the first centre to which they were admitted in the primary analysis.

We then described reported variation in invasive cardiac management by plotting (1) the proportion of STEMI cases reported as being managed with an invasive cardiac strategy and (2) the proportion of NSTEMI cases reported as being managed with an invasive cardiac strategy with respect to the proportion of individuals with kidney impairment, with each point representing a centre. We distinguished centres in these plots by PCI availability to understand the impact of service availability on centre-level variation. To better understand how clinically important comorbidities may impact variation in invasive cardiac intervention among the sub-population where clinical equipoise is greatest (NSTEMI cases with kidney impairment), we also plotted centre-level variation in invasive cardiac treatment restricted to this population including the following as the independent variable (x-

axis): (1) % of cases aged ≥80 years, (2) % of cases with previous myocardial infarction, (3) % of cases with prevalent diabetes mellitus, and (4) % of cases with prevalent heart failure.

Objective 2 – Individual-level variation

We described the whole study population hospitalised for AMI by eGFR stage according to the study covariates. We also described the population with coded chronic renal failure according to their eGFR at admission to better understand the characteristics of this population. We then quantified the individual-level association between eGFR stage and odds of each outcome using multivariable logistic regression, overall and also stratified by (1) AMI subtype and (2) PCI availability at the centre, using eGFR stage 2 as the reference group. We adjusted for all covariates in our multivariable models. We pre-specified adjusting for centre as a random effect; however, because our models would not converge in the overall population, we instead adjusted for centre as a fixed effect.[25] We calculated the adjusted predicted percentages of each outcome from the multivariable logistic regression models, since odds ratios can be exaggerated when the outcome is common.[26, 27]

RESULTS

Aggregated hospital population characteristics

Of the 450,364 hospitalisations for acute coronary syndromes at 209 hospitals in England between 1 January 2014 to 31 March 2019, we included 361,259 people with a first hospitalisation for AMI (STEMI or NSTEMI) at 209 hospitals for centre-level analyses, and 292,572 people with complete covariable data at 207 hospitals for individual-level analyses (**Figure 9.1.**). Of the 361,259 people included in the centre-level analyses, 26,351 (7%) were transferred to at least one other centre within the same AMI event (thesis Appendix 9.2. and also Supplementary table 4 in the full supplementary materials excluded from this thesis).

Of the 209 hospital centres included in the dataset, 120 did not offer PCI services, 38 offered PCI sometimes, and 51 offered PCI all the time. Aggregated at the centre-level, a higher median proportion

of people were female who first presented to hospitals with PCI not available (0.37, IQR 0.34 to 0.39) compared with hospitals offering PCI services all the time (0.29, IQR 0.26 to 0.30) (**Table 1**). The median proportion of people aged 80-89 and 90+ years was higher among centres with PCI not available (0.25, IQR 0.22 to 0.26 and 0.08, IQR 0.06 to 0.11, respectively) compared with centres with PCI available all the time (0.16, IQR 0.14 to 0.20; and 0.03, IQR 0.02 to 0.06, respectively). The median proportion of people with comorbidities also tended to be higher among centres with PCI not available compared with centres with PCI available all the time (e.g., angina (0.26, IQR 0.19 to 0.32 [PCI not available] versus 0.13, IQR 0.09 to 0.20 [PCI available all the time] and previous myocardial infarction (0.23, IQR 0.21 to 0.26 [PCI not available] versus 0.15, IQR 0.14 to 0.18 [PCI available all the time]).

Individual-level population characteristics

At the individual-level, 44,861 (15%) of people had an eGFR corresponding to stage 3a-5 and/or a chronic renal failure diagnosis (thesis Appendix I.3. and also Supplementary table 5 in the full supplementary materials excluded from this thesis). People with stages 1 and 2 were on average younger (58 (10 SD) and 71 (11 SD) years, respectively) compared with people with stages 3a, 3b, 4, and 5 (78 (11 SD), 81 (10 SD), 82 (10 SD), and 76 (13 SD)). Comorbidity prevalence tended to be highest among people with eGFR stages 3b and 4 compared with other eGFR stages.

Of the 18,924 people with coded renal failure, 12,883 people (68%) had an eGFR corresponding to stage 3b-5 (thesis Appendix I.4. and also Supplementary table 6 in the full supplementary materials excluded from this thesis). Overall, the mean age of this subgroup was 78 years (SD 12) and 38% were female.

Centre variation in invasive cardiac investigation and interventions

We observed substantial variation across centres in England in the proportion of people reported as receiving invasive cardiac treatment versus conservative treatment for both STEMI and NSTEMI (**Figure 9.2.**). For both STEMI and NSTEMI, there was a negative association between the proportion of AMI cases reported as receiving an invasive cardiac strategy and the proportion of AMI cases with an admission eGFR<60mL/min/1.73m^2. However, the variation in invasive cardiac treatment between

centres for STEMI hospitalisations was dependent on PCI availability: an average (SD) proportion of 0.58 (0.19) and 0.77 (0.16) with STEMI were reported as being managed with an invasive cardiac strategy at centres with PCI services not available or available sometimes, respectively. In contrast, an average proportion of 0.96 (SD 0.03) with STEMI were reported as being managed with an invasive cardiac strategy at centres with PCI always available. For NSTEMI hospitalisations, there was substantial variation in reporting of invasive cardiac management across all levels of PCI availability: the mean (SD) proportion of people reported as being managed with an invasive cardiac strategy was 0.63 (0.15), 0.71 (0.12), and 0.80 (0.11) for centres with PCI not available, PCI sometimes available, and PCI always available, respectively.

When restricting to people hospitalised with NSTEMI and kidney impairment, we did not see strong centre-level associations between the proportion reported as being managed with an invasive cardiac strategy and the proportion with previous myocardial infarction, prevalent diabetes mellitus, or prevalent heart failure (Supplementary figure 1 in the full supplementary materials excluded from this thesis). However, we observed some centre-level association with proportion aged ≥80 years.

When restricting to centres with PCI available all the time (N=51), we observed variation in the percent treated with invasive cardiac treatment versus conservative treatment across centres, particularly for people hospitalised for NSTEMI with kidney impairment (**Figure 9.3.**).

Individual variation in invasive cardiac investigation and interventions

At the individual-level, the proportion of people who received invasive cardiac treatment decreased with worsening eGFR stage overall: 80,092 (93%) of people received cardiac investigation and/or intervention with eGFR stage 1 versus 464 (40%) of people with eGFR stage 5 (**Supplementary table 7**). Among people with STEMI, this trend was the same, although a higher proportion of people received invasive cardiac treatment across all levels of kidney function (98% to 60% for stages 1 and 5, respectively). Among people with NSTEMI, the proportions receiving cardiac investigation and/or

intervention were lower across all levels of kidney function (90% to 33% for stages 1 and 5, respectively).

After adjusting for measured confounders and using people with eGFR stage 2 as the reference group, we observed a consistent decrease in the odds of receiving invasive cardiac treatment with worsening kidney function (**Appendix 1.5.,** and also Supplementary figure 2 and table 7 in the full supplementary materials excluded from this thesis). This pattern was similar overall, and for both STEMI and NSTEMI hospitalisations. For example, restricting to people with STEMI, the odds ratio decreased from 0.63 (95% CI 0.59 to 0.68) [stage 3a] to 0.16 (95% CI 0.13 to 0.21) [stage 5]. Similarly, restricting to people with NSTEMI, the odds ratio decreased from 0.73 (95% CI 0.70 to 0.75) [stage 3a] to 0.16 (95% CI 0.13 to 0.18) [stage 5]. In both STEMI and NSTEMI, the odds of being treated with invasive cardiac treatment for people with coded chronic renal failure were higher (OR 0.31, 95% CI 0.28 to 0.34; OR 0.42, 95% CI 0.40 to 0.44, respectively) than for people with eGFR stages 4 and 5.

When converted to the predicated probability scale, there were significant differences between the adjusted probability of being treated with invasive cardiac treatment with worsening eGFR stage for both STEMI and NSTEMI (**Figure 9.4.**). However, the adjusted probability of being treated with invasive cardiac treatment was high across all levels of kidney function among people with STEMI (e.g., Stage 1: 93.7% [95% CI 93.3 to 94.1]; Stage 5: 77.6% [95% CI 74.0 to 81.1]; coded renal failure: 85.2% [95% CI 84.3 to 86.2]). The adjusted probabilities were lower among people with NSTEMI (e.g., Stage 1: 73.2% [95% CI 72.7 to 73.8]; Stage 5: 44.5% [95% CI 41.2 to 47.5]; coded renal failure: 62.5% [95% CI 61.8 to 63.1]).

Adjusted predicted probabilities of invasive cardiac treatment when stratified by PCI service availability among included centres followed similar association by eGFR stage. Adjusted probabilities ranged from: 70.2% (95% CI 65.9 to 67.4) [eGFR stage 2] to 32.3% (95% CI 28.0 to 36.7) [eGFR stage 5] among centres with PCI never available; 78.9% (95% CI 78.4 to 79.4) [eGFR stage 2] to 45.5% (95% CI 39.9 to 51.1) [eGFR stage 5] among centres with PCI available sometimes; and 92.6% (95% CI 92.4 to

92.8) [eGFR stage 2] to 76.0% (95% CI 73.0 to 79.0) [eGFR stage 5] among centres with PCI available all the time (Supplementary table 8 in the full supplementary materials excluded from this thesis).

DISCUSSION

Summary of findings

We described substantial variation in the proportion of people hospitalised with AMI receiving invasive cardiac treatment versus conservative treatment across hospitals in England in MINAP data. This variation was more pronounced among people with NSTEMI and with an admission serum creatinine indicating kidney impairment. At the individual-level, although we observed relative inequalities in AMI treatment by eGFR stage, these inequalities were less extreme when considered on the absolute scale. We observed lower adjusted probabilities of being treated with invasive cardiac treatment with decreasing eGFR stage, although the adjusted probability was relatively high across all eGFR levels for STEMI hospitalisations (77-93%) and above 50% for eGFR stages 1-4 and for those with coded renal failure for NSTEMI hospitalisations.

Strengths

This study used data from MINAP, part of the National Institute of Cardiovascular Outcomes Research (NICOR) audit programme.[19, 21, 28] Our dataset was large and nationally representative. The data included granular AMI treatment and covariate data, which enabled us to observe AMI treatment variation across AMI subtypes and according to kidney function – an important clinical characteristic when considering AMI treatment strategies. We found similar variation across hospitals in England as a previous study which used the same data source, but only covered AMI hospitalisations between 2004 to 2010.[20] Further, we investigated variation specifically among people with kidney impairment, a high-risk group which are largely excluded from clinical trials,[9] leading to ambiguity in nation-wide clinical guidelines from NICE.[10]

A previous study investigated individual-level associations between particular comorbidities captured in MINAP, including chronic renal failure, and receipt of optimal guideline-recommended AMI treatment. This study found no evidence that people with coded renal failure had lower odds of optimal AMI treatment.[29] Our study investigated a broader subgroup of people hospitalised for AMI with kidney impairment and found that people with coded chronic renal failure had a lower probability of receiving invasive cardiac treatment compared to people with no evidence of kidney impairment. Moreover, people with coded renal failure had a higher probability of receiving this treatment compared with people with eGFR stages 4-5 for both STEMI and NSTEMI hospitalisations. We were unable to distinguish between people with kidney transplants, people on dialysis, and people with other types of kidney disease within this subgroup, which would help in understanding why people with coded renal failure are more likely to receive invasive cardiac care compared with people with eGFR stages 4 and 5.

Limitations

These data were unlinked to other routinely collected health datasets. We know from previous work that AMI case ascertainment is incomplete with MINAP alone,[30-32] particularly since MINAP focuses on capturing type 1 AMI,[21] meaning our study population is only a selection of all AMI hospitalisations in England. These results are therefore vulnerable to collider bias. We also relied on serum creatinine recorded at the time of AMI hospitalisation to determine baseline kidney function and eGFR stage. Thus, misclassification of kidney disease status is likely, since many patients may experience acute declines in kidney function at the time of AMI hospitalisation due to intercurrent illness, or as previous work demonstrated, random error in the estimation of eGFR defined in secondary care versus eGFR defined in primary care.[31] Yet it is likely to be the creatinine level on hospitalisation (or soon after) that is taken into account by the clinician referring for (or accepting for) angiography.

Further, these data demonstrate substantial variation in the proportion of patients recorded in MINAP as receiving invasive cardiac management after AMI. This is likely to reflect true variations in treatment

but also regional variations in AMI pathways and reporting issues (such as non-PCI centres failing to report procedures carried out off-site and reporting more type 2 AMI or medically managed AMI in highly co-morbid patients). Thus, we are careful to explain that the variation we observe is the reported variation in the MINAP data and may not reflect true variation in AMI care.

Future research

The AMI treatment variation across cardiology centres in England should be further explored in linked secondary care data to understand if this variation could be exploited in a natural experiement, comparing invasive versus conservative NSTEMI treatment strategies in people with impaired kidney function/chronic kidney disease. Linking these audit data with other routinely collected health data is important to improve AMI case ascertainment and reliably estimate baseline kidney function.

Conclusions

We highlighted that substantial variation in AMI treatment exists across hospitals in England, particularly among people hospitalised for NSTEMI and with kidney impairment.

REFERENCES

1. Santopinto JJ, Fox KA, Goldberg RJ, et al. Creatinine clearance and adverse hospital outcomes in patients with acute coronary syndromes: findings from the global registry of acute coronary events (GRACE). *Heart (British Cardiac Society)*. Sep 2003;89(9):1003-8. doi:10.1136/heart.89.9.1003

2. Hanna EB, Chen AY, Roe MT, Saucedo JF. Characteristics and in-hospital outcomes of patients presenting with non-ST-segment elevation myocardial infarction found to have significant coronary artery disease on coronary angiography and managed medically: stratification according to renal function. *Am Heart J*. Jul 2012;164(1):52-7.e1. doi:10.1016/j.ahj.2012.04.009

3. Fox CS, Muntner P, Chen AY, et al. Use of evidence-based therapies in short-term outcomes of ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction in patients with chronic kidney disease: a report from the National Cardiovascular Data Acute Coronary Treatment and Intervention Outcomes Network registry. *Circulation*. Jan 26 2010;121(3):357-65. doi:10.1161/circulationaha.109.865352

4. Nitsch D, Caplin B, Hull S, Wheeler D. *National Chronic Kidney Disease Audit: National Report (Part 1)*. 2017. https://www.lshtm.ac.uk/files/ckd_audit_report.pdf

5. Major RW, Cheng MRI, Grant RA, et al. Cardiovascular disease risk factors in chronic kidney disease: A systematic review and meta-analysis. *PLoS One*. 2018;13(3):e0192895. doi:10.1371/journal.pone.0192895

6. Weaver WD, Simes RJ, Betriu A, et al. Comparison of primary coronary angioplasty and intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review. *Jama*. Dec 17 1997;278(23):2093-8.

7. A clinical trial comparing primary coronary angioplasty with tissue plasminogen activator for acute myocardial infarction. *The New England journal of medicine*. Jun 5 1997;336(23):1621-8. doi:10.1056/nejm199706053362301

8. Fox KA, Clayton TC, Damman P, et al. Long-term outcome of a routine versus selective invasive strategy in patients with non-ST-segment elevation acute coronary syndrome a meta-analysis of individual patient data. *J Am Coll Cardiol*. Jun 1 2010;55(22):2435-45. doi:10.1016/j.jacc.2010.03.007

9. Konstantinidis I, Nadkarni GN, Yacoub R, et al. Representation of Patients With Kidney Disease in Trials of Cardiovascular Interventions: An Updated Systematic Review. *JAMA internal medicine*. Jan 2016;176(1):121-4. doi:10.1001/jamainternmed.2015.6102

10. NICE guideline [NG185]: Acute coronary syndrome. Web. NICE. Accessed March 10, 2022. https://www.nice.org.uk/guidance/ng185

11. Byrne RA, Rossello X, Coughlan JJ, et al. 2023 ESC Guidelines for the management of acute coronary syndromes: Developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC). *European heart journal*. 2023;44(38):3720-3826. doi:10.1093/eurheartj/ehad191

12. Lawton JS, Tamis-Holland JE, Bangalore S, et al. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. Jan 18 2022;145(3):e4-e17. doi:10.1161/cir.0000000000001039

13. Szummer K, Lundman P, Jacobson SH, et al. Influence of renal function on the effects of early revascularization in non-ST-elevation myocardial infarction: data from the Swedish Web-System for Enhancement and Development of Evidence-Based Care in Heart Disease Evaluated According to Recommended Therapies (SWEDEHEART). *Circulation*. Sep 8 2009;120(10):851-8. doi:10.1161/circulationaha.108.838169

14. Bhatia S, Arora S, Bhatia SM, et al. Non-ST-Segment-Elevation Myocardial Infarction Among Patients With Chronic Kidney Disease: A Propensity Score-Matched Comparison of Percutaneous Coronary Intervention Versus Conservative Management. *J Am Heart Assoc*. Mar 10 2018;7(6)doi:10.1161/jaha.117.007920

15. Shaw C, Nitsch D, Steenkamp R, et al. Inpatient coronary angiography and revascularisation following non-ST-elevation acute coronary syndrome in patients with renal impairment: a cohort study using the Myocardial Ischaemia National Audit Project. *PLoS One*. 2014;9(6):e99925. doi:10.1371/journal.pone.0099925

16. Shaw C, Nitsch D, Lee J, Fogarty D, Sharpe CC. Impact of an Early Invasive Strategy versus Conservative Strategy for Unstable Angina and Non-ST Elevation Acute Coronary Syndrome in Patients with Chronic Kidney Disease: A Systematic Review. *PLOS ONE*. 2016;11(5):e0153478. doi:10.1371/journal.pone.0153478

17. Scott J, Bidulka P, Taylor DM, et al. Management and outcomes of myocardial infarction in people with impaired kidney function in England. *BMC Nephrology*. 2023/11/02 2023;24(1):325. doi:10.1186/s12882-023-03377-x

18. Huang HD, Alam M, Hamzeh I, et al. Patients with severe chronic kidney disease benefit from early revascularization after acute coronary syndrome. *International Journal of Cardiology*. 2013/10/09/ 2013;168(4):3741-3746. doi:https://doi.org/10.1016/j.ijcard.2013.06.013

19. Weston C, Perwaiz S, Wang J, Kerr J. *Management of Heart Attack: analyses from the Myocardial Ischaemia National Audit Project (MINAP) and the National Audit of Percutaneous Coronary Intervention (NAPCI) 2023 Summary Report (2021/22 data)*. Web. 2023.

https://www.nicor.org.uk/wp-content/uploads/2023/06/10633-NICOR-Annual-Summary_Reports_MINAP_v4_AC.pdf

20. Chung S-C, Sundström J, Gale CP, et al. Comparison of hospital variation in acute myocardial infarction care and outcome between Sweden and United Kingdom: population based cohort study using nationwide clinical registries. *BMJ : British Medical Journal*. 2015;351:h3913. doi:10.1136/bmj.h3913

21. Herrett E, Smeeth L, Walker L, Weston C, on behalf of the MAG. The Myocardial Ischaemia National Audit Project (MINAP). *Heart (British Cardiac Society)*. 2010;96(16):1264. doi:10.1136/hrt.2009.192328

22. Levey AS, Stevens LA, Schmid CH, et al. A New Equation to Estimate Glomerular Filtration Rate. *Annals of Internal Medicine*. 2009/05/05 2009;150(9):604-612. doi:10.7326/0003-4819-150-9-200905050-00006

23. Stevens PE, Levin A. Evaluation and Management of Chronic Kidney Disease: Synopsis of the Kidney Disease: Improving Global Outcomes 2012 Clinical Practice Guideline. *Annals of Internal Medicine*. 2013/06/04 2013;158(11):825-830. doi:10.7326/0003-4819-158-11-201306040-00007

24. Iwagami M, Tomlinson LA, Mansfield KE, et al. Validity of estimated prevalence of decreased kidney function and renal replacement therapy from primary care electronic health records compared with national survey and registry data in the United Kingdom. *Nephrol Dial Transplant*. Apr 1 2017;32(suppl_2):ii142-ii150. doi:10.1093/ndt/gfw318

25. Peracha J, Pitcher D, Santhakumaran S, et al. Centre variation in mortality following post-hospitalization acute kidney injury: analysis of a large national cohort. *Nephrol Dial Transplant*. Oct 19 2022;37(11):2201-2213. doi:10.1093/ndt/gfab348

26. Norton EC, Dowd BE, Maciejewski ML. Odds Ratios—Current Best Practice and Use. *Jama*. 2018;320(1):84-85. doi:10.1001/jama.2018.6971

27. Bidulka P, Mathur R, Lugo-Palacios DG, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes, Obesity and Metabolism*. 2023;25(1):282-292. doi:https://doi.org/10.1111/dom.14874

28. Myocardial Ischaemia/MINAP (Heart Attack audit). Web. The National Institute for Cardiovascular Outcomes Research. Accessed 18 May, 2021. https://www.nicor.org.uk/national-cardiac-audit-programme/myocardial-ischaemia-minap-heart-attack-audit/

29. Yadegarfar ME, Gale CP, Dondo TB, Wilkinson CG, Cowie MR, Hall M. Association of treatments for acute myocardial infarction and survival for seven common comorbidity states: a nationwide cohort study. *BMC Med*. Aug 24 2020;18(1):231. doi:10.1186/s12916-020-01689-5

30. Herrett E, Shah AD, Boggon R, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ : British Medical Journal*. 2013;346:f2350. doi:10.1136/bmj.f2350

31. Bidulka P, Scott J, Taylor DM, et al. Impact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. *BMJ Open*. 2022;12(3):e057909. doi:10.1136/bmjopen-2021-057909

32. Coles B, Teece L, Weston C, et al. Case-ascertainment of acute myocardial infarction hospitalisations in cancer patients: a cohort study using English linked electronic health data. *European Heart Journal - Quality of Care and Clinical Outcomes*. 2021;doi:10.1093/ehjqcco/qcab045

Table 9.1. Aggregate population characteristics at the centre-level for people hospitalised for AMI between 2014 and 2019 in England.

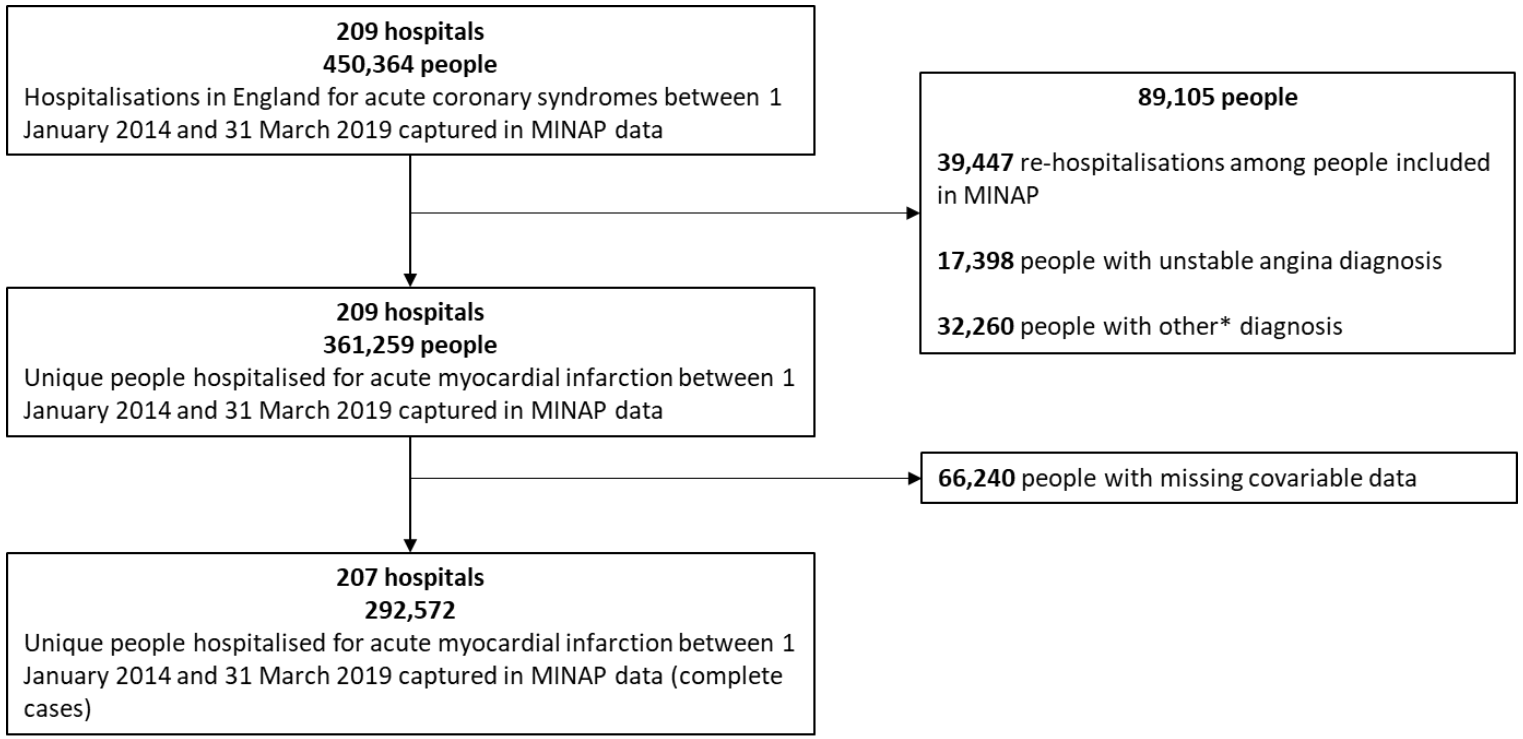
Reported as the median (interquartile range) proportion across centres of people with each characteristic, unless otherwise specified. Presented across all centres, and stratifying centres according to PCI availability.

	All hospitals N=209	PCI not available N=120	PCI available sometimes N=38	PCI available all the time N=51
Number of people admitted per year, median number of people (IQR)				
2014	271 (129-436)	175 (102-280)	317 (188-432)	601 (437-841)
2015	282 (137-443)	167 (99-282)	306 (222-424)	603 (460-853)
2016	253 (131-414)	168 (93-258)	305 (192-386)	595 (480-873)
2017	258 (147-454)	173 (77-252)	306 (208-408)	643 (526-882)
2018	244 (153-460)	174 (103-237)	316 (254-410)	641 (501-911)
2019	63 (33-108)	42 (17-63)	82 (63-100)	161 (114-221)
Female	0.35 (0.30-0.38)	0.37 (0.34-0.39)	0.35 (0.32-0.37)	0.29 (0.26-0.30)
Age (years)				
50-59	0.16 (0.13-0.19)	0.14 (0.13-0.16)	0.16 (0.14-0.18)	0.21 (0.19-0.23)
60-69	0.22 (0.20-0.24)	0.21 (0.19-0.23)	0.21 (0.19-0.23)	0.25 (0.23-0.26)
70-79	0.25 (0.23-0.27)	0.25 (0.23-0.27)	0.25 (0.24-0.27)	0.24 (0.22-0.25)
80-89	0.23 (0.19-0.26)	0.25 (0.22-0.26)	0.23 (0.21-0.26)	0.16 (0.14-0.20)
90+	0.07 (0.04-0.09)	0.08 (0.06-0.11)	0.08 (0.05-0.10)	0.03 (0.02-0.06)
Missing	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Ethnicity				
White	0.86 (0.64-0.96)	0.88 (0.64-0.96)	0.86 (0.64-0.94)	0.85 (0.63-0.95)
Black, Asian, Mixed, or other	0.03 (0.01-0.11)	0.02 (0.01-0.07)	0.05 (0.01-0.13)	0.04 (0.01-0.11)
Missing	0.04 (0.01-0.15)	0.05 (0.00-0.16)	0.02 (0.01-0.15)	0.04 (0.01-0.13)
eGFR stage at AMI hospitalisation				
1-2	0.67 (0.62-0.72)	0.65 (0.62-0.69)	0.66 (0.61-0.72)	0.74 (0.71-0.78)
3a-3b	0.23 (0.19-0.26)	0.25 (0.22-0.27)	0.24 (0.21-0.27)	0.17 (0.16-0.20)
4-5	0.06 (0.04-0.07)	0.06 (0.05-0.08)	0.07 (0.05-0.08)	0.04 (0.03-0.05)
Missing	0.02 (0.01-0.04)	0.02 (0.01-0.04)	0.02 (0.01-0.03)	0.02 (0.01-0.05)

Comorbidities				
Angina	0.23 (0.16-0.29)	0.26 (0.19-0.32)	0.23 (0.18-0.29)	0.13 (0.09-0.20)
Cerebrovascular disease	0.09 (0.06-0.11)	0.10 (0.07-0.12)	0.08 (0.07-0.10)	0.06 (0.03-0.08)
COPD	0.16 (0.13-0.19)	0.18 (0.15-0.21)	0.16 (0.15-0.19)	0.13 (0.11-0.16)
Type 2 diabetes mellitus	0.25 (0.23-0.27)	0.26 (0.24-0.27)	0.27 (0.24-0.29)	0.20 (0.19-0.23)
Heart failure	0.06 (0.04-0.09)	0.07 (0.06-0.10)	0.07 (0.06-0.08)	0.03 (0.02-0.06)
Hypercholesterolaemia	0.30 (0.21-0.40)	0.30 (0.21-0.42)	0.31 (0.25-0.36)	0.29 (0.22-0.40)
Hypertension	0.52 (0.47-0.57)	0.54 (0.49-0.60)	0.52 (0.48-0.56)	0.48 (0.44-0.52)
Myocardial infarction	0.22 (0.17-0.25)	0.23 (0.21-0.26)	0.22 (0.20-0.26)	0.15 (0.14-0.18)
Peripheral vascular disease	0.04 (0.03-0.06)	0.05 (0.03-0.06)	0.04 (0.03-0.05)	0.03 (0.02-0.05)
Renal failure	0.08 (0.04-0.10)	0.08 (0.05-0.11)	0.09 (0.07-0.10)	0.04 (0.02-0.07)
Previous coronary interventions				
PCI	0.11 (0.09-0.14)	0.11 (0.09-0.15)	0.12 (0.10-0.16)	0.11 (0.09-0.13)
CABG	0.07 (0.06-0.09)	0.08 (0.06-0.09)	0.08 (0.07-0.09)	0.05 (0.04-0.06)
Prescriptions pre-AMI hospitalisation				
Beta-blocker	0.28 (0.23-0.32)	0.30 (0.27-0.34)	0.30 (0.28-0.33)	0.20 (0.15-0.23)
RASi	0.37 (0.30-0.41)	0.38 (0.35-0.42)	0.39 (0.35-0.43)	0.28 (0.21-0.32)
Statin	0.42 (0.35-0.48)	0.44 (0.40-0.49)	0.44 (0.39-0.48)	0.32 (0.27-0.36)
Smoking status				
Non-smoker	0.37 (0.34-0.43)	0.37 (0.34-0.44)	0.38 (0.35-0.45)	0.37 (0.32-0.40)
Ex-smoker	0.33 (0.28-0.37)	0.34 (0.31-0.37)	0.33 (0.27-0.37)	0.29 (0.26-0.32)
Current smoker	0.21 (0.18-0.27)	0.20 (0.17-0.24)	0.20 (0.16-0.24)	0.29 (0.24-0.33)
Missing	0.04 (0.01-0.09)	0.05 (0.02-0.10)	0.04 (0.01-0.09)	0.04 (0.01-0.09)

AMI: acute myocardial infarction; CABG: coronary artery bypass graft; COPD: chronic obstructive pulmonary disease; eGFR: estimated glomerular filtration rate; IQR: interquartile range; PCI: percutaneous coronary intervention; RASi: renin-angiotensin system inhibitor

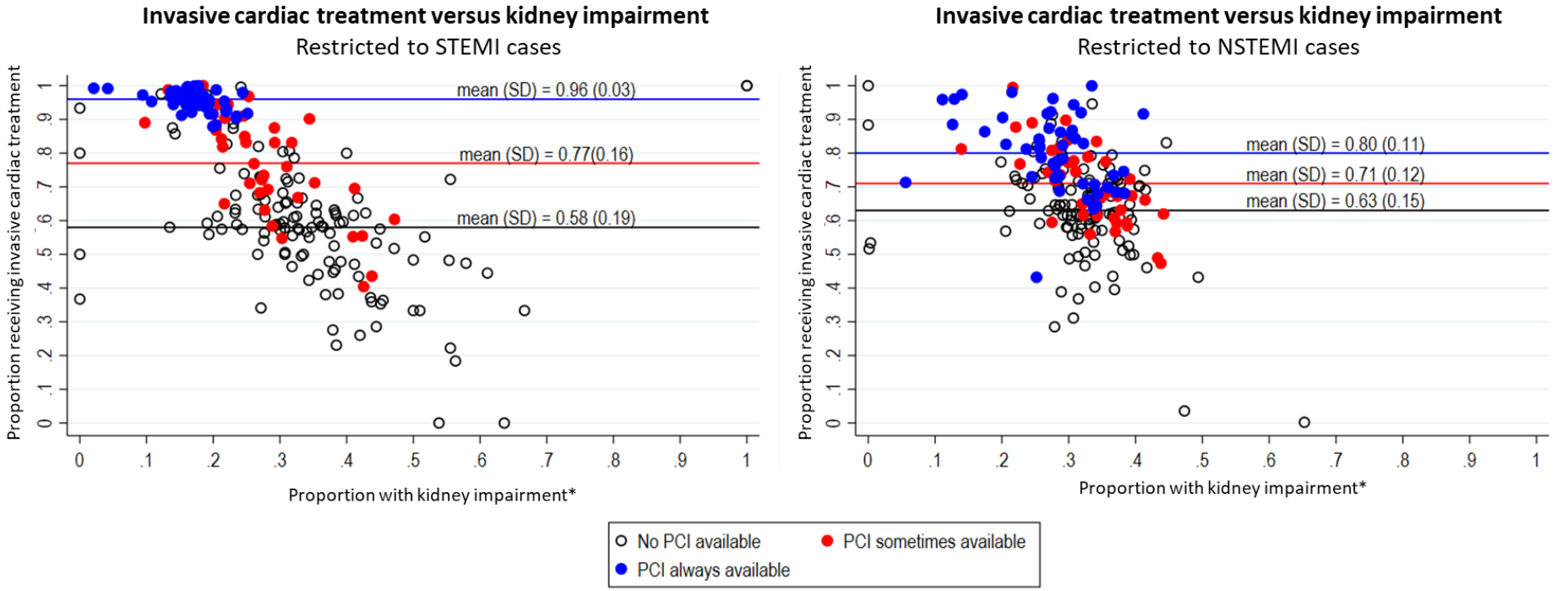
Figure 9.1. Flow diagram showing study population selection



AMI: Acute Myocardial Infarction; MINAP: Myocardial Ischaemia National Audit Project

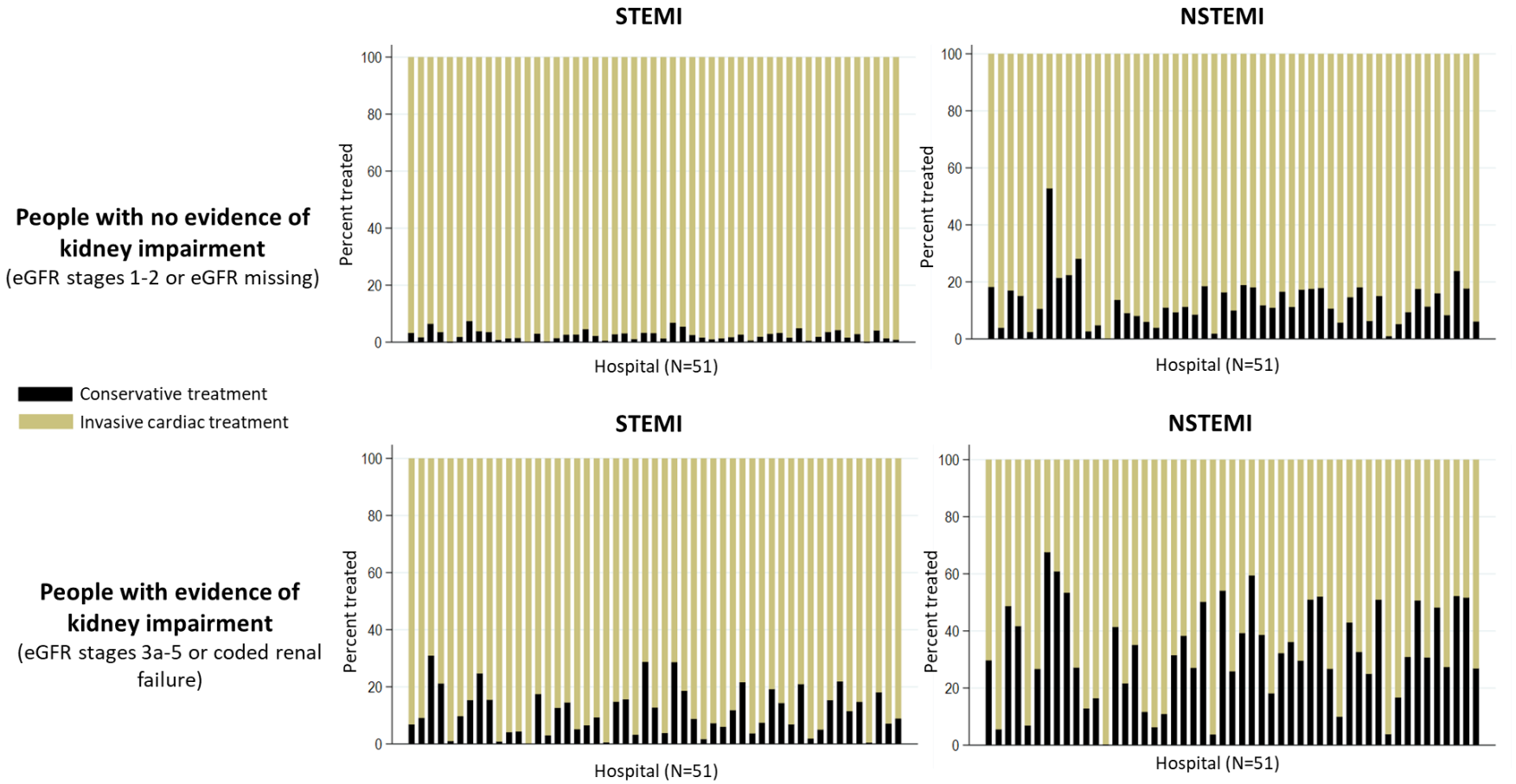
*Other diagnoses include threatened myocardial infarction, chest pain (uncertain cause), myocardial infarction unconfirmed, other diagnosis (not myocardial infarction)

Figure 9.2. Variation in AMI treatment (angiography and/or PCI) at the centre-level, for (a) STEMI and (b) NSTEMI.



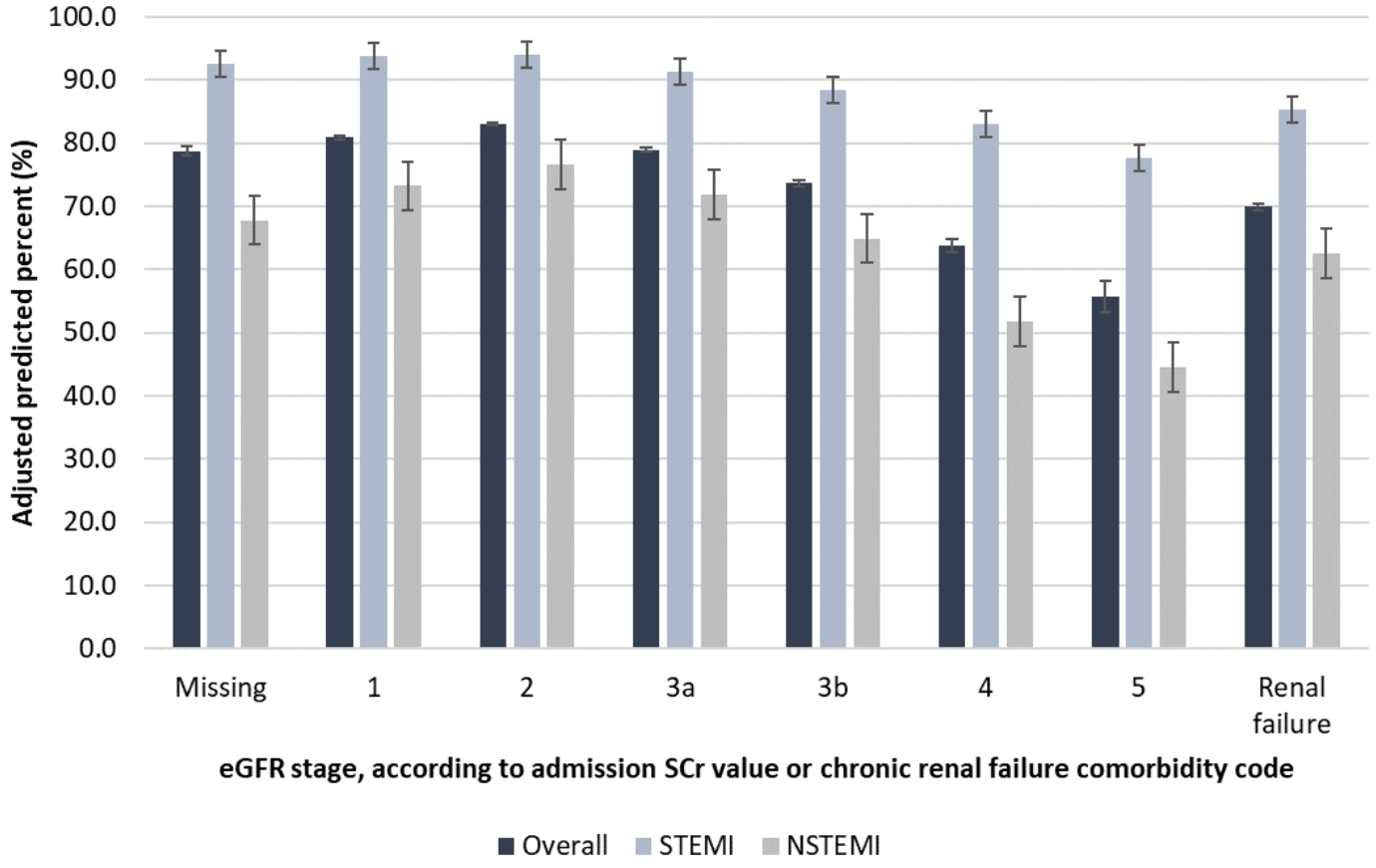
*Kidney impairment defined as eGFR stages 3a-5 or coded renal impairment

Figure 9.3. Centre variation in invasive cardiac treatment versus conservative treatment among centres with PCI available all the time, stratified by AMI subtype (STEMI and NSTEMI) and level of kidney function (no evidence of kidney impairment and evidence of kidney impairment)



AMI: acute myocardial infarction; eGFR: estimated glomerular filtration rate; PCI: percutaneous coronary intervention; NSTEMI: non-ST elevated myocardial infarction; STEMI: ST-elevated myocardial infarction

Figure 9.4. Adjusted predicted percentages of people who receive angiography and/or PCI by eGFR stage, overall and stratified by AMI subtype



AMI: acute myocardial infarction; eGFR: estimated glomerular filtration rate; NSTEMI: non-ST-elevated myocardial infarction; PCI: percutaneous coronary intervention; STEMI: ST-elevated myocardial infarction

9.2. Relevance to my thesis

In this manuscript, I partially address objective 2C: to inform the selection of potential preference-based instruments for a comparative effectiveness analysis of alternative AMI treatments. I demonstrate substantial variation across cardiology centres in England in the *recorded* treatment of AMI, particularly among people with kidney impairment hospitalised for NSTEMI. This is true when looking at crude proportions of people recorded as receiving invasive cardiac treatment strategies by centre as well as in multivariable logistic regression models modelling cardiology centre with a fixed effect.

In this chapter, I pre-specified the analysis to use a mixed model using a random effect to account for clustering of the study population within cardiology centres, similar to my analyses in Chapter 3. However, when implementing this modelling strategy, the full multivariable mixed effect model would not converge, possibly due to complex correlation structures within the data. I therefore included cardiology centre as a fixed effect in an alternative modelling strategy to overcome this issue.[1] This alternative modelling strategy does not rely on assumptions about random effects distribution, making it more robust to misspecification of these random effects.[2] However, including centres as fixed effect dummy variables within the multivariable model will reduce the degrees of freedom and result in less precise effect estimates. Further, these models will only estimate within-cluster effects, rather than population-level effects, and cannot consider the role of specific centre-level characteristics (e.g., teaching hospital).

In **Table 9.2.**, I present additional analyses not included in the draft manuscript. This analysis restricted to cardiology centres which offer PCI all the time (n=51) and fitted a mixed effects logistic regression model with cardiology centre as a random effect. I add fixed effect variables to the model in 3 steps: model 1 includes only the cardiology-centre as a random effect; model 2 includes the cardiology-centre as a random effect and eGFR stage as a fixed effect; and model 3 adds all other covariates, including year of admission and other important patient-level characteristics. Of interest is the intracluster coefficient, which is the ratio of the between-cluster variability and the sum of the within-cluster and between-cluster variabilities[3] (i.e., the proportion of the variation in the outcome explained by the centre-level variation). This exploratory analysis shows that there is a substantial proportion of variation in the reported provision of an invasive cardiac strategy for AMI between cardiology centres, even after accounting for individual characteristics and time period.

Table 9.2. Additional data describing the cardiology centre variation in reported invasive cardiac treatment strategy in cardiology centres which offer PCI services all the time

		Model 1*			Model 2**			Model 3***		
			95% CI		OR	95% CI		OR	95% CI	
STEMI	**Missing**	-	-	-	0.51	0.25	1.05	0.31	0.14	0.67
	Stage 1	-	-	-	2.57	2.06	3.2	0.89	0.75	1.05
	Stage 2	-	-	-	1	-	-	1	-	-
	Stage 3a	-	-	-	0.37	0.32	0.42	0.58	0.51	0.67
	Stage 3b	-	-	-	0.21	0.17	0.25	0.42	0.36	0.5
	Stage 4	-	-	-	0.11	0.09	0.13	0.24	0.2	0.29
	Stage 5	-	-	-	0.11	0.07	0.17	0.17	0.11	0.27
	Coded renal failure	-	-	-	0.17	0.14	0.2	0.32	0.26	0.39
	Rho (intracluster coefficient)	0.27	0.17	0.41	0.24	0.17	0.42	0.29	0.18	0.43
NSTEMI	**Missing**	-	-	-	0.86	0.55	1.34	0.62	0.43	0.89
	Stage 1	-	-	-	2.7	2.17	3.35	0.85	0.75	0.96
	Stage 2	-	-	-	1	-	-	1	-	-
	Stage 3a	-	-	-	0.48	0.43	0.53	0.75	0.68	0.83
	Stage 3b	-	-	-	0.26	0.22	0.3	0.48	0.43	0.53
	Stage 4	-	-	-	0.13	0.1	0.16	0.25	0.2	0.31
	Stage 5	-	-	-	0.17	0.13	0.24	0.19	0.13	0.26
	Coded renal failure	-	-	-	0.26	0.22	0.31	0.4	0.34	0.45
	Rho (intracluster coefficient)	0.26	0.16	0.39	0.28	0.17	0.42	0.28	0.17	0.43

*Model 1: Cardiology centre as random effect
**Model 2: Cardiology centre as random effect and eGFR stage as fixed effect
***Model 3: Cardiology centre as random effect, eGFR stages, sex, age, admission year, ethnicity, comorbidites (previous MI, angina, hypertension, hypercholesterolaemia, peripheral vascular disease, COPD, heart failure, type 2 diabetes), co-prescriptions (RASi, beta-blocker, statin)

CI: confidence interval; COPD: chronic obstructive pulmonary disease; eGFR: estimated glomerular filtration rate; MI: myocardial infarction; OR: odds ratio; RASi: renin-angiotensin system inhibitors

I am cautious in claiming this observed centre-level variation represents true centre-level variation in AMI treatment in England since I am using unlinked MINAP data in this study, which I demonstrated in Chapter 7 to be insufficient on its own to ascertain AMI hospitalisations in England.

In further exploratory analyses which I have not included in the draft manuscript, I used the variation in AMI treatment observed in this cohort to define the potential instrument which could be used in a future comparative effectiveness IV analysis: the tendency to treat with an invasive versus conservative cardiac strategy at the cardiology centre level in the 1-year prior to the individual's AMI hospitalisation.

This instrument was defined similarly to the instrument in Case Study 1 (the TTP alternative second-line oral antidiabetic treatments).

In defining this instrument, I restricted the study population to those hospitalised for NSTEMI with an eGFR<60mL/min/1.73m^2, since this is the study population of interest in future comparative effectiveness work (more details provided in the thesis discussion section 10.5.2.). I considered the relevance and the exchangeability assumptions by calculating the partial F-statistic to measure the strength of the association between this instrument and the treatment and by plotting the standardised levels of covariates across deciles of the instrument (Appendix I.6. to I.8.). These preliminary analyses suggest that this variation in treatment could be a valid instrument. However, these analyses are premature, and the assumptions will need to be re-considered in linked secondary care data (MINAP and HES).

This research serves as pilot work to understand cardiology centre-level variation in invasive cardiac strategies for people with impaired kidney function hospitalised for AMI, particularly NSTEMI. In future work, I will use nationally representative linked data from primary and secondary care (including HES and MINAP data, see thesis discussion section 10.5.2.). I will confirm that the centre-level variation in AMI treatment among people with impaired kidney function observed in this chapter persists in these linked data. I then plan to use this variation, if present, as a preference-based instrument in an IV analysis, using a similar IV design as I presented in Chapters 4 to 6 for Case Study 1.

9.3. References

1. Peracha J, Pitcher D, Santhakumaran S, et al. Centre variation in mortality following post-hospitalization acute kidney injury: analysis of a large national cohort. *Nephrol Dial Transplant*. Oct 19 2022;37(11):2201-2213. doi:10.1093/ndt/gfab348

2. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge University Press; 2006.

3. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med*. May-Jun 2004;2(3):204-8. doi:10.1370/afm.141

CHAPTER 10. DISCUSSION

10.1. Overview

This thesis aimed to advance the use of routinely collected health data to study the comparative effectiveness of treatments. Within this aim, I presented research from two case studies: (1) alternative second-line oral antidiabetic treatments for people with T2DM managed in primary care, and (2) alternative AMI treatments among people with kidney impairment managed in secondary care. Below, I repeat the specific objectives of this thesis from the introduction, by case study.

Case Study 1: Alternative second-line antidiabetic treatments for people with T2DM

1A: To examine inequalities in second-line antidiabetic treatment prescribing according to sociodemographic characteristics which are likely to be important potential confounders when studying the comparative effectiveness of these drugs (Chapter 3).

1B: To inform the selection of potential instruments for natural experiments comparing alternative second-line antidiabetic treatments by investigating treatment variation at the clinical commissioning group (CCG) level (Chapters 3 to 5).

1C: To design and conduct an IV analysis to estimate the relative effectiveness of alternative second-line antidiabetic treatments with respect to outcomes important to patients, healthcare providers, and policymakers (Chapter 4 to 5).

1D: To investigate heterogeneous treatment effects across the target population of people with T2DM in English primary care (Chapter 6).

Case Study 2: Alternative AMI treatments for people with kidney impairment

2A: To investigate potential biases in defining a study population of people hospitalised for AMI with reduced kidney function in English secondary care using primary (NCKDA) and secondary care (MINAP and HES) data sources (Chapter 7).

2B: To examine inequalities in AMI treatment and outcomes by kidney function (Chapter 8).

2C: To explore variation in AMI treatment strategies at the individual and cardiology centre-level to inform the selection of a preference-based instrument for future comparative effectiveness studies using an IV analysis (Chapter 9).

In the methods overview (Chapter 2), I outlined the approach I followed in both case studies to meet the aim and objectives of this thesis (**Figure 2.1.**). Below, I repeat these steps, which are each mapped to the specific chapters and objectives of this thesis in **Figure 2.1.**

1. Ensure the main data source is suitable to define the study population and treatment.
2. Understand important patient-level characteristics that influence treatment receipt.
3. Examine variation in treatment receipt to inform the selection of a preference-based instrument.
4. Design an IV analysis with input from scientists and patients.
5. Conduct an IV analysis to estimate the comparative treatment effects overall and by subgroups.

The remainder of the discussion section will be structured as follows: In Section 10.2., I summarise the main findings from the research papers I presented in **Chapters 3-9** and relate them back to the steps I used to study comparative treatment effects in routinely collected health data (**Figure 2.1.**). In Section 10.3., I discuss the main contributions of this thesis related specifically to (i) advancements in using routinely collected health data for pharmacoepidemiological research, and (ii) advancements in clinical understanding of treatments for CVD, kidney disease, and T2DM. In Section 10.4., I discuss the main limitations of this thesis and the efforts I made to minimise the impact of these limitations on my results. In Section 10.5., I discuss the clinical implications of my research and areas for future work, and summarise in general how the approach I took to study the comparative effectiveness of treatment can be used to inform HTA decision-making. In Section 10.6., I summarise my personal learning and development during my PhD studies. Finally, I conclude this thesis in Section 10.7.

10.2. Summary of main findings

10.2.1. Case Study 1: Second-line antidiabetic treatment for people with T2DM

In this case study, I first investigated sociodemographic inequalities in second-line oral antidiabetic treatment prescribing in a general population of people with T2DM managed in English primary care (Chapter 3). I found statistically significant but small differences in the probabilities of being prescribed the three main alternative second-line treatments in the UK. Those of ethnic minorities and from more deprived areas compared with those of white ethnicity and those from more deprived areas, respectively, had lower predicted probabilities of being prescribed SGLT2i versus DPP4i and SU, after adjusting for potential confounders as fixed effects and clustering by CCG as a random effect.

Following this study, I designed a comparative effectiveness analysis using an IV to compare the causal effect of the same alternative oral antidiabetic treatments on clinical measures (HbA1c, BMI, SBP, and eGFR) and kidney and CVD outcomes (Chapter 4). The IV design was informed by previous work which demonstrated variation in second-line antidiabetic prescribing across CCGs in England[1] and confirmed in the data used in this thesis (Chapter 5). I used this treatment variation to define the TTP for each alternative second-line antidiabetic treatment (Chapters 5 to 6).

In the comparative effectiveness analysis using an IV analysis (Chapter 5), I emulated an ideal trial to compare SGLT2i, DPP4i, and SU, all added to metformin, as second-line oral antidiabetic treatments. I found evidence that SGLT2i added to metformin monotherapy as second-line oral antidiabetic treatment was better than the alternatives (DPP4i and SU) on average at reducing mean HbA1c, BMI, and SBP across a general population of people with T2DM managed in English primary care. There was also some evidence that SGLT2i were better than the alternatives (DPP4i and SU) on average at reducing the hazard of heart failure hospitalisation and MAKE. I did not find evidence that SGLT2i were better than the same alternatives at reducing mean change in eGFR or reducing the hazards of MACE and all-cause death. Nor did I find evidence of heterogenous treatment effects when stratifying by prevalent CVD status. Because there was no published trial which directly compared all three of the antidiabetic drugs of interest in this study, it was challenging to triangulate these results with RCT data, since RCTs do not make the same strong, mostly untestable IV assumptions and are usually unbiased by confounding.

Thus, in the second IV analysis (Chapter 6), I emulated a published trial[2] to confirm that the IV analysis can emulate the findings of an RCT, and to explore heterogenous treatment effects. For this study, I focused on the comparison of DPP4i and SU as second-line oral antidiabetic treatments, since there are several trials which directly compare these treatments. After selecting a suitable trial,[2] I found that the LIV analysis successfully emulated the ATE for the primary outcome of the published trial when restricting to the trial eligible subgroup of the cohort. This trial ATE and the CATE for the trial eligible subgroup of the cohort suggested no evidence of a difference in effect for DPP4i and SU at reducing mean HbA1c at 1-year. The CATEs for the trial ineligible subgroup of the cohort were similar in direction but different in magnitude and with greater uncertainty compared with the trial eligible subgroup of the cohort and the trial population. I also found some evidence of heterogeneity according to baseline HbA1c: there was some evidence that people grouped into the highest level of baseline HbA1c had greater reductions in HbA1c at 1-year when treated with DPP4i versus SU.

10.2.2. Case Study 2: AMI treatment for people with kidney impairment

In this case study, I first investigated differential AMI case ascertainment in MINAP and HES by level of kidney function (Chapter 7). I found that people with moderate to severe kidney impairment were more likely to have an AMI captured in both MINAP and HES compared with people with mild kidney impairment. I also found that eGFR captured in MINAP data within 24 hours of AMI hospitalisation varies compared with the most recent eGFR captured in primary care data; however, there was no systematic bias resulting in either a consistent under- or over-estimation in MINAP data compared with primary care data. Agreement between these two data sources when categorising people as having moderate to severe kidney impairment (eGFR<60mL/min/1.73m^2) was moderate to good.

In the same cohort, I then investigated the association between kidney impairment and AMI treatment and outcomes (Chapter 8). I found that people with lower levels of kidney function defined using the most recent eGFR in primary care data had lower odds and predicted probability of being treated with an invasive cardiac strategy compared with people with higher levels of kidney function. I also found that people with lower levels of kidney function had worse outcomes during the AMI hospitalisation and post-discharge (among those who survived the AMI hospitalisation).

Finally, I present pilot work to describe variation in AMI treatment strategies across cardiology centres in England using a larger copy of MINAP data (Chapter 9). In this cross-sectional study, I found substantial variation in invasive versus conservative cardiac strategies to treat AMI. This variation was particularly pronounced for people hospitalised for NSTEMI and with impaired kidney function. Further work is needed to confirm this variation in linked secondary care data (MINAP and HES). This variation could then be used as a preference-based IV to investigate the comparative effectiveness of alternative AMI treatment strategies among people with kidney impairment hospitalised for NSTEMI.

10.3. Contributions

In this section, I outline the main original contributions I make to the pharmacoepidemiological literature in this thesis. I organise my contributions according to 2 main categories:

1. Advancements in using routinely collected health data for pharmacoepidemiological research.
2. Advancements in clinical understanding of treatments for CVD, kidney disease, and T2DM

The individual contributions are summarised in **Table 10.1.** I provide greater detail to accompany this table in Sections 10.3.1. to 10.3.2.

Table 10.1. Summary of main original contributions of this thesis to pharmacoepidemiological research

Contribution	Details	Chapter(s) associated with this contribution
Advancements in using routinely collected health data for pharmacoepidemiological research		
Applying an IV analysis within the TTE framework.	I illustrate an example of how the TTE framework can be combined with an IV analysis to generate rigorous evidence on the comparative effectiveness of alternative antidiabetic treatments.	4, 5, 6
Demonstrating the importance of epidemiological triangulation when using a TTE framework.	I highlight the importance of triangulating evidence from my TTE and IV analyses with RCT data to support the IV analysis and transport results from an RCT to a general population, allowing for treatment heterogeneity.	5, 6
Illustrating challenges in the English secondary care setting for pharmacoepidemiology research.	I highlight the need for using linked data in the English secondary care setting to avoid selection bias due to conditioning on a collider.	7
Advancements in clinical understanding of treatments for CVD, kidney disease, and T2DM		
New evidence on the comparative effectiveness of SGLT2i versus DPP4i versus SU, all added to metformin, as second-line oral antidiabetic treatments for people with T2DM.	I provide observational evidence that SGLT2i are more effective at reducing mean HbA1c, BMI, and SBP (versus DPP4i and SU), and reducing the hazards of heart failure hospitalisation (versus DPP4i) and MAKE (versus SU). I also demonstrate treatment effect heterogeneity when comparing the effectiveness of DPP4i versus SU at reducing mean HbA1c at 1-year follow-up.	5, 6
Inequalities in alternative second-line antidiabetic treatment prescribing by sociodemographic characteristics.	I demonstrate that people of ethnic minorities and from more deprived areas have a lower probability of being prescribed SGLT2i as second-line antidiabetic treatment compared with white people and people from less deprived areas, although these disparities were small.	3

Contribution	Details	Chapter(s) associated with this contribution
Inequalities and variation in invasive versus conservative cardiac strategies to treat AMI among people with impaired kidney function.	I demonstrate that people with lower levels of kidney function have lower probabilities of invasive cardiac management for AMI compared with people with higher levels of kidney function. I also demonstrate variation across cardiology centres in invasive versus conservative cardiac strategies.	8, 9

ATE: average treatment effect; BMI: body-mass index; DPP4i: dipeptidyl peptidase-4 inhibitors; HbA1c: haemoglobin A1c; IV: instrumental variable; MAKE: major adverse kidney event; RCT: randomised controlled trial; T2DM: type 2 diabetes mellitus; TTE: target trial emulation; SBP: systolic blood pressure; SGLT2i: sodium-glucose co-transporter 2 inhibitors; SU: sulfonylureas

10.3.1. Advancements in using routinely collected health data for pharmacoepidemiological research

Applying an IV analysis within a target trial emulation framework

In this thesis, my research was motivated by areas of genuine clinical uncertainty, where available RCT evidence is lacking due in part to limited generalisability to general patient populations and the use of placebo comparators rather than active comparators. To address this evidence gap, I used the TTE framework[3,4] to design and conduct comparative effectiveness analyses in routinely collected health data using an IV analysis.[5] This framework involves taking the key elements of a trial design and articulating how they can be emulated in an observational study. The goal of this exercise is for observational researchers to clearly define key aspects of the observational study's protocol to recognise and avoid risks of bias when comparing exposures/treatments such as confounding, immortal time bias, selection bias, etc.[3,4,6]

Previous research has described the challenges and opportunities of using IV analyses within the target trial framework to generate evidence on the comparative effectiveness of treatments.[7,8] These challenges include identifying a suitable instrument which meets the IV assumptions and specifying the causal estimand (in many IV studies, the LATE is estimated under the monotonicity assumption, which relates only to the 'compliers' – a difficult subgroup to define[9,10] (see section 2.4.4.)). However, the 2SRI and LIV analyses can estimate the ATE to better emulate an RCT.[11-13] Moreover, the LIV can estimate the ATE in the presence of essential heterogeneity, i.e. where unobserved characteristics act as effect modifiers of the treatment effect. As demonstrated in Chapter 6, using the LIV can improve the transportability of results from an RCT with strict inclusion/exclusion criteria to general populations. Finally, the IV analysis allows researchers to avoid assuming no unmeasured confounding, which is a major limitation of most pharmacoepidemiological studies using routinely collected health data to estimate causal treatment effects.

Both the TTE framework and the IV analysis face criticism. Pearce and Vandenbroucke (2023) argue that the TTE framework narrows the focus of causal inference in observational research and should not be considered gold-standard for observational studies seeking to uncover causal effects. Alternative methods could be more suited to address causal research questions.[14] In response, De Stavola et al (2023) agree that the TTE framework should not be considered the gold-standard, but argue it is a useful framework to be applied where appropriate to enhance the quality of comparative effectiveness studies.[15] Both agree that the framework benefits from 'mapping' complementary causal inference methods and analyses to it, including IV analyses.[14,15]

Similarly, debate exists about the appropriateness of IV analyses in pharmacoepidemiology. Soumerai and Koppel (2017) argue that IV analyses make strong assumptions which cannot be tested and result in low-quality evidence which is not useful to policymakers and not easily interpretable by the general public.[16] In response, Keele and Small (2019) argue that IV analyses can generate useful evidence which does not rely on the assumption of no unmeasured confounding, as most other observational studies must assume.[17] The authors acknowledge the alternative assumptions the IV analysis must make and recommend careful scrutiny when designing and conducting this type of analysis.[17] Further, a commentary by Baiocchi (2019) commended the accessibility of the reporting of an IV analysis[18] in a study comparing the mortality risks for being transferred to an intensive care unit versus continued treatment on a general ward.[19] This demonstrates the possibility of conveying the complex theories underlying IV analyses in a way that is interpretable by a non-expert.

In this thesis, the combination of a TTE and IV analysis necessitated the careful consideration and articulation of the area of clinical uncertainty, the data sources to be used, the methods, analysis, and estimand of greatest clinical interest, and presentation and triangulation of results. Cogent descriptions of these analyses were important in publishing the research paper presented in Chapter 5 at a general medical journal and will be important to disseminate this work to a broad audience of patients, clinicians, statisticians, and HTA agencies.

Epidemiological triangulation

The critiques by Pearce and Vandenbroucke (2023), De Stavola et al (2023), and Keele and Small (2019) agree that epidemiological triangulation across different studies prone to different biases is important to appraise and synthesise research aiming to make causal inferences.[14,15,17] In this thesis, I contribute an applied example where epidemiological triangulation is possible and necessary when using the TTE framework to (i) contextualise the findings from the IV analysis in Chapter 5 with published RCT data, and (ii) to verify reproducibility of the ATE in a published RCT using an IV analysis in routinely collected health data, and to transport the results to the target population in which clinical uncertainty remains. These triangulation exercises strengthen the conclusions of my comparative effectiveness analyses and allow me to explore treatment heterogeneity where trial evidence is not necessarily generalisable.

Illustrating challenges in the English secondary care setting for pharmacoepidemiology research

Unlike primary care, secondary care in the UK is organised by disease specialty. Pharmacoepidemiological studies using routinely collected health data from this setting face

challenges due to the fragmentation of data across sources like HES and disease-specific audits. In this thesis, I illustrate the potential collider bias[20,21] which can be induced by using unlinked secondary care data.

In Chapter 7, I highlighted the association between AMI case ascertainment in two secondary care datasets (MINAP and HES) and (i) kidney function and (ii) death during the AMI hospitalisation and post-discharge (among survivors). With a simple DAG, I illustrated how selecting on, for example, capture in MINAP data, could induce a spurious association between an exposure/treatment and outcome, like invasive cardiac treatment strategy and death in-hospital, because of this particular form of selection bias.

Other studies have demonstrated the pervasiveness of collider bias in observational research (e.g., risk factor studies for severe COVID-19 disease[22]) as well as in RCTs (by not accounting for loss to follow-up).[20] I add to the literature highlighting the benefits of using linked routinely collected health data for observational research with an example of where collider bias could impact observational studies set in UK secondary care. This work informs my own future research, where I will minimise the risk of this selection bias by using linked secondary care data (further details in section 10.5.2.).

10.3.2. Advancements in clinical understanding of treatments for CVD, kidney disease, and T2DM

New evidence on the comparative effectiveness of SGLT2i versus DPP4i versus SU, all added to metformin, as second-line oral antidiabetic treatments for people with T2DM

In Case Study 1 of this thesis (Chapter 5), I demonstrated that initiating second-line oral antidiabetic treatment with SGLT2i were more effective at reducing mean HbA1c, BMI, and SBP (versus DPP4i and SU), and at reducing the hazards of heart failure hospitalisation (versus DPP4i) and MAKE (versus SU) among a general population of people with T2DM. There was no evidence of heterogenous treatment effects by cardiovascular disease status at baseline. This observational evidence adds to other observational[23-25] and randomised evidence,[26-31] and a network meta-analysis[32] supporting more general uptake of SGLT2i among people with T2DM and other cardiovascular and kidney diseases to improve outcomes and reduce future healthcare burdens to the NHS. I also demonstrated potential heterogenous treatment effects for DPP4i versus SU among the subgroup of people with the highest levels of HbA1c at baseline (≥88mmol/mol) (Chapter 6).

Inequalities in alternative second-line oral antidiabetic treatment prescribing by sociodemographic characteristics

I demonstrated statistically significant but small inequalities in second-line antidiabetic treatment according to ethnicity (ethnic minorities having a lower predicted probability of receiving SGLT2i compared with white people) and deprivation (people from more deprived areas having a lower predicted probability of receiving SGLT2i compare with people from less deprived areas).[33] This was encouraging as more pronounced disparities in other settings (e.g., USA) by ethnicity and SES have been reported,[34] as well as disparities in other aspects of T2DM treatment in the UK.[35-41] In this research paper, I also demonstrated a lower proportion of people being prescribed SGLT2i with prevalent CVD compared with people without prevalent CVD during the study period (2014-2020), which had been observed in a previous study.[42] This finding was incongruous with NICE guidelines (2015 and updated in 2022),[43,44] which focused recommendations for SGLT2i in the 2015 guideline for people with prevalent CVD based on CVD safety trial data.[26] Since the guideline update in 2022, which more strongly recommends SGLT2i among people at high risk of or with CVD and/or kidney disease, the prescribing of these drugs should be more common in these patient groups.

Inequalities and variation in invasive versus conservative cardiac strategies to treat AMI among people with impaired kidney function

I added to the available literature[45-48] demonstrating a persistently lower probability of invasive versus conservative cardiac management strategy among people with kidney impairment in English secondary care[49] (Chapter 8). These inequalities were particularly evident among people hospitalised for NSTEMI and with moderate to severe kidney impairment (eGFR<60mL/min/1.73m^2) (Chapters 8-9).

European and USA cardiology guidelines recommend invasive management for NSTEMI cases at higher risk of adverse outcomes.[50,51] These guidelines recognise the lack of RCT data addressing this question, but cite observational research supporting invasive management among people with CKD.[52,53] NICE guidelines suggest caution when managing people with NSTEMI with comorbidities that could increase the risk of adverse events from an invasive management strategy,[54] which may perhaps contribute to the variation observed at the individual- and cardiology-centre level which I present in Chapter 9. Future work is needed to confirm this centre-level variation across England in linked secondary care data (see section 10.5.).

10.4. Limitations

10.4.1. IV assumptions

Like any modelling strategy, the IV analysis must make assumptions which cannot be fully tested. Of the 4 main IV assumptions I could only formally test one (that the instrument was strongly associated with the treatment).

The exchangeability assumption assumes that the IV is not associated with unmeasured confounders. I cannot observe unmeasured confounders, so I could not formally test this assumption. However, I used a falsification test,[55] plotting the standardised values of the covariates by deciles of the IV (the TTP) to observe any associations between the IV and measured confounders. Only year of second-line treatment initiation showed a strong association with deciles of the IV. I adjusted for all measured confounders in the IV models, which would have accounted for confounding by year of second-line treatment initiation. But I cannot know if unmeasured confounders impacted the analysis. Violations are common in prescriber preference IV studies where patients with specific characteristics 'doctor shop' based on the known preferences of the provider.[56] However, in the UK it is unlikely that people choose their GP (which are grouped into CCGs) based on their prescribing preference.

I could not test that the IV affects the outcome only via the treatment (the exclusion restriction). I relied on DAGs and clinical reasoning to assume that the IV only affected the outcome via the treatment. However, it's possible that those CCGs which preferred SGLT2i could have also preferred delivering other aspects of care which are independently associated with improved outcomes for people with T2DM.[5] I could have explored some measure of practice quality (e.g., meeting Quality and Outcomes Framework (QOF) targets for other aspects of T2DM unrelated to second-line antidiabetic prescribing) to explore this assumption; however, this would require a bespoke arrangement with CPRD to generate the data.

Finally, I could not test for the monotonicity assumption as I could not observe counterfactual treatments within the same individuals. Swanson et al conducted a survey of physicians with hypothetical patients demonstrating potential violations of this assumption with a prescriber-preference IV.[9] I could not do a similar survey, since these data are de-identified.

The target trial emulation of the Nauck et al (2017) published trial[2] was not only important to investigate treatment heterogeneity, but also to support the IV assumptions I made in both research papers presented in Chapters 5 and 6. Successfully emulating the ATE from the published trial within the trial eligible cohort offers support for the IV assumptions; however, this does not prove these assumptions are valid. The ATE estimated using the IV in Chapter 6 was subject to uncertainty demonstrated by wide 95% CI. Although I applied the same inclusion and exclusion criteria to the

published trial, the RCT population and the trial eligible population were quite different in terms of important markers like mean baseline HbA1c, mean age, etc.

10.4.2. Residual confounding

In Case Study 1, I rely on the assumptions of the IV being valid to minimise the risk of residual confounding. However, violations of these assumptions, particularly the exchangeability assumption, could bias the treatment effects. I compared the IV analysis with other methods which measure the same estimand but do not rely on the IV assumptions. These included a traditional multivariable regression and a doubly robust propensity score-IPTW-RA analysis. They showed similar results, but also assume no unmeasured confounding. My triangulation exercise with RCTs was reassuring in that the results from the IV analysis agreed with those from RCTs, albeit the study populations and contrasts were substantially different between the RCTs and my own study.

In Case Study 2, I used traditional multivariable regression to study the association between kidney function and invasive versus conservative cardiac strategies and other outcomes such as death in hospital and post discharge. In these analyses, I adjusted for measured confounders including age, sex, ethnicity, comorbidities, and year of admission. However, these covariates were likely not perfectly measured in the routinely collected data, and I omitted several unmeasured confounders likely to bias the association such as frailty. The assumption of no unmeasured confounding is most likely violated in these analyses, hence the need for future research in linked datasets applying IV methods to account for residual confounding (section 10.5.2.).

10.4.3. Chance

The IV analysis is inefficient (i.e., the statistical power is generally less than that of similar analyses using traditional multivariable regression approaches).[5] This problem can be reduced by identifying an instrument which strongly predicts the treatment (the relevance assumption). We observe the relative inefficiency of the IV compared with other approaches when we compare the results from the main analysis in Chapters 5 and 6 with those from traditional multivariable regression – the latter have narrower 95% confidence intervals. Thus, the lack of evidence for any difference in the comparative effectiveness of the three alternative antidiabetic treatments we report for certain outcomes could simply be due to chance.

10.4.4. Selection bias

As I explained in the methods section 2.2.1., selection bias is unlikely to have impacted Case Study 1 since people with T2DM are managed mainly in primary care, which is the 'gateway' to the NHS healthcare service. The primary care data we use are representative of the UK population in terms of age and sex.[57,58]

In contrast, selection bias is a problem in Case Study 2 which is set in English secondary care. These healthcare services are generally organised by organ systems delivered by specialist consultants.[59] This results in many health datasets which are organ or disease-specific. These disease audits, like MINAP, are rich data sources describing the treatment pathway for acute coronary syndrome hospitalisations, including AMI.[60] However, these disease-focused audits make defining people with impaired kidney function who are hospitalised for AMI challenging. As I explain in the research papers in Chapters 7 to 9, the MINAP audit was designed to study particular types of acute coronary syndromes. Selection into this dataset has changed as the audit has evolved since its inception in 2000, but the persistent under-reporting of NSTEMI cases has been recognised as a limitation, possibly due to variation in resources allocated to data capture across cardiology centres.[60]

This selection is described in the general population in previous work,[61] and in the impaired kidney function population in Chapter 7. As a consequence of this selection, collider bias risks inducing spurious associations between NSTEMI treatments and outcomes when using unlinked MINAP data. This bias can occur if these treatments, or upstream variables which influence treatment (e.g., kidney function), and outcomes (e.g., death in hospital) directly influence being captured in MINAP data, which I observed in the research paper in Chapter 7.

In Chapter 9, I used an unlinked MINAP dataset (2014-19) to describe AMI treatment variation across cardiology centres in England. Because of the potential impact of collider bias in these unlinked data, I must confirm this variation in AMI treatment among people with impaired kidney function exists across hospitals in England. Using linked secondary care data to ascertain AMI cases will be important in my future work (see section 10.5.2.).

10.4.5. Treatment misclassification

In Case Study 1, I assumed a prescription recorded by the GP meant a person was taking the drug as prescribed. English routinely collected health data are limited by the lack of dispensing data from pharmacies, which are available in other countries like Scotland[62] and Denmark.[63] It is possible that

people in my study did not take their treatments as prescribed, which would likely have diluted the true treatment effect in the comparative effectiveness analyses (Chapters 5 to 6).

In Case Study 2, AMI treatment pathways are complex and not necessarily provided during the index hospitalisation, particularly for NSTEMI. MINAP and HES data each provide data describing angiography and PCI services delivered in secondary care. However, I did not include the National Audit of Percutaneous Coronary Interventions (NAPCI) audit dataset.[64] Similar to AMI case ascertainment, capture of angiography and PCI services may vary across secondary care data. Differential capture based on kidney function could lead to biases in future comparative effectiveness work for these treatments among people with impaired kidney function. In future analyses, I will explore the ascertainment of angiography and PCI services across these secondary care data to understand the potential bias due to AMI treatment misclassification.

10.4.6. Outcome misclassification

SGLT2i have been shown in trials to be particularly beneficial at preventing kidney function decline.[30,31,65] It was therefore important to highlight kidney-related outcomes in the comparative effectiveness work (Case Study 1) in Chapter 5. I followed previous RCT and observational studies in using time-to-a 40% decline in eGFR as a proxy for substantial kidney function decline.[66] Further, I used a composite outcome, MAKE, including 40% decline in eGFR, kidney replacement therapy (including dialysis and transplantation), and all-cause death.[67] Misclassification of MAKE could have been reduced by incorporating data from kidney-specialist clinics held by the UK Renal Registry,[68] an audit dataset capturing incident dialysis and AKI events at secondary and tertiary nephrology clinics across England. However, linkage with CPRD is not routine for these data and therefore was not feasible within my PhD registration period.

Furthermore, in Case Study 1, I adjusted for the baseline outcome measure when investigating the change score for clinical measures as an outcome (mean change in HbA1c, BMI, SBP, and eGFR). This introduces a risk of bias due to measurement error and regression to the mean since extreme values of baseline clinical measures are more likely to be less extreme in subsequent measures.[69,70] An added concern is adjusting for baseline outcome measures when these measures are mediators, and not confounders.[71] In my study, the baseline clinical measures were measured prior to or on the same day as second-line treatment initiation, implying these baseline outcome measures were confounders and not mediators of the treatment prescribed on the outcomes of interest. Excluding these clinical measures from the regression model as a covariate would draw serious concerns from the target

clinical audience when presenting these results. Thus, I chose to focus my interpretation of the mean change in these outcomes conditional on the baseline values (i.e., an ANCOVA analysis).

10.4.7. Missing data

I used a complete case analysis under the complete case assumption[72,73] in most studies included in this thesis (Chapters 3, 6, 7, 8, 9). However, in the comparative effectiveness analysis for SGLT2i, DPP4i, and SU in Chapter 5, the primary analysis used MICE to impute missing covariate and outcome data with 5 imputations.[74,75] This approach similarly assumes data were missing at random; however, a complete case analysis was not favourable in this case due to the substantial amount of missing outcome data for continuous clinical measures, particularly the primary outcome (HbA1c – 34% of the study population were missing this measure at 1-year follow-up). While HbA1c should be measured at least once per year by the GP,[44] in practice in our data this target was missed or else HbA1c measures were not captured in the routine primary care data. The MICE added significant computational intensity to the analysis but yielded similar results compared with the alternative complete case analysis. Thus, in the subsequent IV analysis presented in Chapter 6, a complete case approach to handle missing data was taken.

10.4.8. Internal validity

I evaluated the assumptions made in the primary analyses of both case studies with a variety of alternative and sensitivity analyses to understand the internal validity of my studies.

In Case Study 1, I had to balance the need for thorough evaluations of assumptions with the computational intensity of the analyses. For example, I could have considered alternative definitions of the IV. The impact of using different time periods to define the instrument, e.g., prescribing history in the 6-months prior to the index, could have been considered. However, I focused the alternative analyses in Chapter 5 on handling missing data and comparing the IV analysis with alternative analyses which do not rely on the IV assumptions (e.g., traditional multivariable regression which assumes no unmeasured confounding). There were no major differences across these alternative analyses which supported the internal validity of the study. The internal validity of the IV analysis in Chapter 5 was strengthened by the target trial emulation in Chapter 6, where the ATE from a published trial agreed with the ATE from an IV analysis which used the same instrument as was used in Chapter 5.

In Case Study 2, I relied heavily on the assumption of no unmeasured confounding when investigating associations between kidney function and outcomes like AMI case ascertainment, AMI treatment

received, and death. As discussed in section 10.4.2., the results are likely biased by residual confounding. Further, as discussed in section 10.4.4., the selection on the collider (being captured in MINAP data) for the research paper presented in Chapter 9 must be explored in future work using linked national data.

10.4.9. What I would do differently

Writing this thesis and considering its limitations prompts reflection and consideration of what could have been done to improve this body of work. Cutting across both case studies is the limitation that the key assumptions of the analysis (either the IV assumptions or the assumption of no unmeasured confounding) cannot be fully tested. I thread arguments as to why these assumptions may or may not be violated throughout this work; however, a falsification test, namely, the incorporation of a negative and/or positive control outcome,[76,77] would be a useful addition to this work. As I explain in section 1.2.5. of the thesis introduction, a negative control outcome is an outcome for which the treatment is known to have no causal effect, and the confounding structure of the treatment and negative control outcome is the same as the treatment and outcome of primary interest. Any treatment effect observed with the negative control outcome would suggest the analysis has not minimised the risk of confounding bias. A positive control outcome is similar, but an outcome for which the treatment is expected to have some causal effect. These falsification tests are considered useful tools in pharmacoepidemiological studies, and are useful in IV analyses to support the plausibility of the IV assumptions.[76]

Previous work has used incident cancer diagnosis and cataract surgery when comparing SGLT2i versus GLP1-RA, assuming these outcomes are not caused by either treatment and share a similar confounding structure.[78] These could have been similarly applied in my own analyses, although people with a history of cancer would need to be excluded from the analysis which is a substantial proportion of the study population. Another study investigating similar treatments used road traffic accidents as a negative control.[23] However, this negative control outcome may not be suitable in the context of the IV analysis, since the IV analysis may lack the power to detect any confounded association between the treatments and road traffic accidents. Further, it is not clear that road traffic accidents would share a similar confounding structure as my primary causal effect of interest.

A potential positive control outcome that could also have been useful in Case Study 1 would have been genital infection. SGLT2i are known to increase the risk of genital infections.[79] Observing this adverse treatment effect in Chapter 5 would have been helpful to assuage concerns about the IV analysis strategy.

10.5. Implications for clinical practice and future work

10.5.1. Case Study 1: Second-line antidiabetic treatment for people with T2DM

The results from this thesis support other observational and randomised evidence suggesting that SGLT2i are efficacious and effective second-line oral antidiabetic treatments to improve key clinical measures (e.g., HbA1c) and reduce the risk of adverse CVD and kidney outcomes.[78,80-83] NICE guidelines were most recently updated in 2022 to recommend SGLT2i as the preferred second-line antidiabetic treatment added to metformin for people at risk of or with CVD, and also for people with T2DM and CKD.[44] Future updates should consider the evidence supporting benefits of SGLT2i beyond these subgroups at reducing the risk of adverse CVD and kidney disease progression, particularly as SGLT2i reach the end of their patented period and prices decrease.[84]

In Chapters 5 and 6, I focused analyses on outcomes at 1- and 2-years follow-up. I made this decision because by 2-years follow-up, many people had switched antidiabetic treatments. The ITT approach of my study did not account for this treatment switching, but instead considered the impact of initiating second-line treatment with each drug, irrespective of intercurrent events.[85] Beyond 2-years follow-up, many people in my study were censored (administratively or because of death). Thus, my findings are not easily interpretable beyond the 2-year follow-up period. In line with NICE recommendations for future research,[44] I am co-leading a study not included in this thesis which is investigating the counterfactual outcomes for people initiating second-line antidiabetic treatment with SU, DPP4i, or SGLT2i added to metformin, in the same primary care dataset. Using a microsimulation model developed using US routinely collected health data,[86] the dynamic antidiabetic treatment patterns over follow-up are taken into account to predict the long-term impact of second-line antidiabetic drugs for outcomes important to T2DM patients. This study is addressing a limitation of the research presented in Case Study 1, which excluded GLP1-RA as a treatment of interest. This was necessary because this drug class is not currently recommended as an option for second-line antidiabetic treatment.[44] However, in this microsimulation modelling work, I am able to consider antidiabetic treatment beyond second-line therapy.

NICE also recommends studying the effectiveness of SGLT2i in different ethnic groups, since the risk of certain micro and macrovascular events is different across ethnicities.[44] While I described ethnic differences in SGLT2i prescribing versus the alternatives in Chapter 3, I did not investigate heterogenous treatment effects by ethnicity in the comparative effectiveness of SGLT2i versus the alternatives in Chapter 5. Investigating these treatment disparities was unrealistic in my IV analysis, which would have been underpowered to detect clinically important differences in these subgroups

due to the small proportion of people of non-white ethnicity in the study population. However, our PPI panellists stressed that this is an important research area, and as NICE recommends,[44] should be prioritised in future research.

10.5.2. Case Study 2: AMI treatment for people with kidney impairment

Inequalities in AMI treatment, particularly NSTEMI, has been described previously.[45-48] In this case study, I demonstrated that these inequalities persist in more recent years, and these inequalities appear to exist for STEMI cases (to a much lesser extent) as well as NSTEMI in a population of people with kidney impairment. However, more work is needed to (i) understand these complex care pathways and how they impact treatment decisions in people with kidney impairment, and (ii) understand misclassification of treatment status and selection of patients being captured across different secondary care datasets (e.g., MINAP, NAPCI, HES), particularly since data from these sources are already being used by NICE to recommend AMI treatment.[54] This work will facilitate higher quality observational research studying these treatment inequalities and comparative effectiveness analyses for alternative AMI treatments, in the persistent absence of randomised data on AMI treatment for people with impaired kidney function.[87]

In my future work, I will use nationally representative and linked primary and secondary care data to confirm variation in NSTEMI treatment across hospital centres in England. If confirmed, I will exploit this variation as an instrument in an LIV analysis to investigate the comparative effectiveness of invasive versus conservative cardiac strategies for people with impaired kidney function hospitalised for NSTEMI. Outcomes of interest will include mortality, kidney outcomes (e.g., AKI, dialysis) and CVD outcomes (e.g., AMI readmission). At the time of writing this thesis, I have permissions in place to access two datasets: (1) the 'NHS Data Lake' used for health services planning and auditing, and (2) the British Heart Foundation (BHF) Data Science Centre CVD-COVID-UK/COVID-IMPACT Secure Data Environment[88] (referred to as the BHF Data Science Centre SDE hereafter) (**Table 10.2.**)

Table 10.2. Datasets and their purposes for future analyses to study the comparative effectiveness of alternative AMI treatment among people with kidney impairment.

Type of data	**Database 1: NHS Data Lake**	**Database 2: BHF Data Science Centre SDE**	**Purpose**

Primary care	CVD-PREVENT[89] (linkage not yet in place)	GDPPR[90]	• Identify those with eGFR <60mL/min/1.73m^2 (derived from serum creatinine lab test results) • Confounders (relevant comorbidities, demographic information, etc)
Secondary care	HES APC[91]	HES APC[91]	• Identify study population (people admitted to hospital for ACS, and subtypes of ACS (particularly NSTEMI)) • Confounders (relevant comorbidities, demographic information, etc) • Define exposure status (invasive cardiac strategy used?) • Define outcomes (death in hospital, AMI readmission, etc.)
	MINAP[60] and NAPCI[64] (NICOR audits)	MINAP[60] and NAPCI[64] (NICOR audits)	• Identify study population (people admitted to hospital for ACS, subtype of ACS) • Confounders (relevant comorbidities, demographic information, etc) • Define exposure status (invasive cardiac strategy used?) • Define outcomes (death in hospital, AMI readmission, etc.)
Kidney-specific audit data	UKRR[68]	N/A (linkage not available)	• Identify kidney-related outcomes (AKI, dialysis, kidney transplant) • Identify those with CKD4/5 in renal units.
Death data	Civil Registry Deaths	Civil Registry Deaths	• Outcomes (death)

ACS: acute coronary syndrome; AKI: acute kidney injury; CKD: chronic kidney disease; GDPPR: General Practice Extraction Service (GPES) Data for Pandemic Planning and Research; HES: Hospital Episode Statistics; MINAP: Myocardial Ischaemia National Audit Project; UKRR: UK Renal Registry

The main advantages and disadvantages of using each dataset for future work are described in **Table 10.3.**

Table 10.3. Summary of the main advantages and disadvantages of the databases for future work investigating the comparative effectiveness of alternative AMI treatments in people with impaired kidney function

Data environment	Advantages	Disadvantages
Database 1: NHS Data Lake	• Linkage to UKRR to minimise misclassification of kidney outcomes.	• Primary care data linkage pending. • No direct access to these data for non-NHS analysts.
Database 2: BHF Data Science Centre SDE	• Linkage for primary care and secondary care data (including HES and NICOR audits) available.	• No linkage to UKRR. • SDE still only available for COVID-19 related studies and incident cohort begins follow-up in 2019.

BHF: British Heart Foundation; HES: Hospital Episode Statistics; NHS: National Health Service; NICOR: National Institute for Cardiovascular Outcomes Research; SDE: Secure Data Environment; UKRR: UK Renal Registry

The NHS Data Lake (dataset 1 in **Table 10.2.**) is limited because of the lack of an active linkage to primary care data. Without a primary care linkage, it is challenging to accurately define a study population with kidney impairment as I demonstrated in Chapter 7 (HES does not have laboratory data to derive eGFR and MINAP eGFR is subject to substantial random error). Although commissioners hope that CVD-PREVENT data[89] will be linked in this data environment, this linkage is not guaranteed. Further, I do not have direct access to these data. My analysis code must be transferred and implemented by an internal NHS analyst with permission to access these data. However, this data resource has the particular advantage of being linked to the UKRR, which will reduce the misclassification of kidney outcomes (see also section 10.4.6.) that will be particularly important in Case Study 2.

The BHF Data Science Centre SDE (dataset 2 in **Table 10.2.**) is limited by the lack of a linkage to the UKRR to define kidney outcomes. Further, this SDE was created in response to the COVID-19 pandemic; only COVID-related studies are currently permitted in this data environment. Further, the cohort included in these data was defined in 2019 – thus, follow-up for incident events can only begin after this date. Because it is possible that the variation in AMI treatment across hospitals in England was impacted by the COVID-19 pandemic, I adapted the protocol to meet the requirements for the BHF Data Science Centre SDE to include studying the changes in AMI treatment variation before and after

the COVID-19 pandemic began in 2020. A summary of my approved study in this database is published on the BHF website.[92]

Triangulating results from these two databases, each with its own advantages and disadvantages, will be helpful in supporting the causal inferences made in the planned comparative effectiveness analyses. Alternative nationally representative databases like CPRD[57,58] and OpenSAFELY[93] could also be useful, but have similar disadvantages, with no routine linkage available to the UKRR (CPRD) and MINAP (CPRD and OpenSAFELY) at the time of writing this thesis.

10.5.3. Interconnectivity of Case Study 1 and 2 and implications for clinical practice

CVD, kidney disease, and T2DM are interrelated conditions that share common risk factors and are commonly co-prevalent among people living with multiple long-term conditions.[94] Interventions, including pharmacological therapies, that improve risk factors and outcomes for one of these conditions may also directly or indirectly improve risk factors and outcomes for another. As the prevalence of people living with multiple long-term conditions increases,[59,95,96] it is becoming increasingly important to identify and understand therapies whose pharmacological actions directly or indirectly improve outcomes across these commonly co-prevalent disease areas.

In Case Study 1 of this thesis, I found evidence that SGLT2i were better than the alternative second-line antidiabetic treatments at improving important clinical risk factors (HbA1c, BMI, SBP) and also reducing the hazards of CVD and kidney outcomes. This case study restricted the study population to people with an eGFR≥30mL/min/1.73m^2, since NICE guidelines recommend alternative treatments for people with an eGFR<30, although the drug license is for anyone with eGFR>15, and the EMPA Kidney trial showed benefits of SGLT2i among people with eGFR as low as 15.[30] My general conclusions from this case study were that increasing SGLT2i prescribing among this study population would improve patient outcomes and reduce pressure on the NHS by, for example, reducing the risks of heart failure hospitalisations and severe kidney function decline.

These findings are important in the context of Case Study 2, where I studied those living with kidney impairment who are at substantially increased risk of AMI and adverse outcomes following AMI compared with people with normal kidney function.[97,98] Increasing SGLT2i prescribing among people living with and without T2DM[30,65,99,100] and thereby reducing the risk of heart failure hospitalisation and kidney function decline could improve outcomes for those that are hospitalised for AMI. While my own findings (Chapter 5) and findings from other trials do not demonstrate that SGLT2i have a strong protective effect on AMI incidence itself,[28,29] reducing the proportion of people who go on to develop

moderate to severe CKD would indirectly improve outcomes by way of maintaining higher kidney function in this high-risk patient population.

10.5.4. A role for qualitative research when designing a natural experiment using an IV analysis

In this thesis, an important assumption I made was that the TTP alternative second-line antidiabetic treatments is largely due to CCG-level directives that are not associated with patient-level factors, and that this introduced an element of randomness in treatment allocation that could be exploited in a natural experiment. While I was able to demonstrate the lack of any strong association with patient-level factors across levels of the instrument (to investigate the exchangeability assumption of the IV analysis), I could not evaluate whether the TTP of the CCG was independent of unmeasured or more accurately measured patient-level factors and other variables that could have acted as confounders of the instrument and outcome.

Qualitative research in the form of standardised interviews and focus groups to understand CCG, hospital, and GP prescribing/treatment patterns, and what factors are most important in predicting these treatment decisions would be helpful to understand whether the CCG-level prescribing preference for alternative antidiabetic treatments (Case Study 1) or hospital-level treatment preference for AMI hosptialisations (Case Study 2) are suitable and valid instruments. This type of research was beyond the scope of this PhD thesis, but would add valuable context and support in the major assumptions made in the IV analyses.

Previous work published in 2007 by RAND Europe (commissioned by the National Audit Office) described factors that influence GP prescribing in England by conducting qualitative research in the form of interviews, focus groups, and workshops at two primary care trusts in England with relatively high and relatively low adherence to NICE guidance on statin prescribing.[101] This research found that prescribing behaviour at the primary care trust level was influenced by "local guidelines, newsletters, site visits by prescribing advisers, personali[s]ed contacts, and recommendations from specialist consultants in the secondary health setting."[101] At the GP practice level, factors which seemed to most strongly influence prescribing were "the professional experience of the GP, the clinical needs of the patient, patient demand, peer networks, and drug company representatives".[101] Updated qualitative research focused on antidiabetic treatment prescribing and AMI treatment decisions in primary and secondary care, respectively, would be useful to understand whether the variation in these clinical treatment settings is suitable to be used as an instrument in natural experiments to evaluate the comparative effectiveness of these drugs.

10.5.5. General discussion on incorporating 'real-world evidence' in HTA

The research I present in this thesis has illustrated the challenges and opportunities in generating high-quality observational evidence which can be used by HTA agencies to inform healthcare services and policy. The challenges I experienced align with the barriers towards increasing the acceptability of evidence on comparative treatment effects using routinely collected health data outlined in a recent perspective piece by Gomes et al (2024).[102] These challenges are grouped into 4 major themes: (i) "limitations of guidelines" and best practices acceptable to HTA agencies, (ii) "challenges with following best practice, data quality, and access", (iii) "the potential for unexplained biases", and (iv) "ingrained reluctance or inability to accept even high-quality NRS by HTA agencies". [102] In **Table 10.4.** I explain how the research in this thesis maps to these challenges and also the NICE real-world evidence framework to conduct and report studies using real-world data published in 2022.[103]

Table 10.4. Mapping the barriers to uptake of observational evidence in HTA[102] to (i) the NICE real-world evidence framework to conduct and report studies using real-world data[103] and (ii) to the approach I used in this thesis to study comparative treatment effects in routinely collected health data

Barriers to uptake of observational evidence in HTA (Gomes et al, 2024)[102]	**NICE real-world evidence framework for conducting and reporting studies using real-world data[103]**	**Approach taken in this thesis**	**Details**
"Limitations of guidelines" and best practices acceptable to HTA agencies	NICE has published a 'living' framework which outlines how to conduct and report studies using real-world data to generate high-quality observational evidence.	I followed an approach, illustrated in **Figure 2.1.**, which aligns with guidance from NICE to study comparative treatment effects using routinely collected health data.	**Figure 2.1.** clearly illustrates the approach I took to study comparative treatment effects in routinely collected health data. This approach aligns with the NICE guidance on how to conduct and report 'real-world evidence studies'.[103]
"Challenges with following best practice, data quality, and access"	First principle for evidence generation: "Ensure data is of good and known provenance, relevant and of sufficient quality to answer the research question".	Step 1 to 3 of **Figure 2.1.**: 1. Ensure main data source is suitable to define the study population. 2. Understand important patient-level characteristics that influence treatment receipt. 3. Examine variation in treatment to inform the selection of a preference-based instrument.	I carefully considered the suitability of the routinely collected health data I used in both case studies of this thesis before conducting any comparative effectiveness analyses. I described patient-level (e.g., socioeconomic characteristics) and group-level (e.g., CCG or cardiology centre) variation in treatment. • Inequalities in health and healthcare services are highlighted by NICE as an area where routinely collected health data can make important contributions.[103] • Exogenous variation in treatment was needed to apply an IV analysis in a comparative effectiveness study to reduce bias from confounding.
"The potential for unexplained biases"	Third principle for evidence generation: "Use analytical methods that minimise the risk of	Steps 4 to 5 of **Figure 2.1.:** 4. Design a comparative effectiveness analysis with input from a multidisciplinary team.	I published a study protocol and SAP to clearly articulate the analytical methods to be used to minimise the risk of bias and characterise uncertainty in my comparative effectiveness analyses. The research papers presented in Chapters 5 to 6 followed recommendations from NICE to use a TTE framework, combined with

Barriers to uptake of observational evidence in HTA (Gomes et al, 2024)[102]	NICE real-world evidence framework for conducting and reporting studies using real-world data[103]	Approach taken in this thesis	Details
	bias and characterise uncertainty".	5. Conduct a comparative effectiveness analysis to estimate comparative treatment effects.	advanced methods to account for confounding. The IV analysis minimised the risk of measured and unmeasured confounding subject to meeting several untestable assumptions (Chapters 5 to 6) and accounted for essential heterogeneity (Chapter 6) in transporting the results of a published RCT to the target population of people with T2DM treated in primary care. I use a variety of alternative and sensitivity analyses to consider the risk of bias impacting the conclusions of the studies. I also emphasise the need for epidemiological triangulation to compare the findings of my own research, which are vulnerable to certain biases, with that from other studies vulnerable to different biases. I specifically triangulate to other RCTs; but I highlight the challenges in making comparisons due to differences in the study populations and treatment comparisons.
"Ingrained reluctance or inability to accept even high-quality NRS by HTA agencies"	Second principle for evidence generation: "Generate evidence in a transparent way and with integrity from study planning through to study conduct and reporting".	Steps 4-5 of **Figure 2.1.:** 4. Design a comparative effectiveness analysis with input from a multidisciplinary team. 5. Conduct a comparative effectiveness analysis to estimate comparative treatment effects.	I published a study protocol and SAP to clearly articulate the analytical methods to be used to minimise the risk of bias and characterise uncertainty in my comparative effectiveness analyses. I wrote the research paper in an accessible format to be widely understood by non-experts in causal inference and IV analyses. The paper was accepted by a general medical journal (the *BMJ*), indicating this paper will be of interest to a wide audience and potentially considered in future NICE guideline updates.

CCG: Clinical Commissioning Group; HTA: health technology assessment; IV: instrumental variable; NICE: National Institute for Health and Care Excellence; NRS: non-randomised studies; RCT: randomised controlled trial; SAP: statistical analysis plan; TTE: target trial emulation

Despite the challenges outlined by Gomes et al (2024), there are many examples of real-world evidence being used in HTA to develop and update guidelines for drug treatments.[104,105] Traditionally, observational research studying new interventions has been limited by the slow accrual and access to data on the newly rolled-out intervention. This creates a lag period which makes it difficult to study new treatments in populations not included in RCTs. However, rapid analyses of novel COVID-19 vaccines in observational research using sources like OpenSAFELY have tackled this issue and paved the way for more timely and rapid evidence generation for consideration by HTA agencies.[93,106]

The recent publication of the NICE real-world evidence framework (2022) signals a desire for HTA agencies to make better use of high-quality 'real-world evidence' generated from increasingly available large routinely collected health data[88,93] and advanced research and analytical methods. This framework is not overly prescriptive, since the datasets, methods, and analyses suitable for individual studies will vary based on the research context. However, the framework does highlight the TTE framework[3,4] as particularly suited for studies using RWD to estimate comparative treatment effects. Recent studies have demonstrated the usefulness of this framework in the HTA-context.[8,107] My research adds to this literature to demonstrate how the TTE framework can be combined with other advanced study designs to reduce the risk of bias in the observational research setting.

I agree with NICE guidance that the TTE framework is particularly suited to inform HTA; the advantages of designing an observational study by mapping key aspects of the study design to the ideal trial are clear.[6] Applications of TTE studies which successfully emulate,[108] or even predict results from RCTs,[109,110] build confidence in high-quality comparative effectiveness research. These studies also highlight opportunities to develop methods to transport RCT results to target populations for which HTA need evidence. Moreover, this framework may improve understanding and confidence in the observational methods and analyses applied in RWD for HTA committee members and stakeholders who are more familiar with evidence generated by RCTs. This improved understanding and confidence would hopefully alleviate the reluctance of HTA agencies to accepting high-quality observational evidence (the fourth barrier defined by Gomes et al (2024), **Table 10.4.**).

While I share the enthusiasm with NICE for the TTE framework, I agree with Pearce and Vandenbrouke (2023) that this framework is not optimal in all causal inference research, and that ultimately epidemiological triangulation is key to drawing causal inferences. The NICE real-world evidence framework mentions "triangulation" twice in the downloadable corporate document, while "target trial" is mentioned 13 times at the time of writing this thesis.[111] Future iterations should consider enhancing the guidance on epidemiological triangulation, with case studies, as a cornerstone to causal inference.

10.6. Personal learning

I experienced substantial development as an epidemiologist over the course of my PhD studies. This development can be grouped into four broad categories: research design and conduct, team science, statistical analyses and health econometrics, and presentation and communication skills.

10.6.1. Research design and conduct

Over the course of my PhD registration, I gained considerable experience in research design by leading ethics applications, protocol and SAP writing, successful grant applications, and national and international collaborations. I appreciated the chance to develop my skills in data management and analyses in Stata and R, interpreting and presenting complex analyses, and drafting manuscripts and responding to peer reviewers. My research skills development greatly benefited from the formal and informal training I received from the research teams I worked with over the course of my PhD.

10.6.2. Team science

I was fortunate to work in collaborative multidisciplinary teams to enhance the rigour and relevance of my research outputs for diverse audiences and stakeholders. Specifically, I have benefited from being a team member of the LSHTM Electronic Health Records (EHR) research group, the QECKD study and NACARAI study team, and the PERMIT study team (see project descriptions in section 2.1.).

The collaborative and relatively horizontal management structure of the LSHTM EHR research group facilitated ample opportunities for me to learn from other group members. I was able to draw upon a pool of expertise in routine health data, epidemiology, medicine, and statistics from team members at all seniority levels to improve my own work using routinely collected health data for pharmacoepidemiological research.

The QECKD and NACARAI teams include a diverse group of clinicians, statisticians, database experts (NCKDA, MINAP, and the UK Renal Registry), and patients based throughout the UK. These experts filled my own knowledge gaps related to the clinical realities of the complex AMI care pathways for people with kidney disease and the resultant ‘messy’ data collected from primary and secondary care.

Similarly, the PERMIT team includes a diverse group of clinicians, statisticians, health economists, policymakers leading HTA, and patients based throughout the UK and the US. I received substantial formal and informal training on applying IV methodology to answer pharmacoepidemiological

questions from this team, as well as insight into clinical care pathways for people with T2DM. In return, I offered the team formal and informal training on the datasets used in this project, and informal training on how to contextualise our findings to the broader epidemiological landscape. The intensity of the analyses included in Chapters 5 and 6, as well as other research not included in this thesis, required a division of labour. In particular, the IV analyses, bootstrapping, and MI in Chapter 5 required a parallelisation of work and computing across our virtual and physical machines to develop analysis code and reduce its running time from weeks to days.

My PhD experience is one example where team science was important to producing better research. Other examples were particularly evident during the COVID-19 pandemic, where collaborations between industry, academia, regulators, software engineers, etc facilitated rapid responses to the pandemic. These responses popularised new ways of working (e.g., trusted research environments/SDE) and produced high-quality evidence which saved lives and improved the research landscape for RCTs and observational research alike.[112]

It seems intuitive that cross-disciplinary teamwork should lead to higher impact research. However, traditional academic incentives and structures often place focus on the individual and their own success at obtaining first or last author publications and success in obtaining grants.[113] I was fortunate during this PhD to be part of several multidisciplinary teams which worked in a collaborative and collegial manner to share knowledge and boost the impact of our research without a myopic focus on the success of any one individual.

10.6.3. Statistical analyses and health econometrics

Over the course of my PhD, my knowledge and appreciation of statistical methods in epidemiology and health econometrics grew substantially. I received formal and informal training from the PERMIT team on IV theory and analyses. My supporting role in the target trial emulation and transportability research paper on which I am second author presented in Chapter 6 introduced me to the concept of essential heterogeneity which is more widely acknowledged in health econometrics. Through my formal and informal training and self-directed learning, I enhanced my knowledge of other statistical concepts which I supported and interpreted in the research papers presented in Chapters 5 to 6 including MI, propensity scores, IPTW, and doubly robust methods.

10.6.4. Presentation and communication skills to scientists and general audiences

Fortunately, I had many opportunities to enhance my presentation and communication skills during my PhD studies. I presented research from this PhD at 3 national and international conferences, as well as several meetings and events, both internal (LSHTM) and external (Imperial College London, McMaster University, and the UK Renal Registry). I was the PERMIT project lead on a PPI panel, which was a useful session both to inform the research methods and to contextualise our findings to a non-scientific expert audience. For both the PERMIT project (Case Study 1) and the QECKD and NACARAI projects (Case Study 2), I was most often the lead analyst presenting results to the grant holders and advisory committees. Finally, as part of my employment at LSHTM, I teach on several MSc modules. Over the course of the PhD, I have progressed from leading practicals for introductory epidemiology and statistics modules to developing and delivering lectures and module organising for advanced courses in analysis of EHR data and advanced research methods.

10.7. Conclusions

This thesis aimed to advance the use of routinely collected health data to study the comparative effectiveness of treatments. This thesis carefully considered two areas of clinical uncertainty, namely second-line oral antidiabetic treatment among people with T2DM and alternative AMI treatment among people with kidney impairment. I described inequalities in treatment, treatment variation at the health care provider group level, and the comparative effectiveness of treatments according to outcomes important to patients, clinicians, and policy makers.

In this thesis, I applied the TTE framework with an IV analysis to minimise the risk of bias, including confounding, in comparative effectiveness analyses of treatments. I made original contributions to the epidemiological literature to advance the use of routinely collected health data for pharmacoepidemiological studies and advance clinical knowledge regarding treatments for CVD, kidney disease, and T2DM. These original contributions can be triangulated with other high-quality observational and randomised evidence to contribute to future iterations of NICE guidelines.

In the discussion, I have outlined important areas of future research which can consider the approach of applying the TTE framework and IV analysis, as well as other advanced quantitative methods, to draw useful inferences from routinely collected health data. Essential to this work was multidisciplinary teamwork which enhanced the rigour and relevance of this research to the clinical and HTA context. Applying these principles of team science with advanced methodologies and analyses in causal inference can help improve the quality of observational research using routinely collected health data to improve clinical practice and outcomes for patients.

10.8. References

1. Wilkinson S, Douglas I, Stirnadel-Farrant H, et al. Changing use of antidiabetic drugs in the UK: trends in prescribing 2000–2017. *BMJ Open*. 2018;8(7):e022768. doi:10.1136/bmjopen-2018-022768

2. Nauck MA, Meininger G, Sheng D, Terranella L, Stein PP. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes, obesity & metabolism*. Mar 2007;9(2):194-205. doi:10.1111/j.1463-1326.2006.00704.x

3. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

4. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

5. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. Jun 15 2014;33(13):2297-340. doi:10.1002/sim.6128

6. Fu EL. Target Trial Emulation to Improve Causal Inference from Observational Data: What, Why, and How? *Journal of the American Society of Nephrology*. 2023;34(8)

7. Swanson SA. Instrumental Variable Analyses in Pharmacoepidemiology: What Target Trials Do We Emulate? *Current Epidemiology Reports*. 2017/12/01 2017;4(4):281-287. doi:10.1007/s40471-017-0120-1

8. Moler-Zapata S, Hutchings A, O'Neill S, Silverwood RJ, Grieve R. Emulating Target Trials With Real-World Data to Inform Health Technology Assessment: Findings and Lessons From an Application to Emergency Surgery. *Value in Health*. 2023;26(8):1164-1174. doi:10.1016/j.jval.2023.04.010

9. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and evaluation of the monotonicity condition for preference-based instruments. *Epidemiology (Cambridge, Mass)*. May 2015;26(3):414-20. doi:10.1097/ede.0000000000000279

10. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*. 1996;91(434):444-455. doi:10.2307/2291629

11. Basu A, Coe NB, Chapman CG. 2SLS versus 2SRI: Appropriate methods for rare outcomes and/or rare exposures. *Health Econ*. Jun 2018;27(6):937-955. doi:10.1002/hec.3647

12. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*. 2008/05/01/ 2008;27(3):531-543. doi:https://doi.org/10.1016/j.jhealeco.2007.09.009

13. Basu A. ESTIMATING PERSON-CENTERED TREATMENT (PeT) EFFECTS USING INSTRUMENTAL VARIABLES: AN APPLICATION TO EVALUATING PROSTATE CANCER TREATMENTS. *J Appl Econ (Chichester Engl)*. June/July 2014;29(4):671-691. doi:10.1002/jae.2343

14. Pearce N, Vandenbroucke JP. Are Target Trial Emulations the Gold Standard for Observational Studies? *Epidemiology (Cambridge, Mass)*. Sep 1 2023;34(5):614-618. doi:10.1097/ede.0000000000001636

15. De Stavola BL, Gomes M, Katsoulis M. Transparency and Rigor: Target Trial Emulation Aims to Achieve Both. *Epidemiology (Cambridge, Mass)*. Sep 1 2023;34(5):624-626. doi:10.1097/ede.0000000000001638

16. Soumerai SB, Koppel R. The Reliability of Instrumental Variables in Health Care Effectiveness Research: Less Is More. *Health Serv Res*. Feb 2017;52(1):9-15. doi:10.1111/1475-6773.12527

17. Keele L, Small D. Instrumental variables: Don't throw the baby out with the bathwater. *Health Serv Res*. 2019;54(3):543-546. doi:10.1111/1475-6773.13130

18. Baiocchi M. Which Deteriorating Ward Patients Benefit From Transfer to the Intensive Care Unit?: Critically Engaging Methods in a Well-Designed Natural Experiment. *JAMA Network Open*. 2019;2(2):e187698-e187698. doi:10.1001/jamanetworkopen.2018.7698

19. Grieve R, O'Neill S, Basu A, Keele L, Rowan KM, Harris S. Analysis of Benefit of Intensive Care Unit Transfer for Deteriorating Ward Patients: A Patient-Centered Approach to Clinical Evaluation. *JAMA Network Open*. 2019;2(2):e187704-e187704. doi:10.1001/jamanetworkopen.2018.7704

20. Holmberg MJ, Andersen LW. Collider Bias. *Jama*. 2022;327(13):1282-1283. doi:10.1001/jama.2022.1820

21. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol*. Feb 1 2018;47(1):226-235. doi:10.1093/ije/dyx206

22. Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun*. Nov 12 2020;11(1):5749. doi:10.1038/s41467-020-19478-2

23. Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. *The Lancet Diabetes & Endocrinology*. 2023;11(9):644-656. doi:10.1016/S2213-8587(23)00171-7

24. Khunti K, Charbonnel B, Cooper A, et al. Associations between second-line glucose-lowering combination therapies with metformin and HbA1c, body weight, quality of life, hypoglycaemic events

and glucose-lowering treatment intensification: The DISCOVER study. *Diabetes, Obesity and Metabolism*. 2021;23(8):1823-1833. doi:https://doi.org/10.1111/dom.14400

25. Wilkinson S, Williamson E, Pokrajac A, et al. Comparative effects of sulphonylureas, dipeptidyl peptidase-4 inhibitors and sodium-glucose co-transporter-2 inhibitors added to metformin monotherapy: a propensity-score matched cohort study in UK primary care. *Diabetes, Obesity and Metabolism*. 2020;22(5):847-856. doi:https://doi.org/10.1111/dom.13970

26. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *New England Journal of Medicine*. 2015/11/26 2015;373(22):2117-2128. doi:10.1056/NEJMoa1504720

27. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *New England Journal of Medicine*. 2017/08/17 2017;377(7):644-657. doi:10.1056/NEJMoa1611925

28. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine*. 2019/01/24 2018;380(4):347-357. doi:10.1056/NEJMoa1812389

29. Cannon CP, Pratley R, Dagogo-Jack S, et al. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. *New England Journal of Medicine*. 2020;383(15):1425-1435. doi:10.1056/NEJMoa2004967

30. Empagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2022;388(2):117-127. doi:10.1056/NEJMoa2204233

31. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *New England Journal of Medicine*. 2019/06/13 2019;380(24):2295-2306. doi:10.1056/NEJMoa1811744

32. Type 2 diabetes [B] Pharmacological therapies with cardiovascular and other benefits in people with type 2 diabetes, NICE Guideline NG28. Web. National Institute for Health and Care Excellence (NICE). 2022. https://www.nice.org.uk/guidance/ng28/evidence/b-pharmacological-therapies-with-cardiovascular-and-other-benefits-in-people-with-type-2-diabetes-pdf-10956473392

33. Bidulka P, Mathur R, Lugo-Palacios DG, et al. Ethnic and socioeconomic disparities in initiation of second-line antidiabetic treatment for people with type 2 diabetes in England: A cross-sectional study. *Diabetes, Obesity and Metabolism*. 2023;25(1):282-292. doi:https://doi.org/10.1111/dom.14874

34. Eberly LA, Yang L, Eneanya ND, et al. Association of Race/Ethnicity, Gender, and Socioeconomic Status With Sodium-Glucose Cotransporter 2 Inhibitor Use Among Patients With

Diabetes in the US. *JAMA Network Open*. 2021;4(4):e216139-e216139. doi:10.1001/jamanetworkopen.2021.6139

35. Eastwood SV, Mathur R, Sattar N, Smeeth L, Bhaskaran K, Chaturvedi N. Ethnic differences in guideline-indicated statin initiation for people with type 2 diabetes in UK primary care, 2006–2019: A cohort study. *PLoS medicine*. 2021;18(6):e1003672. doi:10.1371/journal.pmed.1003672

36. Mathur R, Farmer RE, Eastwood SV, Chaturvedi N, Douglas I, Smeeth L. Ethnic disparities in initiation and intensification of diabetes treatment in adults with type 2 diabetes in the UK, 1990–2017: A cohort study. *PLoS medicine*. 2020;17(5):e1003106. doi:10.1371/journal.pmed.1003106

37. Khunti K, Wolden ML, Thorsted BL, Andersen M, Davies MJ. Clinical Inertia in People With Type 2 Diabetes. *Diabetes care*. 2013;36(11):3411. doi:10.2337/dc13-0331

38. Khunti K, Godec TR, Medina J, et al. Patterns of glycaemic control in patients with type 2 diabetes mellitus initiating second-line therapy after metformin monotherapy: Retrospective data for 10 256 individuals from the United Kingdom and Germany. *Diabetes, obesity & metabolism*. Feb 2018;20(2):389-399. doi:10.1111/dom.13083

39. Khunti K, Seidu S. Therapeutic Inertia and the Legacy of Dysglycemia on the Microvascular and Macrovascular Complications of Diabetes. *Diabetes care*. 2019;42(3):349. doi:10.2337/dci18-0030

40. Khunti K, Gomes MB, Pocock S, et al. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: A systematic review. https://doi.org/10.1111/dom.13088. *Diabetes, Obesity and Metabolism*. 2018/02/01 2018;20(2):427-437. doi:https://doi.org/10.1111/dom.13088

41. Khunti S, Khunti K, Seidu S. Therapeutic inertia in type 2 diabetes: prevalence, causes, consequences and methods to overcome inertia. *Therapeutic Advances in Endocrinology and Metabolism*. 2019/01/01 2019;10:2042018819844694. doi:10.1177/2042018819844694

42. Wilkinson S, Douglas IJ, Williamson E, et al. Factors associated with choice of intensification treatment for type 2 diabetes after metformin monotherapy: a cohort study in UK primary care. *Clin Epidemiol*. 2018;10:1639-1648. doi:10.2147/CLEP.S176142

43. NG28: Type 2 diabetes in adults: management. NICE. https://www.nice.org.uk/guidance/ng28

44. NICE guideline [NG28]: Type 2 diabetes in adults: management. Web. NICE. Accessed 10 June, 2024. https://www.nice.org.uk/guidance/ng28/chapter/Recommendations#reviewing-drug-treatments

45. Wong JA, Goodman SG, Yan RT, et al. Temporal management patterns and outcomes of non-ST elevation acute coronary syndromes in patients with kidney dysfunction. *European heart journal*. 2009;30(5):549-557. doi:10.1093/eurheartj/ehp014

46. Panchal HB, Zheng S, Devani K, et al. Impact of Chronic Kidney Disease on Revascularization and Outcomes in Patients with ST-Elevation Myocardial Infarction. *American Journal of Cardiology*. 2021;150:15-23. doi:10.1016/j.amjcard.2021.03.057

47. Fox CS, Muntner P, Chen AY, et al. Use of evidence-based therapies in short-term outcomes of ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction in patients with chronic kidney disease: a report from the National Cardiovascular Data Acute Coronary Treatment and Intervention Outcomes Network registry. *Circulation*. 2010/01// 2010;121(3):357-365. doi:10.1161/circulationaha.109.865352

48. Shaw C, Nitsch D, Steenkamp R, et al. Inpatient Coronary Angiography and Revascularisation following Non-ST-Elevation Acute Coronary Syndrome in Patients with Renal Impairment: A Cohort Study Using the Myocardial Ischaemia National Audit Project. *PLOS ONE*. 2014;9(6):e99925. doi:10.1371/journal.pone.0099925

49. Scott J, Bidulka P, Taylor DM, et al. Management and outcomes of myocardial infarction in people with impaired kidney function in England. *BMC Nephrology*. 2023/11/02 2023;24(1):325. doi:10.1186/s12882-023-03377-x

50. Lawton JS, Tamis-Holland JE, Bangalore S, et al. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. Jan 18 2022;145(3):e4-e17. doi:10.1161/cir.0000000000001039

51. Byrne RA, Rossello X, Coughlan JJ, et al. 2023 ESC Guidelines for the management of acute coronary syndromes: Developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC). *European heart journal*. 2023;44(38):3720-3826. doi:10.1093/eurheartj/ehad191

52. Szummer K, Lundman P, Jacobson SH, et al. Influence of renal function on the effects of early revascularization in non-ST-elevation myocardial infarction: data from the Swedish Web-System for Enhancement and Development of Evidence-Based Care in Heart Disease Evaluated According to Recommended Therapies (SWEDEHEART). *Circulation*. Sep 8 2009;120(10):851-8. doi:10.1161/circulationaha.108.838169

53. Huang HD, Alam M, Hamzeh I, et al. Patients with severe chronic kidney disease benefit from early revascularization after acute coronary syndrome. *International Journal of Cardiology*. 2013/10/09/ 2013;168(4):3741-3746. doi:https://doi.org/10.1016/j.ijcard.2013.06.013

54. NICE guideline [NG185]: Acute coronary syndrome. Web. NICE. Accessed March 10, 2022. https://www.nice.org.uk/guidance/ng185

55. Pizer SD. Falsification Testing of Instrumental Variables Methods for Comparative Effectiveness Research. *Health Serv Res*. Apr 2016;51(2):790-811. doi:10.1111/1475-6773.12355

56. Zhang L. *Using physician's prescribing preference as an instrumental variable in comparative effectiveness research*. University of Glasgow; 2023. https://theses.gla.ac.uk/84016/

57. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. Jun 2015;44(3):827-36. doi:10.1093/ije/dyv098

58. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology*. 2019;48(6):1740-1740g. doi:10.1093/ije/dyz034

59. Whitty CJM, MacEwen C, Goddard A, et al. Rising to the challenge of multimorbidity. *BMJ (Clinical research ed)*. 2020;368:l6964. doi:10.1136/bmj.l6964

60. Herrett E, Smeeth L, Walker L, Weston C, on behalf of the MAG. The Myocardial Ischaemia National Audit Project (MINAP). *Heart (British Cardiac Society)*. 2010;96(16):1264. doi:10.1136/hrt.2009.192328

61. Herrett E, Shah AD, Boggon R, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ : British Medical Journal*. 2013;346:f2350. doi:10.1136/bmj.f2350

62. Prescribing practice and dispensing pharmacy open data. Web. Public Health Scotland. Accessed 8 April, 2024. https://publichealthscotland.scot/publications/prescribing-practice-and-dispensing-pharmacy-open-data/prescribing-practice-and-dispensing-pharmacy-open-data-data-from-january-to-march-2023/

63. Johannesdottir SA, Horváth-Puhó E, Ehrenstein V, Schmidt M, Pedersen L, Sørensen HT. Existing data sources for clinical epidemiology: The Danish National Database of Reimbursed Prescriptions. *Clin Epidemiol*. 2012;4:303-13. doi:10.2147/clep.S37587

64. National Audit of Percutaneous Coronary Intervention (NAPCI). Web. National Institute of Cardiovascular Outcomes Research (NICOR). Accessed 1 April, 2024. https://www.nicor.org.uk/~documents/ncap/pci/10633-nicor-annual-summary-reports-napci-v5-ac/?layout=default

65. Heerspink HJL, Stefánsson BV, Correa-Rotter R, et al. Dapagliflozin in Patients with Chronic Kidney Disease. *New England Journal of Medicine*. 2020;383(15):1436-1446. doi:10.1056/NEJMoa2024816

66. Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug

Administration. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. Dec 2014;64(6):821-35. doi:10.1053/j.ajkd.2014.07.030

67. Levin A, Agarwal R, Herrington WG, et al. International consensus definitions of clinical trial outcomes for kidney failure: 2020. *Kidney Int*. Oct 2020;98(4):849-859. doi:10.1016/j.kint.2020.07.013

68. Registry UR. *UK Renal Registry 25th Annual Report - data to 31/12/2021*. 2023. https://ukkidney.org/audit-research/annual-report/25th-annual-report-data-31122021

69. Glymour MM. Commentary: Modelling change in a causal framework. *Int J Epidemiol*. Oct 13 2022;51(5):1615-1621. doi:10.1093/ije/dyac151

70. Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM. When is baseline adjustment useful in analyses of change? An example with education and cognitive change. *Am J Epidemiol*. Aug 1 2005;162(3):267-78. doi:10.1093/aje/kwi187

71. Tennant PWG, Arnold KF, Ellison GTH, Gilthorpe MS. Analyses of 'change scores' do not estimate causal effects in observational data. *Int J Epidemiol*. Oct 13 2022;51(5):1604-1615. doi:10.1093/ije/dyab050

72. Bartlett JW, Carpenter JR, Tilling K, Vansteelandt S. Improving upon the efficiency of complete case analysis when covariates are MNAR. *Biostatistics*. Oct 2014;15(4):719-30. doi:10.1093/biostatistics/kxu023

73. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med*. Dec 10 2010;29(28):2920-31. doi:10.1002/sim.3944

74. Powney M, Williamson P, Kirkham J, Kolamunnage-Dona R. A review of the handling of missing longitudinal outcome data in clinical trials. *Trials*. Jun 19 2014;15:237. doi:10.1186/1745-6215-15-237

75. Lee KJ, Roberts G, Doyle LW, Anderson PJ, Carlin JB. Multiple imputation for missing data in a longitudinal cohort study: a tutorial based on a detailed case study involving imputation of missing outcome data. *International Journal of Social Research Methodology*. 2016/09/02 2016;19(5):575-591. doi:10.1080/13645579.2015.1126486

76. Davies NM, Thomas KH, Taylor AE, et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. *International Journal of Epidemiology*. 2017;46(6):2067-2077. doi:10.1093/ije/dyx014

77. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology (Cambridge, Mass)*. 2010;21(3):383-388. doi:10.1097/EDE.0b013e3181d61eeb

78. Htoo PT, Buse J, Cavender M, et al. Cardiovascular Effectiveness of Sodium-Glucose Cotransporter 2 Inhibitors and Glucagon-Like Peptide-1 Receptor Agonists in Older Patients in Routine Clinical Care With or Without History of Atherosclerotic Cardiovascular Diseases or Heart Failure. *J Am Heart Assoc*. Feb 15 2022;11(4):e022376. doi:10.1161/jaha.121.022376

79. Liu J, Li L, Li S, et al. Effects of SGLT2 inhibitors on UTIs and genital infections in type 2 diabetes mellitus: a systematic review and meta-analysis. *Scientific reports*. 2017/06/06 2017;7(1):2824. doi:10.1038/s41598-017-02733-w

80. D'Andrea E, Wexler DJ, Kim SC, Paik JM, Alt E, Patorno E. Comparing Effectiveness and Safety of SGLT2 Inhibitors vs DPP-4 Inhibitors in Patients With Type 2 Diabetes and Varying Baseline HbA1c Levels. *JAMA internal medicine*. 2023;183(3):242-254. doi:10.1001/jamainternmed.2022.6664

81. Zelniker TA, Wiviott SD, Raz I, et al. SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. *The Lancet*. 2019/01/05/ 2019;393(10166):31-39. doi:https://doi.org/10.1016/S0140-6736(18)32590-X

82. Wilding J, Fernando K, Milne N, et al. SGLT2 Inhibitors in Type 2 Diabetes Management: Key Evidence and Implications for Clinical Practice. *Diabetes therapy : research, treatment and education of diabetes and related disorders*. Oct 2018;9(5):1757-1773. doi:10.1007/s13300-018-0471-8

83. Ni L, Yuan C, Chen G, Zhang C, Wu X. SGLT2i: beyond the glucose-lowering effect. *Cardiovascular Diabetology*. 2020/06/26 2020;19(1):98. doi:10.1186/s12933-020-01071-y

84. Taylor SI. The High Cost of Diabetes Drugs: Disparate Impact on the Most Vulnerable Patients. *Diabetes care*. Oct 2020;43(10):2330-2332. doi:10.2337/dci20-0039

85. *ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials*. 2020. 17 Februrary. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf

86. Basu A, Sohn MW, Bartle B, Chan KCG, Cooper JM, Huang E. Development and Validation of the Real-World Progression in Diabetes (RAPIDS) Model. *Medical decision making : an international journal of the Society for Medical Decision Making*. Feb 2019;39(2):137-151. doi:10.1177/0272989x18817521

87. Konstantinidis I, Nadkarni GN, Yacoub R, et al. Representation of Patients With Kidney Disease in Trials of Cardiovascular Interventions: An Updated Systematic Review. *JAMA internal medicine*. Jan 2016;176(1):121-4. doi:10.1001/jamainternmed.2015.6102

88. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ (Clinical research ed)*. Apr 7 2021;373:n826. doi:10.1136/bmj.n826

89. Network OfHIDOaNB. *CVDPREVENT: First Annual Audit Report*. 2021. https://s3.eu-west-2.amazonaws.com/nhsbn-static/CVDPREVENT/2021/CVDPREVENT_First%20Annual%20Audit%20ReportFINAL.pdf

90. General Practice Extraction Service (GPES) Data for pandemic planning and research: a guide for analysts and users of the data. Web. NHS England. Accessed 1 April, 2024. https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research/guide-for-analysts-and-users-of-the-data

91. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *International journal of epidemiology*. 2017;46(4):1093-1093i. doi:10.1093/ije/dyx015

92. Adlam D, Nitsch D, Bidulka P. CCU066: Changes in acute cardiac care of patients with reduced kidney function during the COVID-19 pandemic. Web. British Heart Foundation Data Science Centre. Accessed 2 April, 2024. https://bhfdatasciencecentre.org/projects/ccu066-changes-in-acute-cardiac-care-of-patients-with-reduced-kidney-function-during-the-covid-19-pandemic-2/

93. Andrews C, Schultze A, Curtis H, et al. OpenSAFELY: Representativeness of electronic health record platform OpenSAFELY-TPP data compared to the population of England. *Wellcome Open Res*. 2022;7:191. doi:10.12688/wellcomeopenres.18010.1

94. Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: a comparative risk assessment. *The Lancet Diabetes & Endocrinology*. 2014;2(8):634-647. doi:10.1016/S2213-8587(14)70102-0

95. Multimorbidity: clinical assessment and management: NICE guideline[NG56]. Web. National Institute for Health and Care Excellence. Accessed 22 September, 2021. https://www.nice.org.uk/guidance/ng56

96. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*. 2012;380(9836):37-43. doi:10.1016/S0140-6736(12)60240-2

97. Bagai A, Lu D, Lucas J, et al. Temporal Trends in Utilization of Cardiac Therapies and Outcomes for Myocardial Infarction by Degree of Chronic Kidney Disease: A Report From the NCDR Chest Pain-MI Registry. *J Am Heart Assoc*. Dec 18 2018;7(24):e010394. doi:10.1161/jaha.118.010394

98. Santopinto JJ, Fox KA, Goldberg RJ, et al. Creatinine clearance and adverse hospital outcomes in patients with acute coronary syndromes: findings from the global registry of acute coronary events (GRACE). *Heart (British Cardiac Society)*. Sep 2003;89(9):1003-8. doi:10.1136/heart.89.9.1003

99. McMurray JJV, Solomon SD, Inzucchi SE, et al. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. *New England Journal of Medicine*. 2019/11/21 2019;381(21):1995-2008. doi:10.1056/NEJMoa1911303

100. Packer M, Anker SD, Butler J, et al. Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure. *New England Journal of Medicine*. 2020;383(15):1413-1424. doi:10.1056/NEJMoa2022190

101. Scoggins A, Tiessen J, Ling T, Rabinovich L. *Prescribing in primary care: Understanding what shapes GP's prescribing choices and how might these be changed*. 2007. https://www.rand.org/content/dam/rand/pubs/technical_reports/2007/RAND_TR443.pdf

102. Gomes M, Turner AJ, Sammon C, et al. Acceptability of Using Real-World Data to Estimate Relative Treatment Effects in Health Technology Assessments: Barriers and Future Steps. *Value in Health*. doi:10.1016/j.jval.2024.01.020

103. NICE real-world evidence framework. Web. National Institute for Health and Care Excellence. Accessed 2 April, 2024. https://www.nice.org.uk/corporate/ecd9/chapter/overview

104. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open*. Jun 30 2016;6(6):e011666. doi:10.1136/bmjopen-2016-011666

105. Bullement A, Podkonjak T, Robinson MJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. *International journal of technology assessment in health care*. 2020;36(4):388-394. doi:10.1017/S0266462320000434

106. Hulme WJ, Williamson EJ, Green ACA, et al. Comparative effectiveness of ChAdOx1 versus BNT162b2 covid-19 vaccines in health and social care workers in England: cohort study using OpenSAFELY. *BMJ (Clinical research ed)*. 2022;378:e068946. doi:10.1136/bmj-2021-068946

107. Gomes M, Latimer N, Soares M, et al. Target Trial Emulation for Transparent and Robust Estimation of Treatment Effects for Health Technology Assessment Using Real-World Data: Opportunities and Challenges. *PharmacoEconomics*. Jun 2022;40(6):577-586. doi:10.1007/s40273-022-01141-x

108. Wang SV, Schneeweiss S, Initiative R-D. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *Jama*. 2023;329(16):1376-1385. doi:10.1001/jama.2023.4221

109. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride. *Diabetes care*. Dec 2019;42(12):2204-2210. doi:10.2337/dc19-0069

110. Riddle MC. A Verdict for Glimepiride: Effective and Not Guilty of Cardiovascular Harm. *Diabetes care*. 2019;42(12):2161-2163. doi:10.2337/dci19-0034

111. Excellence NIfHaC. *NICE real-world evidence framework: corporate document*. 2022. 23 June. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

112. Grieve R, Yang Y, Abbott S, et al. The importance of investing in data, models, experiments, team science, and public trust to help policymakers prepare for the next pandemic. *PLOS Global Public Health*. 2023;3(11):e0002601. doi:10.1371/journal.pgph.0002601

113. Kucharski AJ, Funk S, Eggo RM. The COVID-19 response illustrates that traditional academic reward structures and metrics do not reflect crucial contributions to modern science. *PLOS Biology*. 2020;18(10):e3000913. doi:10.1371/journal.pbio.3000913

APPENDICES

APPENDIX A: Ethics approvals

Appendix A1: Ethics approvals for the studies included in this thesis

Study	**Ethics applications/permissions**
Case Study 1 (Chapters 3 to 6)	• Approval from the Independent Scientific Advisory Committee of the Medicines and Health Regulatory Agency (reference 20_064). • LSHTM approval (reference 21395).
Case Study 2 (Chapters 7-8)	• Approval granted from NCKDA Steering Committee. • LSHTM approval (reference 16988). • NCKDA has NHS Research Ethics Committee and Confidentiality Advisory Group section 251 on-going approval (2017-2027).
Case Study 2 (Chapter 9)	• LSHTM ethics approval not needed for variation analyses since it is classed as audit/service evaluation.
Overarching PhD	• LSHTM ethics approval (reference 28740).

APPENDIX B: Chapter 2 – Supplementary methods

Appendix B.1. Key assumptions and study requirements which I relied on in this thesis when estimating the average treatment effect (ATE) using alternative methods to adjust for confounding

Key features of the methods	Details
Main assumptions	Causal inference in pharmacoepidemiology in general: 1. Exchangeability/ignorability – no confounding. 2. Non-interference – no spill-over effects (the outcome in one individual is influenced by their own treatment and not the treatment of others). 3. Consistency – exposure definition is specific enough that different levels will not have different effects on the outcome. 4. Positivity – being treated or untreated is possible across all combinations of covariates. Instrumental variable analysis: 1. Relevance assumption – the instrument strongly predicts the treatment. 2. Exogeneity assumption – the instrument is not associated with unmeasured confounders. 3. Exclusion restriction – the instrument acts on the outcome only via the treatment. 4. Monotonicity – changes in the instrument yield changes in the treatment receipt across all levels of the instrument. Traditional multivariable regression and propensity scores with IPTW-RA: 1. No unmeasured confounding (due to measurement error of measured confounders and unmeasured confounders).
Main requirements	Instrumental variable analysis (using a preference-based instrument): • There is an area of ambiguity in clinical guidelines which leads to genuine clinical uncertainty in how best to treat people with a particular indication. • This clinical uncertainty leads to substantial variation in treatment which can be used as a preference-based instrument in an IV analysis (e.g., group of general practitioners (GP), hospitals). • The data exist to define the preference-based instrument for each individual in the study, the treatment, the outcome, and measured confounders. • The study population is sufficiently large to adequately power the analyses. Traditional multivariable regression and propensity scores with IPTW-RA: • There is an area of ambiguity in clinical guidelines which leads to genuine clinical uncertainty in how best to treat people with a particular indication. • This clinical uncertainty leads to substantial variation in treatment. • The data exist to define the treatment, the outcome, and measured confounders. • The study population is sufficiently large to adequately power the analyses.

Appendix B.2. Table originally presented by Swanson et al 2015[1] to illustrate compliance types which must be considered when using a preference-based instrumental variable

		Seen by a prescriber who prefers treatment A	
		Prescribed treatment A	**Prescribed treatment B**
Seen by a prescriber who prefers treatment B	**Prescribed treatment A**	Always taker	Defier
	Prescribed treatment B	Complier	Never taker

APPENDIX C: Chapter 3 – Key supplementary materials from the published research paper

Appendix C.1. Adjusted predicted percentages of second-line treatment prescribed by ethnicity and IMD (Supplementary table 5 in the published research paper)

Second-line antidiabetic prescribed	Ethnicity or IMD quintile	[1]Adjusted predicted probability (95% CI)	P-value (Wald test)	P-value (joint-test)
Ethnicity				
SU	White	0.36 (0.34-0.37)	-	
	South Asian	0.37 (0.34-0.39)	0.37	
	Black	0.37 (0.34-0.40)	0.25	
	Mixed/other	0.36 (0.32-0.40)	0.96	0.61
DPP4i	White	0.43 (0.40-0.45)	-	
	South Asian	0.44 (0.41-0.47)	0.47	
	Black	0.44 (0.40-0.48)	0.40	
	Mixed/other	0.47 (0.42-0.51)	0.05	0.21
SGLT2i	White	0.21 (0.19-0.23)	-	
	South Asian	0.20 (0.18-0.22)	0.04	
	Black	0.19 (0.16-0.22)	0.02	
	Mixed/other	0.17 (0.14-0.21)	0.01	0.003
IMD quintile				
SU	1 (least deprived)	0.36 (0.34-0.38)	-	
	2	0.35 (0.33-0.37)	0.15	
	3	0.36 (0.34-0.38)	0.78	
	4	0.36 (0.34-0.38)	0.87	
	5 (most deprived)	0.36 (0.35-0.38)	0.53	0.26
DPP4i	1 (least deprived)	0.42 (0.39-0.45)	-	
	2	0.43 (0.40-0.46)	0.07	
	3	0.42 (0.40-0.45)	0.33	
	4	0.43 (0.40-0.46)	0.14	
	5 (most deprived)	0.44 (0.42-0.47)	0.003	0.04

Second-line antidiabetic prescribed	Ethnicity or IMD quintile	[1]Adjusted predicted probability (95% CI)	P-value (Wald test)	P-value (joint-test)
SGLT2i	1 (least deprived)	0.22 (0.20-0.25)	-	
	2	0.22 (0.20-0.24)	0.65	
	3	0.22 (0.20-0.24)	0.40	
	4	0.21 (0.19-0.23)	0.05	
	5 (most deprived)	0.19 (0.17-0.21)	<0.001	<0.001

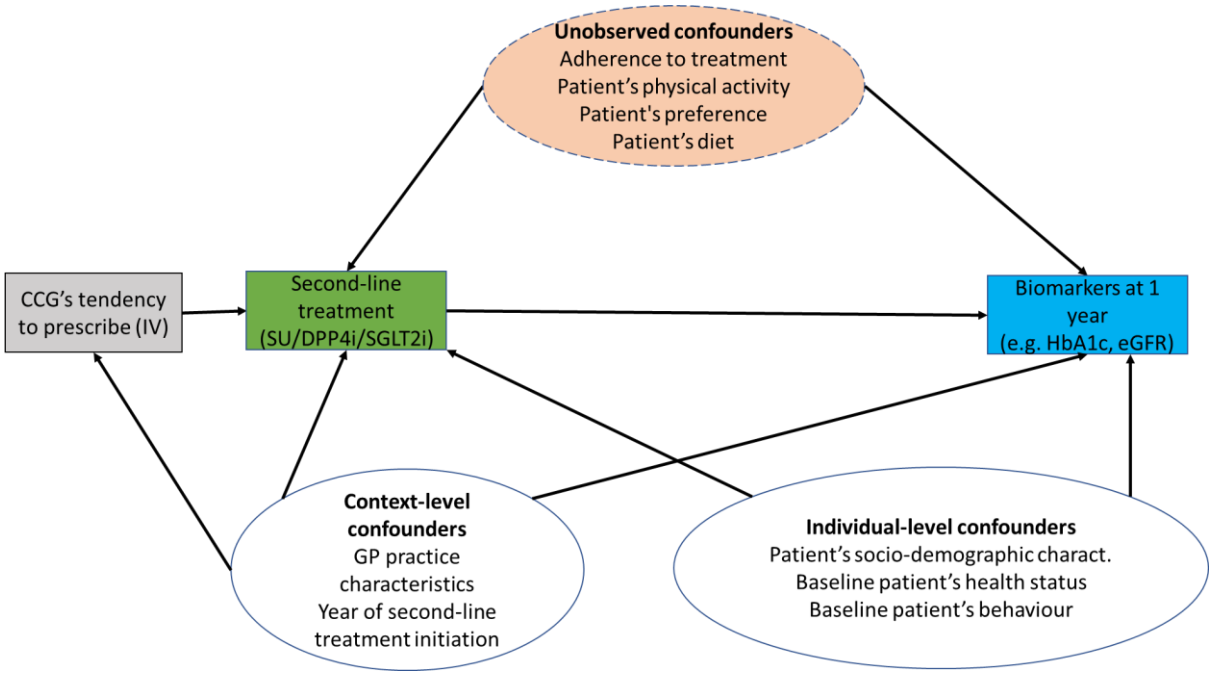
[1]Mutually adjusted for deprivation (ethnicity estimates) and ethnicity (deprivation estimates), as well as number of patients registered at the patients' GP practice, years on first-line category, age category, sex, last HbA1c prior to second-line initiation category, BMI, prevalent heart failure, ischaemic heart disease, myocardial infarction, stroke, unstable angina, RASi and/or statin co-prescription, CKD category, blood pressure category, history of proteinuria, blindness, cancer (any), hospitalisation (any) in past year, smoking status, alcohol status, region, all as fixed effects, and CCG-clustering as a random effect.

APPENDIX D: Chapter 4 – Key supplementary materials from the published research paper

The PERMIT study statistical analysis plan (SAP), which includes table 4.1. describing the amendments made to the published protocol, can be found at this link: https://www.lshtm.ac.uk/media/72276.

APPENDIX E: Chapter 5 – Key supplementary materials from the research paper *in press*

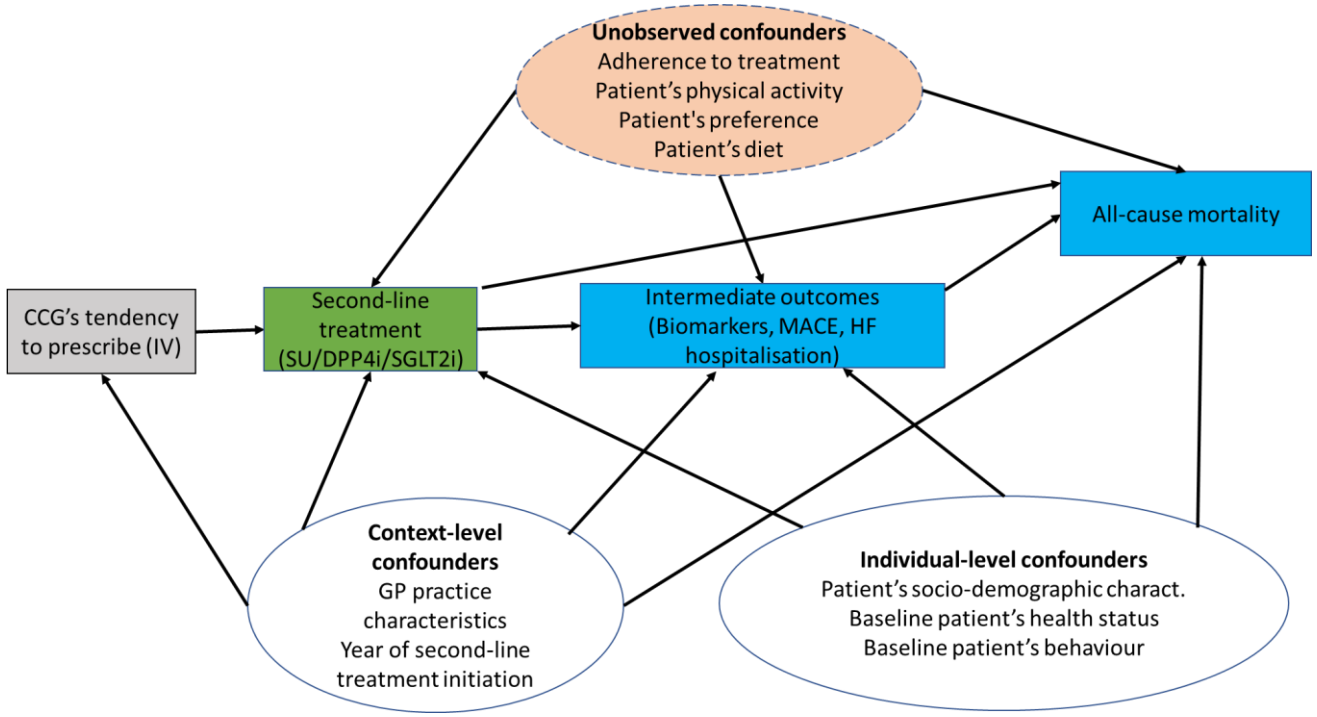
Appendix E.1. Directed acyclic graph (DAG) illustrating the causal relationship between the instrument, exposure, and primary outcome (change in HbA1c from baseline to 1-year follow-up) (Supplementary figure 1A in the research paper)



Commentary: This directed acyclic graph illustrates that the receipt of second-line treatment is subject to unmeasured and (context- and individual-level) observed factors that confounds the link between treatment and the outcome of interest (biomarkers at 1 year). This figure suggests that the CCG's tendency to prescribe (the proposed instrumental variable) predicts the second-line treatment received by a patient registered in that CCG, but does not have a direct effect on the health outcome of interest. That is, it is assumed that the only path through which the CCGs tendency to prescribe influences the biomarkers at 1 year is through its influence in the treatment received. Thus, this reflects the explicit assumption of an IV design that the instrument is not independently associated with outcomes, unobserved confounders and individual-level confounders. The DAG allows for an association between context-level confounders (such as GP practice list size) and the IV as larger practices may have different prescription patterns compared to smaller practices. The

individual-level confounders considered in this hypothesised causal diagram were classified in three broad categories: patient's socio-demographic characteristics (age, sex, etc.), baseline health status (e.g. relevant comorbidities, biomarkers and medications such as statins and renins), and baseline behaviour (alcohol and smoking status). For simplicity, this figure does not reflect all the existing correlations amongst different factors, for example the one existing between unobserved and individual-level confounders.

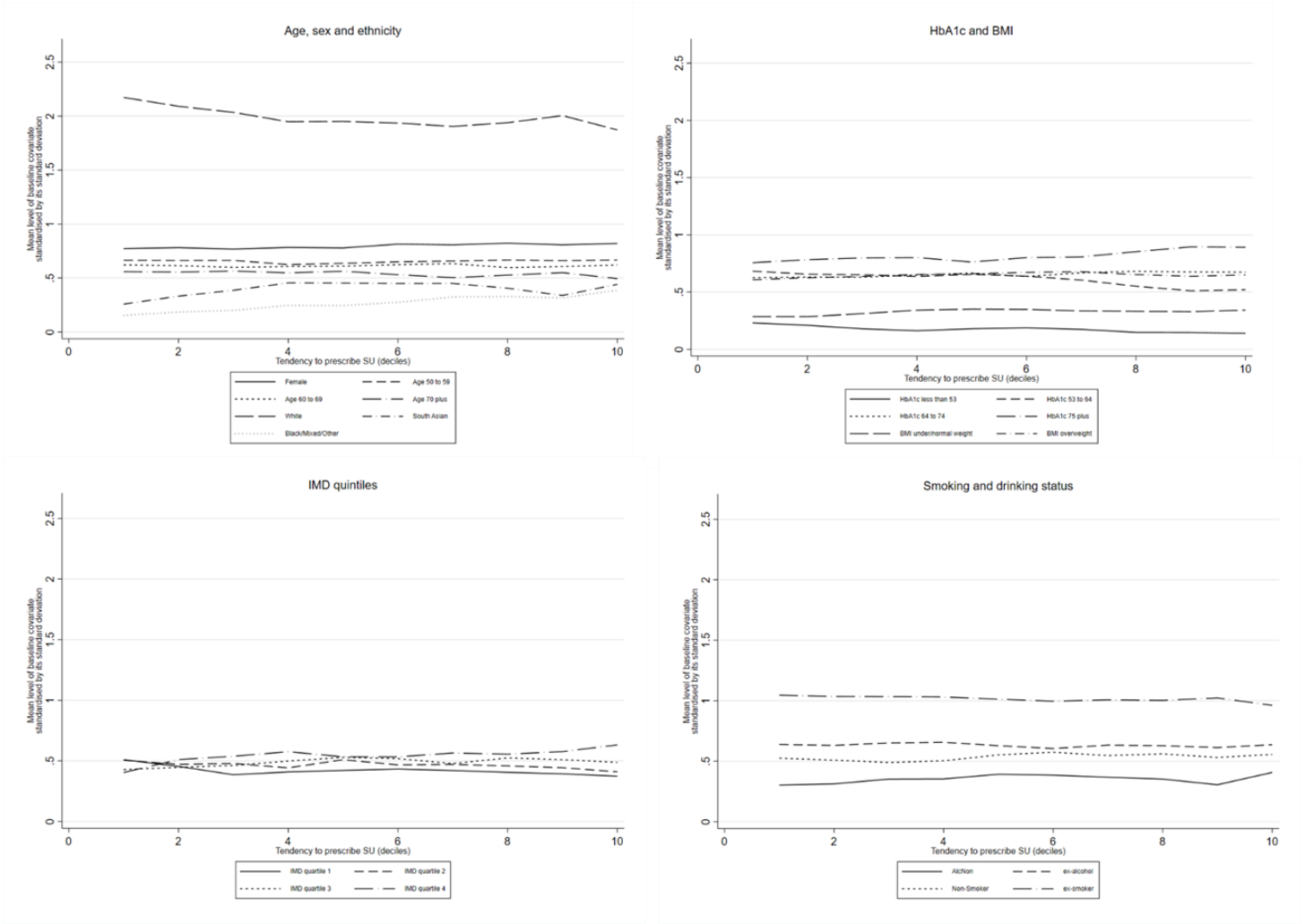
Appendix E.2. Directed acyclic graph (DAG) illustrating the causal relationship between the instrument, exposure, and all-cause mortality (secondary outcome) (Supplementary figure 1B in the research paper)

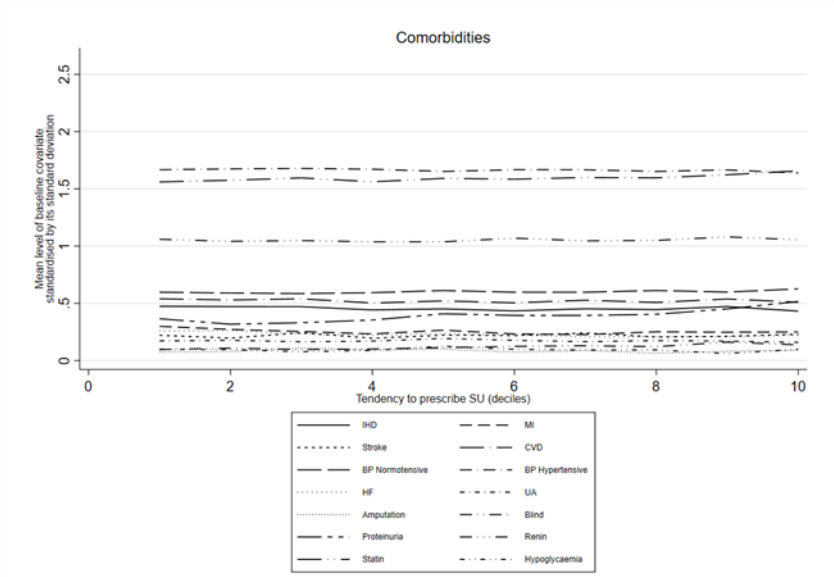


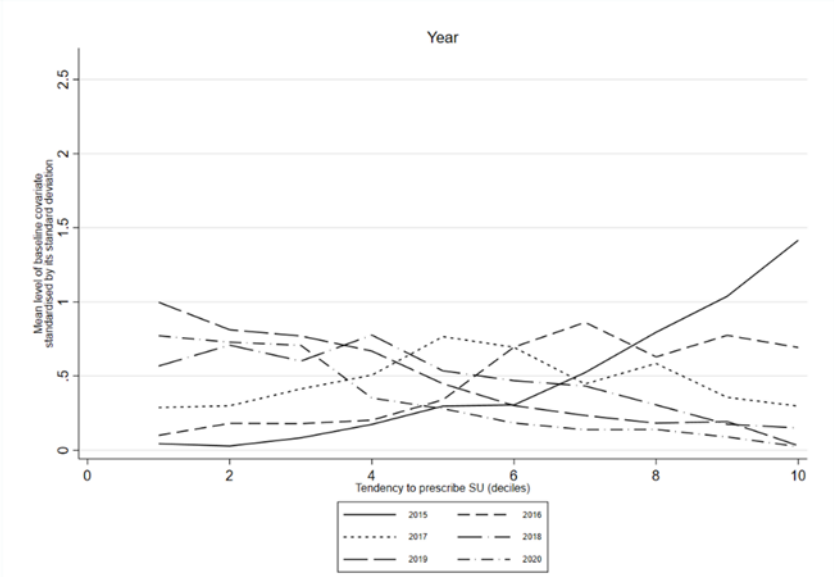
Commentary:

This DAG builds on the same structure as the previous one, but considers biomarkers at one year from intensification along with other adverse events such as MACE and hospitalisations due to heart failure as intermediate outcomes on the pathway from treatment to all-cause mortality. It can be seen that all-cause mortality, the long-term outcome of interest, is also subject to unobserved and measured confounders, but that the only path through which the IV influences both intermediate and long-term outcomes is through its influence in the treatment received.

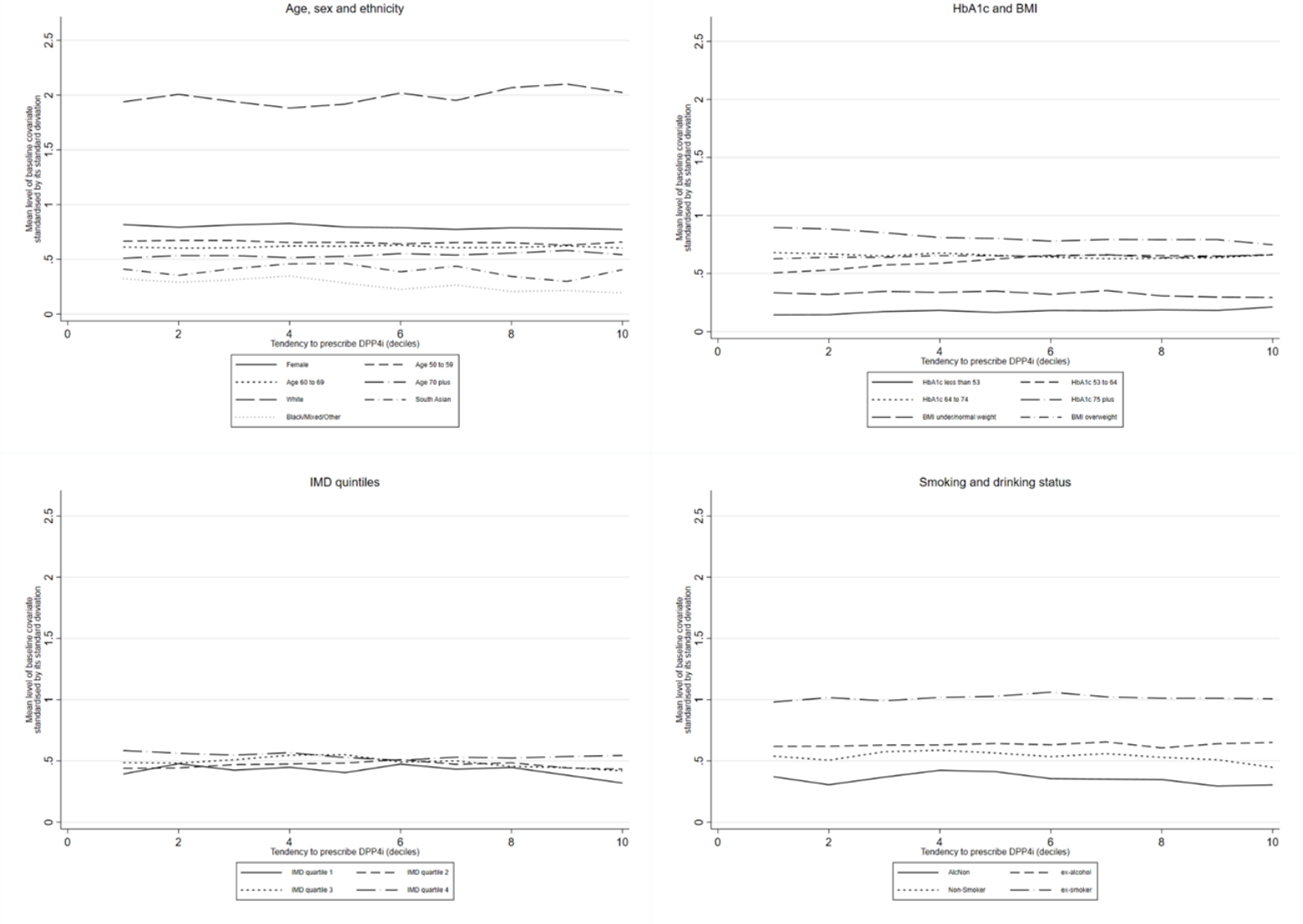
Appendix E.3. Covariate balance plots according to levels of the instrumental variable for SU (Supplementary figure 2A in the research paper)

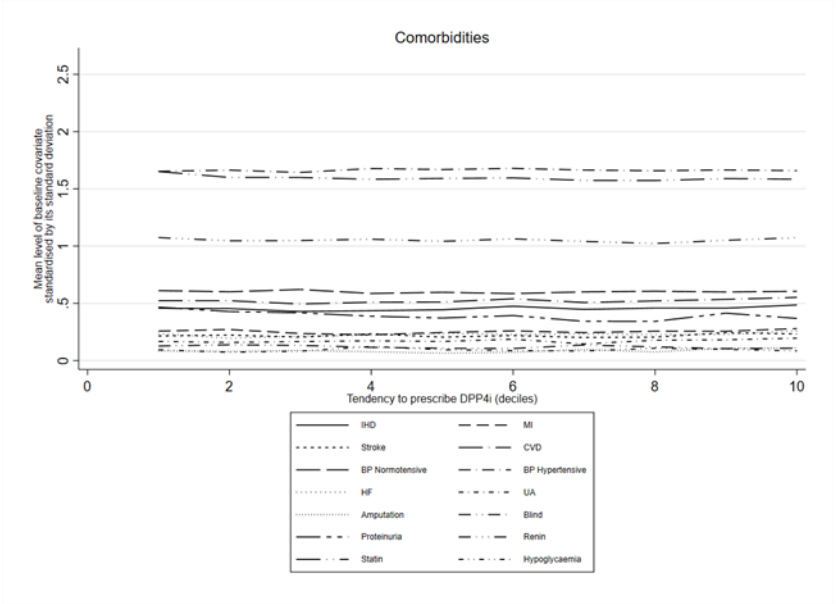


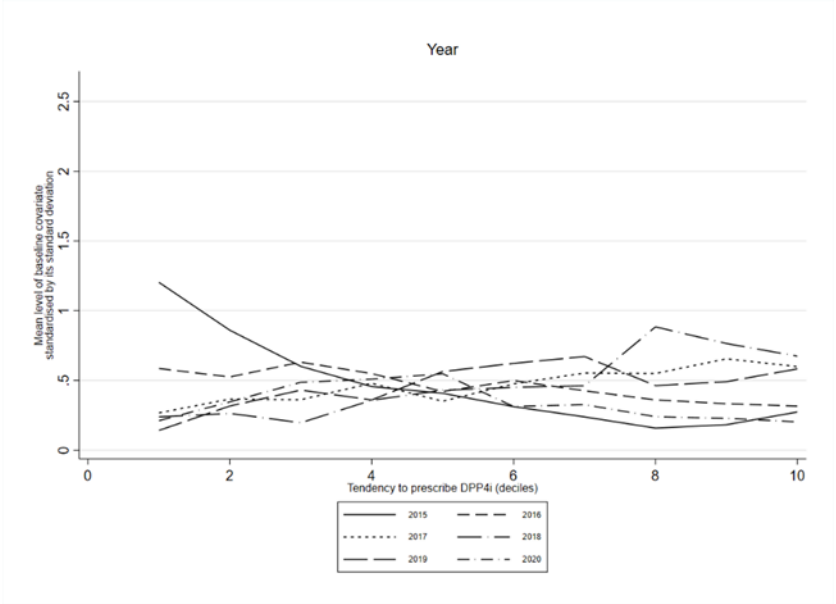

Comorbidities
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe SU (deciles)
IHD
MI
Stroke
CVD
BP Normotensive
BP Hypertensive
HF
UA
Amputation
Blind
Proteinuria
Renin
Statin
Hypoglycaemia


Year
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe SU (deciles)
2015
2016
2017
2018
2019
2020

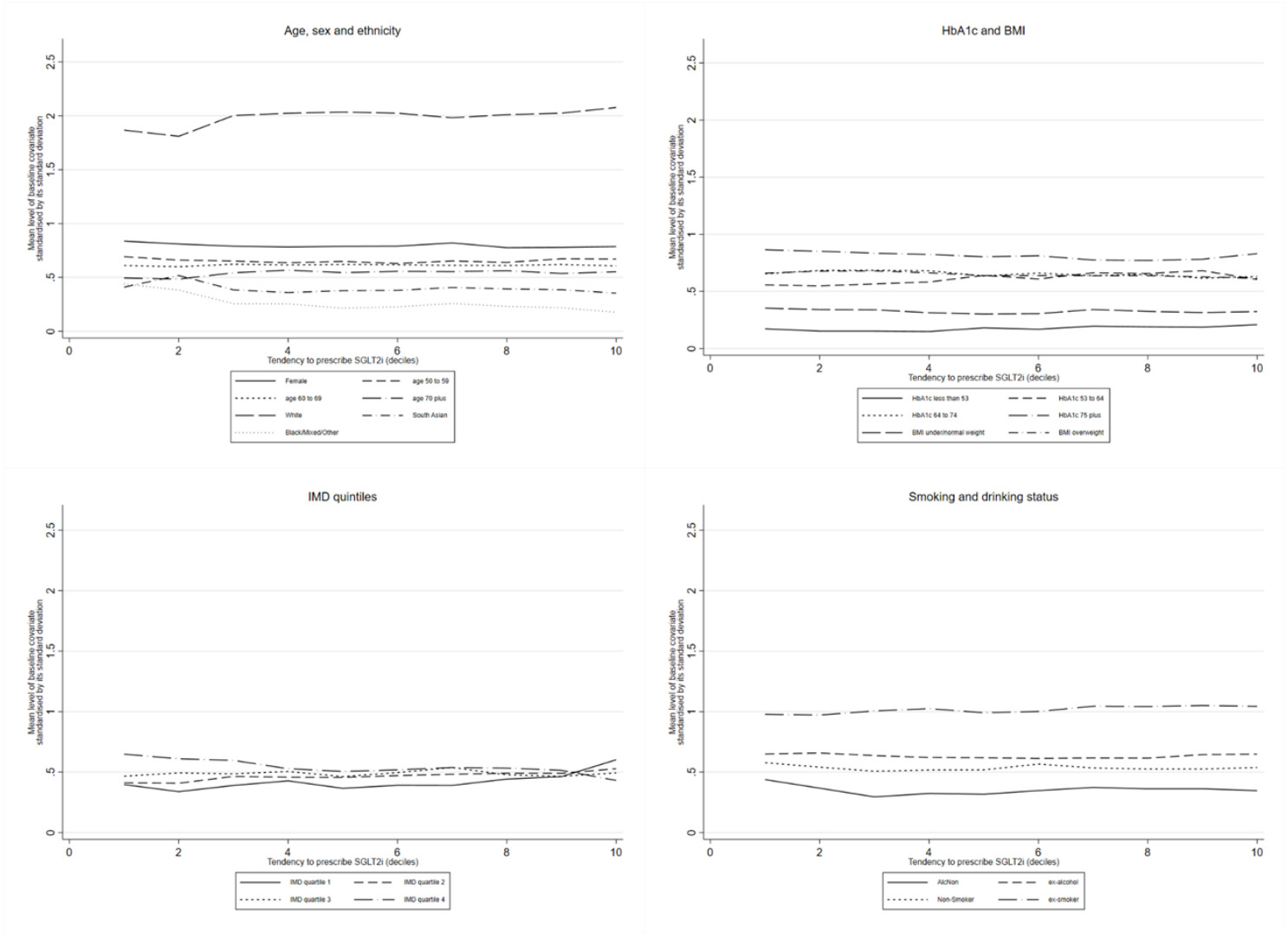
Appendix E.4. Covariate balance plots according to levels of the instrumental variable for DPP4i (Supplementary figure 2B in the research paper)

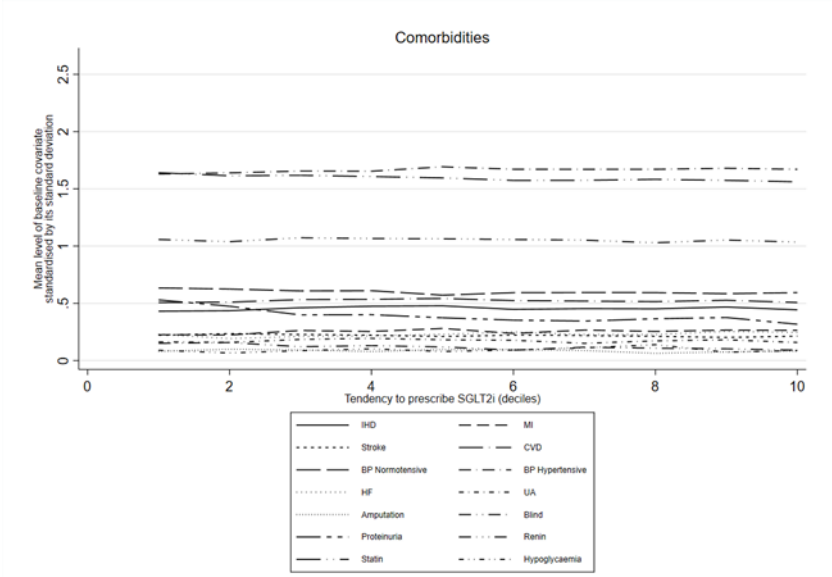


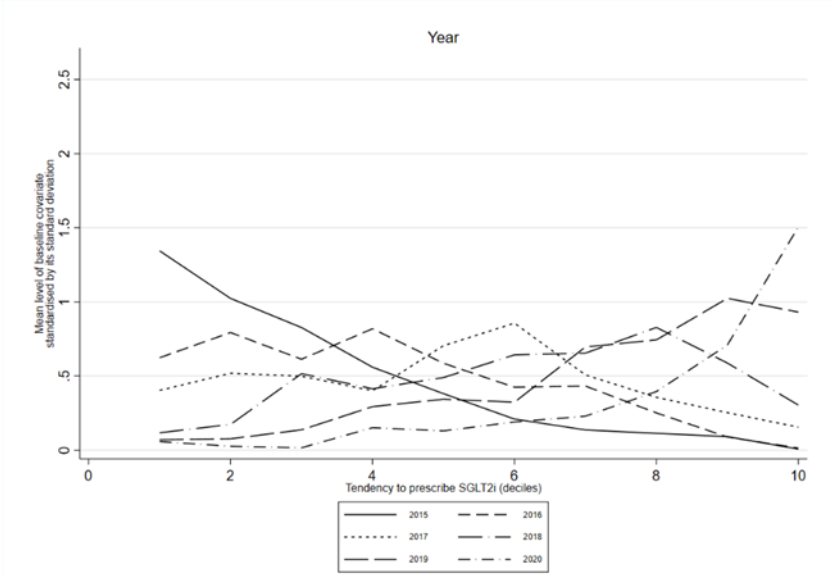

Comorbidities
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe DPP4i (deciles)
IHD
MI
Stroke
CVD
BP Normotensive
BP Hypertensive
HF
UA
Amputation
Blind
Proteinuria
Renin
Statin
Hypoglycaemia


Year
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe DPP4i (deciles)
2015
2016
2017
2018
2019
2020

Appendix E.5. Covariate balance plots according to levels of the instrumental variable for SGLT2i (Supplementary figure 2C in the research)




Comorbidities
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe SGLT2i (deciles)
IHD
MI
Stroke
CVD
BP Normotensive
BP Hypertensive
HF
UA
Amputation
Blind
Proteinuria
Renin
Statin
Hypoglycaemia


Year
Mean level of baseline covariate standardised by its standard deviation
0
.5
1
1.5
2
2.5
0
2
4
6
8
10
Tendency to prescribe SGLT2i (deciles)
2015
2016
2017
2018
2019
2020

Appendix E.6. Supplementary methods

INSTRUMENTAL VARIABLE ANALYSIS

We followed a two-stage residual inclusion (2SRI) approach in the main analysis to reduce the risk of bias from unmeasured confounding and allow us to identify the average treatment effects (ATEs). We compared these estimates to those from using 2-stage least squares (2SLS) (alternative analysis).

2-stage least squares (2SLS):

We estimated a linear probability model for each of the treatments of interest (DPP4i or SGLT2i using SU as the reference category) on the instruments (TTP for either treatment), and the covariates. In the second stage, we regressed the outcome of interest on the covariates and the predicted probability capturing the propensity, for each treatment obtained from the first stage. The two stages were estimated jointly so that standard errors reflected the uncertainty of both stages. However, when effects are heterogeneous, that is they vary with respect to observed or unobserved covariates, this approach estimates the *local* average treatment effects (LATE) since they relate to the compliers whose treatment assignment is altered by the instrument (or in the case of a continuous IV, a weighted average of LATEs), which is less relevant for decision making compared with the overall average treatment effect. We therefore consider 2SRI to be the more informative approach here since estimates relate to the full population provided the first stage model is correctly specified.

2-stage residual inclusion:

First, for DPP4i and SGLT2i, we estimated first-stage probit models for whether or not the patient was prescribed this treatment, as a function of the covariates and the tendency to prescribe that treatment. For continuous outcomes (e.g., HbA1c at 12 months), we then estimated a second stage regression model, using ordinary least squares including the generalised residuals[2] from the first stage models, all measured baseline covariates, NHS region and time period. For the censored time-to-event outcomes (e.g., time to 3-point MACE), in the second stage we estimated Cox proportional hazards models that account for individual frailty,[3] in addition to covariates and the observed residuals from the first stage models.[4]

Variable selection:

In addition to the covariates listed, the 2SRI models also considered the quadratic forms of age and baseline HbA1c as well as two sets of interactions. The first set of interactions are those between baseline HbA1c with age, sex and baseline BMI. The second set of interactions are the products of the IV (for the first stage models) or the treatment indicator variables (for the second stage models) with baseline HbA1c, eGFR, BMI, systolic blood pressure and age.

To prevent overfitting the models, we used the Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm[5, 6] to inform which of the interactions above are relevant in each case. The LASSO aims to find the set of coefficients that minimise the sum-of-squares loss function subject to a constraint on the sum of absolute values of coefficients. This results in a linear regression in which only a small number of covariates have non-zero coefficients that can then be included in the model in question. In particular, we used the 'rigorous LASSO' approach[7] which places a high priority on controlling overfitting, thus often producing parsimonious models.

By partialling out variables prior to penalisation, we ensure these variables are always included in the selected models and only penalised (and potentially discard) variables in the interaction sets. We use the 'rigorous LASSO' approach[7] which places a high priority on controlling overfitting, thus often producing parsimonious models.[8] By partialling out variables prior to penalization, we ensure these variables are always included in the selected models and only penalise (and potentially discard) variables in the interaction sets.[8]

To select the variables included in the estimation of the 2SRI models, we ran the rigorous LASSO for the first and second stage models for each outcome. The final set of covariates used to estimate effects for each outcome included all the covariates mentioned in the *Covariates* section plus the interactions that were selected in at least one model of the respective outcome. For the Cox proportional hazards models, all final selected covariates were then assessed for violation of the proportional hazards assumption using Schoenfeld residuals.

Estimating treatment effects:

From the second stage models using the selected variables, we calculated the difference in the average absolute change in predicted outcomes (continuous measures) and times to event between the comparison groups, providing estimates of the treatment effect according to individual-level covariates. We aggregate these estimates to report results overall and according to whether or not patients had pre-existing CVD (at least one of previous MI, previous stroke, CHF, IHD, or unstable angina).

All standard errors were calculated with non-parametric bootstrapping as described below, and accounted for clustering of individuals within NHS region.

HANDLING OF MISSING DATA, CENSORING, AND LOSS TO FOLLOW-UP

The PERMIT study uses routine linked data (CPRD and HES) which raises several challenges for the statistical analysis. Missing data may occur due to:

- Non-attendance at a GP within the requisite time period for the study outcome definition (+/- 3 months either side of timepoints 0.5, 1, 2, 3, 4 and 5 years)
- Information not being recorded during a GP visit
- Tests not being done during a GP visit

In addition to missing data, all patients are not fully followed from baseline to five years. For example, a patient enrolled in December 2020 will have 12 months follow-up to 2021 and can only be included in the analysis models for the continuous outcomes for the periods between baseline and 6 months or 1 year, as they are unobserved for subsequent timepoints. The final challenge is related to 'loss to follow-up' or 'dropouts', as a patient may stop attending GP appointments before death or censoring (end of follow-up or patient/GP stops contributing to CPRD) has occurred.

Supplementary methods table 1 presents the full list of covariates which are adjusted for in both sets of analyses, summarises the seven survival outcomes and the four continuous clinical outcomes measures at timepoints 0.5, 1, 2, 3, 4 and 5 years.

Supplementary methods table S1: Summary of (i) analysis variables adjusted for in the continuous and survival analyses, (ii) survival information included in the imputation models for the analysis variables, and (iii) continuous outcome information included in the imputation models for the analysis variables

Analysis Covariates		
Baseline age	Sex	Ethnicity
Index of multiple deprivation	Days since 2nd line treatment assignment	Practice size in 2014
Renin	Statin	Myocardial Infarction
Unstable Angina	Stroke	Hypoglycaemia
Heart Failure	Cancer history	Proteinuria history
Advanced eye disease	Lower extremity amputation	Lower extremity amputation
CKD	Baseline HbA1c	Baseline systolic blood pressure
Baseline diastolic blood pressure	Baseline eGFR	Baseline BMI
Smoking status	Alcohol status	Year of first 2 line initiation

Practice region	IHD	Hospital attendance in last year
Age-squared	HbA1c-squared	HbA1c*Baseline Age
HbA1c*BMI	HbA1c*Sex	
Nelson-Aalen Estimates & Event indicator information		
MACE	MI	Stroke
Heart Failure hospitalisation	Death	End stage kidney disease (ESKD)
eGFR decline from 40%	Composite kidney	
Continuous clinical measures for time t= 0.5, 1, 2, 3, 4 and 5 years		
Change in HbA1c at time *t*	Change in eGFR at time *t*	Change in BMI at time *t*
Change in SBP at time *t*		

Missing data is present in in both the continuous outcomes, and also several covariates which are used in the analysis of the continuous and survival outcomes. The percentage of missing values in the analysis covariates are available in Table 1 of the main paper. **Supplementary methods table 2** presents the percentage missing for those not censored at time *t* for each clinical measure. The longitudinal clinical measures (HbA1c, BMI, SBP, eGFR) could be unavailable at any timepoint (6 months, 1, 2, 3 or 5 years). For example, a patient could have observed HbA1c values at all time points except for year 2.

Supplementary methods table S2. Percentage of observations which are missing at time point t after accounting for censoring (Supplementary methods table S2 in the research paper)

Time *t*	**N**	**Percentage (%) missing in continuous measures at timepoint *t***			
		HbA1c	**BMI**	**SBP**	**eGFR**
6 months	72,066	32.1	50.0	39.0	42.2
1 year	66,702	33.7	44.7	33.6	37.4
2 years	52,962	36.4	47.8	37.2	40.0
3 years	39,099	38.6	50.0	39.6	42.1
4 years	26,366	40.6	52.9	43.5	43.7
5 years	15,651	46.9	59.4	51.0	48.1

Multiple imputation

We used Multiple imputation by chained equations[9] to handle both missing values in analysis covariates and missingness in the continuous outcomes,[10, 11] which will generate five imputed datasets.

Predictive mean matching with 10 donors[12] was used to impute all categorical and continuous partially-observed variables to improve robustness to misspecification of the imputation model. We assumed that data were `missing at random' (MAR). For missing values in the continuous outcome measures, this assumption implies that this missingness is at random, i.e. at random conditional on all other measures in the model including all preceding and subsequent levels of the measure in question, and the levels of any measures that were available at the timepoint in question. Some measurements taken repeatedly over time, e.g., HbA1c and BMI, were missing at baseline for some individuals, and the same rationale for supporting the underlying MAR assumption would apply here as for outcomes with intermittent missingness, given that measurements for time periods prior to baseline and during the subsequent follow-up periods were available for the imputation models.

For the analysis covariates, ethnicity had the greatest proportion of missing values. Previous literature has shown that conducting a MAR analysis for ethnicity can lead to similar point estimates as implementing missing data methods under the `missing not at random' assumption.[13, 14] Here, our base case analysis used multiple imputation for ethnicity, along with the other covariates, and we examined robustness to the assumed missing data mechanism by undertaking complete-case analysis in a sensitivity analysis.

Imputation model specification

Due to the non-linear trajectory of the continuous outcomes, all continuous measures from 6 months to 5 years were used when imputing a continuous outcome at time *t.* For example, a patient's observed HbA1c values from baseline, 6 months and 2-5 years would be used to impute their unobserved year 1 HbA1c value, in addition to any auxiliary information which would improve the imputed value. The imputation models for the analysis covariates included information on both the survival and continuous outcomes to ensure congeniality[15] between the covariates and each continuous and survival outcome. The imputation models for all partially-observed covariates are specified in **Supplementary methods table 3**. Interactions included in the analysis models were treated as `just another variable'[16] and imputed using MICE with PMM.

Appendix E.11. Differences in the change in continuous clinical measures for the three second-line antidiabetic treatment comparisons (inverse probability of treatment weighting - regression adjustment (IPTW-RA) analysis, complete cases) (Supplementary table 19 in the research paper)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in HbA1c (mmol/mol) from baseline		**N =**	46,900	42,441	32,364	23,082	15,126	8,022
	DPP4i vs SU	**Estimate**	3.29	1.31	0.09	-0.29	-0.75	-0.56
		(95% CI)	(2.97, 3.62)	(0.94, 1.68)	(-0.35, 0.52)	(-0.79, 0.20)	(-1.36, -0.14)	(-1.42, 0.29)
	SGLT2i vs SU	**Estimate**	1.73	-1.23	-1.30	-1.84	-1.13	-1.00
		(95% CI)	(1.30, 2.17)	(-1.73, -0.73	(-2.00, -0.60)	(-2.65, -1.03)	(-2.23, -0.03)	(-2.41, 0.41)
	SGLT2i vs DPP4i	**Estimate**	-1.56	-2.54	-1.38	-1.55	-0.37	-0.44
		(95% CI)	(-1.99, -1.13)	(-3.02, -2.05)	(-2.06, -0.71)	(-2.34, -0.76)	(-1.47, 0.72)	(-1.87, 0.99)
Difference in the change in BMI (kg/m^2) from baseline		**N =**	33,508	34,431	25,809	18,253	11,577	5,903
	DPP4i vs SU	**Estimate**	-0.53	-0.59	-0.46	-0.52	-0.49	-0.35
		(95% CI)	(-0.60, -0.46)	(-0.67, -0.52)	(-0.54, -0.38)	(-0.62, -0.41)	(-0.62, -0.37)	(-0.55, -0.14)
	SGLT2i vs SU	**Estimate**	-1.23	-1.30	-1.08	-0.95	-0.81	-0.70
		(95% CI)	(-1.31, -1.15)	(-1.39, -1.22)	(-1.18, -0.97)	(-1.09, -0.81)	(-1.00, -0.63)	(-0.97, -0.43)
	SGLT2i vs DPP4i	**Estimate**	-0.70	-0.71	-0.62	-0.43	-0.32	-0.35
		(95% CI)	(-0.76, -0.64)	(-0.78, -0.64)	(-0.71, -0.52)	(-0.56, -0.31)	(-0.50, -0.14)	(-0.60, -0.10)
Difference in the change in eGFR (mL/min/1.73m^2) from baseline		**N =**	39,113	39,337	30,034	21,398	14,060	7,659
	DPP4i vs SU	**Estimate**	-0.17	-0.16	0.07	-0.09	0.54	0.42
		(95% CI)	(-0.39, 0.04)	(-0.38, 0.05)	(-0.19, 0.34)	(-0.43, 0.25)	(0.12, 0.96)	(-0.18, 1.03)
	SGLT2i vs SU	**Estimate**	0.10	0.12	1.21	1.42	1.59	2.79
		(95% CI)	(-0.24, 0.43)	(-0.23, 0.47)	(0.75, 1.67)	(0.80, 2.04)	(0.70, 2.49)	(1.72, 3.85)
	SGLT2i vs DPP4i	**Estimate**	0.27	0.28	1.14	1.51	1.06	2.36
		(95% CI)	(-0.04, 0.58)	(-0.05, 0.61)	(0.70, 1.57)	(0.90, 2.12)	(0.16, 1.95)	(1.29, 3.44)

Supplementary methods table S3. Fitted imputation models for each partially-observed variable used in the continuous and survival analyses. Analysis covariates, Survival information and Continuous outcomes are specified in Supplementary methods table 1. (Supplementary methods table S3 from the research paper)

Partially-observed	**Covariates adjusted for in each imputation model***
Covariates	
Ethnicity	Analysis Covariates All Nelson-Aalen Estimates & Event indicators All continuous outcomes for t=0.5, 1, …, 5yrs
Index of multiple deprivation	
Baseline Hba1c	
Baseline eGFR	
BMI	
Smoking status	
Alcohol status	
Systolic blood pressure	
Diastolic blood pressure	
Baseline Hba1c squared	
Baseline Hba1c * Age at baseline	
Baseline Hba1c * BMI	
Baseline Hba1c * Sex	
Continuous clinical outcomes	
Hba1c at time *t*	Analysis Covariates All Continuous outcomes for t=0.5, 1, …, 5yrs
eGFR at time *t*	
BMI at time *t*	
Systolic blood pressure at time *t*	
*For patients who died, their imputation models also included time to death from baseline.	

Imputation model stratification and follow-up status

The imputation models were stratified by (i) treatment assignment (SGLT2i, DPP4i, SU), and (ii) status across the follow-up period ((i) a patient has died during follow-up; (ii) a patient is fully followed for 5 years, and (iii) censored as a patient no longer contributes to the study due to either reaching the end of

the study monitoring period, or the patient/practice no longer contributing to CPRD). For patients who are censored, we assume that this is `completely at random' as censoring pertains to administrative reasons or due to the end of the follow-up period, which are unlikely to be related to the patient's characteristics of interest in our analyses, such as their prognosis. For people who have died, their corresponding missing values prior to death are likely to differ to those who were alive at a given follow-up timepoint, and hence missing values for these people were imputed separately from those with full follow-up, or who stopped contributing before full follow-up was reached.

In total, 9 imputation models were used to impute each partially-observed variable depending on which treatment and follow-up status strata they belonged to. The three imputation models for patients who died (Death-SU; Death-SGLT2i; Death-DPP4i) also included a "time to death from baseline", in addition to the variables specified in **Supplementary methods table 3**, to recognise that this may be predictive of the missing outcome.

Due to the specification of the imputation models (**Supplementary methods table 3**) and the MICE[17] package in R, it was not possible to restrict imputing missing continuous outcome values up to the point of death or the point of no longer contributing to the study. Instead, missing values in the continuous clinical measures were imputed for all timepoints t=0.5, 1,…, 5 years. Censoring rules were then applied post-imputation before running the statistical analyses.

Post-imputation estimation of treatment effects and confidence intervals

The relative treatment effects were estimated in each imputed dataset using two-stage residual inclusion IV (with a frailty inclusion for time-to-event outcomes when using Cox proportional hazards).[4] Rubin's rules[18] was applied to obtain an overall treatment effect:

$$\hat{\bar{\theta}}_d = M^{-1} \sum_{m=1}^{M} \hat{\theta}_{m,d}$$

where *d* = (i) SGLT2i vs. SU, (ii) SGLT2i vs. DPP4i or (iii) DPP4i vs. SU. For Cox proportional hazards, Rubin's rules were applied on the log-hazard scale. The analysis model of interest was applied to each of the five multiply imputed datasets (M=5). This number of imputations was chosen as the overall analytical framework (IV residual inclusion) required that standard errors were estimated with the non-parametric bootstrap i.e. each of the nine imputation models were applied within each of the 500 bootstrap replications. The choice of M=5[19, 20] was a balance between recognising the importance of the number of imputed datasets for improved inference and the impact on computational time when running MI, non-parametric bootstrapping, IV residual inclusion and a Cox proportional hazards model with a frailty inclusion term.

Confidence intervals for the treatment effects were estimated using bootstrap sampling (BS), stratifying by region, treatment group, death and censoring status to maintain similar sampling patterns within each bootstrap sample. The original unimputed data were bootstrapped 500 times, and within each bootstrap sample, MI was applied (BS-then-MI).[21, 22]

Within each bootstrap sample $b = 1, \ldots, 500$, we took the same approach to handling missing data and implementing the analysis model, as previously specified. Rubin's rules were applied to the *M* imputed datasets of bootstrap sample *b* to get an overall treatment effect for each drug comparison *d*:

$$\hat{\bar{\theta}}_{b,d} = M^{-1} \sum_{m=1}^{M} \hat{\theta}_{m,d}$$

The 500 estimates of $\hat{\bar{\theta}}_{b,d}$ was used to estimate variance and calculate *t*-based confidence intervals.

Appendix E.7. Differences in the change in continuous clinical measures for the three second-line antidiabetic treatment comparisons for the main analysis (2SRI, bootstrap-multiple imputation) (Supplementary table 8 in the research paper)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in HbA1c (mmol/mol) from baseline		**N =**	72,066	66,702	52,962	39,099	26,366	15,651
	DPP4i vs SU	**Estimate**	2.89	0.68	-1.40	-1.84	-2.33	-2.78
		(95% CI)	(1.99, 3.80)	(-0.31, 1.68)	(-2.55, -0.24)	(-3.21, -0.47)	(-3.94, -0.72)	(-5.10, -0.45)
	SGLT2i vs SU	**Estimate**	2.11	-2.51	-4.95	-6.50	-5.35	-1.77
		(95% CI)	(1.09, 3.12)	(-3.72, -1.30)	(-6.46, -3.45)	(-8.47, -4.52)	(-7.98, -2.73)	(-6.89, 3.35)
	SGLT2i vs DPP4i	**Estimate**	-0.79	-3.20	-3.56	-4.66	-3.02	1.01
		(95% CI)	(-1.93, 0.35)	(-4.58, -1.81)	(-5.28, -1.84)	(-6.90, -2.41)	(-5.95, -0.09)	(-4.75, 6.76)
Difference in the change in BMI (kg/m^2) from baseline		**N =**	72,066	66,702	52,962	39,099	26,366	15,651
	DPP4i vs SU	**Estimate**	-0.52	-0.70	-0.64	-0.57	-0.79	-0.68
		(95% CI)	(-0.68, -0.37)	(-0.83, -0.56)	(-0.80, -0.48)	(-0.78, -0.35)	(-1.05, -0.53)	(-1.09, -0.28)
	SGLT2i vs SU	**Estimate**	-1.41	-1.55	-1.50	-1.55	-1.32	-1.83
		(95% CI)	(-1.57, -1.25)	(-1.72, -1.37)	(-1.74, -1.23)	(-1.92, -1.19)	(-1.81, -0.83)	(-2.80, -0.85)
	SGLT2i vs DPP4i	**Estimate**	-0.89	-0.85	-0.85	-0.99	-0.52	-1.14
		(95% CI)	(-1.06, -0.71)	(-1.03, -0.66)	(-1.12, -0.58)	(-1.39, -0.59)	(-1.07, 0.02)	(-2.20, -0.09)
Difference in the change in eGFR (mL/min/1.73m^2) from baseline		**N =**	72,066	66,702	52,962	39,099	26,366	15,651
	DPP4i vs SU	**Estimate**	-0.14	0.14	1.44	2.85	3.40	4.01
		(95% CI)	(-0.71, 0.44)	(-0.46, 0.73)	(0.69, 2.19)	(1.97, 3.74)	(2.26, 4.55)	(2.42, 5.61)
	SGLT2i vs SU	**Estimate**	-0.21	0.44	1.39	1.99	3.66	5.99
		(95% CI)	(-0.87, 0.46)	(-0.29, 1.18)	(0.49, 2.30)	(0.69, 3.28)	(1.97, 5.36)	(2.83, 9.15)
	SGLT2i vs DPP4i	**Estimate**	-0.07	0.31	-0.04	-0.87	0.26	1.98
		(95% CI)	(-0.87, 0.69)	(-0.53, 1.14)	(-1.09, 1.01)	(-2.35, 0.61)	(-1.81, 2.33)	(-1.56, 5.51)
Difference in the change in SBP (mm Hg) from baseline		**N =**	72,066	66,702	52,962	39,099	26,366	15,651
	DPP4i vs SU	**Estimate**	-0.80	-0.31	-0.43	-0.94	-1.28	-0.55

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
		(95% CI)	(-1.70, 0.09)	(-1.29, 0.66)	(-1.47, 0.62)	(-2.13, 0.24)	(-2.70, 0.13)	(-2.76, 1.66)
	SGLT2i vs SU	**Estimate**	-2.57	-2.07	-2.97	-3.11	-0.96	-5.64
		(95% CI)	(-3.60, -1.54)	(-3.10, -1.04)	(-4.31, -1.62)	(-4.83, -1.40)	(-3.19, 1.26)	(-9.73, -1.56)
	SGLT2i vs DPP4i	**Estimate**	-1.77	-1.76	-2.54	-2.17	0.32	-5.09
		(95% CI)	(-2.91, -0.62)	(-2.99, -0.53)	(-4.05, -1.03)	(-4.24, -0.11)	(-2.27, 2.91)	(-9.82, -0.36)

Appendix E.8. Summary of results from main analysis for kidney, cardiovascular, and mortality time-to-event outcomes, as well as summary of results for alternative analyses for kidney, cardiovascular, and mortality outcomes (Supplementary table 10 in the research paper)

Outcome	Treatment comparison	Analysis method*	Hazard ratio	95% CI (lower)	95% CI (upper)
≥40% decline in eGFR	DPP4i vs SU	Base (2-years follow-up max)	0.66	0.37	1.17
		Complete case (CC)	0.78	0.39	1.58
		Multivariable regression – CC	0.98	0.85	1.12
		Base (5-years follow-up max)	0.65	0.42	0.99
	SGLT2i vs SU	Base (2-years follow-up max)	0.42	0.22	0.81
		Complete case (CC)	0.40	0.17	0.91
		Multivariable regression – CC	0.78	0.61	0.99
		Base (5-years follow-up max)	0.47	0.24	0.92
	SGLT2i vs DPP4i	Base (2-years follow-up max)	0.64	0.29	1.43
		Complete case (CC)	0.51	0.18	1.39
		Multivariable regression – CC	0.80	0.63	1.01
		Base (5-years follow-up max)	0.73	0.33	1.59
MAKE	DPP4i vs SU	Base (2-years follow-up max)	0.72	0.50	1.03
		Complete case (CC)	0.86	0.56	1.32
		Multivariable regression – CC	0.74	0.68	0.80
		Base (5-years follow-up max)**	-	-	-
	SGLT2i vs SU	Base (2-years follow-up max)	0.79	0.51	1.23
		Complete case (CC)	0.81	0.48	1.39
		Multivariable regression – CC	0.59	0.51	0.68
		Base (5-years follow-up max)**	-	-	-
	SGLT2i vs DPP4i	Base (2-years follow-up max)	1.11	0.66	1.84
		Complete case (CC)	0.94	0.50	1.79
		Multivariable regression – CC	0.80	0.69	0.92
		Base (5-years follow-up max)**	-	-	-
Heart failure hospitalisation	DPP4i vs SU	Base (2-years follow-up max)	1.41	0.73	2.71
		Complete case (CC)	1.26	0.63	0.63
		Multivariable regression – CC	0.74	0.63	0.87
		Base (5-years follow-up max)	1.25	0.74	2.11

Outcome	Treatment comparison	Analysis method*	Hazard ratio	95% CI (lower)	95% CI (upper)
	SGLT2i vs SU	Base (2-years follow-up max)	0.46	0.20	1.05
		Complete case (CC)	0.43	0.16	1.11
		Multivariable regression – CC	0.50	0.37	0.69
		Base (5-years follow-up max)	0.63	0.28	1.42
	SGLT2i vs DPP4i	Base (2-years follow-up max)	0.32	0.12	0.85
		Complete case (CC)	0.34	0.11	1.06
		Multivariable regression – CC	0.68	0.50	0.93
		Base (5-years follow-up max)	0.51	0.20	1.30
MACE	DPP4i vs SU	Base (2-years follow-up max)	1.09	0.70	1.69
		Complete case (CC)	1.02	0.64	1.63
		Multivariable regression – CC	0.82	0.74	0.92
		Base (5-years follow-up max)	0.90	0.63	1.27
	SGLT2i vs SU	Base (2-years follow-up max)	0.99	0.61	1.62
		Complete case (CC)	0.93	0.53	1.63
		Multivariable regression – CC	0.83	0.71	0.96
		Base (5-years follow-up max)	1.12	0.70	1.80
	SGLT2i vs DPP4i	Base (2-years follow-up max)	0.91	0.51	1.63
		Complete case (CC)	0.91	0.48	1.72
		Multivariable regression – CC	1.00	0.87	1.16
		Base (5-years follow-up max)	1.25	0.72	2.16
All-cause mortality	DPP4i vs SU	Base (2-years follow-up max)	0.82	0.51	1.32
		Complete case (CC)	0.93	0.55	1.58
		Multivariable regression – CC	0.63	0.56	0.70
		Base (5-years follow-up max)	0.90	0.63	1.29
	SGLT2i vs SU	Base (2-years follow-up max)	1.14	0.64	2.03
		Complete case (CC)	1.17	0.61	2.27
		Multivariable regression – CC	0.50	0.41	0.62
		Base (5-years follow-up max)	1.25	0.77	2.05
	SGLT2i vs DPP4i	Base (2-years follow-up max)	1.39	0.71	2.74
		Complete case (CC)	1.26	0.59	2.68

Outcome	Treatment comparison	Analysis method*	Hazard ratio	95% CI (lower)	95% CI (upper)
		Multivariable regression – CC	0.80	0.66	0.98
		Base (5-years follow-up max)	1.39	0.77	2.50
Myocardial infarction	DPP4i vs SU	Base (2-years follow-up max)	1.41	0.74	2.68
		Complete case (CC)	1.19	0.60	2.35
		Multivariable regression - CC	0.83	0.70	0.98
		Base (5-years follow-up max)	0.87	0.52	1.46
	SGLT2i vs SU	Base (2-years follow-up max)	1.35	0.67	2.71
		Complete case (CC)	1.18	0.53	2.61
		Multivariable regression - CC	0.92	0.75	1.13
		Base (5-years follow-up max)	1.82	0.92	3.59
	SGLT2i vs DPP4i	Base (2-years follow-up max)	0.95	0.38	2.40
		Complete case (CC)	0.99	0.37	2.64
		Multivariable regression - CC	1.11	0.90	1.36
		Base (5-years follow-up max)	2.09	0.92	4.75
Stroke	DPP4i vs SU	Base (2-years follow-up max)	1.26	0.62	2.56
		Complete case (CC)	1.26	0.63	2.54
		Multivariable regression - CC	0.82	0.70	0.97
		Base (5-years follow-up max)	0.92	0.54	1.57
	SGLT2i vs SU	Base (2-years follow-up max)	0.63	0.30	1.31
		Complete case (CC)	0.73	0.32	1.69
		Multivariable regression - CC	0.74	0.58	0.96
		Base (5-years follow-up max)	0.65	0.30	1.39
	SGLT2i vs DPP4i	Base (2-years follow-up max)	0.50	0.21	1.21
		Complete case (CC)	0.58	0.22	1.51
		Multivariable regression - CC	0.90	0.70	1.16
		Base (5-years follow-up max)	0.71	0.30	1.64

***** Base method is the main analysis (2 stage-residual inclusion (2SRI) instrumental variable analysis with multiple imputation to account for missing data, assuming data are missing at random)

** Models could not converge for MAKE outcome extended to 5-years follow-up

Appendix E.9. Differences in the change in continuous clinical measures for the three second-line antidiabetic treatment comparisons for 2 stage-least squares (2SLS) instrumental variable analysis on complete cases (Supplementary table 14 in the research paper)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in HbA1c (mmol/mol) from baseline		**N =**	46,900	42,441	32,364	23,082	15,126	8,022
	DPP4i vs SU	**Estimate**	1.84	-0.24	-1.81	-2.91	-1.46	-1.00
		(95% CI)	(0.43, 3.25)	(-1.76, 1.28)	(-3.61, -0.01)	(-4.96, -0.86)	(-4.12, 1.21)	(-4.62, 2.62)
	SGLT2i vs SU	**Estimate**	1.42	-4.86	-6.80	-11.58	-11.20	-13.80
		(95% CI)	(-0.36, 3.21)	(-7.19, -2.52)	(-10.00, -3.59)	(-15.97, -7.18)	(-17.16, -5.24)	(-24.31, -3.30)
	SGLT2i vs DPP4i	**Estimate**	-0.41	-4.61	-4.98	-8.67	-9.74	-12.81
		(95% CI)	(-2.52, 1.69)	(-7.30, -1.93)	(-8.74, -1.23)	(-13.67, -3.67)	(-16.51, -2.98)	(-24.637, -0.98)
Difference in the change in BMI (kg/m^2) from baseline		**N =**	33,508	34,431	25,809	18,253	11,577	5,903
	DPP4i vs SU	**Estimate**	-0.41	-0.71	-0.73	-0.58	-1.24	-0.83
		(95% CI)	(-0.68, -0.14)	(-0.95, -0.46)	(-1.05, -0.40)	(-1.02, -0.14)	(-1.80, -0.67)	(-1.72, 0.06)
	SGLT2i vs SU	**Estimate**	-1.41	-1.80	-1.78	-1.86	-2.11	-0.98
		(95% CI)	(-1.71, -1.12)	(-2.14, -1.47)	(-2.29, -1.28)	(-2.65, -1.07)	(-3.17, -1.06)	(-3.35, 1.39)
	SGLT2i vs DPP4i	**Estimate**	-1.00	-1.10	-1.06	-1.28	-0.88	-0.15
		(95% CI)	(-1.36, -0.64)	(-1.45, -0.74)	(-1.62, -0.49)	(-2.13, -0.43)	(-2.05, 0.29)	(-2.86, 2.56)
Difference in the change in eGFR (mL/min/1.73m^2) from baseline		**N =**	39,113	39,337	30,034	21,398	14,060	7,659
	DPP4i vs SU	**Estimate**	-0.06	0.28	1.28	3.73	3.87	4.70
		(95% CI)	(-0.76, 0.64)	(-0.50, 1.07)	(0.26, 2.30)	(2.57, 4.88)	(2.33, 5.41)	(2.43, 6.97)
	SGLT2i vs SU	**Estimate**	0.40	1.10	2.03	3.35	5.36	9.03
		(95% CI)	(-0.59, 1.39)	(0.05, 2.15)	(0.70, 3.36)	(1.21, 5.49)	(2.69, 8.03)	(5.00, 13.07)
	SGLT2i vs DPP4i	**Estimate**	0.46	0.82	0.75	-0.38	1.49	4.33
		(95% CI)	(-0.50, 1.42)	(-0.25, 1.88)	(-0.80, 2.30)	(-2.76, 2.00)	(-1.61, 4.59)	(-0.43, 9.09)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in SBP (mm Hg) from baseline		**N =**	40,588	41,049	30,967	21,972	13,832	7,100
	DPP4i vs SU	**Estimate**	-0.77	-0.02	-1.33	-0.79	-1.57	1.72
		(95% CI)	(-1.99, 0.44)	(-1.24, 1.20)	(-2.72, 0.06)	(-2.40, 0.82)	(-3.54, 0.40)	(-1.25, 4.69)
	SGLT2i vs SU	**Estimate**	-3.35	-1.57	-2.42	-4.12	-1.10	-3.14
		(95% CI)	(-4.89, -1.81)	(-3.11, -0.03)	(-4.35, -0.50)	(-6.64, -1.60)	(-4.36, 2.16)	(-8.44, 2.15)
	SGLT2i vs DPP4i	**Estimate**	-2.58	-1.56	-1.09	-3.33	0.47	-4.86
		(95% CI)	(-4.25, -0.91)	(-3.32, 0.21)	(-3.29, 1.10)	(-6.20, -0.46)	(-3.37, 4.30)	(-11.48, 1.76)

Appendix E.10. Differences in the change in continuous clinical measures for the three second-line antidiabetic treatment comparisons for ordinary least squares (OLS) regression adjusted for measured confounders (complete cases) (Supplementary table 15 in the research paper)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in HbA1c (mmol/mol) from baseline		**N =**	46,900	42,441	32,364	23,082	15,126	8,022
	DPP4i vs SU	**Estimate**	3.30	1.30	0.09	-0.22	-0.57	-0.67
		(95% CI)	(3.00, 3.60)	(0.96, 1.65)	(-0.32, 0.51)	(-0.71, 0 .28)	(-1.18, 0.03)	(-1.52, 0.18)
	SGLT2i vs SU	**Estimate**	1.72	-1.38	-1.73	-2.04	-1.66	-0.89
		(95% CI)	(1.36, 2.09)	(-1.78, -0.98)	(-2.26, -1.21)	(-2.71, -1.38)	(-2.59, -0.73)	(-2.27, 0.49)
	SGLT2i vs DPP4i	**Estimate**	-1.58	-2.68	-1.83	-1.83	-1.09	-0.22
		(95% CI)	(-1.94, -1.22)	(-3.07, -2.29)	(-2.34, -1.32)	(-2.49, -1.16)	(-1.99, -0.19)	(-1.59, 1.15)
Difference in the change in BMI (kg/m^2) from baseline		**N =**	33,508	34,431	25,809	18,253	11,577	5,903
	DPP4i vs SU	**Estimate**	-0.53	-0.60	-0.52	-0.53	-0.50	-0.42
		(95% CI)	(-0.59, -0.47)	(-0.66, -0.54)	(-0.59, -0.44)	(-0.63, -0.44)	(-0.62, -0.38)	(-0.60, -0.23)
	SGLT2i vs SU	**Estimate**	-1.22	-1.27	-1.12	-0.98	-0.76	-0.83
		(95% CI)	(-1.28, -1.15)	(-1.35, -1.20)	(-1.21, -1.03)	(-1.11, -0.85)	(-0.94, -0.58)	(-1.11, -0.56)
	SGLT2i vs DPP4i	**Estimate**	-0.68	-0.67	-0.60	-0.45	-0.26	-0.42
		(95% CI)	(-0.74, -0.63)	(-0.73, -0.61)	(-0.69, -0.52)	(-0.57, -0.33)	(-0.44, -0.09)	(-0.67, -0.16)
Difference in the change in eGFR (mL/min/1.73m^2) from baseline		**N =**	39,113	39,337	30,034	21,398	14,060	7,659
	DPP4i vs SU	**Estimate**	-0.21	-0.19	0.08	-0.08	0.57	0.46
		(95% CI)	(-0.41, -0.01)	(-0.40, 0.02)	(-0.17, 0.33)	(-0.40, 0.24)	(0.18, 0.97)	(-0.15, 1.07)
	SGLT2i vs SU	**Estimate**	-0.10	0.02	0.79	1.06	1.65	1.90
		(95% CI)	(-0.35, 0.14)	(-0.25, 0.29)	(0.45, 1.13)	(0.61, 1.51)	(1.01, 2.28)	(1.04, 2.75)
	SGLT2i vs DPP4i	**Estimate**	0.10	0.21	0.71	1.14	1.07	1.44
		(95% CI)	(-0.11, 0.32)	(-0.03, 0.44)	(0.40, 1.02)	(0.70, 1.59)	(0.49, 1.66)	(0.54, 2.33)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in SBP (mm Hg) from baseline		**N =**	40,588	41,049	30,967	21,972	13,832	7,100
	DPP4i vs SU	**Estimate**	-0.71	-0.86	-0.54	-0.40	0.01	-0.06
		(95% CI)	(-1.02, -0.40)	(-1.16, -0.56)	(-0.90, -0.18)	(-0.84, 0.04)	(-0.51, 0.54)	(-0.79, 0.68)
	SGLT2i vs SU	**Estimate**	-2.35	-2.09	-1.83	-1.53	-0.86	-1.69
		(95% CI)	(-2.76, -1.94)	(-2.47, -1.71)	(-2.30, -1.37)	(-2.12, -0.93)	(-1.68, -0.04)	(-2.81, -0.57)
	SGLT2i vs DPP4i	**Estimate**	-1.64	-1.23	-1.30	-1.12	-0.87	-1.63
		(95% CI)	(-2.02, -1.26)	(-1.57, -0.88)	(-1.74, -0.85)	(-1.69, -0.56)	(-1.63, -0.11)	(-2.71, -0.55)

			Year of follow-up					
Outcome	**Comparison**		**0.5**	**1**	**2**	**3**	**4**	**5**
Difference in the change in SBP (mm Hg) from baseline		**N =**	40,588	41,049	30,967	21,972	13,832	7,100
	DPP4i vs SU	**Estimate**	-0.66	-0.93	-0.37	-0.52	-0.06	-0.08
		(95% CI)	(-0.99, -0.33)	(-1.26, -0.59)	(-0.74, 0.00)	(-0.95, -.085)	(-0.59, 0.47)	(-0.83, 0.67)
	SGLT2i vs SU	**Estimate**	-2.47	-2.29	-1.84	-1.34	-0.90	-1.67
		(95% CI)	(-2.97, -1.96)	(-2.83, -1.75)	(-2.45, -1.23)	(-2.06, -0.63)	(-1.80, 0.09)	(-3.15, -0.19)
	SGLT2i vs DPP4i	**Estimate**	-1.81	-1.36	-1.48	-0.83	-0.84	-1.58
		(95% CI)	(-2.29, -1.34)	(-1.86, -0.86)	(-2.06, -0.90)	(-1.53, -0.14)	(-1.82, 0.15	(-3.06, -0.10)

APPENDIX F: Chapter 6 – Supplementary materials

Appendix F.1. Description of the trial review screening process to identify suitable trials for this target trial emulation (Note S1 in the research paper)

We used a list of published trials (*reference to be added once pre-print is published*) which directly compared DPP4i vs SU among people with type 2 diabetes mellitus (T2DM). This list included 35 published trials. We screened for trials eligible for this target trial emulation in the following two-step process (see also **Table S3a**):

1. Screening the title, abstract, and methods sections to ensure the trials were phase 3, double-blind trials among a general population of people with T2DM and reported HbA1c at baseline and at 52 weeks follow-up (1-year) as an outcome. Those which did not meet these criteria were excluded at this stage.
2. Among those passing the first screening, the methods were reviewed in further detail to ensure the trial met the following criteria:
    a. Results were reported as intention-to-treat in the primary analysis.
    b. Missing outcome data (HbA1c at 1-year) were accounted for with rigorous methodology (e.g., not last observation carried forward).
    c. Did not exclude participants based on criteria we determined in advance were difficult to define in the Clinical Practice Research Datalink (CPRD) (e.g., liver laboratory test results, family history of disease, clinical judgements).
    d. Did not exclude a significant proportion of people based on a clinical characteristic so as to make the study population not representative of a general T2DM population (e.g., exclude those with a history of cancer).
    e. Outcome data were reported in tables with exact numbers (i.e., HbA1c at baseline and 1-year were not only reported in figure format).
    f. Peer-reviewed publication was accessible online.

Of the 35 published trials which directly compared DPP4i vs SU, 9 trials passed the first screening and only 1 trial passed the second screening (**Table S3b**). Reasons for not passing the first and second screening are presented in **Table S3b**.

Briefly, of the 26 trials which did not pass the first screening, 18 did not report HbA1c at 1-year follow-up, 8 were not in a general T2DM population, 3 were not double-blinded, 10 were not phase 3 trials, and 1 compared adding on DPP4i to SU vs only SU. Reasons for exclusion were not mutually exclusive.

Briefly, of the 8 trials which did not pass the second screening, 1 was a per protocol analysis, 1 used last HbA1c measure carried forward to impute missing HbA1c outcome data, 1 only presented HbA1c outcome data in a figure (no exact numbers reported in tables), 4 applied exclusion criteria which reduced the representativeness of the study population or were difficult to define in CPRD data (history of cancer, family history of medullary thyroid carcinoma, liver function tests, clinical judgement on risk of dehydration of volume depletion), and 1 trial did not have an accessible publication.

One trial by Nauck et al passed the two screens and was therefore used in this target trial emulation.

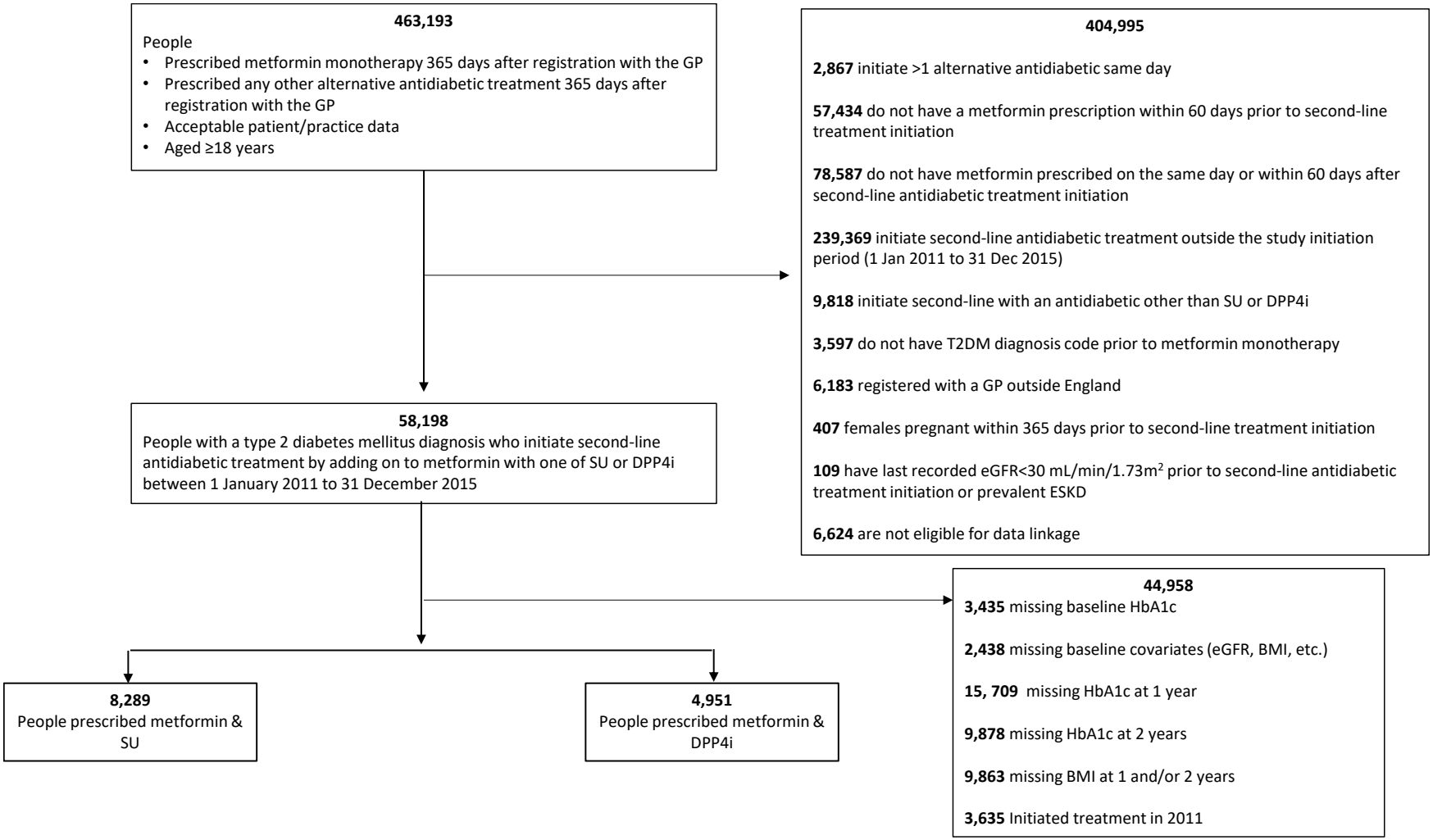
Appendix F.2. Summary of randomised controlled trial (RCT) search for target trial emulation (Table S3b in the research paper)

Trial ID	Trial phase	Study population	Change in HbA1c at 1 year reported	Screen 1	Exclusion justification	Screen 2	Exclusion justification
NCT00094770	3	T2DM	1	1	-	1	-
NCT00575588	3	T2DM	1	1	-	0	Per protocol analysis.
NCT00622284	3	T2DM	1	1	-	0	Used last measure carried forward to impute missing outcome values in primary analysis.
NCT00838903	3	T2DM	1	1	-	0	Exclusion criteria include family history of medullary thyroid carcinoma or multiple endocrine neoplasia type 2 and liver laboratory test results difficult to define in CPRD.
NCT00856284	3	T2DM	1	1	-	0	Exclusion criteria include history of cancer.
NCT01682759	3	T2DM	1	1	-	0	Cannot access article.
NCT02471404	4	T2DM	1	1	-	0	Exclusion criteria include clinical judgement for risk of dehydration or volume depletion.
NCT00102466	3	T2DM	1	1	-	0	Liver laboratory test results difficult to define in CPRD.
NCT01794143	3	T2DM	0	1	-	0	Precise HbA1c at 52 weeks not available (figure only).
NCT00102388	3	T2DM	0	0	No HbA1c at 1 year.	-	-
NCT00707993	3	T2DM and elderly (65-90 years)	1	0	Not general T2DM population.	-	-
NCT01006603	phase 3b/4	T2DM and elderly (65+ years)	1	0	Not general T2DM population.	-	-
NCT01204294	3 (open-label)	T2DM	1	0	Open-label trial.	-	-
NCT01243424	3	T2DM and CVD	0	0	Not general T2DM population, and no HbA1c	-	-

Trial ID	Trial phase	Study population	Change in HbA1c at 1 year reported	Screen 1	Exclusion justification	Screen 2	Exclusion justification
					at 1 year (48 or 64 weeks only reported as a figure).		
NCT00509236	3	T2DM and ESKD	1	0	Not general T2DM population.	-	-
NCT00509262	3	T2DM and CKD	1	0	Not general T2DM population.	-	-
NCT00701090	3	T2DM	0	0	No HbA1c at 1 year.	-	-
NCT01183104	4 (open-label)	T2DM and elderly (60+ years) and not on any other antidiaebtic, incl metformin	1	0	Not general T2DM population and open-label.	-	-
NCT01189890	3	T2DM and elderly (65-85 years)	0	0	Not general T2DM population.	-	-
NCT01822548	3		0	0	No HbA1c at 1 year.	-	-
NCT01871558	3		0	0	No HbA1c at 1 year.	-	-
NCT02007278	4		0	0	No HbA1c at 1 year and phase 4.	-	-
NCT00106340	3	T2DM	0	0	No HbA1c at 1 year.	-	-
NCT00957060	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
NCT01099137	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
NCT01341717	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
NCT01547104	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
NCT01847144	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
NCT02280486	4 (open-label)	T2DM	0	0	No HbA1c at 1 year and phase 4 open-label.	-	-
NCT03693560	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-

Trial ID	Trial phase	Study population	Change in HbA1c at 1 year reported	Screen 1	Exclusion justification	Screen 2	Exclusion justification
Eudra CT 2004-004559-21	3	T2DM	0	0	No HbA1c at 1 year.	-	-
UMIN000004791	3	T2DM	1	0	Adding on DPP4i to SU, not comparing.	-	-
UMIN000006986	4	T2DM	0	0	No HbA1c at 1 year and phase 4.	-	-
UMIN000009544	4 (open-label)	T2DM	0	0	No HbA1c at 1 year and phase 4 open-label.	-	-
UMIN000013356	4 (open-label)	T2DM with BMI >=25kg/m2 or fatty liver	0	0	No HbA1c at 1 year and phase 4 (open-lavel) and not general T2DM population.	-	-

Appendix F.3. Flow diagram illustrating the *general population* inclusion and exclusion criteria (Figure S2 of the research paper)



APPENDIX G: Chapter 7 – Supplementary materials

Appendix G.1. ICD-10 codes for AMI identified in HES (Supplementary table 1 in the research paper)

AMI subtype	ICD-10 codes
Overall	I21, I22, I23
STEMI	I21.0, I21.1, I21.2, I21.3, I22.0, I22.1, I22.8
NSTEMI	I21.4, I21.9, I22.9
Missing	I23

AMI: acute myocardial infarction; HES: Hospital Episode Statistics; ICD-10: International Classification of Diseases 10th Edition; STEMI: ST elevated myocardial infarction; NSTEMI: Non-ST elevated myocardial infarction

Appendix G.2. CALIBER definition of AMI subtypes (STEMI, NSTEMI) using MINAP data (Supplementary table 2 in the research paper)

	MINAP variable		
AMI subtype	**Discharge diagnosis**	**Markers elevated?**	**ECG result**
Other	Threatened MI, Chest pain uncertain cause, MI unconfirmed, other diagnosis	-	-
STEMI	STEMI	Raised or missing	ST elevation, LBBB, or ST elevation
	NSTEMI/Troponin positive ACS	Raised or missing	ST elevation
	ACS troponin negative	Raised	ST elevation
	ACS troponin unspecified	Raised	ST elevation
NSTEMI	NSTEMI/Troponin positive ACS	Raised or missing	ST depression, T wave changes only, other abnormality, Normal ECG, or LBBB
	ACS troponin negative	Raised	LBBB, ST depression, T wave changes only, Other abnormality, normal ECG, or missing
	ACS troponin unspecified	Raised	LBBB, ST depression, T wave changes only, Other abnormality, normal ECG, or missing
Unstable angina	*Any remaining hospitalisations not assigned as STEMI, NSTEMI, or other diagnosis		

ACS: acute coronary syndrome, AMI: acute myocardial infarction, ECG: electrocardiogram, LBBB: left bundle branch block, MI: myocardial infarction, MINAP: Myocardial Ischaemia National Audit Project, NSTEMI: Non ST-elevation myocardial infarction, STEMI: ST-elevation myocardial infarction

Appendix G.3. Details on data sources for covariates (Supplementary table 4 of the research paper)

Category	Covariate	Data source
Sociodemographic and lifestyle variables	Age at AMI admission	NCKDA
	Sex	NCKDA
	IMD quintiles	NCKDA
	Smoking status	NCKDA, MINAP
Comorbidities	Angina	MINAP and HES
	Cerebrovascular disease	MINAP and HES
	COPD	MINAP and HES
	Diabetes mellitus	MINAP and HES
	Heart failure	MINAP and HES
	Hypertension	NCKDA, MINAP, and HES
	Previous myocardial infarction	MINAP, and HES
	Peripheral vascular disease	NCKDA, MINAP, and HES
	Dialysis	NCKDA
	Kidney transplant	NCKDA

AMI: Acute Myocardial Infarction, COPD: Chronic Obstructive Pulmonary Disease, HES: Hospital Episode Statistics, IMD: Index of Multiple Deprivation, NCKDA: National Chronic Kidney Disease Audit, MINAP: Myocardial Ischaemia National Audit Project

APPENDIX H: Chapter 8 – Supplementary materials

Appendix H.1. Definitions for processes of AMI care in MINAP and HES datasets (Additional table 6 in the research paper)

Process of care	MINAP definition	HES definition
Angiography	Angiography defined using data from the following variables: • Interventional hospital procedure • Coronary angiography • Why no angiography • Procedure performed • Additional reperfusion treatment • Why no intervention • Coronary intervention	Angiography defined by searching the opertn_XX variables (XX=01-24) using the following OPCS codes: • K63
PCI	PCI defined using data from the following variables: • Why no intervention • Coronary intervention • Interventional hospital procedure • Procedure performed at admission • Initial reperfusion treatment • Additional reperfusion treatment	PCI defined by searching the opertn_XX variables (XX=01-24) using the following OPCS codes: • K49 • K50 • K75
CABG	CABG defined using data from the following variables: • Why no intervention • Coronary intervention • Interventional hospital procedure	CABG defined by searching the opertn_XX variables (XX=01-24) using the following OPCS codes: • K40 • K41 • K42 • K43 • K44 • K45 • K46

APPENDIX I: Chapter 9 – Supplementary materials

Appendix I.1. Variables used to define outcomes* (Supplementary table 1 in the manuscript)

Process of care	**MINAP definition**
Angiography	Angiography defined using data from the following variables: • Interventional hospital procedure • Coronary angiography • Why no angiography • Procedure performed • Additional reperfusion treatment • Why no intervention • Coronary intervention
Percutaneous coronary intervention (PCI)	PCI defined using data from the following variables: • Why no intervention • Coronary intervention • Interventional hospital procedure • Procedure performed at admission • Initial reperfusion treatment • Additional reperfusion treatment
Coronary artery bypass graft (CABG)	CABG defined using data from the following variables: • Why no intervention • Coronary intervention • Interventional hospital procedure
In-hospital death	Death in hospital defined using data from the following variables: • Reason for no angiography • No intervention • Admission ward • Discharge destination

*Table adapted from previous publications:

Bidulka P, Scott J, Taylor DM, et alImpact of chronic kidney disease on case ascertainment for hospitalised acute myocardial infarction: an English cohort study. *BMJ Open* 2022;12:e057909. doi: 10.1136/bmjopen-2021-057909

Scott, J., Bidulka, P., Taylor, D.M. et al. Management and outcomes of myocardial infarction in people with impaired kidney function in England. *BMC Nephrol* 24, 325 (2023). https://doi.org/10.1186/s12882-023-03377-x

Appendix I.2. Transfers between hospitals during the same AMI event in the centre-level and individual-level (complete cases) analysis (Supplementary table 4 in the manuscript)

Number of transfers during AMI event	Centre-level analysis Number of patients, n (% of total patients)	Individual-level analysis (complete cases) Number of patients, n (% of total patients)
0	334,908 (93)	270, 382 (92)
1	26,256 (7)	24,548 (8)
2	92 (0)	86 (0)
3	<5 (0)	<5 (0)
Total	361,259 (100)	295,019 (100)

AMI: acute myocardial infarction

Appendix I.3. Baseline characteristics at first AMI hospitalisation during the study period between 2014-2019 (Supplementary table 5 in the manuscript)

	Total	Missing eGFR*	Stage 1	Stage 2	Stage 3a	Stage 3b	Stage 4	Stage 5	Coded chronic renal failure**
N=	**292,572**	**13,013**	**85,825**	**116,290**	**32,583**	**18,531**	**6,247**	**1,159**	**18,924**
Female	95,044 (32)	3,672 (28)	19,926 (23)	38,019 (33)	13,597 (42)	8,879 (48)	3,252 (52)	534 (46)	7,165 (38)
Age in years, mean (SD)	69 (14)	63 (16)	58 (10)	71 (11)	78 (11)	81 (10)	82 (10)	76 (13)	78 (12)
Age category in years (%)									
<50	25,314 (9)	2,782 (21)	17,264 (20)	4,163 (4)	435 (1)	172 (1)	65 (1)	37 (3)	396 (2)
50-59	53,038 (18)	2,583 (20)	32,205 (38)	14,768 (13)	1,510 (5)	511 (3)	181 (3)	88 (8)	1,192 (6)
60-69	67,367 (23)	2,795 (21)	26,812 (31)	28,291 (24)	4,563 (14)	1,713 (9)	495 (8)	181 (16)	2,517 (13)
70-79	71,787 (25)	2,579 (20)	8,402 (10)	38,838 (33)	10,203 (31)	4,784 (26)	1,351 (22)	335 (29)	5,295 (28)
80+	75,066 (26)	2,274 (17)	1,142 (1)	30,230 (26)	15,872 (49)	11,351 (61)	4,155 (67)	518 (45)	9,524 (50)
Ethnicity									
White	263,181 (90)	11,156 (86)	75,375 (88)	106,620 (92)	29,897 (92)	17,090 (92)	5,761 (92)	1,006 (87)	16,276 (86)
Black	3,051 (1)	164 (1)	742 (1)	1,078 (1)	355 (1)	206 (1)	73 (1)	23 (2)	410 (2)
Asian	19,205 (7)	1,277 (10)	6,874 (8)	6,219 (5)	1,692 (5)	919 (5)	320 (5)	109 (9)	1,795 (9)
Mixed/other	7,135 (2)	416 (3)	2,834 (3)	2,373 (2)	639 (2)	316 (2)	93 (1)	21 (2)	443 (2)
Comorbidities									
Angina	57,417 (20)	1,919 (15)	10,000 (12)	22,877 (20)	8,295 (25)	5,127 (28)	1,595 (26)	213 (18)	7,391 (39)
Cerebrovascular disease	22,015 (8)	641 (5)	2,913 (3)	8,436 (7)	3,520 (11)	2,350 (13)	816 (13)	104 (9)	3,235 (17)
COPD	45,547 (16)	1,565 (12)	11,253 (13)	18,206 (16)	5,667 (17)	3,337 (18)	1,062 (17)	179 (15)	4,278 (23)
Diabetes mellitus	70,726 (24)	2,885 (22)	15,873 (18)	24,036 (21)	9,406 (29)	6,463 (35)	2,476 (40)	494 (43)	9,093 (48)
Heart failure	16,607 (6)	404 (3)	1,435 (2)	4,853 (4)	2,683 (8)	2,231 (12)	940 (15)	99 (9)	3,962 (21)
Hypercholesterolaemia	91,157 (31)	3,586 (28)	25,909 (30)	37,227 (32)	10,222 (31)	5,415 (29)	1,590 (25)	255 (22)	6,953 (37)
Hypertension	143,109 (49)	5,100 (39)	32,274 (38)	57,574 (50)	18,852 (58)	11,172 (60)	3,723 (60)	641 (55)	13,773 (73)
Previous MI	55,768 (19)	1,880 (14)	10,837 (13)	21,273 (18)	8,051 (25)	4,994 (27)	1,623 (26)	229 (20)	6,881 (36)

Peripheral vascular disease	12,729 (4)	381 (3)	2,187 (3)	4,408 (4)	1,749 (5)	1,091 (6)	401 (6)	76 (7)	2,436 (13)
Kidney failure	18,924 (6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	18,924 (100)
Co-prescriptions									
Beta-blocker	72,263 (25)	1,897 (15)	13,888 (16)	27,859 (24)	10,433 (32)	6,745 (36)	2,228 (36)	324 (28)	8,889 (47)
RASi	93,634 (32)	2,467 (19)	20,650 (24)	37,684 (32)	13,286 (41)	8,023 (43)	2,487 (40)	324 (28)	8,713 (46)
Statin	113,325 (39)	3,562 (27)	25,198 (29)	45,691 (39)	15,117 (46)	9,052 (49)	2,999 (48)	508 (44)	11,198 (59)
Smoking status									
Non-smoker	118,602 (41)	4,882 (38)	25,686 (30)	50,178 (43)	15,389 (47)	9,264 (50)	3,422 (55)	589 (51)	9,192 (49)
Ex-smoker	95,914 (33)	3,580 (28)	20,957 (24)	41,639 (36)	12,478 (38)	7,079 (38)	2,159 (35)	404 (35)	7,618 (40)
Current smoker	78,056 (27)	4,551 (35)	39,182 (46)	24,473 (21)	4,716 (14)	2,188 (12)	666 (11)	166 (14)	2,114 (11)

* Missing eGFR and no code indicating prevalent chronic renal failure

** Anyone with a code indicating prevalent chronic renal failure, regardless of what eGFR stage the SCr recorded within 24 hours of admission corresponds to

Appendix I.4. Description of people with coded chronic renal failure, stratified by the eGFR stage corresponding to the SCr measured within 24 hours of hospitalisation (Supplementary table 6 in the manuscript)

	Total	Missing eGFR	Stage 1	Stage 2	Stage 3a	Stage 3b	Stage 4	Stage 5
N= (row %)	**18,924 (100)**	**542 (3)**	**295 (2)**	**2,114 (1)**	**3,090 (2)**	**5,116 (27)**	**4,953 (26)**	**2,814 (15)**
Female	7,165 (38)	167 (31)	71 (24)	851 (40)	1,186 (38)	2,044 (40)	1,871 (38)	975 (35)
Age in years, mean (SD)	78 (12)	73 (13)	61 (11)	75 (12)	78 (11)	80 (10)	80 (11)	73 (12)
Age category in years (%)								
<50	396 (2)	24 (4)	40 (14)	58 (3)	43 (1)	57 (1)	74 (1)	100 (4)
50-59	1,192 (6)	59 (11)	92 (31)	166 (8)	140 (5)	186 (4)	230 (5)	319 (11)
60-69	2,517 (13)	118 (22)	99 (34)	367 (17)	389 (13)	484 (9)	516 (10)	544 (19)
70-79	5,295 (28)	135 (25)	48 (16)	685 (32)	907 (29)	1,392 (27)	1,266 (26)	862 (31)
80+	9,524 (50)	206 (38)	16 (5)	838 (40)	1,611 (52)	2,997 (59)	2,867 (58)	989 (35)
Ethnicity								
White	16,276 (86)	424 (78)	248 (84)	1,886 (89)	2,751 (89)	4,557 (89)	4,291 (87)	2,119 (75)
Black	410 (2)	33 (6)	7 (2)	27 (1)	45 (1)	73 (1)	98 (2)	127 (5)
Asian	1,795 (9)	62 (11)	30 (10)	156 (7)	238 (8)	392 (8)	456 (9)	461 (16)
Mixed/other	443 (2)	23 (4)	10 (3)	45 (2)	56 (2)	94 (2)	108 (2)	107 (4)
Comorbidities								
Angina	7,391 (39)	209 (39)	102 (35)	800 (38)	1,236 (40)	2,152 (42)	1,906 (38)	986 (35)
Cerebrovascular disease	3,235 (17)	85 (16)	56 (19)	341 (16)	542 (18)	880 (17)	881 (18)	450 (16)
COPD	4,278 (23)	130 (24)	92 (31)	596 (28)	742 (24)	1,135 (22)	1,091 (22)	492 (17)
Diabetes mellitus	9,093 (48)	255 (47)	117 (40)	773 (37)	1,289 (42)	2,431 (48)	2,594 (52)	1,634 (58)
Heart failure	3,962 (21)	105 (19)	54 (18)	329 (16)	559 (18)	1,196 (23)	1,259 (25)	460 (16)
Hypercholesterolaemia	6,953 (37)	220 (41)	140 (47)	837 (40)	1,192 (39)	1,892 (37)	1,701 (34)	971 (35)
Hypertension	13,773 (73)	399 (74)	195 (66)	1,497 (71)	2,276 (74)	3,748 (73)	3,597 (73)	2,061 (73)
Previous MI	6,881 (36)	208 (38)	94 (32)	695 (33)	1,092 (35)	1,930 (38)	1,866 (38)	996 (35)
Peripheral vascular disease	2,436 (13)	80 (15)	54 (18)	234 (11)	352 (11)	653 (13)	655 (13)	408 (14)
Kidney failure	18,924 (100)	542 (100)	295 (100)	2,114 (100)	3,090 (100)	5,116 (100)	4,953 (100)	2,814 (100)
Co-prescriptions								

Beta-blocker	8,889 (47)	236 (44)	113 (38)	841 (40)	1,393 (45)	2,469 (48)	2,446 (49)	1,391 (49)
RASi	8,713 (46)	231 (43)	135 (46)	1,018 (48)	1,637 (53)	2,626 (51)	2,207 (45)	859 (31)
Statin	11,198 (59)	294 (54)	148 (50)	1,183 (56)	1,847 (60)	3,021 (59)	2,999 (61)	1,706 (61)
Smoking status								
Non-smoker	9,192 (49)	252 (46)	94 (32)	980 (46)	1,474 (48)	2,491 (49)	2,434 (49)	1,467 (52)
Ex-smoker	7,618 (40)	195 (36)	101 (34)	828 (39)	1,293 (42)	2,137 (42)	2,032 (41)	1,032 (37)
Current smoker	2,114 (11)	95 (18)	100 (34)	306 (14)	323 (10)	488 (10)	487 (10)	315 (11)

COPD: chronic obstructive pulmonary disease; eGFR: estimated glomerular filtration rate; MI: myocardial infarction; RASi: renin-angiotensin system inhibitor

Appendix I.5. Association between eGFR stage and angiography/PCI intervention, overall and by AMI subtype (Supplementary table 7 in the manuscript)

eGFR stage at first hospital	No. (row %)	Total	Age and sex adjusted OR (95% CI)	Adjusted* OR (95%CI)	Adjusted** with centre as fixed effect OR (95% CI)	Adjusted*** predicted percent (95% CI)
OVERALL						
Missing eGFR	10,980 (84)	13,013	0.73 (0.69-0.77)	0.70 (0.67-0.74)	0.58 (0.54-0.63)	78.7 (78.0-79.4)
1	80,092 (93)	85,825	0.88 (0.85-0.91)	0.86 (0.83-0.89)	0.84 (0.81-0.87)	81.1 (80.8-81.5)
2	95,979 (82)	116,290	1 (reference)	1 (reference)	1 (reference)	83.0 (82.8-83.2)
3a	22,047 (68)	32,583	0.68 (0.66-0.70)	0.72 (0.69-0.74)	0.70 (0.67-0.72)	78.9 (78.6-79.2)
3b	9,916 (54)	18,531	0.45 (0.43-0.46)	0.49 (0.47-0.51)	0.45 (0.44-0.47)	73.6 (73.1-74.1)
4	2,354 (38)	6,247	0.24 (0.22-0.25)	0.26 (0.25-0.28)	0.23 (0.21-0.24)	63.7 (62.7-64.7)
5	464 (40)	1,159	0.16 (0.14-0.18)	0.17 (0.15-0.19)	0.15 (0.13-0.17)	55.7 (53.3-58.1)
Chronic renal failure	9,778 (52)	18,924	0.31 (0.30-0.32)	0.39 (0.37-0.40)	0.39 (0.37-0.40)	69.9 (69.4-70.5)
STEMI only						
Missing eGFR	6,192 (94)	6,573	0.76 (0.67-0.85)	0.77 (0.68-0.86)	-	92.6 (91.9-93.2)
1	38,565 (98)	39,514	0.94 (0.87-1.03)	0.93 (0.86-1.02)	-	93.7 (93.3-94.1)
2	38,588 (93)	41,315	1 (reference)	1 (reference)	-	94.0 (93.8-94.2)
3a	8,246 (85)	9,718	0.62 (0.57-0.66)	0.63 (0.59-0.68)	-	91.3 (90.9-91.7)
3b	3,811	5,082	0.40	0.43	-	88.4

eGFR stage at first hospital	No. (row %)	Total	Age and sex adjusted OR (95% CI)	Adjusted* OR (95%CI)	Adjusted** with centre as fixed effect OR (95% CI)	Adjusted*** predicted percent (95% CI)
	(75)		(0.37-0.44)	(0.40-0.47)		(87.8-89.0)
4	1,062 (62)	1,726	0.24 (0.21-0.26)	0.25 (0.22-0.28)	-	83.0 (81.8-84.3)
5	189 (60)	314	0.15 (0.12-0.20)	0.16 (0.13-0.21)	-	77.6 (74.0-81.1)
Chronic renal failure	2,198 (70)	3,130	0.25 (0.23-0.28)	0.31 (0.28-0.34)	-	85.2 (84.3-86.2)
NSTEMI only						
Missing eGFR	4,788 (74)	6,440	0.58 (0.54-0.61)	0.57 (0.53-0.61)	-	67.8 (66.7-68.9)
1	41,527 (90)	46,311	0.79 (0.75-0.82)	0.80 (0.77-0.83)	-	73.2 (72.7-73.8)
2	57,391 (77)	74,975	1 (reference)	1 (reference)	-	76.6 (76.3-76.8)
3a	13,801 (60)	22,865	0.70 (0.68-0.73)	0.73 (0.70-0.75)	-	71.8 (71.3-72.2)
3b	6,105 (45)	13,449	0.45 (0.43-0.47)	0.48 (0.46-0.50)	-	64.9 (64.3-65.6)
4	1,292 (29)	4,521	0.21 (0.20-0.17)	0.23 (0.22-0.25)	-	51.8 (50.5-53.2)
5	275 (33)	845	0.15 (0.13-0.38)	0.16 (0.13-0.18)	-	44.5 (41.2-47.5)
Chronic renal failure	7,580 (48)	15,794	0.37 (0.35-0.38)	0.42 (0.40-0.44)	-	62.5 (61.8-63.1)

* Adjusted for age (continuous), sex, ethnicity (White, Black, Asian, Mixed/other), comorbidities (previous MI, angina, hypertension, hypercholesterolaemia, peripheral vascular disease, cerebrovascular disease, COPD, heart failure, type 2 diabetes mellitus), prevalent prescriptions (RASi, beta blockers, statins), and smoking status

** Adjusted for all covariates plus hospital at first admission as a fixed effect (to account for clustering at the centre-level)

*** Adjusted predicted percentages derived using the adjusted model, without adjusting for hospital at first admission as a fixed effect

Appendix I.6. Partial F-statistics summarising the strength of association between the proportion of people with NSTEMI receiving cardiac investigation and/or intervention in the 1-year prior to a person's admission date (the proposed instrumental variable) and the treatment actually received, overall and stratified by hospital's PCI availability (not included in the manuscript in Chapter 9).

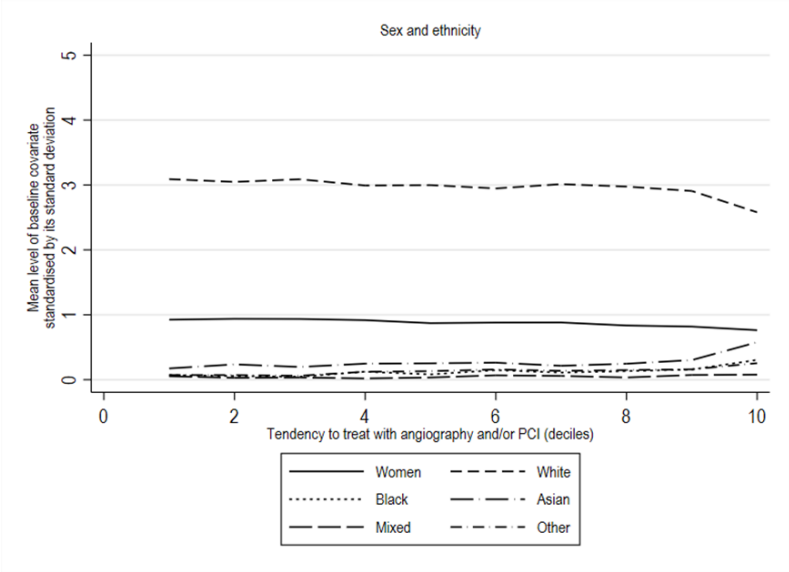
PCI services availability at admission hospital	**Partial F-statistic**
Overall	4,619
PCI not available	1,533
PCI sometimes available	688
PCI always available	1,077

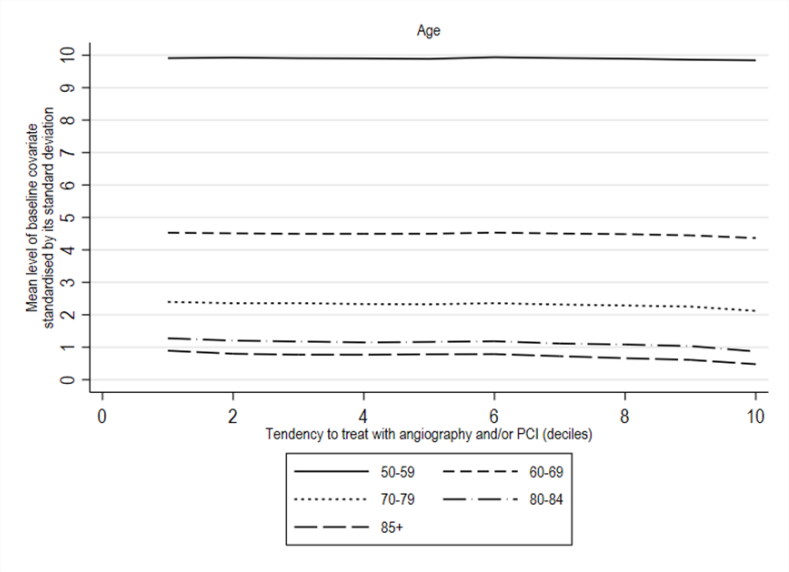
Appendix I.7. Balance of standardised covariates across levels of the proposed instrumental variable for invasive cardiac treatment among people hospitalised for NSTEMI with kidney impairment, including all hospital centres in the study (not included in the manuscript in Chapter 9).

(3A) Tendency to treat with invasive cardiac treatment by sex, ethnicity, and age

IV: tendency to treat with invasive cardiac treatment

All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI

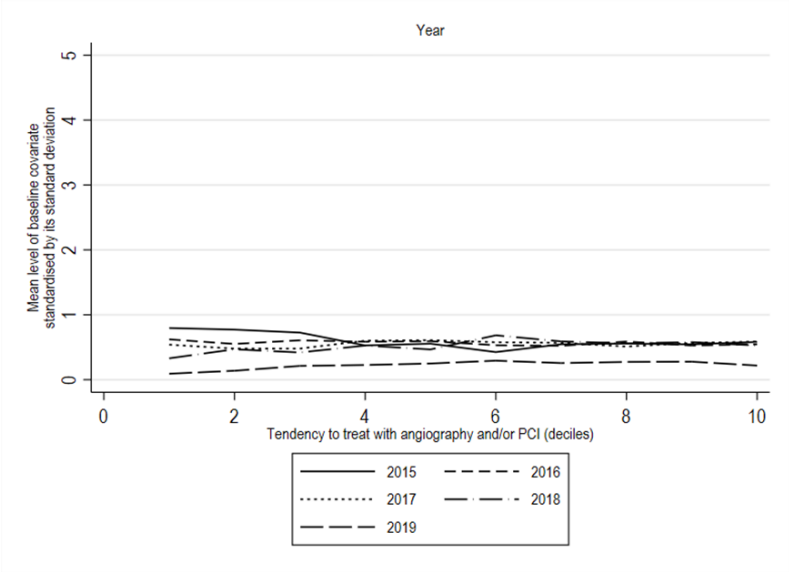


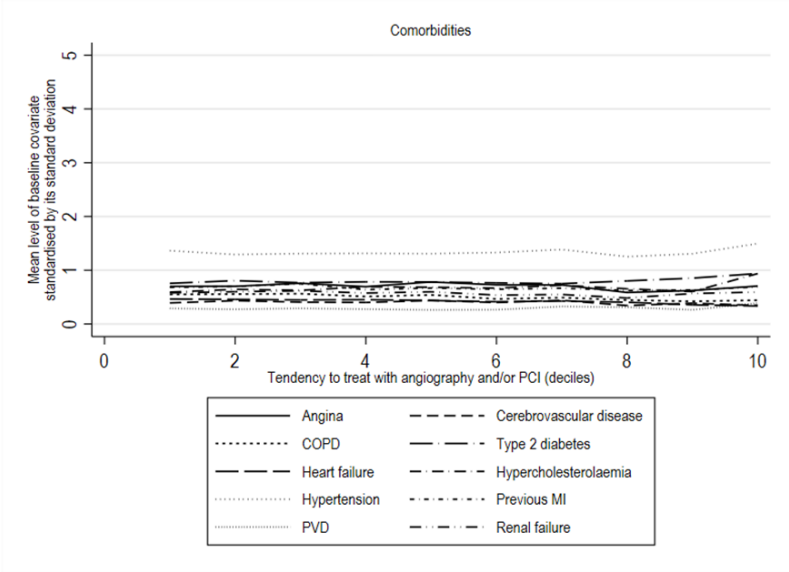


(3B) Tendency to treat with invasive cardiac treatment by year of AMI admission and comorbidities

IV: tendency to treat with invasive cardiac treatment

All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI

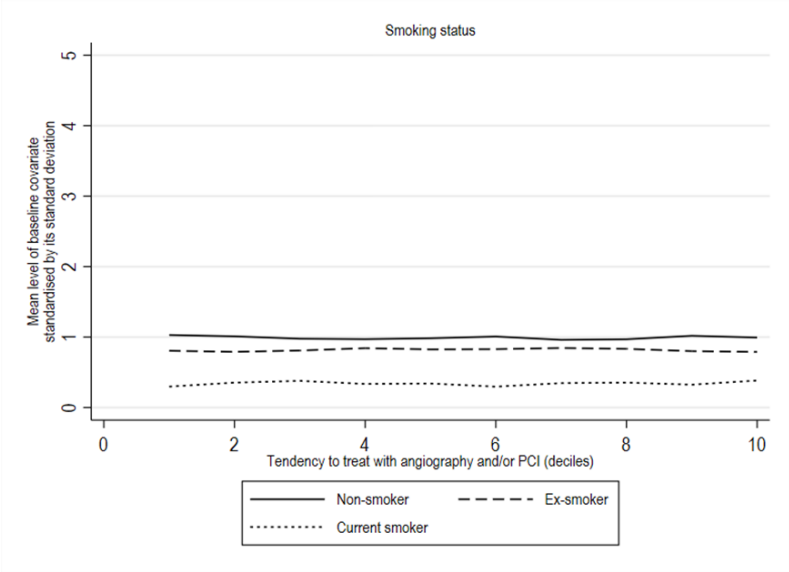


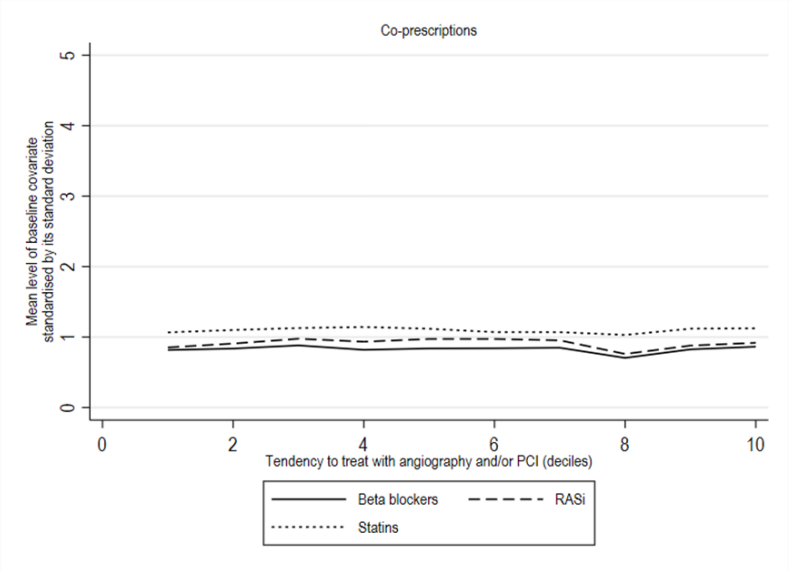


(3C) Tendency to treat with invasive cardiac treatment by smoking status and co-prescriptions (prescribed prior to admission)

IV: tendency to treat with invasive cardiac treatment

All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI

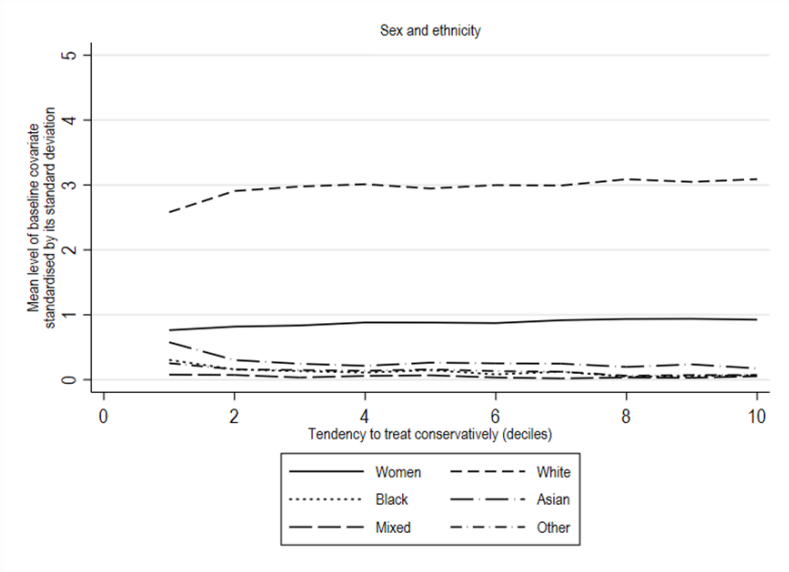


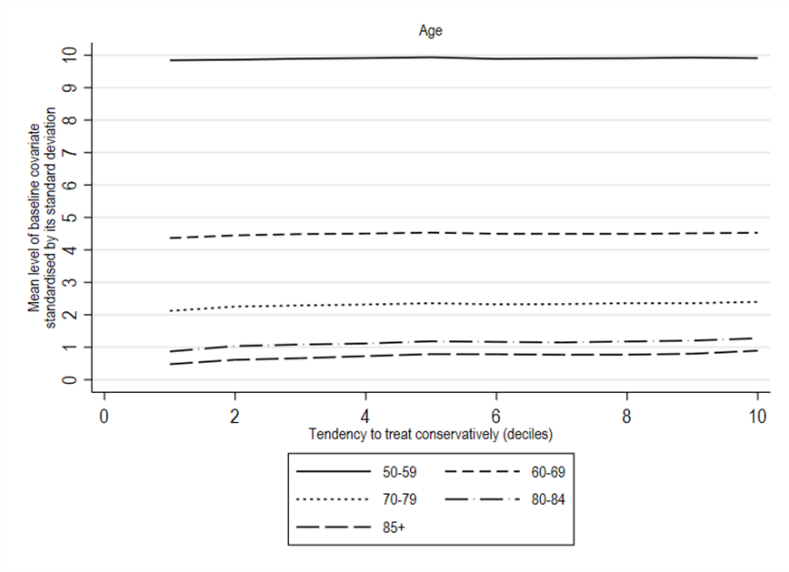


(3D) Tendency to treat with conservative treatment by sex, ethnicity, and age

IV: tendency to treat conservatively

All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI

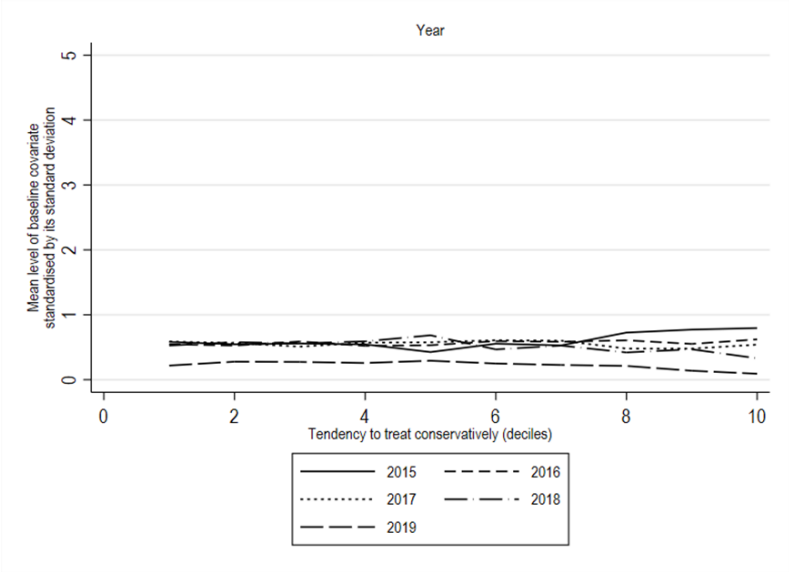


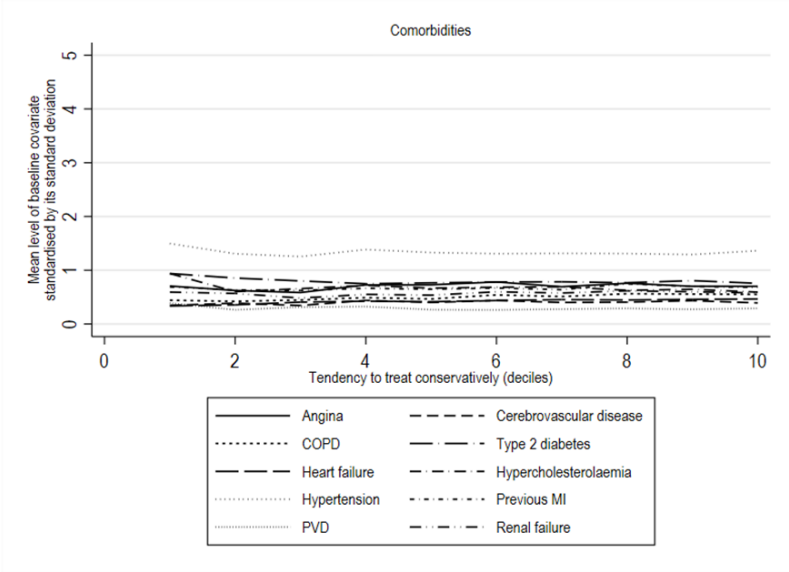


(3E) Tendency to treat with conservative treatment by year of AMI admission and comorbidities

IV: tendency to treat conservatively

All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI

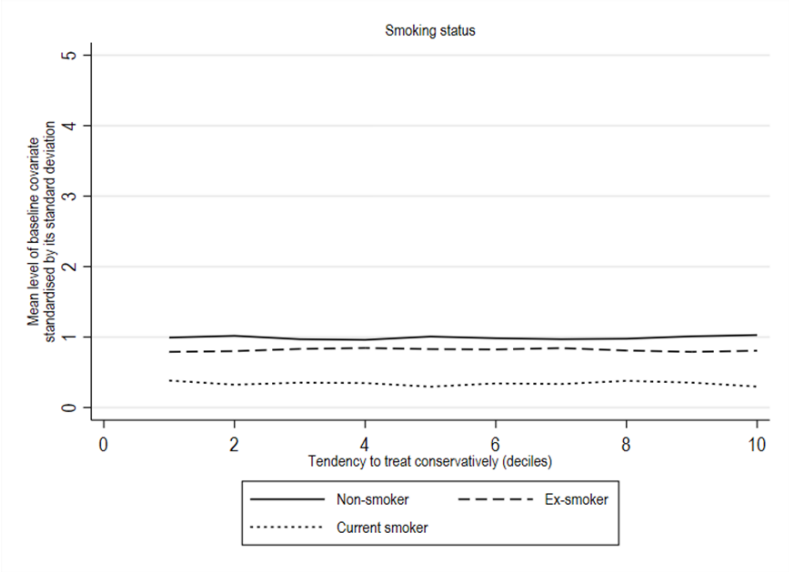


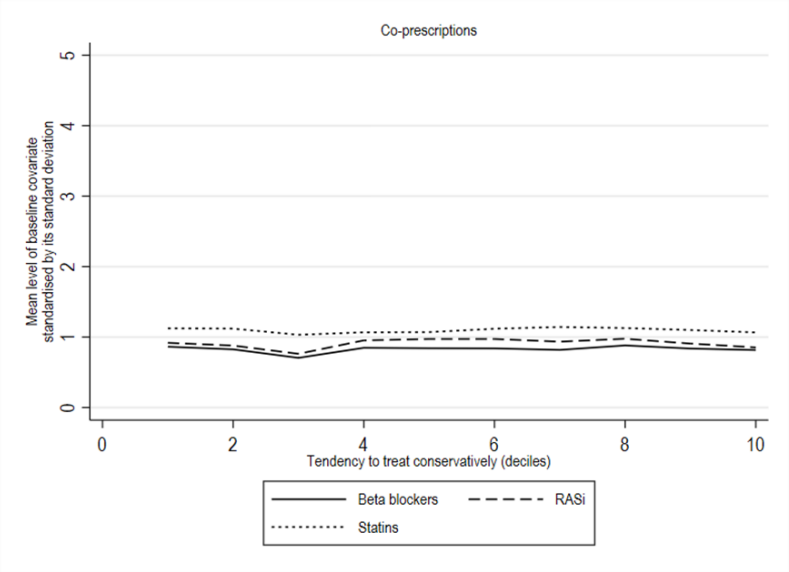


(3F) Tendency to treat with conservative treatment by smoking status and co-prescriptions

IV: tendency to treat conservatively

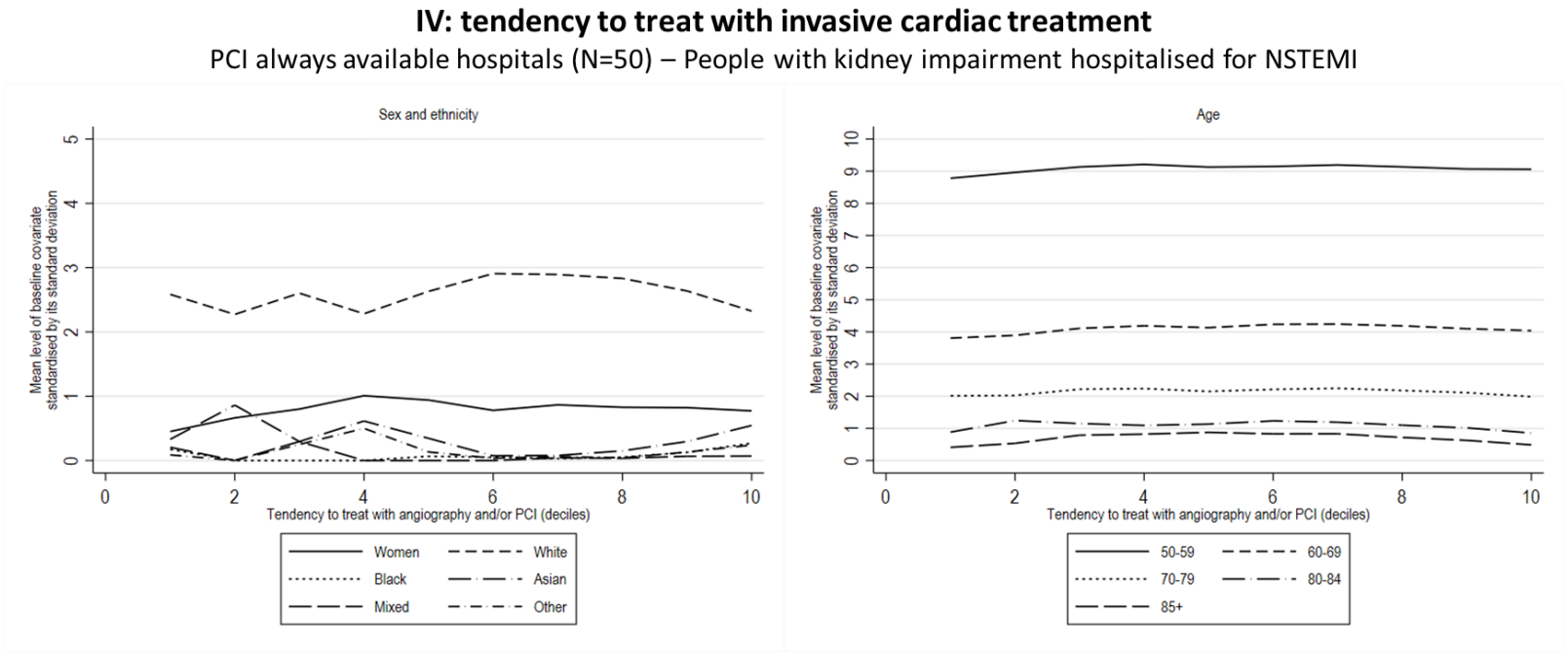
All hospitals (N=207) – People with kidney impairment hospitalised for NSTEMI





Appendix I.8. Balance of standardised covariates across levels of the proposed instrumental variable for invasive cardiac treatment among people hospitalised for NSTEMI with kidney impairment, including only hospital centres with PCI available all the time (not included in the manuscript in Chapter 9)

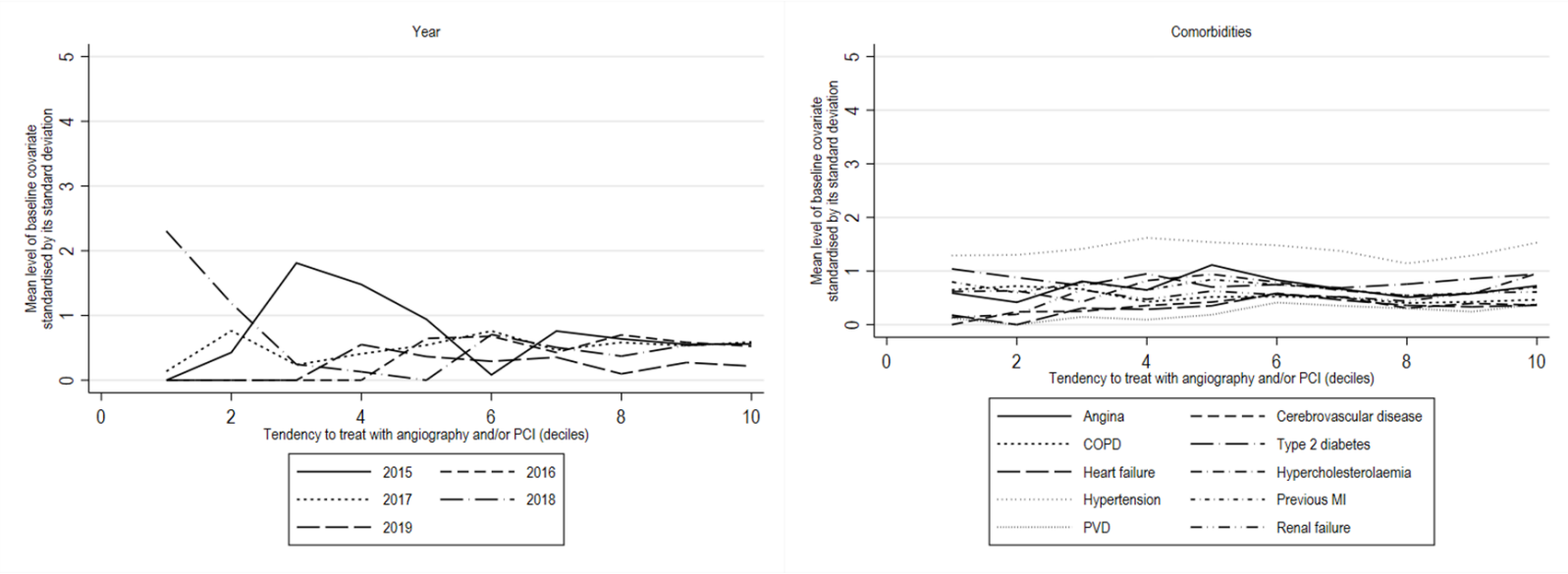
(4A) Tendency to treat with invasive cardiac treatment by sex, ethnicity, and age



(4B) Tendency to treat with invasive cardiac treatment by year of AMI admission and comorbidities

IV: tendency to treat invasive cardiac treatment

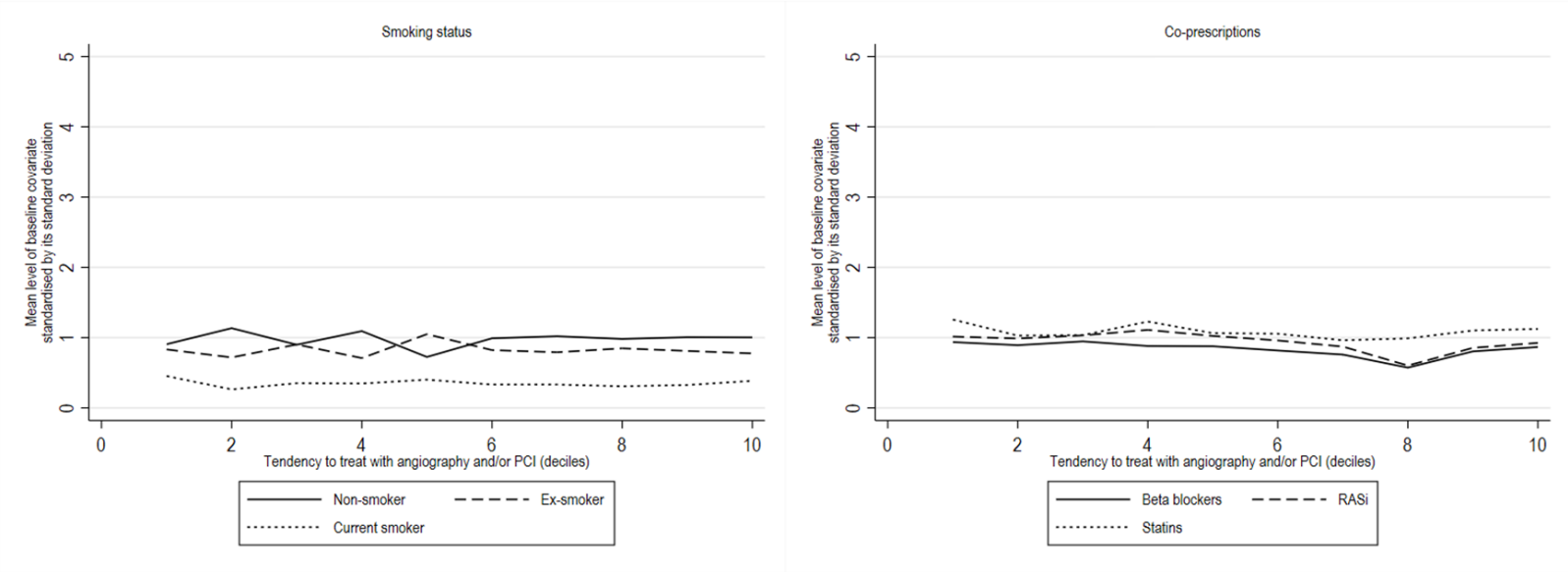
PCI always available hospitals (N=50) – People with kidney impairment hospitalised for NSTEMI



(4C) Tendency to treat with invasive cardiac treatment by smoking status and co-prescriptions (prescribed prior to admission)

IV: tendency to treat with invasive cardiac treatment

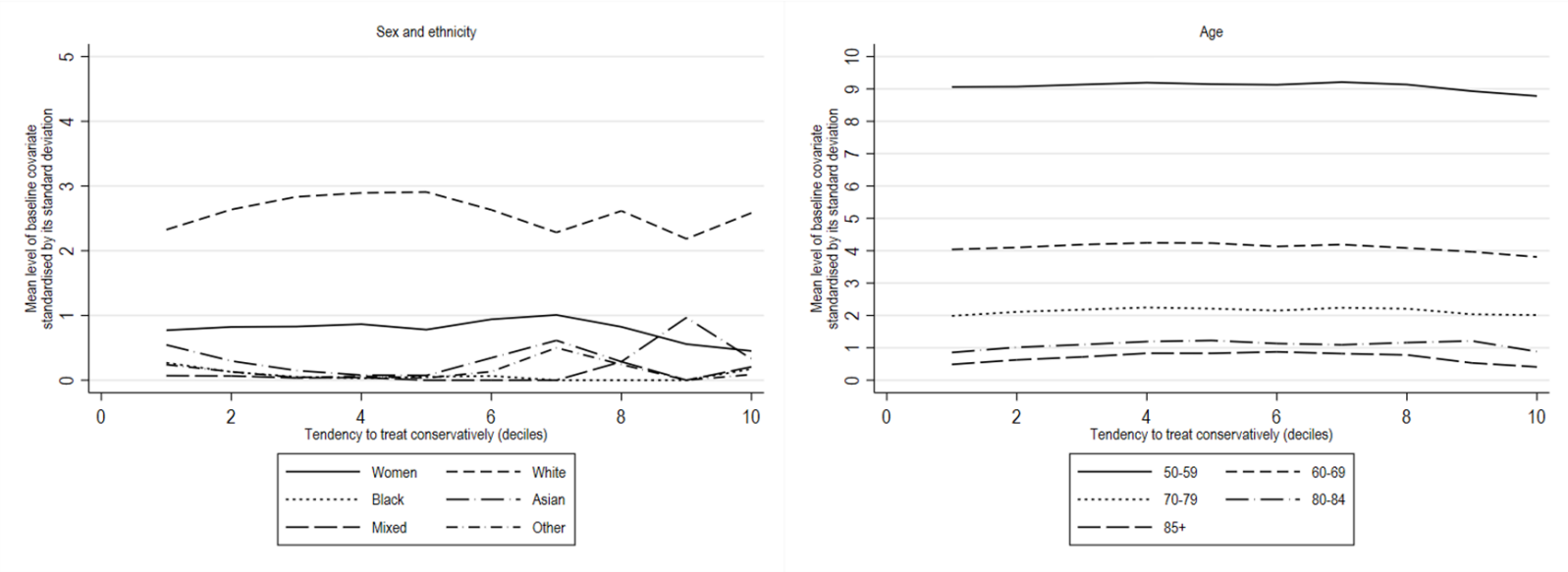
PCI always available hospitals (N=50) – People with kidney impairment hospitalised for NSTEMI



(4D) Tendency to treat with conservative treatment by sex, ethnicity, and age

IV: tendency to treat conservatively

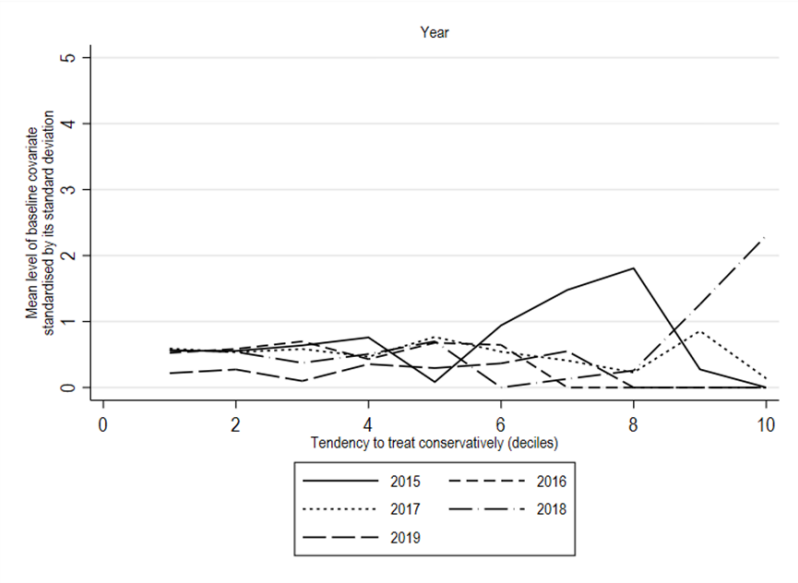
PCI always available hospitals (N=50) – People with kidney impairment hospitalised for NSTEMI

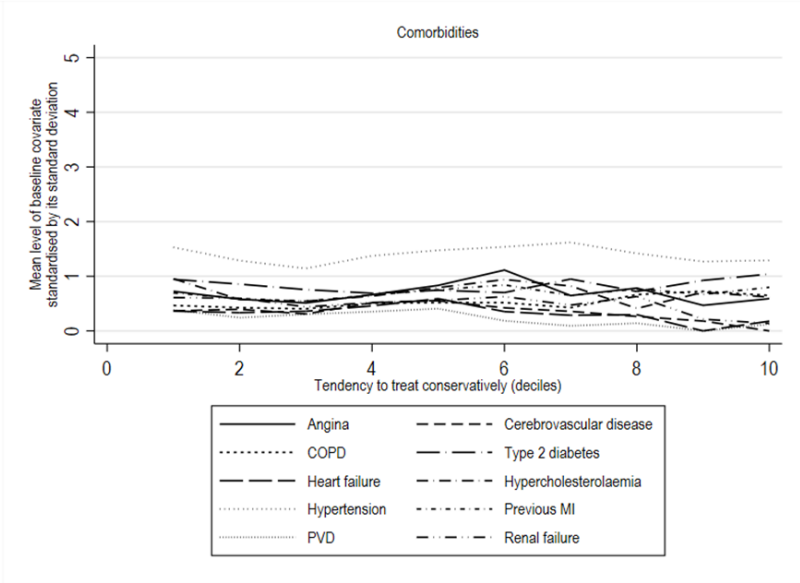


(4E) Tendency to treat with conservative treatment by year of AMI hospitalisation and comorbidities

IV: tendency to treat conservatively

PCI always available hospitals (N=50) – People with kidney impairment hospitalised for NSTEMI

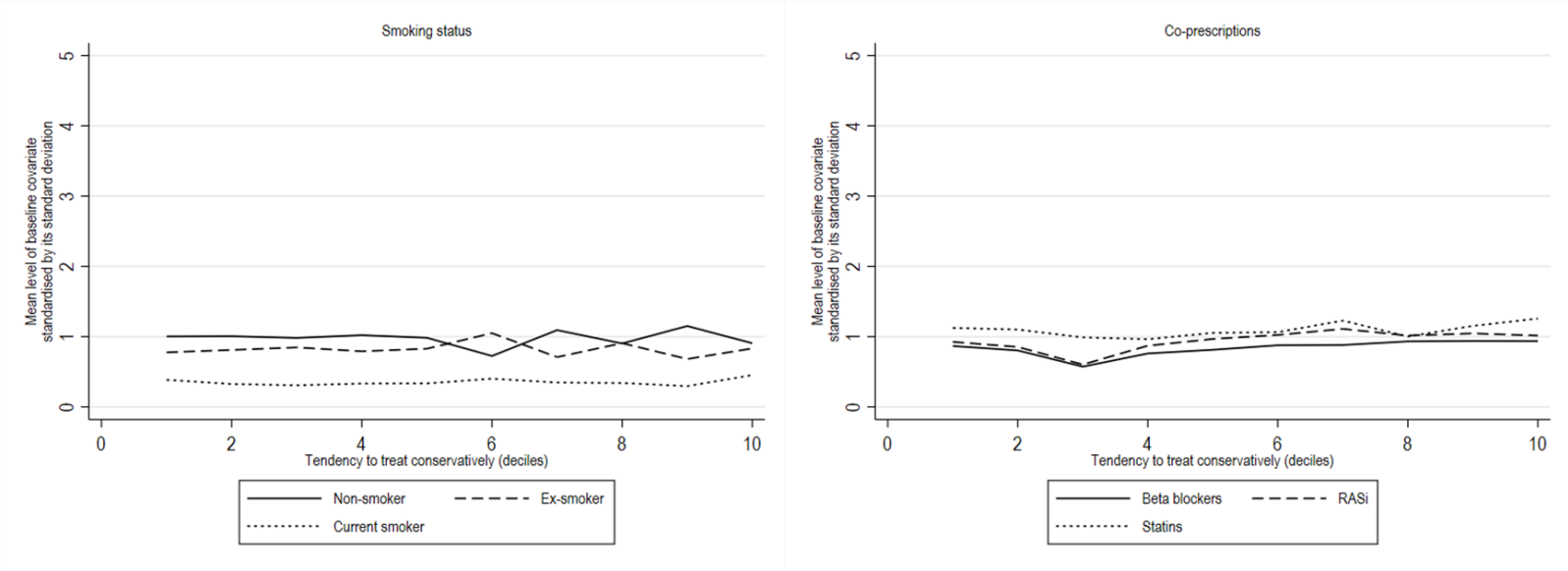




(4F) Tendency to treat with conservative treatment by smoking status and co-prescriptions (prescribed prior to admission)

IV: tendency to treat conservatively

PCI always available hospitals (N=50) – People with kidney impairment hospitalised for NSTEMI



APPENDIX REFERENCES

1. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and evaluation of the monotonicity condition for preference-based instruments. *Epidemiology (Cambridge, Mass)*. May 2015;26(3):414-20. doi:10.1097/ede.0000000000000279

2. Gourieroux C, Monfort A, Renault E, Trognon A. Generalised residuals. *Journal of Econometrics*. 1987/01/01/ 1987;34(1):5-32. doi:https://doi.org/10.1016/0304-4076(87)90065-0

3. Martínez-Camblor P, Mackenzie T, Staiger DO, Goodney PP, O'Malley AJ. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostatistics*. Jan 1 2019;20(1):80-96. doi:10.1093/biostatistics/kxx062

4. Martínez-Camblor P, Mackenzie T, Staiger DO, Goodney PP, O'Malley AJ. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostatistics*. 2017;20(1):80-96. doi:10.1093/biostatistics/kxx062

5. Frank lE, Friedman JH. A Statistical View of Some Chemometrics Regression Tools. *Technometrics*. 1993/05/01 1993;35(2):109-135. doi:10.1080/00401706.1993.10485033

6. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996;58(1):267-288. doi:https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

7. Belloni A, Chen D, Chernozhukov V, Hansen C. Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica*. 2012;80(6):2369-2429. doi:https://doi.org/10.3982/ECTA9626

8. Ahrens A, Hansen CB, Schaffer ME. lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*. 2020;20(1):176-235. doi:10.1177/1536867x20909697

9. van Buuren S, Oudshoorn, C.G.M. *Multivariate Imputation by Chained Equations: MICE V1.0 User's manual. TNO Report PG/VGZ/00.038*. 2000. http://www.multiple-imputation.com/

10. Powney M, Williamson P, Kirkham J, Kolamunnage-Dona R. A review of the handling of missing longitudinal outcome data in clinical trials. *Trials*. Jun 19 2014;15:237. doi:10.1186/1745-6215-15-237

11. Lee KJ, Roberts G, Doyle LW, Anderson PJ, Carlin JB. Multiple imputation for missing data in a longitudinal cohort study: a tutorial based on a detailed case study involving imputation of missing outcome data. *International Journal of Social Research Methodology*. 2016/09/02 2016;19(5):575-591. doi:10.1080/13645579.2015.1126486

12. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. *BMC Medical Research Methodology*. 2014/06/05 2014;14(1):75. doi:10.1186/1471-2288-14-75

13. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020/08/01 2020;584(7821):430-436. doi:10.1038/s41586-020-2521-4

14. Mathur R, Rentsch CT, Morton CE, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet (London, England)*. May 8 2021;397(10286):1711-1724. doi:10.1016/s0140-6736(21)00634-6

15. Meng X-L. Multiple-Imputation Inferences with Uncongenial Sources of Input. *Statistical Science*. 1994;9(4):538-558, 21.

16. Seaman SR, Bartlett JW, White IR. Multiple imputation of missing covariates with non-linear effects and interactions: an evaluation of statistical methods. *BMC Medical Research Methodology*. 2012/04/10 2012;12(1):46. doi:10.1186/1471-2288-12-46

17. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 12/12 2011;45(3):1 - 67. doi:10.18637/jss.v045.i03

18. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. Feb 20 2011;30(4):377-99. doi:10.1002/sim.4067

19. van Buuren S. *Flexible Imputation of Missing Data, Second Edition*. 2 ed. Chapman and Hall/CRC; 2018:444.

20. Carpenter JR, Kenward MG. *Multiple Imputation and its Application*. John Wiley & Sons, Ltd; 2012.

21. Bartlett JW, Hughes RA. Bootstrap inference for multiple imputation under uncongeniality and misspecification. *Statistical Methods in Medical Research*. 2020;29(12):3533-3546. doi:10.1177/0962280220932189

22. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. *Stat Med*. Jun 30 2018;37(14):2252-2266. doi:10.1002/sim.7654