COMPARATIVE EFFECTIVENESS AND SAFETY OF DPP-4 INHIBITORS

Jil Billy Mamza, BPharm, MPH

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

December 2015

Abstract

Approximately 3 million people throughout the UK suffer with Type 2 diabetes mellitus (T2DM), and are 32% more likely to die early. There remains a lack of evidence for the long-term effectiveness and safety of anti-diabetic drugs in preventing diabetes-related complications, making it unclear how to optimally manage diabetes. Work to date includes observational studies subject to bias, and randomised controlled trials (RCTs), which may not reflect the 'real world' situation. The aim of this thesis is to combine such findings via systematic reviews, meta-analyses and retrospective cohort studies to provide more water-tight evidence of the effectiveness and safety of glucose-lowering therapies (GLT) in the long term.

Firstly, a systematic review of observational studies was performed, identifying and providing a simple description of the types of biases and control measures employed in retrospective cohort studies on treatment outcomes of GLTs. Secondly, retrospective cohort studies were conducted to strengthen the evidence of the clinical effectiveness of DPP-4 inhibitors, compare their durability when combined with other anti-diabetic drugs and assess their cardiovascular safety when used in patients with T2DM, using data from The Health Improvement Network (THIN) database. Linear and logistic regression, Cox proportional hazard regression models and propensity score techniques were used to analyse routine clinical data. Thirdly, a meta-analysis was conducted on RCTs investigating the risk of bone fracture following the

administration of DPP-4 inhibitor, based on data from an extensive literature search.

Conducting this research has led to a better understanding of how biases may have influenced retrospective cohort studies on oral anti-diabetic drugs. An algorithm was developed to illustrate strategies for addressing biases. Potential clinical factors associated with 'response' to DPP-4 inhibitor treatment were found to include the addition of DPP-4 inhibitor to ongoing metformin (MET), or MET plus sulphonylurea (SU) therapy. High HbA1c at the time of treatment intensification and longer duration of diabetes were associated with the lack of HbA1c target attainment. In terms of the durability of second-line glucose-lowering agents, the co-administration of thiazolidinedione with MET was associated with the most durable glycaemic response, followed by a SU and then a DPP-4 inhibitor. Compared with a SU, adding a DPP-4 inhibitor to MET was associated with an increased need for earlier treatment intensification with a third agent. The use of statins, being a female, a smoker, having longer duration of diabetes and higher HbA1c at baseline were identified to be associated with earlier dual therapy failure.

In terms of cardiovascular safety, routine clinical data showed patients who intensified MET + SU dual therapy with a DPP-4 inhibitor were associated with a decreased risk of a composite of non-fatal cardiovascular outcomes and all-cause mortality compared to those who added insulin. Furthermore, the results from meta-analysis showed DPP-4 inhibitors are not associated with increased bone fracture risks in patients with T2DM.

This research is valuable in informing the choices of healthcare professionals in prescribing treatments for T2DM. For the users of this treatment, it is good news that DPP-4 inhibitors are not generally associated with fracture incidence, and that findings support the use of DPP-4 inhibitors as a second line therapeutic option, especially among non-obese patients whose glucose control remains suboptimal following MET treatment. It is recommended that treatment should be characterised on an individual basis. There remains a need for robust RCTs to investigate the influence of obesity and longer treatment durations on the efficacy of co-administering DPP-4 inhibitors to patients who are unresponsive to other oral GLTs.

Peer-reviewed publications and presentations from thesis

Mamza J. et al., Comparative Efficacy of Adding Sitagliptin to Metformin, Sulfonylurea or Dual Therapy: A Propensity Score-Weighted Cohort Study. Diabetes Therapy, 2015. 6(2): p. 213-226.

Mamza J. et al., Determinants of Glycaemic Response to Add-on Therapy with a DPP-4 inhibitor: A Retrospective Cohort Study using a UK Primary Care database. Diabetes Tech. and Therap. 2016 Feb;18(2):85-92

Mamza J. et al., Important differences in the durability of glycaemic response among second-line treatment options when added to metformin in type 2 diabetes: a retrospective cohort study. Annals of Medicine, 2016 Jun; 48(4):224-34

Mamza J. et al., DPP-4 inhibitor therapy and bone fractures in people with type 2 diabetes: a systematic review and meta-analysis. Diabetes Research and Clinical Practice, 2016 (Accepted: DIAB_DIAB-D-16-00168)

Mamza J. et al., Effects of Dual Therapy Intensification with Insulin or a DPP-4 Inhibitor on Cardiovascular Events and All-Cause Mortality in Patients with Type 2 Diabetes. (currently under review)

Oral Presentations

2015 American Diabetes Association Scientific Meeting, Boston, USA Durability of Second-Line Oral Glucose-Lowering Therapy in Type 2 Diabetes: Results of a Large UK Cohort Study

2015 Diabetes UK Professional Conference, London, UK. Intensification of dual therapy with insulin vs DPP-4 inhibitor and the risk of cardiovascular events and deaths in patients with diabetes

Peer-reviewed publications outside thesis

Anyanwagu U, Mamza J, Mehta R, Donnelly R, Idris I. Cardiovascular events and all-cause mortality with Insulin vs Glucagon-like peptide-1 analogue in Type 2 Diabetes. BMJ Heart, 2016 (Accepted: heartjnl-2015-309164)

Wang C, Mamza J, Idris I. Biphasic vs basal bolus insulin regimen in Type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. Diabetic Medicine, 2015. 32(5): p. 585-594

Mungreiphy NK, Mamza J, Lakhdar AF, Bannister M, Elliott J, Idris I. Clinical use and efficacy of biphasic insulin lispro 50/50 in people with insulin treated diabetes - a nationwide evaluation of clinical practice. Curr Med Res Opin. 2015 Feb 9:1-9

Mamza J, Mehta R, Idris I (2014) Obesity independently predicts responders to biphasic insulin 50/50 (Humalog Mix50 and Insuman Comb 50) following conversion from other insulin regimens: a retrospective cohort study. BMJ Open Diabetes Research & Care 2 (doi:10.1136/bmjdrc-2014)

Acknowledgements

My utmost gratitude goes to God for giving me the wisdom, inspiration and sufficient grace to complete this course. All glory and praise to You!

To Iskandar Idris and Richard Donnelly, my supervisors: thank you for your tireless commitment in providing me with the right guidance for my research. You have supported me intellectually by helping me shape my ideas and providing extensive comments on my work. Your career advice and the opportunities you created for me to gain further experience through involvement with other research projects are highly valued.

To Raj Mehta: your statistical advice and encouragements throughout my study period will not be forgotten.

To my wonderful family, friends and all who were there for me: your prayers, encouragements and loving support are deeply felt and treasured.

I would also like to thank the Sherwood Forest Hospital for partly funding my fellowship. I would also like to thank Carol Marlin who extracted and summarised the data I used for meta-analysis.

To all my gospel choir friends, thank you for making my time memorable.

And lastly to my girlfriend Lindsey for her emotional and loving support throughout the PhD process, and for making sure I had Toad in the hole Wednesdays ☺

Contents

List of Figures

List of Tables

List of Appendices

Abbreviations

BNF	British National Formulary
DCCT	Diabetes Control and Complication Trial
DPP-4	Dipeptidyl peptidase-4
TZD	Thiazolidinedione
GFR	Glomerular filtration rate
GLP-1	Glucagon-like peptide-1
HDL-C	High-density lipoprotein cholesterol
LDL-C	Low-density lipoprotein cholesterol
MHRA	Medicines and Health products Regulatory Agency
NICE	National Institute of Clinical Exccellence
OAD	Oral antidiabetic drug
RCT	Randomised Controlled Trials
THIN	The Health Improvement Network
T2DM	Type 2 Diabetes Mellitus
UKPDS	United Kingdom Prospective Diabetes Study

Chapter 1: Introduction

Type 2 diabetes mellitus (T2DM) is a progressive disease which requires continuing medical care to reduce the risk of long-term complications, and combination therapy is widely used in patients with T2DM in order to achieve optimal glucose levels.[1-3] Achieving and maintaining glycaemic control are steered by guidelines which include initial intervention with metformin followed by continuing timely addition of a different class of glucose-lowering agent while monitoring individual patient's risk of hypoglycaemia, weight gain or other complications.[4,5] The assessment of glycaemic response and safety of glucose-lowering agents may best be addressed using well-designed long-term observational studies.[6] However, these studies are flawed with bias and other methodological challenges.[7] Given the lack of evidence for long-term effectiveness of glucose-lowering agents, it is unclear how to optimally manage patients who fail therapy with metformin[8] or predict the risks of other clinical outcomes that may occur from the administration of different combination therapy regimens. Therefore, this study was undertaken to improve the understanding of the effectiveness of dipeptidyl-peptidase-4 (DPP-4) inhibitor regimens and their associated adverse events in the management of T2DM in routine clinical practice.

1.1 The burden of type 2 diabetes

Approximately 3.2 million people are known to be diagnosed with diabetes in the United Kingdom, with Type 2 diabetes accounting for about 90% of all people with diabetes. The prevalence of the disease is currently estimated at 6.2% in England among people aged 17 years and above who are registered with a general practice,[9] and people with T2DM are 32% more likely to die early.[10] There has been a significant increase in diabetes cases among younger people, an age group which have been excluded in most diabetes studies [11].

T2DM cost the National Health Service (NHS) about £10 billion each year, driven by the increased risk of complications such as cardiovascular diseases, retinopathy, kidney failure and lower limb amputation, and is projected to worsen with time. Prevention and early treatment of these complications could limit their impact on an individual and save costs.[12]

In Europe, more than 5% of the total health care cost is spent on morbidity associated with diabetes, the bulk of which relates to treatment of complications occurring from particularly T2DM.[11] The incidence, mortality and cost associated with diabetes around the globe have been increasing at an alarming rate. The International Diabetes Federation (IDF) estimated approximately 382 million people worldwide were afflicted by diabetes in 2013. This is predicted to increase to 592 million by 2035. More than 80% of people diagnosed with diabetes live in low- and middle-income countries of the world and about fifty percent of these individuals go undiagnosed until complications have developed.[13]

1.2 Diagnosis and treatment of type 2 diabetes

1.2.1 How is type 2 diabetes diagnosed?

The World Health Organization (WHO) Consultation concluded that glycated haemoglobin (HbA1c) can be used as a diagnostic test for diabetes, provided that strict quality assurance measures are in place and assays are standardised in accordance with the international reference.[14,15] This program was implemented by the National Glycohemoglobin Standardization Program (NGSP) Steering Committee,[16] and a value of 6.5% is recommended as the cut point for diagnosing diabetes. However, HbA1c < 6.5% does not necessarily exclude diabetes diagnosed using glucose tests. In the UK, consensus statements have suggested patients at most risk of diabetes to have HbA1c of 6-6.4%.[17] However, the American Diabetes Association (ADA) defines a population with pre-diabetes as HbA1c of 5.7-6.4%.[18] The WHO has recommended some diagnostic criteria which most clinicians find helpful (Table 1).[14]

Two fasting plasma glucose (FPG) ≥7mmol/L, a random plasma glucose ≥11.1mmol/L (in a symptomatic patients) and a 2-hour post 75g glucose load ≥11.1mmol/L after an oral glucose tolerance test (OGTT), are considered to be diagnostic of diabetes.[19,20] Individuals with intermediate plasma glucose levels between those considered normal and those considered diabetic are termed to have impaired glucose tolerance (IGT).[21] In situations where there is impaired fasting glucose (IFG) and impaired glucose tolerance (IGT), an OGTT is considered. IFG and IGT are both intermediate forms of hyperglycaemia and

are defined by an FPG level between 6 and 7mmol/L, and plasma glucose between 7.8 and 11.1mmol/L at 2 hours post OGTT respectively.[20]

There have been debates on the reliability of HbA1c as the diagnostic criteria for type 2 diabetes.[20] Although in some laboratories the precise measurement is similar to that of plasma glucose, global consistency with both assays is problematic.[22] Therefore, it is imperative that international standards are met whether plasma glucose or HbA1c assay is used.[14] The International Federation of Clinical Chemists (IFCC) established a working group on HbA1c and introduced an international standardization program with the establishment of reference method procedures for HbA1c in mmol/mol.[23] HbA1c test has been reported as less sensitive to fasting glucose measurements, however this disadvantage may be offset at a population level by conducting multiple testing.[18] The use of HbA1c has proven more reliable for capturing long-term exposure to hyperglycaemia which usually paves the way for complications – an advantage it has over single measures of glucose concentration. It also has the ability to show 8 to 12 weeks of average plasma glucose level and can be measured without fasting anytime.[22,24]

In reality, medical practices make decision as to which criteria to use (e.g. some practices use a combination of a single raised fasting glucose and a single raised HbA1c, taken on the same day). Two random HbA1c values >6.5% (48mmol/mol) has also been accepted to be diagnostic of diabetes in some practices.

Table 1.1 Diagnostic criteria for T2DM

Type 2 Diabetes (units are mmol/L) except specified	
Symptomatic (e.g. polyuria, polydipsia, unexplained weight loss)	**Asymptomatic**
A single fasting plasma glucose ≥7 OR A single random plasma glucose ≥11.1	A fasting glucose ≥7 on two separate occasions OR A random glucose ≥11.1 on two separate occasions OR An HbA1c ≥6.5% (48mmol/mol) on two separate occasions OR HbA1c ≥ 6.5% + an elevated plasma glucose (fasting ≥7 or random ≥11.1)

Impaired Fasting Glucose	**Pre-Diabetes**	**Impaired Glucose Tolerance**
FPG 6.1–6.9 mmol/l (WHO criteria)	HbA1c 6–6.4% (42–47mmol/mol) (NICE) The American Diabetes Association uses wider criteria (HbA1c 5.7–6.4%, 39–47mmol/mol)	Fasting plasma glucose <7.0mmol/l AND 2h plasma glucose (after 75g oral glucose load) 7.8–11mmol/l (WHO criteria) Oral glucose tolerance tests are complex, expensive and less reproducible. Still used in pregnancy (where HbA1c is inaccurate).

Source: GP Update [25]

Within routine therapy, new HbA1c results are found to be lower than unstandardized results, and HbA1c values are now translated into estimated average glucose (eAG) values.[26] Thus the benchmark for diagnosis of 6.5% will be equivalent to 140mg/dl (7.8mmol/L) eAG. In this thesis, reference will be made to the NGSP's unit HbA1c (%) and IFCC's mmol/mol only.

1.2.2 How is type 2 diabetes managed?

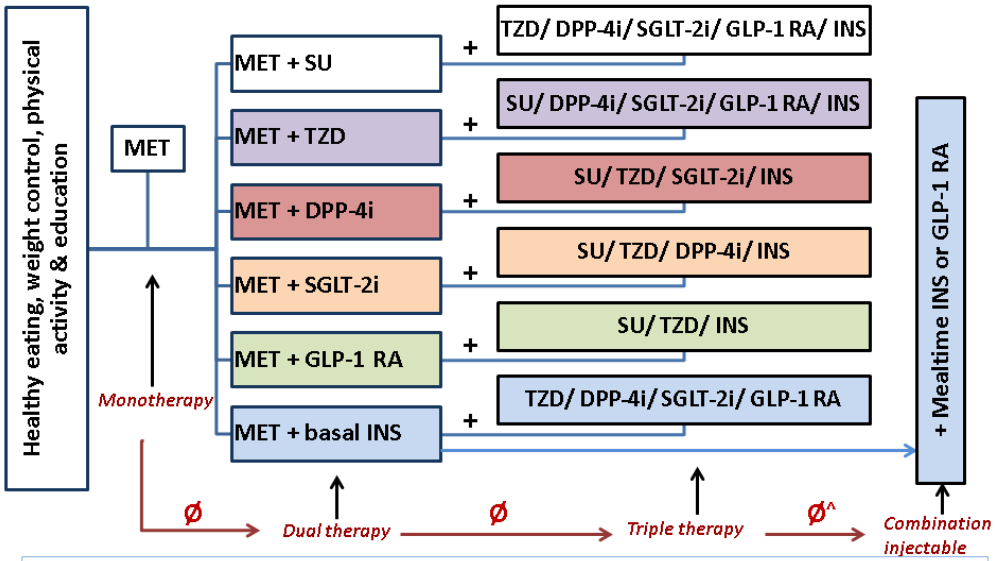
The main goal of diabetes management is to return the HbA1c value to a normal range in order to reduce vascular complications. However, evidence from the ADVANCE (Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation) and ACCORD (Action to Control Cardiovascular Risk in Diabetes) trials give conflicting reports on the best target HbA1c suitable to mitigate adverse outcomes.[27,28]

The initial steps for treating patients with T2DM consist of tailored dietary and physical exercise intervention to improve metabolic control and reduce patient's risk of cardiovascular events.[29,30] Initiating therapy with metformin (MET) when lifestyle intervention fails, or the addition of sulphonylurea (SU) and other classes of glucose-lowering drugs when these drugs fail to maintain metabolic targets have led to the consideration of alternative combination regimens tailored towards individual patient's needs.[5] Studies have shown the benefits of combining oral GLT and insulin to achieve metabolic targets and reduce the risks of hypoglycaemia, weight gain and insulin insensitivity.[31-34] These therapies yield reasonable durable glycaemic control, however they are unable to tackle the progression of diabetes due to the deterioration in b-cells

function, or the patient's eventual need for insulin.[35] Newer therapeutic agents are being used as part of combination therapies to intensify glycaemic control when other glucose-lowering therapies have failed. As newer classes of treatment options have become available, clinicians and patients are faced with a range of medications with different mechanisms of action.

The ADA and the European Association for the Study of Diabetes (EASD) have recently published a position statement on the management of hyperglycaemia in patients with type 2 diabetes.[5] The general recommendations for glucose-lowering therapy in T2DM, as depicted in Figure 1.1 have been summarized in the literature.[5] This was needed because of the increasing uncertainty of the proper treatment selection and sequence for administration. The need to individualize both glucose control targets and treatment strategies has been emphasized with patient-centered care and shared decision making.

Figure 1.1 Recommendations for therapy



Abbreviations: HbA1c (glycated haemoglobin); MET (metformin); SU (sulphonylurea); TZD (thiazolidinedione); DPP-4i (dipeptidyl peptidase-4 inhibitor); SGLT-2i (Sodium-glucose co-transporter-2 inhibitor); GLP-1 RA (glucagon-like peptide-1 receptor agonist); INS (insulin)
Ø – If HbA1c target not achieved after ~3 months
Ø^ – If HbA1c not achieved after 3 months of triple therapy and patient is on: 1) Oral therapy, move to injectable; 2) GLP-1 RA, add basal insulin; 3) Optimally titrated insulin, add GLP-1 RA or mealtime insulin
NOTE: Order of therapy not meant to denote any specific preference – choice dependent on a variety of patient- and disease-specific factors

1.2.3 Dipeptidyl-peptidase-4 (DPP-4) inhibitor therapy

Incretin-based therapies such as DPP-4 inhibitors are novel treatments, compared to the traditional GLTs such as metformin, sulphonylurea, thiazolidinedione and insulin.[36] Incretin hormones are secreted after the ingestion of a meal, and lead to increased glucose-dependent insulin production and secretion.[37] These hormones are secreted into the gastrointestinal tract (GIT) and are cleaved by the proteolytic enzyme, DPP-4 to produce inactive metabolites.[37-39] Glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP), increase insulin production and secretion and GLP-1 decreases glucose production.[38] An insufficiency of GLP-1 has been shown to exist in T2DM while DPP-4 inhibitors competitively inhibit the action of DPP-4.[39] DPP-4 inhibitors prevent

the breakdown of the incretin hormones, thereby increasing levels of GLP-1, GIP and glucose-stimulated insulin release from the β-cells of the pancreas.[36] Available DPP-4 inhibitors include sitagliptin, alogliptin, saxagliptin, vildagliptin and linagliptin.[39]

DPP-4 inhibitors serve as an alternative treatment intensification option to a number of currently available GLTs and the consensus position statement of the ADA and the EASD[5] suggest the use of DPP-4 inhibitors in combination with other oral GLT and insulin. However, the role of DPP-4 inhibitors in the treatment of T2DM is debatable. DPP-4 inhibitors do not cause hypoglycaemia due to their mode of action, they are not associated with weight gain and their insulin-releasing effects are glucose-dependent.[40]

An area of concern with incretin-based therapies has been possible the risk of acute pancreatitis, particularly among individuals with a prior history of history of pancreatitis. However, emerging data from real-world observational studies[41] and RCTs involving DPP-4 inhibitors[42,43] have shown no statistically significant increased rates of pancreatic disease. Furthermore, concerns have also been raised about the increased risk of bone fractures seen with TDZ, predominantly among women.[44] Experimental studies of animal models suggest that the incretin hormone GLP-1 and GIP are capable of increasing bone density.[38,45] This has led researches to investigate the influence of DPP-4 inhibitor therapy on bone fracture in humans. A review and meta-analysis of RCT studies conducted in 2011 suggests that treatment with DPP-4 inhibitors could be associated with a reduced risk of bone fractures.[46]

1.2.4 Intermediate treatment outcomes

Systematic reviews of oral glucose-lowering therapies have highlighted the short-term and long-term clinical outcomes of glucose-lowering therapies.[47-49] These outcomes are important when evaluating and comparing the effectiveness of GLTs. They include changes in HbA1c and lipid levels, blood pressure, body weight, microvascular and macrovascular disease, adverse events and mortality. It is crucial that adverse events are evaluated, as they may affect adherence to therapy as well as morbidity and mortality.

The following texts summarize some of the outcomes of interest in this thesis and introductory evidence-based findings associated with GLTs. Additional evidence of treatment outcome measures are discussed in further detail in the relevant chapters of this thesis.

Glycated haemoglobin (HbA1c)

Results from randomized controlled trials have shown oral diabetes medications had varying effects on HbA1c reductions. The ADOPT (A Diabetes Outcomes Progression Trial) study,[50] for example, reported rosiglitazone, a TZD no longer widely used showed a significant greater reduction in HbA1c to SU (between-group absolute difference of -0.4%) and MET (between-group absolute difference of -0.1%). A recent review of 140 clinical trials and 26 observational studies on head-to-head comparisons of GLTs (MET, second-generation SU, TZD, meglitinides, DPP-4 inhibitors, and glucagon-like peptide-1 (GLP-1) receptor agonists) as monotherapy and combination therapy[51] reported most medications decreased the HbA1c level

by about 1% (absolute reduction). Bennett et al reported MET as the most efficacious among the monotherapies. However, combination therapies had additive effects and were better at reducing HbA1c compared with monotherapy regimens.[51] Examples of comparative effectiveness of GLTs from different studies are illustrated in as forest plots in another study, showing the relative differences in HbA1c between different combination GLTs.[6]

Body Weight

Placebo-controlled trials have shown oral GLTs such as TZD, SU and repaglinide increased body weight by 1 to 5kg, whereas, MET was steadily associated with neutrality or weight loss compared with most other GLTs.[52] In a more recent study of direct comparisons of monotherapies with TZD and second-generation sulphonylurea, MET was associated with weight loss. SU and meglitinides increased weight in a similar fashion, SU increased weight less than TZD, and GLP-1 agonist decreased weight compared with SU. Combinations of MET plus a TZD or MET plus SU increased weight more than MET monotherapy. The co-administration of a DPP-4 inhibitor to MET had similar effect on weight compared with MET monotherapy, although the strength of evidence was low due to fewer studies on DPP-4 inhibitors and other standard combinations.[51]

Serum lipid levels

Oral glucose-lowering therapies have been reported to have small to moderate effects of on lipid levels, which varied across medication type.[6] Overall, MET

had favourable effects on all the lipid classes compared with TZD, SU, and DPP-4 inhibitors, for example, MET consistently decreased low-density lipoprotein (LDL-C) and triglyceride levels and modestly increased high-density lipoproteins (HDL-C) levels. However, TZD increased LDL-C levels, while SU and DPP-4 inhibitors had little effect on them. In addition, combination therapy involving MET plus TZD increased LDL-Clevels compared with MET monotherapy, or MET plus SU.[51]

Pioglitazone, a TZD monotherapy increased HDL-C levels more than MET or SU. Compared with MET monotherapy, MET plus pioglitazone increased HDL-C levels, whereas, similar levels were reported with MET plus DPP-4 inhibitors. Metformin plus TZD increased HDL-C levels more than the combination of MET and SU.[51]

TZD decreased triglyceride levels compared with MET monotherapy, whereas MET decreased triglyceride levels compared with SU. MET monotherapy decreased triglyceride levels compared with the combination of MET and TZD, while MET plus TZD decreased triglyceride levels compared with the combination of MET plus SU.[51]

1.2.5 Distal diabetes-related complications

Cardiovascular events and all-cause mortality

Patients with T2DM are associated with two to fourfold increased risk of developing cardiovascular (CV) diseases compared to non-diabetics, and CV diseases are a major cause of morbidity in these patients.[53] Intense glycaemic control has been shown to improve diabetes-related microvascular

complications but has not been able to demonstrate major effects of macrovascular complications such as myocardial infarction and stroke.[48,54] The strength of evidence available to support comparative effectiveness findings from diabetes medications on all-cause mortality, CV disease and mortality is insufficient.[6] And it is unclear whether effects on mortality differed between different combination therapies, due to poor adjustment for key confounders in cohort studies and lack of long-term studies or studies assessing multiple medication use.[51]

The risks of CV diseases associated with GLTs are currently being evaluated by researchers. Previous smaller sized trials have demonstrated that MET and insulin (INS) are not associated with increased CV risks, while the SUs remain controversial.[53] The ADOPT study which involved 4,360 patients followed for a median of 4 years reported similar rates of all-cause mortality, CV disease mortality and morbidity, and stroke between MET, SU or rosiglitazone monotherapy.[50] However, conflicting results were recorded in observational studies compared with the trial data, for example, MET was associated with a lower risk for all-cause mortality and CV events than SU.[6] A meta-analysis of five prospective RCTs comprising 33,040 participants assessed the effect of intensive glucose-lowering regimen on death and cardiovascular outcomes compared with standard regimen.[55] The data showed intensive glycaemic control resulted in a 17% reduction in non-fatal myocardial infarction (OR, 0·83; 95%CI, 0·75–0·93), 15% reduction in coronary heart disease events (0·85, 0·77–0·93) and no significant effect on stroke events (0·93, 0·81–1·06) or all-cause mortality (1·02, 0·87–1·19).[55] This implies intensive compared

with standard glycaemic control significantly reduces CV events without increasing the risk of death. However, the extent of HbA1c reduction might differ across populations.

RCT studies have been conducted on the CV morbidity and mortality outcome of a newer glucose-lowering agent – empagliflozin (a SGLT-2 inhibitor).[56] The study was conducted among 7,020 T2DM patients considered to be at high CV risk and the results showed patients who received empagliflozin, as compared with placebo, had a lower rate of the primary composite outcome and CV events and all-cause death.[56] In terms of combination therapy, shorter duration RCTs have reported lower risk of CV outcome with MET than with the combination of MET and rosiglitazone, although the confidence intervals (CI) of the pooled odds ratio overlapped, 0.43 (95% CI, 0.17-1.10).[51]

DPP-4 inhibitors and CVD outcome trials

With the increasing number of newly licenced glucose-lowering agents used for treatment of T2DM, questions have been raised regarding the long-term cardiovascular safety of some of these agents.[57,58] In response, international regulatory agencies now require new antidiabetic agents to not only show glucose-lowering ability but also require the agents are not associated with clinically meaningful increases in rates of major adverse cardiovascular events (MACE).[59,60] DPP-4 inhibitors have been widely investigated for their CV safety in large-scale outcomes trial in patients with T2DM who are at risk of, or with established coronary disease.[53]

TECOS

Results from the Trial Evaluating Cardiovascular Outcomes with Sitagliptin (TECOS) study was presented at the ADA 2015 scientific meeting and published. TECOS was a randomized double-blind study of 14,671 patients assigned either sitagliptin or placebo as add-on to their existing therapy, and was designed to determine the CV safety of long-term sitagliptin use in T2DM patients with established CV disease.[61] The primary CV outcome after a median follow-up of 3 years was a composite of CV death, non-fatal MI, non-fatal stroke, or hospitalization for unstable angina. The results showed sitagliptin was non-inferior to placebo for the primary composite CV outcome (HR 0.98; 95% CI, 0.88-1.09; $P < 0.001$), while rates of hospitalization for heart failure (HF) did not differ between the two groups (HR 1.00; 95% CI, 0.83–1.20; $P = 0.98$).

SAVOR-TIMI

Another study investigated the safety and efficacy of DPP-4 inhibitor, saxagliptin used in the management of T2DM. The Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus (SAVOR)–Thrombolysis in Myocardial Infarction (TIMI) 53 study was a double-blind, placebo-controlled trial examining the CV effect of saxagliptin vs placebo in 16,492 patients with T2DM.[42] Participants were followed up for a median 2.1 years to a primary composite endpoint of CV death, non-fatal MI, or non-fatal stroke. Results from SAVOR-TIMI 53 showed saxagliptin was non-inferior to placebo for the primary composite endpoint (HR 1.00, 95 % CI 0.89–1.12,

$p < 0.001$), although saxagliptin use was associated with an increase in secondary outcome of hospitalization for heart failure compared to placebo (HR 1.27, 95 % CI 1.07–1.51, $p = 0.007$).[42]

EXAMINE

The Examination of Cardiovascular Outcomes with Alogliptin versus Standard of Care in Patients with T2DM and Acute Coronary Syndrome (EXAMINE) reported a different outcome on heart failure.[43] The EXAMINE study evaluated the effects of DPP-4 inhibitor, alogliptin in 5,380 high-risk patients reported no increase in the risk of hospitalization with heart failure.

Therefore, with the exception of increased risk of HF observed in SAVOR, the use of DPP-4 inhibitor has not shown any increase in CV risks in large scale trials.

Furthermore, results of a pooled analysis of DPP-4 inhibitor data did not show any increase in MACE compared to placebo or SU.[62] And a retrospective cohort study in a Danish population[63] showed no significant difference between DPP-4 inhibitor and MET use on all-cause mortality (Hazard Ratio, HR 1.25, 95 % CI 0.92–1.71, $P = 0.153$) or the composite event of stroke, MI, and all-cause mortality (HR 1.22, 95 % CI 0.92–1.61, $P = 0.164$).

Nonetheless, there is limited information on head-to-head comparison of CV safety associated with the use of DPP-4 inhibitors vs other GLTs, when used as a second- or third-line intensification regimens.

Bone fractures

Increased interest has emerged regarding the risk of bone fracture from different GLTs. People with diabetes are more likely to experience an osteoporotic fracture, compared to the rest of the population.[64] There is a greater chance that hypoinsulinaemia and hyperglycaemia in diabetes (type 1 and type 2, respectively) alter bone tissue composition and decrease bone production, leading to the weakening of bone.[65] Long standing hyperglycaemia caused by insulin resistance in T2DM greatly affects multiple tissues and the frequency of complications in this group of patients.[65] Therefore, optimal glycaemic control may potentially have a positive effect on bone, thereby, protecting against bone fracture. This action could be as a result of the increase in circulating levels of GLP-1 and gastric intestinal polypeptide, which are both involved in the regulation of bone metabolism.[38,66,67] DPP-4 inhibitors may in fact be protective to bone, reducing the rate of fractures.[64] A review of RCTs reported that DPP-4 inhibitors, when compared with placebo or comparator treatments, were associated with fewer fracture events.[46] However, other studies have showed that TZD, either as monotherapy or in combination with another GLT was associated with a higher risk of bone fractures than was MET alone or combined with SU.[51] Two large clinical trial studies ADOPT[50] and RECORD (rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes),[68] reported elevated fracture risk among women receiving the no longer used TZD, rosiglitazone compared with regimens containing SU or MET. Hypoglycaemia induced by some GLTs may increase the risk of falls and potentially lead to higher fracture risk,

especially in older patients.[69] Most clinical trials were not designed to evaluate the risk of SU's, for example, on fractures or falls and clinical trials reporting fracture as a primary endpoint are scarce.[70] However, there is the potential for insulin and SU induced hypoglycaemia surrogate bone fractures occurring as a result of falls. In addition, combination GLTs had varying associated risks for fractures, which may affect the choice of a second-, or third-line agents. Hence, further research to assess fracture risk in T2DM patients will be useful.

1.3 Aims and objectives

Clinical outcomes that can be assessed when investigating the comparative effectiveness and safety of GLTs used in the management of T2DM have been described. It is very important that continuous monitoring of clinical and metabolic parameters is undertaken to enable quality investigations into the treatment outcomes in patients with T2DM. There is limited information on head-to-head efficacy and safety comparison studies on the use of DPP-4 inhibitors vs other GLTs, when administered as intensification regimen within routine clinical practice. The growing number of observational studies investigating the outcomes of treatment, however, suggests that a more robust approach is needed to addressing methodological challenges such as bias and confounding. This thesis proposes that understanding the different types of biases and strategies that can be employed to minimise or control biases can strengthen the evidence from comparative effectiveness studies. Investigations are carried out on the effectiveness and safety of DPP-4 inhibitors when added as intensification treatment option following failure of other oral glucose-lowering therapies to achieve glycaemic control. The objectives of the thesis

are achieved through literature review of epidemiological studies, meta-analysis of RCTs and statistical analyses of a large UK general practice (GP) primary care database. The patient population chosen for analysis include patients with T2DM who are 18 years and older. The objectives of each research question answered were pre specified based on existing evidence from routine clinical practice and discussions held with a consultant diabetologist.

Results generated from this thesis should be interpreted with caution and not be seen as definitive in clinical practice. This thesis evaluates both short-term and long-term outcomes of DPP-4 inhibitors use in the management of T2DM, therefore, it will strengthen evidence-based treatment outcomes in the field. Furthermore, it provides policymakers and researchers with new insight into the comparative effectiveness and safety of oral glucose-lowering medications used in the treatment of T2DM.

The aim of this thesis was attained by addressing the following objectives:

1. To review evidence of biases associated with retrospective cohort analyses of administrative health records of diabetes patients
2. To identify clinical and metabolic parameters that predict glycaemic response or non-response to DPP-4 inhibitors
3. To determine the glycaemic effectiveness of intensifying ongoing GLTs with DPP-4 inhibitors

4. To assess the durability of DPP-4 inhibitor as 2nd line GLT following MET therapy failure

5. To investigate the CV disease and mortality risks associated with dual therapy intensification

6. To review the evidence of bone fractures associated with DPP-4 inhibitors

1.4 Thesis overview

Chapter 2 examines the literature for biases that may be associated with retrospective cohort studies of glucose-lowering therapies from healthcare databases. Here, a systematic review of observational studies is conducted and examples from the relevant literature are identified to describe the biases and strategies for minimising them. It is hypothesised that not many studies report biases as methodological challenge or highlight methods that can be employed to address them in the literature. An algorithm was outlined to illustrate approaches that can be adopted to reduce these biases.

Chapter 3 describes the applications of epidemiological and statistical techniques in the analysis of observational data. Database management protocols for extracting and organizing GP primary care data using the Health Improvement Network (THIN) database were developed. In addition, description of the advanced statistical and epidemiological techniques such as logistic regression, Cox proportional hazard regression, propensity score

analyses and sensitivity analyses, which are employed in this thesis are provided.

In Chapter 4 it is assumed that apart from baseline HbA1c, little is known about clinical parameters that determine the glycaemic response to a DPP-4 inhibitor when it is co-administered with other GLTs in routine clinical practice. Research is conducted to identify independent predictors of response and non-response among T2DM patients receiving DPP-4 inhibitors as add-on therapy.

In Chapter 5 the glycaemic effectiveness and body weight responses of co-administering DPP-4 inhibitor to patients with inadequate glycaemic control following treatment with metformin (MET), sulphonylurea (SU) or MET plus SU regimens is assessed. Investigation is conducted on HbA1c changes within subgroups of different categories of baseline HbA1c levels to determine whether DPP-4 inhibitors produce additional glucose-lowering effect in one treatment group compared to another.

In Chapter 6 a 5-year follow-up data is examined to determine the glycaemic durability of DPP-4 inhibitor compared with SU and TZD, when used as 2^{nd} line treatment options following MET monotherapy failure to achieve glucose control. In addition, the risk of treatment failure in patients who had their MET therapy intensified with the aforementioned 2^{nd} line GLT is assessed. Analysis is conducted to determine clinical and demographic factors that may be associated with increased risk of treatment failure, and the rate of HbA1c goal attainment of these agents.

Chapter 7 investigates the concerns large observational studies have raised about the CV safety and mortality risks of treatment intensification with glucose-lowering agents. It argues that attention should not only be drawn to the risks associated with monotherapy and dual therapy alone, but also to the risk associated with intensification with a 3rd line treatment option as commonly experienced in practice. A 5-year routine clinical practice data is examined to compare the time to a composite of non-fatal acute myocardial infarction (AMI), non-fatal stroke or all-cause mortality in patients who had their diabetes treatment intensified with a DPP-4 inhibitor vs. insulin (INS) following dual therapy failure with MET plus SU.

In Chapter 8 a systematic review is conducted based on the evidence that DPP-4 inhibitor use may be protective to bone, thereby reducing fracture incidence. In this chapter, a meta-analysis of randomized controlled trial (RCT) studies is conducted to further assess the safety of DPP-4 inhibitors. Updates to the previous review of evidence are provided to generate more insight on the role of DPP-4 inhibitor in fracture incidence among patients with T2DM.

Chapter 9 draws the research projects together and relates the preceding chapters to the aim and objectives presented in Chapter 1. It summarises the findings of the thesis and outlines the importance of the research and recommends opportunities for future research.

Chapter 2: Systematic review of biases in retrospective cohort studies

2.1 Summary

Objective

To identify and provide a simple description of the types of biases and control measures employed in observational studies of treatment outcomes from GLTs used in type 2 diabetes.

Method

A series of searches were performed on comprehensive electronic databases using search terms that included synonyms of 'glucose-lowering medications', 'bias', 'observational studies' and 'diabetes mellitus'. The data sources which include EMBASE, MEDLINE, PubMed and SCOPUS were searched for retrospective cohort studies investigating treatment outcomes of GLTs. The types of biases described were based on examples identified from relevant literature. A number of approaches were suggested to help address these challenges.

Results

658 relevant publications were found through electronic database search. Overall, there was a low rate of bias reporting as only 27 studies met the inclusion

criteria for reporting a bias that may be associated with retrospective cohort study of administrative databases and involved GLTs. Four of the 27 reviewed articles were designed to demonstrate how selection (n=3) and measurement (n=1) biases could be minimised. In general, biases identified and the numbers of studies that reported them include selection (n=11) and misclassification (n=6) biases. Others are prevalent user, immortal time, measurement, start-time, prescription, surveillance, reverse causation, allocation, and missing data biases. Strategies employed to minimise bias in include new-user design, stratified analyses, rigorous selection criteria, regression, propensity score analyses, sensitivity and fixed cohort analyses. An algorithm was used to illustrate approaches that can be taken to minimise bias.

Conclusion

There is a growing number of real-world studies. However, the rate at which biases commonly associated with retrospective cohort studies are reported is very low. These biases have been described to increase their awareness and strategies to minimise them.

2.2 Introduction

Healthcare databases constitute an important tool in pharmacoepidemiology and researchers are increasingly using them to assess the relationships between drug exposures and health outcomes in large populations.[71] Researchers find these databases attractive because they reflect routine clinical practice, they are sufficiently large to allow the study of rare events, and are readily available for analysis at limited cost and within a reasonable time compared with large RCTs.[7] In addition, database use complements safety information often associated with the strict inclusion criteria and limited sample sizes of clinical trial studies.

However, if observational studies from databases are not rigorously conducted, they will be prone to methodological problems that may compromise their validity. Many database studies suffer from major biases leading to misleading findings, especially when the biases are not properly understood and addressed.[72] The importance of observational studies in understanding the risks and benefits of drug treatment cannot be overemphasised;[73] therefore, understanding potential biases associated with databases of longitudinal records and the appropriate methods for addressing them are essential in the study design and analysis of these data.[72]

The number of observational studies which assess the effect of various glucose-lowering regimens in patients with diabetes is on the increase. However, time-related biases have not been extensively described in studies involving the treatment of diabetes. And if these biases are not adequately

controlled, the effect of the treatments under investigation may be overestimated thereby making a drug appear protective when in reality may have no effect or may worsened health outcome.[74] Given the frequency with which biases are encountered in observational studies, this Chapter was aimed at assessing the types of biases reported in retrospective cohort studies of treatment outcomes of antidiabetic medications from healthcare records. The primary objective was to estimate the frequency of bias reporting in the literature and describe the types of biases identified in order to increase awareness. In addition, suggestions were provided for strategies that can be employed to minimise or control the influence of these biases using examples from the relevant literature.

2.3 Methods

2.3.1 Information source and search strategy

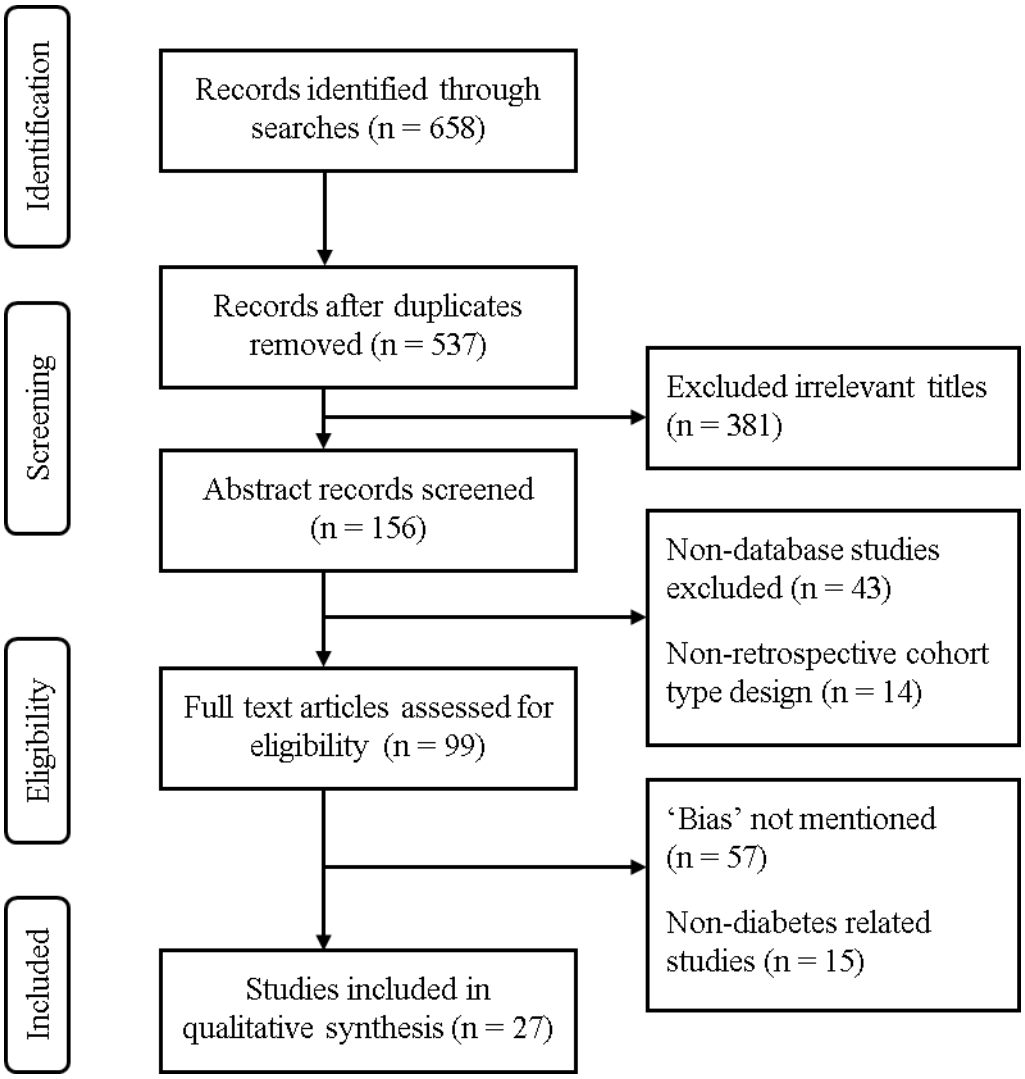
Systematic searches were carried out to identify observational studies which investigated treatment outcomes of GLTs from healthcare records. The following electronic databases were searched up to January, 2015; MEDLINE (from 1996), PubMed, EMBASE (from 1980), and SCOPUS. Keywords used in the search included synonyms of 'glucose-lowering medications', 'bias', 'observational studies' and 'diabetes mellitus'. Further search was also conducted on the reference lists of some retrieved articles. The search was restricted to retrospective cohort studies from administrative databases or healthcare records of patients with diabetes. Studies were restricted to English language reports in adults.

2.3.2 Study selection

The studies were assessed first by the titles, then by the abstract and followed by full text review. Titles that were included mentioned synonyms for any glucose-lowering medication, diabetes and treatment outcomes. Titles contradicting the search terms, by mentioning study designs not considered, or population groups that were not T2DM were excluded. After collecting the articles, the abstracts were assessed according to eligibility criteria. The aim of the search was to identify retrospective cohort studies that investigated the effectiveness and treatment outcomes of GLTs using healthcare records of patients with T2DM. The number of these potentially relevant studies was noted and the studies were manually searched for the term 'bias' to identify if bias section was included or reference was made to bias in the methods or limitation of the studies. If this was not present but the study met other criteria listed above, then the study would remain in the group of potentially relevant studies. Non-English articles were not included. Only human studies conducted in adult population were used, as these were the most relevant to application of results to patients with T2DM. There was no restriction on the publication date of articles, as all studies up until January 2015 were included.

The numbers of papers included or excluded at each stage of the selection process are outlined in Figure 2.1.

Figure 2.1 Study selection flow chart



2.3.3 Data abstraction and quality assessment

A standard form was used to abstract data on general study characteristics, for example, data source, study purpose, study participants (population, age, sex and race), study outcome, biases reported and bias control methods employed. These parameters were collated and tabulated in Appendix A-2. The studies included in the review were assessed for their methodological rigour guided by recommendations for appraising the quality of retrospective cohort studies Appendix A-1.[75] However, the quality of one eligible study could not be

compared against another because the studies differed on their settings, population, exposures and outcomes. The types of biases reported in each study were assessed in order to gain more understanding and provide a description of steps that have been taken to address them.

In the absence of an a priori classification structure for the different biases associated with retrospective cohort studies, narrative summaries of biases and figurative illustrations on how bias influence a study were presented. In addition, suggestions were made on strategies that can be employed in minimising or controlling for the biases with the aid of an algorithm.

2.4 Results

From the primary search output of 658 articles, 84 observational studies were identified to be potentially eligible retrospective cohort studies set out in the criteria for inclusion. From the eighty-four studies, 27 studies were identified as relevant and included in the review as outlined in Figure 2.1.

2.4.1 General characteristics of the studies

The characteristics of the 27 (32%) eligible studies selected for evaluation are summarized in Appendix A-2. Four of the 27 reviewed studies were specifically designed with the aim of addressing bias in retrospective cohort studies.[76-79] The number (n) of studies and types of biases reported in the literature include; selection bias[76-78,80-91] (n=15), misclassification bias[84,87,92-95] (n=6), healthy (prevalent) user bias,[84] immortal time,[96] measurement bias,[79] start-time bias,[97] prescription bias,[98] surveillance bias,[99] detection bias,[100] indication bias,[92,101] reverse causation bias,[102] informative censoring,[103,104]

allocation bias,[102] lead time[94] and lag time bias.[104] A description of these biases, how they influence retrospective cohort studies, and suggested strategies for minimising them are provided below. A brief definition of some of the biases is outlined in Appendix A.

2.4.2 Major methodological challenges

Selection bias

Selection bias is the systematic difference that occurs when individuals or groups of patients are selected for analysis where proper randomisation is not achieved. Healthcare data of treatment exposure groups may be subject to selection bias due to unmeasured patient characteristics which are unavailable in the database. Selection bias may also be introduced when participants' selection criteria results in a large proportion of patients being excluded from the analysis, for example, when there's a transient interruption in patients' therapy as a result of a CV event.[83] To ensure a large proportion of patients are not lost from such cohort, a minimum inclusion criteria that will allow the capture of the truly exposed patients can help reduce selection bias.[83]

To minimize selection bias in observational studies of healthcare records, researchers have used statistical regression models to appropriately adjust for confounding variables.[80] A confounder is a factor that is associated with both the outcome and exposure or treatment, but is not an intermediate variable. However, in situations where there is incomplete or missing data on patient information such as treatment histories, length of time since diagnoses of diabetes, duration of prior GLT, and other important patient parameters,

selection bias may affect the interpretation of the results even when regression models are employed.[80]

One approach to reduce or eliminate the effect of treatment selection bias and confounding effects is the use of propensity score analytical technique, which allows a researcher to analyse an observational (non-randomized) study so that it mimics some of the characteristics of a randomised controlled trial.[105] Propensity scores were employed to reduce any potential selection bias in five studies.[78,81,86,88,89] Propensity scores for initiating a treatment versus comparison drug is calculated from a logistic regression model that estimates the likelihood of initiating that treatment based on the observed patient characteristics. Covariates are selected based on their hypothesised confounding relationship with the outcome variables.

Sensitivity analysis preceded propensity score analysis in three studies, [88] or was conducted separately to confirm the consistency of the findings. In one study, sensitivity analysis was conducted by excluding an affected patient group and thereafter examining to see if the results differ from that of primary analysis.[106]

Misclassification bias

Actual adherence to glucose-lowering therapies cannot be confirmed from databases; therefore database studies rely on prescription information from general practitioners or dose estimates from pharmacy refill data as a proxy for medication taking, which may result in exposure misclassification. In addition, certain over-the-counter medications such as lipid-lowering drugs may have

been used by those classified as not using a lipid-lowering drug, thereby resulting in misclassification bias. Misclassification may also be introduced where there is no medical record to confirm an outcome event, e.g., AMI, stoke or cause of death. A meta-analysis of studies comparing the cause of death noted on a death certificate with the cause of death found on autopsy reported that at least one-third of death certificates are likely to be incorrect.[107] Misclassification bias from treatment crossover may occur when individuals are switched from one therapy to the other.[95,99] Most databases are dependent upon physician entry into the computerised medical record systems and patient adherence to prescribed medication. Hence, some reassurance may exist from the fact that many of the diagnoses, exposure, and outcome defining variables would have been validated in previous studies, minimising the possibility of misclassification bias.[94] Records in some databases may be more robust particularly when patient or prescription records are updated directly as issued by a general practitioner or electronic prescription records.[94]

Approaches to avoid misclassification bias may include the use of clearly defined exposure and outcome variables. For example, by segregating and restricting participants into treatment groups of patients taking specific treatments.[95] In addition, strict inclusion and exclusion criteria may be applied. For example, patients with 'lesser' amounts of treatment exposure may be excluded, participants may be segregated to specific treatment groups, participants who develop certain events or whose record terminated before an index or event date could be excluded. Another approach may be to utilise a

window for exposure classification to allow for prescription and diagnosis entry.[94]

With stringent inclusion criteria, differential misclassification of a study outcome with respect to drug comparison groups is unlikely to occur.[92] Therefore, the effect of non-differential outcome misclassification may depend on the sensitivity and specificity of the outcome assessment. Sensitivity analysis could be conducted to investigate bias toward the null arising from non-specific outcome assessment.[92] Sensitivity analysis may be conducted to examine the risk of developing a disease outcome within different exposure periods.

Measurement bias

Measurement bias may be introduced when medical information influences decision for recording laboratory information or therapeutic data. For example, a bolded warning on a drug may have a potential to introduce measurement bias by making the drug monitoring more likely in affected patients.[108] It may also occur if the ascertainment of a disease onset relies on patient self-report, or if assigned treatments and the potential for adverse effects are known to the prescriber and patients, in which case, diagnostic tests may become associated with an ongoing treatment, and thus bias the detection of a disease outcome.[79] There is also the possibility that a clinician may less likely prescribe an exposure drug to patients who may have been more prone to developing a complication.[109] Overcoming this type of bias is difficult in healthcare records especially when datasets are not able to provide detailed reports of circumstances surrounding disease diagnoses. Furthermore, the drug indication

itself might influence diagnosis, for example, when there are concerns regarding the presence of metabolic syndrome in patients treated with antihypertensive drugs.[79] In situations where measurement bias is suspected, particularly with poor quality healthcare records, subgroup and sensitivity analyses may be carried out to investigate if results differ across pre-defined subgroups.

While measurement bias might explain some inconsistencies, it might be of different magnitude across clinical settings (depending on diabetes screening practices or protocols) and the validity of database laboratory results.[79] Measurement bias can be reduced by exploring endpoints used for recordings. Similar time periods of observation in the database for treatment and comparisons groups can be used to ensure participants have comparable medical surveillance for ascertainment and treatment.[79,110] In addition, where detailed information on diabetes testing is available, unbiased estimates can be obtained by using propensity scores.[79]

Figure 2.2 Lead time bias

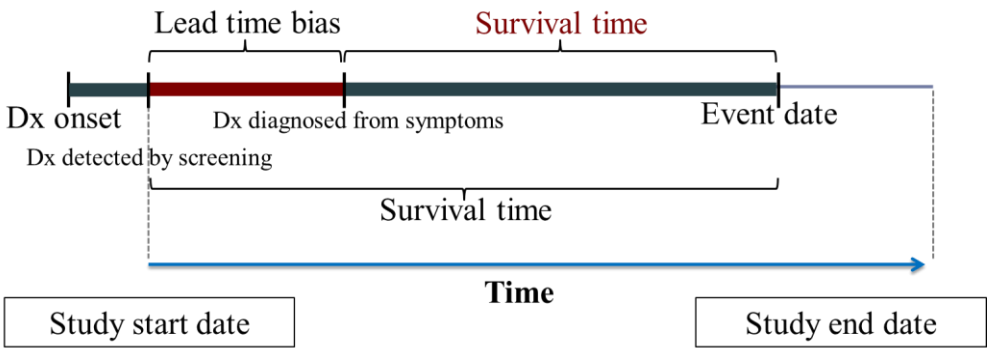


Figure 2.2 Abbreviation: Dx (disease)

Lead time, lag time and reverse causation bias

The date of diagnosis of a disease (Dx) e.g. cancer, may vary based on many factors. For example, a patient may be diagnosed after a screening and before any presentation of symptoms. In which case, the date of diagnosis is likely to be earlier than if screening was not done (Figure 2.2). If a patient commences cancer treatment and is diagnosed only after symptoms develop, his survival time might be less than the patient who was screened before treatment. A longer survival time implies that the duration of time the patients is known to have the disease (lead time) is longer. The treatment(s) the patient might receive may not change his total survival. Lead time bias may prompt patients to undergo cancer screening in order to increase their survival time. In an attempt to minimise potential lead time bias from incidentally discovered cancer, Hwang et al,[94] excluded a small number of patients who may have had an early stage disease in order to make the study population homogenous, thereby involving only a cohort with higher stage disease.

In terms of cardiovascular diseases, glucose-lowering therapies may have an influence on cardiovascular disease outcomes through different biological mechanisms of action that can result in increased or decreased risk of cardiovascular events. The time window chosen for identifying an outcome should reflect the period of disease onset. In the event a minimum time is required before a particular cardiovascular outcome is established, a reasonable approach will be to consider a lag time between the exposure and start of follow-up. However, this should be done carefully and tailored to individual studies. Employing the lag time approach reduces the possibility of another methodological challenge known as reverse causation or protopathic bias, where treatment of symptom appears to cause a disease outcome. This will occur when a patient's treatment plan is selected based on some preclinical conditions. For example, early symptoms of heart failure may influence the choice or outcome of the next line of GLT, hence, a medication may appear to cause cardiovascular outcomes because it was prescribed for those with increased cardiovascular disease risk. Therefore, treatment will increase risk because of prescribing pattern rather than the harmful effect of the treatment.

Another challenge to consider when assessing the effects of GLT on cardiovascular diseases is the appropriate duration of the exposure risk window. There's a possibility that patients may stop or change therapy shortly before the occurrence of a cardiovascular event. This often results in the exposure risk window being extended to accommodate an exposure lag or latency period. The latency period after a treatment is discontinued refers to a time period during which a specific outcome can still be attributed to the

discontinued medication. This time period should be included in the follow-up time. Two of the reviewed studies reported how latency period after drug discontinuation was considered in a bid to minimise bias.[92,104]

Sensitivity analyses can be conducted to assess the different exposure risk window definitions.[92,104]. When assessing the clinical effects of GLTs, it is suitable to clearly identify the biological hypothesis to test and to choose the appropriate exposure risk window in line with this, considering lag or latency periods and sensitivity analyses to address any uncertainties.

Due to the progressive nature of diabetes, treatment discontinuation, switching or intensification is often encountered during the course of therapy. Treatment may change in response to advancing diabetes, or as a result of adverse effects associated with specific agents. These situations could potentially lead to an increased risk of cardiovascular events shortly after a regimen is altered or discontinued, or lack of apparent effectiveness in glucose control, which may result to bias. A common approach for assessing drug safety in observational study setting is by terminating the drug exposure once a medication is discontinued, a method referred to as-treated (AT) analysis.[91] However, this may be prone to bias (informative censoring) if stopping a treatment could determine future cardiovascular outcomes. Therefore, censoring after terminating a treatment in this setting eliminates potential outcomes of an exposure category. Assessing the preclinical manifestations of cardiovascular events shortly after discontinuation, for example may also help to detect if informative censoring is present.[103] Sensitivity analyses can also be carried out to assess potential informative censoring.[91,92,103,104] In one of the reviewed

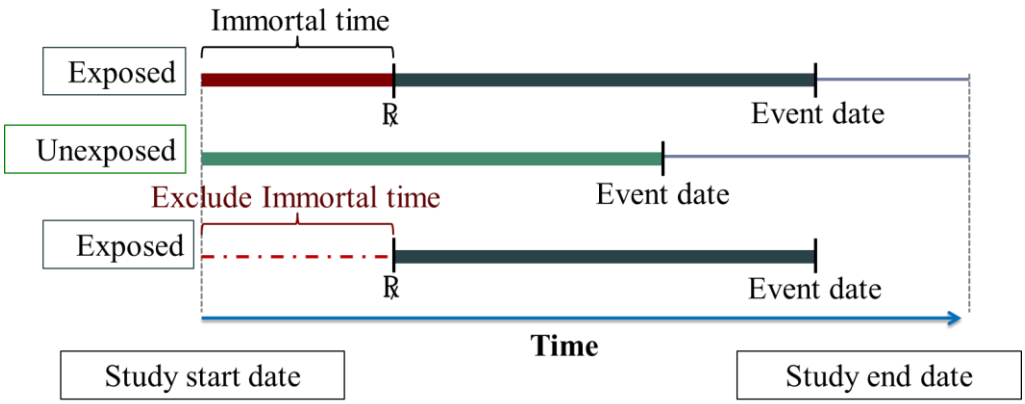
studies, for example, patients' exposure status was to either insulin glargine or none over a 4 month period, and participants were then followed-up without regard to any subsequent change in exposure status.[102] By overlooking the transition to other exposures, this can minimise possible reverse causation bias in some circumstances but can bias towards the null.[102] A few studies accounted for time-varying exposures by using Cox proportional models and analysing exposure to drugs as a time dependent variable.[102,104] Such models make the assumption that treatment changes are independent of CV outcomes and may result in biased results in the presence of patient characteristics that vary over time, affecting both diabetes treatment choice and cardiovascular risk.

Intention-to-treat (ITT) analysis may be employed as it is not affected by informative censoring bias in the same way as as-treated analysis. The ITT method carries forward the initial exposure status and ignores changes in treatment status over time. However, this approach may be prone to exposure misclassification bias, which may be increased or decreased depending on the length of the follow-up period before treatment termination.[7] Techniques that have been employed to minimise potential bias that may result from the changes of glucose-lowering drugs during treatment include, restricting or segregating the cohort into treatment groups of patients taking the same treatments to allow study participants to share similar characteristics,[95] or categorising the periods of exposure.[92] Time-dependent confounders can be accounted for in Cox proportional hazard models, however other analytical techniques that require more extensive programming, such as marginal

structural models can also be used to address this problem. This methods are based on the assumption of no unmeasured confounding, hence all important covariates that predict treatment change should be identified and be available in the data source.

Employing the intention-to-treat analysis ensures a fixed cohort is maintained such that, patients are categorised on the basis of their exposure across their entire follow-up period.[86,102] Colhoun et al., for example segregated patients on insulin glargine only from those who never received insulin glargine concomitantly with any other type of insulin, those on non-glargine insulin only and never had any insulin glargine at any time during follow-up, and those on non-glargine plus glargine insulin who were using insulin glargine concomitantly with another type of insulin. This analytical approach uses the data available more completely and defines actual exposure more accurately, but at the cost of being more prone to reverse causation bias when there is no clear or complete account of participants' transition from one treatment exposure to another. Cox proportional hazards models can be used to examine the primary hypothesis of whether the incidence of a disease outcome varies by exposure to a specific treatment.[102]

Figure 2.3 Immortal time bias



Immortal time

Immortal time as illustrated in

Figure 2.3 is the period of observation time during which the event cannot occur, i.e. person-time that is event-free.[111] This occurs, for example, when the inclusion criteria requires all exposed participants to have survived for a specific period before an event, and the event-free person-time is incorrectly included in the exposed. It is the type of bias that is introduced when exposure status depends on information that is unavailable at the time of cohort entry, and later becomes available during study follow-up.[82,83,90,95] This can bias the results by reducing the estimated incidence rate of cardiovascular events, for example, by diluting the person-time of the treated with some other person-time that has no risk for study outcomes. Identifying the potential influence of immortal time bias in the interpretation of results can be challenging.

Therefore, examples on how immortal time bias may influence observational studies will help researchers in addressing them.[7]

In one of the reviewed studies, exposure status was defined on the basis of the patient being administered routine GLT at the time of diagnosis of the outcome.[96] However, because the outcome involved death as well as other competing events which occurred after the initial diagnosis, the continued use of GLT could introduce immortal time bias. To address the possible role of immortal time bias in the observed association of metformin, Wu et al.,[96] performed multivariate regression analyses of the survival data based on the Cox proportional hazards modelling (as proposed by Levesque et al)[73]. After which they employed landmark analyses at different time intervals to confirm the beneficial association of metformin use with overall survival of patients was unlikely to be as a result of immortal time bias.[96]

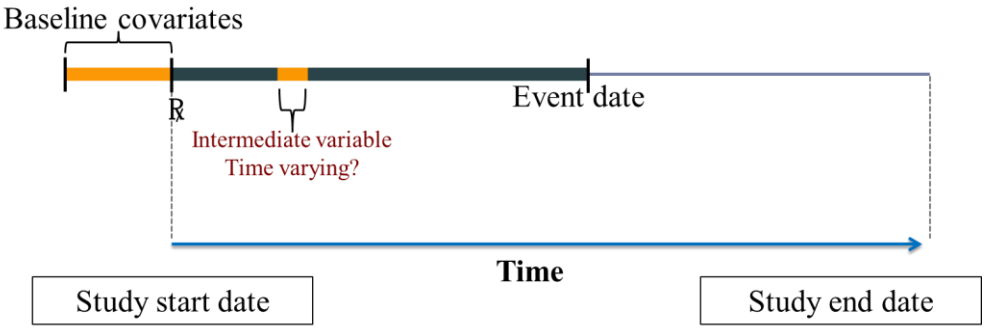
Immortal time bias also occurs when participants assessed from the time of enrolment are compared against groups defined by an outcome event occurring sometime during the follow-up. The types of events associated with this bias vary and may include the occurrence of a disease(s) or onset of an adverse drug effect. Comparative analyses using events that occur during follow-up are different from analyses using baseline characteristics that are specified completely before the occurrence of any outcome event.[7]

Specific calendar dates[95] or disease indication[90] for example, may be used to identify cohort entry and follow-up, however future exposures and events that occur during follow-up should not be used as criteria for cohort entry. Defining

cohort entry by exposure to a drug of interest substantially reduces the opportunity for immortal time bias to affect a study, even though it does not remove the bias completely.[7] For example, if a study follows a cohort of drug initiators, but then compares monotherapy to combination therapy, where these subgroups are identified 'at any time during follow-up' then immortal person-time is created by this exposure definition. Alternatively, if cohort entry is defined by more than one prescribed drug or more than a certain duration of treatment, however follow-up was initiated at the first prescription, then immortal person-time may be a source of bias.[7] Furthermore, conditioning the inclusion or exclusion of patients on the basis of information collected during follow-up, such as use of insulin after cohort entry,[82] adding-on other GLT[93] or switching to other GLT have the potential for bias arising from immortal time.

Overadjustment bias

Figure 2.4 Overadjustment bias

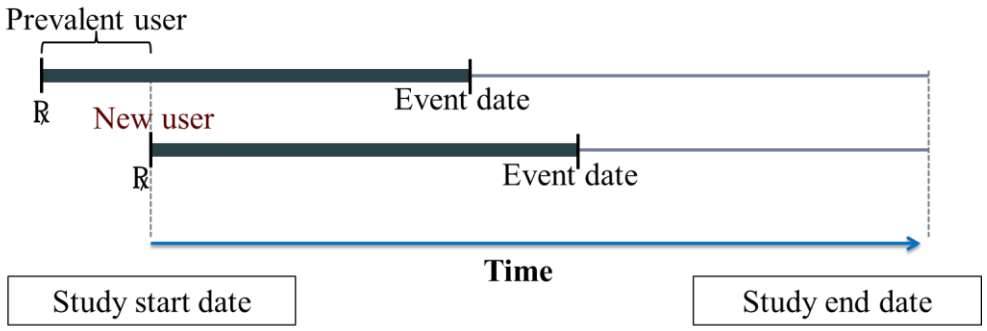


This may arise from adjustments made on covariates that are intermediate between treatment exposure and outcome of interest.[112] (Figure 2.4) These are covariates that are regularly assessed and reported during the follow-up period for monitoring purpose. The intermediate variables should be included as time-dependent covariates in the adjustment model. Appropriate sequencing of

covariate assessment, exposure status, cohort entry and follow-up will alleviate the potential for bias that might arise through adjustment of covariates that have been affected by a treatment exposure.

Healthy user or prevalent user bias

Figure 2.5 Prevalent user bias



The addition of a time-varying variable in a time-dependent Cox model or logistic regression model may introduce bias, which in turn may not provide the causal effect of treatment on an outcome. Treatments can be initiated sometime before, during or after the observation period in a cohort study. The inclusion of individuals who are prescribed a glucose-lowering therapy for some time before the study follow-up commences may introduce 'prevalent user' bias.[113] In addition, when a new GLT is initiated as intensification therapy, comparing new users of the drug with prevalent drug users may lead to false findings, thereby making new user patients to appear as having a higher baseline CV risk than a survivor cohort of prevalent users with adequate glucose control.[7]

Ray[114] described two types of prevalent user bias. The first type occurs when the risk of the outcome of interest varies over time as a result of the prevalent users being survivors of the initial period of drug therapy. Here, users who had experienced an outcome before the start of a study period might have discontinued the drug under study, thereby giving persistent users additional survival benefits. The second type relates to the alteration of risk factors that could influence outcomes of the treatment drug when started sometime before or at the study entry point.

The 'new user' design was proposed to address prevalent user bias.[114] It involves moving the time point of start of follow-up to the end of immortal time period or start of drug therapy. To avoid prevalent user bias, Lund et al[84] created a prevalent user cohort and new user sub-cohort during the study design (Figure 2.5). The prevalent user cohort included individuals who may have been using oral GLT for a period of time prior to study entry. Including these subjects in the analysis increased the statistical power; however it may have introduced bias if the risk of the outcome varied with time on treatment.[84]

New user design has been shown to minimise the risk of bias from including treatment groups with different durations of previous therapy.[114] This can be done by selecting all drug users who fulfil the treatment criteria of new-user group, and comparing against an appropriate non-users group. For example, if the treatment under investigation is normally used as second or third-line therapy, then a comparator medication that is similarly used at that stage is likely to be the most appropriate. Covariates at a time just before the start of the drug therapy among drug users and at a corresponding time among non-

users should be ascertained and included in the analysis. The new-user cohort design may also address problems of immortal time bias among the users. In which case, the start point of follow-up among the comparison group should include the 'immortal time' of the user cohort since differential exclusion of immortal time periods may introduce bias.[73]

Prescription bias

Prescription bias may occur as a result of failure to capture prescriptions for diabetes medications, or the absence of a complete longitudinal prescription data for all participants, or inability to confirm if patients took their medication. This form of misclassification bias makes it extremely difficult to define true exposure with clinical data and may also lead to selection bias if an exposure cohort has an incomplete prescription data.[98] However, this situation would underestimate rather than overestimate the risk of a drug. In addition, because GLTs are long-term medications for a chronic condition, the likelihood of inaccurate exposure assessment is reduced. Other diabetes medications could be used as comparators to reduce or eliminate potential biases.

Start-time bias

To address potential start-time bias, one of the reviewed studies compared a treatment cohort to an unexposed cohort by randomly matching each exposed subject for age and gender, with their start time defined as that of the exposed matched subject.[97] Analyses were performed by comparing outcomes in each exposed and unexposed cohort and contrasting them to one another, and also by direct comparison between the exposed cohorts. This eliminated potential

start-time bias in the unexposed cohort. However, it is important to note that baseline comparison between exposed and unexposed may not be meaningful when immortal time exists. Therefore careful consideration of methodology is needed in observational analyses with time-dependent treatment.

Surveillance bias

Surveillance bias is a type of information bias which occurs when some patients have more diagnostic tests performed on them or are followed up more closely than others, often leading to the more closely monitored group having a more frequent diagnosis of outcome, or measurement of parameters. One of the reviewed studies prevented surveillance bias by excluding patients who developed an event after regulatory warnings about a possible association between TZDs and bladder cancer were issued.[99] Surveillance bias could be investigated where a period of an event trigger is known. Subjects can be regrouped according to the rates of reporting before and after event trigger periods and analysed.

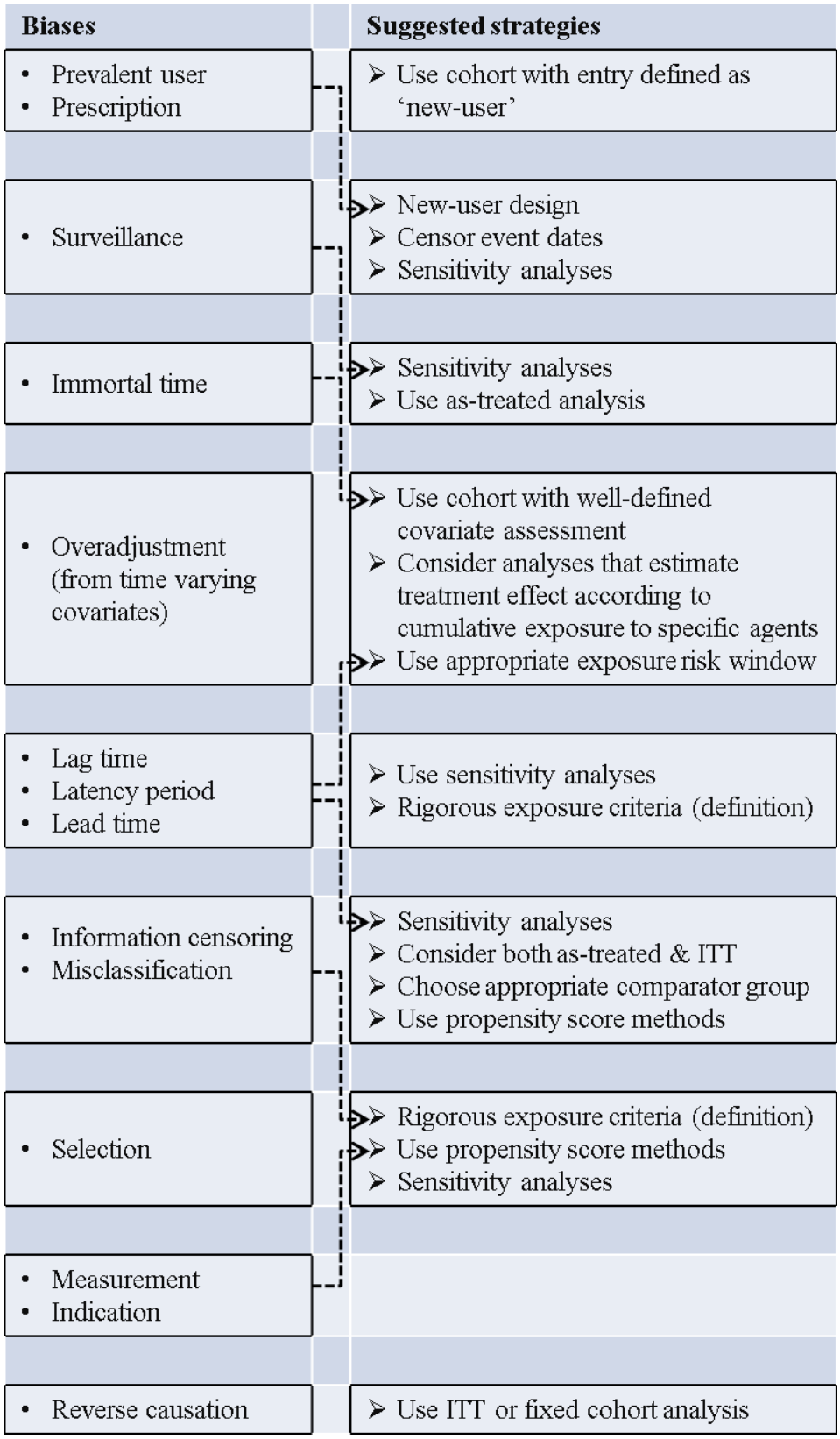
Detection biases

Detection bias is similar to surveillance bias. It occurs when patients on a particular treatment or having a particular disease are more frequently monitored than others, thereby leading to more opportunity for screening, early detection of a an event or a biased overestimation of the incidence of an event.[100] Tseng et al showed detection bias may have had an influence in their study after new-onset diabetes showed a 30% higher risk of developing colon cancer. However, whether diabetes duration played a causal role in developing cancer remained debatable.[100]

Allocation or indication bias

Confounding by indication (allocation bias) may be sources of bias in retrospective observational studies.[88] One of the reviewed studies demonstrated the high potential for this problem to occur.[102] For example, less healthy patients were more likely to have been prescribed a simple-to-use once daily insulin glargine regimen than other insulins that require more frequent injections or have a greater risk of nocturnal hypoglycaemia. Less healthy patients might also be less likely to be prescribed additional non-glargine insulin on top of insulin glargine. As a result, confounding by indication could not be excluded as the most likely reason for observing a higher cancer rate in insulin glargine users. Results of observational analysis of drug effects may not be a substitute for results of randomised trials because, fundamentally, it is impossible to rule out indication bias except by random allocation.[102] Therefore, propensity scores analyses were derived to exclude allocation bias and produce comparable subgroups for a stratified analysis, as have also been used to adjust for selection bias.[81] The propensity score analysis mitigates against indication bias and strengthens the interpretation of a true association between an event outcome and exposure.[81]

Figure 2.6 Algorithm of biases and control strategies



2.5 Conclusion

There is a growing number of retrospective cohort studies being reported from assessment of healthcare records of patients with type 2 diabetes. However, only about 32% of studies in this area reported at least one type of bias and provided an explanation on how the bias was addressed. The types of biases identified from the systematic literature search have been described to increase their awareness.

Observational studies play a vital role in evaluating associations between glucose-lowering therapies and treatment outcomes. The assessment of retrospective cohort analysis of diabetes-related healthcare records demonstrates the low frequency of bias-reporting. This review builds on previous assessment of methodological challenges in observational studies by describing biases that have been mentioned or addressed in observational studies and suggests strategies that can be implemented to reduce biases in order to minimize spurious findings and ensure robust study conclusions (Figure 2.6).

As more observational studies attempting to assess the effect of glucose-lowering therapies on clinical outcomes emerge, attention should be drawn to the fact that different biases may have implications in different clinical areas, and a slightly varying approach may be taken in dealing with these biases depending on the study hypothesis, for example, when assessing the association between GLTs and CV outcomes, the biological hypothesis being tested should be clearly identified and the appropriate exposure risk window

chosen.[7] In addition, appropriate techniques for dealing with other methodological challenges such as confounding should be explored.

A systematic approach using standard methods was applied while evaluating the observational studies. Despite a rigorous search and study selection process, one limitation of this review is it's restriction to retrospective cohort studies where biases associated with the study design or analyses have been reported. Therefore, the problem of channelling bias, which influences case-control studies, for example was not discussed in this review. Patorno et al., 2014[7] had conducted an extensive review, whereby, 86 studies were investigated for any methodological challenge that may be associated with glucose-lowering medications and cardiovascular outcome studies (whether methodological issues were reported or not). The review demonstrated large numbers of observational studies employ improper practices with regard to study design and analysis. On the other hand, the current review provides an estimate of the frequency of bias reporting and the most commonly reported ones. In addition, the present review suggests strategies that can be employed in minimising and controlling biases to strengthen the evidence of findings in future analysis conducted in this thesis. Limitations of this review include the possibility of missing out valuable studies written in other languages, authors of reviewed articles were not contacted to clarify ambiguous methodological details or results. In additions, the types of biases and strategies described are not a perfect classification structure to follow as some of the biases are interrelated and bias control measures will vary depending on study

hypothesis. The summary estimates on measure of effects for the association of GLTs with outcomes obtained from the reviewed articles were not consistent.

Chapter 3: Methodology

3.1 Introduction

An increasing volume of patient information is being recorded electronically during the course of treatment in primary care. The quality of collected data has improved overtime due to evidence-based guidelines, financial incentives to GPs and advanced computer programming, which has made large volumes of data to be easily processed and analysed by researchers.[115]

As demonstrated in Section 2.1, a large number of studies have been conducted using electronic records obtained from hospitals, insurance companies or general practices. Sometimes, these data are not representative of the general population. However, a major drive for this research is being able to assess nationally-representative information from primary practices throughout the UK.

The UK has several databases of electronic medical records from practices nationwide. The most popular of the databases that are being utilised by researchers include the Clinical Practice Research Datalink (CPRD),[116] The Health Improvement Network (THIN)[117] and QRESEARCH.[118] The large size of these databases yields a better precision, especially when the population is split into subgroups in order to explore varying information across the UK. Nonetheless, the use of these databases is not without its problems. There may

be inaccuracies in the recorded information or incompleteness. In addition, some practices which have agreed to contribute their data may not be representative of the entire primary care. As a result, practices are being selected to receive training in order to improve the quality of data recording and data may only be included in the databases once it has been proven that their records exceed a certain standard.[115]

This chapter describes the source of data utilised in this project and the various statistical and epidemiological methods that have been employed in the extraction and organisation of data, as well as advanced statistical techniques employed. It also describes the meta-analysis technique employed in Chapter 8:

3.2 Data source: The Health Improvement Network

The Health Improvement Network (THIN) database contains anonymous patient data from over 500 general practices and over 12 million patients, with about half of the patients still alive and actively contributing prospective data throughout England and Wales. Historical information of patients who have died or transferred out of a practice is available in the database. THIN has been validated at both practice and dataset level by comparing its demographic information, morbidity, mortality, prevalence, and geographical rates with various national data sources, including Department of Health's issued Read codes for the Quality and Outcomes Framework (QOF), 2001 Census and the National Statistics and the Office for National Statistics (ONS). The database contains information on all past and current medical diagnosis, prescribed

medications, referrals to specialist, laboratory results, lifestyle characteristics and other health outcomes measurements.

Clinical information of patients is recorded in THIN using Read Codes – a hierarchical dictionary of medical nomenclature,[119] which provides detailed records of disease management at every point in time. THIN releases updates of their database to researchers every four months, with each new release containing additional information from practices. The research presented in this thesis was based on THIN versions 1305 and 1405, which contain data from 477 practices up to the end of May 2013 and 2014, respectively.

3.3 Study design and population

Retrospective cohort study design is employed in the analysis of data. Longitudinal studies follow study subjects through time to find out who develops an event and who does not. However, the cohorts being examined are based on existing data on exposure and outcomes. In addition, the data source already contains information on the diagnoses made for each patient, their prescriptions, and other important demographic and metabolic parameters. Therefore, unlike the conventional prospective cohort study, the exposure and outcomes of interest have already occurred. In these studies, it can be difficult to attain complete and accurate exposure data, which can result in misclassification of the exposure. Hence, additional statistical methods and strict selection criteria are used to address these methodological challenges.

The study population include all patients in THIN who have been identified to have T2DM and are 18 years or older. Information extracted on the cohorts of

patients who have commenced both oral and injectable glucose-lowering medications are described in subsequent chapters according to research questions being addressed.

The index date, or sometimes referred to as baseline date connotes the date that a medication under investigation is initiated.[114] The criteria for eligibility of patients in each study are described in subsequent sections. Depending on the study, patients are excluded if they have very serious medical conditions that may influence the outcome of the study, for example, end-stage kidney disease or liver disease, organ transplant, HIV/AIDS and cancer except for non-melanoma skin cancer.

3.4 Study variables

3.4.1 Exposure definition

Incident exposures are to the following glucose-lowering medications; MET, SU, TZD, DPP-4 inhibitor or insulin (INS). Various combination regimens involving the aforementioned medications are assessed for effectiveness and safety depending on research question. The follow-up period in all cohorts commences from the index date (the date an additional glucose-lowering medication is added to the regimen) until a switch to or another glucose-lowering medication is added, or the 90th day post index date when HbA1c level is recorded, or the end of a study period.

3.4.2 Outcome definitions

The primary and secondary outcomes being investigated differ between studies therefore each study outcome(s) are described in the relevant sections.

3.4.3 Covariates

The covariates have been selected a priori on the basis of clinical significance. These include baseline demographic and medical parameters of the patients and include the following: age (years), gender (male or female), socio-economic status (measured using Townsends Index of Deprivation(113)), body weight (kg), Body Mass Index (BMI, kg/m^2), baseline HbA1c (%, mmol/mol), total cholesterol levels (mmol/L), low-density lipoprotein cholesterol (LDL-C, mmol/L), high-density lipoprotein cholesterol (HDL-C mmol/L), triglycerides (mmol/L), glomerular filtration rate (GFR, $mls/min/1.73m^2$), urinary albumin creatinine ratio (ACR, mg/mol), systolic and diastolic blood pressures (mmHg), smoking status and duration of diabetes (years) drug treatment. Others include the use of lipid-lowering drugs (mainly statins), antihypertensive drugs, aspirin and the following comorbidities at baseline: coronary heart diseases (CHD), peripheral arterial disease (PAD), hypoglycaemia and heart failure. A standardised computerised algorithm was developed and used to extract these data from THIN using Read Codes and British National Formulary (BNF) Codes (for the medications). Appendix B provides further detail about the definition of comorbidities and classes of medications.

3.5 Statistical Analyses

3.5.1 Baseline comparison of treatment groups

Baseline characteristics that might distinguish between categories of users of medication and other patient groups are analysed using the chi-square test for

categorical variables and *t* test for continuous variables. Values of resulting categorical data are expressed as numbers or proportions (in percentage, %), whereas mean and standard deviation (SD) or median and interquartile range (IQR) are used to describe continuous data. Standardized differences are also used to compare baseline characteristics between the treatment groups. For a continuous covariate, the standardized difference is defined as

$$d = \frac{(\bar{x}_{treatment} - \bar{x}_{control})}{\sqrt{\frac{S^2_{treatment} + S^2_{control}}{2}}}$$

Where $\bar{x}_{\text{treatment}}$ and $\bar{x}_{\text{control}}$ denote sample mean of the covariate in treated and untreated subjects, respectively, whereas $S^2_{treatment}$ and $S^2_{control}$ denote sample variance of covariate in treated and untreated subjects, respectively.

The standardized difference for dichotomous variables is defined as

$$d = \frac{(\hat{P}_{treatment} - \hat{P}_{control})}{\sqrt{\frac{\hat{P}_{treatment}(1 - \hat{P}_{treatment}) + \hat{P}_{control}(1 - \hat{P}_{control})}{2}}}$$

Where $\hat{P}_{treatment}$ and $\hat{P}_{control}$ denote the prevalence or mean in treated and untreated subjects, respectively.

The standardized difference d compares mean differences in units of the pooled standard deviation. It also provides a framework for comparing the mean or prevalence of baseline covariates between treatment groups in a propensity score matched cohort. To indicate serious imbalance, $d < 0.1$ has been used as a criterion to indicate a negligible difference.[120]

3.5.2 Linear and multiple regression

Linear regression is used for modelling the relationship between a continuous outcome and a continuous exposure, for example, the relationship between change in HbA1c and change in weight. A scatterplot is used to illustrate this relationship. The Wald's P value which tests the null hypothesis of no association is taken into consideration. The assumptions underlying linear regressions are tested by examining the scatterplot to see if the association is approximately linear. Where the outcome variable is normally distributed for each of the explanatory variable, the assumption is tested by assessing whether the distribution of residuals is reasonably normally distributed. The spread (standard deviation) is assessed to check whether it remains constant across a range of explanatory variables.[121]

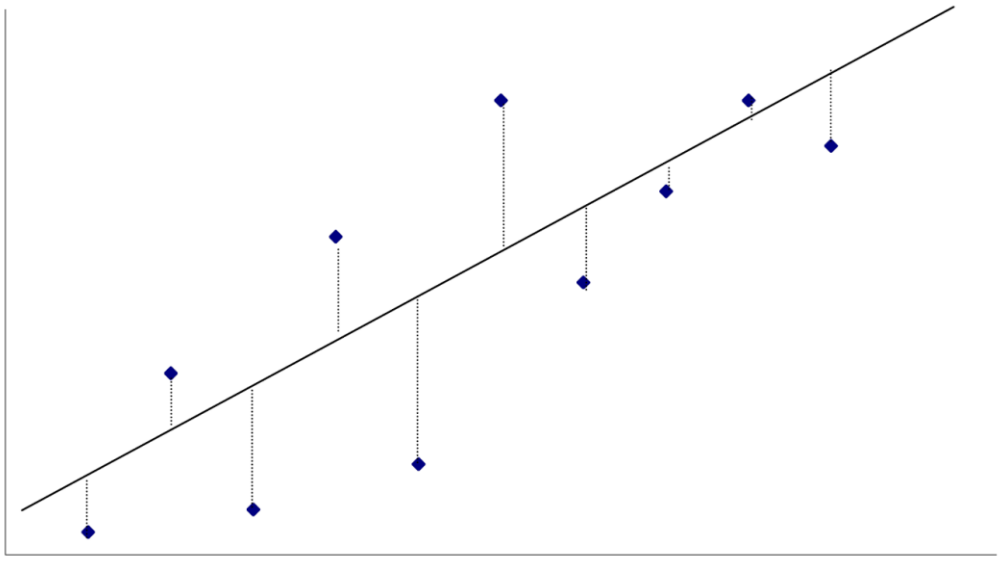
Multivariate regression is employed to examine the effect of GLTs on the outcome of interest, having adjusted for potential confounders. This model is developed firstly by constructing a basic univariate analysis between the GLT and outcome of interest so as to establish the unadjusted or crude estimate of the measure of effect. A list of a priori confounders and potential confounding factors described in Section 3.4.3 are investigated for their confounding effect. A priori confounders refer to variables that are thought to have strong reasons to likely confound the association between GLT and outcome of interest.

All the variables that are identified as potential confounders are added into the regression model. The effect of the exposure of interest is examined and potential confounders not deemed fit are removed from the model one at a

time, which noting the changes they make to the model as they are removed. If the exclusion of the potential confounder changes the effect of the exposure on the model by >10% then it is considered as an important confounder and included in the fully adjusted model. The fully-adjusted model with the exposure of interest is then fit with all of the important confounding variables which have been identified. The model is re-examined and results from the fully adjusted model are reported as adjusted odds ratio. Binary and categorical variables are fitted into the multiple regression model and their effects are assessed by choosing one of the categories as the reference (baseline) category. Potential a priori confounding variables are also assessed for possible interactions using the likelihood ratio tests.[122]

3.5.3 Least squares

Figure 3.1 Graphical illustration of least squares



Perfect linear relationships between variables are unlikely in observational studies. This may be due to experimental error or the underlying relationship may be approximately linear.[121] Instead of finding the best fit line, the method of least squares is used to determine the best fit line to data. As illustrated in Figure 3.1, least squares method finds the values of the regression coefficients which minimise the vertical distances between the points and the line. Least squares means (LS Means) or marginal means are the group means after adjusting for confounders. LS Mean represents the average of actual mean values, having adjusted for multiple factors including both categorical and continuous covariates (baseline characteristic measures). The LS mean is used when an inferential comparison between treatments is required.

3.5.4 Logistic regression model

Logistic regression is used for the analysis of studies with a binary outcome variable, for example responders to treatment (yes or no). The rationale for using logistic regression is to assess binary outcomes which are collected at specific time points during follow-up. This produces estimates of the Mantel-Haenszel (MH) odds ratios (OR) and 95% confidence interval around the MH-OR. Logistic regression is a more flexible approach than the MH method as it can cope with a large number of confounding variables and interaction terms and it can also include continuous exposure variables as well as binary or categorical exposure variables.[121]

Logistic regression models are developed for binary outcomes while taking account of confounding variables. Models that are based on odds are used,

where the odds of having an outcome of interest are calculated and compared according to a priori confounders. The effects of these characteristics are then assessed on the odds of the outcome, for example, the ratio between the odds of responding to intensification treatment in T2DM patients who have used a DPP-4 inhibitor to the odds of response in those who used a comparator GLT serves as the estimated odds ratio. Multivariate regression is used to examine the effect of an exposure variable of interest on an outcome while adjusting for potential confounders. Adjusted odds ratios from the regression models are expressed as point estimates with 95% confidence intervals (CI) at the significance level of 0.05.

Logistic regression model also serves as the basis for estimating the probability of having an outcome of interest. For example, if p is the probability of responding to treatment, then the odds of response is $\frac{p}{1-p}$ which is the probability or risk of responders divided by the risk of not responding. The natural logarithm of this is used in the logistic model, which is log $(\frac{p}{1-p})$.

Linear prediction function

Results are made clearer by computing predicted or expected values for hypothetical cases. For example, to get a practical understanding for the impact of DPP-4 inhibitor intensification therapy in a logistic regression model, comparisons are made on the predicted probabilities of treatment response for patients on DPP-4 inhibitor vs comparator treatment group who have different baseline HbA1c values in the model. Such predictions are referred to as

predictive margins. In addition, the marginal effects can succinctly show, for example, how the adjusted predictions within a categorical variable differ.[123]

3.5.5 Longitudinal data analyses

The incidence of a disease or an event is measured and expressed as the rate of the disease. Incidence rates are presented as per 1,000 person years. The person-time at risk is the sum of the time each person remains 'at risk' during the study period.

It is possible that event rates in longitudinal studies can change rapidly over time, especially when measured from a specific event such as commencement of a treatment or diagnosis of disease.[121] This information would be lost if incidence rates are calculated for the year following initiation of treatment. Therefore, another approach known as survival analysis is used to analyse data where the outcome of interest is the time until the event of interest occurs. Survival analysis is used to assess time to treatment failure as well as time to diagnosis of cardiovascular diseases or death.

Survival analysis

The primary challenge with survival data is that not all patients in the cohort have an actual survival time. Some patients could be lost to follow-up before the end of study, whereas others may not have had the event of interest at the end of the study; these are referred to as censored patients and the follow-up times employed for such participants are called right-censored observations.[121] With survival data, every patient in the study has a follow-up time and also a

censoring outcome variable which indicates whether the follow-up time was censored or just an observed time until the event of interest occurred.

Kaplan-Meier (KM) survival estimation

The summary and display of survival data are based on the calculations of survival probabilities throughout the duration of the study and the median survival time. The survival probability also known as survival function is the probability of surviving until at least a certain time or equivalently the probability of not having the event of interest by the set time.[121] With censored survival data, the Kaplan-Meier method is used to calculate the survival probabilities. This method uses the actual and the censored survival times to calculate the estimated probability of survival for all times throughout the observational period. The KM estimates are plotted against time to show a survival curve.

To determine whether there is a significant difference in survival between two or more treatment groups, for example, to see if a new treatment improves survival compared with a standard treatment, special statistical tests are conducted to examine this. The log-rank test is used to compare survival curves across the entire study period rather than just at one specific point in time. It compares the observed survival in the separate groups with the survival pattern that would be expected if all groups had the same underlying survival experience (this is the null hypothesis).

Cox regression

The log-rank test provides a relatively straightforward way of comparing survival curves. However, it is not able to cope with more complex survival data with several confounders including continuous variables and possible interactions. Therefore, a statistical model with more flexible and powerful approach is developed to take account of censored observations and handle many explanatory variables and interactions. The Cox proportional hazards model is therefore used to derive hazard ratios which compare the hazard rates in different exposure treatment groups.

The Cox model is based on the hazard rate $h(t)$ which is the instantaneous rate of the outcome at time *t* during follow-up. It is not necessary for the hazard function to be constant over time in the Cox model, but can take any form. This makes the model more flexible.

The Cox proportional hazards model takes the form of:

$$\log\big(h(t)\big) = \log\big(h_0(t)\big) + \beta_1 X_1 + \beta_2 X_2 + \ldots$$

Where, $h(t)$ = the hazard rate at any time; $h(t)$ = the baseline hazard at time t, it is the hazard rate when all of the exposure variables are equal to 0. For an exposure variable coded 0 = unexposed and 1= exposed, it can be shown that at any time t:

$$\frac{h(t)\ in\ exposed\ people}{h(t) in\ unexposed\ people} = \frac{h_0(t)\exp(\beta_1)}{h_0(t)} = \exp(\beta_1)$$

This is called the hazard ratio (HR). The HR is interpreted in a similar way to the rate ratio described above. Values over 1 indicate a higher hazard rate in

the exposed group, meaning their survival is lower. Values less than 1 indicate a lower hazard rate in the exposed group, meaning higher survival. The adjusted hazard ratios of associations are calculated and reported where applicable. Time dependent covariates are also included in the Cox regression model during analysis. Adjusted hazard ratios (aHR) are also expressed as point estimates with 95% confidence interval. Likelihood ratio tests (LRT) and Wald's tests are calculated for Cox proportional hazards model to determine statistical significance, using the same approach as for logistic regression models.

3.5.6 Propensity score analyses

Estimation

As patients are not randomly assigned to treatments in this study, factors that influence an outcome may also influence the assigned treatment. Therefore, analyses are conducted to adjust for the confounding factors and properly estimate their influence. A common approach to dealing with multiple confounding factors that affect both the treatment and outcome is propensity scoring.[124] A propensity score (PS), e is the probability of having the factor of interest, for example, receiving a particular treatment given the confounding factors present at baseline.

$$e = \Pr(Z_i = 1|X_i)$$

Where Z is an indicator variable denoting the treatment received (Z=0 for control or comparison treatment vs. Z =1 for active treatment) by each subject (i), and conditional on observed baseline covariate (X).[124]

In observational studies, treatment assigned subjects often differ systematically from untreated subjects. Therefore, an unbiased estimate of the average treatment effect cannot be obtained by directly comparing outcomes between the two treatment groups. Propensity score is used to estimate average treatment effects.

Propensity score is estimated using the study data and statistical modelling techniques. For example, logistic regression and multivariable logistic regression models are used to obtain predicted probabilities or propensity score values for both the probability of being treated (p) and the probability of not being treated ($1 - p$). Where more than 2 treatment groups were examined, a model with multinomial outcome or multiple logistic regressions was used.

The effect of treatment for each subject is defined to be $Y_i(1) - Y_i(0)$, where Y represents the potential outcomes for treatment (1) vs control (0) observed for each patient i.The average treatment effect (ATE) is the average effect of moving an entire population from untreated to treated, which is defined as $E[Y_i(1) - Y_i(0)]$.[125] A related measure of treatment effect used is the average treatment effect for the treated (ATT). The ATT is the average effect that would be seen if everyone in the treated group received the treatment, compared to if no one in the treated group received the treatment;[126] defined as $E[Y_i(1) - Y_i(0)|Z = 1]$.

The treatment effects vary across individuals in the treated and comparison groups in observational studies hence the ATT and ATE may not be equal since different methods may yield different estimates of the treatment effect.

Therefore, the most appropriate causal estimate was used to answer the research questions in subsequent chapters and the appropriate propensity score method was used to estimate that quantity. These methods are briefly described in Sections 7.2 and 8.2 of this thesis.

Propensity score application

The estimated PS was applied to the treated and comparison individuals in preparation for the final outcome model. Below is the description of two PS application methods employed in this thesis: full PS matching (PSM) and inverse probability of treatment weights (IPTW). The following is an overview of the PS application methods employed.

a) PS Matching

Full PSM analysis uses all individuals in the study population (treatment and comparison) to form matched sets that have at least one treated and at least one comparison subject in each matched set.[127] The matched sets are created a way that ensures a minimal difference in the global PS is present, defined as the sum of the distances between the PS of all pairs of treated vs comparison subjects within each matched set and across all matched sets.[127] After the matched sets are formed, each treated subject receives a weight of 1, while subjects in the comparison group within each matched set receive a weight that is proportional to the number of treated individuals in the matched set divided by the number of comparison subject in the matched set. As a result of this weighting approach, the ATT is estimated in the full PSM method employed.

The standardised bias of each measured covariate included in the model is then calculated using weighted proportions and weighted standard deviations.

Logistic regression was used to regress dichotomous outcome on an indicator variable in the matched sample and standard errors were obtained using a robust variance estimator.[128] Modifications were made on logistic regression models by adjusting for the baseline covariates described in Section 3.4.3. For studies where outcome was time-to-event, a Cox proportional hazards model was fitted. Kaplan-Meier survival curves were estimated separately for the treated and untreated participants in the PSM sample. The log-rank test is not appropriate for comparing the Kaplan-Meier survival curves between treatment groups because the test assumes two independent samples.[129,130] However, the stratified log-rank test is appropriate for data involving matched pairs.[129] Finally, Cox proportional hazards model was used to regress survival time on an indicator variable denoting the treatment status.

b) Inverse probability of treatment weights

IPTW is a weighting technique that is used when ATE is the desired estimand. In the event that a measure of treatment effect is required from the comparison of more than two treatment groups, a multinomial propensity score was estimated based on all the baseline covariates in the study population.[131] The entire study population was weighted by inverse probability of treatment weights derived from the propensity score. The IPTW for each subject was defined as $\frac{Z_i}{e_i} + \frac{1-Z_i}{1-e_i}$, where Z_i denotes treatment status for whether or not the ith participant was and e denotes the estimated propensity score. Assuming

that Y_i denotes the outcome variable measured on the ith subject, the first estimate of the ATE is then defined as

$$\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i Y_i}{e_i} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-Z_i)Y_i}{1-e_i}$$

Where n denotes the number of participants in the full sample.[132] To adjust for baseline covariates, weighted estimator from the family of doubly robust estimators was used, as described by Lunceford and Davidian. The estimator requires specifying the PS model and regression models relating the expected outcome to baseline covariates in treated and untreated individuals separately.[133] IPTW using the PS score belongs to the class of models called marginal structural models that allow one to account for time-varying confounders when estimating the effect of time-varying exposures.[134]

Logistic regression was used to regress dichotomous outcome on an indicator variable in the weighted cohort, while standard errors were obtained using a robust variance estimator.[128] Modifications were made on logistic regression models by adjusting for the baseline covariates in each study.

For studies where outcome was time-to-event, Cox proportional hazards model was fitted using IPTW. In addition, a robust variance estimator was employed to account for the weighted nature of the sample. Unlike PSM where stratified log-lank test was conducted. Here, the adjusted Kaplan-Meier estimates of survival curves was obtained in an approach similar to the methods employed by Xie and Liu,[135] who proposed a weighted version of the log-rank test to test the null hypothesis that the survival curves are equal to one another.[135]

Balance diagnostics

A true propensity score is a balancing score whereby, the cohort of all subjects with the same probability of treatment will have similar distributions of baseline factors among those who were treated and those who were not treated.[105] As this condition is not fully met in observational studies (unlike RCTs in which the true propensity score is often defined by the study design). Thus, balance diagnostics allow one to investigate whether the propensity score model has been adequately specified. Standardized difference was used to compare the difference in means and prevalence of the baseline covariates. Unlike t-tests and other statistical tests of hypothesis, the standardized difference is not influenced by sample size. Therefore, standard difference was used to compare balance in measured variables between treatment group subjects as demonstrated in the literature.[136]

3.5.7 Multiple imputation using chained equations

Routinely collected data are rich data sources. However, missing data among key variables are a substantial problem to many clinical databases and can cause bias in observational studies depending on the level of missingness associated with the variables.[137] Many variables in primary care databases are recorded mainly for their clinical relevance. For example, patients with T2DM and cardiovascular diseases are more likely to have health indicators recorded compared to the rest of the primary care population.[137]

Multiple imputation (MI) technique is used to deal with missing data in this project. The rationale behind the use of MI is its applicability even if data are

missing not at random (MNAR).[138] In addition, estimating the variance and finding confidence intervals is relatively straight forward with MI. Once missing data are imputed, the imputed data sets are easily analysed as though they are complete. MI is particularly helpful when there are missing covariate information. For example, assuming there are two variables to assess, DBP and SBP, but SBP has some missing values. Data from individuals with SBP and DBP is used to estimate the conditional distribution of SBP given DBP. That means for individuals with missing SBP record, imputation is done by drawing randomly from SBP|DBP, where this conditional distribution has been estimated in complete cases. This is done multiple (M) times (typically ≥ 5 times), giving rise to M 'complete' data records. These data are then analysed in using the appropriate rules that has been set out for analysis. The main obstacle for application of MI in large clinical databases lies with limited information made available, giving rise to missing data as procedures or motivation for recording data change over time.[138]

Imputation by chained equations (ICE), also known as full conditional specification (FCS) involves the specification of univariate imputation models for each partially observed variable. For example, suppose X represents fully observed variables, and Y_1, Y_2 and Y_3 have missing values, models are specified for:

$$f(Y_1 | Y_2, Y_3, X)$$

$$f(Y_2 | Y_1, Y_3, X)$$

$$f(Y_3 | Y_1, Y_2, X)$$

Specific number of imputation, are conducted by initially imputing missing values in Y_1, Y_2 and Y_3 by random sampling from the observed values. MI with chained equations has the flexibility of allowing different model types to be specified for each variable, for example, linear regression for normally distributed variable, logistic for binary variables and multinomial logistic for unordered categorical variables. MI is employed by ensuring careful selection of variables for inclusion into MI models. The outcome variable in the final model of interest is also included in the imputation model. In the survival analysis for example, the outcome variable *t*, which represents time to event of interest and the event indicator d are incorporated into the model. Where d = 0, t records the censoring time (the last time at which a subject was seen, and had still not had the event).[139]

Incorporating time to event outcomes in imputation of Cox proportional hazards model which assumes the hazard at time t, given covariates X is given by:

$$h(t|X) = h_0(t)\exp(\beta^t X)$$

Where $h_0(t)$ denotes an arbitrary baseline hazard function and β a vector of (log) hazard ratios.[140]

3.5.8 Sensitivity and subgroup analyses

Multiple sensitivity and subgroup analyses are performed to assess the reliability of the outcomes, the impact of missing data and potential unmeasured confounders.[141,142] Subgroup analyses are conducted to provide greater understanding of each research question. These are demonstrated

according to study hypothesis in the relevant sections of this thesis and include descriptive analyses or stratification analyses.

3.5.9 Statistical software

Stata version 13[143] and R[144] are the statistical software packages used for all calculations and analyses. In addition, statistical software programmes 'teffects' (treatment effects estimation for observational data) and 'pbalchk' (balance checking of the covariates) written for Stata are used to estimate parametric propensity score models and propensity score techniques which involve matching and regression adjustments on IPTW. Another programme, 'twang' written in R is used to estimate nonparametric PS models, weighting application methods and balance assessment. Standardized bias are examined through both graphical and numerical methods and reported in the relevant sections of this thesis. The automated features of both R and Stata software are useful for facilitating balance check on measured covariates.

3.6 Potential biases and limitations

A possible weakness to the electronic health record studies could be attributed to the identification of outcome measures due to variable coding practices. For example, not all recorded diagnoses of diabetes in primary care are explicitly linked to a confirmatory test results. However, THIN data has been validated at both practice and dataset against various national data sources, including Department of Health's issued Read codes for the Quality and Outcomes Framework (QOF).

There is also the possibility that an outcome measure in an observational study may be rare and so may not have adequate power to demonstrate statistical significance. Sample size calculations based on some very conservative assumptions demonstrate that fewer than 5000 patients are needed to demonstrate statistically significant results. However, the datasets explored in this thesis contain relevant information of over 25,000 patients. This suggests that the study is adequately powered and more so that the study period under assessment of more than five years provides additional number of exposure cases and outcomes.

The limitations of the studies conducted using data from THIN are discussed in more detail in the respective chapters.

3.7 Meta-analysis

Meta-analysis is used in Chapter 8 to assess and determine whether the use of DPP-4 inhibitor in the treatment of T2DM has a causal contribution to the development of bone fracture. Meta-analysis is used to calculate the average measure of effect by assembling quantitative results from several RCT studies together. The number of fracture events reported in the exposed and unexposed (comparator) treatment groups of each primary study are extracted and used to calculate a new single measure of effect, referred to as the pooled result or summary statistic. The pooled data from the studies are expressed as an odds ratio, together with its precision (95% confidence interval). The results from all of the studies are summarized and displayed using forest plots.

3.7.1 Assessing heterogeneity between studies

Analyses are conducted to determine whether the results from each of the studies are similar. The term heterogeneity is used to describe the degree to which the studies vary. The I^2 statistical test is used to quantify the effect of heterogeneity between the results of the studies. I^2 value range from 0% to 100%, representing the percentage of total variation across the studies that is due to true heterogeneity rather than to chance. I^2 by itself does not explain the actual range of effects.[145] A value of 0% would indicate that there is no variability between studies that cannot be explained by chance, whereas, a value of 50% would indicate that 50% of the total variability in the meta-analysis is due to heterogeneity rather than to chance, and a value of 100% would indicate that all of the variation in the meta-analysis is due to heterogeneity rather than to chance. I^2 is used together with the observed effects to provide a sense of the true effects. In addition, the forest plot and I^2 are both used together to get a sense of the absolute dispersion.

3.7.2 Pooling meta-analysis data

In pooling the effect estimates of the studies together to generate the odds ratio, fixed effect method is used to calculate the summary statistic. The fixed effect method calculates a weighted average of the OR from all of the different studies – the weight being proportional to the size of the study. Therefore the bigger the sample size of a study, the more influence it will have on the pooled odds ratio. It is important to note that the goal of a meta-analysis is not simply to report the mean effect size, but also to report how the effect sizes in the various studies are dispersed about the mean.[146] The calculations are carried

out by calculating a weighted average of the log odds ratios using the inverse of the variance of each log odds ratio as the weight. The fixed effect method assumes that all of the available studies are trying to estimate the same true value, that is, it does not vary according to where or when or in whom the study was conducted.[147]

3.7.3 Bias in meta-analysis

A potential problem that can occur during the primary study search process is publication bias. Publication bias occurs when published studies that have found "interesting" (usually positive) results are more likely to be identified in during search and more likely to be published earlier than the "less interesting" (usually negative) ones. The extent of publication bias is assessed through the inspection of the magnitude of published effects in relation to the order of publication. If publication bias is present, then the earlier reports will usually tend to find larger effects than more recent studies. Another approach that is not dependent on the time sequence of publication is used to address bias. This is by using a funnel plot of the magnitude of the odds ratio against the study precision.[148,149]

3.8 Conclusion

Chapter 3 provides the description of advanced statistical methods and techniques employed in the analysis of THIN data as well as meta-analysis data to answer research questions in this thesis. The application of these methods, other methods used in the selection of study participants and the results of analyses are described and discussed in Chapters 4 to 8.

Chapter 4: Determinants of glycaemic response to DPP-4 inhibitor

4.1 Summary

Background and Aim

Apart from baseline HbA1c, little is known about clinical parameters that affect glycaemic response to a DPP-4 inhibitor when used in routine clinical practice. THIN database was explored in order to assess the variability in response to DPP-4 inhibitor when used as add-on therapy.

Methods

Data on 25,386 patients with type 2 diabetes (T2DM), newly treated with a DPP-4 inhibitor (2007-2013), was sourced from UK General Practices via THIN database. Baseline clinical parameters of patients (n=13,525) in whom a DPP-4 inhibitor was added because of suboptimal glucose control (HbA1c>7%) were compared with 12-months follow-up data. An optimum response to the DPP-4 inhibitor was defined as HbA1c <7.0% at 12 months. Descriptive analyses and unadjusted comparisons using Chi squared and t tests were carried out to ascertain glycaemic and body weight responses to treatment intensification with a DPP-4 inhibitor. Predictor of response analyses were performed using multivariate logistic regression.

Results

Overall, 1,708 (13%) of our study population achieved Hba1c of <7%, Intensification with a DPP-4 inhibitor was associated with significant reductions in HbA1c (-0.5%), body weight (-0.9kg) and total cholesterol (-0.1mmol/L), p<0.001. Independent predictors of achieving optimal HbA1c target of <7% included the use of MET (Adjusted Odds Ratio: 2.58, 95%CI: 2.18-3.04) and use of MET plus SU (1.42, 95%CI: 1.21-1.68) as oppose to none use. The independent predictors of suboptimal glucose control included higher baseline HbA1c (OR: 0.64; 0.61-0.68; i.e., 1% increase in Hba1c was associated with a 36% reduced likelihood of response), longer diabetes duration (per every year increase) (0.85; 0.83-0.88) and intensification therapy below 9 months compared with 9-12 months.

Conclusion

There is a significant variability in glycaemic response to a DPP-4 inhibitor in routine practice. The best effect is achieved as add-on to MET and MET plus SU, but responses are significantly lower with increased diabetes duration and among patients with high HbA1c at baseline.

4.2 Introduction

Randomised clinical trials (RCT) have examined the efficacy and safety of different glucose-lowering therapy (GLT) either as mono- or combination therapy in patients with T2DM.[27,28,50] Most patients require gradual escalation of therapy and multiple treatment options are becoming more widely available,[5] but there are few head-to-head clinical trials to compare outcomes

in routine clinical practice using different dosing and drug sequencing options.[150] In particular, the comparative effectiveness of a DPP-4 inhibitor as second, third or fourth line therapy beyond MET is unclear.

Baseline HbA1c is a well-recognised determinant of glycaemic response to many different therapies, including DPP-4 inhibitors, but beyond this little is known about which clinical or biochemical factors influence the glycaemic response in everyday practice when DPP-4 inhibitor is added to mono, dual or triple glucose-lowering regimen.

The aim of this study was to evaluate the variability and determinants of glycaemic response to DPP-4 inhibitor therapy in routine clinical practice when added to MET or SU monotherapy, and when used as add-on to dual (MET+SU) or triple (MET+SU+TZD) therapy.

4.3 Method

4.3.1 Study design and data source

Retrospective cohort analysis of data from THIN database was conducted. The study population comprised a cohort of patients identified to have T2DM and registered to a GP practice for >12 months before the index date. The index date (June 2007–May 2013) was defined as the date of initiation of DPP-4 inhibitor therapy. The cohort included patients who were >18 years old with suboptimal glucose control (HbA1c> 7.0%) 6 months or more after using other GLT. Patients were prescribed a DPP-4 inhibitor as add-on to other GLT.

4.3.2 Treatment exposure

Exposure was to at least two prescriptions for a DPP-4 inhibitor, from the index date (the date of the first prescription) until there was a switch to, or addition of, another GLT, or the 90th day post index date when HbA1c is recorded, or 12 months after the index date. Patients were segregated into the following treatment groups based on the oral GLT they received at baseline: MET monotherapy, SU monotherapy, MET+SU as dual therapy and those on triple therapy (MET+SU+TZD).

4.3.3 Outcome

The primary outcome was to determine the glycaemic effect of intensification with a DPP-4 inhibitor in terms of achieving HbA1c target of <7% after exposure to a DPP-4 inhibitor, and the factors that may influence this response or non-response according to the use of DPP-4 inhibitor as mono- or as add on therapies.

4.3.4 Statistical analysis

Baseline characteristics that were assumed to influence 'responders' and 'non-responders' to DPP-4 inhibitor therapy are described as Covariates in 3.4.3. Descriptive analysis was used to identify the baseline characteristics that might distinguish between responders and non-responders, as appropriate. t-test and Chi squared test were used to describe continuous and categorical variables, respectively. Multivariate logistic regressions were carried out to identify covariates that were associated with a response within 12 months.

4.3.5 Secondary analysis

Tests for interaction were carried out to compare the metabolic effects of a DPP-4 inhibitor as add-on therapy to MET-only, SU-only, MET+SU and MET+SU+TZD regimens. Comparative analysis on changes in HbA1c at 12 months was carried out and glycaemic response endpoint changes in HbA1c were assessed based on baseline HbA1c categories 7-7.5%, 7.5-8.0%, 8.0-9.0% and ≥9% respectively. In addition, the proportion of patients who achieved glycaemic targets (<7.0%) were also described for the full cohort and for those with HbA1c ≥7.5% at baseline.

4.3.6 Bias

To ensure individuals included in the study have not been treated with DPP-4 inhibitor for some time before the follow-up commenced, thereby introducing prevalent user bias, new users of DPP-4 inhibitor (individuals who initiated DPP-4inhibitors as intensification therapy) were assessed at baseline. The "new users" design was used to minimise bias that may be associated with prevalent use of DPP-4 inhibitors.[114]

Selection bias was minimised by ensuring all participants in the exposed treatment groups examined initiated DPP-4 inhibitors as new users. In addition, individuals whose HbA1c was <7% at the time of intensification with DPP-4 inhibitor were excluded as shown in Figure 4.1. This is because lower HbA1c may have been achieved as a result the effect of previous GLTs. Missing information of baseline characteristics on all individuals involved in the study were accounted for to further curtail selection bias.

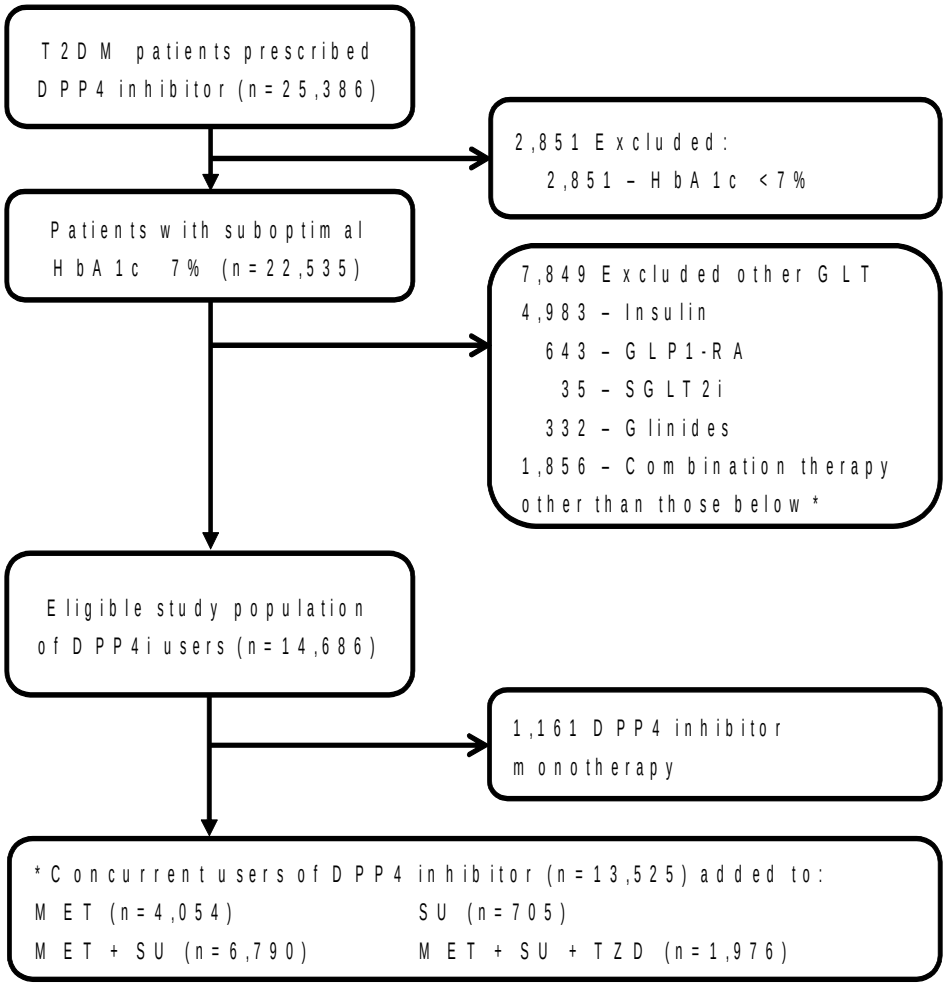
Post index date exposure to any other glucose-lowering therapy other than a DPP-4 inhibitor was not permitted. This was to reduce confounding by indication or misclassification bias. In addition, individuals were segregated into separate combination treatment groups to prevent confounding by co-medication. The cohort was restricted to an estimated 12 months follow-up to reduce the risk of bias introduced by an overlapping treatment effect.[114]

4.4 Results

4.4.1 Patient characteristics

Of the 25,386 users of a DPP-4 inhibitor, 13,525 patients fulfilled the criteria for cohort entry (Figure 4.1). The cohort had a mean age of 62 years (60% Male) and predominantly obese (61% with BMI $\geq$ 30kg/m^2). (Table 4.1) Treatment groups included patients prescribed a DPP-4 inhibitor as add-on therapy to MET alone (30%), SU alone (5%), MET+SU (50%) and MET+SU+TZD (15%). Baseline HbA1c was significantly lower among responders compared to non-responders (8.2% vs. 8.9% respectively, $p<0.001$) (Table 4.1).

Figure 4.1 Study participant selection



4.4.2 Response to DPP-4 inhibitor therapy

Overall, the addition of a DPP-4 inhibitor resulted in a 0.5% reduction in HbA1c ($p<0.001$). Approximately 13% of patients achieved HbA1c <7% following co-administration of a DPP-4 inhibitor, based on the criteria for response described previously. This response was not significantly different across gender, social deprivation or among patients using antihypertensive or lipid lowering medication. It was also not different with weight, BMI or smoking. Co-administration of a DPP-4 inhibitor was associated with a 0.9kg

reduction in body weight and a 0.1 mmol/l reduction in total cholesterol at 12 months (p<0.001).

Table 4.1 Baseline characteristics of patients prescribed DPP-4 inhibitor

Variables	Total* (N=13,525)	Responders (n=1,708)*	Non-responders (n=11,817)*	P value
Age (Yrs)	62.3 (12.2)	62.9 (12.0)	62.2 (12.2)	0.04
HbA1c (%)	8.8 (1.5)	8.2 (1.3)	8.9 (1.5)	<0.001
SBP (mmHg)	134.7 (15.3)	134.3 (15.0)	134.8 (15.3)	0.3
DBP (mmHg)	77.5 (9.6)	76.9 (9.6)	77.6 (9.6)	0.004
TC (mmol/l)	4.3 (1.1)	4.2 (1.0)	4.3 (1.1)	<0.001
HDL-C (mmol/l)	1.1 (0.3)	1.1 (0.3)	1.1 (0.3)	0.2
LDL-C (mmol/l)	2.2 (0.9)	2.2 (0.9)	2.3 (0.9)	0.01
TGC (mmol/L)	2.2 (2.3)	2.1 (3.8)	2.2 (2.0)	0.2
Weight (Kg)	93.2 (21.0)	93.5 (21.3)	93.1 (21.0)	0.6
Duration (Yrs) $^{\infty}$	1.5 (2.7)	1.0 (2.6)	1.7 (2.7)	<0.001
BMI (kg/m^2)	32.5 (6.5)	32.5 (6.8)	32.5 (6.5)	0.9
Gender				
Male	8113 (60)	1036 (61)	7077 (60)	-
Female	5412 (40)	672 (39)	4740 (40)	0.5
HbA1c level (%)				
7-7.5 (53-58)	2306 (17)	541 (32)	1765 (15)	.
7.5-8 (58-64)	2621 (19)	454 (27)	2167 (18)	<0.001
8-9 (64-75)	4051 (30)	445 (26)	3606 (31)	<0.001
≥ 9 (75)	4547 (34)	268 (16)	4279 (36)	<0.001
BMI (kg/m^2)				
Normal (<25kg/m^2)	1248 (9)	169 (10)	1079 (9)	.
Overweight (25-29.9kg/m^2)	4060 (30)	504 (30)	3556 (30)	0.3
Obese (≥ 30kg/m^2)	8217 (61)	1035 (61)	7182 (61)	0.4
Smoking Status				
Non-smoker	5238 (39)	665 (39)	4573 (39)	.
Current	2053 (15)	246 (14)	1807 (15)	0.4
Ex-smoker	6234 (46)	797 (47)	5437 (46)	0.9
Deprivation				
Least deprived	2992 (22)	402 (24)	2590 (22)	.
Less	2875 (21)	373 (22)	2502 (21)	0.6
Average	2811 (21)	355 (21)	2456 (21)	0.3
More	2732 (20)	323 (19)	2409 (20)	0.1
Most deprived	2115 (16)	255 (15)	1860 (16)	0.1

Variables	Total* (N=13,525)	Responders (n=1,708)*	Non-responders (n=11,817)*	P value
Comorbidity				
CHD	7822 (58)	986 (58)	6836 (58)	0.9
PAD	2277 (17)	255 (15)	2022 (17)	0.02
Stroke	3071 (23)	402 (24)	2669 (23)	0.4
Heart Failure	1595 (12)	198 (12)	1397 (12)	0.8
Hypoglycaemia	2478 (18)	277 (16)	2201 (19)	0.02
Other Medication				
Aspirin	5270 (39)	705 (41)	4565 (39)	0.04
Antihypertensive	9869 (73)	1265 (74)	8604 (73)	0.2
LLT	10537 (78)	1339 (78)	9198 (78)	0.6
Oral GLTs				
MET alone	4054 (30)	794 (46)	3260 (28)	<0.001
SU alone	705 (5)	62 (4)	643 (5)	0.002
MET+SU	6790 (50)	703 (41)	6087 (52)	<0.001
MET+SU+TZD	1976 (15)	149 (9)	1827 (15)	<0.001
Follow-up (Mths)				
9 to 12	8740 (65)	1388 (81)	7352 (62)	-
6 to <9	1484 (11)	96 (6)	1388 (12)	<0.001
3 to <6	1627 (12)	110 (6)	1517 (13)	<0.001
0 to <3	1674 (12)	114 (7)	1560 (13)	<0.001

* Mean (standard deviation) for continuous variables; Frequency (percentage) for categorical variables
∞ Estimated diabetes duration as time from first glucose-lowering therapy

4.4.3 Factors influencing outcomes

After adjusting for baseline HbA1c, duration of diabetes, previous diagnosis of peripheral arterial disease and treatment duration, the odds of responding to intensification with a DPP-4 inhibitor is approximately 2.6 times more when DPP-4 inhibitor is co-administered with MET than when it is not (Adjusted Odds Ratio, 2.58; 95%CI, 2.18-3.04). The odds of response is also increased by 42% when DPP-4 inhibitor is added to MET + SU dual therapy as oppose to none (1.42, 95%CI: 1.21-1.68). On the other hand, the odds of not responding to DPP-4 inhibitor independently decreased by 36% (0.64; 95%CI:

0.61-0.68) for each % unit increase in HbA1c and also decreased by 15% (0.85; 95% CI: 0.83-0.88) for each unit increase in diabetes duration (years). (Table 4.2)

Table 4.2 Logistic regression model for attaining <7.0% HbA1c target

	Unadjusted		Adjusted*	
	OR (95% CI)	**p value**	**OR (95% CI)**	**p value**
Age (Yrs)	1.00 (1.00,1.01)	0.3		
HbA1c (%)	0.65 (0.62,0.69)	<0.001	0.64 (0.61,0.68)	<0.001
DBP (mmHg)	0.99 (0.99,1.00)	0.01		
TC (mmol/l)	0.97 (0.90,1.05)	0.5		
LDL-C(mmol/l)	1.01 (0.92,1.10)	0.9		
Duration (Yrs) [∞]	0.85 (0.82,0.88)	<0.001	0.85 (0.83,0.88)	<0.001
Comorbidity				
PAD	0.85 (0.73,0.98)	0.8	0.88 (0.76,1.01)	0.08
Hypoglycaemia	0.92 (0.80,1.06)	0.3		
Medication				
Aspirin	1.00 (0.90,1.12)	0.9		
Oral GLTs				
MET alone	2.89 (2.39,3.50)	<0.001	2.58 (2.19,3.04)	<0.001
SU alone	1.35 (0.98,1.85)	0.07		
MET+SU	1.55 (1.29,1.87)	<0.001	1.42 (1.21,1.68)	<0.001
MET+SU+TZD	1 (1.00,1.00)	1.0		
Treatment, Mth				
9 to 12	1 (1.00,1.00)	-	1.00	.
6 to <9	0.42 (0.34,0.52)	<0.001	0.42 (0.34,0.52)	<0.001
3 to <6	0.47 (0.38,0.58)	<0.001	0.47 (0.38,0.57)	<0.001
0 to <3	0.47 (0.38,0.58)	<0.001	0.47 (0.38,0.58)	<0.001

Abbreviations: OR – Odds Ratio of Predictors of response or non-response;
[∞] Estimated as time from first glucose-lowering therapy

4.4.4 Effectiveness of DPP-4 inhibitor as add-on therapy

The glycaemic effectiveness of DPP-4 inhibitor when added to different oral glucose-lowering regimens was examined. The baseline glucose-lowering medications that the patients received prior to intensification differed remarkably. Therefore, a head-to-head comparison on the effectiveness across

treatment groups was not assessed. The probability of response was predicted based on DPP-4 inhibitor treatment follow-up time in months (Figure 4.2).

Figure 4.2 Predicted probability of response to DPP-4 inhibitor over time

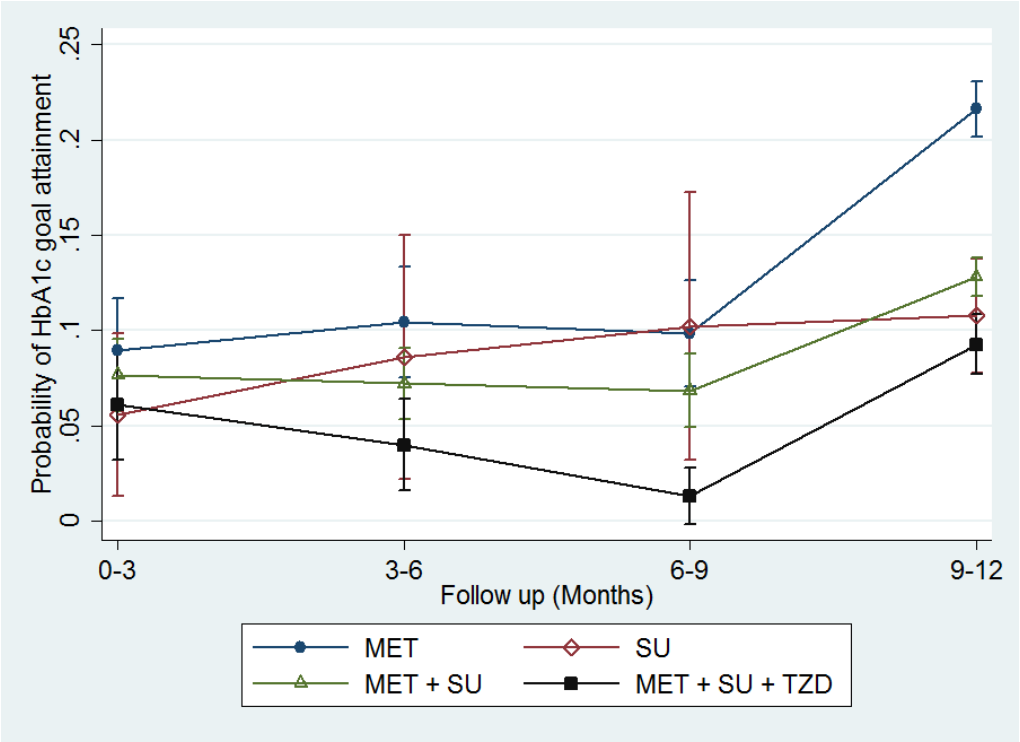


Figure 4.2 shows the probability of responding to intensification with a DPP-4 inhibitor was consistently higher when added to MET vs SU monotherapies, or MET + SU dual therapy, with the highest probability of responders after 9 months of treatment. Response to DPP-4-inhibitor co-administration with triple regimen was poor.

Table 4.3 Descriptive analysis for HbA1c, weight and cholesterol responses

Parameter	MET alone	SU alone	MET + SU	MET + SU + TZD
Number of patients (%)	4,054 (30)	705 (5)	6,790 (50)	1,976 (15)
Mean (SE) Age, Years	60 (0.2)	70 (0.5)	63 (0.1)	63 (0.3)
Mean (SE) HbA1c, %	8.5 (0.02)	9.0 (0.06)	9.0 (0.02)	8.8 (0.03)
Mean (SE) Duration of Diabetes, Years$^{\infty}$	1.7 (0.04)	1.7 (0.1)	1.5 (0.03)	1.2 (0.06)
Pr. Response (SE)	0.18 (0.01)	0.10 (0.01)	0.11 (0.01)	0.07 (0.01)
Overall Mean (SE) Change in HbA1c, %$^{\wedge}$	-0.58 (0.02)	-0.42 (0.05)	-0.48 (0.02)	-0.21 (0.03)
Subgroup HbA1c (SE) Change in HbA1c, %				
7 to <7.5%	-0.19 (0.04)	0.25 (0.1)	0.06 (0.04)†	0.35 (0.1)
7.5 to <8.0%	-0.32 (0.04)	-0.25 (0.10)	-0.16 (0.03)	0.15 (0.05)
8.0 to 9.0%	-0.61 (0.04)	-0.23 (0.08)	-0.33 (0.03)	-0.14 (0.05)
≥ 9.0%	-1.18 (0.04)	-0.88 (0.08)	-0.93 (0.02)	-0.77 (0.05)
Mean (SE) Change in Weight (kg)$^{\wedge}$	-1.0 (0.07)	-0.20 (0.2)†	-0.74 (0.05)	-1.46 (0.1)
Mean (SE) Change in TC (mmol/l)$^{\wedge}$	-0.17 (0.01)	-0.06 (0.03)†	-0.1 (0.01)	-0.12 (0.02)

Abbreviations: Pr. (Predicted probability of response); SE (Standard Error); TC (Total cholesterol);
$^{\wedge}$ Absolute change; † P value > 0.05; $^{\infty}$ Estimated as time from first glucose-lowering therapy

The probability of DPP-4 inhibitor treatment response according to different baseline HbA1c levels was also examined. In terms of absolute changes in HbA1c at 12 months, intensification with a DPP-4 inhibitor was associated with HbA1c reduction between 0.2 and 0.6% across the treatment groups. Table 4.3 summarises the overall reductions in HbA1c, body weight and total cholesterol across the respective treatment groups.

Figure 4.3 Proportion of patients achieving <7% HbA1c target at 1 year

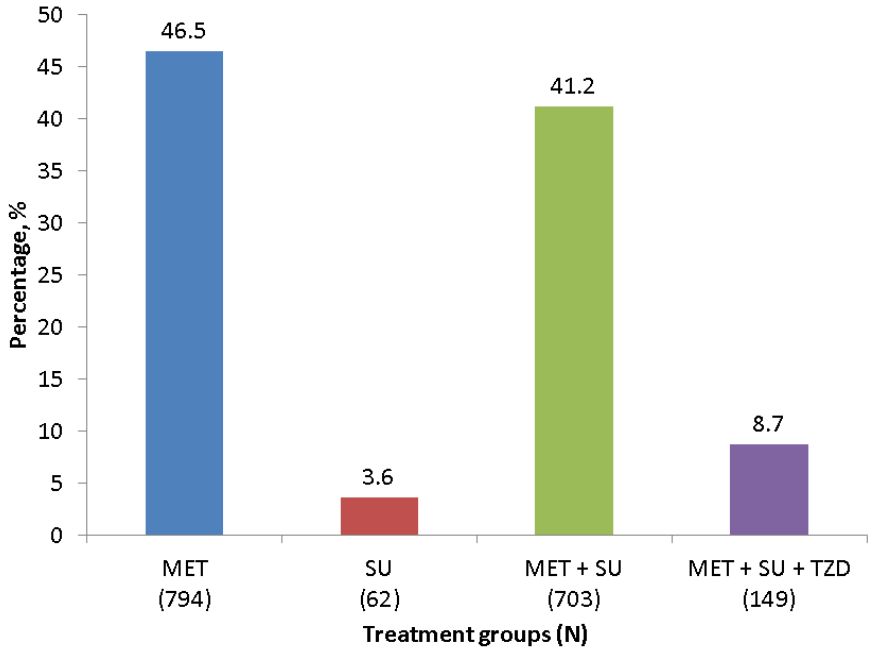


Figure 4.3 shows the proportion of patients who achieved HbA1c target <7%, after intensification with DPP-4 inhibitor. The addition of DPP-4 inhibitor to ongoing metformin monotherapy vs metformin plus sulphonylurea accounted for 47% vs 41% achieving target, respectively.

Descriptive analysis show the proportion of patients who achieved HbA1c target of <7.0% at 1 year. The data showed adding DPP-4 inhibitor to monotherapy involving MET vs MET + SU resulted in 47 vs 41%, respectively of users met the target as compared to 4% of SU only users (Figure 4.3). In a subgroup of patients with suboptimal HbA1c above 7.5%, results show similar proportion of patients met a target below 7% when DPP-4 inhibitor was added to dual MET + SU regimen and MET only regimen (45 vs 43%, respectively, Figure 4.4).

Figure 4.4 Proportion with HbA1c ≥7.5% achieving <7% HbA1c target

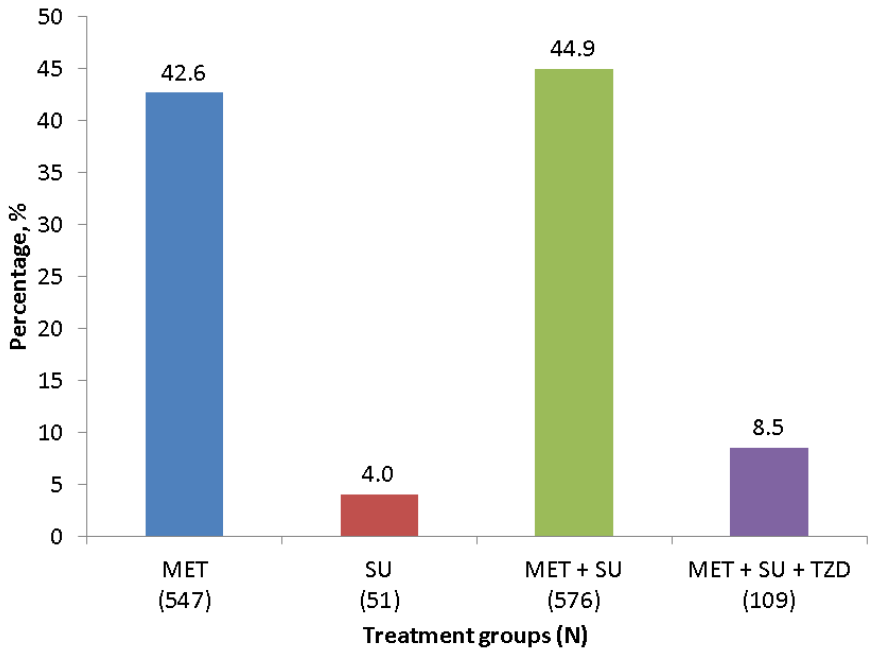


Figure 4.4 shows the proportion of patients who achieved HbA1c target <7%, after intensification with DPP-4 inhibitor among patients with HbA1c ≥7.5% at baseline. The addition of DPP-4 inhibitor to ongoing metformin monotherapy vs metformin plus sulphonylurea therapy resulted in similar proportion of patients achieving HbA1c target (45% vs 43%, respectively).

4.5 Discussion

Overall, this large cohort study in primary care showed a significant 0.5% reduction in HbA1c up to 12 months after patients with suboptimal HbA1c from various oral glucose-lowering therapies were co-administered a DPP-4 inhibitor as add-on treatment. The addition of DPP-4 inhibitor to MET was found to be the most effective in terms of glycaemic response. Co-administering DPP-4 inhibitor with MET + SU therapy was also associated with responders. Conversely, higher HbA1c at baseline and longer diabetes duration was independently associated with less likelihood of achieving Hba1c target of <7%.

Despite the improvement in mean HbA1c following the addition of a DPP-4 inhibitor to on-going treatment, only 13% of patients 'responded' to treatment intensification. This relatively low percentage of responders reflects the difficulties in achieving HbA1c target in a challenging cohort of patients who have failed to achieve optimal glucose levels with other oral glucose-lowering drugs. The co-administration of a DPP-4 inhibitor with MET independently predicted response to therapy and resulted in a significant ($p<0.001$) reduction of HbA1c (-0.6%), body weight (-1.0 kg) and total cholesterol (0.2 mmol/l) at 12 months. Similar results were obtained in previous systematic reviews and meta-analysis of RCTs where treatment with sitagliptin + MET alone was found to be more effective at improving HbA1c levels than MET alone.[151] A review by Chatterjee[152] compared DPP-4 inhibitors with Metformin monotherapy, and DPP-4 inhibitors + Metformin with other glucose-lowering

drugs (e.g., sulfonylurea, basal insulin, pioglitazone, and GLP-1 agonist) . The study reported that DPP-4 inhibitor monotherapy was less effective in reducing HbA1c levels and weight than Metformin alone.[152] This, and data derived from our study show that DPP-4 inhibitor have a better response when prescribed in combination with Metformin. Evidence has shown that Metformin increases GLP-1 secretion,[153] which may account for the observed synergistic effects of a DPP-4 inhibitor with Metformin.

Interestingly, despite the neutral effects of DPP-4 inhibitors on body weight[154] and evidence showing ~90% inhibition of plasma DPP-4 activity and an approximate 3 fold increased in active GLP-1 level with sitagliptin in obese patients with diabetes,[155] BMI did not play any role in determining whether adding DPP-4 inhibitor in routine clinical practice would result in achieving Hba1c target.

In contrast to RCT evidence showing the efficacy of DPP-4 inhibitors when used as an add-on therapy to SU,[156,157] real world data observed in this study suggests that concurrent use of a SU is a predictor of non-response to DPP-4 inhibitor treatment intensification. In addition, this study has shown the addition of DPP-4 to SU compared to the other regimen examined is associated with the lowest proportion of patients achieving HbA1c target below 7% and the least HbA1c reduction. These discordant results may be explained by the delay in treatment intensification and longer diabetes duration that is often seen in routine clinical practice as oppose to RCT recruits. In a previous study using the THIN database, in patients with T2D, after failure of glycaemic

control with oral GLT, insulin initiation was delayed for at least 1.8 years in 25% of cases, and for almost 5 years in 50% of cases.[158]

Studies examining the use of DPP-4 inhibitor as a third or fourth line therapy, for example, regimens involving the combination of MET, SU and TZD, are lacking. A recent study [159] showed initial combination therapy with sitagliptin and pioglitazone yielded significantly greater reductions in HbA1c (between 0.4 and 0.7%) than monotherapy of either drug. Combination therapy was found to be generally well tolerated; however, hypoglycaemia and weight gain were reported in all treatment groups compared with the sitagliptin monotherapy group over the 54 weeks of the study. This study showed that addition of a DPP-4 inhibitor to the MET+SU+TZD regimen resulted in the least reduction of HbA1c among patients with baseline HbA1c levels above 8%. Crude odds ratios suggest adding a DPP-4 inhibitor to this triple therapy regimen was not associated with any significant response. This may reflect increased disease duration, where the use of insulin may be the most appropriate treatment choice in many patients.

Limitations of study

The analysis conducted in this study was subject to some limitations inherent to observational studies; for example, the exposure data relates to prescriptions so it could not be ascertained whether glucose-lowering medications were actually used. However, should there be any overestimation of exposure to the medications in our analysis, such a misclassification would be non-differential and only bias results towards unity. Potential residual confounders such as ethnicity, compliance, indications for different drug treatments, compliance

and differences in dosages administered to patient groups were not accounted for. In addition, data obtained for this study lacked information on the date of diagnosis of diabetes. Therefore, a proxy for diabetes duration was used by considering earliest date of first glucose-lowering therapy prescription as a surrogate of date of first diagnosis, and include in the model for analysis. The estimated duration of diabetes calculated suggests individuals on triple regimen (MET + SU + TZD) appeared to have lower diabetes duration (1.2 years) compared to individuals on monotherapy or dual therapy at the time of intensification. This disparity raises concern on the accuracy of using estimated diabetes duration as oppose to actual diabetes duration in this model. A sensitivity analysis after excluding the proxy diabetes duration from the model yielded similar results. Despite these limitations, the study highlights the effectiveness of DPP-4 inhibitor therapy as an add-on to MET in real-world practice. This study has demonstrated how simple clinical and demographic parameters may influence outcomes following DPP-4 inhibitor therapy among patients with suboptimal glucose control.

Conclusion

This retrospective cohort study based on UK primary care data assessed a range of factors that may be associated with glycaemic response or non-response to treatment intensification with a DPP-4 inhibitor. The study hypothesised that HbA1c may not be the only factor that could determine glycaemic responsiveness. The results showed that the addition of DPP-4 inhibitor to MET or MET + SU therapy was associated with positive

glycaemic response, while higher HbA1c at baseline and longer diabetes duration were independently associated with less response. With the ongoing uncertainty regarding the optimal second or third-line treatment option, data from this study presents additional evidence on the benefit of DPP-4 inhibitor and support the use of a DPP-4 inhibitor as a 2nd or 3rd line therapeutic option among patients whose glucose control remains suboptimal despite MET or MET plus SU treatment. In view of the potential long-term beneficial effects of DPP-4 inhibitor on β-cell function[160] as well previous study in a different ethnic group,[161] this study supports the use of DPP-4 inhibitor in patients with T2DM.

Chapter 5: Comparative effect of a DPP-4 inhibitor as add-on therapy

5.1 Summary

Aim

The aim of this study was to assess the glycaemic effectiveness of co-administering sitagliptin to patients with inadequate glycaemic control following treatment with MET, SU, or MET + SU.

Methods

A cohort of 25,386 patients with T2DM (HbA1C > 7%), newly treated with sitagliptin between 2007 and 2013, was sourced from UK general practices via THIN database. Among these, eligible patients were segregated into three groups: MET (n = 3,364), SU (n = 509), or MET + SU therapy (n = 5,929). The relative efficacy of sitagliptin added to SU or MET + SU compared with sitagliptin added to MET monotherapy was assessed with regards to HbA1c and body weight changes from baseline up to 52 weeks. The glycaemic efficacy was a measure of average treatment effects obtained from multivariable linear regression models and propensity score weighting analysis.

Results

A total of 9,802 patients were included in the study. Overall, addition of sitagliptin 100 mg once daily resulted in 0.5% (5.5mmol/mol) HbA1c reduction ($P < 0.001$) and 0.8 kg weight reduction at 1 year ($P < 0.001$). Efficacy was similar across the treatment groups, but in patients with baseline HbA1c ≥ 9% adding sitagliptin to MET + SU produced a significantly smaller reduction in HbA1c when compared to the reference group MET (MET + SU vs. MET only: −0.5% vs. −0.7%, $P < 0.001$). The mean HbA1c reduction from baseline within this subgroup of patients was not significantly different between SU and MET monotherapies (−0.8% vs. −0.7%, respectively, $P = 0.4$). Across treatment groups, HbA1c reductions with add-on sitagliptin occurred after 24 weeks of treatment with a peak reduction occurring between 36–48 weeks, and receded after week 48.

Conclusion

In real world general practice setting, sitagliptin was effective in patients with suboptimal glycaemic control with MET, SU or dual therapy, maximum between 36-48 weeks, but in patients with HbA1c of >9% receiving MET+SU therapy, adding sitagliptin as a 3rd agent conferred minimal benefit.

5.2 Introduction

The majority of patients with T2DM eventually require combination therapy to control hyperglycaemia as their disease progress.[1] To this end, the use of combination therapies from different classes that have complementary mechanisms of action is recommended to facilitate more effective lowering of

blood glucose levels.[162] The combination of MET and SU is the most widely used dual combination GLT in patients with T2DM.[163] However, combination therapy with these two agents may also not achieve or maintain glycaemic control,[164] necessitating the need for further treatment intensification. In such situations, the use of injectable therapy such as insulin or GLP-1 receptor agonist is often the next therapeutic step, although triple GLT (e.g., adding a TZD to ongoing dual therapy with MET and a SU) is also used in clinical practice as shown in Chapter 4:. However, many patients find the need for insulin injection or the adverse effects of oedema and/or an increase in body weight with TZD to be undesirable, which may adversely affect treatment compliance and glycaemic response.[165] Hence, there is a need for additional options that can be added to MET and SU to avoid the need to switch to insulin. While RCTs have examined the efficacy of various combination therapies, comparative efficacy data from routine real-world clinical practice could yield important and complimentary clinical information that needs be taken into account when determining treatment strategies.[166]

Sitagliptin is a once-a-day orally-active DPP-4 inhibitor which has been administered to improve glycaemic control in patients with T2DM treated as add-on therapy to MET or to SU monotherapy as well as add-on to MET + SU combination therapy.[167-169] Real-world studies on the comparative efficacy of the co-administration of sitagliptin with MET, SU, or dual MET and SU therapy have not been reported. This is relevant in view of the fact that, although both sitagliptin and SU stimulate insulin secretion from pancreatic β-cells,[131,170] sitagliptin, unlike SU, also lowers glucagon concentrations,[171]

which is likely to also contribute to the glucose-lowering obtained with this agent. A previous RCT has shown that sitagliptin is effective when used as add-on combination treatment with MET and SU therapy.[168] The study showed that, the co-administration of sitagliptin with MET + SU resulted in a HbA1c reduction of 0.9% relative to placebo, and a reduction of 0.6% when sitagliptin is added to SU monotherapy, relative to placebo. However, the effectiveness of sitagliptin in real-world practice has not been reported. Furthermore, within this setting, if sitagliptin is effective in combination with an SU then triple combination therapy with MET and an SU is likely to be effective as well.

The aim of this study, therefore, is to report the glycaemic response and treatment effect of sitagliptin when added to MET, SU, or MET + SU combination therapy in routine clinical practice. In order to address the influence of bias from confounders, the glycaemic efficacy of sitagliptin co-administration was evaluated using multivariable linear regression and propensity score weighting analysis.

5.3 Methods

5.3.1 Study design and data Source

Retrospective cohort analyse was conducted on the primary care data, which contains anonymous patient information from general practices throughout England and Wales.[170]

Study population

The study population comprised a cohort of patients identified to have T2DM and registered to a practice for more than 12 months before the index date (i.e.,

between January 1, 2006 and the end of the study on May 30, 2013). The cohort included patients who had inadequate glycaemic control (HbA1c levels ≥ 53 mmol/mol (7%) after 6 months of MET monotherapy, SU monotherapy, or dual therapy consisting of both MET and SU. Clinicians have debated whether DPP-4 inhibitors exert better glycaemic effectiveness as second- or third-line therapy in the absence of a consensus treatment strategy. Clinical evidence has also shown that sitagliptin is the most widely used DPP-4 inhibitor in UK primary care, especially as a second line option to MET and as third-line option after MET + SU combination therapy has failed to achieve glycaemic control. The latter might also be as a result of the extensive use of SU as a second-line option for along period. These reasons influenced the decision to exclude individuals using TZD as well as patients who were concurrently taking other GLTs such as GLP-1 receptor agonist, SGLT-2 inhibitors, glinides, and acarbose from this study. In addition, patients with any records of insulin prescription and those taking sitagliptin as monotherapy or those taking another type of DPP-4 inhibitor were excluded. Sitagliptin was selected because of its high prescription rate (nearly 75% of DPP-4 inhibitor prescription) in primary care. Concurrent lipid-lowering drugs, aspirin, and antihypertensive medications were allowed.

5.3.2 Treatment exposure

Patients were administered an average of 100 mg/day of sitagliptin and the follow-up period commenced from the index date (the date of the first sitagliptin prescription) until a switch to or addition of another GLT, or the 90th day post index date when HbA1c level is recorded, or 52 weeks after the

index date. Patients were segregated into the following treatment groups based on the oral antidiabetic treatments they received at baseline: MET, SU monotherapy and MET + SU. A parallel-group study involving the underlying treatment groups was set up with MET monotherapy group serving as the comparison or reference group.

5.3.3 Outcome

The primary glycaemic effectiveness outcome was change from baseline in HbA1c at 52 weeks. Secondary outcome was change from baseline in body weight. The glycaemic effectiveness of a treatment regimen is a measure of average treatment effect (ATE) exhibited by SU and MET + SU treatment groups when compared with MET monotherapy, the reference group.

Covariates

Covariates were selected a priori on the basis of clinical significance. These have been previously described in Section 3.4.3.

5.3.4 Statistical analysis

Analysis on the primary outcome of sitagliptin as an add-on therapy assessed the treatment groups for superiority with regard to the average HbA1c change from baseline at their respective endpoints. Multinomial propensity scores on the baseline covariates were estimated.[131] Balance in baseline covariates was assessed between the treatment groups using absolute standardized differences before and after propensity score weighting. A standardized effect size $\geq 20\%$ indicated serious imbalance. The variations in mean and frequency distribution of measured baseline covariates between treatment groups with the same estimated propensity score was examined and summarized.

Propensity score model

Inverse probability of treatment weighting (IPTW) using the propensity score was employed to estimate the measures of effect.[105,172] The method permitted the ATE on the population to be estimated, thereby enabling the ascertainment of glycaemic responses if patients receiving SU + sitagliptin had been assigned to receive MET + SU before the addition of sitagliptin, relative to if they had all received MET + sitagliptin (reference group). Propensity score was considered as a prognostic covariate and included in the multivariable linear regression model. Average changes in HbA1c were calculated and expressed as point estimates with 95% CI, at the conventional statistical significance level of 0.05. Analyses were conducted using R[144] and Stata[143] packages.

5.3.5 Secondary and subgroup analyses

Baseline HbA1c was categorized into four strata: 7 to <7.5%, ≥7.5 to <8%, ≥8 to <9%, and ≥9%. Subgroup analysis for efficacy in end point changes from baseline in HbA1c was performed across the treatment groups. In addition, correlation and linear regression analysis was performed to assess the relationship between changes in HbA1c and changes in weight at 52 weeks in the study population.

5.3.6 Bias

New users of sitagliptin were assessed at baseline to ensure individuals included in the study have not been treated with sitagliptin or any other DPP-4 inhibitor before the follow-up commenced, thereby introducing prevalent user bias. Selection bias was minimised by ensuring all participants in the exposed treatment groups examined initiated sitagliptin as new users.

In addition, individuals whose HbA1c was <7% at the time of intensification with sitagliptin were excluded as shown in Figure 5.1. This is because lower HbA1c may have been achieved as a result the effect of previous GLTs. Missing information of baseline characteristics on all individuals involved in the study were accounted for to further curtail selection bias.

To address the influence of bias from confounding, the glycaemic efficacy of sitagliptin co-administration was evaluated using multivariable linear regression and propensity score-weighted analysis.

Post index date exposure to any other glucose-lowering therapy other than a sitagliptin was not permitted. This was to reduce confounding by indication or misclassification bias. In addition, individuals were segregated into separate combination treatment groups to prevent confounding by co-medication. The cohort was restricted to an estimated 12 months follow-up to reduce the risk of bias introduced by an overlapping treatment effect.[114]

5.4 Results

5.4.1 General patient characteristics

Figure 5.1 Study population screening and selection process

Of the 25,386 users of DPP-4 inhibitor who were screened, 9,802 (39%) patients fulfilled the criteria for cohort entry and were assigned to one of three treatment groups as outlined in Figure 5.1. The number of patients assigned to each treatment group include: 3,364 (34%) previously on MET alone, 509 (5%) on SU alone, and 5,929 (61%) previously on dual therapy (MET + SU) regimen.

Table 5.1 Baseline characteristics of patients treated with sitagliptin

	Cohort				
Baseline variable	**MET alone (*n* = 3,364)**	**SU alone (*n* = 509)**	**MET + SU (*n* = 5,929)**	**ES[a]**	**ES[b]**
Demographics					
Age (years), mean (SD)	61.7 (12.2)	61.5 (12.5)	61.7 (12.5)	0.01	0.00
Gender, *n* (%)					
Male	1,988 (59)	294 (58)	3,516 (59)	0.03	0.00
Female	1,376 (41)	215 (42)	2,413 (41)	0.03	0.00
Townsend deprivation, *n* (%)					
Least deprived	738 (22)	104 (20)	1,225 (21)	0.04	0.01
Less	707 (21)	98 (19)	1,247 (21)	0.04	0.00
Average	708 (21)	116 (23)	1,261 (21)	0.04	0.00
More	645 (19)	104 (20)	1,233 (21)	0.03	0.02
Most deprived	566 (17)	87 (17)	963 (16)	0.02	0.01
Clinical parameters, mean (SD)					
HbA1c (%)	8.8 (1.4)	8.8 (1.4)	8.8 (1.4)	0.01	0.00
HbA1c category, % (mmol/mol)					
7–7.5 (53–58)	610 (18)	91 (18)	1,012 (17)	0.02	0.01
7.5–8 (58–64)	629 (19)	102 (20)	1,133 (19)	0.03	0.01
8–9 (64–75)	1,001 (30)	134 (26)	1,754 (30)	0.05	0.02
≥ 9 (75)	1,124 (33)	182 (36)	2,030 (34)	0.01	0.01
BMI (kg/m^2)	32.8 (6.8)	32.5 (6.9)	32.6 (6.6)	0.05	0.01
Weight (kg)	93.9 (21.6)	92.8 (21.5)	93.3 (21.1)	0.05	0.01

	Cohort				
Baseline variable	**MET alone (*n* = 3,364)**	**SU alone (*n* = 509)**	**MET + SU (*n* = 5,929)**	**ES[a]**	**ES[b]**
SBP (mmHg)	134 (15.1)	133.5 (15.2)	134.5 (15.1)	0.05	0.00
DBP (mmHg)	77.4 (9.4)	76.7 (9.1)	77.2 (9.5)	0.08	0.01
TC (mmol/L)	4.3 (1.1)	4.3 (1.1)	4.3 (1.1)	0.02	0.01
HDL-C (mmol/L)	1.1 (0.3)	1.1 (0.3)	1.1 (0.3)	0.02	0.01
LDL-C(mmol/L)	2.3 (0.9)	2.3 (0.9)	2.2 (0.9)	0.01	0.00
Triglyceride (mmol/L)	2.2 (2.6)	2.2 (1.8)	2.1 (1.7)	0.02	0.02
GLT duration (years)	1.6 (2.7)	1.6 (2.7)	1.6 (2.7)	0.06	0.00
Smoking status, *n* (%)					
Non-smoker	1,333 (40)	195 (38)	2,379 (40)	0.01	0.01
Current smoker	494 (15)	76 (15)	859 (14)	0.01	0.01
Ex-smoker	1,537 (46)	238 (47)	2,691 (45)	0.01	0.01
BMI category, *n* (%)					
Normal	289 (9)	47 (9)	556 (9)	0.01	0.01
Overweight	985 (29)	161 (32)	1,690 (29)	0.08	0.01
Obese	2,090 (62)	301 (59)	3,683 (62)	0.03	0.00
Use of medications, *n* (%)					
Aspirin	1,306 (39)	208 (41)	2,361 (40)	0.02	0.01
Antihypertensive	2,482 (74)	363 (71)	4,332 (73)	0.06	0.00
LLT	2,636 (78)	394 (77)	4,658 (79)	0.05	0.01
Comorbidities, *n* (%)					
CHD	1,936 (58)	293 (58)	3,392 (57)	0.04	0.02

	Cohort				
Baseline variable	**MET alone (*n* = 3,364)**	**SU alone (*n* = 509)**	**MET + SU (*n* = 5,929)**	**ES[a]**	**ES[b]**
PAD	536 (16)	85 (17)	940 (16)	0.06	0.02
Cerebrovascular	767 (23)	126 (25)	1,340 (23)	0.02	0.00
Heart failure	350 (10)	56 (11)	679 (11)	0.00	0.01
Hypoglycaemia	667 (20)	105 (21)	1,130 (19)	0.02	0.00
Follow-up (weeks)					
0–12	383 (11)	51 (10)	676 (11)	0.05	0.02
12–24	370 (11)	49 (10)	644 (11)	0.02	0.00
24–36	339 (10)	48 (9)	593 (10)	0.02	0.00
36–48	826 (25)	140 (28)	1,502 (25)	0.04	0.00
48–52	1,446 (43)	221 (43)	2,514 (42)	0.07	0.01

GLT duration is the duration of treatment from first GLT.

ES is the absolute standardized mean difference of means or percentages divided by the standard deviation

[a]ES in unweighted; [b]ES in propensity score weighted cohort based on average treatment effect in the population

Table 5.2 ATE of sitagliptin as add-on to ongoing oral GLT

			Estimated treatment difference (95% CI)			
Variables	**MET (reference)**	***P* value**	**SU vs. MET**	***P* value**	**MET + SU vs. MET**	***P* value**
HbA1c change[a],						
%	−0.49 (−0.53, −0.45)	<0.001	−0.03 (−0.14, 0.09)	0.6	0.03 (−0.02, 0.08)	0.3
mmol/mol	−5.4 (−5.8, −4.9)		−0.3 (−1.3, 1.0)		0.3 (0.2, 0.9)	
HbA1c subgroup[b]						
7–7.5%	−0.33 (−0.44, −0.22)	<0.001	0.05 (−0.21, 0.31)	0.7	0.03 (−0.08, 0.15)	0.6
53–58 mmol/mol	−3.6 (−4.8, 2.4)		0.6 (−2.5, 3.7)		0.3 (−1.0, 1.8)	
7.5–8%	−0.37 (−0.46, −0.27)	<0.001	-0.05 (−0.30, 0.19)	0.7	−0.01 (−0.12, 0.11)	0.9
58–64 mmol/mol	−4.0 (−5.0, −3.0)		-0.6 (−3.6, 2.3)		−0.1 (−1.4, 1.3)	
8–9%	−0.46 (−0.53, −0.38)	<0.001	−0.01 (−0.22, 0.20)	0.9	−0.02 (−0.11, 0.07)	0.6
64–75 mmol/mol	−5.0 (−5.8, −4.2)		−0.1 (−2.6, 2.4)		−0.2 (−1.3, 0.8)	
≥9%	−0.68 (−0.77, −0.59)	<0.001	0.08 (−0.11, 0.26)	0.4	0.18 (0.16, 0.31)	0.01
≥75 mmol/mol	−7.4 (−8.4, −6.4)		1.0 (−1.3, 3.1)		2.2 (1.9, 3.7)	
ATE			0.02 (−0.09, 0.12)	0.7	0.03 (−0.02, 0.08)	0.2
Weight change, kg	−0.93 (−1.09, −0.78)	<0.001	0.14 (−0.28,0.56)	0.5	0.14 (−0.05,0.33)	0.2

[a] Change in HbA1c from PS weighted linear regression model; [b] Least square mean difference from PS weighted linear regression model

The patients had a mean age of 62 years and were predominantly male (60%), obese (BMI > 30 kg/m^2, 62%), and on various anti-hypertensive medication (73%). The average follow-up time was 38 weeks and there was no significant difference in baseline demographic and metabolic characteristics of patients between the treatment groups (Table 5.1).

5.4.2 Effectiveness

Overall, the co-administration of sitagliptin to patients who had inadequate glycaemic control from ongoing MET, SU, and MET + SU regimen resulted in a significant 5.5 mmol/mol (0.5%) reduction in HbA1c (P < 0.001) and a 0.8 kg reduction in body weight (P < 0.001) (Table 5.2). The average HbA1c and weight reductions across the treatment groups were generally similar.

Figure 5.2 Changes in HbA1c up to 52 weeks after adding sitagliptin

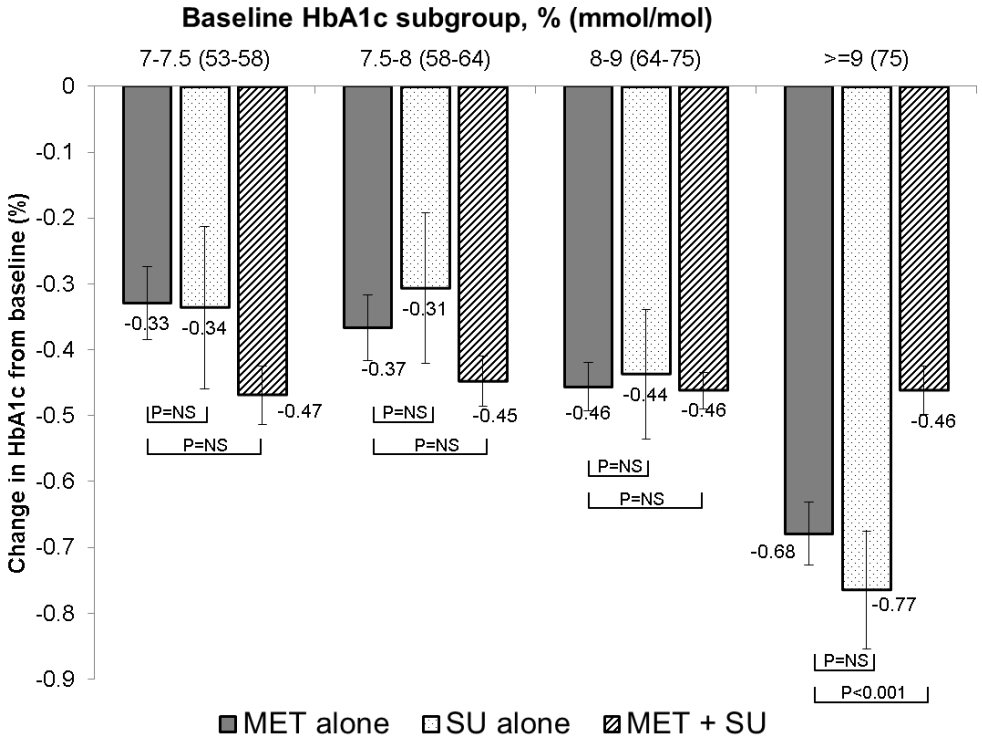


Figure 5.2 shows changes in HbA1c at 52 weeks after the co-administration of sitagliptin

Treatment effect

The ATEs with regards to HbA1c reduction produced by the co-administration of sitagliptin with SU monotherapy and with MET + SU did not show any change in HbA1c value (0.02%, P = 0.7 and 0.03%, P = 0.2, respectively; Table 5.2) However, when stratified according to levels of HbA1c at baseline, a significant difference in the treatment efficacy was observed in the subgroup of HbA1c ≥ 9% at baseline. In this HbA1c subgroup, after adjusting for confounders which include duration of GLT prior to starting sitagliptin, glycaemic efficacy was significantly greater among patients on MET monotherapy compared with their counterparts using MET + SU (−0.7% vs. −0.5%, respectively, P < 0.001; Figure 5.2). The mean reduction from baseline in HbA1c was not significantly different between SU monotherapy users and the MET monotherapy reference group (−0.8% vs. −0.7%, P = 0.4; Table 5.2). Hence, adding sitagliptin to MET + SU dual therapy did not confer additional glucose-lowering effects compared with co-administration of sitagliptin with MET or SU monotherapies.

Overall, after adjusting for confounders, the co-administration of sitagliptin produced a glycaemic effect that appeared to increase over time in both treatment and reference groups. However, this effect was not sustained throughout the study period, independent of all treatment groups (Figure 5.3). HbA1c reduction was observed to take effect after 24 weeks of treatment with sitagliptin, with a peak reduction between week 36 and 48 and receded after week 48. Although adding sitagliptin to MET monotherapy initially appears to produce a better onset of effect compared with MET+SU data from this study

show that, the adjusted mean changes from baseline were not significantly different between the treatment and reference groups.

Figure 5.3 Changes in HbA1c over time during the 52-week follow-up

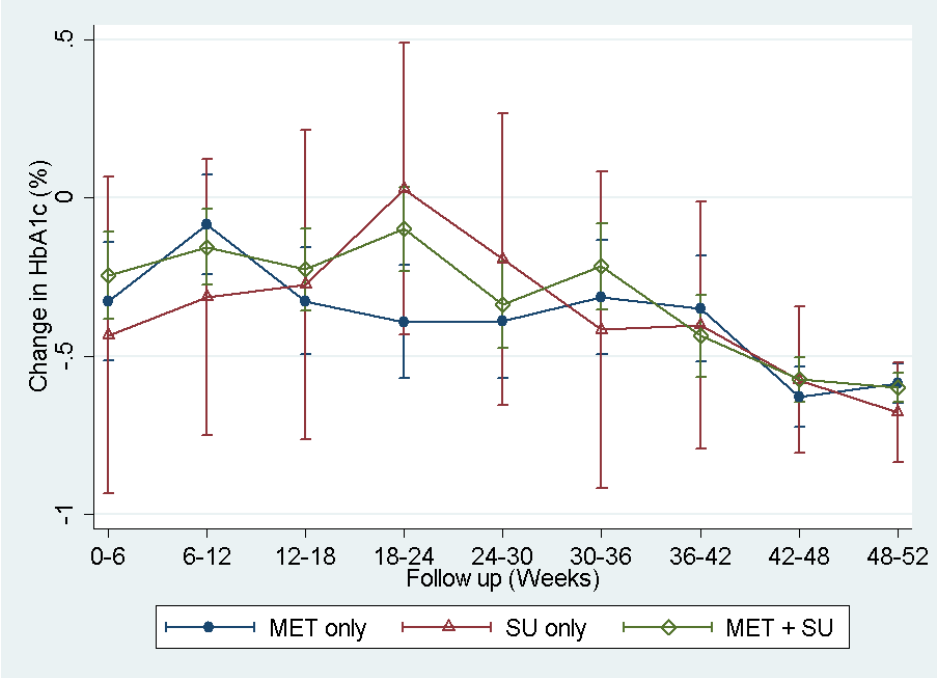


Figure 5.3 show the glycaemic response after intensification with sitagliptin appeare to improve after 24 weeks of co-administration with SU only and after 36 weeks of co-administration with MET only or MET + SU dual therapy. Wide error bars also highlight the small sample size of SU only group.

5.4.3 Other analyses

Figure 5.4 Probability density function of PS weighted cohort

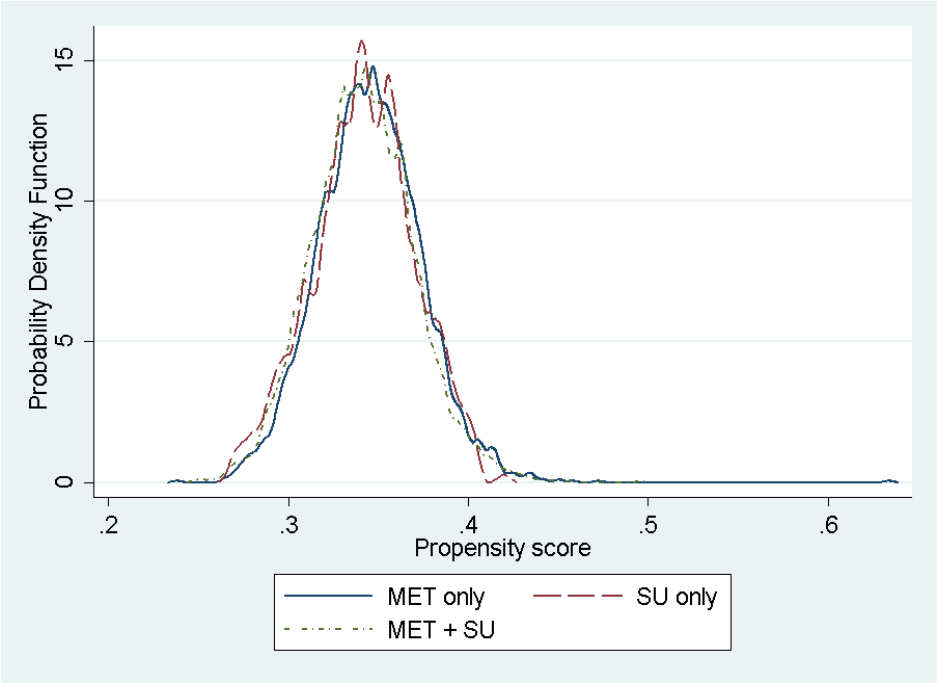
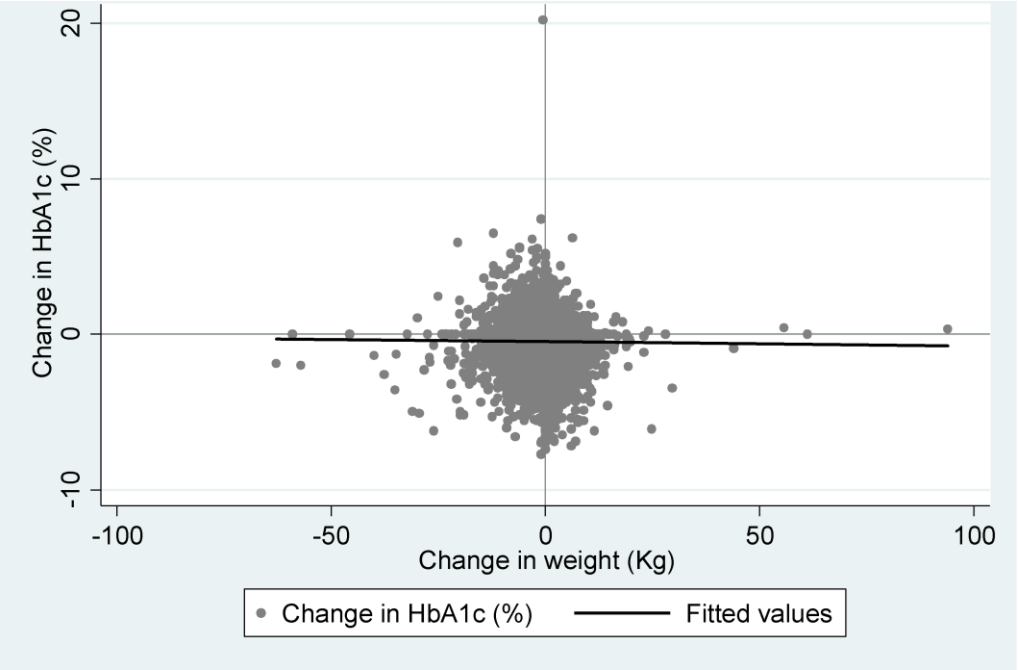


Figure 5.4 shows the distribution of the propensity score weighting for treatment groups MET (n=3,364) vs. SU (n=509) and MET + SU (n=5,929).

The distribution of the propensity score in Figure 5.4 shows no indication of probability mass near 1 or 0 and the respective estimated densities of propensity score in the covariate distribution between the treated groups overlap each other. There is therefore no violation of the overlap assumption.[173]

Figure 5.5 Relationship between changes in HbA1c and body weight



The scatter plot of individual patient data shows no association between change in HbA1c and change in weight from baseline to endpoints. (Pearson's correlation coefficient, r = −0.01; P = 0.3; Figure 5.5) Therefore, the changes in HbA1c observed in the population does not account for the variation in weight change. The sensitivity analysis after multiple imputation showed similar results to our complete case models which indicate findings are unlikely attributable to bias from missing information. (Appendix C-1)

5.5 Discussion

Comparative effectiveness studies which examine the efficacy of the co-administration of sitagliptin to SU or MET + SU are not widely reported. Even where RCTs were carried out, the lack of rigorous patient inclusion and exclusion criteria such as the ones explored in this study may limit the generalizability of study findings. Overall, this study showed the addition of

100 mg/day of sitagliptin to patients with T2DM with inadequate glycaemic control following MET monotherapy, SU monotherapy or both, resulted in a 5.5 mmol/mol (0.5%) reduction in HbA1c and a 0.8 kg weight loss at endpoint. The average HbA1c and weight reductions across the treatment groups were generally similar except within a subgroup of patients who had HbA1c ≥ 9% at baseline, where the co-administration of sitagliptin with MET+SU did not confer additional significant glucose-lowering, even after adjusting for a proxy of diabetes duration. Thus, adding sitagliptin to SU confers equivalent benefit in Hba1c lowering compared with adding to MET, but the use of sitagliptin in combination with SU and MET therapy, is not efficacious. Since the glycaemic effectiveness of sitagliptin co-administration was analysed using multivariable linear regression, absolute comparison between treatment groups was not able to be performed.

Interestingly, the latter finding contradicts findings from a RCT which showed additional HbA1c reduction with sitagliptin when added to MET plus glimepiride therapy.[168] The 24-week trial conducted on 229 individuals with T2DM showed that, the co-administration of sitagliptin with glimepiride plus metformin resulted in a HbA1c reduction of 0.9% relative to placebo, and a reduction of 0.6% when sitagliptin is co-administered with glimepiride monotherapy (n=212), relative to placebo. This discrepancy may be as a result of longer follow-up period and other unexplainable factors that may be associated with real-world data. In addition, results from this study were based on active comparator MET monotherapy group as oppose to placebo used in the RCT. Based on the availabilities of newer therapeutic agents and other

injectable therapies such as insulin or GLP-1 analogue, the merit of using sitagliptin to manage hyperglycaemia as a triple therapy in routine practice remains questionable. However, the observed equal benefit in HbA1c reduction when sitagliptin was added to patients who have failed SU therapy (compared with MET-sitagliptin combination therapy) implies an additional mechanism of action of sitagliptin therapy, above and beyond its ability to stimulate insulin secretion from an already exhausted pancreatic β-cells, such as GLP-1 and glucose-dependent insulinotropic peptide (GIP)-induced suppression glucagon secretion.[171] However, results obtained from previous systematic reviews and meta-analysis of RCTs studies have shown greater reduction in HbA1c (-0.55%, 95% CI: -0.63 to -0.46%) among individuals administered DPP-4 inhibitor plus MET compared with MET monotherapy.[174] The synergistic effect of sitagliptin with MET is increasingly well recognized and may be explained by the fact that MET enhances the expression and production of GLP-1 from the terminal ileum.[175]

Another important and novel observation from this study relates to the durability of sitagliptin therapy. As a whole, across the treatment group, HbA1c reduction was observed to take effect after 24 weeks of treatment with sitagliptin, with a peak reduction between week 36 and 48 and receded after week 48. This is in contrast to most findings from RCT, where peak HbA1c reduction seemed to occur earlier, at approximately six weeks post initiation of sitagliptin. It is assumed that the reason for the delay may be due to the deterioration of β-cell function as diabetes progresses. It was not possible to adjust for β-cell markers and other factors that might indicate a more certain

reason for the slower response observed. Nonetheless, most RCTs investigating the efficacy of DPP-4 inhibitors with SU or MET have reported outcomes from shorter follow-up period compared to this study, although one study using saxagliptin in combination with glyburide, followed up patients for 76 weeks.[176] In the study by Chacra et al, HbA1c reduction occurred immediately upon initiation of saxagliptin, peak reduction after 8–12 weeks, with a further rise in Hba1c thereafter, returning back to baseline at 76 weeks, which may reflect the progressive nature of diabetes.[176] However, in the two longest-running trials of DPP-4 inhibitors the 'escape phenomenon', assessed by a secondary increase in HbA1c levels between weeks 24 and 104 following a good initial HbA1c reduction, was significantly less pronounced with sitagliptin 100 mg or vildagliptin 100 mg than with glipizide or glimepiride, respectively,[177,178] suggesting better β-cell protection and durability of glucose control with a DPP-4inhibitor. The 'escape phenomenon' refers to a resistance that develops from the effect of a continuously present stimulus. A more recent 52-week RCT comparing sitagliptin versus canagliflozin (a SGLT-2 inhibitor) when added to MET + SU, showed maximum HbA1c reduction at 12 weeks and a progressive rise in Hba1c thereafter.[179] Data from the present study has shown glycaemic response to sitagliptin for up to 52 weeks after treatment intensification. Assessing a 52-week outcome does not fit with the NICE criteria which suggest treatment should be discontinued after 6 months, unless the HbA1c has fallen by 0.7%. The 0.5% reduction in HbA1c obtained from this study is below the recommended threshold and raises the concern for adherence to NICE guidelines. This data also shows that HbA1c reduction was

improved after 6 months of intensification with sitagliptin and highlights the need for further investigation into extended treatment duration in some patient groups.

Limitations of study

This study was subject to limitations highlighted in Section 4.5. Potential residual confounders such as ethnicity, compliance, indications for different drug treatments, compliance and differences in dosages administered to patient groups were not accounted for. In addition, data obtained for this study also lacked information on the date of diagnosis of diabetes. Therefore, a proxy for diabetes duration was used by considering earliest date of first glucose-lowering therapy prescription as a surrogate of date of first diagnosis, and include in the model as GLT duration. Furthermore, the number of events of hypoglycaemia after MET use highlights challenges on the reliability of data on hypoglycaemia in the database. Hypoglycaemia was defined as any previous episodes of hypoglycaemia recorded prior to intensification with DPP-4 inhibitor. Despite these limitations, the study highlights the glycaemic response of DPP-4 inhibitor therapy as add-on therapy in routine clinical practice.

In summary, the addition of sitagliptin to MET monotherapy, SU monotherapy, and MET + SU regimens in patients with inadequate glycaemic control is a good therapeutic option for achieving efficacy in patients with T2DM. However, adding sitagliptin to an ongoing MET + SU regimen appears

to be less efficacious among patients whose HbA1c is above 9% at the time of administration.

Chapter 6: Durability of glycaemic response among 2nd line treatments

6.1 Summary

Importance

There is limited information about the durability of glycaemic control when different oral GLTs are used as add-on treatments to MET in patients with T2DM.

Aim

To compare time to treatment failure between different classes of oral GLT when used as second line (add-on) treatments to MET monotherapy at HbA1c ≥ 7.5%.

Methods

A retrospective cohort study on 20,070 patients who were newly treated with a SU, DPP-4 inhibitor or TZD following MET therapy failure (2007-2014). Patients' data was sourced from UK General Practices via THIN database. The risk of dual therapy failure was compared between 3 treatment groups: MET+SU (reference group, n=15,508), MET+DPP-4 inhibitor (n=3,080) and MET+TZD (n=1,482). Follow-up was until treatment substitution or

intensification with a 3rd GLT, or for up to 5 years (totalling 46,430 person-years). Propensity score weighting and Cox proportional hazard regression analyses were employed.

Results

Unadjusted survival analysis showed the incidence of dual therapy failure at 1 year was 15% with SU, 23% with DPP-4 inhibitor and 8% with TZD. Corresponding failure rates at 2 years were 26%, 38% and 12% respectively. Adjusted multivariate models showed that, compared to the SU group, adding a DPP-4 inhibitor was associated with an increased risk of treatment failure (adjusted hazard ratio, aHR, 1.58; 95% CI: 1.48-1.68), while adding a TZD was associated with a reduced hazard (aHR, 0.45; 95% CI: 0.41-0.50). Baseline parameters associated with an increased hazard of intensification included HbA1c, diabetes duration, gender, smoking status and the use of statins.

Conclusions

In routine clinical practice, adding a DPP-4 inhibitor to MET is associated with an increased, earlier requirement for treatment intensification compared to adding a SU or TZD. Adding a TZD to MET resulted in the most durable glycaemic response.

6.2 Introduction

There is evidence that tight glucose control, especially in the early years after diagnosis, reduces the risk of long-term cardiovascular (CV) complications in

patients with T2DM).[55,180] International guidelines therefore recommend an individualized treatment strategy to achieve and maintain target levels of glycaemic control.[5] MET is the usual first-line therapy when diet and exercise are insufficient, but due to the progressive decline in beta cell function many patients fail to maintain adequate levels of HbA1c with monotherapy and require treatment intensification by adding a second oral agent.[4,5]

For most patients in whom MET alone is no longer sufficient, the options include adding a SU, DPP-4 inhibitor or a TZD. While these drugs have shown broadly similar reductions in HbA1c in randomized trials, the durability of glycaemic responses when added as dual therapy with MET in everyday practice is unknown. Recent observational studies,[181,182] and RCTs,[50] have mainly reported on the durability of GLTs when used as initial monotherapy rather than as add-on treatments in patients with longer duration T2DM.

Thus, the aim of this study was to compare the time to treatment failure among patients who added a DPP-4 inhibitor, SU or TZD to MET monotherapy in routine clinical practice, and to assess the glycaemic and body weight responses over time.

6.3 Methods

6.3.1 Study design and data source

Retrospective cohort analysis was conducted using patient records from THIN database. The study population comprised a cohort of patients identified to have T2DM and registered to a practice for more than 12 months before the index date (January 1st 2007 - May 30th 2014). The cohort included patients

who were newly treated with an SU, DPP-4 inhibitor or TZD following MET therapy failure. Patients who were administered other GLTs such as GLP-1 receptor agonists, SGLT2 inhibitors, glinides and acarbose were excluded from the study due to the small numbers of user cases. Also excluded were patients who added insulin treatment to MET monotherapy in order to enable us to compare different oral GLTs.

6.3.2 Treatment exposure

The exposures were incident intensification prescription of any SU (gliclazide, glimepiride, glipizide, glibenclamide or tolbutamide), a DPP-4 inhibitor (sitagliptin, vildagliptin, saxagliptin or linagliptin), or TZD (pioglitazone) as 2nd line GLT following MET monotherapy failure. Pioglitazone constitute 100% of TZDs in the data. The follow-up period commenced from the index date (the date of incident intensification prescription) through to the date of a censoring addition or substitution of another GLT at HbA1c $\geq$ 7.5%, up to 5 years after the index date. The study end date was May 30th 2014. Patients were segregated into three treatment groups based on the GLTs they received at baseline: MET + SU (reference/control group) vs. MET + DPP-4 inhibitor or MET + TZD.

6.3.3 Outcome

The primary composite outcome was time to dual therapy failure. This was defined as time to substitution or intensification of treatment with a 3rd agent at HbA1c $\geq$ 7.5%. Secondary outcomes included the glycaemic effectiveness

and body weight responses. The risks of treatment failure in the study population were compared across the three treatment groups.

6.3.4 Statistical analyses

Primary analyses include descriptive statistical analysis using Chi squared tests and logistic regression to assess all variables. A multinomial propensity score was assigned based on all the baseline covariates in our study.[131] This was designed to estimate the probability that a patient's initial 2nd line therapy was an SU (MET+SU was the treatment group with the largest number of patients).[183] Propensity score (PS) was estimated via inverse-probability-weighted regression adjustments (IPWRA)[184] using a logistic regression model in which the treatment status (indicator variable) was regressed on the baseline covariates.[185]

Balance in baseline covariates was assessed between the treatment groups by estimating the absolute standardized differences before and after propensity score weighting. A standardized effect size ≥ 10% indicated serious imbalance.[120] The variations in mean and frequency distribution of measured baseline covariates between treatment groups with the same estimated propensity score was summarized.

Crude and adjusted Kaplan–Meier estimates of survival functions were calculated to evaluate the association of the treatment groups, and differences in survival were assessed via the log rank tests. From the survival curves, the probability of dual therapy failure occurring within a 5-year follow-up was computed. Cox proportional hazards model was constructed, adjusting for all

covariates while including propensity score as a prognostic covariate. The marginal hazard ratios were estimated to allow for the quantification of the adjusted risk of requiring intensification with a 3rd line glucose-lowering agent in DPP-4 inhibitor or TZD treated groups compared to the SU group.

6.3.5 Sensitivity analyses

Additional sensitivity analyses were carried out to evaluate the robustness of the results by examining the assumption of no unmeasured confounding variable.[114,186] Assumption was made for an unmeasured covariate that would influence the measure of effect.[185] In addition, sensitivity analysis was carried out to compare results of covariates with missing data with those having multiple-imputed data to assess the reliability of the outcomes and the impact of missing data.

6.3.6 Bias

Selection and prevalent user biases were minimised by ensuring all participants in the exposed treatment groups examined initiated DPP-4 inhibitor as new users. A thoughtful and thorough specification of the selection model was employed to allow the successful application of the propensity score weighting technique, as shown in Figure 6.1. In an approach similar to that used in as-treated analyses, the intensification regimen, DPP-4 inhibitor was used to define drug exposure. To reduce confounding by indication, post index date exposure to any GLT other than MET, SU, DPP-4 inhibitor or TZD was not permitted. In addition, bias that may occur from Kaplan–Meier estimates of survival functions due to an unbalanced distribution of covariates was

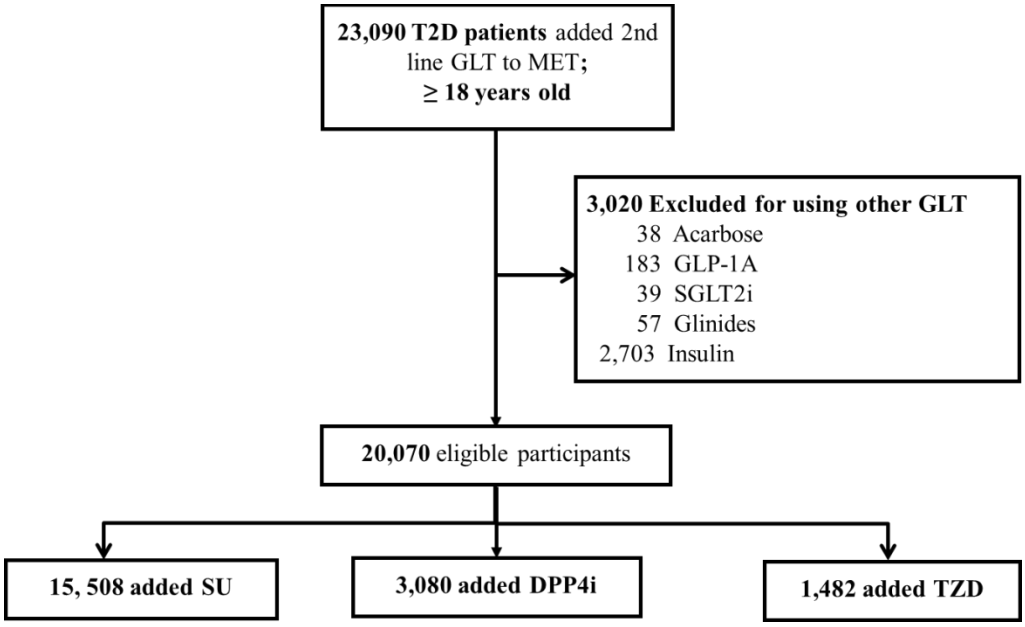
addressed. Adjusted log-rank test was used to compare the equality of the survival curves in the propensity score weighted cohort.[135]

6.4 Results

6.4.1 General patient characteristics

After screening 23,090 patients who intensified MET treatment with a 2nd line therapy, 20,070 patients made the criteria for cohort entry and were assigned to one of three treatment groups as outlined in Figure 6.1.

Figure 6.1 Study population screening and selection process



Abbreviations: DPP-4i (dipeptidyl-dipeptidase-4 inhibitors); MET (metformin); SU (sulphonylurea); TZD (thiazolidinedione); GLP1-RA (glucagon-like peptide 1 receptor agonist; SGLT-2i (sodium-glucose cotransporter-2 inhibitor)

Table 6.1 Baseline characteristics of patients

Baseline variable	Total (N = 20,070)	MET + SU (n = 15,508)	MET + DPP-4i (n = 3,080)	MET + TZD (n = 1,482)	Effect Size [a]	Effect Size [b]
Demographics						
Age (yrs), Mean (SD)	59.2 (12.9)	59.8 (13.1)	57.2 (12.0)	56.8 (11.5)	0.23	0.07
Gender, No. (%)						
Male	11741 (59)	9097 (59)	1767 (57)	877 (59)	0.04	0.02
Female	8329 (41)	6411 (41)	1313 (43)	605 (41)	0.04	0.02
Townsend deprivation, No. (%)						
Least deprived	4210 (21)	3140 (20)	751 (24)	319 (22)	0.10	0.03
Less	3950 (20)	3036 (20)	625 (20)	289 (20)	0.02	0.01
Average	4328 (22)	3388 (22)	644 (21)	296 (20)	0.05	0.04
More	4297 (21)	3381 (22)	588 (19)	328 (22)	0.07	0.02
Most deprived	3285 (16)	2563 (17)	472 (15)	250 (17)	0.04	0.02
Clinical Parameters, Mean (SD)						
HbA1c (%)	9.0 (2.2)	9.1 (2.3)	8.6 (1.6)	8.6 (1.7)	0.24	0.07
BMI (kg/m^2)	31.7 (6.6)	31.1 (6.5)	33.8 (6.8)	33.3 (6.6)	0.41	0.05
Weight (Kg)	90.9 (21.4)	89.0 (20.9)	97.8 (22.0)	96.0 (21.6)	0.41	0.06
SBP (mmHg)	134.8 (15.6)	135.0 (15.9)	133.7 (14.6)	135.3 (14.4)	0.10	0.04
DBP (mmHg)	79.4 (9.7)	79.4 (9.8)	79.6 (9.3)	79.8 (9.0)	0.05	0.01
TC (mmol/l)	4.7 (1.3)	4.8 (1.4)	4.6 (1.2)	4.5 (1.2)	0.19	0.06
HDL-C (mmol/l)	1.2 (0.3)	1.2 (0.3)	1.2 (0.3)	1.2 (0.3)	0.05	0.02
LDL-C(mmol/l)	2.6 (1.1)	2.7 (1.1)	2.5 (1.0)	2.5 (1.0)	0.20	0.05
Triglyceride (mmol/L)	2.4 (2.5)	2.5 (2.7)	2.3 (1.9)	2.4 (2.5)	0.05	0.03
GFR (mls/min/1.73m^2)	73.8 (17.2)	73.4 (17.3)	75.7 (16.4)	74.7 (16.8)	0.13	0.02
ACR (mg/mol)	4.1 (9.3)	4.3 (9.4)	3.9 (9.3)	3.2 (8.5)	0.12	0.05

Baseline variable	Total (N = 20,070)	MET + SU (n = 15,508)	MET + DPP-4i (n = 3,080)	MET + TZD (n = 1,482)	Effect Size [a]	Effect Size [b]
Diabetes duration (yrs) [c]	2.9 (3.6)	2.8 (3.6)	3.4 (3.3)	3.0 (3.2)	0.17	0.02
Smoking status, No. (%)						
Non-smoker	8706 (43)	6787 (44)	1315 (43)	604 (41)	0.06	0.04
Current smoker	3750 (19)	2935 (19)	534 (17)	281 (19)	0.04	0.01
Ex-smoker	7614 (38)	5786 (37)	1231 (40)	597 (40)	0.06	0.03
Use of Medications, No. (%)						
Aspirin	3803 (19)	2861 (18)	544 (18)	398 (27)	0.23	0.06
Antihypertensive	10592 (53)	7985 (51)	1803 (59)	804 (54)	0.14	0.02
LLT	11588 (58)	8506 (55)	2138 (69)	944 (64)	0.29	0.06
Comorbidities, No. (%)						
Hypoglycaemia	509 (3)	463 (3)	32 (1)	14 (1)	0.13	0.10
CHD	270 (1)	230 (1)	21 (1)	19 (1)	0.07	0.06
PAD	210 (1)	178 (1)	14 (0)	18 (1)	0.07	0.04
Heart Failure	458 (2)	406 (3)	41 (1)	11 (1)	0.13	0.11
Stroke	297 (1)	261 (2)	19 (1)	17 (1)	0.09	0.08

Effect size (ES) is the absolute standardised mean difference of means or percentages divided by the standard deviation.

[a] Unweighted ES

[b] ES after propensity score weighted cohort based on average treatment effect in the population (ATE). Differences between treatment groups have been reduced by weighting using the propensity score

[c] Diabetes duration is time from first diagnosis of diabetes to date of intensification with 2nd line drug (index date)

The number (proportion) of patients assigned to each treatment group included 15,508 (77%) who added a SU 3,080 (15%) added a DPP-4 inhibitor and 1,482 (7%) added a TZD.

Overall, patients had a mean age of 59 yrs and were predominantly male (59%). Compared with patients who added other GLTs, those who added an SU appeared to be older (60 vs 57 yrs, respectively), had higher mean HbA1c levels (9.1% vs 8.6%, respectively), lower BMI and lower diabetes duration (Table 6.1). The patients' socioeconomic status was similar across the treatment groups. Before PS weighting, many of the measured covariates had a standardized difference above the 0.10 level (Table 6.1). However, the application of PS weighting brought into balance the distributions of the measured covariates. Apart from previous hypoglycaemia and a diagnosis of other coronary heart diseases (CHD), the baseline characteristics of the weighted sample were not statistically different; as a result, the systematic differences between subjects in the treatment groups in the original cohort have been substantially reduced in the weighted sample (Table 6.1). This shows that the differences between the treatment groups have been reduced by PS weighting and adequate balance on baseline covariates has been induced by the specification of the PS model used.

6.4.2 Estimating survival curves for treatment failure

Figure 6.2 Full cohort Kaplan–Meier survival curves

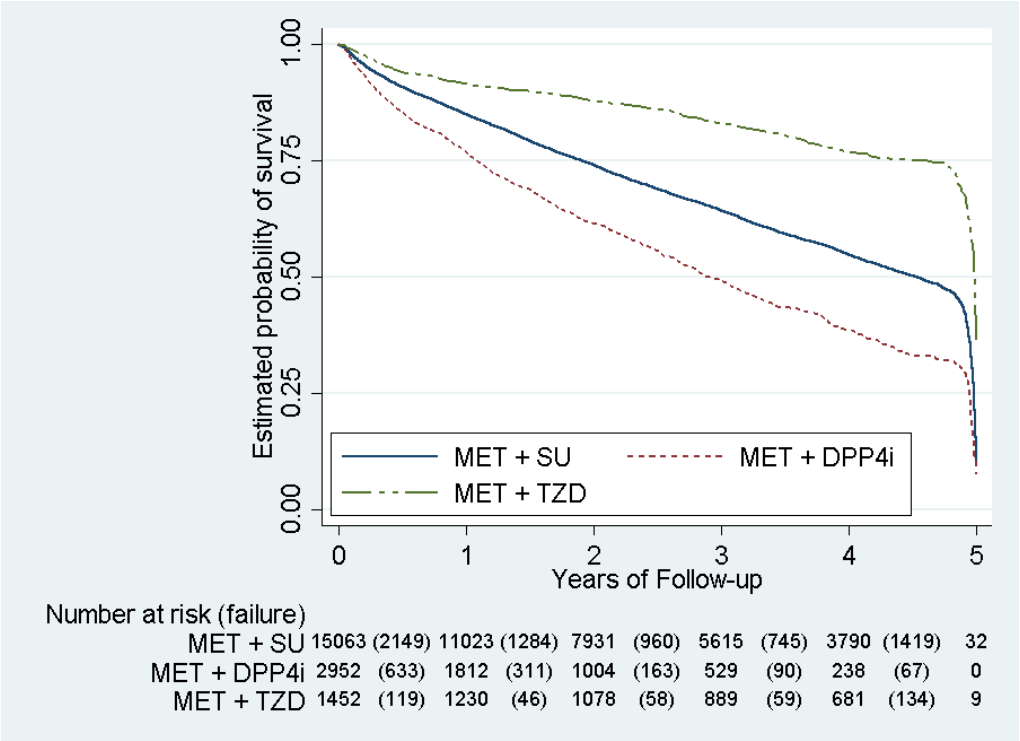
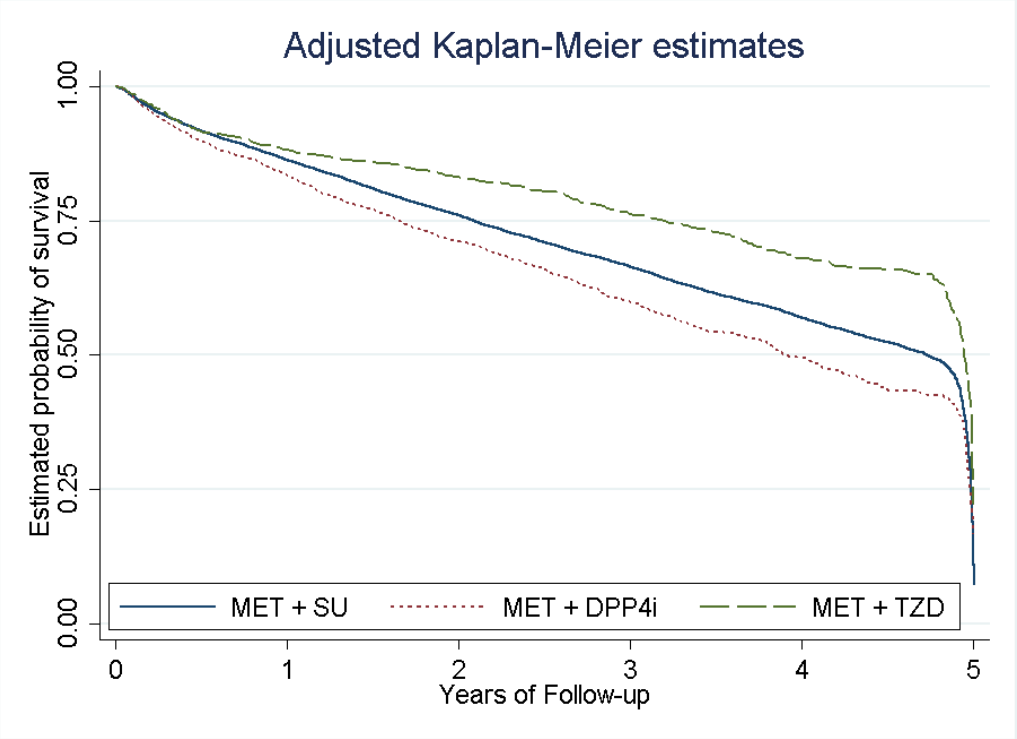


Figure 6.2 depicts the Kaplan-Meier survival curves for the treatment group (SU, DPP-4 inhibitor and TZD) participants in the original sample. The three survival curves are significantly different from one another (log-rank test P value < 0.001)

Figure 6.3 Propensity weighted corhot KM survival curves



6.4.3 Time to dual therapy failure

The average time to treatment failure in the cohort of patients who added a 2nd line oral glucose-lowering agent after MET is summarised in Table 2. Overall, 6,891 (44%) of patients who received a SU, 1,360 (44%) who received a DPP-4 inhibitor and 438 (30%) of patients who received TZD had to add or switch to another glucose-lowering therapy during the study period. The mean time to treatment failure among TZD users was the longest at 3.3 years, followed by SU users (2.4 years) and then DPP-4 inhibitor (1.6 years). The unadjusted survival analysis showed the incidence of dual therapy failure at 1 year was 15% with SU, 23% with DPP-4 inhibitor and 8% with TZD. Corresponding failure rates at 2 years were 26%, 38% and 12%, respectively (Figure 6.2).

Crude Kaplan–Meier (KM) survival curves for subjects who added SU, DPP-4 inhibitor or TZD in the unadjusted cohort are shown in Figure 6.2. The result shows there was a significant difference between the three curves; log-rank test $p<0.001$. The adjusted KM survival curves obtained from the PS weighted cohort were also significantly different (adjusted log-rank test $p< 0.001$) (Figure 6.3). From the estimated survival curves, this data showed that the rates of dual therapy failure were significantly different. Second line use of a TZD was associated with the most durable glycaemic response, followed by the SU and then a DPP-4 inhibitor. Thus, patients who added a DPP-4 inhibitor or SU to MET monotherapy were more likely to require a 3rd line glucose-lowering agent than those who added a TZD.

These results were consistent and remained significant in the adjusted multivariable Cox proportional hazards models, with DPP-4 inhibitor use (aHR, 1.58; 95%CI, 1.48-1.68) being associated with an increased hazard of dual therapy failure and TZD use (aHR, 0.45; 95%CI, 0.41-0.50) associated with a decreased hazard of treatment failure compared with SU, respectively (Table 6.2).

In addition, factors predicting earlier failure of dual therapy on any of the glucose-lowering agents were led by use of a lipid lowering drug, mainly statins (aHR = 1.57). Other significant risk factors included being female (aHR = 1.38), current smoking status (aHR = 1.07), T2DM duration (aHR = 1.07), body weight (aHR = 1.02) and HbA1c (aHR = 1.02) (Appendix C-2).

Table 6.2 Rates, hazard ratios and glycaemic and body weight responses

	MET + SU	MET + DPP-4i	MET + TZD
Person-years (n)	36,643	4,964	4,823
Av time to treatment failure (yrs)	2.4 (1.7)	1.6 (1.3)	3.3 (1.7)
Unadjusted failure rate (95% CI)			
Year 1	0.15 (0.14-0.16)	0.23 (0.22-0.25)	0.08 (0.07-0.10)
Year 2	0.26 (0.25-0.27)	0.38 (0.37-0.41)	0.12 (0.10-0.14)
Year 3	0.36 (0.35-0.37)	0.51 (0.48-0.53)	0.17 (0.15-0.19)
Year 4	0.45 (0.44-0.46)	0.61 (0.59-0.64)	0.23 (0.21-0.26)
Year 5	0.90 (0.88-0.91)	-	0.64 (0.56-0.72)
Adjusted hazard ratio (95% CI) [a]	Reference	1.58 (1.48-1.68)	0.45 (0.41-0.50)
Mean (SD) HbA1c change, % [b]	-1.3 (2.4)	-0.9 (1.6)	-1.2 (1.9)
Year 1	-1.0 (2.0)	-0.6 (1.4)	-0.6 (1.7)
Year 2	-1.2 (2.3)	-0.8 (1.6)	-0.9 (1.9)
Mean (SD) Weight change, Kg [b]	-0.2 (6.7)	-1.8 (6.3)	+1.8 (8.8)
Year 1	-0.4 (5.1)	-0.7 (3.9)	+0.2 (4.7)
Year 2	-0.1 (5.8)	-1.1 (4.8)	+0.5 (6.1)

[a] Adjusted for all baseline covariates and propensity score
[b] Overall change in absolute value **(values are running average)**

6.4.4 Glycaemic and body weight responses

Results of the descriptive analysis showed that, overall, the co-administration of SU, DPP-4 inhibitor and TZD to patients who had inadequate glycaemic control with MET were associated with significant HbA1c reductions of -1.3%, -0.9% and -1.2%, respectively ($p<0.001$). Over the course of therapy, the addition of a SU produced between 0.3 and 0.5% greater reduction in HbA1c compared to the addition of a DPP-4 inhibitor whereas the addition of

a TDZ appears to show a fluctuating pattern of reduction that was not significantly different from the SU (Figure 6.4).

Figure 6.4 Glycaemic responses – HbA1c over time

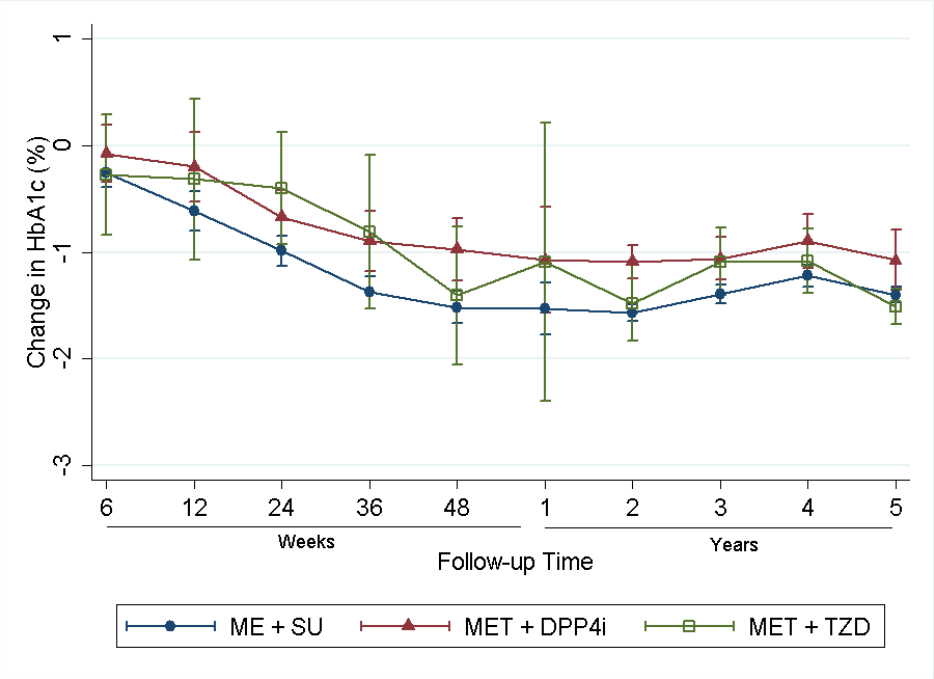


Figure 6.4 shows SU added to MET maintained a 0.3 to 0.5% greater reduction in HbA1c compared to MET plus DPP-4 inhibitor. There was no clear or consistent difference in HbA1c changes between SU and TZD when added to MET over time. (Note: Mean HbA1c changes are not running averages)

Figure 6.5 Patients achieving HbA1c targets at 1 year

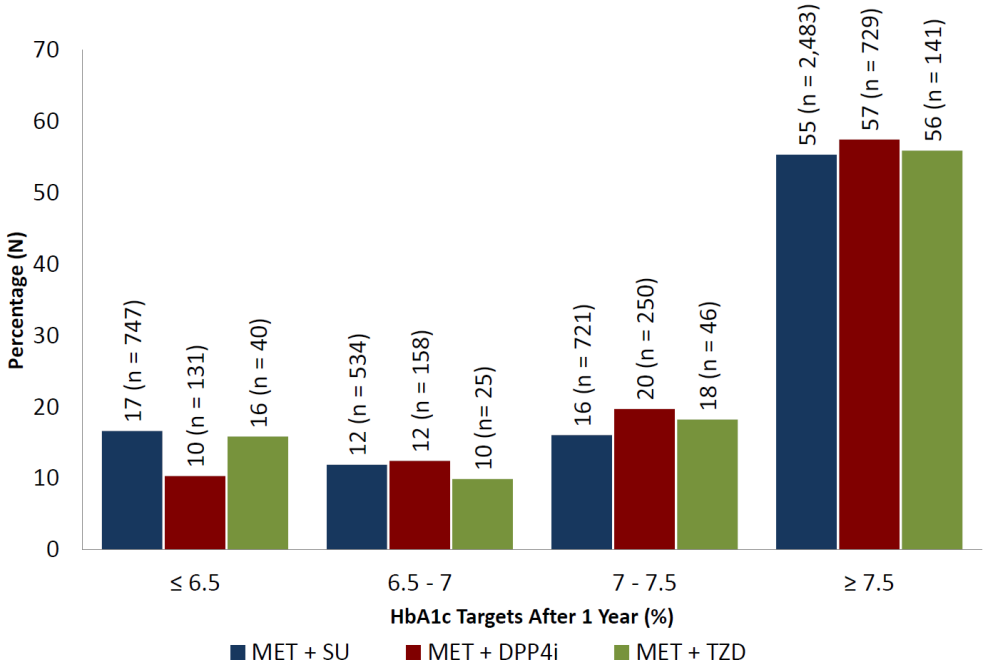


Figure 6.5 shows HbA1c goal achievement was quite low for all dual therapies studied. The addition of SUs second line to MET had the best rate of HbA1c goal attainment, followed by DPP-4 inhibitors, and then TZDs

In addition, the data show that addition of a 2nd line oral agent to MET was associated with an overall 15% of patients meeting a HbA1c target ≤ 6.5% and about 27% meeting a HbA1c target < 7% after 1 year of dual therapy. In terms of comparative responses at 1 year, the proportion of patients attaining HbA1c targets below 7% after using SU, DPP-4 inhibitor and TZD include 29%, 22% and 26%, respectively (Figure 6.5).

Figure 6.6 Body weight changes over time

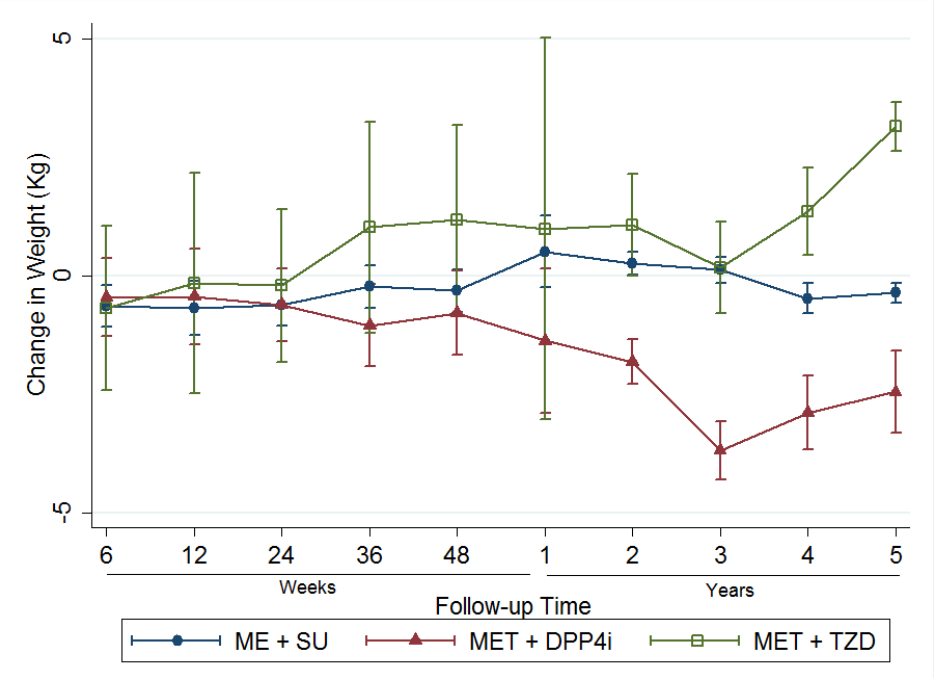


Figure 6.6 shows intensifying MET therapy with TZD was associated with weight gain over time, while DPP-4 inhibitor was associated with weight loss over time. SU was associated with neutral weight changes over time

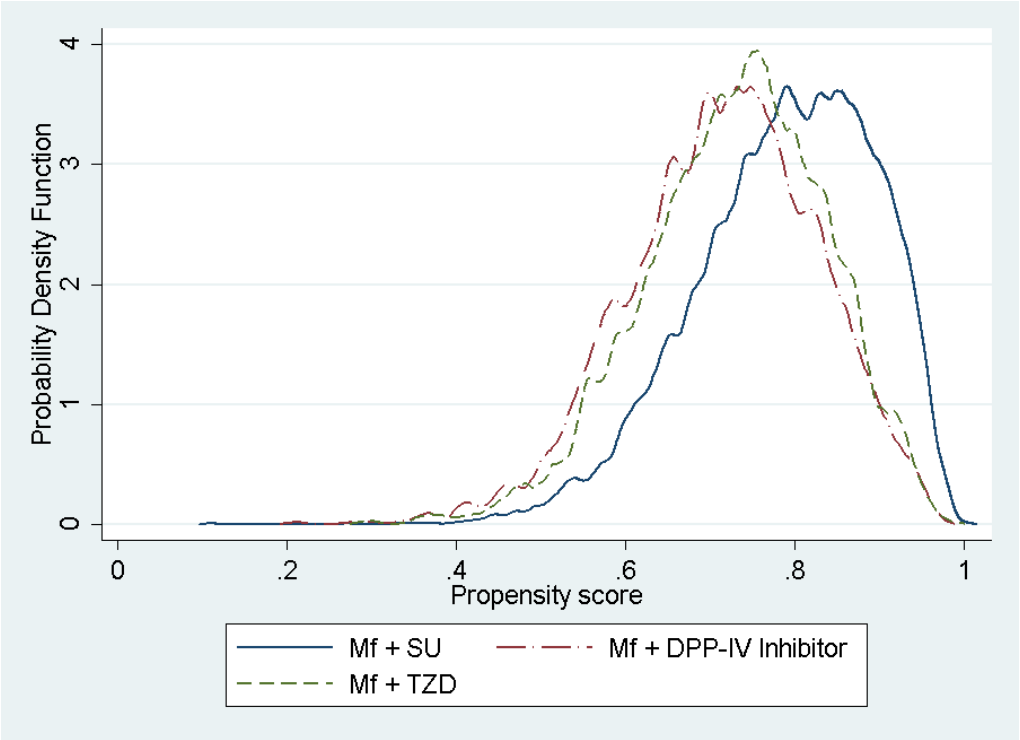
In terms of overall changes in body weight, the addition of a TZD was associated with significant weight gain (1.8kg, p<0.001), while the addition of DPP-4 inhibitor produced a significant weight reduction (-1.8kg, p<0.001). A very small reduction in body weight was observed with the addition of SU (-0.2kg, p<0.001) (Figure 6.6).

6.4.5 Sensitivity analyses

Violations in the proportional hazards assumptions were assessed using Schoenfeld residuals test, which tests the null hypothesis that the hazard ratio is constant over time. There is no evidence (P=0.5) to reject the assumption of

proportional hazards for the treatment groups. The sensitivity analysis on missing data yielded comparable results to complete case models; the estimated aHR for DPP-4 inhibitor was 1.47 (95% CI: 1.34-1.60) and 0.50 for TZD (95% CI: 0.43-0.58), which reflects results that are unlikely to be attributable to bias from missing information. The probability density functions of the PS weighting of the treatment groups show there was no violation of the overlap assumption either.[173] (Figure 6.7)

Figure 6.7 Distribution of propensity score by treatment group



The probability density functions of the propensity score weighting for treatment groups MET + SU (n=15,508) vs. MET + DPP-4 inhibitor (n=3,080) and MET + TZD (n=1,482). The distribution of the propensity score show no indication of too much probability mass near 0 or 1 and the respective estimated densities of propensity score in the covariate distribution between the treatment groups overlap each other.

6.5 Discussion

This study has shown that in routine clinical practice, among patients with T2DM receiving a 2nd line GLT as add-on to MET, the addition of a TZD is associated with the most durable glycaemic response, followed by a SU and then a DPP-4 inhibitor. Compared with an SU, adding a DPP-4 inhibitor to MET was associated with an increased risk of dual therapy failure and/or an earlier requirement for treatment intensification with a 3rd agent. Factors associated with earlier dual therapy failure included concomitant use of statin therapy, being female, a smoker, having longer duration of diabetes and higher baseline HbA1c. Adding an SU to MET as the 2nd line treatment gave the best chance of attaining an HbA1c goal of 7.5.

The Agency for Healthcare Research and Quality has suggested that the durability of glycaemic response after treatment intensification is best investigated using well-designed long-term observational studies.[6] Previous studies, however, have mainly focused on the durability of initial monotherapies, often in drug-naïve patients.[50,181,182] The present study has focused on the most commonly prescribed add-on therapies to MET. The results are similar to those in the ADOPT (A Diabetes Outcomes Progression) trial,[50] which showed that 'time to monotherapy failure' was longer with rosiglitazone, a TZD no longer widely used compared with MET and a SU, glyburide.[50]

The ADOPT study which involved 4,360 patients followed for a median of 4 years also reported rosiglitazone showed a significant greater reduction in

HbA1c compared to SU (between-group absolute difference of -0.4%), which contrasts the HbA1c changes observed in our study. Combination therapies have been shown to have additive effects and are better at reducing HbA1c compared with monotherapy regimens. A recent review of 140 clinical trials and 26 observational studies on head-to-head comparisons of GLTs (MET, second-generation SU, TZD, meglitinides, DPP-4 inhibitors, and glucagon-like peptide-1 (GLP-1) receptor agonists) as monotherapy and combination therapy[51] reported most medications decreased the HbA1c level by about 1% (absolute reduction). A study conducted by the Quartet study group examined the long-term sustained effects of adding pioglitazone or gliclazide to failing MET monotherapy.[187] The 2-year, randomised, multicentre trial were performed in patients with inadequately controlled HbA1c (7.5-11% inclusive) and the mean reduction in HbA1c from baseline was 0.9% for TZD vs 0.8% for SU ($p = 0.2$). The SU group showed an initial better efficacy compared to TZD users, however, a progressive rise in HbA1c was observed by week 24 in both groups, with the SU group showing a more prominent increase. A similar initial pattern of reduction was observed in our study. However, SU and TZD groups showed progressive increase after 2 years of treatment.

Observational studies have also shown that patients initially prescribed MET are significantly less likely to require treatment intensification than those who initiated treatment with other GLTs.[51,181,182] However, therapeutic responses to GLTs may be different when used as add-on to MET compared with monotherapy.[188] The results of this real-world observational study are different to those of a previous randomized, controlled trial in which better durability of

glycaemic response was observed over 2 years with a DPP-4 inhibitor (sitagliptin) added to MET compared with a SU (glipizide).[178] In addition, a recent real-world study by Inzucchi et al[189] showed the opposite result to our study; patients treated with MET + sitagliptin showed a 24% lower risk of insulin initiation over a 6-year period compared with MET + SU users (HR 0.76; 95% CI 0.65–0.90). We speculate that one of the reasons for this disparity could be as a result of the different A1c levels characterised by the different cohorts. The study population examined by Inzuuchi et al had a mean A1c of approximately 8% compared to our study population with approximately 9% at baseline, moreover, the outcome of their stratification analysis which examined patients with A1c ≥9% was not statistically different between sitagliptin vs SU users, and results from their Cox model also showed that a 1% increase in A1c level was associated with a 20% increase in the risk of insulin initiation. The pattern of HbA1c reduction observed across the treatment groups over time was similar for SU and DPP-4 users, even though SU users maintained between -0.3% and -0.5% more reduction in HbA1c compared to DPP-4 inhibitor users. However, there was no consistent pattern seen with TZD users, which showed fluctuations in glycaemic response after 48 weeks. It appears the initial decline in HbA1c among TZD users was not as prominent as that seen with SU or DPP-4 inhibitor within the first 24 weeks of treatment. We assume the reason for this initial slow response could be explained by certain uncontrollable factors. For example, SU is an insulin secretagogue that acts to increase insulin secretion, whereas TZD acts by activating a nuclear receptor thereby altering genetic transcription, however,

TZDs also act without increasing insulin secretion. Therefore, the slow initial glycaemic response observed with TZD is consistent with the response that should occur within the first 3 months of TZD administration.

All assessed classes of glucose-lowering agents were associated with statistically significant reductions in HbA1c. The overall absolute mean change in HbA1c between SU, DPP-4 inhibitor and TZD was -1.3%, -0.9% and -1.2%, respectively. The respective 2-year mean change obtained in our study (-1.2%, -0.8% and –0.9%) are consistent with previous evidence that an SU is superior to a TZD when added to MET[188] – data on DPP-4 inhibitor was not available in this study. Conversely, other reviews have concluded the glycaemic efficacy of SU was not superior to DPP-4 inhibitor[190] or TZD[51,190] when added to MET, which corresponds with this 5-year data for TZD users. In terms of goal attainment , most prospective randomized head-to-head trials of SU vs DPP-4 inhibitor co-administered with MET have 1 to 2 years study duration and have similar proportions of patients who are lost to follow-up due to lack of efficacy. Similar proportions of patients attained glycaemic efficacy (HbA1c < 7%) in our data; SU (29%) compared to DPP-4 inhibitor (22%) and TZD (26%). A similar result was obtained from a randomised active-comparator study which showed the proportion of patients who achieved this target at 1 year was similar between SU and DPP-4 inhibitors as add-on to MET (22.7% vs 23.1%, respectively).[191] It is well known that the higher the HbA1c, the greater the reduction in HbA1c with all agents. Therefore, the higher HbA1c recorded for SU users at baseline (9.1%) compared to DPP-4 inhibitor vs TZD (8.6% vs 8.6%) may add more weight to the absolute mean change observed

among SU users. Results obtained from RCTs and real-world data in the area of superiority between SU and DPP-4inhibitor as the add-on therapies to MET vary, meaning that a definitive conclusion cannot be made on the superiority of either SU over DPP-4 inhibitor in controlling HbA1c in patients with T2DM. The results of this real-world data imply that future robust research should examine efficacy in subgroup of patients over time and what influences a clinician's choice of treatment option.

When used in combination with MET, weight loss was observed with DPP-4 inhibitor (-1.8kg) vs. weight gain with TZD (1.8kg). TZD was associated with weight increase after 24 weeks of treatment, and excessive weight gain was observed after 3 years. Placebo-controlled trials have shown TZD and SU increased body weight by 1 to 5kg.[52] In a study of direct comparisons of monotherapies with TZD and SU, increased body weight was recorded with SU, even though this was lesser than that observed with TZD. Combinations of MET plus a TZD or MET plus SU increased weight more than MET monotherapy.[51] In contrast, SU was associated with borderline weight loss in this study. The reason for this disparity cannot be explained from the data. However, underlying factors such as education, lifestyle changes and combination therapy with MET are assumed to have contributed to the weight loss observed. Moreover, the greater amount of weight gain induced by TZD compared with SU in the study is consistent with data from another RCT – the ADOPT study, where patients administered SU monotherapy gained weight during the first year of treatment and thereafter experienced a gradual decline

in body weight during the subsequent years.[50] On the other hand, the weight loss that appeared consistent 3 years from commencement of therapy the weight loss accompanying use of a DPP-4 inhibitor appeared to be consistent 3 years from commencement of therapy. A previous study showed the co-administration of a DPP-4 inhibitor with MET is associated with similar weigh loss effect when compared with MET monotherapy, although the strength of evidence was low due to fewer studies on DPP-4 inhibitors.[51] In contrast, previous overviews have concluded there is no significant weight change with a DPP-4 inhibitor (-0.14kg, 95%CI: -0.94 to 0.63kg), while SU and TZD were associated with 2.06kg vs 2.08kg weight gain, when used in combination with MET, respectively.[190] Additional analysis conducted to investigate the correlation between change in body weight and change in HbA1c showed a significantly negative but weak association between change in HbA1c and change in weight in the cohort. (Pearson's correlation coefficient, $r = -0.03$; $p<0.001$). Change in weight accounted for approximately 0.1% of the total variation in HbA1c change; for every 1kg increase in weight, HbA1c increased by an estimated 0.01%, which is clinically irrelevant.

The risk for hypoglycaemia with an SU has been reported to increase by 6-fold, compared with other GLT.[51] Newer agents such as DPP-4 inhibitor when added to MET was also reported to reduce HbA1c levels, but without additional risk for hypoglycaemia.[51] Unfortunately, the risk of hypoglycaemia in this data could not be assessed due to inadequate reporting of hypoglycaemic events by general practices. The assessment of this study data resulted in the identification of factors that may independently predict earlier

need for treatment intensification other than HbA1c. These include diabetes duration, gender, smoking status, body weight and the use of statins. These findings may be particularly relevant for evaluating whether the adherence to GLTs could be influenced by individual patient characteristics and foster research on the characteristics of patients that benefit most from SU, DPP-4 inhibitors and other newer second line agents. The observation that the use of statins is an independent predictor of the need for treatment intensification is interesting. However, there is the possibility that this effect may be behavioural, for example, pro-active general practices may prescribe more statins for early intervention. Adverse effects of statins in increasing the risk of developing new onset diabetes and/or reducing insulin secretion and insulin sensitivity is being recognised.[192] Further studies are required to clarify the hyperglycaemic effects of statins and the need for intensification of GLT.

Limitations of the study

The analyses were subject to a number of limitations that are inherent to observational studies. Firstly, there is no certainty that the patients were fully compliant with their medication. Other factors apart from HbA1c may also influence the decision to intensify treatment in everyday practice. These may include unknown compliance, tolerability, safety, cost, physician's reason for adding/substituting with a third oral agent and a patient's preference. Unfortunately, the study was unable to evaluate how these factors might have influenced the findings. These information are useful in the management of type 2 diabetes in routine clinical practice and can be best obtained through qualitative research studies. Other newer agents being added to MET

monotherapy such as GLP-1receptor agonists and SGLT-2 inhibitors were not assessed due to limited number of patients using them in our data. The failure to obtain similar data on SGLT-2 inhibitors and GLP-1receptor agonists is a major limitation of the present study. Although we could not account for potential residual confounders such as compliance, indications for intensification treatments, markers of β-cell deterioration and differences in dosages, we were able to account for differences in the observed covariates and used robust analytical techniques to control confounding that may bias the results of the estimated treatment effects. The use of propensity score analysis to estimate average treatment effect in a large dataset contributed to the balancing of treatment and comparison groups on the available covariates without the loss of observations. However, the limitation to this technique is that it only accounts for observed covariates. Hence, other factors that may influence a physician's choice of therapy that cannot be accounted for in the study or any other hidden biases that may remain after PS weighting l overlap between treatment and control groups.

Conclusion

Comparative effectiveness studies and RCTs which examine the durability of 2nd line glucose-lowering agents when added to MET are limited. Conducting RCTs at this level of treatment is not without its numerous challenges. It was observed that in routine clinical practice, among patients with T2DM, the rates of dual therapy failure was significantly different; with second line use of a TZD being associated with the most durable glycaemic response at the expense of greatest amount of weight gain, followed by SU, and then a DPP-4 inhibitor,

the latter associated with the greatest amount of weight loss. In the absence of real-world evidence of comparative effectiveness studies on the durability of 2nd line glucose-lowering agents following MET monotherapy failure to maintain adequate glycaemic control, this study highlights the differences in HbA1c end points between three oral treatment options that are frequently used in patients as dual therapy.

Chapter 7: Dual therapy intensification and the risk of CV events

7.1 Summary

Aim

To compare time to a composite endpoint of non-fatal acute myocardial infarction, non-fatal stroke or all-cause mortality in patients with T2DM who had their treatment intensified with a DPP-4 inhibitor or INS following dual therapy (MET plus SU) failure.

Method

A retrospective cohort study was conducted on 5,238 patients who were newly treated with either a DPP-4 inhibitor or INS following dual therapy failure (2007-2014). Data was sourced from UK General Practices via THIN database. The risk of the composite outcome was compared between 2 treatment groups: MET+SU+INS (n=1,584) and MET+SU+DPP-4 inhibitor (reference group, n=3,654), while adjusting for baseline covariates. Follow-up was for 5 years (totalling 16,887 person-years). Propensity score matching analysis and Cox proportional hazard models were employed.

Results

Overall, there were 123 and 171 composite outcome events among patients who added INS and a DPP-4 inhibitor, respectively (44.5 vs 14.6 events per 1000 person-years). Addition of INS was associated with a significantly higher hazard ratio (HR) compared to the addition of a DPP-4 inhibitor (aHR 2.6 (95% CI: 1.9–3.4; p<0.01), an effect that was even more pronounced in obese (BMI 30-34.9kg/m^2) patients (corresponding aHR 3.6, 95%CI: 2.3-5.6, p< 0.01).

Conclusion

In routine clinical practice, intensification of dual oral therapy (MET+SU) by adding INS is associated with an increased risk of CV events and death compared with adding a DPP-4 inhibitor. These findings are in line with suggestion from previous studies regarding the cardiovascular safety of insulin in T2DM, but need to be interpreted with caution due to the observational nature of the study.

7.2 Introduction

There is evidence that tight glucose control, especially in the early years after diagnosis, reduces the risk of long-term CV complications in patients with T2DM.[55,180] International guidelines therefore recommend an individualized treatment strategy to achieve and maintain target levels of glycaemic control.[5] Metformin is the usual first-line drug therapy in T2DM.[4,5] The joint EASD/ADA position statements[5] as well as the NICE guidelines recommend the use of a range of GLTs such as SU, DPP-4 inhibitors, SGLT2 inhibitors, TZD, Insulin and GLP-1 RA as second line to MET therapy. Routine clinical data from UK primary care shows that MET + SU is the most widely used dual oral GLT combination. Because of the progressive decline in β-cell function, many patients fail to maintain adequate levels of HbA1c despite up-titration to maximum tolerated doses of dual therapy with MET and SU. Previous real-world data studies have examined the cardiovascular safety of MET therapy intensification with available second line treatment options.[193]

Several treatment options are also available and recommended when MET plus SU therapy is insufficient, but there is very limited data on CV and diabetes-related outcomes in this group of patients to inform decision-making about third-line treatments. For many patients the choice includes adding basal insulin (INS) or a DPP-4 inhibitor as a third oral agent, but the mortality risk and comparative outcomes when these treatments are used in dual therapy failure are unknown. There are concerns about the CV safety of INS in T2DM,[27,193-197] but these studies have mainly investigated the use of INS *per se*, as monotherapy or in combination with metformin.[27,193-197] On the other

hand, the UKPDS[198] and ORIGIN[199] trials have demonstrated the safety of INS, while recent prospective RCTs have shown the CV outcomes of DPP-4 inhibitors are non-inferior to placebo.[42,43] However, no RCT has compared INS with DPP-4 inhibitors either in terms of their CV safety or effectiveness as a third option after MET plus SU fails. Further work is needed to explore the CV safety of INS when used as a third line therapy, often in patients with longer duration disease and higher CV risk. Insulin is known to exert antiatherogenic effects[200] and many patients prefer to delay INS treatment because of fear of injections, weight gain and the risk of hypoglycaemia. Therefore, adding a DPP-4 inhibitor to MET+SU is an effective alternative to lower HbA1c. Prior to recent RCTs which have demonstrated the safety of DPP-4 inhibitor, there has been some uncertainty about CV outcomes with DPP-4 inhibitors[42,201] and till date, there are no comparative outcome studies available on DPP-4 inhibitor versus INS in patients with dual therapy failure.

Therefore, the aim of the present study is to compare CV outcomes and mortality among patients with T2DM who, in routine clinical practice, intensified their treatment with the addition of INS or a DPP-4 inhibitor following dual therapy (MET+SU) failure.

7.3 Methods

7.3.1 Study design and data source

A retrospective cohort analysis was conducted using data from THIN. The study population comprised a cohort of patients identified to have T2DM and registered to a practice for more than 12 months before the index date (January

1st 2007 - May 30th 2014). The cohort included patients who were newly treated with a DPP-4 inhibitor or INS following MET+SU therapy failure. Patients who were administered other GLTs were excluded from the study. Also excluded were patients with a baseline diagnosis of a CV condition.

7.3.2 Treatment exposure

The exposures were incident intensification prescription of INS (long-acting, short or fast-acting, or biphasic) or a DPP-4 inhibitor (sitagliptin, vildagliptin, saxagliptin, linagliptin) as 3rd line GLT following dual (MET+SU) treatment failure. The follow-up period commenced from the index date through to the date of a censoring outcome event until a switch to, or addition of, another GLT, up to 5 years after the index date. The study end date was May 30th 2014. Patients were segregated into two treatment groups based on the GLTs they received at baseline: MET + SU + INS vs. MET + SU + DPP-4 inhibitor (reference/control group).

7.3.3 Outcome

The primary composite outcome was time to diagnosis of predefined events. These included non-fatal acute myocardial infarction (AMI), non-fatal stroke and all-cause mortality. Secondary outcomes included CV events (non-fatal AMI, non-fatal stroke and CV-related deaths combined), all-cause deaths and CV-related deaths. Read codes used for identifying AMI and strokes are included in Appendix B-1. CV-related deaths were included where the cause of death was documented. Subjects whose cause of death could not be verified were ignored in the CV deaths analysis. The risks of events in the study

population were compared between the two treatment groups. In addition, descriptive analysis of the glycaemic and body weight responses of patients in each treatment group was conducted.

7.3.4 Statistical analyses

Primary analysis was time to the composite outcome of non-fatal AMI, non-fatal stroke or all-cause death in a propensity score-matched cohort. A propensity score (PS) model was estimated using a logistic regression model in which the treatment status was regressed on the baseline covariates.[185] The balance in baseline covariates between the treated (INS) and reference (DPP-4 inhibitor) subjects was assessed using standardized differences before and after matching.[202] An absolute standardized difference > 10% indicated serious imbalance. The mean and frequency distribution of measured baseline covariates between treatment groups with the same estimated PS was examined and summarized. Pairs of treated group and reference subjects were matched based on their estimated treatment probabilities using logistic regression. The average treatment effect on the treated (ATT) was estimated by finding at least 1 match for each of the treated subjects from the reference group. PS was considered as a prognostic covariate and included in a Cox proportional hazards regression model.

Crude and adjusted KM estimates of survival functions were obtained for the treatment groups in the full cohort and PS-matched cohort. From these survival functions, the absolute reduction in the probability of an event occurring within a 5-year follow-up was estimated.

7.3.5 Subgroup and sensitivity analyses

The analysis included an assessment of the hazard ratio of an event occurring in subgroups of obese patients with BMI between 30 to 34.9kg/m^2 and those considered as severely with BMI ≥ 35kg/m^2. Sensitivity analysis was aimed at examining the assumption of no unmeasured binary confounding variable.[114,186] Assumption was made for an unmeasured covariate that would increase the odds of assigned treatment.[185] Sensitivity analysis was applied to KM survival functions in the PS-matched cohort.

7.3.6 Biases

Data analysis employed the "new user" design to minimize biases associated with prevalent use of intensification regimens.[114] In an approach similar to that used in as-treated analyses, treatment exposure was defined using the intensification regimen; therefore, post index date exposure to any other GLT other than MET, SU, DPP-4 inhibitor or INS was not permitted in the study to reduce confounding by indication.

Immortal-time bias was addressed by ensuring subjects diagnosed with outcome events on or before the index date were excluded. In addition, to eliminate bias that may occur from KM estimates of survival functions due to an unbalanced distribution of covariates, a stratified log-rank test was used to compare the equality of the survival curves in the matched sets (KM survival curves were estimated separately for INS treated and compared with DPP-4 inhibitor treated participants in the PS-matched sample).[129]

7.4 Results

7.4.1 General patient characteristics

After screening 8,654 patients who intensified MET+SU treatment with a 3rd line drug, 5,238 patients made the criteria for cohort entry and were assigned to one of two treatment groups as outlined in Figure 7.1. The number (proportion) of patients assigned to each treatment group included: n=1,584 (30%) for MET + SU + INS and n=3,654 (70%) for MET + SU + DPP-4 inhibitor.

Figure 7.1 Study population screening and selection process

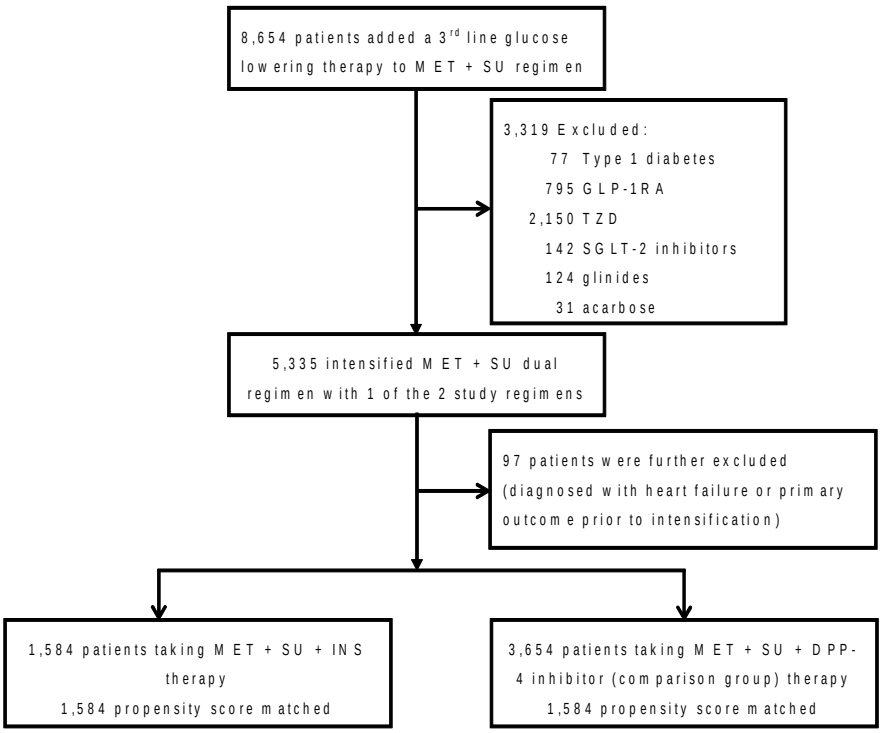


Table 7.1 Baseline characteristics of patients

	Cohort					
	Full			Propensity Matched		
Baseline variable	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 3654)	Std. diff [a]	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 1584)	Std. diff [b]
Demographics						
Age (yrs), Mean (SD)	53.7 (14.8)	56.6 (11.5)	-0.18	53.7 (14.8)	54.4 (12.8)	-0.05
Gender, No. (%)						
Male	770 (49)	2163 (59)	-0.16	770 (49)	768 (48)	0.00
Female	814 (51)	1491 (41)	0.16	814 (51)	816 (52)	0.00
Deprivation, No. (%)						
Least deprived	294 (19)	802 (22)	-0.07	294 (19)	302 (19)	-0.01
Less	300 (19)	733 (20)	-0.03	300 (19)	336 (21)	-0.06
Average	333 (21)	782 (21)	0.00	333 (21)	295 (19)	0.06
More	352 (22)	711 (19)	0.05	352 (22)	349 (22)	0.01
Most deprived	305 (19)	626 (17)	0.04	305 (19)	302 (19)	0.01
Parameters, Mean (SD)						
HbA1c (%)	9.9 (2.9)	9.2 (2.7)	0.18	9.9 (2.9)	9.8 (3.7)	0.03
BMI (kg/m2)	29.8 (6.7)	32.3 (6.3)	-0.30	29.8 (6.7)	29.8 (6.1)	0.00
Weight (Kg)	84.6 (20.4)	93.1 (20.3)	-0.32	84.6 (20.4)	84.5 (19.7)	0.00
SBP (mmHg)	132.6 (17.5)	135.0 (15.2)	-0.11	132.6 (17.5)	132.7 (15.6)	-0.01
DBP (mmHg)	79.6 (10.5)	80.4 (9.3)	-0.07	79.6 (10.5)	79.8 (9.2)	-0.03
TC (mmol/l)	5.1 (1.6)	4.8 (1.3)	0.16	5.1 (1.6)	5.1 (1.8)	0.00
HDL-C (mmol/l)	1.2 (0.4)	1.1 (0.3)	0.12	1.2 (0.4)	1.2 (0.3)	0.00

	Cohort					
	Full			Propensity Matched		
Baseline variable	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 3654)	Std. diff [a]	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 1584)	Std. diff [b]
LDL-C(mmol/l)	2.8 (1.1)	2.7 (1.0)	0.14	2.8 (1.1)	2.8 (1.1)	0.04
TGC (mmol/L)	2.9 (5.8)	2.5 (3.2)	0.06	2.9 (5.8)	2.8 (6.0)	0.01
Serum albumin (g/L)	42.0 (4.3)	42.8 (3.7)	-0.17	42.0 (4.3)	42.0 (3.9)	-0.01
eGFR (mls/min/1.73m2)	74.1 (19.0)	75.8 (16.7)	-0.08	74.1 (19.0)	74.4 (17.3)	-0.02
ACR (mg/mol)	4.8 (11.4)	3.6 (9.0)	0.11	4.8 (11.4)	4.2 (9.7)	0.06
Diabetes duration (yrs) [c]	2.6 (4.6)	2.7 (3.1)	-0.03	2.6 (4.6)	2.5 (3.8)	0.01
Smoking status, No. (%)						
Non-smoker	619 (39)	1604 (44)	-0.07	619 (39)	602 (38)	0.02
Current smoker	435 (27)	669 (18)	0.16	435 (27)	458 (29)	-0.03
Ex-smoker	530 (33)	1381 (38)	-0.07	530 (33)	524 (33)	0.01
BMI Categories, No. (%)						
$\leq 30 kg/m^2$	918 (58)	1447 (40)	0.29	918 (58)	910 (57)	0.01
$30\text{-}34.9 kg/m^2$	354 (22)	1167 (32)	-0.18	354 (22)	368 (23)	-0.02
$\geq 35 kg/m^2$	312 (20)	1040 (28)	-0.16	312 (20)	306 (19)	0.01
Medications, No. (%)						
Aspirin	220 (14)	734 (20)	-0.14	220 (14)	242 (15)	-0.04
Antihypertensive	587 (37)	1873 (51)	-0.22	587 (37)	581 (37)	0.01
LLT	608 (38)	2136 (58)	-0.31	608 (38)	585 (37)	0.03
Comorbidities, No. (%) [d]						
Other CHD [e]	38 (2)	71 (2)	0.00	38 (2)	63 (4)	-0.10*
PAD	29 (2)	49 (1)	0.02	29 (2)	40 (3)	-0.05

	Cohort					
	Full			Propensity Matched		
Baseline variable	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 3654)	Std. diff [a]	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 1584)	Std. diff [b]
Heart Failure	31 (2)	36 (1)	0.05	31 (2)	39 (2)	-0.03
Hypoglycaemia [e]	124 (8)	142 (4)	0.10	124 (8)	158 (10)	-0.09*

[a] Standardized differences are the absolute difference in means or percentage divided by the standard deviation of the treated group
[b] Resulting standardized difference after 1:1 matching based on average treatment effect on treated (ATT) propensity score technique and robust variance estimation. See Appendix C-3 for graphical illustration of balance
[c] Diabetes duration is time from first diagnosis of diabetes to date of intensification with 3rd line drug (index date)
[d] Comorbidities are defined in Appendix B-1
[e] In the matched cohort, only CHD and hypoglycaemia had statistically significant standardized difference at 0.10 level

Patients had a mean age of 56 yrs and were 56% male. Compared with patients who added a DPP-4 inhibitor, those who added INS to MET+SU had higher mean HbA1c levels (9.9% vs 9.2%). PS matching resulted in the inclusion of 3,168 patients (1,584 MET+SU+INS matched 1:1 with MET+SU+DPP-4 inhibitor). Before PS matching, many of the measured covariates had a standardized difference above the 0.10 level (Table 7.1). However, the application of PS matching brought into balance the distributions of the measured covariates. Apart from previous hypoglycaemia and a diagnosis of other CHD, the baseline characteristics of the matched sample were not statistically different; as a result, the systematic differences between INS and DPP-4 inhibitor subjects in the original sample have been substantially reduced or eliminated in the matched sample (Table 7.1). This shows that the differences between the treatment groups have been reduced by PS matching and adequate balance on baseline covariates has been induced by the specification of the PS model used.

7.4.2 Time to composite outcome

The time to a composite outcome in the cohort of patients is summarised in Table 2. The median time before the composite outcome among the DPP-4 inhibitor users was longer at 2.4 years (IQR: 1.1-3.8) compared to INS users (2.1 years, IQR: 0.9-3.6). The survival analysis showed the 5-year cumulative incidence of composite outcome was 9% with DPP-4 inhibitor and 23% with INS.

7.4.3 Estimating survival curves and survival effects

Crude KM survival curves for INS treated subjects and the reference DPP-4 inhibitor treated subjects in the full original sample are reported in Figure 7.2. The result showed there was a significant difference between the two curves; log-rank test $p<0.001$. The KM survival curves obtained from the PS matched sample are also summarized, with stratified log-rank test $p<0.001$ (Figure 7.3).

Figure 7.2 Full cohort KM survival curves

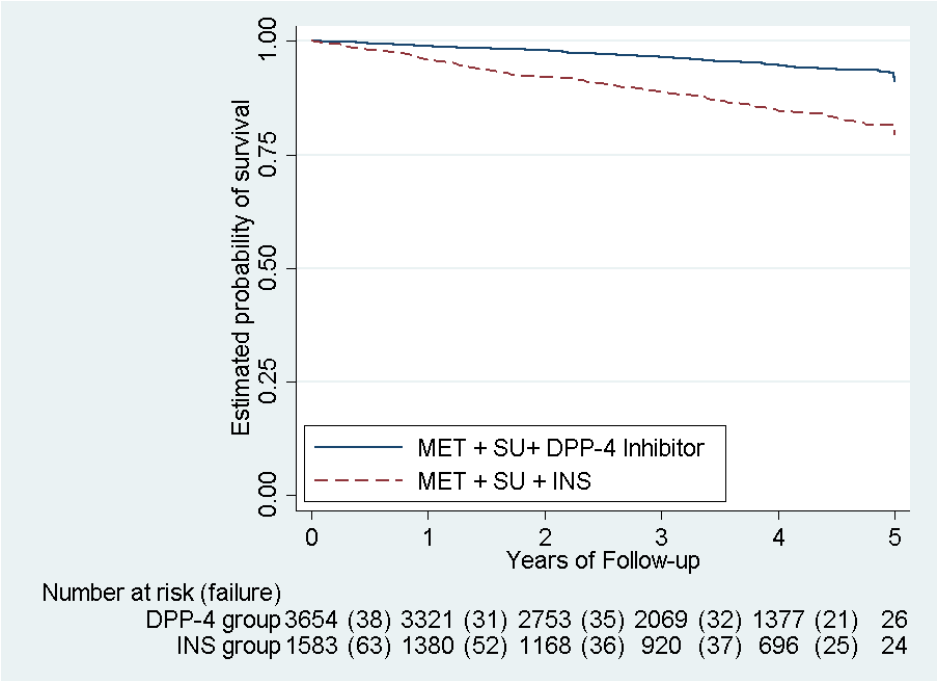


Figure 7.2 shows the Kaplan-Meier survival curves in treated (MET + SU + INS) and reference treatment group (MET + SU + DPP-4 inhibitor) participants in the original sample. The two survival curves are significantly different from one another (log-rank test P value < 0.001)

Figure 7.3 KM survival curves for PS matched cohort

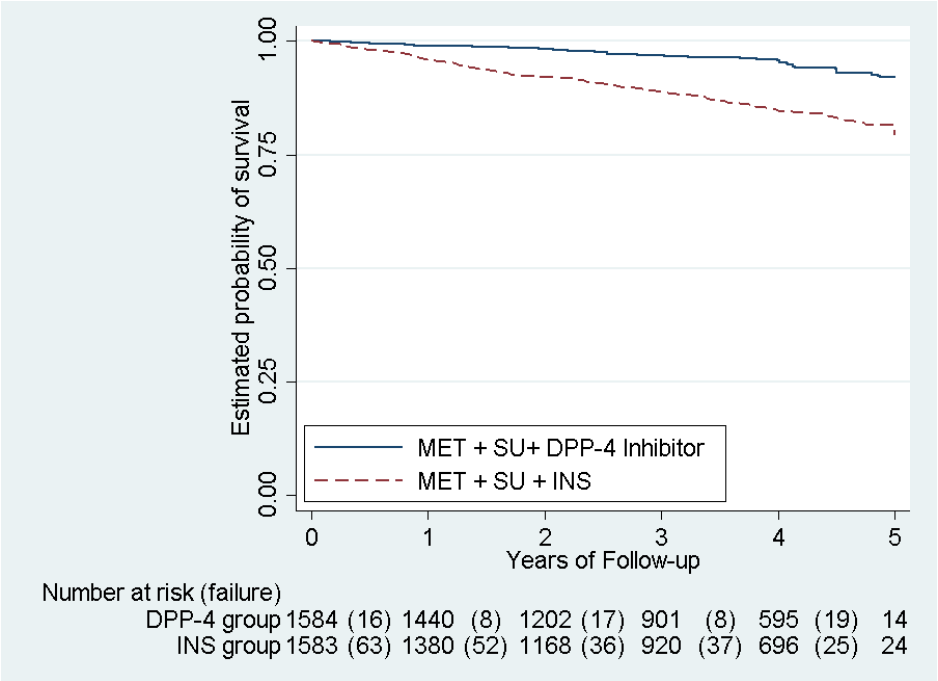


Figure 7.3 depicts the Kaplan-Meier survival curves in treated (MET + SU + INS) and reference treatment group (MET + SU + DPP-4 inhibitor) in the propensity score matched sample. The survival curves are significantly different from one another (stratified log-rank test P value < 0.001)

From the estimated survival curves, this data show that patients who intensified treatment with INS were significantly more likely to experience a composite outcome than those who added a DPP-4 inhibitor. For example, from the matched sample, the probability of dying or experiencing a CV event at 3 yrs was 0.11 (95% CI, 0.10-1.03) with INS and 0.03 (95% CI, 0.02-0.04) with a DPP-4 inhibitor.

Table 7.2 Events, rates and hazard ratios of outcomes

	MET + SU + INS (n = 1584)	MET + SU + DPP-4 (n = 1584)
Person-years	5,193	11,694
Sample population		
Composite outcome (No. of events) [a]	231	171
Unadjusted rate (95% CI)	44.5 (39.1-50.6)	14.6 (12.6-17.0)
Adjusted hazard ratio (95% CI)	2.6 (1.9-3.4)	1.0 (Reference)
CV events (No. of events) [b]	95	94
Unadjusted rate (95% CI)	18.4 (15.1-22.5)	8.1 (6.6-9.9)
Adjusted hazard ratio (95% CI)	2.0 (1.5-2.8)	1.0 (Reference)
All-cause deaths (No. of events) [c]	124	64
Unadjusted rate (95% CI)	23.2 (19.5-27.7)	5.4 (4.2-6.9)
Adjusted hazard ratio (95% CI)	3.7 (2.7-5.2)	1.0 (Reference)
CV deaths (No. of events) [d]	9	5
Unadjusted rate (95% CI)	1.7 (0.9-3.3)	0.4 (0.2-1.0)
Adjusted hazard ratio (95% CI)	2.6 (0.8-8.9)	1.0 (Reference)
Subgroup population		
BMI 30-34.9Kg/m^2		
Composite outcome (No. of events) [a]	54	44
Unadjusted rate (95% CI)	46.4 (35.5-60.5)	11.8 (8.8-15.9)
Adjusted hazard ratio (95% CI)	3.6 (2.3-5.6)	1.0 (Reference)
BMI ≥ 35Kg/m^2		
Composite outcome (No. of events) [a]	30	37
Unadjusted rate (95% CI)	29.6 (20.7-42.4)	11.2 (8.1-15.5)
Adjusted hazard ratio (95% CI)	2.4 (1.4-4.0)	1.0 (Reference)

[a] Composite outcome include: non-fatal AMI, non-fatal stroke or all-cause death. For full regression model of sample population, see Appendix C-4
[b] CV events relate to non-fatal AMI, non-fatal stroke or CV related deaths
[c] All-cause deaths only include records with confirmed cause of death.
[d] CV deaths refer to deaths from all cardiovascular causes

Rates are calculated per 1000 person-years in all cases.

Overall, there were 123 and 171 composite outcome events among patients who added INS vs a DPP-4 inhibitor, respectively (44.5 vs 14.6 events per 1000 person-yrs). The rate of occurrence remained the same after matching. The adjusted hazard ratio (aHR) from the PS-matched model was 2.6 (95% CI: 1.9–3.4; p< 0.01, Table 7.2). A breakdown of the number of component outcome events showed the following: the number of CV events (non-fatal AMI, non-fatal stroke or CV-related deaths) was 95 and 94 among patients who added INS vs. a DPP-4 inhibitor, respectively (18 vs 8 events per 1000 person-yrs; p<0.01); all-cause deaths were 124 vs 64 events, respectively (23 vs 5 events per 1000 person-yrs, Table 7.2).

7.4.4 Subgroup and other analyses

In a subgroup of patients with BMI 30-34.9kg/m^2, the composite outcomes were 54 and 44 among patients who added INS vs a DPP-4 inhibitor, respectively (46 vs 12 events per 1000 person-yrs; aHR 3.6 (2.3-5.6); p< 0.01). The subgroup with BMI ≥ 35kg/m^2 had 30 and 37 composite events from intensification with INS vs a DPP-4 inhibitor (30 vs 11 events per 1000 person-yrs; aHR 2.4 (1.4-4.0); p< 0.01, Table 7.2).

Stratification analysis across baseline BMI categories showed the risk of composite outcome among obese (BMI 30-34.9kg/m^2) patients was not significantly different to normal BMI status (BMI ≤ 30kg/m^2) patients.

In terms of glycaemic response, INS vs DPP-4 inhibitor users showed absolute mean reduction in HbA1c of -1.3% vs -1.0%, respectively (P < 0.001). With the exception of the period between 48 weeks and 1 year, the mean reduction

in HbA1c was not significantly different between INS and DPP-4 inhibitor over time. (Figure 7.4)

Figure 7.4 Glycaemic effectiveness – HbA1c over time

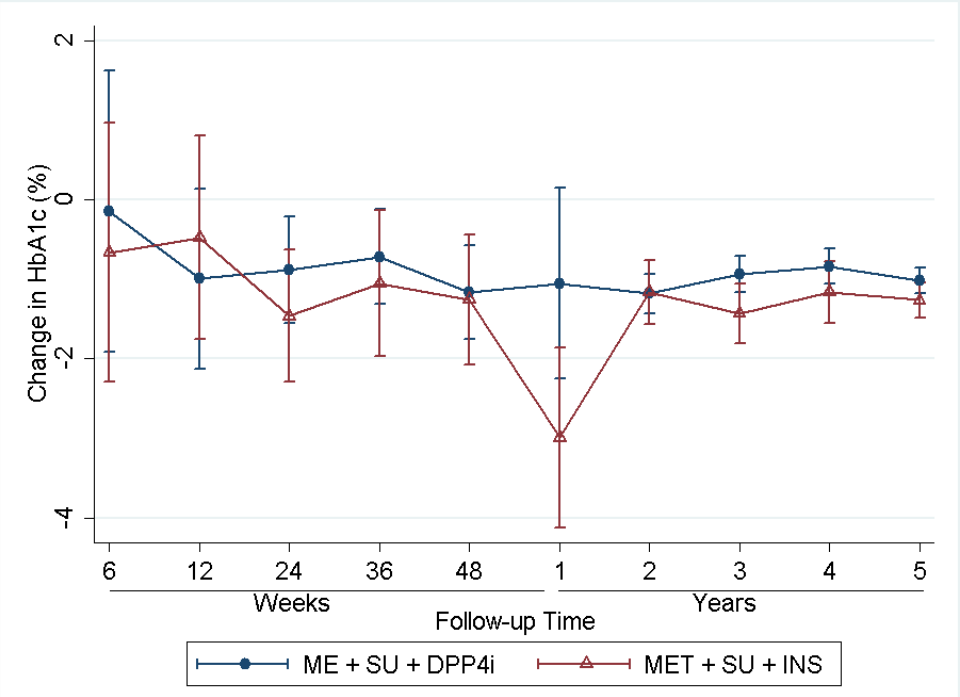


Figure 7.4 shows the addition of INS appearing to have a lower HbA1c reduction however, this was not statistically different to the addition of a DPP-4i, with the exception of changes at 1 year (-3% vs -1%, respectively)

Figure 7.5 Body weight changes over time

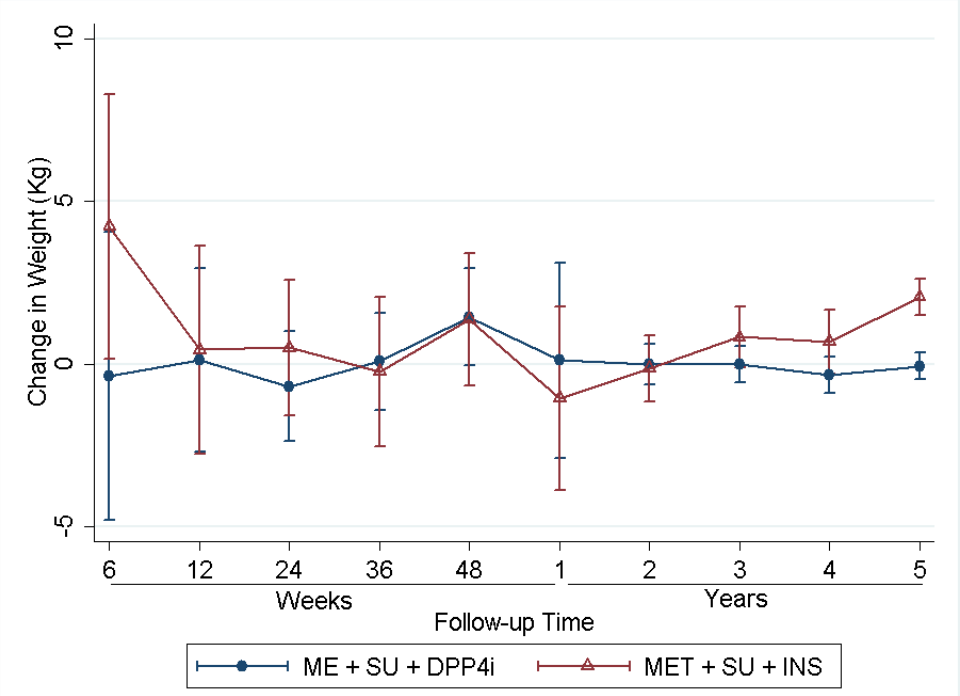


Figure 7.5 shows fluctuations in body weight changes within first 2 years of treatment between INS and DPP-4i, which were not statistically significantly different. INS users appeared to steadily put on weight after 1 year of therapy, while DPP-4i users remained neutral

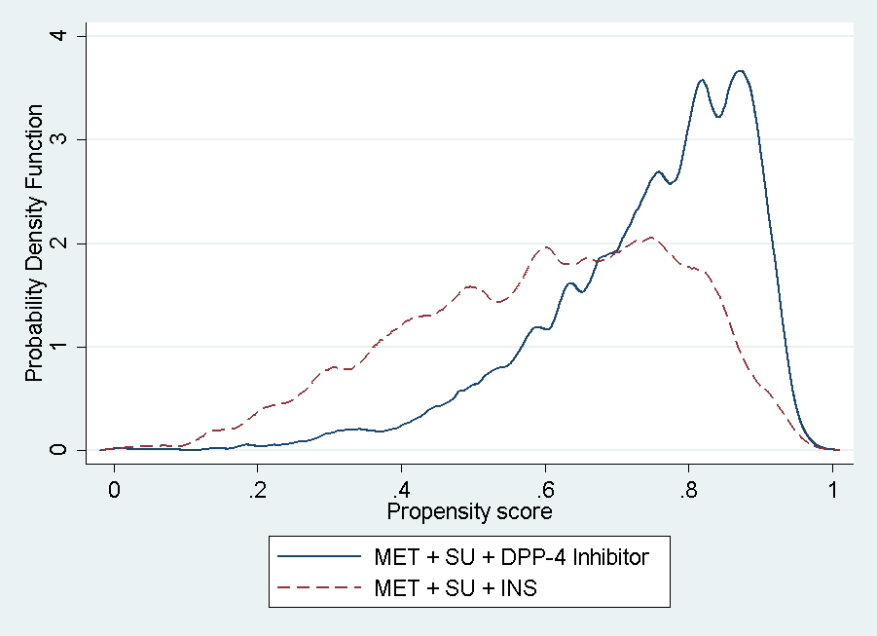
An absolute significant body weight increase was observed with INS (1.2kg, P < 0.001), whereas DPP-4 inhibitor showed a non-significant weight loss (-0.1kg, P = 0.5). From this data, INS users appeared to have consistently gained weight after the first year of treatment intensification. (Figure 7.5)

Sensitivity analyses

One of the analyses was to assess how strongly an unmeasured confounder would have to be associated with treatment selection in order for a previously statistically significant treatment effect to become statistically non-significant

if the unmeasured confounder had been accounted for. However, a large majority of estimated effects of covariates in the study were not statistically significant. Therefore, this sensitivity analysis was not employed. Moreover, the P value for the stratified log-rank test in the matched cohort was p<0.001. Hence, the small p-value obtained in the primary analysis cannot be taken as an indication that the study is insensitive to unmeasured confounders.

Figure 7.6 Distribution of propensity score by treatment group



The sensitivity analysis on missing data yielded comparable results to complete case models (aHR 2.3 (1.7-3.0); p< 0.001), which reflects results that are unlikely to be attributable to bias from missing information. The probability density functions of the PS matching of the treatment groups show there was no violation of the overlap assumption.[173] (Figure 7.6)

7.5 Discussion

This study has shown that in routine clinical practice, among patients with T2DM receiving MET+SU as dual therapy, intensification of diabetes treatment with the addition of INS compared with the addition of a DPP-4 inhibitor was associated with a significantly increased risk of a composite outcome of non-fatal AMI, non-fatal stroke and all-cause death. Furthermore, the increased risk with INS was even higher among the subgroup of patients who were obese (BMI>30). In the absence of a consensus on which third-line treatment is most appropriate when maximum tolerated doses of MET+SU fail to maintain adequate glycaemic control, this study identifies important differences in CV and mortality outcomes between two treatment options that are frequently used in patients with dual therapy failure.

The risk-benefit balance and overall safety of a more intensive treatment strategy in T2DM has recently been questioned,[203] and in particular the use of INS has been associated with an increase in life-threatening hypoglycaemia risk and mortality.[27] Other observational studies have also raised concerns about INS use in T2DM. For example, a dose-response relationship between INS exposure and all-cause deaths was reported in a large Canadian population,[194] and worse survival was reported among INS treated patients (relative to those on MET+SU) in a study exploring the relationship between HbA1c and CV disease.[195] More recently, adverse CV events and increases in all-cause mortality were reported in a cohort of patients who

received INS compared with other agents,[197] and among those whose treatment was intensified to INS (compared with adding a SU) following failure of metformin monotherapy.[193] However, an important limitation of these observational studies was their inability to control for differences in HbA1c,[194] hidden confounders or allocation bias[193,195-197] because they compared INS therapy with either MET or SU, both of which are often used much earlier in the course of the disease. The present study overcomes many of these limitations and specifically compared outcomes in a cohort of dual therapy failure patients without prior evidence of CV disease.

The observation of an even higher hazard ratio for the composite of non-fatal AMI, non-fatal stroke and all-cause death among the obese subgroup is clinically important. Insulin therapy is associated with weight gain, thereby increasing the amount of insulin required to control hyperglycaemia,[204,205] at the expense of further weight gain, increased insulin resistance and potentially increased risk of CHD.[206] This study population still had suboptimal glucose control (HbA1c > 8%) despite treatment intensification. The study data shows the participants have poor response to glucose-lowering treatment intensification, although this may not apply to most other populations. Thus, patient factors associated with persistently high HbA1c might be an important determinant of increased mortality risks and may need to be further investigated. It is assumed that patients in this study may have required high dose insulin treatment in order to achieve glycaemic targets. A previous study

has shown that the effectiveness of insulin therapy to lower HbA1c levels among overweight patients with diabetes is reduced.[164]

Limitations of the study

The analyses were subject to a number of limitations that are inherent to observational studies. For example, it was not possible to establish if patients were fully compliant with their medication. Factors that may influence the decision to treatment intensification such as tolerability and safety, or cost could not be assessed due to unavailability of these records. In addition, covariates were mainly included as baseline parameters and their effects were not assessed during the follow-up period. Some of these time-varying variables are relevant during the entire observation period for monitoring outcomes. Although potential residual confounders such as compliance, indications for intensification treatments, markers of β-cell deterioration and differences in dosages were not assessed, robust analytical techniques were used to account for differences in the observed covariates and to control for confounding that may bias the results of the estimated treatment effects. The use of propensity score matching to estimate average treatment effect in the dataset contributed to the balancing of treatment and comparison groups on the available covariates. However, this technique only accounts for observed covariates. Even though a thoughtful and thorough specification of the selection model was employed in successfully applying the propensity score matching technique and minimising bias, the study findings must be interpreted with caution in light of the above limitations.

Conclusion

Comparative effectiveness studies and RCTs which examine the risks of cardiovascular events or deaths from the co-administration of INS or DPP-4 inhibitor as 3rd line regimens are not reported. This data has shown that among patients with diabetes who are receiving metformin and sulfonylurea therapy, the addition of insulin compared with DPP-4 inhibitor was associated with an increased risk of a composite of non-fatal cardiovascular outcomes and all-cause mortality. The observed excess risk of adverse cardiovascular events was increased in patients who are obese. Insulin is a still a very important treatment option in the management of T2DM and outcomes from this study are not clinically applicable until RCTs comparing insulin with DPP-4 inhibitors have been conducted. These findings require further investigation to clarify the risk associated with insulin regimen in view of the availability of newer GLTs.

Chapter 8: DPP-4 inhibitors and bone fracture: review and meta-analysis

8.1 Summary

Aim

Fracture risk is higher in older adults with T2DM. Oral GLTs have different effects on bone metabolism. The purpose of this study is to appraise the evidence from literature and determine the effect of DPP-4 inhibitor on the risk of developing bone fractures.

Methods

Using Boolean search terms, the search strategy combined synonyms of 'fracture' and 'DPP-4 inhibitor'. Comprehensive electronic databases which include EMBASE (1974-2015), MEDLINE (1946-2015), the EMA and the WHO ICTRP databases were searched for RCTs which compared a DPP-4 inhibitor with an active comparator or placebo amongst patients with type 2 diabetes. A meta-analysis was performed to compare DPP-4 inhibitor with either an active comparator or a placebo. The outcome measure was the presence or absence of fracture.

Results

The search yielded 5,061 records relating to fractures and DPP-4 inhibitor, from which 51 eligible RCTs were selected for meta-analysis (N=36,402). Thirty-seven (37) studies compared DPP-4 inhibitor with placebo (n=23,974), while fourteen (14) studies (n=12,428) compared DPP-4 inhibitor with an active comparator. The mean age of patients was 57.5±5.4 years, the average HbA1c was 8.2%, while the average BMI was 30±2 kg/m^2. Overall, there was no significant association of fracture events with the use of DPP-4 inhibitor when compared with placebo (OR; 0.82, 95%CI 0.57-1.16, P = 0.9) or when DPP-4 inhibitor was compared against an active comparator (OR; 1.59, 95% CI 0.91-2.80, P=0.9).

Conclusion

This study offers a larger, up-to-date review of the subject. The meta-analysis showed that there was no significant association between DPP-4 inhibitor use and the incidence of fractures.

8.2 Introduction

Patients with T2DM are associated with an increased risk of developing bone fractures.[207,208] This is related to a variety of factors, such as recurrent falls[209] due to diabetes-related co-morbidities such as retinopathy, loss of balance, neuropathy and hypoglycaemic events,[207,210,64] as well as hyperglycaemia induced alterations in tissue composition, leading to osteoporosis and bone fragility.[64] Glucose-lowering medications such as TZDs have been reported to

reduce bone density[211,212] and increase the risk of fractures.[44,50,68] In addition, Insulin therapy is also associated with an increased fracture risk,[69,213,214] despite the neutral effect it exhibits on bone density.[215]

Experimental studies suggest that the incretin hormone glucagon like peptide-1 (GLP-1) and gastric intestinal polypeptide (GIP) are capable of increasing bone density in animal models.[38,45] GLP-1 has also been reported to induce osteoblast differentiation and inhibits osteoclast activity.[14,15] DPP-4 inhibitor is a widely used GLT that inhibits the breakdown of these incretin hormones and therefore induce a rise in the level of these incretin hormones and may exert protective effects on the bone.[36] Although bone mass density (BMD) has been shown to predict fracture incidence, BMD does not necessarily give a full picture of bone quality and strength.[64] For example, a meta-analysis of over 65 studies showed that BMD was decreased in patients with Type 1 diabetes mellitus (T1DM), but increased in those with T2DM.[216] However, despite this higher than normal BMD, patients with T2DM remains at an increased risk of fractures, by around 20% in both sexes.[217]

A recent review and meta-analysis of trial studies suggests that treatment with DPP-4 inhibitors could be associated with a reduced risk of bone fractures.[46] Monami et al indicated that DPP-4 inhibitors, when compared with placebo or comparator treatments, were associated with fewer fracture events. The basis for this association may be explained by the protective effect of DPP-4 inhibitors on the bone. In a 2014 animal model study by Glorie et al.,[218] the use of sitaglitpin in diabetic male rats increased trabecular bone volume, cortical bone volume and BMD. The loss of bone strength was attenuated, and bone

biomarkers indicated a decrease in bone resorption. These findings are also supported by a Kyle et al. study, where high-fat diet-fed mice treated with sitagliptin showed an increase in vertebral BMD.[219]

A previous review and meta-analysis of 28 RCTs enrolled 11,880 and 9,175 patients for DPP-4 inhibitors and comparators, respectively.[46] Monami et al[46] examined the risk of bone fractures. DPP-4 inhibitors, compared with placebo or other treatments, were reported to be associated with a reduced risk of fractures (MH-OR: 0.60, 95% CI: 0.37-0.99) at a borderline P value of 0.045. At the time of previous systematic reviews, there was scarce research directly investigating the effect of DPP-4 inhibitors on fracture incidence. Fractures are often noted as adverse events, rather than primary endpoints. All other OADs reduce HbA1c, however, they have not been shown to have the same protective effect on bone as DPP-4 inhibitors. Therefore, other factors independent of glycaemic control may influence fracture risk.

Further investigation into the trials included in Monami et al[46] suggested 3 of the RCTs did not mention if fracture events occurred during the study.[220-222] In addition, considering that bone fracture takes time to occur, the majority of RCTs included in the previous review had a follow-up averaging 24 weeks, with only 7 RCTs having ≥52 weeks included in the meta-analysis, the limitations associated with the previous review by Monami et al emphasised the necessity of an updated meta-analysis. Therefore, it was necessary to undertake an updated systematic review and meta-analysis. This is especially considering that since the 2011 meta-analysis, a number of new robust RCTs investigating DPP-4 inhibitors have been published. The purpose of this study

is to obtain an updated review and meta-analysis of the literature, to identify if DPP-4 inhibitors are associated with a decreased risk of bone fractures. Therefore, it is hypothesised that DPP-4 inhibitors will have a protective effect against fractures, as already demonstrated in an earlier review.

8.3 Methods

8.3.1 Data sources and search strategy

A series of searches were performed, investigating the association between fracture incidence and the use of DPP-4 inhibitors in patients with T2DM. Comprehensive electronic databases were searched. Using Boolean search terms, the search strategy combined synonyms of 'fracture' and 'DPP-4 inhibitor'. The data sources include EMBASE (1974-2015) and MEDLINE (1946-2015). Reviews of approved drugs were identified manually on the European Medicines Agency (EMA) database. WHO International Clinical Trials Registry Platform (ICTRP), including clinicaltrials.gov, was searched to identify clinical trials and unpublished studies. There was also manual searching of the supplementary data within the Monami et al., 2011 study.[46] The full database search strategies employed are found in Appendix A.

8.3.2 Study selection and eligibility

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) 2009 checklist was used to structure the method of the review.

The studies were assessed first by the titles, then by the abstract and followed by full text review. Titles that were included mentioned synonyms for DPP-4 inhibitors, and the patient group, i.e., T2DM. Titles contradicting the search terms, by mentioning the wrong intervention, and conditions that were not T2DM were excluded. Within the clinicaltrials.gov searches, only the titles with the recruitment statuses, 'not recruiting' and 'authorised' were selected. Studies that were 'not yet recruiting', unfinished studies and studies without any data published were excluded.

After collecting the relevant abstracts, the articles were assessed according to eligibility criteria. The aim of the search was to identify RCTs that compared the intervention with a comparator drug or placebo, amongst T2DM patients. At this stage, non-RCT, randomised or controlled trials were excluded. Only RCTs were used as these are higher quality studies with better control and replicability. The intervention had to be a DPP-4 inhibitor, and the outcome measure recorded was the number of fractures. The studies were manually searched for the term, 'fracture' to identify if fractures had occurred during the trial. If this was not present, then the study would be excluded. Non-English articles were not included. Only human studies were used, as these were the most relevant to application of results to patients with T2DM. There was no restriction on the publication date of articles, as all studies up until May, 2015 were included.

8.3.3 Data abstraction and quality assessment

The data of interest were the number of fracture events during the trials. A data extraction form was utilised to record different characteristics of each study. This included: the author, year of publication, clinical trial number, location of the population, the duration of the trial (weeks), the intervention and comparator used, the number of participants in each group, sex (% of females), the mean age, the mean baseline HbA1c (%), mean baseline BMI (kg/m^2), hypoglycaemic events (% of participants), and the number of fractures in both the intervention and the comparator groups.

The JADAD scale was used to assess the quality of each RCT. It evaluates the method and appropriateness of randomisation, blinding and following up of each participant in the study. This produces a number for every study, with a higher score indicating a higher quality of study design.[223]

8.3.4 Data synthesis and analysis

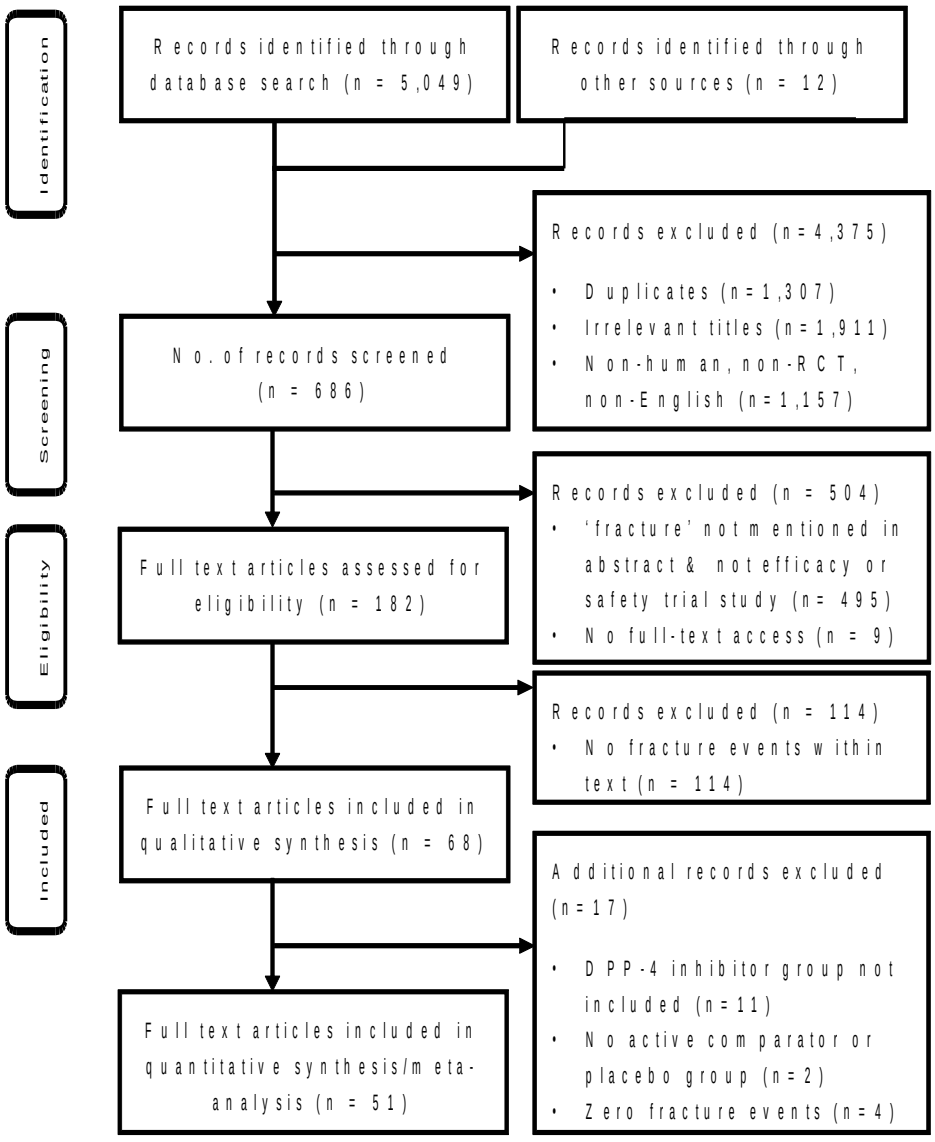
Eligible studies that were used in the meta-analysis were based on whether they included a DPP-4 inhibitor treatment, in addition to an active comparator or placebo treatment. Studies were also included that involved combination therapy within the intervention, active comparator or placebo arms. A meta-analysis was performed to compare DPP-4 inhibitors with either an active comparator or a placebo. The outcome measure was dichotomous, with the presence or absence of fracture. The Mantel-Haenszel odds ratio (MH-OR) model was used to calculate odds ratios for all the studies, as well as 95%

confidence interval (CI) levels. Heterogeneity was assessed by the calculation of value.

8.4 Results

The initial database searches identified 5,049 records relating to fractures and DPP-4 inhibitors. In addition to this, 12 studies were manually identified from a previous meta-analysis.[224] From these, 51 studies met the criteria for inclusion in the meta-analysis as outlined in Figure 8.1. The patient characteristics of the eligible studies are summarised in Appendix A-4 through A-6.

Figure 8.1 Study screening



8.4.1 Study characteristics

The searches resulted in 51 RCTs with 36,402 participants. Thirty-seven (37) studies compared a DPP-4 inhibitor with a placebo, involving 23,974 patients. Fourteen (14) studies including 12,428 patients were used in the comparison of a DPP-4 inhibitor against an active comparator. The mean (SD) age of patients was 57.5 (5.4) years and 47% of the entire population was female. Participants had an average (SD) BMI of 30.2 (2.0) kg/m^2. The average HbA1c was 8.2%

and the percentage of patients who experienced a hypoglycaemic event was lower in the intervention compared to the comparator group (4.3% vs 5.3%, respectively.

There were 39 multi-country trials, nine multisite trials, and two trials in the US. The multisite trials occurred in Europe, a number of Asian countries, Puerto Rico and the US. The durations of the studies ranged from 12 to 205 weeks, with an average of 37.5 weeks. There were 29 RCTs with a duration < 52 weeks, and 22 studies were >52 weeks. The earliest dated study was in 2006,[225] with the most recent study being published in 2014.[159] Further information are summarised in Appendix A-4.

Intervention

The DPP-4 inhibitors investigated in each RCT are as follows: 5 studies with alogliptin, 3 with linagliptin, 12 studies with saxagliptin, 29 with sitagliptin and two studies with vildagliptin. Considering the placebo groups, a total of 28 studies (55%) had placebo in combination with another drug. Within the active comparator groups, five studies (35%) were combination therapies.[179,191,226-228] Eight studies involved the use of MET, and another eight studies involved the use of a sulfonylurea. Two studies included a TZD,[179,229] while one study included a GLP-1[228] and one study involved voglibose, an alpha-glucosidase inhibitor (AGI).[230] Overall, there were 86 cases of fracture in the intervention group, and 64 cases with the comparators. The most common types of fracture were lower limb (12 cases), followed by ankle (10 cases) and then rib (8 cases).

Outcomes

All studies had a primary outcome measure of HbA1c change from baseline. This is with the exception of Alba et al.,[231] who measured changes in α-cell and β-cell function; insulin secretion rate in the clinical trial study NCT00374907;[232] the percentage of individuals experiencing a primary major cardiac event (MACE) in NCT00968708;[233] and the proportion of patients under a HbA1c of 7.0% without any sings of severe hypoglycaemia, as measured in NCT01006603.[234] The secondary outcomes were varied, but mostly included efficacy assessments such as change in glucose, insulin, proinsulin, fasting plasma glucose (FPG), 2-hour post-meal glucose (PMG), body weight and fasting lipids. Two studies included safety assessments such as blood pressure and cases of secondary MACE.[233,235]

RCT exclusion and inclusion criteria

The studies had mostly similar inclusion criteria, with no restrictions on the gender of the patient. Studies were restricted to patients with T2DM with inadequate glycaemic control either with diet and exercise alone. The mean range of requirements, inclusively were an HbA1c level from 7 to 10%; a BMI between 22 to 43kg/m^2; and age 22 to 78 years. Eleven studies required the patients to be unlikely to conceive, or to use contraception. Two studies set limitations for blood pressure (BP),[236,237] only including participants with a BP up to 170/105mmHg. On average, three studies required patients to have had diabetes duration for a minimum of 2 years.[231,238,239] Four studies included patients with renal impairment.[240-243] One study,[233] used patients with acute coronary syndrome between 15 and 90 days prior to the study. Two studies

only included patients that would have the ability to use home blood glucose monitoring,[226,244] and one study included the patients able to use an injection device.[228]

Exclusion criteria differed across the study, but the most commonly encountered criteria were; a history of CV event; females who are pregnant or breastfeeding; patients that refused the use of contraception; the presence of hepatic and renal disease; secondary forms of diabetes; T1DM; a history of diabetic ketoacidosis; symptoms of poorly controlled diabetes; recent gastrointestinal surgery; the prior use of a weight loss drug treatment; the use of any other hyperglycaemic agents; alcohol and substance use; hypersensitivity to any of the treatments; and uncontrolled hypertension.

8.4.2 Bias assessment

Quality assessment

The JADAD scale was used to assess the quality of each study. As shown in Appendix A-8, thirteen (13) studies obtained a score of 5, thirty one (31) studies obtained a score of 4, while seven (7) studies obtained a score of 3, with no study obtaining a score below 3. The average score was 4, therefore the overall quality of RCTs was good. The scores for each item are fully expanded in Appendix A-7.

Randomisation and blinding

All studies were randomised, with 9 studies being randomised by computer generated allocation. 36 studies did not mention the method used for randomisation. 7 studies were stratified according to certain factors.

Forty three of the studies were double-blind with a matching placebo. Seven studies did not use a placebo, and one study was open label.[228]

Publication bias

The effect of publication bias was reduced by ensuring that the systematic review of published literature was as thorough as possible. Additional hand searching of references quoted by each paper was done to make sure all relevant literature was found. Unpublished data were also obtained and included, which may reduce publication bias, although this may also decrease validity as unpublished work has not been peer reviewed.

8.4.3 Meta-analysis

DPP-4 inhibitor and fracture incidence

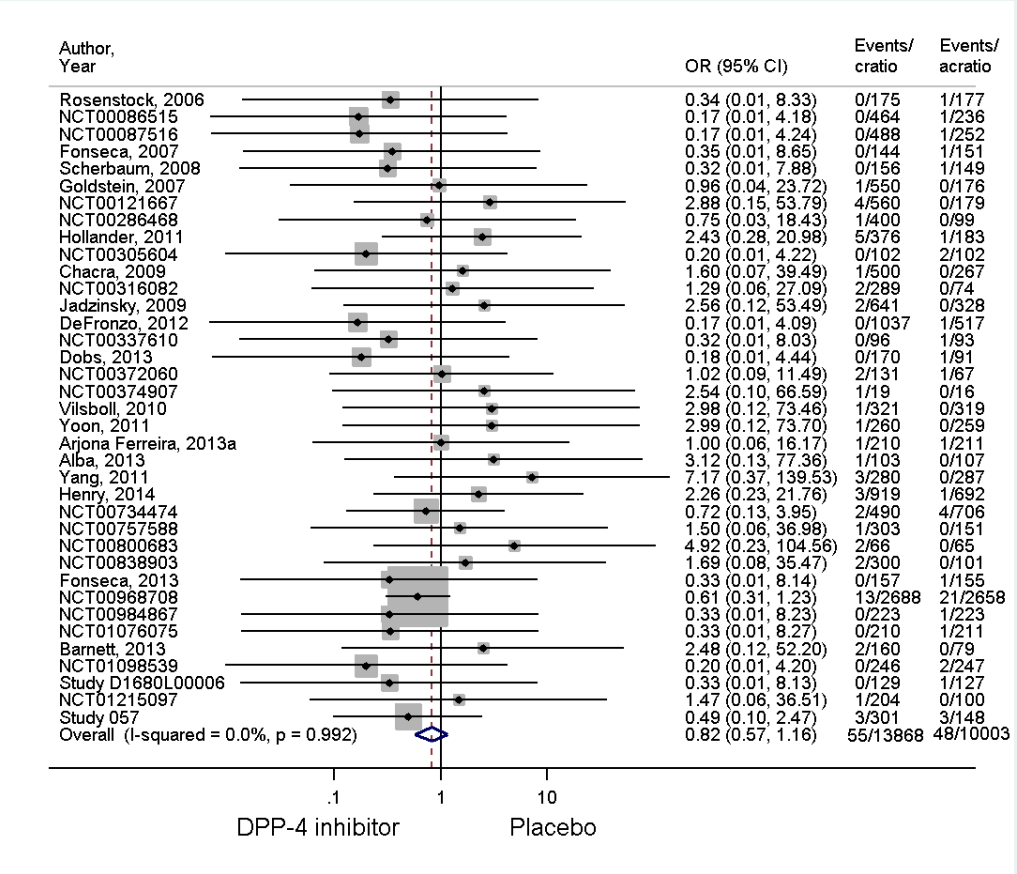
Statistical analysis was carried out in order to investigate the association of fractures with the use of DPP-4 inhibitors. The primary outcome was the effect of DPP-4 inhibitors on the incidence of bone fractures. Odds ratio (OR) was the measure of effect calculated based on the number of fractures that occurred in intervention and comparator or placebo groups. Two subgroup analyses were performed, comparing the incidence of fractures with DPP-4 inhibitors; and active comparators or a placebo. Overall there was no significant association of fracture events with the use of DPP-4 inhibitors.

DPP-4 inhibitor vs placebo

A subgroup analysis comparing DPP-4 inhibitors to placebo is displayed by the forest plot in Figure 8.2. The I^2 value was equal to 0.0%, indicating no presence of heterogeneity. In view of this, a fixed effects model was used. The

analysis gave a non-significant P value of 0.9. The OR value was 0.82 with a 95% confidence interval of 0.57-1.16.

Figure 8.2 Forest plot of fracture cases among DPP-4 inhibitor vs. placebo



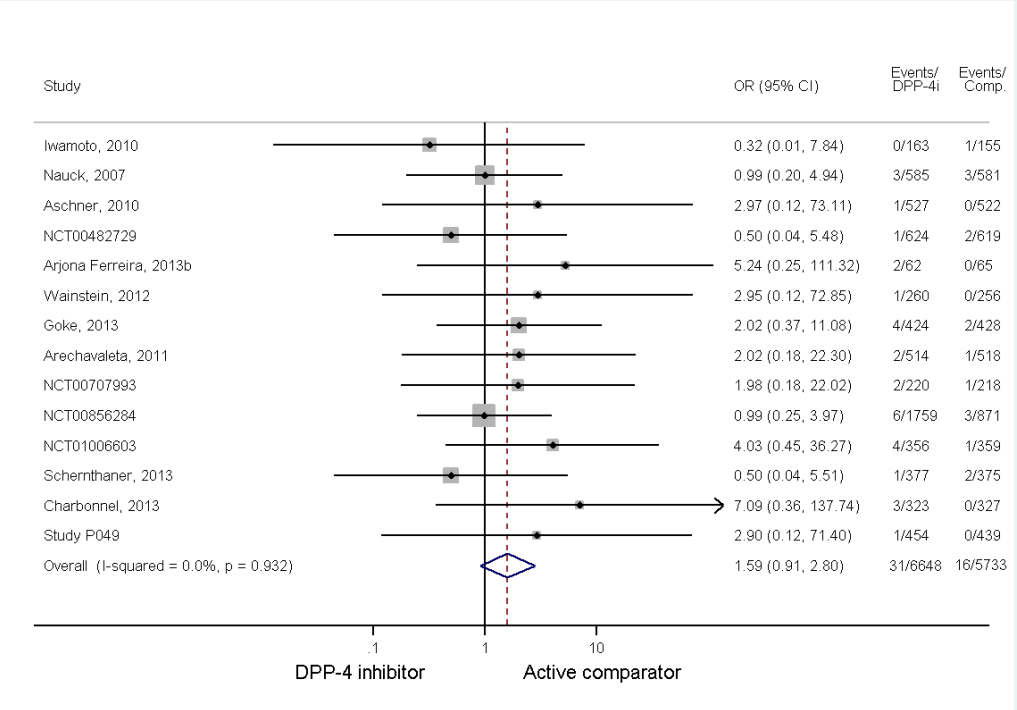
DPP-4 inhibitor vs active comparator

A subgroup analysis on DPP-4 inhibitors compared with an active comparator is displayed by the forest plot in Figure 8.3. The I^2 value was equal to 0.0%, indicating homogeneity. A fixed effects model was also used. The OR value was 1.59 with a 95% confidence interval of 0.91-2.80 (P value, 0.9)

This systematic review and meta-analysis studied the effect of DPP-4 inhibitors on fracture incidence in T2DM. The meta-analysis of 51 RCTs showed that there was no significant association between DPP-4 inhibitor use

and the occurrence of fractures when DPP-4 inhibitor is compared with placebo or active comparator.

Figure 8.3 Forest plot for DPP-4 inhibitor vs. active comparator



From the quality assessments, it can be concluded that the RCTs were well-designed, giving rise to evidence of a good strength. The forest plots and I^2 values of 0.0% in both meta-analyses indicated that there was no statistical heterogeneity. This means that any variability across the studies is attributable to chance, not to the heterogeneity of the studies themselves.[245] This consistency amongst studies should provide a confidence in applying these results.[246] Thus, this review can be considered to have a good level of internal validity. This review had a large, multinational sample, reflecting its external validity.

In the context of the research surrounding the role of incretin hormones in bone metabolism, this study demonstrates no significant reduction of fracture incidence with the use of DPP-4 inhibitors. This is in contrast to Monami et al.,[46] who reported that there was a statistically significant association between DPP-4 inhibitor use and reduced occurrence of fracture. From Monami et al., 17 of the 20 trials comparing DPP-4 inhibitors with placebo were used in this meta-analysis. Three trials were not included, since these did not mention that a fracture event occurred during the study.[220-222] There was also no mention that authors of the original studies were contacted for this information. Furthermore, the present meta-analysis included 22 trials with study duration of 52 weeks and above, whereas the previous review included only 7 trials with ≥52 weeks.[46]

Another meta-analysis of RCT studies by Su et al[247] examined the risk of bone fractures associated with GLP-1 receptor agonists (liraglutide and exenatide) when compared with placebo or active comparator treatment. Incident fracture data from 11,206 patients was pooled across 16 RCTs and the results showed treatment with liraglutide was associated with a significant reduction in the risk of bone fracture (MH-OR: 0.38, 95% CI: 0.17–0.87), while treatment with exenatide was associated with an increased risk of bone fractures (MH-OR: 2.09, 95% CI: 1.03–4.21).

The contrasting results that have emerged from these incretin-based therapy review studies show the need for further investigations into the role of incretin hormones in bone metabolism across various populations. The reason for this disparity between results from previous reviews and this current review study

cannot be explained from our data. However, we assume that other underlying factors such as lifestyle changes might play a significant role. The disparity in the results obtained from previous reviews of RCTs implies that that a definitive conclusion cannot be made on the effect of DPP-4 inhibitor on bone fractures in the long-term.

The SAVOR-TIMI 53 study by Scirca et al.,[42] was conducted across 16,492 T2DM patients for a median of 2.1 years. This large, multisite, double-blind RCT investigated cardiovascular outcomes in T2D patients, comparing the effects of saxagliptin with placebo. In keeping with this meta-analysis, it was found that there was no significant difference in the number of patients experiencing a fracture, between the saxagliptin and placebo group (P = 1.0). In addition, a large-scale retrospective cohort study by Driessen et al[248] directly investigated the effect of DPP-4 use on fracture risk. In this study, information from the Clinical Practice Research Datalink (CPRD) database on 216,816 patients was examined. It demonstrated no significant difference in the hazard risk of fractures, between DPP-4 inhibitor users and matched control patients (adjusted hazard ratio 0.89; 95% CI of 0.71-1.13).

The studies included in this review are not without their limitations. For example, RCTs had an average duration or follow-up of 37.5 weeks. This may be too short of a follow-up time to observe fracture events. In addition to this, fracture was not reported as a primary end-point in any of the studies, but only as an adverse event.

As in the previous review,[46] combination therapy was also included within the intervention, active comparator and placebo arms. Therefore, the inclusion of these combination therapies may have exerted an effect on the outcome, and may result to null as others increase fracture risk. In future, a meta-regression could be performed, in order to identify the significance of this effect. Nonetheless, the presence of combination therapies in many patients accurately reflects the reality that single drug therapy is often insufficient to control glycaemic levels.

Another limitation of this review was the unavailability of some data, especially within unpublished, yet disclosed trials from the clinicaltrials.gov website. If a published article referring to the same clinical trial identifier number could be identified, missing data were obtained from that source. In the case that there was no published article for a clinical trial, or no additional information within the published article, data were marked as 'not reported' (NR). It is acknowledged that this missing data may have given an incomplete picture of patient characteristics. However, the data that were missing did not relate to the outcome measure of this review.

Furthermore, it was assumed that the risk of fracture may be influenced by the use of TZD since this has been previously established. Therefore, a sensitivity analysis was conducted by excluding the study with participants being administered TZD as active comparator; however, there was no difference observed on the outcome. It was also thought that non-osteoporotic fractures such may not be clinically significant as these fractures are mainly caused by trauma rather than the result of osteoporosis. Therefore, bone fractures

incidents which were assumed to be caused by trauma such as ribs, skull and sternum fractures were excluded in a sensitivity analysis. The forest plots obtained showed a similar dispersion and the corresponding I^2 values and odds ratios were not different to the estimates obtained from those inclusive of all fracture types.

Figure 8.4 Funnel plot for DPP-4 inhibitor - Placebo

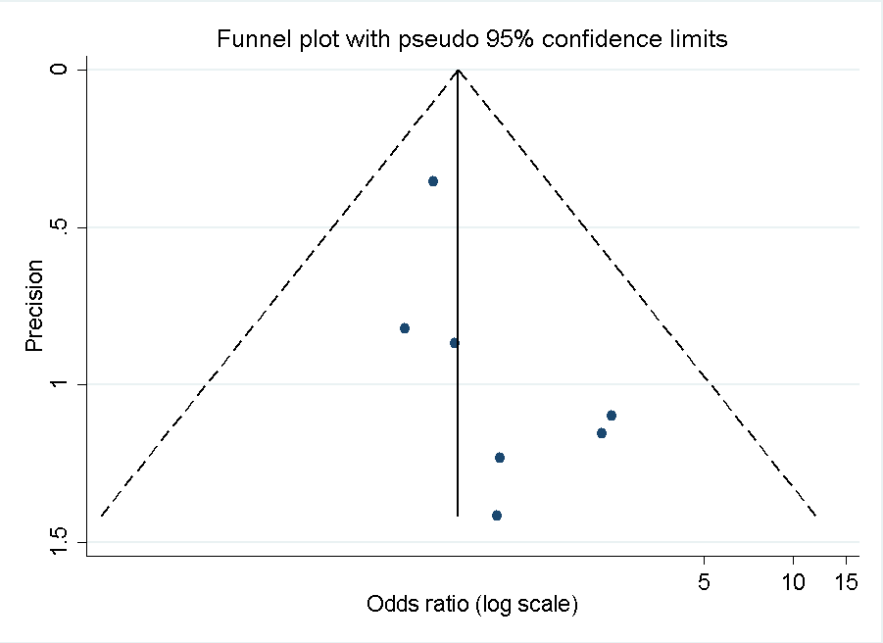
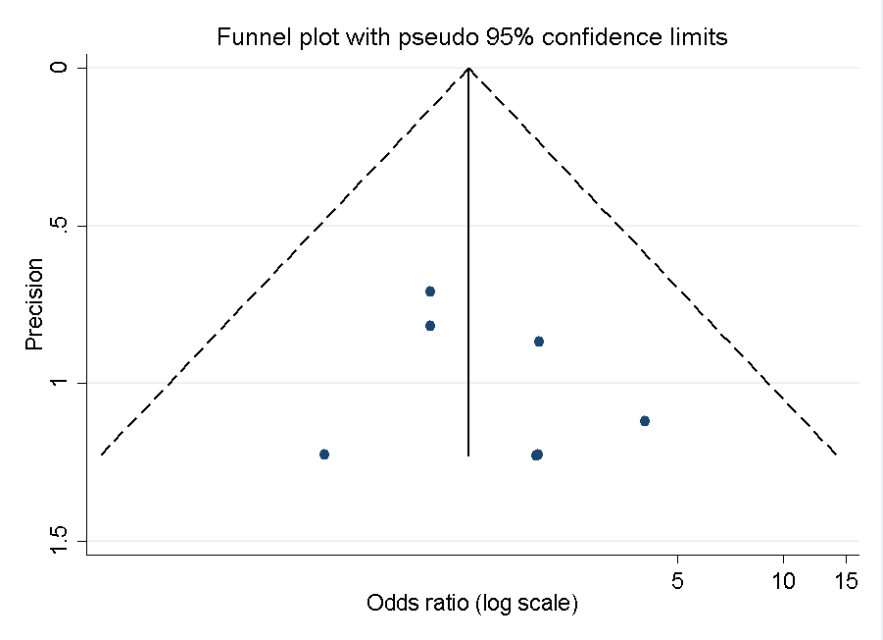


Figure 8.5 Funnel plot for DPP-4 inhibitor - Active comparator



The funnel plots in Figure 8.4 and Figure 8.5 show the funnels are approximately symmetrical, indicating the absence of publication bias.

The use of RCTs provides a high level of internal validity, improving the strength of recommendation for the practice of evidence-based medicine. Still, RCTs may not have a high level of external validity. This is since they may not accurately reflect a real-world environment where factors are not so tightly controlled. However, these rigorous methods allow the conclusion to be made, that the difference in the outcome being measured, is in fact due to a change in the independent variable. This systematic review has demonstrated the absence of a significant association between DPP-4 inhibitor and fracture incidence.

Chapter 9: Conclusions and recommendations

This thesis focused on DPP-4 inhibitor because it represents a major new class of oral GLT with substantial longitudinal data from routine clinical practice. DPP-4 inhibitor's metabolic profile for reducing HbA1c and maintaining a weight neutral effect potentially offer a number of unique clinical advantages for the management of T2DM. The fundamental argument of this thesis is that given the lack of evidence for long-term effectiveness of DPP-4 inhibitors as well as other glucose-lowering medications, the assessment of glycaemic response and safety of GLTs may best be addressed using well-designed long-term observational studies. The thesis was aimed at exploring how methodological problem of bias can influence retrospective cohort study design and employ strategies that can strengthen the quality of evidence from administrative or healthcare database studies which reflect the 'real-world' comparative effectiveness and safety situation where dual, triple or quadruple therapies are often used to achieve glucose target levels. In addition, meta-analysis was undertaken to complement observational data on the safety of DPP-4 inhibitors in clinical practice.

Chapter 9 highlights the major findings of the primary care database studies reported in previous Chapters of this thesis and how the Aims and objectives set out in 1.3 were met. It also outlines the importance of the studies, application to clinicians and researchers, and recommends some areas for future research.

9.1 Summary of evidence and recommendations

The review of literature conducted in Chapter 2 highlighted the growing number of observational studies being reported from healthcare records of patients with T2DM. It was observed that an estimated 32% of studies in this area reported at least one type of bias and provided an explanation on how the bias was addressed. A summary of the different types of biases identified in the review, and strategies that can be employed to minimise these biases were illustrated in the form of an algorithm in Figure 2.6. The eligible articles assessed were heterogeneous in their methodological approach and treatment outcomes reported. Therefore, a narrative summary of how the types of biases may influence retrospective cohort analyses was provided using examples from the relevant literature, which is tailored around individual study hypothesis.

Chapters 3 to 8 of this thesis fall into three categories; 1) description of methods used for the analysis of large healthcare database; 2) applications of advanced statistical and epidemiological methods in answering research questions around the clinical effectiveness and safety of DPP-4 inhibitors and 3) strengthening the evidence of the effect of DPP-4 inhibitors on the risk of bone fractures. These objectives were met through the results reported in

Chapters 4 to 8. Below are the key findings from these chapters, their implications and recommendation for future work.

In Chapter 4, a logistic regression model was developed to identify the predictors of response and non-response to the addition of DPP-4 inhibitor as an intensification option when glycaemic levels are not achieved with other oral GLTs (Objective 2). It was hypothesised that clinical, metabolic and demographic parameters may influence treatment response following DPP-4 inhibitor therapy among patients with suboptimal glucose control. Two potential factors that are associated with treatment response were identified: the addition of DPP-4 inhibitor to MET and the co-administration of DPP-4 inhibitor with MET + SU therapy. However, higher HbA1c at the time of treatment intensification and longer duration of diabetes were associated with the likelihood of not achieving Hba1c target. The findings of this study support the use of a DPP-4 inhibitor as a second line therapeutic option in patients whose glucose control remains suboptimal following MET treatment. However, robust RCTs are required to fully investigate the effectiveness of DPP-4 inhibitors as an add-on to various combination therapies in patients unresponsive to various oral GLTs.

The glycaemic effectiveness and body weight responses of co-administering DPP-4 inhibitor to patients with inadequate glycaemic control following treatment with MET, SU or combination of both MET plus SU was investigated (Objective 3). In addition, the HbA1c changes according to categories of HbA1c levels at baseline were examined. This was achieved through the set-up of a parallel-group study involving the underlying treatment

groups and estimated a multinomial propensity score on the baseline covariates in order to obtain the average treatment effect (ATE) of DPP-4 inhibitor in the population. The intensification treatment with DPP-4 inhibitor following MET monotherapy, SU monotherapy or both, was observed to result in an average 0.5% reduction in HbA1c and a 0.8 kg weight loss within a year. However, the reductions across the treatment groups were generally similar. Findings from this study also support the addition of DPP-4 inhibitor as a good therapeutic option for achieving effective responses in patients with T2DM. However, adding DPP-4 inhibitor to an ongoing MET + SU regimen appeared to be less efficacious among patients whose HbA1c was above 9% at the time of administration. It is therefore recommended that treatment should be characterized on an individual basis and robust RCTs should be carried out to investigate the influence of obesity and longer treatment durations on the efficacy of co-administering DPP-4 inhibitors to patients unresponsive to other oral GLTs. This finding may lend some basis to utilising 'personalised therapy', to offer the right treatment to the right patient at the right time.

In Chapter 6, a 5-year follow-up data was assessed to determine the glycaemic durability of DPP-4 inhibitor compared with SU and TZD when used as 2nd line treatment options following MET monotherapy failure (Objective 4). The risk of treatment failure in patients who had their MET therapy intensified with these 2nd line glucose-lowering agents was assessed through the application of multinomial propensity scores, survival analysis and Cox proportional hazards regression models. The addition of a TZD was observed to be associated with the most durable glycaemic response, followed by a SU and then a DPP-4

inhibitor. The findings on TZD and SU mirrors that from the ADOPT study, however the ADOPT study was published before the availability of DPP-4 inhibitors. It was interesting to observe, compared with a SU, adding a DPP-4 inhibitor to MET was associated with an increased need of an earlier requirement for treatment intensification with a 3rd agent. In addition, some factors that may be associated with earlier dual therapy failure were identified. These included concomitant use of statin therapy, being a female, a smoker, having longer duration of diabetes and higher baseline HbA1c.

The observed associated effect of statins in reducing durability of DPP-4 inhibitor intensification is in light of current research interest in the role of statin on glucose metabolism, diabetes risk, and glycaemic control among patients with pre-existing diabetes.[192] Further research is required to see whether statin is associated with progression of diabetes.

Furthermore, it was observed that adding a SU to MET as the 2nd line treatment resulted in similar proportion of patients attaining an HbA1c goal of 7.5% compared with DPP-4 inhibitor and TZD. Conducting RCTs at this level of treatment and follow-up period is not without its numerous challenges. Therefore, in the absence of real-world evidence of comparative effectiveness studies on the durability of 2nd line glucose-lowering agents following MET monotherapy failure to maintain adequate glycaemic control, this thesis highlights the differences in HbA1c end points between three oral treatment options that are frequently used as dual therapy in routine clinical practice.

In Chapter 7 investigation was made on some concerns large observational studies have raised regarding cardiovascular safety and mortality risks of treatment intensification with glucose-lowering agents (Objective 5). International regulatory agencies require all new antidiabetic agents to possess glucose-lowering ability and not be associated with increased cardiovascular events risks.[59,60] This has raised questions about the most appropriate choice of therapy for treatment intensification. Insulin is still one of the most established and effective glucose-lowering therapies available and its use in people with T2DM has grown markedly over recent years. More recently however, the effectiveness and safety of insulin therapy has been a subject of intense discussion.[197,203]

Chapter 7 highlighted that attention should not only be drawn to the risks that may be associated with monotherapy and dual therapy alone, but also intensification with a 3rd line treatment option. The clinical practice data was examined and comparisons were made on the time to a composite of non-fatal acute myocardial infarction (AMI), non-fatal stroke or all-cause mortality in patients who had their diabetes treatment intensified with a DPP-4 inhibitor versus insulin following dual therapy failure with MET plus SU. This was conducted through the application of propensity score matching analysis to minimise sampling bias. It was observed that among patients with diabetes who are receiving metformin and sulfonylurea therapy, the addition of insulin compared with DPP-4 inhibitor was associated with an increased risk of a composite of non-fatal cardiovascular outcomes and all-cause mortality. The observed excess risk of adverse cardiovascular events was increased in patients

who are obese (BMI > 30kg/m^2). In the absence of a consensus on which 3rd line treatment is most appropriate when maximum tolerated doses of MET+SU fail to maintain adequate glycaemic control, some important differences in cardiovascular and mortality outcomes between two treatment options that are frequently used in patients with dual therapy failure were identified in routine clinical practice data. These findings require further investigation to clarify the risk associated with insulin, especially among obese patients with T2DM, in view of the availability of other therapies available for treating hyperglycaemia in these patients.

In Chapter 8 a review of the evidence that DPP-4 inhibitor use may be protective to bone, thereby reducing fracture incidence was conducted. This was on the basis of previous smaller meta-analysis on incretin therapy and fracture risk. This thesis has added to the literature a larger, more up-to-date meta-analysis on this safety concern. While DPP-4 inhibitors have not been shown to significantly protect against fracture, this research is still valuable in informing the choices of healthcare providers in prescribing treatments from this class of drugs. For the users of this treatment, it is good news that DPP-4 inhibitors are not generally associated with fracture incidence, in contrast to TZD, which are known to be associated with increased fracture risk. However, the results drawn from RCTs in the field are varied, meaning that a definitive conclusion cannot be made on the role of DPP-4 inhibitors in protecting against bone fractures. The results of this review imply that future research should include studies of a longer duration. In addition, more studies are

required to directly investigate the number of fractures with the use of DPP-4 inhibitors, as a primary endpoint, rather than an adverse event.

9.2 Application to health outcomes

Data explored in this thesis has the potential to transform population health. The quality and outcomes framework has been a key step for progress in providing benefits to health outcomes of patients with T2DM and other diseases. This has led to the emergence of registers of a wide range of chronic diseases, and estimates of disease frequencies, intermediate outcome measures and data on some aspects of lifestyle (such as general practice recorded obesity and lifestyle intervention options). There is now a better understanding of the variation in quality of care provided in general practices and practice profile tools have helped identify some of these variations. For example, there is considerable variation between practices in the prescribing of glucose-lowering therapies in T2DM. Data from the current research can provide additional information on lifestyle factors such as smoking, levels of obesity, referrals for exercise programmes, etc, as well as clinical and metabolic parameters associated with treatment response and non-response. The availability of these records can assist in the evaluation of interventions offered to reduce the burden of diabetes and diabetes-related complications. The possibility of combining these information with health outcomes data could provide an endless potential. Certainly, data on its own does not save lives, however, appropriate analysis of good quality data can support actions that will.

9.3 Areas for future research

An important area for future research would be to explore the methods presented within this thesis to assess and analyse primary care data to increase the evidence of effectiveness and safety of GLTs. A number of newer therapeutic agents that are either used as single or combination therapies, for example, GLP-1 receptor agonists and SGLT-2 inhibitors have not been assessed for their long-term effectiveness and safety. These glucose-lowering medications need to be assessed for their long-term efficacy in glucose control and consequences in future research.

The risk of developing T2DM has been linked to hereditary factors. The associated risk is estimated to be 40% if a single parent is affected with the disease. Studies of patients with maturity-onset diabetes of the young (MODY) or inherited diabetes have identified the evidence of single gene defects in these individuals, and have shown that it takes the individual's environment to express the genes.[249,250] Further research that could incorporate biobank data and other secondary care data may provide meaningful insights into other factors that may predict better health outcome, and to advance 'personalised medicine'.

Other diabetes-related outcomes

Other outcomes include hypoglycaemia, quality of life and functional status. Conclusions could not be drawn regarding the comparative effects of DPP-4 inhibitors on the quality of life and functional status of patients with T2DM because of a lack of robust data for assessing quality of life.

Hypoglycaemia event was defined in the study as any recording of hypoglycaemia by the GP prior to treatment intensification with a DPP-4 inhibitor. The frequency of hypoglycaemic episodes at baseline was observed to vary greatly within the data. For example, hypoglycaemia appeared to be more frequent (about 20%) in subjects taking MET, SU or MET plus SU earlier on during diabetes therapy than in subjects taking combination oral glucose-lowering medications over a longer period of time (3%). The proportions of subjects who experienced any hypoglycaemic episodes ranged from 1 to 20% for subjects on MET prior to index treatment date, and 4 to 19% for MET plus SU users. The disparity is assumed to be attributable to poor reporting of hypoglycaemic episodes in general practices during the study period. Although it was not viable to examine the comparative risks of hypoglycaemia across treatment groups in this thesis, there is the evidence of ongoing work to increase the validity of hypoglycaemia reporting in routine clinical database. Khunti et al[251] assessed whether there is an association between hypoglycaemia, the risk of CV events, and all-cause mortality among insulin-treated patients with a diagnosis of diabetes using data from the CPRD database (a nationally representative population).[251]

9.4 Concluding remarks

Observational studies investigating the effectiveness and safety of glucose-lowering therapies from administrative or healthcare database studies are increasing and becoming more widely used. As newer glucose-lowering therapies are increasingly being prescribed in routine clinical practice, it is also important to ensure that the analytical methods employed in revealing

associated effects of treatments are not flawed with methodological challenges such as bias and confounding. A number of biases which influence retrospective cohort analyses and strategies for minimizing them have been discussed in this thesis. Other methodological challenges that influence observational studies have been described elsewhere.[7]

This research has investigated patient records from a large UK primary care database for the effectiveness and safety of DPP-4 inhibitor when added as intensification glucose-lowering agent following failure of other oral regimens to achieve glycaemic control. A 'real-world' situation where dual, triple or quadruple therapies are often used to achieve glucose target levels was explored by applying robust analytical techniques to control for bias. Given the lack of evidence for long-term effectiveness and safety of these treatment combinations, it is unclear how to optimally manage patients who fail to achieve glycaemic targets. Therefore, this research was undertaken to improve the understanding of the effectiveness of DPP-4 inhibitor regimens and their associated safety concerns in the management of T2DM. This thesis has demonstrated that the addition of DPP-4 inhibitor is a good therapeutic option for achieving glycaemic targets in patients with T2DM. However, the study also observed the effectiveness of adding DPP-4 inhibitor to an ongoing oral glucose-lowering regimen may be influenced by some clinical and metabolic parameters, which suggests that treatment should be characterised on an individual basis.

References

1. Turner, R.C., et al., *Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus: progressive requirement for multiple therapies (UKPDS 49). UK Prospective Diabetes Study (UKPDS) Group.* JAMA, 1999. **281**(21): p. 2005-12.
2. American Diabetes Association, *Standards of Medical Care in Diabetes.* Diabetes Care, 2012. **35**(Supplement 1): p. S11-S63.
3. Dods, R.F. *Understanding diabetes a biochemical perspective*. 2013; Available from: http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=542878.
4. National Institute for Health and Clinical Excellence. *Type 2 diabetes: newer agents*. 2009 [cited 2013 19th July]; Available from: http://www.nice.org.uk/guidance/CG87.
5. Inzucchi, S.E., et al., *Management of Hyperglycemia in Type 2 Diabetes, 2015: A Patient-Centered Approach: Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes.* Diabetes Care, 2015. **38**(1): p. 140-149.
6. Bennett, W.L., et al., *Oral diabetes medications for adults with type 2 diabetes: an update*. 2011, Agency for Healthcare Research and Quality: Rockville, MD.
7. Patorno, E., et al., *Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations.* Diabetologia, 2014. **57**(11): p. 2237-50.
8. Inzucchi, S.E., et al., *Management of Hyperglycemia in Type 2 Diabetes: A Patient-Centered Approach: Position Statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD).* Diabetes Care, 2012. **35**(6): p. 1364-1379.
9. Health and Social Care Information Centre, *Quality and Outcomes Framework - Prevalence, Achievements and Exceptions Report 2013-14.* 2014.

10. Health and Social Care Information Centre, *National Diabetes Audit 2012–2013*, in *Report 2: Complications and Mortality*. 2015.
11. Meeking, D., *Understanding diabetes & endocrinology : a problem-orientated approach*. 2011, London: Manson Pub.
12. Diabetes UK. *The Cost of Diabetes Report*. 2014 23 July 2015]; Available from: https://www.diabetes.org.uk/Documents/Diabetes%20UK%20Cost%20of%20Diabetes%20Report.pdf.
13. International Diabetes Federation. *Diabetes: facts and figures*. 2013 5th February, 2014]; Available from: http://www.idf.org/worlddiabetesday/toolkit/gp/facts-figures.
14. World Health Organization *Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus*. 2011.
15. Weykamp, C., et al., *The IFCC Reference Measurement System for HbA1c: a 6-year progress report.* Clin Chem, 2008. **54**(2): p. 240-8.
16. Little, R.R., et al., *The national glycohemoglobin standardization program: a five-year progress report.* Clin Chem, 2001. **47**(11): p. 1985-92.
17. National Institute for Health and Clinical Excellence *Preventing type 2 diabetes - risk identification and interventions for individuals at high risk: guidance*. 2012.
18. American Diabetes Association, *Standards of Medical Care in Diabetes.* Diabetes Care, 2010. **33**(Supplement 1): p. S4-S10.
19. American Diabetes Association, *Standardization of the Oral Glucose Tolerance Test: Report of the Committee on Statistics of the American Diabetes Association, June 14, 1968.* Diabetes, 1969. **18**(5): p. 299-310.
20. Genuth, S., et al., *Follow-up report on the diagnosis of diabetes mellitus.* Diabetes Care, 2003. **26**(11): p. 3160-7.
21. Sayetta, R.B. and R.S. Murphy, *Summary of Current Diabetes-related Data from the National Center for Health Statistics.* Diabetes Care, 1979. **2**(2): p. 105-119.
22. Committee, T.I.E., *International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-1334.
23. Hoelzel, W., et al., *IFCC Reference System for Measurement of Hemoglobin A1c in Human Blood and*

the National Standardization Schemes in the United States, Japan, and Sweden: A Method-Comparison Study. Clinical Chemistry, 2004. **50**(1): p. 166-174.
24. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, *Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus.* Diabetes Care, 1997. **20**(7): p. 1183-97.
25. GP Update Limited, *Type 2 diabetes*, in *The GP Update Handbook*. 2013.
26. Nathan, D.M., et al., *Translating the A1C Assay Into Estimated Average Glucose Values.* Diabetes Care, 2008.
27. Gerstein, H.C., et al., *Effects of intensive glucose lowering in type 2 diabetes.* N Engl J Med, 2008. **358**(24): p. 2545-59.
28. Patel, A., et al., *Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes.* N Engl J Med, 2008. **358**(24): p. 2560-72.
29. Franz, M.J., et al., *Evidence-based nutrition principles and recommendations for the treatment and prevention of diabetes and related complications.* Diabetes Care, 2003. **26 Suppl 1**: p. S51-61.
30. Knowler, W.C., et al., *Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin.* N Engl J Med, 2002. **346**(6): p. 393-403.
31. Riddle, M.C., J. Rosenstock, and J. Gerich, *The Treat-to-Target Trial: Randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients.* Diabetes Care, 2003. **26**(11): p. 3080-3086.
32. Fritsche, A., M.A. Schweitzer, and H.U. Haring, *Glimepiride combined with morning insulin glargine, bedtime neutral protamine hagedorn insulin, or bedtime insulin glargine in patients with type 2 diabetes. A randomized, controlled trial.* Ann Intern Med, 2003. **138**(12): p. 952-9.
33. Johnson, J.L., S.L. Wolf, and U.M. Kabadi, *Efficacy of insulin and sulfonylurea combination therapy in type II diabetes. A meta-analysis of the randomized placebo-controlled trials.* Arch Intern Med, 1996. **156**(3): p. 259-64.
34. Yki-Järvinen, H., *Combination Therapies With Insulin in Type 2 Diabetes.* Diabetes Care, 2001. **24**(4): p. 758-767.

35. Nathan, D.M., et al., *Management of hyperglycaemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy. A consensus statement from the American Diabetes Association and the European Association for the Study of Diabetes.* Diabetologia, 2006. **49**(8): p. 1711-21.
36. Baetta, R. and A. Corsini, *Pharmacology of dipeptidyl peptidase-4 inhibitors: similarities and differences.* Drugs, 2011. **71**(11): p. 1441-67.
37. Russell-Jones, D. and S. Gough, *Recent advances in incretin-based therapies.* Clin Endocrinol (Oxf), 2012. **77**(4): p. 489-99.
38. Baggio, L.L. and D.J. Drucker, *Biology of Incretins: GLP-1 and GIP.* Gastroenterology, 2007. **132**(6): p. 2131-2157.
39. Jarvis, C.I., A. Cabrera, and D. Charron, *Alogliptin: a new dipeptidyl peptidase-4 inhibitor for type 2 diabetes mellitus.* Ann Pharmacother, 2013. **47**(11): p. 1532-9.
40. Ndefo, U.A., O. Okoli, and G. Erowele, *Alogliptin: A new dipeptidyl peptidase-4 inhibitor for the management of type 2 diabetes mellitus.* Am J Health Syst Pharm, 2014. **71**(2): p. 103-9.
41. Egan , A.G., et al., *Pancreatic Safety of Incretin-Based Drugs — FDA and EMA Assessment.* New England Journal of Medicine, 2014. **370**(9): p. 794-797.
42. Scirica, B.M., et al., *Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus.* New England Journal of Medicine, 2013. **369**(14): p. 1317-1326.
43. White, W.B., et al., *Alogliptin after Acute Coronary Syndrome in Patients with Type 2 Diabetes.* New England Journal of Medicine, 2013. **369**(14): p. 1327-1335.
44. Colhoun, H.M., et al., *Hospitalised hip fracture risk with rosiglitazone and pioglitazone use compared with other glucose-lowering drugs.* Diabetologia, 2012. **55**(11): p. 2929-37.
45. Seino, Y., M. Fukushima, and D. Yabe, *GIP and GLP-1, the two incretin hormones: Similarities and differences.* Journal of Diabetes Investigation, 2010. **1**(1-2): p. 8-23.
46. Monami, M., et al., *Dipeptidyl peptidase-4 inhibitors and bone fractures: a meta-analysis of randomized clinical trials.* Diabetes Care, 2011. **34**(11): p. 2474-6.
47. Waugh, N., et al., *Newer agents for blood glucose control in type 2 diabetes: systematic review and*

economic evaluation. Health Technol Assess, 2010. **14**(36): p. 1-248.
48. Boussageon, R., et al., *Effect of intensive glucose lowering treatment on all cause mortality, cardiovascular death, and microvascular events in type 2 diabetes: meta-analysis of randomised controlled trials*. Vol. 343. 2011.
49. Inzucchi, S.E., *Oral antihyperglycemic therapy for type 2 diabetes: scientific review.* JAMA, 2002. **287**(3): p. 360-72.
50. Kahn, S.E., et al., *Glycemic Durability of Rosiglitazone, Metformin, or Glyburide Monotherapy.* New England Journal of Medicine, 2006. **355**(23): p. 2427-2443.
51. Bennett, W.L., et al., *Comparative effectiveness and safety of medications for type 2 diabetes: an update including new drugs and 2-drug combinations.* Ann Intern Med, 2011. **154**(9): p. 602-13.
52. Bolen, S., et al., *Comparative Effectiveness and Safety of Oral Diabetes Medications for Adults With Type 2 Diabetes: Comparative Effectiveness Review No. 8*. 2007, Agency for Healthcare Research and Quality: Rockville, MD.
53. Azim, S., W.L. Baker, and W.B. White, *Evaluating cardiovascular safety of novel therapeutic agents for the treatment of type 2 diabetes mellitus.* Curr Cardiol Rep, 2014. **16**(11): p. 541.
54. Giorgino, F., A. Leonardini, and L. Laviola, *Cardiovascular disease and glycemic control in type 2 diabetes: now that the dust is settling from large clinical trials.* Ann N Y Acad Sci, 2013. **1281**: p. 36-50.
55. Ray, K.K., et al., *Effect of intensive control of glucose on cardiovascular outcomes and death in patients with diabetes mellitus: a meta-analysis of randomised controlled trials.* Lancet, 2009. **373**(9677): p. 1765-72.
56. Zinman, B., et al., *Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes.* New England Journal of Medicine, 2015. **373**(22): p. 2117-2128.
57. Nissen, S.E. and K. Wolski, *Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes.* New England Journal of Medicine, 2007. **356**(24): p. 2457-2471.
58. Holman, R.R., H. Sourij, and R.M. Califf, *Cardiovascular outcome trials of glucose-lowering drugs or strategies in*

type 2 diabetes. The Lancet, 2014. **383**(9933): p. 2008-2017.
59. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research. *Guidance for industry: diabetes mellitus — evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes*. 2008; Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf.
60. European Medicine Agency Committee for Medicinal Products for Human Use. *Guideline on clinical investigation of medicinal products in the treatment of diabetes mellitus*. 2010 [cited 2015 27th November]; Available from: http://www.ema.europa.eu/docs/en_GB/documentlibrary/Scientific_guideline/2010/02/WC500073570.pdf.
61. Green, J.B., et al., *Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes.* New England Journal of Medicine, 2015. **373**(3): p. 232-242.
62. Engel, S.S., et al., *Cardiovascular safety of sitagliptin in patients with type 2 diabetes mellitus: a pooled analysis.* Cardiovasc Diabetol, 2013. **12**: p. 3.
63. Scheller, N.M., et al., *All-cause mortality and cardiovascular effects associated with the DPP-IV inhibitor sitagliptin compared with metformin, a retrospective cohort study on the Danish population.* Diabetes, Obesity and Metabolism, 2014. **16**(3): p. 231-236.
64. Montagnani, A. and S. Gonnelli, *Antidiabetic therapy effects on bone metabolism and fracture risk.* Diabetes Obes Metab, 2013. **15**(9): p. 784-91.
65. de Paula, F.J.A., M.C. Horowitz, and C.J. Rosen, *Novel Insights into the Relationship between Diabetes and Osteoporosis.* Diabetes/Metabolism Research and Reviews, 2010. **26**(8): p. 622-630.
66. Sanz, C., et al., *Signaling and biological effects of glucagon-like peptide 1 on the differentiation of mesenchymal stem cells from human bone marrow.* American Journal of Physiology - Endocrinology and Metabolism, 2010. **298**(3): p. E634-E643.
67. Yamada, C., et al., *The Murine Glucagon-Like Peptide-1 Receptor Is Essential for Control of Bone Resorption.* Endocrinology, 2008. **149**(2): p. 574-579.
68. Home, P.D., et al., *Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination*

therapy for type 2 diabetes (RECORD): a multicentre, randomised, open-label trial. Lancet, 2009. **373**(9681): p. 2125-35.
69. Schwartz, A.V., et al., *Older women with diabetes have a higher risk of falls: a prospective study.* Diabetes Care, 2002. **25**(10): p. 1749-54.
70. Lapane, K., et al., *Sulfonylureas and Risk of Falls and Fractures: A Systematic Review.* Drugs & Aging, 2013. **30**(7): p. 527-547.
71. Ray, W.A., *Improving Automated Database Studies.* Epidemiology, 2011. **22**(3): p. 302-304 10.1097/EDE.0b013e31820f31e1.
72. Yang, X.L., et al., *Addressing different biases in analysing drug use on cancer risk in diabetes in non-clinical trial settings--what, why and how?* Diabetes Obes Metab, 2012. **14**(7): p. 579-85.
73. Levesque, L.E., et al., *Problem of immortal time bias in cohort studies: Example using statins for preventing progression of diabetes.* BMJ (Online), 2010. **340**(7752): p. 907-911.
74. Suissa, S. and L. Azoulay, *Metformin and the risk of cancer: Time-related biases in observational studies.* Diabetes Care, 2012. **35**(12): p. 2665-2673.
75. Elwood, J.M., *Critical appraisal of epidemiological studies and clinical trials*. 2007, Oxford; New York: Oxford University Press.
76. Segal, J.B., et al., *Using Propensity Scores Subclassification to Estimate Effects of Longitudinal Treatments: An Example Using a New Diabetes Medication.* Medical Care Comparative Effectiveness and Safety: Emerging Methods, 2007. **45**(10): p. S149-S157.
77. Suh, H.S., et al., *Comparative effectiveness of statin plus fibrate combination therapy and statin monotherapy in patients with type 2 diabetes: Use of propensity-score and instrumental variable methods to adjust for treatment-selection bias.* Pharmacoepidemiol Drug Saf, 2012. **21**(5): p. 470-484.
78. Sato, D., et al., *Effects of a sitagliptin safety alert on prescription behaviour for oral antihyperglycaemic drugs: A propensity score-matched cohort study of prescription receipt data in japan.* Drug Safety, 2013. **36**(8): p. 605-615.
79. Winterstein, A.G., et al., *Misclassification in assessment of diabetogenic risk using electronic health records.*

Pharmacoepidemiology and Drug Safety, 2014. **23**(8): p. 875-881.
80. Blonde, L., et al., *Greater reductions in A1C in type 2 diabetic patients new to therapy with glyburide/metformin tablets as compared to glyburide co-administered with metformin.* Diabetes, Obesity and Metabolism, 2003. **5**(6): p. 424-431.
81. Idris, I., G. Warren, and R. Donnelly, *Association between thiazolidinedione treatment and risk of macular edema among patients with type 2 diabetes.* Arch Intern Med, 2012. **172**(13): p. 1005-1011.
82. Johnson, J.A., et al., *Reduced cardiovascular morbidity and mortality associated with metformin use in subjects with Type 2 diabetes.* Diabetic Medicine, 2005. **22**(4): p. 497-502.
83. Loebstein, R., et al., *Database evaluation of the effects of long-term rosiglitazone treatment on cardiovascular outcomes in patients with type 2 diabetes.* J Clin Pharmacol, 2011. **51**(2): p. 173-80.
84. Lund, J.L.M., et al., *Thiazolidinedione use and ulcerative colitis-related flares: An exploratory analysis of administrative data.* Inflammatory Bowel Diseases, 2011. **17**(3): p. 787-794.
85. Masica, A.L., et al., *Comparative effectiveness research using electronic health records: Impacts of oral antidiabetic drugs on the development of chronic kidney disease.* Pharmacoepidemiology and Drug Safety, 2013. **22**(4): p. 413-422.
86. Miao, R., et al., *Real world outcomes of adding rapid-acting insulin versus switching to analog premix insulin among US patients with type 2 diabetes treated with insulin glargine.* Patient Preference and Adherence, 2013. **7**: p. 951-960.
87. Roumie, C.L., et al., *The effect of incident antidiabetic regimens on lipid profiles in veterans with type 2 diabetes: a retrospective cohort.* Pharmacoepidemiology & Drug Safety, 2011. **20**(1): p. 36-44.
88. Wang, L., et al., *Real-world outcomes of US employees with type 2 diabetes mellitus treated with insulin glargine or neutral protamine Hagedorn insulin: A comparative retrospective database study.* BMJ Open, 2013. **3**(4).
89. Wei, W., et al., *Much ado about nothing? A real-world study of patients with type 2 diabetes switching basal*

insulin analogs. Advances in Therapy, 2014. **31**(5): p. 539-560.
90. Yeh, H.T., et al., *Effects of thiazolidinediones on cardiovascular events in patients with type 2 diabetes mellitus after drug-eluting stent implantation: a retrospective cohort study using the national health insurance database in Taiwan.* Clin Ther, 2012. **34**(4): p. 885-93.
91. Winkelmayer, W.C., et al., *Comparison of cardiovascular outcomes in elderly patients with diabetes who initiated rosiglitazone vs pioglitazone therapy.* Arch Intern Med, 2008. **168**(21): p. 2368-75.
92. Johannes, C.B., et al., *The risk of coronary heart disease in type 2 diabetic patients exposed to thiazolidinediones compared to metformin and sulfonylurea therapy.* Pharmacoepidemiology & Drug Safety, 2007. **16**(5): p. 504-12.
93. Gerrits, C.M., et al., *A comparison of pioglitazone and rosiglitazone for hospitalization for acute myocardial infarction in type 2 diabetes.* Pharmacoepidemiol Drug Saf, 2007. **16**(10): p. 1065-71.
94. Hwang, A.L., et al., *Metformin and survival in pancreatic cancer: A retrospective cohort study.* Pancreas, 2013. **42**(7): p. 1054-1059.
95. Khalangot, M., et al., *Glibenclamide-related excess in total and cardiovascular mortality risks: data from large Ukrainian observational cohort study.* Diabetes Res Clin Pract, 2009. **86**(3): p. 247-53.
96. Wu, W., et al., *The association of diabetes and anti-diabetic medications with clinical outcomes in multiple myeloma.* British Journal of Cancer, 2014. **111**(3): p. 628-36.
97. Tannen, R., et al., *A new "Comparative Effectiveness" assessment strategy using the THIN database: comparison of the cardiac complications of pioglitazone and rosiglitazone.* Pharmacoepidemiology & Drug Safety, 2013. **22**(1): p. 86-97.
98. Brownstein, J.S., et al., *Rapid identification of myocardial infarction risk associated with diabetes medications using electronic medical records.* Diabetes Care, 2010. **33**(3): p. 526-531.
99. Mamtani, R., et al., *Association Between Longer Therapy With Thiazolidinediones and Risk of Bladder Cancer: A Cohort Study.* Journal of the National Cancer Institute, 2012. **104**(18): p. 1411-1421.

100. Tseng, C.H., *Diabetes, metformin use, and colon cancer: a population-based cohort study in Taiwan.* European Journal of Endocrinology, 2012. **167**(3): p. 409-16.
101. Wenten, M., et al., *Relative risk of acute pancreatitis in initiators of exenatide twice daily compared with other anti-diabetic medication: a follow-up study.* Diabetic Medicine, 2012. **29**(11): p. 1412-8.
102. Colhoun, H.M. and S.E. Group, *Use of insulin glargine and cancer incidence in Scotland: a study from the Scottish Diabetes Research Network Epidemiology Group.* Diabetologia, 2009. **52**(9): p. 1755-65.
103. Graham, D.J., et al., *Risk of acute myocardial infarction, stroke, heart failure, and death in elderly Medicare patients treated with rosiglitazone or pioglitazone.* JAMA, 2010. **304**(4): p. 411-8.
104. Andersson, C., et al., *Comparable long-term mortality risk associated with individual sulfonylureas in diabetes patients with heart failure.* Diabetes Res Clin Pract, 2011. **94**(1): p. 119-25.
105. Austin, P.C., *An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.* Multivariate Behavioral Research, 2011. **46**(3): p. 399-424.
106. Rahme, E., et al., *Hospitalization for gastrointestinal bleeding associated with non-steroidal anti-inflammatory drugs among elderly patients using low-dose aspirin: a retrospective cohort study.* Rheumatology, 2007. **46**(2): p. 265-272.
107. Roulson, J., E.W. Benbow, and P.S. Hasleton, *Discrepancies between clinical and autopsy diagnosis and the value of post mortem histology; a meta-analysis and review.* Histopathology, 2005. **47**(6): p. 551-9.
108. Bird, S., et al., *The association between drospirenone and hyperkalemia among women taking oral contraceptives.* Pharmacoepidemiol Drug Saf, 2011. **20**: p. S72-S73.
109. Etminan, M.P.M., et al., *Oral contraceptives and the risk of gallbladder disease: a comparative safety study.* CMAJ Canadian Medical Association Journal, 2011. **183**(8): p. 899-904.
110. Tyczynski, J.E., et al., *Safety Assessment of an Anti-Obesity Drug (Sibutramine): A Retrospective Cohort Study.* Drug Safety, 2012. **35**(8): p. 629-644.

111. Suissa, S., *Immortal Time Bias in Pharmacoepidemiology.* American Journal of Epidemiology, 2008. **167**(4): p. 492-499.
112. Schisterman, E.F., S.R. Cole, and R.W. Platt, *Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies.* Epidemiology (Cambridge, Mass.), 2009. **20**(4): p. 488-495.
113. Gerhard, T., *Bias: Considerations for research practice.* American Journal of Health-System Pharmacy, 2008. **65**(22): p. 2159-2168.
114. Ray, W.A., *Evaluating Medication Effects Outside of Clinical Trials: New-User Designs.* American Journal of Epidemiology, 2003. **158**(9): p. 915-920.
115. de Lusignan, S. and C. van Weel, *The use of routinely collected computer data for research in primary care: opportunities and challenges.* Fam Pract, 2006. **23**(2): p. 253-63.
116. CPRD. *Welcome to The Clinical Practice Research Datalink*. 7th September, 2015]; Available from: http://www.cprd.com/intro.asp.
117. IMS Health. *Epidemiology and Pharmacology Information Core*. 7th September, 2015]; Available from: http://www.epic-uk.org/about-us/about-us.shtml.
118. QRESEARCH.
119. NHS Connecting for Health, *Read Codes*.
120. Normand, S.L.T., et al., *Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores.* Journal of Clinical Epidemiology, 2001. **54**(4): p. 387-398.
121. Kirkwood, B.R. and J.A.C. Sterne, *Essential medical statistics*. 2007, Malden, Mass.: Blackwell Science.
122. Kirkwood, B.R.S.J.A.C.K.B.R., *Essential medical statistics*. 2003, Malden, Mass.: Blackwell Science.
123. Williams, R., *Using the margins command to estimate and interpret adjusted predictions and marginal effects.* Stata Journal, 2012. **12**(2): p. 308-331.
124. Rosenbaum, P.R. and D.B. Rubin, *The central role of the propensity score in observational studies for causal effects.* Biometrika, 1983. **70**(1): p. 41-55.
125. Imbens, G.W., *Nonparametric estimation of average treatment effects under exogeneity: A review.* The Review of Economics and Statistics, 2004. **86**: p. 4-29.
126. Hirano, K., G.W. Imbens, and G. Ridder, *Efficient Estimation of Average Treatment Effects Using the*

Estimated Propensity Score. Econometrica, 2003. **71**(4): p. 1161-1189.
127. Stuart, E.A. and K.M. Green, *Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes.* Dev Psychol, 2008. **44**(2): p. 395-406.
128. Joffe, M.M., et al., *Model selection, confounder control, and marginal structural models: Review and new applications.* American Statistician, 2004. **58**(4): p. 272-279.
129. Klein, J.P. and M.L. Moeschberger. *Survival analysis techniques for censored and truncated data*. 1997; Available from: http://site.ebrary.com/id/10005871.
130. Harrington, D., *Linear Rank Tests in Survival Analysis*, in *Encyclopedia of Biostatistics*. 2005, John Wiley & Sons, Ltd.
131. McCaffrey, D.F., et al., *A tutorial on propensity score estimation for multiple treatments using generalized boosted models.* Stat Med, 2013. **32**(19): p. 3388-414.
132. Lunceford, J.K. and M. Davidian, *Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study.* Stat Med, 2004. **23**(19): p. 2937-60.
133. Austin, P.C., *A Tutorial and Case Study in Propensity Score Analysis: An Application to Estimating the Effect of In-Hospital Smoking Cessation Counseling on Mortality.* Multivariate behavioral research, 2011. **46**(1): p. 119-151.
134. Robins, J.M., M.A. Hernan, and B. Brumback, *Marginal structural models and causal inference in epidemiology.* Epidemiology, 2000. **11**(5): p. 550-60.
135. Xie, J. and C.F. Liu, *Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data.* Statistics in Medicine, 2005. **24**(20): p. 3089-3110.
136. Austin, P.C., *Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples.* Statistics in Medicine, 2009. **28**(25): p. 3083-3107.
137. Marston, L., et al., *Issues in multiple imputation of missing data for large general practice clinical databases.* Pharmacoepidemiology and Drug Safety, 2010. **19**(6): p. 618-626.

138. Carpenter, J.R. and M.G. Kenward. *Multiple imputation and its application*. 2013; Available from: http://dx.doi.org/10.1002/9781119942283.
139. van Buuren, S., H.C. Boshuizen, and D.L. Knook, *Multiple imputation of missing blood pressure covariates in survival analysis.* Stat Med, 1999. **18**(6): p. 681-94.
140. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: Issues and guidance for practice.* Statistics in Medicine, 2011. **30**(4): p. 377-399.
141. Psaty, B.M., et al., *Assessment and control for confounding by indication in observational studies.* J Am Geriatr Soc, 1999. **47**(6): p. 749-54.
142. Schneeweiss, S., *Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics.* Pharmacoepidemiol Drug Saf, 2006. **15**(5): p. 291-303.
143. StataCorp, *Stata Statistical Software*. 2013, StataCorp LP: College Station, TX.
144. R Core Team, *R: A Language and Environment for Statistical Computing*. 2013, R Foundation for Statistical Computing: Vienna, Austria.
145. Higgins, J.P.T., *Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified.* International Journal of Epidemiology, 2008. **37**(5): p. 1158-1160.
146. Borenstein, M. and I. Wiley. *Introduction to meta-analysis*. 2009; Available from: http://site.ebrary.com/id/10297978.
147. Egger, M. and G.D. Smith, *Meta-analysis: Potentials and promise.* BMJ, 1997. **315**(7119): p. 1371-1374.
148. Egger, M., et al., *Bias in meta-analysis detected by a simple, graphical test.* BMJ, 1997. **315**(7109): p. 629-634.
149. Sterne, J.A.C. and M. Egger, *Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis.* Journal of Clinical Epidemiology, 2001. **54**(10): p. 1046-1055.
150. Bolen, S., et al., *Systematic Review: Comparative Effectiveness and Safety of Oral Medications for Type 2 Diabetes Mellitus.* Annals of Internal Medicine, 2007. **147**(6): p. 386-399.
151. Li, J., et al., *Effects of sitaglipin combined with metformin for type 2 diabetes mellitus: A systematic*

review. Chinese Journal of Evidence-Based Medicine, 2013. **13**(7): p. 836-843.
152. Chatterjee, S., *Review: DPP-4 inhibitors are less effective than metformin for reducing HbA1c in type 2 diabetes.* Annals of Internal Medicine, 2012. **157**(2): p. JC2-13.
153. Mulherin, A.J., et al., *Mechanisms underlying metformin-induced secretion of glucagon-like peptide-1 from the intestinal L cell.* Endocrinology, 2011. **152**(12): p. 4610-9.
154. Ahrén, B., et al., *Twelve- and 52-Week Efficacy of the Dipeptidyl Peptidase IV Inhibitor LAF237 in Metformin-Treated Patients With Type 2 Diabetes.* Diabetes Care, 2004. **27**(12): p. 2874-2880.
155. Herman, G.A., et al., *Pharmacokinetics and pharmacodynamics of sitagliptin, an inhibitor of dipeptidyl peptidase IV, in healthy subjects: results from two randomized, double-blind, placebo-controlled studies with single oral doses.* Clin Pharmacol Ther, 2005. **78**(6): p. 675-88.
156. Wong, M.C.S., et al., *Comparative effectiveness of dipeptidyl peptidase-4 (DPP-4) inhibitors and human glucagon-like peptide-1 (GLP-1) analogue as add-on therapies to sulphonylurea among diabetes patients in the asia-pacific region: A systematic review.* PLoS ONE, 2014. **9**(3).
157. Harashima, S.I., et al., *Sitagliptin add-on to low dosage sulphonylureas: efficacy and safety of combination therapy on glycaemic control and insulin secretion capacity in type 2 diabetes.* International Journal of Clinical Practice, 2012. **66**(5): p. 465-476.
158. Rubino, A., et al., *Delayed initiation of subcutaneous insulin therapy after failure of oral glucose-lowering agents in patients with Type 2 diabetes: a population-based analysis in the UK.* Diabet Med, 2007. **24**(12): p. 1412-8.
159. Henry, R.R., et al., *Efficacy and safety of initial combination treatment with sitagliptin and pioglitazone-a factorial study.* Diabetes, Obesity and Metabolism, 2014. **16**(3): p. 223-230.
160. Pospisilik, J.A., et al., *Dipeptidyl Peptidase IV Inhibitor Treatment Stimulates β-Cell Survival and Islet Neogenesis in Streptozotocin-Induced Diabetic Rats.* Diabetes, 2003. **52**(3): p. 741-750.

161. Sudhakaran, C., et al., *Effectiveness of sitagliptin in asian Indian patients with type 2 diabetes-an Indian tertiary diabetes care center experience.* Diabetes Technol Ther, 2011. **13**(1): p. 27-32.
162. Cheruvallath, Z.S., et al., *Design, synthesis and SAR of novel glucokinase activators.* Bioorg Med Chem Lett, 2013. **23**(7): p. 2166-71.
163. Riddle, M.C., *Glycemic management of type 2 diabetes: an emerging strategy with oral agents, insulins, and combinations.* Endocrinol Metab Clin North Am, 2005. **34**(1): p. 77-98.
164. Owen, V., I. Seetho, and I. Idris, *Predictors of responders to insulin therapy at 1 year among adults with type 2 diabetes.* Diabetes Obes Metab, 2010. **12**(10): p. 865-70.
165. Liebl, A., et al., *Comparison of insulin analogue regimens in people with type 2 diabetes mellitus in the PREFER Study: a randomized controlled trial.* Diabetes, Obesity & Metabolism, 2009. **11**(1): p. 45-52.
166. Ahren, B., et al., *Efficacy of vildagliptin versus sulfonylureas as add-on therapy to metformin: comparison of results from randomised controlled and observational studies.* Diabetologia, 2014. **57**(7): p. 1304-7.
167. Charbonnel, B., et al., *Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes inadequately controlled with metformin alone.* Diabetes Care, 2006. **29**(12): p. 2638-43.
168. Hermansen, K., et al., *Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, in patients with type 2 diabetes mellitus inadequately controlled on glimepiride alone or on glimepiride and metformin.* Diabetes Obes Metab, 2007. **9**(5): p. 733-45.
169. Scheen, A.J., *DPP-4 inhibitors in the management of type 2 diabetes: a critical review of head-to-head trials.* Diabetes Metab, 2012. **38**(2): p. 89-101.
170. Lewis, J.D., et al., *Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research.* Pharmacoepidemiol Drug Saf, 2007. **16**(4): p. 393-401.
171. Chen, B., et al., *Sitagliptin lowers glucagon and improves glucose tolerance in prediabetic obese SHROB rats.* Exp Biol Med (Maywood), 2011. **236**(3): p. 309-14.

172. Rosenbaum, P.R., *Model-Based Direct Adjustment.* Journal of the American Statistical Association, 1987. **82**(398): p. 387-394.
173. Busso, M., J. DiNardo, and J. McCrary, *New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators.* Review of Economics and Statistics, 2014.
174. Gao, W., et al., *Efficacy and safety of initial combination of DPP-IV inhibitors and metformin versus metformin monotherapy in type 2 diabetes: A systematic review of randomized controlled trials.* Diabetes, Obesity and Metabolism, 2014. **16**(2): p. 179-185.
175. Kim, M.-H., et al., *Metformin enhances glucagon-like peptide-1 via cooperation between insulin and Wnt signaling.* Journal of Endocrinology, 2013.
176. Chacra, A.R., et al., *Safety and efficacy of saxagliptin in combination with submaximal sulphonylurea versus up-titrated sulphonylurea over 76 weeks.* Diabetes and Vascular Disease Research, 2011. **8**(2): p. 150-159.
177. Matthews, D.R., et al., *Vildagliptin add-on to metformin produces similar efficacy and reduced hypoglycaemic risk compared with glimepiride, with no weight gain: results from a 2-year study.* Diabetes Obes Metab, 2010. **12**(9): p. 780-9.
178. Seck, T., et al., *Safety and efficacy of treatment with sitagliptin or glipizide in patients with type 2 diabetes inadequately controlled on metformin: a 2-year study.* Int J Clin Pract, 2010. **64**(5): p. 562-76.
179. Schernthaner, G., et al., *Canagliflozin compared with sitagliptin for patients with type 2 diabetes who do not have adequate glycemic control with metformin plus sulfonylurea: a 52-week randomized trial.* Diabetes Care, 2013. **36**(9): p. 2508-15.
180. Holman, R.R., et al., *10-year follow-up of intensive glucose control in type 2 diabetes.* N Engl J Med., 2008. **359**(15): p. 1577-89.
181. Berkowitz, S.A., et al., *Initial choice of oral glucose-lowering medication for diabetes mellitus: A patient-centered comparative effectiveness study.* JAMA Internal Medicine, 2014. **174**(12): p. 1955-1962.
182. Ekström, N., et al., *Durability of oral hypoglycemic agents in drug naïve patients with type 2 diabetes: report from the Swedish National Diabetes Register (NDR).* BMJ Open Diabetes Research & Care, 2015. **3**(1).

183. Spreeuwenberg, M.D., et al., *The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health.* Med Care, 2010. **48**(2): p. 166-74.
184. Cole, S.R. and M.A. Hernan, *Adjusted survival curves with inverse probability weights.* Comput Methods Programs Biomed, 2004. **75**(1): p. 45-9.
185. Austin, P.C., *The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments.* Statistics in Medicine, 2014. **33**(7): p. 1242-1258.
186. Rosenbaum, P.R. and D.B. Rubin, *Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome.* Journal of the Royal Statistical Society. Series B (Methodological), 1983. **45**(2): p. 212-218.
187. Charbonnel, B., et al., *Long-term efficacy and tolerability of add-on pioglitazone therapy to failing monotherapy compared with addition of gliclazide or metformin in patients with type 2 diabetes.* Diabetologia, 2005. **48**(6): p. 1093-1104.
188. Monami, M., et al., *Comparison of different drugs as add-on treatments to metformin in type 2 diabetes: a meta-analysis.* Diabetes Res Clin Pract, 2008. **79**(2): p. 196-203.
189. Inzucchi, S.E., et al., *Progression to insulin therapy among patients with type 2 diabetes treated with sitagliptin or sulphonylurea plus metformin dual therapy.* Diabetes, Obesity and Metabolism, 2015. **17**(10): p. 956-964.
190. Phung, O.J., et al., *Effect of noninsulin antidiabetic drugs added to metformin therapy on glycemic control, weight gain, and hypoglycemia in type 2 diabetes.* JAMA - Journal of the American Medical Association, 2010. **303**(14): p. 1410-1418.
191. Goke, B., et al., *Saxagliptin vs. glipizide as add-on therapy in patients with type 2 diabetes mellitus inadequately controlled on metformin alone: long-term (52-week) extension of a 52-week randomised controlled trial.* Int J Clin Pract, 2013. **67**(4): p. 307-16.
192. Anyanwagu, U., I. Idris, and R. Donnelly, *Drug-Induced Diabetes Mellitus: Evidence for Statins and Other Drugs Affecting Glucose Metabolism.* Clinical Pharmacology & Therapeutics, 2015: p. doi:10.1002/cpt.274.

193. Roumie, C.L., et al., *Association between intensification of metformin treatment with insulin vs sulfonylureas and cardiovascular events and all-cause mortality among patients with diabetes.* JAMA, 2014. **311**(22): p. 2288-96.
194. Gamble, J.M., et al., *Insulin use and increased risk of mortality in type 2 diabetes: a cohort study.* Diabetes, Obesity & Metabolism, 2010. **12**(1): p. 47-53.
195. Currie, C.J., et al., *Survival as a function of HbA1c in people with type 2 diabetes: a retrospective cohort study.* The Lancet, 2010. **375**(9713): p. 481-489.
196. Colayco, D.C., et al., *A1C and cardiovascular outcomes in type 2 diabetes: a nested case-control study.* Diabetes Care, 2011. **34**(1): p. 77-83.
197. Currie, C.J., et al., *Mortality and Other Important Diabetes-Related Outcomes With Insulin vs Other Antihyperglycemic Therapies in Type 2 Diabetes.* The Journal of Clinical Endocrinology and Metabolism, 2013. **98**(2): p. 668-677.
198. UK Prospective Diabetes Study (UKPDS) Group, *Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33).* Lancet, 1998. **352**(9131): p. 837-53.
199. Gerstein, H.C., et al., *Basal insulin and cardiovascular and other outcomes in dysglycemia.* N Engl J Med, 2012. **367**(4): p. 319-28.
200. Muniyappa, R., et al., *Cardiovascular Actions of Insulin.* Endocrine Reviews, 2007. **28**(5): p. 463-491.
201. Green, J.B., et al., *Rationale, design, and organization of a randomized, controlled Trial Evaluating Cardiovascular Outcomes with Sitagliptin (TECOS) in patients with type 2 diabetes and established cardiovascular disease.* Am Heart J, 2013. **166**(6): p. 983-989 e7.
202. Haviland, A., D.S. Nagin, and P.R. Rosenbaum, *Combining propensity score matching and group-based trajectory analysis in an observational study.* Psychol Methods, 2007. **12**(3): p. 247-67.
203. Rensing, K.L., et al., *Reducing cardiovascular disease risk in patients with type 2 diabetes and concomitant macrovascular disease: can insulin be too much of a good thing?* Diabetes Obes Metab, 2011. **13**(12): p. 1073-87.

204. van Avendonk, M.J. and G.E. Rutten, *Insulin therapy in type 2 diabetes: what is the evidence?* Diabetes Obes Metab, 2009. **11**(5): p. 415-32.
205. Ryysy, L., et al., *Hepatic fat content and insulin action on free fatty acids and glucose metabolism rather than insulin absorption are associated with insulin requirements during insulin therapy in type 2 diabetic patients.* Diabetes, 2000. **49**(5): p. 749-58.
206. Cho, E., et al., *A prospective study of obesity and risk of coronary heart disease among diabetic women.* Diabetes Care, 2002. **25**(7): p. 1142-8.
207. Vestergaard, P., *Bone metabolism in type 2 diabetes and role of thiazolidinediones.* Curr Opin Endocrinol Diabetes Obes, 2009. **16**(2): p. 125-31.
208. Tuominen, J.T., et al., *Bone mineral density in patients with type 1 and type 2 diabetes.* Diabetes Care, 1999. **22**(7): p. 1196-200.
209. Wallace, C., et al., *Incidence of falls, risk factors for falls, and fall-related fractures in individuals with diabetes and a prior foot ulcer.* Diabetes Care, 2002. **25**(11): p. 1983-6.
210. Carnevale, V., E. Romagnoli, and E. D'Erasmo, *Skeletal involvement in patients with diabetes mellitus.* Diabetes Metab Res Rev, 2004. **20**(3): p. 196-204.
211. Grey, A., *Thiazolidinedione-induced skeletal fragility – mechanisms and implications.* Diabetes, Obesity and Metabolism, 2009. **11**(4): p. 275-284.
212. Yaturu, S., B. Bryant, and S.K. Jain, *Thiazolidinedione Treatment Decreases Bone Mineral Density in Type 2 Diabetic Men.* Diabetes Care, 2007. **30**(6): p. 1574-1576.
213. Monami, M., et al., *Bone fractures and hypoglycemic treatment in type 2 diabetic patients: a case-control study.* Diabetes Care, 2008. **31**(2): p. 199-203.
214. Ivers, R.Q., et al., *Diabetes and risk of fracture: The Blue Mountains Eye Study.* Diabetes Care, 2001. **24**(7): p. 1198-203.
215. Stolk, R.P., et al., *Hyperinsulinemia and bone mineral density in an elderly population: The Rotterdam Study.* Bone, 1996. **18**(6): p. 545-9.
216. Vestergaard, P., *Discrepancies in bone mineral density and fracture risk in patients with type 1 and type 2 diabetes--a meta-analysis.* Osteoporos Int, 2007. **18**(4): p. 427-44.

217. Lipscombe, L.L., et al., *The risk of hip fractures in older individuals with diabetes: a population-based study.* Diabetes Care, 2007. **30**(4): p. 835-41.
218. Glorie, L., et al., *DPP IV inhibitor treatment attenuates bone loss and improves mechanical bone strength in male diabetic rats.* Am J Physiol Endocrinol Metab, 2014. **307**(5): p. E447-55.
219. Kyle, K.A., et al., *Differential Effects of PPAR-γ Activation versus Chemical or Genetic Reduction of DPP-4 Activity on Bone Quality in Mice.* Endocrinology, 2011. **152**(2): p. 457-467.
220. Garber, A.J., et al., *Effects of vildagliptin on glucose control in patients with type 2 diabetes inadequately controlled with a sulphonylurea*.* Diabetes, Obesity and Metabolism, 2008. **10**(11): p. 1047-1056.
221. NCT00396357, *Safety of Vildagliptin and Metformin Combination Therapy in Patients With Type 2 Diabetes Not Well Controlled With Metformin Alone.* 2012.
222. NCT00698932, *Evaluate Efficacy and Safety of Saxagliptin in Adult Patients With Type 2 Diabetes Inadequate Glycemic Control.* 2011.
223. Halpern, S. and M. Douglas, *Appendix: Jadad Scale for Reporting Randomized Controlled Trials*, in *Evidence-based Obstetric Anesthesia*. 2007, Blackwell Publishing Ltd. p. 237-238.
224. Monami, M., et al., *Safety of dipeptidyl peptidase-4 inhibitors: a meta-analysis of randomized clinical trials.* Curr Med Res Opin, 2011. **27 Suppl 3**: p. 57-64.
225. Rosenstock, J., et al., *Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin added to ongoing pioglitazone therapy in patients with type 2 diabetes: a 24-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group study.* Clin Ther, 2006. **28**(10): p. 1556-68.
226. NCT00856284 *Efficacy and Safety of Alogliptin Plus Metformin Compared to Glipizide Plus Metformin in Patients With Type 2 Diabetes Mellitus (ENDURE)*. 2013.
227. Arechavaleta, R., et al., *Efficacy and safety of treatment with sitagliptin or glimepiride in patients with type 2 diabetes inadequately controlled on metformin monotherapy: a randomized, double-blind, non-inferiority trial.* Diabetes, Obesity and Metabolism, 2011. **13**(2): p. 160-168.
228. Charbonnel, B., et al., *Efficacy and safety over 26 weeks of an oral treatment strategy including sitagliptin*

compared with an injectable treatment strategy with liraglutide in patients with type 2 diabetes mellitus inadequately controlled on metformin: a randomised clinical trial. Diabetologia, 2013. **56**(7): p. 1503-1511.
229. Wainstein, J., et al., *Initial therapy with the fixed-dose combination of sitagliptin and metformin results in greater improvement in glycaemic control compared with pioglitazone monotherapy in patients with type 2 diabetes.* Diabetes, Obesity and Metabolism, 2012. **14**(5): p. 409-418.
230. Iwamoto, Y., et al., *Efficacy and safety of sitagliptin monotherapy compared with voglibose in Japanese patients with type 2 diabetes: a randomized, double-blind trial.* Diabetes, Obesity and Metabolism, 2010. **12**(7): p. 613-622.
231. Alba, M., et al., *Sitagliptin and pioglitazone provide complementary effects on postprandial glucose and pancreatic islet cell function.* Diabetes, Obesity and Metabolism, 2013. **15**(12): p. 1101-1110.
232. NCT00374907. *A Study Assessing Saxagliptin Treatment in Subjects With Type 2 Diabetes Who Are Not Controlled With Diet and Exercise*. 2015 7th October, 2015]; Available from: https://clinicaltrials.gov/ct2/show/NCT00374907.
233. NCT00968708 *Cardiovascular Outcomes Study of Alogliptin in Patients With Type 2 Diabetes and Acute Coronary Syndrome (EXAMINE)*. 2014.
234. NCT01006603, *Saxagliptin Compared to Glimepiride in Elderly Type 2 Diabetes Patients, With Inadequate Glycemic Control on Metformin (GENERATION).* 2013.
235. NCT00984867, *Dapagliflozin DPPIV Inhibitor add-on Study.* 2014.
236. NCT00286468, *Study of Alogliptin Combined With Sulfonylurea in Subjects With Type 2 Diabetes Mellitus.* 2012.
237. DeFronzo, R.A., et al., *Efficacy and Tolerability of the DPP-4 Inhibitor Alogliptin Combined with Pioglitazone, in Metformin-Treated Patients with Type 2 Diabetes.* The Journal of Clinical Endocrinology & Metabolism, 2012. **97**(5): p. 1615-1622.
238. NCT00734474, *A Study of LY2189265 Compared to Sitagliptin in Participants With Type 2 Diabetes Mellitus on Metformin.* 2015.
239. Scherbaum, W.A., et al., *Efficacy and tolerability of vildagliptin in drug-naïve patients with type 2 diabetes*

and mild hyperglycaemia. Diabetes, Obesity and Metabolism, 2008. **10**(8): p. 675-682.
240. NCT00800683, *Safety and Efficacy in Type 2 Diabetic Patients With Severe Chronic Renal Impairment.* 2014.
241. NCT01098539, *A Study of the Efficacy and Safety of Albiglutide in Subjects With Type 2 Diabetes With Renal Impairment.* 2014.
242. Arjona Ferreira, J.C., et al., *Efficacy and Safety of Sitagliptin in Patients With Type 2 Diabetes and ESRD Receiving Dialysis: A 54-Week Randomized Trial.* American Journal of Kidney Diseases, 2013. **61**(4): p. 579-587.
243. Arjona Ferreira, J.C., et al., *Efficacy and Safety of Sitagliptin Versus Glipizide in Patients With Type 2 Diabetes and Moderate-to-Severe Chronic Renal Insufficiency.* Diabetes Care, 2013. **36**(5): p. 1067-1073.
244. NCT00707993, *Efficacy and Safety of Alogliptin Compared to Glipizide in Elderly Diabetics.* 2013.
245. Higgins, J.P.T. and S.G. Thompson, *Quantifying heterogeneity in a meta-analysis.* Statistics in Medicine, 2002. **21**(11): p. 1539-1558.
246. Higgins, J.P.T., et al., *Measuring inconsistency in meta-analyses.* BMJ, 2003. **327**(7414): p. 557-560.
247. Su, B., et al., *Risk of bone fractures associated with glucagon-like peptide-1 receptor agonists' treatment: a meta-analysis of randomized controlled trials.* Endocrine, 2014. **48**(1): p. 107-115.
248. Driessen, J.H.M., et al., *Use of dipeptidyl peptidase-4 inhibitors for type 2 diabetes mellitus and risk of fracture.* Bone, 2014. **68**: p. 124-130.
249. Hyttinen, V., et al., *Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study.* Diabetes, 2003. **52**(4): p. 1052-5.
250. Redondo, M.J., P.R. Fain, and G.S. Eisenbarth, *Genetics of type 1A diabetes.* Recent Prog Horm Res, 2001. **56**: p. 69-89.
251. Khunti, K., et al., *Hypoglycemia and Risk of Cardiovascular Disease and All-Cause Mortality in Insulin-Treated People With Type 1 and Type 2 Diabetes: A Cohort Study.* Diabetes Care, 2014.

Appendix A.

Systematic review 1 (Chapter 2)

Appendix A-1 Assessment of retrospective cohort studies

Author:			
Year:	*Short title:*		
DESCRIPTOR	**Yes/No**	**NA**	**Remark**
Description of evidence			
Exposure defined?			
Outcome defined?			
Study population			
Main results			
Internal validity - are results likely affected by:			
Bias?			
Confounders?			
Chance variation?			
Is there a correct time-relationship?			
Is the relationship strong?			
Any dose-response relationship?			
Are results consistent with study?			
Is there any specificity?			
External validity – Can results be applied to:			
Eligible population?			
Source population?			
Other relevant populations?			
Consistency with other evidence			
Consistent with other similar studies?			
Does total evidence suggest any specificity?			
Are they plausible in terms of a pharmacological mechanism?			
Where effect is major, is it coherent with distribution of exposure and outcome?			
Overall assessment of the study			
Relevance	YES	NO	Unsure
Any other comments:			

Source: (Elwood 2007)

Sample search strategy

Search 1: Evidence of bias in pharmacoepidemiological studies of retrospective cohort design

Database: EMBASE

1. Bias.mp.

2. systematic error/

3. drug effect/

4. drug research/

5. cohort analysis/

6. postmarketing surveillance/

7. pharmacoepidemiology/

8. database/

9. electronic medical record/

10. data base/ and general practice/

11. retrospective study/

12. 3 or 4 or 5 or 6

13. 8 or 9 or 10

14. 7 and 12

15. 13 and 14

16. pharmacoepidemiology/

17. 13 and 16

18. 1 or 2

19. 17 and 18

20. 11 and 17

21. 19 or 20

22. limit 21 to human

23. limit 22 to english language

****

Search 2: Evidence of treatment outcomes of glucose lowering therapy from retrospective cohort of administrative databases

Database: EMBASE

1. antidiabetic agent/

2. oral antidiabetic agent/

3. treatment outcome/

4. observational study/

5. cohort analysis/

6. retrospective study/

7. medical record/

8. data base/

9. electronic health record/

10. 1 or 2

11. 3 and 10

12. 4 or 5 or 6

13. 7 or 8 or 9

14. 11 and 12 and 13

15. limit 9 to human

16. limit 10 to english language

****

Appendix A-2 Characteristics of retrospective cohort studies

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
Idris et al., 2012	THIN database, UK	Risk of Diabetic Macular Oedema (DME) from short and long-term TZD use	103,368 patients with T2D and no DME at baseline, and between Jan., 2000 and Nov., 2009	Treatment with TZD was associated with increased risk of DME at 1-year; OR, 2.3 (95%CI, 1.5-3.6) and 10-year follow-up; HR, 2.3 (95%CI, 1.7-3.0)	Selection bias	Propensity score analysis
Suh et al., 2012	Claims database, US	Comparative effectiveness of statin plus fibrate combination therapy reducing risk of CV diseases	Patients with T2D; 318 patients on combination therapy vs. 9,928 on statin monotherapy identified between Jan., 2002 and Dec., 2003	Controlling for bias showed no difference in effect between the two groups; OR, 0.77 (95%CI, 0.58–1.03) Statin-fibrate group was associated with a reduction in CV events; OR, 0.53 (95%CI, 0.34–0.81)	Selection bias	Propensity score; Instrumental variable method
Blonde et al., 2003	Quest Diagnostic & Medco Health pharmacy databases, US	Comparative effectiveness of glyburide/MET combination therapy on change in HbA1c levels	Patients with T2D; 950 patients on glyburide/MET fixed-dose vs. 471 taking glyburide co-administered with MET incident users of OADs between	Lower fixed-doses of glyburide/MET tablets had a significantly greater reduction in A1C compared to glyburide co-administered with MET (2% vs. 1.5% decrease respectively; $p<0.001$)	Selection bias	Regression models used to adjust the change in A1C for important independent covariates

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
			Aug., 2000 and Jul., 2001; Mean age 58 years			
Gerrits et al., 2007	Health care insurer database, US	To compare the risk of hospitalization for AMI in patients treated with pioglitazone relative to rosiglitazone	Patients with T2D; 14,807 patients using pioglitazone vs. 15,104 in the Rosiglitazone group; Patients initiated treatment between 2003 and 2006; follow-up period:	Pioglitazone was associated with 22% relative risk reduction of hospitalization with AMI; aHR, 0.78 (95%CI, 0.63–0.96)	Misclassification bias	Stringent definition of exposure
Wang et al., 2013	MarketScan Commercial Claims and Encounters/Health & Productivity Management Databases, US	To compare real-world outcomes of initiating insulin glargine (GLA) vs. neutral protamine Hagedorn (NPH) insulin	534 patients with T2D previously treated with OADs and/or GLP-1 agonist and initiated GLA (n=356) or NPH (n=178) between 2003 and 2009. Follow-up was for 1 year.	Compared to NPH, insulin GLA users had better persistence or adherence (55% vs. 44%;p=0.002), lower hospitalisation rate (23% vs. 31%; p=0.036) and endocrinologist visit (19% vs. 27%;p=0.038), and no overall cost disadvantages	Selection bias	Propensity score matching and sensitivity analyses
Johnson et	Saskatche-	MET use and the	4,142 patients with	Compared to SU	Selection bias	Propensity

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
al., 2005	wan Health admin. databases, Canada	risk of CV related hospitalization and mortality	T2D, new users of either MET (n, 923), SU (n, 2138) or both (n,1081) between 1991 to 1999; Mean age 65 years	monotherapy, MET was associated with low risk of hospitalization and death; aHR, 0.81(0.68-0.97); MET+SU with lower risk of mortality; aHR, 0.61 (0.46-0.80), but similar rates of hospitalization		scores
Lund et al., 2011	Claims data from the PharMetrics Patient-Centric Database, US	Association between TZD use and ulcerative colitis (UC)-related flares	142 new TZD and 468 other prevalent OADs users with mean follow-up of 7.3 and 6.2 months respectively, between Jan., 2000 to Dec., 2005	TZD use was not associated with UC-related flares; HR, 1.05 (0.66-1.68)	Misclassification bias Selection bias, Prevalent user bias Healthy user bias	Propensity scores
Roumie et al., 2011	Veterans Admin. and Pharmacy Databases, US	Effect of incident OADs on time to initiate a lipid lowering medication (LLM)	6,917 patients not on LLM comprise: users of MET (3,435), SU (3,237) and MET + SU (629) between Jan. 2000 and Dec. 2007	Median (IQR) time to starting LLM following MET was 2.4 (1-4.6) years, but not statistically different for users of SU or combination OADs	Selection bias, Misclassification bias	Sensitivity analysis
Segal et al.,	Ingenix	To compare	From 131,714	No difference in outcome	Selection bias	Propensity

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
2007	LabRx Database, US	outcomes between users of exenatide versus established drugs (insulin and OADs) using propensity score	patients with diabetes, prescribed an antidiabetic drug between Jun & Dec., 2005; 3,225 patients were prescribed exenatide; Aged 18-64 years	relative to existing therapies; Relative odds of hospitalization; 1.02, 95% CI (0.33-1.98)		score
Johannes et al., 2007	Pharmacy claims database, US	The risk of coronary heart disease (CHD) with oral anti-diabetic drugs	25,140 diabetic patients age >18 years had at least one prescription for TZD or MET + SU combination between Jan., 1999 & Jun., 2002	No cardio-protective or deleterious effects between TZDs and MET + SU combination therapy; aHR, 1.02(0.87-1.20)	Misclassification bias	Sensitivity analysis Propensity score matching
Tannen et al., 2013	THIN, UK	To compare CV complications of pioglitazone (PIO) vs. rosiglitazone (ROS)	Patients with T2D; population with established ischemic CVD between Dec., 2005 and Jun., 2008; Aged 35-75 years	PIO replication studies did not affect MI; HR 0.88 (0.49 to 1.42); ROS increased MI in contrast to PIO; whereas in an unselected population, ROS and PIO have reasonably comparable	Start-time bias	

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
				effects		
Mamtani et al., 2012	THIN database, UK	Association between chronic TZD use and risk of bladder cancer	18,459 patients with T2D and used TZD, vs. 41,396 users of SU between Jul., 2000 and Aug., 2010; Incident cancers identified within the period	60 vs. 137 incident bladder cancers; Long-term (≥5yrs) TZD use vs. SU, increased risk of cancer; HR, 3.25 (1.08-9.71), but no difference found between TZD and SU cohorts; HR, 0.93 (0.68-1.29)	Surveillance bias	Censoring events dates, Sensitivity analysis, New user design
Brownstein et al., 2010	Partners Health Care Systems database, US	Association between OADs and myocardial infarction (MI)	Cohort of 34,253 patients treated with OADs; aged > 18 years; between Jan. 2000 and Dec., 2006 Mean age 58 years	Compared with SU, MET and pioglitazone, the RR for MI with rosiglitazone was 1.3 (1.1-1.6), 2.2 (1.6-3.1) and 2.2 (1.5-3.4), respectively	Prescription bias	Used other antidiabetic drugs comparators
Wenten et al., 2012	IMS LifeLink Program Health Plan Claims Database, US	Relative risk of acute pancreatitis in initiators of exenatide vs. other anti-diabetic	Of 482,034 initiators of antidiabetic drugs between Jun., 2005 and Mar., 2009, 24,237 initiated exenatide twice daily	No increased risk of hospitalized acute pancreatitis with exenatide; OR, 0.95 (0.65–1.38)	Indication bias	Propensity score

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
		medication				
Tseng, 2012	National Health Insurance database, Taiwan	Association between diabetes, MET use and colon cancer	265,151 non-diabetic vs. 92,670 diabetic patients; with no colon cancer; Age ≥ 40 years old and followed from Jan., 2003 to Dec., 2005	Risk of colon cancer was 24% higher in patients with diabetes (highest in <1 year of diabetes; aRR, 1.31(1.02–1.68)), but 27% lower among MET users (lowest with ≥3 years use; aRR, 0.643(0.490–0.845)	Detection bias	
Colhoun and Group, 2009	The Scottish Care Information-Diabetes Collaboration (SCI-DC) database	Risk of cancer from the use of insulin glargline	A cohort of 12852 new insulin users across the period Jan., 2002 and Dec., 2005 who were exposed to treatment for ≥ 4 months.	Overall, 447 patients used insulin glargine alone. There was no increase in breast cancer rates associated with insulin glargine use (HR:1.49, 95%CI 0.79–2.83); Insulin glargine-only users had a higher rate than those using non-glargine insulin (HR 3.39, 95% CI 1.46–7.85, p=0.004)	Reverse causation bias Allocation bias	Fixed cohort analysis (ITT) Analysis with exposure classification
Hwang et al., 2013	The Health Improvement	To determine the effect of MET	516 subjects with pre-existing	There was no difference in survival between those	Lead time bias Misclassification	Rigorous exposure

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
	Network (THIN) UK	exposure on survival in patients with advanced pancreatic adenocarcinoma (PAC).	diagnoses for T2DM and PAC. 247 were exposed to MET around the time of PAC diagnosis (between 6 months prior & 1 month after)	exposed and those unexposed to MET; HR, 1.11 (0.89, 1.38), p = 0.367). MET use is not associated with improved survival in subjects with advanced PAC		criteria; Utilised a window for exposure classification to allow for prescription and diagnosis entry
Masica et al., 2013	Baylor Health Care System (BHCS), US Christiana Care Health System (CCHS), Newark, DE	To evaluate the relationships between OAD use and incident chronic kidney disease (CKD)	newly diagnosed T2D cases requiring OADs between 1998 and 2009	798 and 977 patients developed proteinuria and had eGFR <60ml/min /1.73m^2. Compared to MET, SU trended toward association with an increased risk of developing proteinuria; aHR, 1.27 (0.93-1.74), while TZDs had similar association to MET.	Drug selection bias	IPW analysis
Miao et al., 2013	The national US IMPACT database	Compare outcome of intensification	746 patients, ≥18 years with T2DM who had GLA+RAI	There was no significant difference in HbA1c reduction from baseline,	Selection bias	PSM

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
		with rapid-acting insulin (RAI, basal-bolus) vs. premixed insulin (PMX) added to insulin glargine (GLA)	(n=373) or GLA + PMX between 2001 and 2009. Follow-up was for 1 year	the number of patients achieving target HbA1c of 7%, or Healthcare costs and utilization levels in each cohort. Incidence of hypoglycaemic events was also similar in both groups		
Sato et al., 2013	Pharmacies in Japan	To confirm reported changes in OAD prescription behaviour for patients receiving sitagliptin before and after a safety alert using PSM analysis	1,788 patients who had received SU + Sitagliptin vs. 30,963 patients on SU alone before the safety alert were identified between Dec 2009 to Dec 2010	There was a significant reduction in SU dose after the sitagliptin safety alert in patients administered sitagliptin compared to SU alone. PSM is a useful technique to control for selection bias	Selection bias	PSM
Wei et al., 2014	US claims databases (IMPACT and Humana)	To examine outcomes of switching basal insulin analogs among patients with T2DM	Cohort 1 previously on GLA and continued (n=2668) or switched to insulin detemir (n=536) vs Cohort 2	Detemir users in cohort 1 had lower treatment persistence/adherence, with 33–40% restarting GLA; higher rapid-acting insulin use, worse HbA1c	Selection bias	PSM

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
			who were previously on insulin detemir and either continued (n=780) or switched to GLA (n=419) One-year follow-up	outcomes, significantly higher diabetes drug costs, & similar hypoglycaemia rates, health care utilization and total costs. However, the opposite outcomes were observed in cohort 2		
Wells et al., 2013	Cleveland Clinic Database, US	To create a tool that accurately predicts the risk of morbidity and mortality in patients with T2D according to an oral hypoglycaemic agent	33,067 patients with T2D who were prescribed a single OAD between 1998 and 2006. Median follow-up for the mortality outcome was 769 days	The numbers of patients who experienced events were: CHD (n=3062), HF (n=1408), stroke (n=1451), and mortality (n=3661). The prediction tools demonstrated concordance indices (c-statistics) for the specific outcomes: CHD (0.73), HF (0.75), stroke (0.69), and mortality (0.72)	Treatment bias Missing information	Propensity scores MICE
Winterstein et al., 2014	Kaiser Permanente Northwest (KPNW)	To evaluate the validity of electronic health record data in	8183 patients developed DM; aged 35–64 years between 1997 and 2010,	Antipsychotics had the greatest diabetogenic risk; aHR, 1.73(1.44–2.08), the greatest propensity for a	Measurement bias, misclassification	Propensity score matching

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
	health-care system database, US	determining the associations between the use of anti-hypertensive, statins, atypical antipsychotics, and anti-depressants; and two endpoints	entered the cohort at the first negative BG test after ≥6 months without manifest DM. Diagnosis of diabetes was based on a laboratory test in 7583 (92.7%) cases	first negative test; aHR, 1.87(1.74–2.01), and the highest testing rate; aHR, 1.76 (1.72–1.81).		
Wu et al., 2014	University of Texas MD Anderson Cancer Center Tumor Registry database	To examined the association of diabetes, including steroid-induced diabetes (SID), and the impact of anti-diabetic medication on clinical outcomes of multiple myeloma (MM)	1240 consecutive patients with newly diagnosed MM treated at MDACC from 1 January 1996 to 31 December 2010 were identified	DM, SID in particular, is associated with poor clinical outcomes in MM. Insulin/analogues are associated with poor outcomes, whereas MET is associated with improved outcomes	Immortal time bias	Landmark analysis
Khalangot et	System of	To compare	type 2 diabetes	Total mortality was lower	Misclassification	Segregation

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
al., 2009	Diabetes Mellitus Supervision in Ukraine (SINADIAB) database	mortality risks between glibenclamide (GLB), gliclazide (GLC) and glimepiride (GLM)	(T2D) patients being treated with either GLB (n=50341), GLM (n=2479) or GLC (n=11368)	for GLC and GLM, vs. GLB cohort: HRs 0.33 (95%CI: 0.26–0.41), p<0.001 and 0.61 (95%CI: 0.41–0.89), p<0.01 respectively. CVD mortality risk reduction vs. GLB was significant only in GLC cohort: 0.29 (95%CI: 0.21–0.38), p<0.001		
Andersson et al., 2011	Danish Nationwide Registers	To investigate the outcomes of sulfonylurea monotherapy in patients with heart failure, HF	Alive patients 30 days after 1st HF hospitalization in 1997–2006, who received glimepiride (n = 1097), glibenclamide (n = 1031), glipizide (n = 557), gliclazide (n = 251), or tolbutamide (n = 541). Median follow-up (IQR):	Risk of mortality associated with individual SU was similar in patients with HF. Similar HR for mortality obtained: glimepiride 1.10 (0.92–1.33), glibenclamide 1.12 (0.93–1.34), glipizide 1.14 (0.93–1.38), tolbutamide 1.04 (0.85–1.26), and gliclazide (reference)	Selection bias, informative censoring, lag time bias	

Author, Year	Data source	Purpose	Participants	Main outcome	Bias reported	Bias control
			744 (268–1451) days			
Graham et al., 2010	Medicare Database, US	To compare CV risks between rosiglitazone and pioglitazone	227,571 patients aged ≥65 years who initiated treatment with rosiglitazone or pioglitazone from July 2006-June 2009. Follow-up for up to 3 years	Compared with pioglitazone, rosiglitazone was associated with an increased risk of stroke: HR, 1.27 (1.12-1.45), HF: 1.25 (1.16-1.34), and all-cause mortality: 1.14 (1.05-1.24), and an increased risk of the composite of AMI, stroke, heart failure, or all-cause mortality: 1.18 (1.12-1.23)		

Abbreviations: Diabetic Macular Oedema (DME); Type 2 diabetes (T2D); Thiazolidinedione (TZD); Metformin (MET); Sulphonylurea (SU); Confidence Interval (CI); acute myocardial infarction (AMI); Adjusted hazard ratio (aHR); odds ratio (OR); Risk ratio (RR); heart failure (HF); cardiovascular (CV); advanced pancreatic adenocarcinoma (PAC).

Appendix A-3 Glossary of summary description of biases

Immortal time bias	A form of bias that derives from including immortal time in the follow-up. Immortal time is a period of time, during which the event cannot occur, i.e. person-time that is event-free by definition
Exposure-defined cohort	A cohort of patients whose cohort-entry is defined by the first use of a drug of interest
Overadjustment bias	A form of bias that derives from adjusting by variables that are intermediate between exposure and outcome, i.e. outcome risk factors that have been influenced by exposure to the drug of interest
New-user design	A cohort study design that starts following patients at the time they initiate a new drug
Exposure risk window	The time period during which a drug of interest puts a patient at risk of a harmful or beneficial effect with regard to a specific outcome
Lag time after drug initiation	A time period following drug initiation during which a specific outcome cannot be attributed to the initiated drug. This time period should not be included in the follow-up time
Latency period after drug discontinuation	A time period after drug discontinuation during which a specific outcome can still be attributed to the discontinued drug. This time period should be included in the follow-up time
As-treated analysis (AT)	An analytical approach that terminates exposure to a medication when the patient discontinues that medication
Intention-to-treat analysis (ITT)	An analytical approach that carries forward the initial exposure status and disregards changes in treatment status over time
Informative censoring	A form of selection bias arising when discontinuation of a drug of interest is prognostic of (associated with) a future outcome
Confounding	A mixing of effects that arises when patients with different baseline risks are compared—the resulting effect measure is a mix of drug effects and risk factor effects

Systematic review 2 (Chapter 8)

Sample search strategy

Database: EMBASE

1. Fractures, Bone/

2. exp Fractures, Bone/

3. fractur*.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

4. break*.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

5. 1 or 3 or 4

6. 2 or 3 or 4

7. exp Sulfonylurea Compounds/

8. (Sulphonylurea or Sulfonylurea).mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

9. (glibenclamide or Glyburide).mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

10. (glicalzide or tolbutamide or glimepiride or glipizide).mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

11. 7 or 8 or 9 or 10

12. 6 and 11

13. limit 12 to (english and humans)

14. "dipeptidyl-peptidases and tripeptidyl-peptidases"/ or dipeptidyl peptidase 4/

15. DPP-4.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

16. dpp-4.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

17. ("dipeptidyl-peptidase*" or "dipeptidyl peptidase*").mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

18. dpp-iv.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

19. "dpp iv".mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

20. "dpp 4".mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

21. vildagliptin.mp.

22. Dipeptidyl-Peptidase IV Inhibitors/

23. gliptin*.mp. (mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword)

24. sitagliptin.mp.

25. saxagliptin.mp.

26. alogliptin.mp.

27. linagliptin.mp.

28. dutogliptin.mp.

29. dipeptidyl peptidase/

30. dipeptidyl peptidase IV/

31. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30

32. 6 and 31

33. 12 or 32

34. limit 33 to (human and english language and (clinical trial or randomized controlled trial or controlled clinical trial))

35. from 34 keep 16,26,31

36. from 34 keep 34-35,37,39,41-43,46,49,57,105,132

37. from 34 keep 105,132,174,180,184,209,272,309,347-349,354,357-358

38. 35 or 36 or 37

Appendix A-4 Data extraction form for RCTs that compared a DPP-4 inhibitor with placebo

Author, Year	Clinical trial No.	Dur.	Intervention	N	No. of fractures	Comparator	n	No. of fractures	SD
Hollander, 2011	NCT00295633	76	Sitagliptin (2.5mg, 5mg)	381	5	Placebo/TZD	184	1	0.1
Yang, 2011	NCT00661362	24	Saxagliptin (5mg)	283	3	Placebo/Metformin	287	0	0
Scherbaum, 2008	NCT00101712	52	Vildagliptin (50mg)	156	0	Placebo	150	1 cervical	0.1
Fonseca, 2007	NCT00099931	24	Vildagliptin (50mg) BID	144	0	Placebo/Insulin	152	1	1.1
NCT00286468	NCT00286468	26	Alogliptin (12.5mg, 25mg)	401	1 spinal compression	Placebo/Glyburide	99	0	0
Dobs, 2013	NCT00350779	54	Sitagliptin (100mg)	170	0	Placebo/Rosiglitazone/ Metformin	92	1 lower limb	1
NCT01076075	NCT01076075	54	Sitagliptin (100mg)	210	0	Placebo/Pioglitazone	212	1 skull	0.8
Barnett, 2013	NCT01084005	24	Linagliptin (5mg)	162	1 lower limb 1 vertebra	Placebo	79	0	0.1
NCT00316082	NCT00316082	24	Saxagliptin (2.5mg, 5mg)	291	1 femoral neck 1 spinal	Placebo	74	0	0.1
Jadzinsky, 2009	NCT00327015	24	Saxagliptin (5mg, 10mg)	643	1 upper limb 1 wrist	Placebo/Metformin	328	0	0.1
Chacra, 2009	NCT00313313	24	Saxagliptin (2.5mg, 5mg)	501	1 lower limb	Placebo/Glyburide	267	0	0.1

Author, Year	Clinical trial No.	Dur.	Intervention	N	No. of fractures	Comparator	n	No. of fractures	SD
NCT00305604	NCT00305604	24	Sitagliptin (50mg/ 100mg)	102	0	Placebo	104	1 lumbar vertebra, 1 upper limb	0.7
Fonseca, 2013	NCT00885352	26	Sitagliptin (100mg)	157	0	Placebo	156	1 patella	1
NCT00121667	NCT00121667	24	Saxagliptin (2.5mg, 5mg, 10mg)	564	4	Placebo/Metformin /Pioglitazone	179	0	0.1
Goldstein, 2007	NCT00103857	24	Sitagliptin (50mg BID, 100mg QD)	551	1 femur	Placebo/Metformin	176	0	1.2
NCT00087516	NCT00087516	24	Sitagliptin (100mg, 200mg)	488	0	Placebo	253	1 ankle	0.9
NCT00086515	NCT00086515	104	Sitagliptin (100mg)	464	0	Placebo/Glipizide	237	1 traumatic	0.8
NCT01215097	NCT01215097	24	Linagliptin (5mg)	205	1 comminuted	Placebo	100	0	0.8
NCT00984867	NCT00984867	24	Sitagliptin	223	0	Placebo/Sitagliptin/ Metformin	224	1 upper limb	0.8
NCT00968708	NCT00968708	205	Alogliptin (6.25mg/ 12.5mg/ 25mg)	2701	13	Placebo	2679	21	NR
NCT00757588	NCT00757588	24	Saxagliptin (5mg)	304	1 ankle	Placebo/Insulin	151	0	NR
NCT00734474	NCT00734474	104	Sitagliptin (100mg)	492	2	Placebo/Dutaglutide/ Metformin	710	4	1.1
NCT00800683	NCT00800683	52	Linagliptin (5mg)	68	1 femur	Placebo	65	0	1

Author, Year	Clinical trial No.	Dur.	Intervention	N	No. of fractures	Comparator	n	No. of fractures	SD
					1 humerus				
Henry, 2014	NCT00722371	54	Sitagliptin (100mg)	922	2 foot 1 upper limb	Placebo/ Pioglitazone	693	1 foot	1.1
Alba, 2013	NCT00511108	21	Sitagliptin (100mg)	104	1 tibia	Placebo/Pioglitazone	107	0	1
Yoon, 2011	NCT00397631	24	Sitagliptin (100mg)	261	1 humerus	Placebo/ Pioglitazone	259	0	1.2
Vilsboll, 2010	NCT00395343	24	Sitagliptin (100mg)	322	1 pelvic	Placebo/Insulin/ Metformin	319	0	0.9
NCT00374907	NCT00374907	116	Saxgliptin (5mg)	20	1 ankle	Placebo/Metformin	16	0	NR
NCT00372060	NCT00372060	52	Sitagliptin (50mg-100mg)	133	1 patella 1 rib	Placebo/Sitagliptin Pioglitazone	68	1 lower limb	0.8
NCT00337610	NCT00337610	30	Sitagliptin (100mg)	96	0	Placebo/Metformin	94	1 upper limb	0.8
DeFronzo, 2012	NCT00328627	26	Alogliptin (12.5mg, 25mg)	1037	0	Placebo/Pioglitazone	518	1 ankle	0.7
NCT00838903	NCT00838903	156	Sitagliptin (100mg)	302	1 femur 1 spinal	Placebo/Metformin	101	0	NR
Study 057	NR	52	Saxagliptin (5mg)	304	2 foot 1 ankle	Placebo/Insulin	151	1 hand, 1 humerus, 1 lower limb	0.9
Rosenstock, 2006	NCT00086502	24	Saxagliptin (100mg)	175	0	Placebo/Pioglitazone	178	1 lower limb	0.8

Author, Year	Clinical trial No.	Dur.	Intervention	N	No. of fractures	Comparator	n	No. of fractures	SD
Study D1680L00006	NCT01128153	24	Saxagliptin (5mg)	129	0	Placebo/Metformin/ Sulfonylurea	128	1 rib	0.8
NCT01098539	NCT01098539	52	Sitagliptin (25mg-100mg)	246	0	Placebo/Albiglutide	249	1 radius, 1 sternal	NR
Arjona Ferreira, 2013a	NCT00509262	54	Sitagliptin (25mg/ 50mg)	211	1 patella	Placebo/Glipizide	212	1 femur	0.7
The patients had a mean age of 57 years, ranging from 50.9 to 74.9 years. There were 48 cases of fracture amongst 10,051 patients in the placebo group, and 55 fractures amongst 13,923 patients taking a DPP-4 inhibitor									

Appendix A-5 Patient characteristics for RCTs involving DPP-4 inhibitor and placebo

Author, Year	Mean age (years)	SD	Female (%)	Baseline HbA_{1c} (%)	SD	Baseline BMI (kg/m^2)	SD
Hollander, 2011	54	0.9	50.5	8.3	0.1	30	0.3
Yang, 2011	54.1	0.4	51.8	7.9	0	26.2	0.1
Scherbaum, 2008	63.1	0.4	40.6	6.8	0.1	30.2	0.3
Fonseca, 2007	59.3	0.5	48.6	8.4	1.1	33.1	5.6
NCT00286468	56.7	0.4	47.8	8.1	0	30.1	1.1
Dobs, 2013	54.6	0.3	42.4	8.8	1	30.3	6
NCT01076075	54.9	0.7	54.3	8.4	0.8	NR	NR
Barnett, 2013	74.9	0	33.2	7.8	0.1	29.7	0.1
NCT00316082	55	10	54	7.9	0.1	NR	NR
Jadzinsky, 2009	52	11	50.8	9.5	0.1	30.2	4.8
Chacra, 2009	55.1	10	54.9	8.4	0.1	29	4.6
NCT00305604	71.9	6	34.6	7.8	0.7	NR	NR

Author, Year	Mean age (years)	SD	Female (%)	Baseline HbA_{1c} (%)	SD	Baseline BMI (kg/m^2)	SD
Fonseca, 2013	56.1	9	37.7	8.7	1	30	5.2
NCT00121667	54.6	10	49.3	8.1	0.1	31.4	4.9
Goldstein, 2007	53.4	9.9	49.8	9	1.2	32.1	6.6
NCT00087516	54.2	9.9	48.3	8	0.9	NR	NR
NCT00086515	54.5	10	42.9	8	0.8	NR	NR
NCT01215097	55.5	10	50.2	8	0.8	25.6	4
NCT00984867	54.9	10	45.2	7.9	0.8	NR	NR
NCT00968708	60.9	9.9	32.1	NR	NR	29.5	5.6
NCT00757588	57.2	9.4	58.7	NR	NR	NR	NR
NCT00734474	54.1	9.9	53.5	8.1	1.1	31.3	4.4
NCT00800683	64.4	10	39.8	8.2	1	32	5.8
Henry, 2014	51.8	1.1	43.5	8.8	1.1	30.9	5.4
Alba, 2013	53.6	7.9	44.5	7.9	1	30.9	4.8

Yoon, 2011	50.9	11	45.8	9.5	1.2	29.7	5.2
Vilsboll, 2010	57.8	9.2	49.1	8.7	0.9	31	5
NCT00374907	56.5	2.1	61.1	NR	NR	32.8	0.7
NCT00372060	58.4	9.5	35.1	7.7	0.8	NR	NR
NCT00337610	54.8	9.5	53.7	9.2	0.8	NR	NR
DeFronzo, 2012	54.4	9.5	55.1	8.5	0.7	31.2	5.1
NCT00838903	54.5	10	52.4	NR	NR	NR	NR
Study 057	57.2	NR	58.7	8.7	0.9	32.3	NR
Rosenstock, 2006	56.2	11	44.5	8	0.8	31.5	5.1
Study D1680L00006	57	11	40.1	8.3	0.8	29.2	5.1
NCT01098539	63.3	8.7	46.3	NR	NR	NR	NR
Arjona Ferreira, 2013a	64.2	10	40.2	7.8	0.7	26.8	4.8

NR= not reported, BID= twice daily, QD= once daily, SD= standard deviation, n= number of individuals

Appendix A-6 Data extraction form for RCTs that compared a DPP-4 inhibitor with active comparator

Author, Year	Clinical trial No.	Wks	Intervention	n	No. of fractures	Comparator	n	No. of fractures
Iwamoto, 2010	NR	12	Sitagliptin (50mg)	163	0	Voglibose (0.2mg) TID	156	1 foot
NCT00707993	NCT00707993	52	Alogliptin (25mg)	222	1 multiple 1 upper limb	Glipizide (5mg)	219	1 stress
NCT00856284	NCT00856284	104	Alogliptin (12.5mg, 25mg)	1765	3 ankle 1 femur 1 tibia 1 comminuted	Glipizide (5mg-20mg)/ Metformin (1500mg-3300mg)	874	1 ankle 1 facial bone 1 lower limb
NCT01006603	NCT01006603	52	Saxagliptin (5mg)	360	2 lumbar vertebra 1 femur 1 hand	Glimepiride (1mg-6mg)	360	1 ankle
Schernthaner, 2013	NCT01137812	52	Sitagliptin (100mg)	378	1 lower limb	Canagliflozin (300mg)/Metformin/ Sulfonylurea	377	1 hand 1 hip
Nauck, 2007	NCT00094770	52	Sitagliptin (100mg)	588	1 lower limb 1 radius 1 tibia	Glipizide (5mg-20mg)	584	1 ankle 1 lower limb 1 radius
Charbonnel, 2013	NCT01296412	26	Sitagliptin (100mg)	326	1 rib 1 spinal compression 1 sternal	Liraglutide (0.6mg-1.8mg)/ Metformin (>1500mg)	327	0
Arechavaleta, 2011	NCT00701090	30	Sitagliptin (100mg)	516	1 humerus 1 patella	Glimepiride (1mg-6mg)/ Metformin (>1500mg)	519	1 clavicle

Author, Year	Clinical trial No.	Wks	Intervention	n	No. of fractures	Comparator	n	No. of fractures
Wainstein, 2012	NCT00532935	35	Sitagliptin (50mg) BID	261	1 femur	Pioglitazone (30mg-45mg)	256	0
Arjona Ferreira, 2013b	NCT00509236	54	Sitagliptin (25mg)	64	1 hip 1 pelvic	Glipizide (2.5mg-20mg)	65	0
Aschner, 2010	NCT00449930	24	Sitagliptin (100mg)	528	1 tibia	Metformin (500mg-1000mg) BID	522	0
Goke, 2013	NCT00575588	104	Saxagliptin (5mg)	428	1 femoral neck 1 lumbar vertebra 1 patella 1 upper limb	Glipizide (5-20mg)/ Metformin	430	1 femur 1 humerus
Study P049	NR	24	Sitagliptin (100mg)	455	1 tibia	Metformin (2000mg)	439	0
NCT00482729	NCT00482729	44	Sitagliptin (50mg) BID	625	1 rib	Metformin (500mg-1000mg) BID	621	1 rib, 1 skull
The patients had a mean age of 58 years, ranging from 49.7 to 72.6 years. There were 16 cases of fracture amongst 5,749 patients in the active comparator, and 31 fractures with 6,679 patients in the intervention group.								

Appendix A-7 Patient characteristics for RCTs involving DPP-4 inhibitor and active comparator

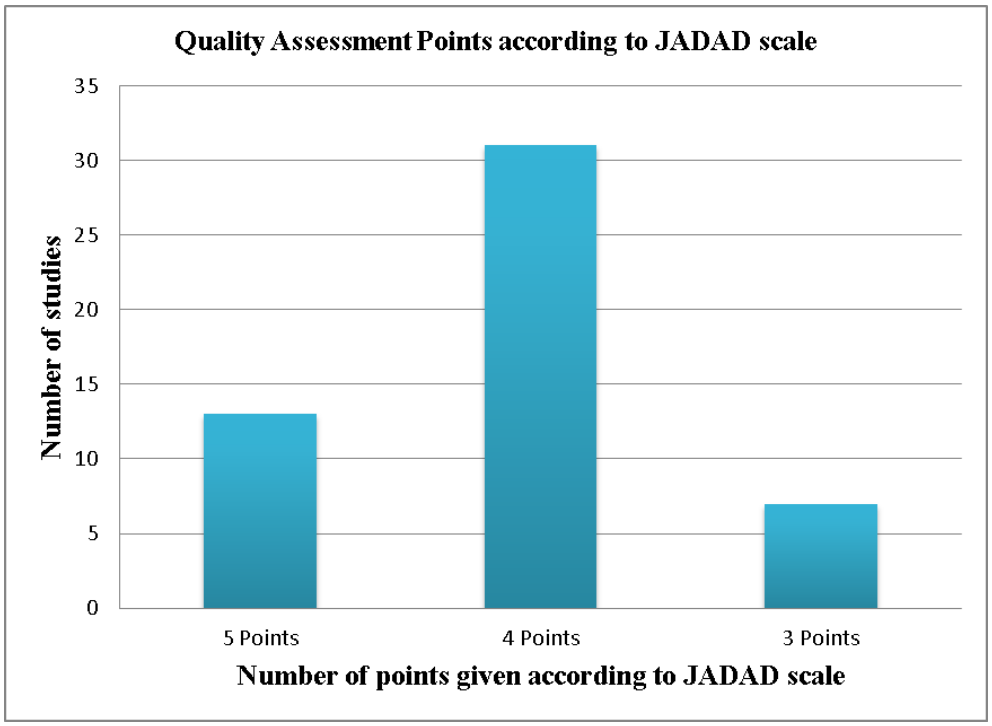
Author, Year	Mean age (years)	SD	Female (%)	Baseline HbA_{1c} (%)	SD	Baseline BMI (kg/m^2)	SD
Iwamoto, 2010	60.7	10	33.5	7.8	0.9	24.7	3.5
NCT00707993	69.9	4.2	55.1	NR	NR	29.8	4.4
NCT00856284	55.4	9.7	50.3	7.6	0.6	31.2	5.4
NCT01006603	72.6	5.6	38.2	NR	NR	NR	NR
Schernthaner, 2013	56.5	9.5	44.1	8.1	0.9	31.6	6.9
Nauck, 2007	56.7	9.6	40.8	7.7	0.9	31.3	5.1
Charbonnel, 2013	57.3	10	45.2	8.2	1	32.7	6
Arechavaleta, 2011	56.3	9.9	45.6	7.5	0.7	30	4.5
Wainstein, 2012	52.3	11	46.4	8.9	1.3	29.8	5.8
Ferreira, 2013b	59.5	9.5	40.3	7.9	0.7	26.8	5
Aschner, 2010	56	11	53.9	7.3	0.7	30.8	4.8
Goke, 2013	57.6	10	48.3	7.5	0.1	31.4	NR
Study P049	56	NR	53.9	7.2	NR	NR	NR
NCT00482729	49.7	11	43.2	9.9	1.8	NR	NR

NR= not reported, BID= twice daily, QD= once daily, TID= thrice daily, SD= standard deviation

Appendix A-8 JADAD scale points for the 51 included RCTs

Author, Year	Randomisation	Blinding	An account of all patients	Total score
Hollander, 2011	1	2	1	4
Iwamoto, 2010	2	2	1	5
Yang, 2011	2	2	1	5
Scherbaum, 2008	1	2	1	4
Fonseca, 2007	1	2	1	4
NCT00286468	1	2	1	4
Dobs, 2013	1	2	1	4
NCT00707993	1	2	1	4
NCT00856284	1	1	1	3
NCT01006603	1	1	1	3
NCT01076075	1	2	1	4
Barnett, 2013	2	2	1	5
Schernthaner, 2013	2	1	1	4
NCT01098539	1	2	1	4
NCT00316082	1	2	1	4
Jadzinsky, 2009	2	2	1	5
Chacra, 2009	2	2	1	5
NCT00305604	1	2	1	4
Fonseca, 2013	1	2	1	4
NCT00121667	1	2	1	4
Goldstein, 2007	1	2	1	4
Nauck, 2007	1	2	1	4
NCT00087516	1	2	1	4
NCT00086515	1	2	1	4
NCT01215097	1	2	1	4
Charbonnel, 2013	2	0	1	3
NCT00984867	2	2	1	5
NCT00968708	1	2	1	4
NCT00757588	1	2	1	4
NCT00734474	1	2	1	4
Arechavaleta, 2011	2	2	1	5
NCT00800683	1	2	1	4
Henry, 2014	1	2	1	4
Wainstein, 2012	1	1	1	3
Alba, 2013	2	2	1	5
Arjona Ferreira, 2013a	2	2	1	5
Arjona Ferreira 2013b	2	2	1	5
NCT00482729	1	1	1	3
Aschner, 2010	2	2	1	5

Appendix A-9 number of RCTs with the point scores 1-5



Appendix B.

Read codes definition of comorbidities

Appendix B-1 Description of comorbidities

Description	Read code
Acute myocardial infarction (AMI)	
Acute ST segment elevation myocardial infarction	G30X000
Acute anterolateral infarction	G300.00
Acute anteroseptal infarction	G301100
Acute coronary syndrome	G311500
Acute inferolateral infarction	G302.00
Acute myocardial infarction	G30..00
Acute myocardial infarction NOS	G30z.00
Acute non-ST segment elevation myocardial infarction	G307100
Attack – heart	G30..11
Coronary thrombosis	G30..12
Heart attack	G30..14
Inferior myocardial infarction NOS	G308.00
MI - acute myocardial infarction	G30..15
Other specified anterior myocardial infarction	G301.00
Posterior myocardial infarction NOS	G304.00
Silent myocardial infarction	G30..17
Subsequent myocardial infarction	G35..00
Coronary heart disease (CHD)	
Acute coronary insufficiency	G31y000
Acute inferoposterior infarction	G303.00
Angina at rest	G311200
Angina on effort	G33z300
Angina pectoris	G33..00
Angina pectoris NOS	G33zz00
Anterior myocardial infarction NOS	G301z00
Arteriosclerotic heart disease	G3...11
Asymptomatic coronary heart disease	G34z000
Atherosclerotic cardiovascular disease	G342.00
Atherosclerotic heart disease	G3...12
Chronic myocardial ischaemia	G34y100
Coronary artery disease	G340.12
Coronary artery spasm	G332.00

Description	Read code
Coronary atherosclerosis	G340.00
Crescendo angina	G311.11
Double coronary vessel disease	G340100
Dressler's syndrome	G310.11
IHD - Ischaemic heart disease	G3...13
Ischaemic cardiomyopathy	G343.00
Ischaemic chest pain	G33z400
Ischaemic heart disease	G3...00
Ischaemic heart disease NOS	G3z..00
Lateral myocardial infarction NOS	G305.00
Mural thrombosis	G30A.00
New onset angina	G33z600
Old myocardial infarction	G32..00
Other acute and subacute ischaemic heart disease	G31..00
Other chronic ischaemic heart disease	G34..00
Other specified ischaemic heart disease	G3y..00
Prinzmetal's angina	G331.00
Single coronary vessel disease	G340000
Stable angina	G33z700
Triple vessel disease of the heart	G340.11
Unstable angina	G311.13
Worsening angina	G311400
Stroke/Cerebrovascular disease	
Amaurosis fugax	F423600
Brain stem stroke syndrome	G663.00
CVA - Cerebrovascular accident unspecified	G66..13
CVA - cerebral artery occlusion	G64..11
CVA - cerebrovascular accid due to intracerebral haemorrhage	G61..11
CVA unspecified	G66..11
Carotid artery occlusion	G631.00
Carotid artery stenosis	G634.00
Cerebellar haemorrhage	G613.00
Cerebellar infarction	G64z.12
Cerebral aneurysm, nonruptured	G673.00
Cerebral arterial occlusion	G64..00
Cerebral embolism	G641.00
Cerebral infarct due to thrombosis of precerebral arteries	G63y000
Cerebral infarction NOS	G64z.00
Cerebral infarction due to embolism of cerebral arteries	G641000
Cerebral infarction due to embolism of precerebral arteries	G63y100
Cerebral infarction due to thrombosis of cerebral arteries	G640000
Cerebral thrombosis	G640.00

Description	Read code
Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs	G6X..00
Cerebrovascular disease	G6...00
Cerebrovascular disease NOS	G6z..00
Chronic cerebral ischaemia	G671100
Drop attack	G65..11
Extradural haemorrhage - nontraumatic	G620.00
Hypertensive encephalopathy	G672.00
Infarction - cerebral	G64..12
Infarction - precerebral	G63..11
Infarction of basal ganglia	G64z400
Intracerebral haemorrhage	G61..00
Intracerebral haemorrhage NOS	G61z.00
Intracerebral haemorrhage, intraventricular	G617.00
Left sided CVA	G667.00
Left sided cerebral infarction	G64z200
Middle cerebral artery syndrome	G660.00
Other transient cerebral ischaemia	G65y.00
Pontine haemorrhage	G614.00
Posterior cerebral artery syndrome	G662.00
Pure motor lacunar syndrome	G665.00
Right sided CVA	G668.00
Right sided cerebral infarction	G64z300
Stenosis, carotid artery	G631.11
Stroke and cerebrovascular accident unspecified	G66..00
Stroke due to cerebral arterial occlusion	G64..13
Stroke due to intracerebral haemorrhage	G61..12
Stroke unspecified	G66..12
Subclavian steal syndrome	G652.00
Subdural haematoma - nontraumatic	G622.00
Subdural haemorrhage - nontraumatic	G621.00
Subdural haemorrhage NOS	G623.00
Transient cerebral ischaemia	G65..00
Transient cerebral ischaemia NOS	G65z.00
Transient global amnesia	G655.00
Transient ischaemic attack	G65..12
Vertebrobasilar insufficiency	G656.00
Heart failure	
Acute congestive heart failure	G580000
Acute heart failure	G582.00
Acute left ventricular failure	G581000
Biventricular failure	G580.14
Cardiac failure	G58..11

Description	Read code
Cardiac failure NOS	G58z.12
Chronic congestive heart failure	G580100
Congestive cardiac failure	G580.11
Congestive heart failure	G580.00
Decompensated cardiac failure	G580200
H/O: aortic aneurysm	14AE.00
H/O: heart failure	14A6.00
Heart failure	G58..00
Heart failure NOS	G58z.00
Heart failure confirmed	1O1..00
Impaired left ventricular function	G581.13
Left ventricular diastolic dysfunction	G5yyA00
Left ventricular failure	G581.00
Left ventricular systolic dysfunction	G5yy900
New York Heart Association classification - class I	662f.00
New York Heart Association classification - class II	662g.00
New York Heart Association classification - class III	662h.00
New York Heart Association classification - class IV	662i.00
Pulmonary oedema - acute	G581.12
Right heart failure	G580.12
Right ventricular failure	G580.13
Hypoglycaemia	
Drug-induced hypoglycaemia without coma	C11y100
Frequency of GP or paramedic treated hypoglycaemia	66A7100
Frequency of hospital treated hypoglycaemia	66A7000
Hypoglycaemia unspecified	C112.00
Hypoglycaemia unspecified NOS	C112z00
Hypoglycaemic attack requiring 3rd party assistance	66Ad.00
Hypoglycaemic coma	C110.00
Hypoglycaemic coma NOS	C110z00
Insulin dependent diabetes mellitus with hypoglycaemic coma	C108E00
Loss of hypoglycaemic warning	66AJ200
Other hypoglycaemia	C116.00
Post-prandial hypoglycaemia	C116000
Reactive hypoglycaemia NOS	C112000
Type 1 diabetes mellitus with hypoglycaemic coma	C10EE00
Type 2 diabetes mellitus with hypoglycaemic coma	C10FD00
X Other hypoglycaemia	Cyu3000
Peripheral arterial disease (PAD)	
AAA - Abdominal aortic aneurysm without mention of rupture	G714.11
Abdominal aortic aneurysm without mention of rupture	G714.00

Description	Read code
Aortic aneurysm	G71..00
Aortic aneurysm repair	7A14.11
Aortic atherosclerosis	G700.00
Arterial embolism and thrombosis	G74..00
Arterial leg ulcer	M271300
Axillo-unifemoral PTFE bypass graft	7A10300
Bypass femoral artery by fem/pop art anast c vein graft NEC	7A48200
Cardiac failure therapy	8B29.00
Chronic peripheral venous hypertension	G8y3.00
Claudication	G73z011
Diabetic peripheral angiopathy	G73y000
Dissecting aortic aneurysm	G710.00
Embolism and thrombosis of the axillary artery	G74y700
Embolism and thrombosis of the femoral artery	G742400
Embolism and thrombosis of the iliac artery unspecified	G74y300
Embolism and thrombosis of the subclavian artery	G74y500
Extremity artery atheroma	G702.00
Femoro-femoral prosthetic cross over graft	7A48E00
Gangrene of finger	G732200
Gangrene of foot	G732100
Gangrene of toe	G732000
Intermittent claudication	G73z000
Ischaemia of legs	G73..12
Ischaemic leg ulcer	M271.12
Ischaemic ulcer diabetic foot	M271000
Mixed venous and arterial leg ulcer	M271400
Other bypass of common femoral artery	7A48.12
Other bypass of femoral artery	7A48.14
Other bypass of femoral artery or popliteal artery	7A48.00
Other bypass of femoral artery or popliteal artery NOS	7A48z00
Other bypass of femoral artery or popliteal artery OS	7A48y00
Other bypass of superficial femoral artery	7A48.16
Other emergency bypass of femoral artery	7A47.16
Other emergency bypass of popliteal artery	7A47.14
Other peripheral vascular disease	G73..00
Other specified peripheral vascular disease NOS	G73yz00
Percutaneous transluminal balloon angioplasty of aorta	7A1A000
Peripheral gangrene	G732.00
Peripheral ischaemia	G73..13
Peripheral ischaemic vascular disease	G73..11
Peripheral vascular disease NOS	G73z.00
Peripheral vascular disease monitoring	662U.00
Raynaud's disease	G730000

Description	Read code
Raynaud's phenomenon	G730100
Raynaud's syndrome	G730.00
Ruptured abdominal aortic aneurysm	G713.11
Ruptured aortic aneurysm NOS	G715.00
Saddle embolus	G740.14
Thoracic aortic aneurysm which has ruptured	G711.00
Thoracic aortic aneurysm without mention of rupture	G712.00
Thrombosis - arterial	G74..12

Appendix C.

Sensitivity analyses, regression models and diagnostic tests

Appendix C-1 Sensitivity analysis comparing glycaemic efficacy in data with missing information and complete cohort

	Missing data cohort							
	SU vs. MET				MET + SU vs. MET			
HbA1c	**Mean**	**SE**	**95% CI**	***P***	**Mean**	**SE**	**95% CI**	***P***
Overall	−0.17	0.12	(−0.38, 0.07)	0.2	0.04	0.05	(−0.07, 0.15)	0.4
Subgroup								
7–7.5%	0.15	0.27	(−0.38, 0.67)	0.6	0.14	0.12	(−0.10, 0.37)	0.3
7.5–8%	−0.01	0.25	(−0.50, 0.48)	0.9	0.03	0.12	(−0.20, 0.26)	0.8
8–9%	0.05	0.22	(−0.39, 0.49)	0.8	−0.08	0.10	(−0.26, 0.11)	0.4
≥ 9%	0.28	0.20	(−0.10, 0.66)	0.2	0.15	0.09	(0.03, 0.33)	0.04

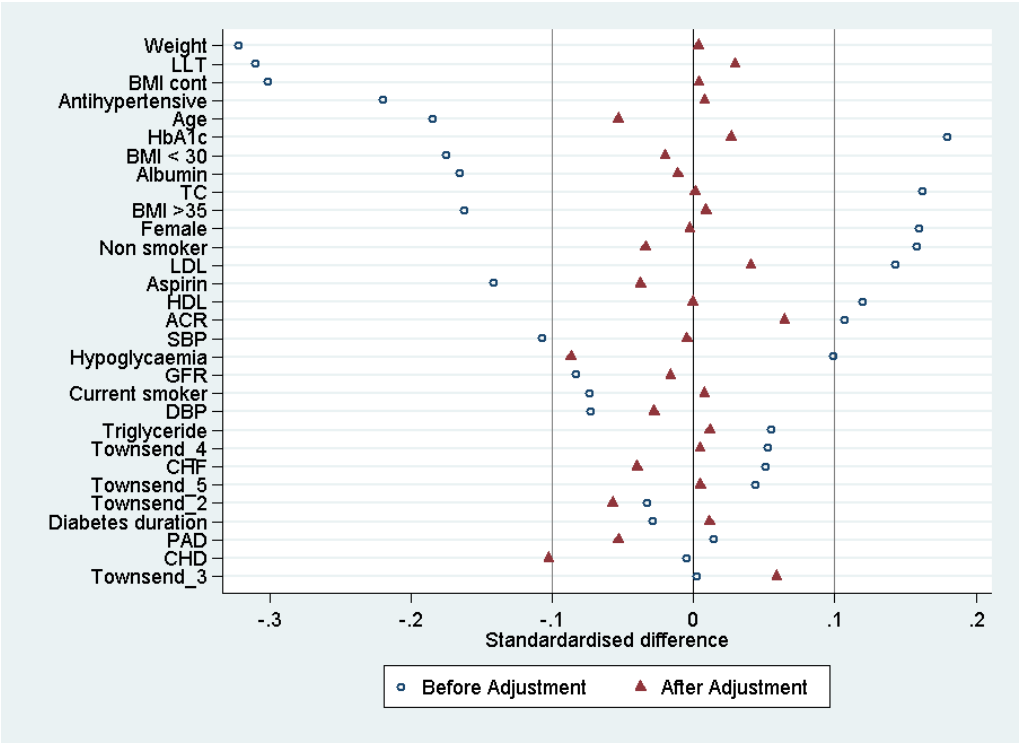
	Complete case cohort							
	SU vs. MET				MET + SU vs. MET			
HbA1c	**Mean**	**SE**	**95% CI**	***P***	**Mean**	**SE**	**95% CI**	***P***
Overall	−0.03	0.06	(−0.14, 0.09)	0.6	0.03	0.03	(−0.02, 0.08)	0.3
Subgroup								
7–7.5%	0.05	0.13	(−0.21, 0.31)	0.7	0.03	0.06	(−0.08, 0.15)	0.6
7.5–8%	−0.05	0.12	(−0.30, 0.19)	0.7	−0.01	0.06	(−0.12, 0.11)	0.9
8–9%	−0.01	0.11	(−0.22, 0.20)	0.9	−0.02	0.05	(−0.11, 0.07)	0.6
≥ 9%	0.08	0.09	(−0.11, 0.26)	0.4	0.18	0.04	(0.16, 0.31)	0.01

Appendix C-2 Cox proportional hazard regression model for the risk of treatment failure

Variable	Adjusted hazards ratio (95% CI)	P value
Treatment groups		
MET + SU	1.00	.
MET + DPP-4i	1.58 (1.48,1.68)	<0.001
MET + TZD	0.45 (0.41,0.50)	<0.001
Age (Years)	0.97 (0.97,0.97)	<0.001
TC (mmol/l)	0.95 (0.93,0.97)	<0.001
Weight (Kg)	1.02 (1.02,1.02)	<0.001
SBP (mmHg)	0.99 (0.99,1.00)	<0.001
HbA1c (%)	1.02 (1.01,1.03)	0.002
Diabetes duration (Years)	1.07 (1.06,1.08)	<0.001
Gender		
Male	1.00	.
Female	1.38 (1.31,1.46)	<0.001
Townsend deprivation		
Least deprived	1.00	.
Less	0.92 (0.86,0.99)	0.03
Average	0.83 (0.78,0.90)	<0.001
More	0.84 (0.78,0.90)	<0.001
Most deprived	0.90 (0.83,0.97)	0.005
Smoking status		
Non-smoker	1.00	.
Current smoker	1.07 (1.01,1.13)	0.03
Ex-smoker	1.10 (1.04,1.15)	0.001
Use of Medications		
Aspirin	0.73 (0.69,0.78)	<0.001
LLT	1.57 (1.46,1.68)	<0.001
Comorbidities		
CHD	0.44 (0.36,0.55)	<0.001
PAD	0.59 (0.47,0.74)	<0.001
Hypoglycaemia	0.30 (0.25,0.36)	<0.001

Adjusted for body mass index (BMI), low-density lipoprotein (LDL), high-density lipoprotein (HDL-C), triglycerides, diastolic BP, glomerular filtration rate (GFR), urinary albumin-creatinine ratio (ACR), the use of antihypertensive drugs and heart failure.

Appendix C-3 Balance diagnostics of standardized differences of covariates between the full and matched cohorts



Appendix C-4 Cox proportional hazard regression model for the risk of composite outcome of in the propensity matched sample population

	Adjusted hazards ratio (95% CI)	P value
Treatment groups		
MET + SU + DPP-4	1 (1.00,1.00)	
MET + SU + INS	2.55 (1.93,3.37)	<0.001
CHD		
No	1 (1.00,1.00)	
Yes	4.44 (3.12,6.31)	<0.001
Hypoglycaemia		
No	1 (1.00,1.00)	
Yes	1.71 (1.20,2.44)	0.003
Propensity score	2.04 (1.05,3.97)	0.04

Abbreviation: MET (metformin); SU (sulphonylurea); DPP-4 (dipeptidylpeptidase-4 inhibitor); INS (insulin); CHD (coronary heart disease)